

УДК004.932

## АЛГОРИТМ ОБНАРУЖЕНИЯ ВИЗУАЛЬНОГО СХОДСТВА ИЗОБРАЖЕНИЙ

*Никифоров М.Б., к.т.н., доцент кафедры «Электронные вычислительные машины» Рязанского государственного радиотехнического университета имени В.Ф. Уткина e-mail: nikiforov.m.b@evm.rsreu.ru*

*Тарасова В.Ю., аспирант кафедры «Электронные вычислительные машины» Рязанского государственного радиотехнического университета имени В.Ф. Уткина, e-mail: Valentina2008.91@mail.ru*

## ALGORITHM FOR DETECTING VISUAL SIMILARITY OF IMAGES

*Nikiforov M.B., Tarasova V.Y.*

*Currently, the volume of processed information is constantly increasing. Finding what you need in such a stream of data is a difficult task. For operational search in systems with a large amount of data, the concept of indexing with keywords, indexes, is used. In the event that it is necessary to detect images with similar content (the image shows the same object, but at different shooting angles, at different times of the year, etc.), the application of this concept is difficult. The developed algorithm for detecting visual similarity of images does not use a priori information, but takes into account only image pixels. It can be used in solving problems of visual search, classification, as well as software for systems for searching attractions from photographs in travel companies, in text and image search systems, in open access electronic catalogs (local history museums, cadastral maps). The algorithm makes it possible to increase the probability of finding images based on their content by hashing each image from the search base using convolutional neural networks.*

**Key words:** visual search, visual similarity, convolutional neural networks, hash.

**Ключевые слова:** визуальный поиск, визуальное сходство, сверточные нейронные сети, хэш.

**Введение**

Визуальный поиск нашел свое применение не только в поисковых системах Яндекс и Google, но и в библиотечных системах (поиск книги по обложке), при поиске изображений в электронных каталогах открытого доступа (краеведческий музей, кадастровые карты), в системах для поиска достопримечательностей по фотографиям в туристических фирмах, в системах поиска текстов и изображений [1]. Алгоритм поиска должен анализировать содержание, например, форму объектов, их текстуру, цвет и быть инвариантным к аффинным преобразованиям: изменениям размеров, ракурса съемки, условий освещенности, к искажениям формы и т.д. Основным требованием, предъявляемым к названным системам является высокая производительность, позволяющая за несколько секунд отвечать на пользовательские запросы.

При решении задач обнаружения изображений со схожим содержанием исходные данные представляют в виде хэша фиксированной длины. Разница между хэшами изображений с одинаковым содержанием должна быть минимальной.

Задачи обнаружения изображений со схожим содержанием являются актуальными, и рассматриваются как в Российской Федерации в научных трудах Игнатов А.К [2], Рогова А.А. [3], Бабенко А.В. [4], так и за рубежом в трудах Эль-сабан М.А., Тавфик А.И., Чалаби А.А.М.Т., Сайед С.Х. [5].

В последнее время было предложено несколько методов для обнаружения схожих по содержанию изображений с использованием метрической и неметрической

*В настоящее время объем обрабатываемой информации постоянно увеличивается. Отыскать в таком потоке данных необходимое является сложной задачей. Для оперативного поиска в системах с большим объемом данных используют концепцию индексирования ключевыми словами, индексы. В том случае если необходимо осуществить обнаружение изображений со схожим содержанием (на изображении представлено один и тот же объект, но под разными углами съемки, в разное время года и д.р.) применение данной концепции затруднительно. Разработанный алгоритм обнаружения визуального сходства изображений не использует априорной информации, а учитывает только пиксели изображения. Он может быть применен при решении задач визуального поиска, классификации, а также в качестве программного обеспечения систем для поиска достопримечательностей по фотографиям в туристических фирмах, в системах поиска текстов и изображений, в электронных каталогах открытого доступа (краеведческие музеи, кадастровые карты). Алгоритм позволяет повысить вероятность нахождения изображений, исходя из их содержания, за счет хэширования каждого изображения из поисковой базы с использованием сверточных нейронных сетей.*

оценки сходства [6]. Применение метрической оценки сходства, рассмотрены в статьях [7 – 9]. В данных работах изображения описаны векторами признаков и визуальное сходство между ними вычисляется, не принимая во внимание возможную взаимосвязь в наборе данных. Поэтому при сильной нелинейной зависимости в наборе данных оценка схожести может пропустить одинаковые по содержанию изображения. Однако, согласно исследованиям авторов [10, 11] метрическая оценка сходства не действительна для восприятия человеком визуального сходства, и, следовательно, функции визуального сходства необязательно должны удовлетворять условиям метрик расстояния [6]. В разработанном алгоритме предлагается использование метрической оценки сходства на основе Евклидова расстояния, однако хэш,

по которому осуществляется сопоставление, формируется на основе нейросетевого подхода, с применением операций свертки и подвыборки [12].

