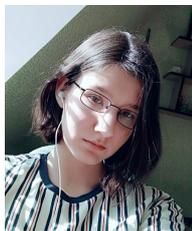


УДК 004.853

ПОИСК ВИЗУАЛЬНО ПОДОБНЫХ ИЗОБРАЖЕНИЙ НА ОСНОВЕ МАШИННОГО ОБУЧЕНИЯ



М. М. Гресик
Студентка БГУИР КСиС
mariya.gm.4@mail.ru



С. Н. Нестеренков
Доцент кафедры программного
обеспечения информационных
технологий, Декан ФКСиС, Кандидат
технических наук
s.nesterenkov@bsuir.by

М. М. Гресик

Студентка 4-ого курса специальности информатика и технологии программирования БГУИР. Проходит производственную практику в Центре информатизации и инновационных разработок в Белорусском государственном университете информатики и радиоэлектроники.

С. Н. Нестеренков.

Кандидат технических наук, декан факультета компьютерных систем и сетей Белорусского государственного университета информатики и радиоэлектроники, доцент кафедры Программного обеспечения информационных технологий. Многократный автор публикаций на тему машинного обучения, алгоритмов принятия решений, искусственных нейронных сетей и автоматизации.

Аннотация. В данной статье рассматриваются вопросы поиска изображений по визуальному образцу. Предлагается методика разработки алгоритма оценки визуальной близости пары изображений на основе машинного обучения. Анализ и использование признаков изображения для построения специализированных поисковых индексов, используемых в алгоритме оценки визуальной близости, а также выбор таких признаков.

Ключевые слова: поиск изображений по визуальному образцу, машинное обучение, нейронные сети.

Введение.

Развитие компьютерной техники и технологий привело к бурному росту количества цифровых изображений. Источников таких изображений множество — цифровые фотоаппараты и видеокамеры, видеорегистраторы, камеры видеонаблюдения, мобильные телефоны, сканеры и т. д. Появились цифровые архивы, содержащие миллионы снимков.

При таких объемах данных важное значение приобретает поиск нужных изображений.

Существует несколько направлений поиска по изображениям: поиск по содержанию (найти фотографию человека или изображение, к примеру, берёзы), поиск по визуальному образцу (найти изображения, похожие на заданное), поиск по описаниям (найти изображение, помеченное как «Совершенно секретно» или «Москва») и т. д.

Каждое из направлений поиска имеет свои особенности и сферы применения.

Целью работы является рассмотрение вопросов построения системы поиска по визуальному образцу в относительно больших (десятки тысяч) коллекциях изображений.

Поиск изображений по визуальному образцу является актуальной задачей, решению которой посвящено множество работ. Данный вид поиска заключается в извлечении существенных свойств изображений и построении на их основе сигнатур, используемых в дальнейшем для сравнения пар изображений. В каждую пару всегда входит изображение из коллекции и изображение-образец, подобное которому стремится найти пользователь.

Результатом сравнения является величина, называемая визуальным подобием

(тематической близостью) изображений.

С точки зрения информационного поиска данная оценка, выполненная человеком-экспертом, будет являться содержательной релевантностью, а рассчитанная в системе — релевантностью формальной.

Свойства оцениваются алгоритмами вычисления значений важных численных величин и называются признаками. Подобных признаков можно выделить достаточно много.

В общем случае они разделяются на две группы: глобальные и локальные. К глобальным признакам можно отнести основные цвета, текстуры, формы, значимые элементы всего изображения. Локальные признаки высчитываются для небольших частей (блоков) изображения. Для обозначения значений признаков также употребляют термин дескриптор. Набор признаков (вектор признаков), описывающий изображение для определенной задачи, называется сигнатурой.

Признаки, отобранные из множества альтернативных для их применения в некоторой задаче анализа изображений и вошедшие в сигнатуру, называют ключевыми. При поиске изображений по образцу анализируются не отдельные экземпляры, а пары изображений, которые сопоставляются друг другу.

