



<http://dx.doi.org/10.35596/1729-7648-2023-21-1-94-97>

Оригинальная статья
Original paper

УДК 004.932.2

МЕТОДЫ КЛАСТЕРИЗАЦИИ ДАННЫХ ПРИ РАСПОЗНАВАНИИ ЭНДОСКОПИЧЕСКИХ СНИМКОВ В ЗАДАЧАХ КОМПЬЮТЕРНОЙ МЕДИЦИНСКОЙ ДИАГНОСТИКИ

Р. В. КОЗАРЬ, Н. С. КОНОЙКО, А. А. НАВРОЦКИЙ

*Белорусский государственный университет информатики и радиоэлектроники
(г. Минск, Республика Беларусь)*

Поступила в редакцию 03.11.2022

© Белорусский государственный университет информатики и радиоэлектроники, 2023
Belarusian State University of Informatics and Radioelectronics, 2023

Аннотация. Представлены результаты анализа существующих методов кластеризации данных, получаемых при эндоскопии гортани. Предложена модификация метода Виолы-Джонса для распознавания изображений с использованием критерия гибкого выхода. Метод Виолы-Джонса исследует все области на изображении и принимает решение о принадлежности распознаваемой области к искомой путем прохождения через классифицированный каскад. Эндоскопические снимки имеют большое количество особенностей, таких как засветка, шумы и другие, которые ухудшают качество распознавания. Для улучшения качества распознавания предложена кластеризация с критерием гибкого выхода, которая удовлетворяет критериям масштабируемости – изменение решения вместо перехода к другой области распознавания. Установлено, что предложенная модификация метода Виолы-Джонса показывает высокие результаты распознавания для эндоскопических снимков.

Ключевые слова: медицинский снимок, критерий, кластеризация, распознавание, метод Виолы-Джонса, классификатор, Хаар, бустинг, вейвлет, алгоритм.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Для цитирования. Козарь, Р. В. Методы кластеризации данных при распознавании эндоскопических снимков в задачах компьютерной медицинской диагностики / Р. В. Козарь, Н. С. Конойко, А. А. Навроцкий // Доклады БГУИР. 2023. Т. 21, № 1. С. 94–97. <http://dx.doi.org/10.35596/1729-7648-2023-21-1-94-97>.

DATA CLUSTERING METHODS FOR RECOGNITION OF ENDOSCOPIC IMAGES IN THE PROBLEMS OF COMPUTER MEDICAL DIAGNOSIS

RAMAN V. KOZAR, NATALIA S. KONOIKO, ANATOLIY A. NAVROTSKY

Belarusian State University of Informatics and Radioelectronics (Minsk, Republic of Belarus)

Submitted 03.11.2022

Abstract. This paper presents the results of the analysis of existing methods for clustering data obtained during endoscopy of a larynx. A modification of the Viola-Jones method for image recognition using the flexible exit criterion is proposed. The Viola-Jones method explores all areas in the image and decides whether the recognized area belongs to the desired one by passing through a classified cascade. Endoscopic images have a large number of features, such as flare, noise, etc., which degrade the quality of recognition. To improve the quality of recognition, clustering with a flexible exit criterion was proposed, which satisfies the scalability criteria: changing the decision of the solution, instead of moving to another recognition area. It has been established that the proposed modification of the Viola-Jones method shows higher recognition results for endoscopic images.

Keywords: medical image, criteria, clustering, recognition, Viola-Jones method, classifier, Haar, boosting, wavelet, algorithm.

Conflict of interests. The authors declare no conflict of interests.

For citation. Kozar R. V., Konoiko N. S., Navrotsky A. A. (2023) Data Clustering Methods for Recognition of Endoscopic Images in the Problems of Computer Medical Diagnosis. *Doklady BGUIR*. 21 (1), 94–97. <http://dx.doi.org/10.35596/1729-7648-2023-21-1-94-97> (in Russian).

