

Распознавание медицинских изображений в задачах компьютерной диагностики

Р.В. КОЗАРЬ, А.А. НАВРОЦКИЙ, А.Б. ГУРИНОВИЧ

В данной работе предлагается модификация алгоритма Виолы-Джонса для распознавания изображений, полученных методом оптической эндоскопией, который повышает точность распознавания. В данной работе представлены результаты проведенных тестов, которые доказывают эффективность и применимость данной методики.

Ключевые слова: эндоскопические снимки, Парето, распознавание, алгоритм Виолы-Джонса.

This paper proposes a modification of the Viola-Jones algorithm for recognizing images obtained by optical endoscopy, which improves the recognition accuracy. This paper presents the results of the conducted tests, which prove the effectiveness and applicability of this technique.

Keywords: endoscopic images, Pareto, recognition, Viola-Jones algorithm.

Введение. В наши дни существует огромное количество методов, подходов, алгоритмов для распознавания изображений, а также объектов на них. Один из них – это метод Виолы-Джонса. Важно отметить, что почти все существующие подобные методы основаны на одном принципе: калькуляция свертки с определенным паттерном.

В момент распознавания изображения с медицинского эндоскопа выше описанный подход ведет к тому, что существенно возрастают вычислительные объемы во время увеличения размеров самой области распознаваемого изображения. Для данного случая – это медицинское изображение. Поэтому в системах медицинской компьютерной диагностики применение подобного подхода малоэффективно.

Стоит отметить, что алгоритм Виолы-Джонса – это один из самых распространенных алгоритмов распознавания снимков, а также различных объектов на распознаваемых изображениях. Основа данного алгоритма – это так называемое сканирующее окно. Однако есть нюанс: использование данного «окна» может резко привести к увеличению нагрузки на алгоритм и как следствие – на физическое оборудование. Также использование данного окна приводит к увеличению самого снимка. Итог: негативные последствия по качеству распознавания и временном периоде, за который это распознавание осуществляется.

В текущей работе предлагается альтернативный способ решения проблемы, описанной выше. В работе были проанализированы результаты опытов, различных тестов и экспериментов, собранных учеными. Было рассмотрено около 50–55 источников: а именно кандидатские и докторские диссертации и их авторефераты, тезисы, статьи, а также книги, материалы научных конференций.

Описание метода распознавания изображений Виолы-Джонса. Метод, о котором уже упомянуто выше, основан на следующих положениях:

- все распознаваемые изображения приводятся в интегральное представление;
- при распознавании используются признаки (каскады) Хаара;
- в работе алгоритма присутствует и используется так называемый бустинг;
- все признаки передаются на вход классификатору, который, в свою очередь, возвращает либо положительный, либо отрицательный результат (истина или ложь);
- в работе алгоритма присутствуют и также используются каскады признаков.

Под классификацией мы понимаем моделирование, основная цель которого – разделение классов данных на различные множества [1, с. 40].

Также необходимо привести изображения в интегральный вид для того, чтобы можно было использовать так называемые каскады (признаки) Хаара.

Эти признаки нужны для нахождения необходимого объекта [2, с. 21]. Также применяется бустинг: это метаалгоритм машинного обучения, необходимый для уменьшения погрешности оценки. Он нужен для выбора наиболее подходящих признаков, которые более присущи стартовому, изначальному объекту на данной части изображения (в нашем случае – изображения, полученного с эндоскопической камеры).

Несколько признаков объединяются в каскады, которые нужны для того, чтобы как можно быстрее отклонить те сканирующие окна, в которых не найдена необходимая часть изображения, которая должна быть распознана. Также отметим, что те части (области) изображения, которые прошли через весь каскад, могут быть классифицированы только тогда, когда все прецеденты были классифицированы корректно и правильно [3, с. 21].

Анализ существующих подходов и решений. В наши дни существует несколько модификаций данного алгоритма, но, к сожалению, для медицинских изображений такие модификации отсутствуют. Их отсутствие объясняется тем, что изображения подобного рода весьма нестандартные. Причин тому может быть множество (например, невысокое разрешение или малый контраст). Также, для того, чтобы использовать те самые решения в области медицинской диагностики, необходимы значительные инженерные усилия. Поэтому оптимизация сводится к тому, чтобы наиболее оптимально реализовать алгоритм на заданной платформе, так как при случайных последующих выборах мы будем наблюдать изменения во всех параметрах оценки распознавания [4, с. 119].