Поэтому в отличие от других задач анализа в текущей проблеме от признаков изображений переходят к признакам пар, значения которых находятся как абсолютные разности значений соответствующих признаков каждого из изображений пары.

Такая несложная формула позволяет моментально их рассчитать для проиндексированных изображений и реализовать на практике поиск с приемлемыми временными характеристиками.

Алгоритм определения визуального подобия изображений.

Схематично предлагаемая методика работ для построения алгоритма определения визуального подобия изображений на основе машинного обучения показана на рисунке 1.

Ее можно представить в виде этапов, которые должны быть последовательно выполнены:

Этап 1. Из коллекции изображений формируется обучающая выборка. Она состоит из множества прецедентов, каждый из которых включает в себя пару изображений и степень их близости, оцененную экспертом.

Этап 2. Программно реализуются (при необходимости и разрабатываются) алгоритмы вычисления всевозможных признаков на изображениях.

Этап 3. Для каждой пары изображений вычисляются значения всех признаков.

Этап 4. Проводится факторный анализ признаков, в результате которого определяются главные признаки, значения которых будут использоваться в обучении нейронной сети и алгоритмы вычисления которых станут частью конечного алгоритма оценки визуального подобия изображений.

Этап 5. Обучается нейронная сеть. По завершению этого процесса запоминаются данные, характеризующие величины ошибок нейронной сети относительно входной выборки. Эти показатели преобразуются в полноту и точность — целевые значения оценки качества выполняемой работы.

Этап 6. Из коллекции изображений формируется тестовая выборка. Ее структура и характеристики (объем, соотношение близких и неблизких пар) должны соответствовать структуре и характеристикам обучающей выборки.

Этап 7. Вычисляются значения ключевых признаков всех изображений, входящих в тестовую выборку.

Этап 8. Прогоняется нейронная сеть на тестовой выборке с определением полноты и точности.

