

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

И. А. Рейер, В. В. Стрижов, М. Потанин, Д. Ожерелков В. А. Шокоров

author@site.ru; co-author@site.ru; co-author@site.ru

<sup>1</sup>Организация, адрес; <sup>2</sup>Организация, адрес

Исследуется метод распознавания растрового представления изображения символа, с помощью скелетного представления. В качестве решения используется Алгоритм Местецкого для получения графа скелетного представления, затем к нему применяется GCNN для классификации символа.

Наша модель ДОЛЖНА повысить скорость и качества обработки изображения.

Так же рассматривается задача привилегированного обучения, когда происходит совместная минимизация ошибки двух сетей, более сложной, но классической CNN и более легкой GCNN. Целью исследования является анализ изменения весов GCNN. Проведены эксперименты по классификации рукописных цифр.

**Ключевые слова:** *распознавание текста, скелетное представление, GCNN, привилегированное обучение*

## 1 Введение

Мы рассматриваем изображения символа, как двумерную фигуру, ограниченной конечным множеством полигонов. Алгоритм Местецкого, основанный на прямом построении обобщенной триангуляции Делоне множества граничных элементов фигуры позволяет нам получить скелетное графовое представление данного символа.

Цель исследования уменьшить вычислительную сложность алгоритма относительно классического подхода решения задачи.

Мы рассматриваем изображения символа, как двумерную фигуру, ограниченной конечным множеством полигонов, полученной из бинаризации изображения (В нашем случае под бинаризацией понимается следующее: все пиксели, цвет которых НЕ черный (то есть  $> 0$  (либо можно установить `threshold`)) в нотации цветов от 0 до 255), становятся белыми)

Алгоритм Местецкого[], основанный на прямом построении обобщенной триангуляции Делоне множества граничных элементов фигуры позволяет нам получить скелетное графовое представление данного символа. Поэтому исследовательская часть нашей работы заключается в том, чтобы научиться качественно преобразовывать графовое представление символа в вектор, для дальнейшего решения задачи классификации. Сложность данного подхода получается из нерегулярности структуры графа, объекты, подлежащие обработке, не упорядочены и могут иметь произвольную размерность и структуру связей. Обстоятельное исследование этой проблемы и возможных подходов к её решению можно найти в статье [1]

Кроме канонического решения задачи обработки изображения CNN, похожая проблема решается в [2] там используется набор контуров и граф скелетного представления с последующим линейным классификатором SVM. Так же существует подход использующий медианное представление [3]. Существует большой набор статей Липкиной А.Л. и Местецкого Л.М. посвященных решению нашей задачи, например [4] [5].

Мы хотим найти оптимальные параметры графового представления (таких как координаты вершин, их валентность, степень, толщина линий и тд.) описывающие символ, для их подбора можно воспользоваться например [6].

## Литература

- [1] Y. LeCun A. Szlam P. Vandergheynst M. Bronstein, J. Bruna. Geometric deep learning: going beyond euclidean data. *Signal Processing Magazine*, 34(4):18–42, 2017.
- [2] Yao C. Bai X Shen W., Wang X. Shape recognition by combining contour and skeleton into a mid-level representation. *Communications in Computer and Information Science*, 483.
- [3] Arseev Sergey Lomov Nikita. Neural networks for shape recognition by medial representation. *3rd International ISPRS Workshop on PSBB Moscow*.
- [4] Л.М. Местецкий А.Л. Липкина. Классификация букв в изображениях на основе медиального представления. *Труды международной конференции по компьютерной графике и зрению "Графикон"*.  
.
- [5] Л.М. Местецкий А.Л. Липкина. Структурный подход к классификации букв в изображениях. *Труды международной конференции по компьютерной графике и зрению "Графикон"*.  
.
- [6] С.Ю. Желтов Ю.В. Визильтер, В.С. Горбацевич. Структурно-функциональный анализ и синтез глубоких конволюционных нейронных сетей. *Компьютерная оптика*, 43(5):886–900, 2019.

*Поступила в редакцию*