В трудах других авторов можно найти решения на основе реализации из библиотеки *OpenCV*. *OpenCV* (*Open Source Computer Vision Library*) – это библиотека машинного зрения с открытым исходным кодом, библиотека методов компьютерного зрения, а также обработки изображений, численных алгоритмов общего назначения с открытым кодом. В этой библиотеке находится более двух тысяч методов, среди которых как классические, так и современные алгоритмы для машинного зрения и машинного обучения. Тем не менее сама реализация изменений практически не претерпевает. Для того, чтобы сравнить результаты распознавания обычно применяют так называемые статистические критерии [5, с. 375].

Имеются также подходы к решению данной проблемы, они основаны на виртуальном (визуальном поиске). Эта методика поиска используется для того, чтобы как можно быстрее распознать объект на изображении.

Модификация алгоритма была сравнена с алгоритмом + детектором Виолы-Джонса, который описан в библиотеке *OpenCV*. В итоге имеем следующее: выигрыш в ускорении в среднем 2,5 раза. Однако точность распознавания пострадала и стала немного ниже. Но для эндоскопических снимков такой подход абсолютно неприемлемый: специфика области вносит свои коррективы, поэтому распознавание должно работать абсолютно точно.

Модификация алгоритма Виолы-Джонса для эндоскопических снимков. В оригинальном алгоритме Виолы-Джонса окно сканирования двигается с шагом Δ . Размеры шага по осям ox и oy обозначим следующим образом (1):

$$\Delta x, \Delta y. \quad (1)$$

Если размеры шагов Δ равны (2), то по-прежнему применим обозначение Δ .

$$\Delta x = \Delta y. \quad (2)$$

Шаг сканирования – крайне важный показатель. Он влияет и на четкость, и на точность, и на качество, и даже на пропускную способность [6, с. 48]. Однако в однородных областях будет наблюдаться следующая зависимость: если увеличивается Δ – вероятность ложного срабатывания уменьшится. Также можно ускорить метод за счет увеличения Δ в тех областях снимков, которые не содержат искомый объект.

В разработанных реализациях Δ выбирают в промежутке $\{1; 2\}$.

Введем новое понятие – степень выхода: индекс классификатора в каскаде, при попадании на который мы исключаем сканирующее окно. В ходе анализа работы алгоритма была выявлена следующая взаимосвязь: распознаваемая область и степень выхода обратно пропорциональны. Рисунок 1 это иллюстрирует. Он показывает обратно пропорциональную зависимость яркости пикселя и степени выхода.

Этот подход дает право увеличить Δ и при минимальной степени выхода, и при приближении к объекту распознавания. Этим мы позволяем избегать лишних затрат физического оборудования и вычислительных мощностей на нормализацию окна поиска [7, с. 5]. Это необходимо для минимизации влияния иных внешних факторов (лишнее освещение и т. д.).

Для того, чтобы все наглядно оценить, были проанализированы изменения Δ при распознавании. Рисунок 2 иллюстрирует изображение постоянного шага $\Delta x = 2$. Из выше представленных иллюстраций прослеживается следующее: неадаптированный детектор Δ не может поместить окно в локальный максимум. А модифицированный шаг Δ – может.

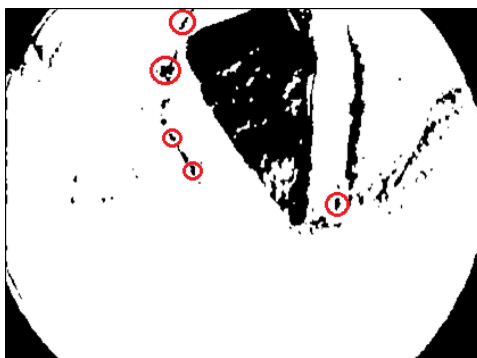


Рисунок 1 – Сканирующее окно на равномерных участках отвергается ранее (при условии, что отсутствует искомый паттерн). При приближении к необходимой области ступень выхода начинает возрастать



Рисунок 2 – Детальное изображение постоянного шага $\Delta x = 2$ при распознавании изображения

Рисунок 3 иллюстрирует детально рассмотренные постоянный шаг $\Delta x = 3$.