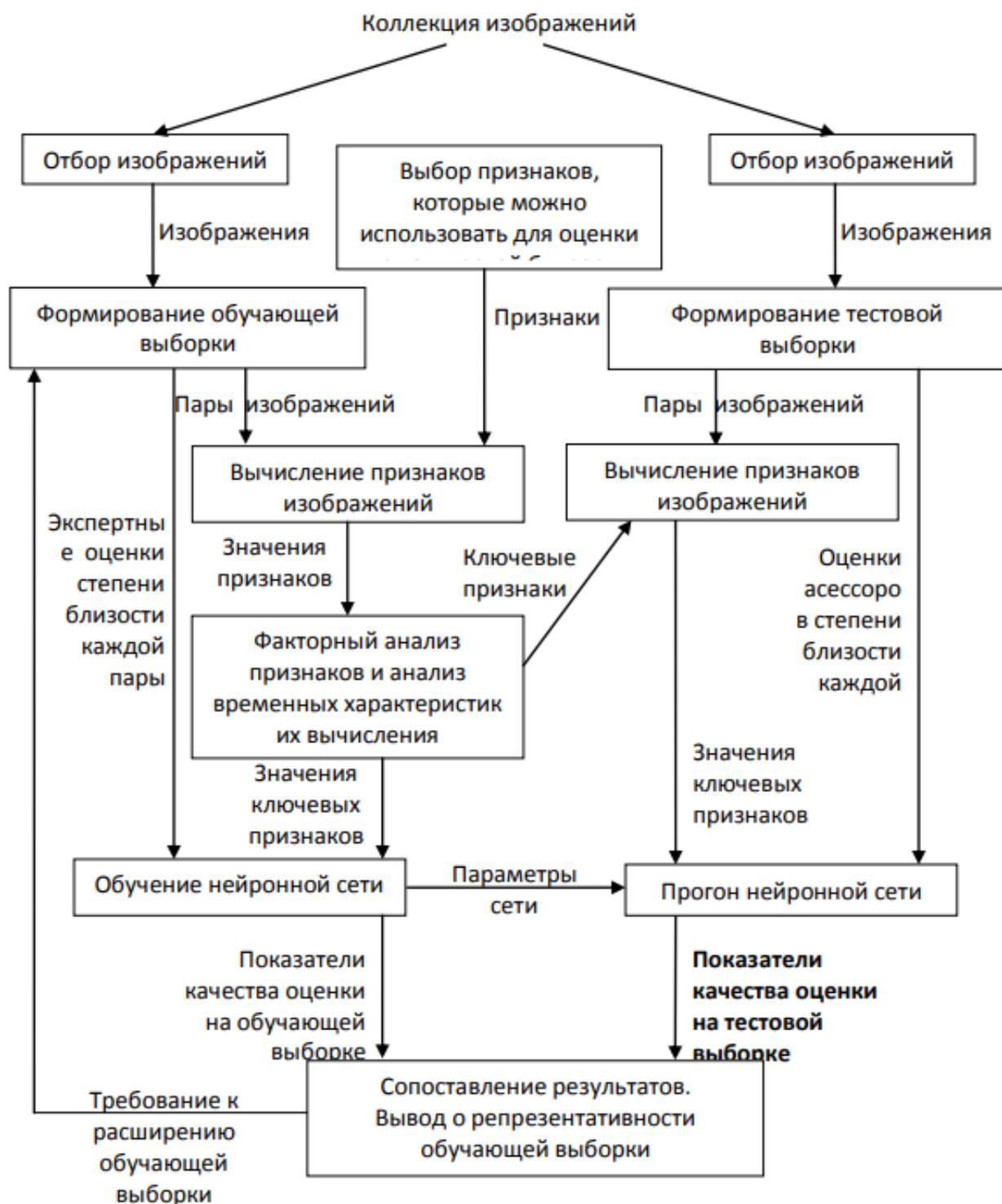


Рисунок 1. Схема построения алгоритма определения визуального подобия изображений

Этап 9. Сравниваются полнота и точность анализа нейронной сетью изображений обучающей выборки с полнотой и точностью анализа нейронной сетью изображений тестовой выборки.

При приблизительно одинаковых значениях соответствующих показателей работа считается завершенной.

Признаки изображений.

В связи с тем, что поиск может осуществляться в больших коллекциях изображений, в качестве признаков мы рассматривали только те, которые можно просчитать для изображений заранее и занести в некое подобие поискового индекса (контрастность, грубость, направленность линейных образов, регулярности и шероховатости текстур).

По этой причине, нами не рассматривались те признаки, которые необходимо рассчитывать по парам изображений непосредственно при осуществлении поиска. Из 25 признаков на этапе выбора главных часть была отсеяна. Таким образом, к ключевым признакам были отнесены следующие:

- Средние значения компонент R, G, B изображения в цветовой модели RGB;
- Медиана яркости;
- Детализованность;
- Признак симметрии изображения;
- Признак наличия текстур;
- Пропорции сторон;
- Наличие лиц на изображении;
- Нормированные 16-ти уровневые гистограммы по компонентам Y (яркости), U, V изображения в цветовой модели YUV;
- Среднеквадратичное отклонение яркостей уменьшенных копий (размером 32x32 пикселя) образца и рассматриваемого изображения.

В качестве набора исходных данных использовалась коллекция Flickr [1].

В ней содержится 20000 фотографий разного качества и без единой темы; имеются фотоснимки людей, пейзажи, городские сцены и т.д. Фотографии сделаны при разном освещении, в помещениях, на улице и т.д.

Из данной коллекции в обучающую выборку были отобраны 500 пар, среди которых 125 являются тематически близкими, а 375 нет. Тестовой выборкой являлось специальное задание дорожки коллекции Flickr с результатами работы группы ассессоров.

