

Верификация авторства арабских рукописей при помощи глубоких свёрточных сетей

Андрей Бояров, Александр Сенов

Санкт-Петербургский Государственный Университет
Математико-Механический Факультет

25 февраля 2016 г.



Задача

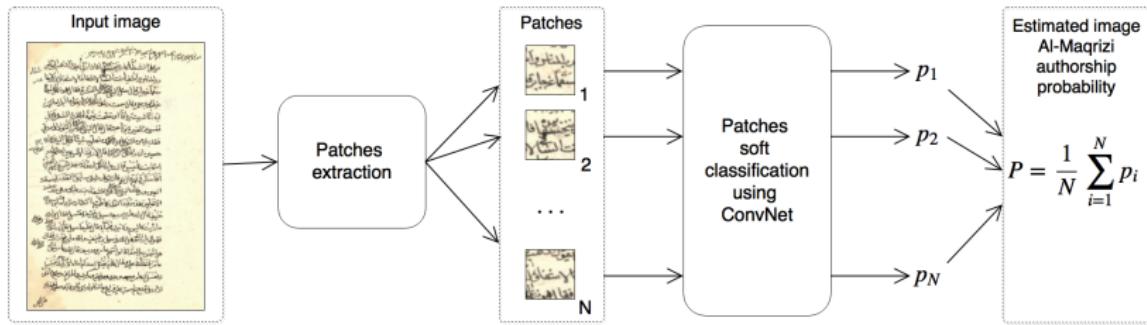
- Верификация авторства аль-Макризи¹ недавно найденной части рукописи *al-Khitat* (историко-географический обзор Египта).

¹Аль-Макризи (1364–1442) — египетский историк и географ периода мамлюков.

Данные

- Тренировочные: 1 манускрипт аль-Макризи, состоящий из 26 страниц и 5 не аль-Макризи манускриптов, по 7 страниц каждый.
- Тестовые: 1 манускрипт аль-Макризи, состоящий из 14 страниц и 3 не аль-Макризи манускриптов, по 7 страниц каждый. Авторы этих трёх манускриптов отличались от авторов пяти манускриптов из тренировочных данных.
- Манускрипт *al-Khitat* состоит из 32 страниц.

Метод



Извлечение патчей

- Скользящее окно: разбиение изображения на квадраты 80×80 с шагом в 20 пикселей.
- Связные компоненты:
 1. Бинаризация изображение с помощью фильтра Otsu².
 2. Выделение компонент связности из бинаризованного изображения.
 3. Фильтрация компонент на основе их размеров
 4. Выделение компонент аутлаеров с помощью алгоритма DBSCAN³.
 5. Извлечение AABB⁴ компонент связности и преобразование их в квадраты 28×28 пикселей.

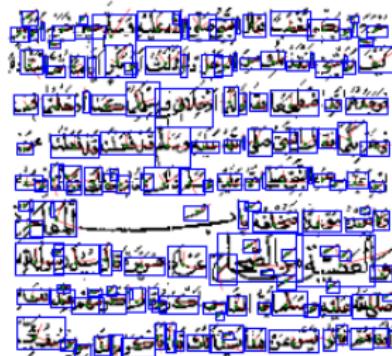
²https://en.wikipedia.org/wiki/Otsu's_method. Реализация в scikit-image

