

Поиск символов в художественных изображениях*

Лемтюжникова Д. В., Апишев М. А., Козинов А. В.

kozinov.av@phystech.edu

Рассматривается проблема распознавания художественного изображения, содержащего символы, в зависимости от контекста. Символ можно распознать многими способами, причём в каждом отдельном случае значение выбирается в зависимости от контекста. Контекст связан не только с окружающими фрагментами на картине, но и с историко-культурными событиями, связанными с изображением.

Ключевые слова: *Анализ изображения, CNN, boxing изображений.*

1 Введение

В данной работе рассматривается задача понимания художественных изображений алгоритмами машинного обучения. Основная цель - это распознать на изображении ключевые символы, а с помощью информации о них и информации об изображении сформировать текстовое описание.

Стоит определиться с тем, что такое **символ**. Существует много определений понятия символ, одно из них следующее: "Символ имеет очень сложное значение, потому что не подчиняется причине; он всегда предполагает много значений, и эта многозначность не может быть сведена к единой логической системе"(В.И. Иванов). И принято решение определять символ с помощью экспертов. Таким образом сформирована база размеченных изображений с выделенными фрагментам, которым сопоставлено название символа и его значение при данном контексте.

На предложенной выборке производится обучение свёрточной нейронной сети. А далее процесс анализа изображения происходит следующим образом: выделяются основные объекты и фон, производится классификация полученных объектов, далее для каждого из объекта выбирается описание на основе того, как элементы связаны друг с другом и с фоном.

Это решение может быть использовано для оценки стоимости картины перед аукционом. Но в отличие от подхода [3], который анализирует картину целиком, представленный подход учитывает наличие специальных смысловых единиц — символов.

* Научный руководитель: Стрижов В. В. Задачу поставила: Лемтюжникова Д. В.

В данной статье рассматривается упрощённая формулировка задачи, в которой нужно определить, присутствует ли символ на изображении или нет.

2 Постановка задачи

2.1 Входные данные

На вход подаётся *RGB* изображение I размера $H \times W$. То есть I — матрица размера $H \times W \times 3$, причём $I_{i,j,k} \in \overline{0, 255}$. Размер изображения заранее не фиксируется и H , и W могут принимать разные значения.

2.2 Выходные данные

Для каждого класса следует выделить все фрагменты изображения, на которых изображены объекты соответствующего класса. А именно получить список множеств B_1, \dots, B_N , где $B_i = \{BB_{gt,1}^{(i)}, \dots, BB_{gt,a_i}^{(i)}\}$ — множество найденных *Bounding boxes*.

2.3 Качество решения

Пусть же набор множеств B'_1, \dots, B'_N , где $B'_i = \{BB_{l,1}^{(i)}, \dots, BB_{l,b_i}^{(i)}\}$ — истинный набор *Bounding boxes* ([1]). Для определения расстояния между двумя *Bounding boxes*, используя метрику отношения пересечения к объединению ([2])

$$IoU(B_l, B_{gt}) = \frac{Area(B_l \cap B_{gt})}{Area(B_l \cup B_{gt})}$$

Тогда качество предсказания класса i будем вычислять следующим образом:

$$IoU_i = \frac{1}{b_i} \sum_{j=1}^{b_i} \min_{k=1 \dots a_i} IoU(B_{l,j}^{(i)}, B_{gt,k}^{(i)})$$

И для подсчёта общего качества решения достаточно просуммировать по всем классам:

$$IoU = \frac{1}{N} \sum_{i=1}^N IoU_i$$

Литература

- [1] Bruno Emile Christophe Rosenberger Baptiste Hemery, Hélène Laurent. Comparative study of localization metrics for the evaluation of image interpretation systems. 2013.
- [2] Jonathan Hui. map (mean average precision) for object detection. 2018.
- [3] Vidush Mukund Rafi Ayub, Cedric Orban. Art appraisal using convolutional neural networks. 2017.