Введение

В настоящее время остается актуальной задача распознавания данных с изображений для последующего принятия решений. Правильность принимаемых решений в значительной степени зависит от качества классификации распознаваемого объекта. Подходы к классификации разнятся в зависимости от вида изображений. В статье рассматриваются особенности классификации объектов на снимках, получаемых методом оптической эндоскопии. Следует отметить, что хорошо себя зарекомендовал метод Виолы-Джонса. Основная его особенность заключается в выделении примитивных признаков каждой распознаваемой области через вейвлеты Хаара, которые крайне эффективны [1, с. 145]. Данный примитив представляет собой трехмерный вектор типа маска-позиция-размер. Каждый вейвлет содержит светлую и темную области, обладает определенными пропорциями, отличается от других вейвлетов по размерам в зависимости от конкретной цели распознавания. Обучение происходит стандартно, цель – ассоциация каждого образца с самим собой [2, с. 121].

При распознавании эндоскопических снимков важно сохранение максимально большого числа деталей, влияющих на принимаемое решение. Во взятом за основу методе Виолы-Джонса распознавание происходит по принципу кадрирования. Искомый объект может иметь разный размер, поэтому необходимо выполнять поиск с использованием признаков различного размера. Применение метода Виолы-Джонса при распознавании медицинских изображений имеет такие недостатки, как крайне длительное время обучения и сложности с обнаружением значительного количества небольших объектов, расположенных близко друг от друга.

Модификация алгоритма при помощи критерия гибкого выхода

В оригинальном методе Виолы-Джонса решение о наличии основного объекта в кадрируемом окне принимается крайне однозначно, что недопустимо при работе с эндоскопическими снимками. Решение производится по следующей формуле:

$$H(x) = 1, \text{ если } \frac{1}{2} \sum_{t=1}^T a_t \leq \sum_{t=1}^T a_t h_j^{(t)}(x). \quad (1)$$

Предлагается модификация алгоритма, которая заключается в комбинации «однозначного» и «нечеткого» подходов. В результате будет принято «неоднозначное» решение, которое позволит улучшить распознавание подобного рода снимков. Установлено, что при распознавании эндоскопических снимков однозначность вела к тому, что часто принималось решение об отсутствии объекта на текущем кадрируемом окне, несмотря на тот факт, что объект на самом деле присутствовал. Использование намеренной «неточной» оценки позволяет классификатору принимать решение, отличное от однозначного. Как следствие – объект может быть распознан корректно. Для решения этой задачи требуется сильный классификатор. Поэтому формула (1) приводится к следующему виду:

$$H_{\text{modifying}}(x) = 1 \leq 2 \frac{\sum_{t=1}^T a^{(t)} h_j^{(t)}}{\sum_{t=1}^T a^{(t)}}. \quad (2)$$

Также предлагается новый критерий – критерий гибкого выхода, описываемый выражением:

$$H_{\text{flex}}(x) = \left\{ 2 \frac{\sum_{t=1}^T a^{(t)} h_j^{(t)}}{\sum_{t=1}^T a^{(t)}} \right\}. \quad (3)$$

Предложенный критерий, исходя из значений $H_{\text{flex}}(x)$, позволяет дать однозначный ответ на вопрос о нахождении искомого объекта на распознаваемом снимке. На рис. 1 представлены результаты распознавания с помощью алгоритма, использующего критерий гибкого выхода, полученные в процессе двух опытов.