³<https://en.wikipedia.org/wiki/DBSCAN>. Реализация в scikit-learn

⁴Axis-aligned minimum bounding box

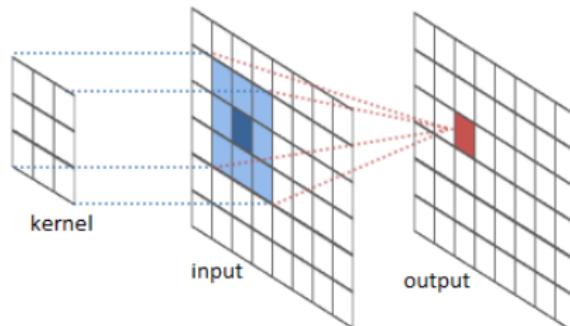
Скользящее окно и связные компоненты

عمر عبد شمس عز الدين
خنزير قفال سودان
بهرمان يحيى الدين
علي صدر حاتم سودان



Глубокие свёрточные сети

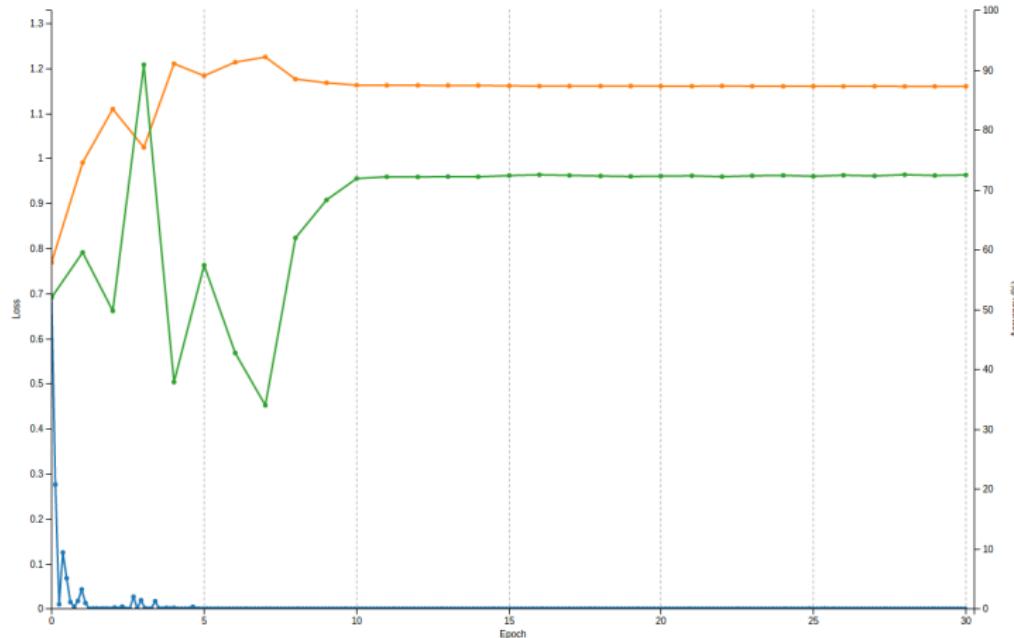
- State of the art в задачах компьютерного зрения (с 2012 года).
- Представитель класса алгоритмов глубокого обучения.
- Представляет собой набор свёрточных, pooling, LRN, полносвязных слоёв. В качестве функции активации используется ReLU: $f(x) = \max(0, x)$.



Скользящее окно, Alexnet

тип	фильтр / шаг	выход
convolution	11 / 4	96
local response norm		96
max pool	3 / 2	96
convolution	5 / 1	256
local response norm		256
max pool	3 / 2	256
convolution	3 / 1	384
convolution	3 / 1	384
convolution	3 / 1	256
max pool	3 / 2	256
fully connected		4096
dropout (50 %)		4096
fully connected		4096
dropout (50 %)		4096
fully connected		2
softmax		2

Скользящее окно, Alexnet, обучение



Связные компоненты, GoogLeNet

тип	фильтр / шаг	выход
convolution	7 / 2	64
max pool	3 / 2	64
local response norm		64
convolution	1 / 1	64
convolution	3 / 1	192
local response norm		192
max pool	3 / 2	192
inception		256
inception		480
max pool	3 / 2	480
inception		512
convolution	1 / 1	128
fully connected		1024
dropout (70 %)		1024
fully connected		2
softmax		2

Результаты

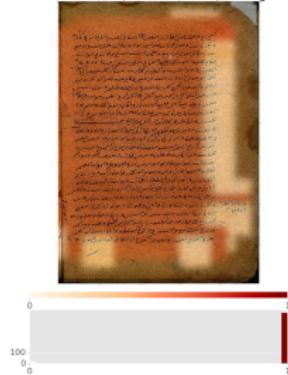
- На тестовом множестве для метода скользящего окна: порог = 0.5, precision = 0.99, recall = 0.92.
- Метод связных компонент менее робастный: он даёт большое число ложно-позитивных результатов.
- Средняя вероятность по страницам того, что манускрипт *al-Khitat* принадлежит перу аль-Макризи составляет 0.86.

Результаты

Estimated Al-Maqrizi authorship probability
Minimum: 4.1e-07, Maximum: 0.98, Average: 0.0085



Estimated Al-Maqrizi authorship probability
Minimum: 5.1e-05, Maximum: 1.0, Average: 0.94



Estimated Al-Maqrizi autorship probability
Minimum: 3.3e-06, Maximum 1.0, Average: 0.77