**Целью работы** является разработка алгоритмического и программного обеспечения для системы обнаружения изображений со схожим содержанием (визуального сходства).

### Материалы и результаты исследований

В том случае, если меру сходства определить как среднеквадратическое отклонение элементов матрицы изображений, то она будет рассчитываться следующим образом:

$$C(A, B) = \sum_{x=0}^N \sum_{y=0}^M \sum_{c=0}^K (A_{x,y,c} - B_{x,y,c})^2, \quad (1)$$

где  $A, B$  – изображения,  $M, N$  – ширина и высота матрицы изображения,  $K$  – количество цветовых каналов изображения,  $x, y, c$  – итерируемые переменные.

Для каждой пары изображений необходимо заново вычислять данную величину. Описанный подход может давать ошибочный результат при съёмке различных по смыслу объектов в одинаковых условиях.

Вместо попиксельного сравнения используются подходы, которые рассчитывают меру сходства с использованием вейвлет-преобразований, перцептивных хэш-функций [13, 14]. Однако данный подход уступает по точности сверточным нейронным сетям [1].

В данной работе предлагается меру сходства двух изображений рассчитывать как разность по модулю между хэшами этих изображений. Чем численная мера меньше, тем наиболее близкими по содержанию являются фотографии. Расчет численной меры сходства вычисляется следующими образом:

$$list\_Score_i = |score_{img} - score_i|, \quad (2)$$

где  $list\_Score_i$  – вычисленная мера сходства для  $i$ -го изображения из поисковой коллекции,  $score_{img}$  – хэш, описывающий изображение-запрос (метрика),  $score_i$  – хэш, описывающий изображение из поисковой коллекции (метрика),  $i$  – порядковый номер изображения в поисковой коллекции.

Хэш формируется с использованием операций свертки и подвыборки. Основной идеей сверточной сети является то, что обработка участка изображения должна не зависеть от расположения этого участка на изображении. Каждое изображение разбивается на 3 цветовых канала:  $R, G, B$ . Входные данные каждого конкрет-

ного значения пикселя нормализуются в диапазоне  $[0; 1]$  по формуле:

$$(p, min, max) = \frac{p - min}{max - min}, \quad (3)$$

где здесь –  $p$  яркость пикселя в одном из каналов  $R, G, B$ ,  $min$  – минимальное значение яркости пикселя,  $max$  – максимальное значение яркости пикселя.

Свертка организована таким образом, что за 6 итераций можно преобразовать изображение в некоторый хэш, состоящий из 512 байт. Данный хэш позволяет получить характеристики изображений. Некоторое ядро свертки обрабатывает предыдущее ядро по фрагментам. На рис. 1 представлена модель алгоритма свертки. Размер следующего слоя свертки рассчитывается по формуле:

$$(w, h) = (mW - kW + 1, mH - kH + 1) \quad (4)$$

здесь  $(w, h)$  – вычисляемый размер следующего слоя свертки,  $mW$  – ширина предыдущего слоя,  $mH$  – высота предыдущего слоя,  $kW$  – ширина ядра,  $kH$  – высота ядра.

Ядро свертки представляет собой фильтр (скользящее окно), которое находит определенные признаки объектов. Ядро (в данной работе применяется ядро размерностью  $3 \times 3$  пикселя с шагом дискретизации 1) скользит по предыдущей итерации и производит операцию свертки. Операция ReLU – это активационная функция, которая обнуляет все отрицательные элементы [15, 16].

Подвыборка представляет собой нелинейное уплотнение карты признаков, при котором группа пикселей ( $2 \times 2$ ) уплотняется до 1 пикселя, проходя операцию подвыборки (применяется максимизирующая функция). Цель операции подвыборки – уменьшение размерности результата предыдущей операции (свертки).

