

Распознавание текста на основе скелетного представления толстых линий и сверточных сетей*

Григорьев А. Д., Местецкий Л. М., Райер И. А, Жариков И. Н,
Стрижов В. В.

grigorev.ad@phystech.edu; mestlm@mail.ru; reyer@forecsys.ru;
zharikov.i.n@yandex.ru; strijov@phystech.edu

Московский физико-технический институт

В работе рассматривается задача распознавания символов на изображении. Предлагается новый способ решения, основанный на медиальном представлении фигуры. Производится построение сверточной нейронной сети, получающей на вход граф скелетного представления объекта. Оценка качества предложенного метода проводится на корпусе изображений арабских цифр в растровом представлении.

Ключевые слова: распознавание символов, сверточная нейронная сеть, медиальное представление, скелетонизация.

1 Введение

В работе решается задача распознавания символов на изображении. Данная проблема возникает при распознавании текста после сегментации на символы, что имеет множество применений на практике, например, при распознавании содержимого документов на фотографии или отцифровке старых книг.

Стандартные существующие методы решения этой задачи предполагают работу с дискретным представлением изображения в виде матрицы пикселей, каждый элемент которой несет информацию о цвете пикселя в формате RGB. Такой формат представления изображений называется растровым. Данный тип графики может быть быстро и эффективно обработан компьютером, в связи с чем широко распространен. Работе с растровыми изображениями посвящено большое число работ по классификации объектов. В последнее время с появлением нейронных сетей качество такой классификации существенно возросло. В частности сверточные нейронные сети в данной задаче позволяют добиться достаточно высокой точности, так ансамбль сверточных нейронных сетей в задаче классификации рукописных цифр набора данных MNIST дает точность 99.77% [1].

В данной работе предлагается развить другой подход работы с изображениями, основанный на непрерывном представлении в виде скелета. Основы данного подхода, а также алгоритмы скелетонизации описаны в работах Местецкого Л. М. [4, 5, 9]. Скелетное представление фигуры может быть представлено в виде графа, алгоритмы его построения были описаны и реализованы [2, 3, 6]. Ранее были предприняты попытки построения классификаторов на основе графового представления скелета фигуры, использовались такие модели, как SVM и случайный лес [8, 10]. Предполагается, что построение графовой сверточной нейронной сети позволит улучшить качество классификации, достигнутое в предыдущих исследованиях.

Существующие сверточные нейронные сети, принимающие на вход граф, можно разделить на два класса: основанные на спектральном и пространственном подходе соответственно. В данной работе рассматриваются пространственные методы, поскольку граф,

*Научный руководитель: Стрижов В. В. Задачу поставил: Местецкий Л. М., Райер И. А. Консультант: Жариков И. Н.

сгенерированный по изображению, имеет произвольный вид. **Предлагается развить подход пространственной фильтрации, описанный в работе [7], адаптировав его к решению рассматриваемой задачи.**

2 Название раздела

TODO

2.1 Название параграфа.

TODO

3 Заключение

TODO

Литература

- [1] Dan Claudiu Cireşan, Ueli Meier, Luca Maria Gambardella, and Jürgen Schmidhuber. Convolutional neural network committees for handwritten character classification. In *ICDAR*, pages 1135–1139. IEEE Computer Society, 2011.
- [2] Michael Dirnberger, Adrian Neumann, and Tim Kehl. Nefi: Network extraction from images. *CoRR*, abs/1502.05241, 2015.
- [3] Kazuhisa Fujita. Extract an essential skeleton of a character as a graph from a character image. *CoRR*, abs/1506.05068, 2015.
- [4] Stanislav Klimenko, Leonid Mestetskiy, and A Semenov. Handwritten fonts modeling based on fat curves. *GraphiCon 2006 - International Conference on Computer Graphics and Vision, Proceedings*, 01 2006.
- [5] Leonid Mestetskiy and Andrey Semenov. Binary image skeleton - continuous approach. volume 1, pages 251–258, 01 2008.
- [6] Juan Nunez-Iglesias, Adam Blanch, Oliver Looker, Matthew W. Dixon, and Leann Tilley. A new python library to analyse skeleton images confirms malaria parasite remodelling of the red blood cell membrane skeleton. *PeerJ*, 6:e4312, 02 2018.
- [7] Felipe Petroski Such, Shagan Sah, Miguel Dominguez, Suhas Pillai, Chao Zhang, Andrew Michael, Nathan Cahill, and Raymond Ptucha. Robust spatial filtering with graph convolutional neural networks. *IEEE Journal of Selected Topics in Signal Processing*, PP, 03 2017.
- [8] Местецкий Л.М. Липкина А.Л. Классификация букв в изображениях на основе медиального представления. *GraphiCon 2018: труды 28-й Междунар. конф. по компьютерной графике и машинному зрению*, pages 362–368, 2018.
- [9] Леонид Моисеевич Местецкий. *Непрерывная морфология бинарных изображений: фигуры, скелеты, циркуляры*. Физматлит, 2009.
- [10] Кушнир О.А. Сравнение формы бинарных растровых изображений на основе скелетизации. *Машинное обучение и анализ данных, 2012. Т. 1, No 3.*, pages 252–263, 2018.