Для формирования выборок и обучения нейронной сети применялся нейроимитатор Сигнейро [2], специализирующийся на задачах обработки и анализа изображений. Была сконструирована двухслойная нейронная сеть, структура которой приведена на рисунке 2.

Количество входов сети соответствует количеству ключевых признаков, выход сети — один. Для осуществления поиска изображений по визуальному подобию на основе полученной нейронной сети была разработана программная оболочка Axioma.

В процессе поиска она извлекает данные сигнатур (значения признаков) из поискового индекса, осуществляет вычисление меры близости образца к изображениям коллекции и ранжирует результаты поиска.

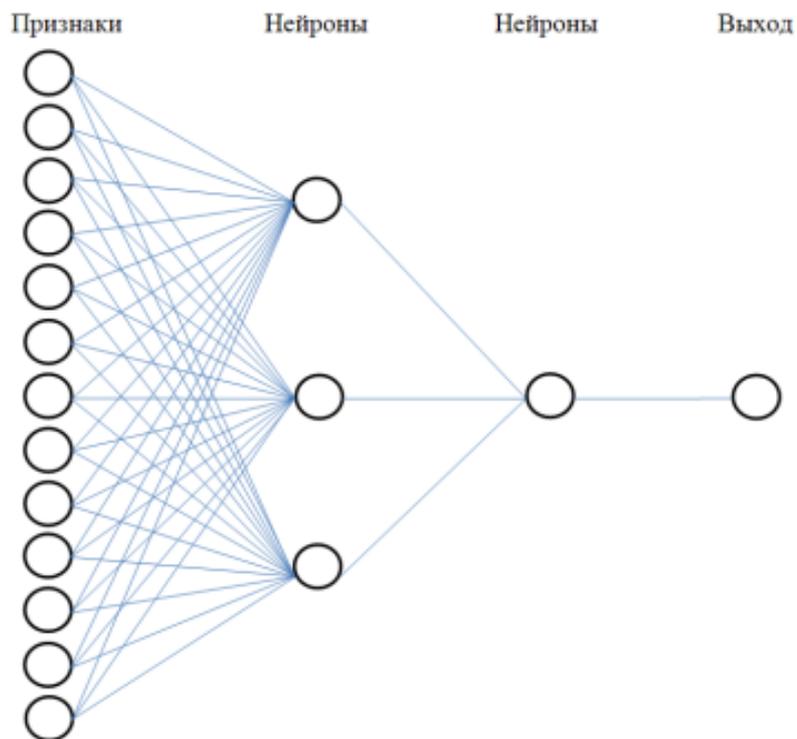


Рисунок 2. Структура двухслойной нейронной сети

График полнота-точность полученного нейросетевого алгоритма приведен на рисунке 3. На графике виден выступ, крайняя точка которого соответствует решению, полученному нейронной сетью.