Обучение алгоритма производилось на наборах данных (изображениях, размерностью  $128 \times 128 \times 3$ ), разбитых по содержанию и условиям съёмки (всего 512 сценариев). Для обучения применялась функция потерь `binary_crossentropy` с оптимизатором Adam.

Алгоритм обнаружения визуального сходства следующий:

1. Формирование поисковой базы.
2. Формирование пользовательского запроса.
3. Формирование метрики запроса по содержанию.
4. Вычисление меры сходства.
5. Вывод результата пользовательского запроса.

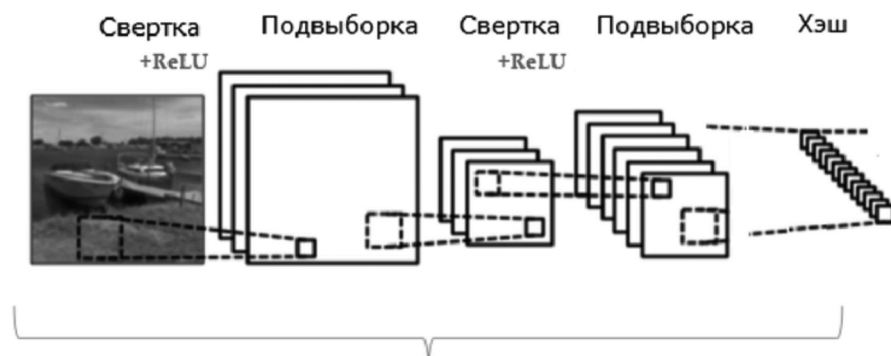


Рис. 1. Модель алгоритма свертки

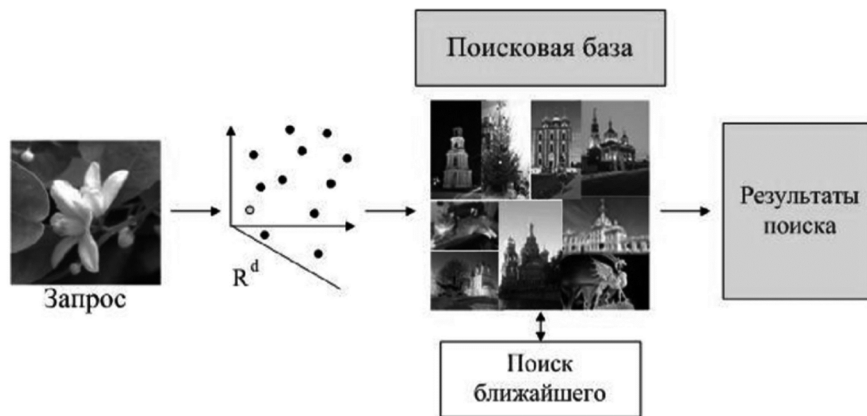


Рис. 2. Схема работы системы визуального поиска

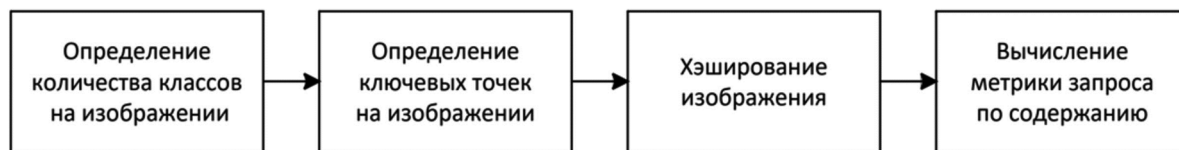


Рис. 3. Схема формирования метрики запроса по содержанию

Поисковая база представляет собой два файла: в первом хранится полный путь к файлу за исключением имени, во втором – ссылка на первый файл, непосредственно имя файла и вычисленная метрика (метрическая оценка сходства двух изображений). Данная структура была разработана для уменьшения размера поисковой базы.

Формирование пользовательского запроса – это загрузка изображения и вычисление для него хэша. На рис. 2 приведена схема работы системы визуального поиска. Сформированные метрики  $R^d$  для корректного поиска должны быть инвариантны к углам поворота, изменению масштаба и угла зрения.

Метрики для корректного поиска должны быть инвариантны к углам поворота, изменению масштаба и угла зрения [17]. Для вычисления метрики использовались:

- метод кластеризации  $k$ -средних;
- гистограмма цветов;
- ключевые точки;
- операции свертки и подвыборки.