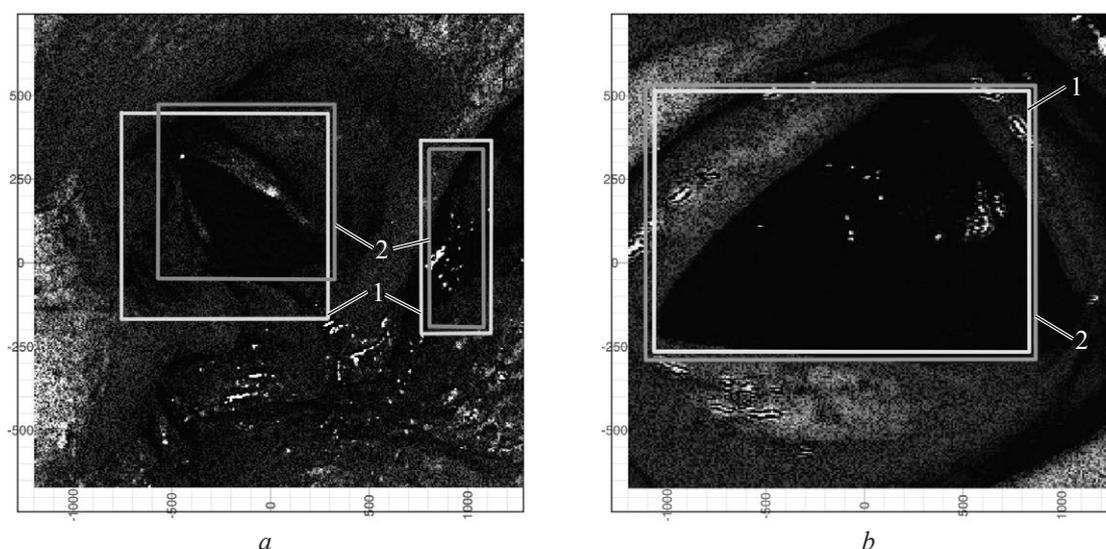


Рис. 1. Результаты распознавания с помощью критерия гибкого выхода области (1) и подобласти (2):
a – опыт № 1; *b* – опыт № 2

Fig. 1. Recognition results using the region (1) and subregion (2) flexible output criterion: *a* – test No 1; *b* – test No 2

На рис. 1 выделены области 1 и подобласти 2. Каждая подобласть – результат работы методики гибкого выхода. Область с большим значением гибкого выхода имеет достаточно хорошее срабатывание и может быть классифицирована в качестве искомого объекта. Полученные результаты демонстрируют, что наибольшее значение гибкого выхода в распознаваемой области является искомым объектом распознавания эндоскопического снимка.

Для повышения качества распознавания использовалась разделительная кластеризация с помощью плотностного метода пространственной кластеризации с присутствием шумов *DBSCAN* [3, с. 438]. Такой подход обусловлен следующим: он масштабируемый, у него минимальные требования к начальным данным о природе объектов, а также он способен отбрасывать шумы при распознавании, поскольку в модели линейной регрессии шум распределен нормально с центром в нуле [4, с. 56]. Применение кластеризации в данном случае достаточно эффективно потому, что при уменьшении порога срабатывания метода увеличивается значение гибкого выхода, что приводит к повышению его точности, и как следствие – к эффективности. Влияние порога распознавания на корректность распознавания представлено в табл. 1.

Таблица 1. Вероятность некорректного распознавания в зависимости от порога распознавания
Table 1. Probability of incorrect recognition depending on the recognition threshold

Порог алгоритма распознавания / Recognition algorithm threshold	0,83	0,85	0,87	0,89	0,91	0,94	0,96	0,97	0,99	1,00
Вероятность ложного срабатывания, % / False positive probability, %	97	76	20	17	16	13	7	3	1	0

Результаты исследований и их обсуждение

Анализ результатов исследований производился с использованием 200 снимков. Сравнивали алгоритм с гибким выходом с немодифицированным методом распознавания Виолы-Джонса. Установлено, что модифицированный метод с применением гибкого выхода и кластеризацией дает при распознавании эндоскопических снимков лучшие результаты. Параметры для проведения кластеризации, а также пороговые значения распознавания с использованием критерия гибкого выхода работают абсолютно одинаково во всех выборках данных, что свидетельствует об инвариантности модифицированного алгоритма к рассматриваемым исходным данным.