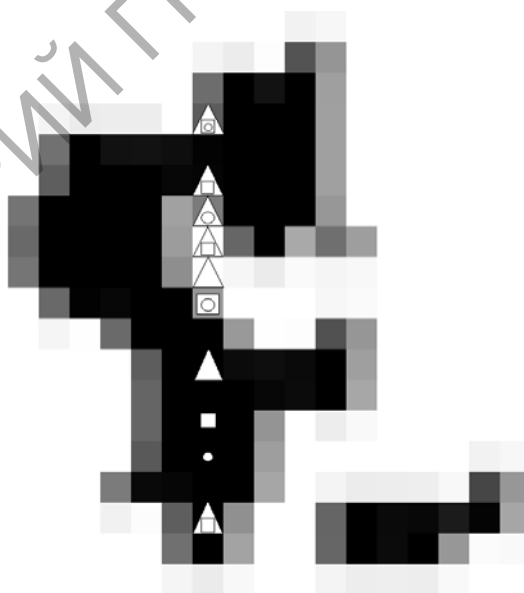


Рисунок 3 – Детальное изображение постоянного шага $\Delta x = 3$ при распознавании изображения

На представленных изображениях рассматриваются самые разные варианты шага Δ . Блоки серого цвета – пиксели от 207 до 216 (в увеличенном масштабе) по оси ox и от 147 до 125 по оси oy . Пиксель черного цвета (также увеличен масштаб) в позиции (212, 120) это правая нижняя часть снимка, представленного на рисунке 1.

Анализ проведенных исследований, результаты. В ходе данной работы для сравнения и анализа были реализованы следующие подходы:

- статический: Δx и Δy неизменны, равны 1, 2 или 3 соответственно;
- *OpenCV* вариант 1: $\Delta x = 2$ для всего распознаваемого эндоскопического снимка. При не обнаружении нужной области Δx уменьшается на 1, а Δy остается равной 1;
- *OpenCV* вариант 2: $\Delta x = 1$, если соотношение всех размеров оригинального снимка и уменьшенного более, чем на 2, иначе $\Delta x = 2$.

Необходимо отметить также, что при проведении экспериментов использовались так называемые масштабирующие коэффициенты (*scaling factor*) s и пороги слияния (*merging threshold*) γ , которые удовлетворяют следующим условиям (3).

$$\begin{aligned} s &\in \{1.1, 1.2, 1.3, 1.4, 1.5\} \\ \gamma &\in \{1, 2, 3, 4, 5\} \end{aligned} \quad (3)$$

Для сравнения результатов были взяты и использованы метрики, описанные следующими выражениями (4).

$$\begin{aligned} recall &= \frac{TP}{TP + EN} \\ precision &= \frac{TP}{TP + FP} \end{aligned} \quad (4)$$

где TP – количество корректно определенных объектов, EN – количество пропущенных объектов, FP – количество некорректных (ложных) срабатываний.

Все полученные результаты первого тестирования сравним по Парето и проиллюстрируем их на рисунке 4.

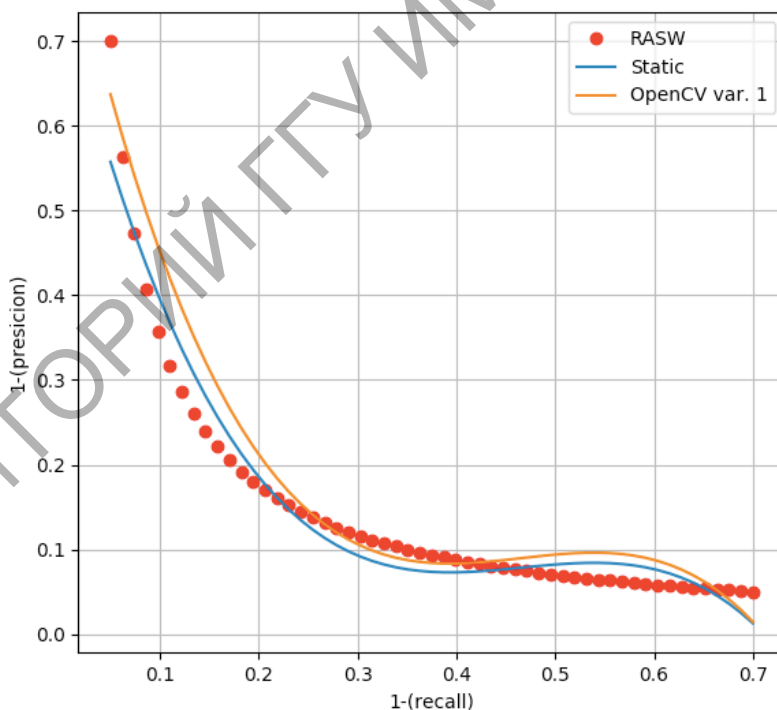
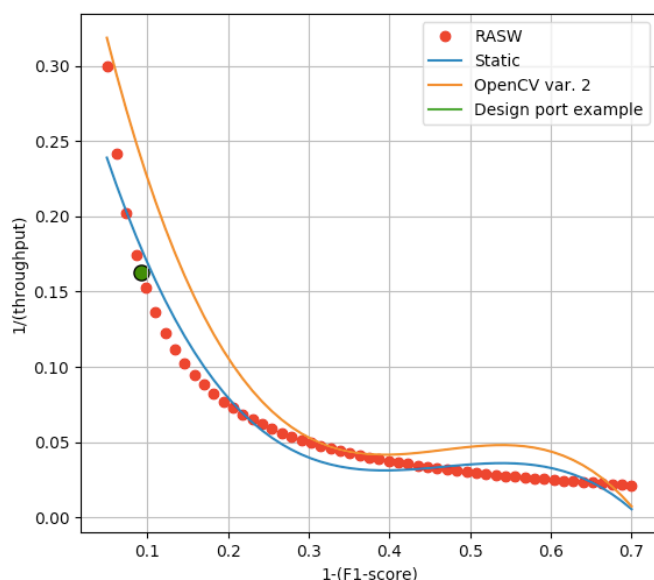