Схема формирования метрики запроса по содержанию представлена на рис. 3.

Вычисление меры сходства осуществляется по формуле (2).

В качестве цветовой составляющей изображения формируется 3 ключевых цвета, определяемых при помощи метода  $k$ -means. Для анализа используется палитра HSV и только Hue-составляющая. На гистограмме (используется гистограмма обесцвеченного изображения) выделяются 5 зон: тень, полутень, основная зона, светлые тона, пересвет. Для каждой из зон рассчитывается количество точек, попавших в данную зону. Данное число отражается в итоговом хэше изображения.

Важной частью анализа изображения является оценка содержания. Для этого применяется свёрточная нейронная сеть MobileNet v2 [5]. Применяемая архитектура относится к классификационным сетям и позволяет определить класс объекта, который отражен на изоб-

ражении. Предлагается использовать данные с предпоследнего слоя.

Хэш формируется на основе представленных методов и представляет собой 136 байт данных:

- ключевые цвета (3 байта);
- зоны гистограммы (5 байт);
- сведения об информационной составляющей (128 байт).

Для определения меры близости изображений, выполняется их побайтовое сравнение. Для каждого из показателей может быть вручную определен коэффициент масштабирования, для того, чтобы отдать приоритет одной из трех составляющих.

### Результаты исследований

На основе предложенного алгоритма обнаружения визуального сходства изображений было разработано программное обеспечение, функциональными возможностями которого являются:

- формирование поисковой коллекции;
- поиск изображений из поисковой коллекции;
- вывод результатов поиска.

На рис. 4 приведен результат работы алгоритма. Время выполнения поиска среди коллекции, в которой содержится более 50 тысяч изображений составляет, 12 миллисекунд. Из рис. 4 видно, что все найденные изображения соответствуют пользовательскому запросу. Все изображения из поисковой выборки содержат деревья, которые изображены на запросе пользователя.

### Обсуждение

Для того, чтобы оценить эффективность предложенного алгоритма обнаружения, необходимо сравнить его с существующими методами. Для этого поисковая база содержала следующие изображения:

- отсканированные изображения с наличием расщеплений;
- нарисованные пользователями с различной степенью точности изображения;



Рис. 4. Результат выполнения поиска

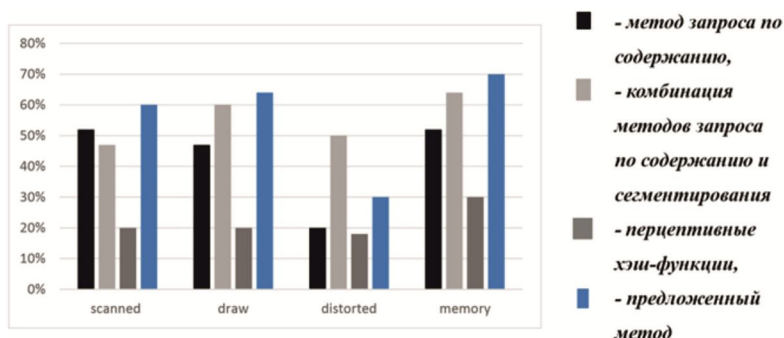


Рис. 5. Гистограмма зависимости верных запросов от типа изображения и метода обработки

– искаженные изображения помехами типа «соль-перец»;

– находящиеся в памяти ПК изображения.

Под рассовмещением понимается отображение одного и того же рисунка, расположенного с некоторым (чаще всего незначительным) смещением (и/или) поворотом.

На рис. 5 приведена гистограмма, столбцы которой показывают какой процент запросов привел к получению верного целевого изображения в зависимости от применяемого метода. Были рассмотрены следующие методы:

- метод запроса по содержанию [18];
- комбинация методов запроса по содержанию с сегментированием [12];
- перцептивные хэш-функции [19];
- предложенный алгоритм.

Предложенный метод является менее эффективным в том случае, если пользовательский запрос представляет собой искаженное изображение помехами типа «соль-перец». Однако для других видов исходных изображений показывает лучший результат.