Заключение

1. Исследована эффективность метода Виолы-Джонса для распознавания эндоскопических снимков. Установлено, что использование комбинации критерия гибкого выхода и кластеризации методом пространственной бустерной кластеризации дает однозначный ответ о принадлежности области распознавания к объекту распознавания.

2. Предложена модификация метода Виолы-Джонса, использующая критерий гибкого выхода и кластеризацию. В процессе проведения экспериментов данная модификация продемонстрировала высокие результаты при распознавании эндоскопических снимков, что на 43 % эффективнее разрешает ошибки первого рода и на 62 % – второго рода.

Список литературы

1. Viola, P. Robust Real Time Face Detection / P. Viola, M. J. Jones // International Journal of Computer Vision. 2004. Vol. 2, No 57. P. 137–154. DOI: 10.1023/B:VISI.0000013087.49260.
2. Каллан, Р. Основные концепции нейронных сетей / Р. Каллан. М.: Вильямс, 2003.
3. Форсайт, Д. Компьютерное зрение. Современный подход / Д. Форсайт, Ж. Понс. М.: Вильямс, 2004.
4. Николенко, С. И. Глубокое обучение / С. И. Николенко, А. А. Кадурич, Е. О. Архангельская. СПб.: Питер, 2018.

References

1. Viola P., Jones M. J. (2004) Robust Real Time Face Detection. *International Journal of Computer Vision*. 2 (57), 137–154. DOI: 10.1023/B:VISI.0000013087.49260.
2. Kallan R. (2003) *Basic Concepts of Neural Networks*. Moscow, Vyllyams Publ. (in Russian).
3. Forsait D., Pons G. (2004) *Computer Vision. Modern Approach*. Mocsow, Vyllyams Publ. (in Russian).
4. Nikolenko S. I., Kadurin A. A., Arkhangelskaya E. O. (2018) *Deep Learning*. Saint-Petersburg, Piter Publ. (in Russian).

Вклад авторов

Козарь Р. В. осуществил разработку модифицированного алгоритма, провел необходимые исследования, подготовил рукопись статьи.

Конойко Н. С. осуществляла подготовку эндоскопических снимков, принимала участие в обсуждении результатов.

Навроцкий А. А. осуществил постановку задачи, вносил предложения по модификации разработанного алгоритма.

Authors' contribution

Kozar R. V. carried out the development of a modified algorithm, conducted the necessary research, prepared the manuscript of the article.

Konoiko N. S. provided data of endoscopic images for research.

Navrotsky A. A. carried out the formulation of the problem, made proposals for modifying the developed algorithm.

Сведения об авторах

Козарь Р. В., аспирант кафедры информационных технологий автоматизированных систем Белорусского государственного университета информатики и радиоэлектроники

Конойко Н. С., врач оториноларинголог-фониатр, заведующая фониатрическим отделением Республиканского научно-практического центра оториноларингологии

Навроцкий А. А., к. ф.-м. н., доцент, заведующий кафедрой информационных технологий автоматизированных систем Белорусского государственного университета информатики и радиоэлектроники

Адрес для корреспонденции

246050, Республика Беларусь,
г. Гомель, ул. Крестьянская, д. 35, кв. 12
Тел.: +375 29 730-13-80
E-mail: pozitron.kozarroman@gmail.com
Козарь Роман Вячеславович

Information about the authors

Kozar R. V., Postgraduate at the Department of Information Technologies of Automated Systems of the Belarusian State University of Informatics and Radioelectronics

Konoiko N. S., Otorhinolaryngologist-phoniatriст, Head of the Phoniatic at the Department of the ENT Center of the Republic of Belarus

Navrotsky A. A., Cand. of Sci., Associate Professor, Head of the Department of Informational Technologies of Automated Systems of the Belarusian State University of Informatics and Radioelectronics

Address for correspondence

246050, Republic of Belarus,
Gomel, Krestjanskaya St., 35, Apt. 12
Tel.: +375 29 730-13-80
E-mail: pozitron.kozarroman@gmail.com
Kozar Raman Viachaslavavich