Рисунок 4 – Результаты первого тестирования: метрики – отзыв распознавания и точность распознавания

Для анализа введем следующую метрику $F1$, которая, по сути, означает средневзвешенное значение отзыва и точности распознавания. Формула метрики представлена ниже (5):

$$F1 = 2 \times \frac{precision \times recall}{precision + recall} \quad (5)$$

Результаты по данной метрике также сравнены по Парето и проиллюстрированы на рисунке 5.

Рисунок 5 – Результаты второго тестирования: метрика $F1$

Выводы. В данной статье был разработан и предложен новый подход для адаптированного сканирующего окна с целью улучшения процесса распознавания объекта при использовании метода Виолы-Джонса, а также результаты его использования и сравнения с существующими методиками распознавания изображений.

Также были представлены результаты для реализации нового детектора распознавания эндоскопических снимков, который можно использовать в любом методе распознавания. Главный критерий – это наличие окна сканирования и каскадного классификатора (для использования признаков Хаара).

Если сравнить модифицированный алгоритм с существующими решениями, методами и подходами – то он дает гораздо более лучшие итоговые показатели (по метрикам $F1$ -score (комбинации *recall*, *precision*)).

И это чрезвычайно важно при распознавании эндоскопических снимков, так как при использовании данной модификации были приняты во внимание критически важные параметры поступающих на вход изображений, такие как неоднородность, с целью большего снижения объема вычислительных нагрузок без потери качества распознавания.

Литература

1. Паттерсон, Дж. Глубокое обучение с точки зрения практика / Дж. Паттерсон, А. Гибсон. – М. : ДМК Пресс, 2018. – 418 с.
2. Brousseau, В. An energy-efficient, fast FPGA hardware architecture for OpenCV compatible object detection / В. Brousseau, J. Rose. – Toronto : Toronto FPT Univ. Press, 2012. – P. 20–21.
3. Гудфеллоу, Я. Глубокое обучение / Я. Гудфеллоу, И. Бенджио, А. Курвилль. – М. : «ДМК Пресс», 2018. – 652 с.
4. Шапиро, Л. Компьютерное зрение / Л. Шапиро, Д. Стокман. – М. : Изд. Дом «Лаборатория знаний», 2015. – 763 с.
5. Местецкий, Л. М. Математические методы распознавания образов / Л. М. Местецкий. – М. : МГУ, ВМиК, 2002–2004. – С. 20–24.
6. Hiromoto, M. Partially parallel architecture for Adaboost based detection with haar-like features / M. Hiromoto, H. Sugano, R. Miyamoto // IEEE Trans. Circuits Syst. Video Technol. – 2009. – Т. 19. – P. 41–52.
7. Lampert, C. Beyond sliding windows : object localization by efficient subwindow search / C. Lampert, M. Blaschko, T. Hofmann. – Anchorage : Anchorage CVPR Press, 2008. – P. 1–8.