### Заключение

Разработанный алгоритм и его программная реализация могут быть использованы пользователем для поиска совпадений в поисковой базе. Алгоритм обнаруже-

ния включает в себя операции свертки и подвыборки, которые позволяют преобразовывать изображения в хэш, инвариантный к различным искажениям, что позволяет с высокой эффективностью находить идентичные изображения в поисковой базе. Малое время работы и высокая эффективность представленного алгоритма позволяют использовать такое приложение на персональном компьютере. Для сокращения времени работы алгоритма обнаружения подвижных объектов планируется применение технологий параллельных вычислений.

### Литература

1. Тарасов А.С., Тарасова В.Ю. Разработка системы поиска похожих изображений в локальных хранилищах. В сборнике: Нейроинформатика-2020. Сборник научных трудов. XXII международная научно-техническая конференция. 2020. С. 286-293.
2. Игнатов А.К. Способ индексации и поиска цифровых изображений. Патент на изобретение RU 2510935 С2, 10.04.2014. Заявка № 2011138862/08 от 23.09.2011.
3. Рогов А.А., Рогова К.А., Спиридонов К.Н., Быстров М.Ю. Информационно-поисковая система Петроглифов Карелии. Вестник компьютерных и информационных технологий. 2008. № 6 (48). С. 6-11.
4. Бабенко А., Лемпицкий В. Эффективный алгоритм

поиска ближайших соседей при больших объемах поисковой базы. Труды 54 научной конференции МФТИ. Инновации и высокие технологии. Долгопрудный: 2011. С. 16-17.

5. Эль-сабан М.А., Тавфик А.И., Чалаби А.А.М.Т., Сайед С.Х. Поиск изображений на естественном языке. Патент на изобретение RU 2688271 С2, 21.05.2019. Заявка № 2016144699 от 14.05.2015.

6. Noa Garcia, George Vogiatzis Learning Non-Metric Visual Similarity for Image Retrieval the IEEE International Conference on Computer Vision Workshops (ICCVW). 2017.

7. Venice Erin Liong, Jiwen Lu, Gang Wang, Pierre Moulin, and Jie Zhou. Deep hashing for compact binary codes learning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015.

8. Albert Gordo, Jon Almazan, Jerome Revaud, and Diane Larlus. Deep image retrieval: Learning global representations for image search. In Proceedings of the IEEE European Conference on Computer Vision. 2016.

9. Wenjie Luo, Alexander G Schwing, and Raquel Urtasun. Efficient deep learning for stereo matching. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.

10. Laurens van der Maaten•Geoffrey Hinton Visualizing non-metric similarities in multiple maps. The Author(s) 2011. This article is published with open access at Springerlink.com DOI 10.1007/s10994-011-5273-4.

11. Tomá's Skopal On Fast Non-metric Similarity

Search by Metric Access Methods. (Eds.): EDBT 2006, LNCS 3896, pp. 718-736 2006.

12. Potapova V., Tarasov A., Grinchenko N. Image Search by Content System Development Published in: 2018 IEEE East-West Design & Test Symposium (EWDTS) 2018. С. 625-629

13. Рудаков И.В., Васютрович И.М. Исследование перцептивных хэшфункций изображения. В сб.: Наука и образование МГТУ им. Н.Э. Баумана. С. 269-280.

14. Tzu-Heng Henry Lee. Wavelet Analysis for Image Processing. [http://disp.ee.ntu.edu.tw/henry/wavelet\\_analysis.pdf](http://disp.ee.ntu.edu.tw/henry/wavelet_analysis.pdf).

15. Николенко С., Кадурин А., Архангельская Е. Глубокое обучение погружение в мир нейронных сетей. СПб.: Питер, 2019. 480 с

16. Тарасова В.Ю., Тарасов А.С., Гринченко Н.Н. Алгоритм обнаружения совпадений в коллекции изображений в сборнике трудов Международного научно-технического форума СТНО-2019. Том 4. С. 164-168

17. Bodrov O.A., Tarasov A.S., Tarasova V.Y., Bodrova I.V. IMAGE SEARCH ALGORITHM IN LOCAL DATA BASE 8th Mediterranean Conference on Embedded Computing, MECO 2019. Proceedings. 2019. С. 453-455.

18. Eric J. Stollnitz, Tony D. DeRose, David H. Salesin Wavelets for Computer Graphics Theory and Applications.

19. F. Sabahi; M. Omair Ahmad; MNS Swamy Content-based Image Retrieval using Perceptual Image Hashing and Hopfield Neural Network. 2018 IEEE 61<sup>st</sup> International Midwest Symposium on Circuits and Systems (MWSCAS).