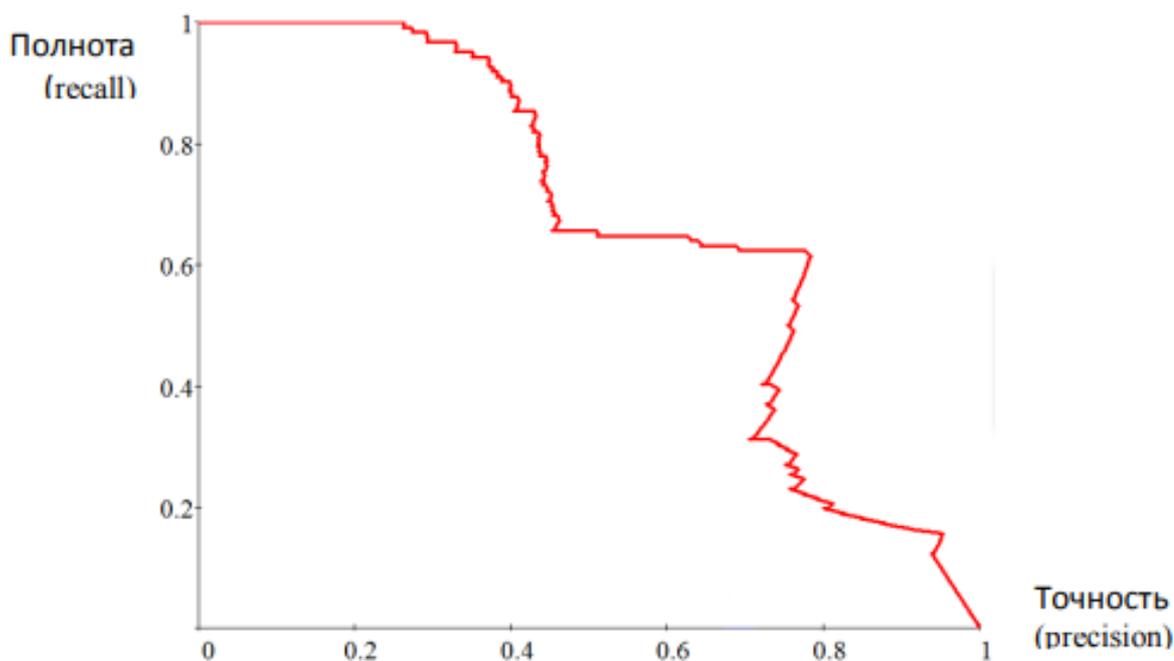


Рисунок 3. График полнота-точность

Заключение.

Анализируя проведенную работу, были сделаны следующие выводы:

Полученные качественные характеристики построенного алгоритма имеют достаточно высокую оценку.

Вошедшая в оптимизационный критерий временная характеристика позволила минимизировать трудоемкость итогового алгоритма, который обладает линейным порядком временной сложности.

В дальнейшем, при увеличении объемов обучающей и тестовой выборок, качественные показатели могут быть улучшены или даже превзойдут значения аналогов, так как на данный момент имеются расхождения в показателях обучающей и тестовой выборок.

Список литературы

[1] Коллекция изображений Flickr [Электронный ресурс]. – Электронные данные – Режим доступа: <http://romip.ru/ru/collections/flickr.html>. – Дата доступа: 24.03.2023

[2] Сигнейро. Нейронные сети в обработке и анализе изображений [Электронный ресурс]. – Электронные данные – Режим доступа: <http://www.signneuro.ru>. – Дата доступа: 24.03.2023

[3] Нестеренков, С.Н. Сравнительный анализ методов бинаризации изображений/ С.Н. Нестеренков, Д.В.Орлов, А. Н.Марков// Восьмая Международная научно-практическая конференция «BIG DATA and Advanced Analytics. BIG DATA и анализ высокого уровня», Минск, Республика, 11-12 мая 2022 год. - С. 104-109

[4] Нестеренков, С.Н. Использование генетического алгоритма для нахождения коэффициентов и структуры нейронной сети / С.Н. Нестеренков, К.П. Белов // Информационные технологии и системы 2018 (ИТС 2018) : материалы междунар. науч. конф., Минск, 25 окт. 2018 г. / Белорус. гос. ун-т информатики и радиоэлектроники ; редкол.: Л. Ю. Шилин [и др.]. - Минск, 2018. - С. 124-125.

[5] Нестеренков, С. Н. Модифицированный генетический алгоритм для обучения нейронной сети/ С. Н. Нестеренков, К. П. Белов // Информационные технологии и системы 2017 (ИТС 2017) : материалы междунар. науч. конф., Минск, 25 окт. 2017 г. / Белорус. гос. ун-т информатики и радиоэлектроники ; редкол.: Л. Ю. Шилин [и др.]. - Минск, 2017. - С. 204-205

SEARCH FOR VISUALLY SIMILAR IMAGES BASED ON MACHINE LEARNING

M. M. Hresik
BSUIR KSiS student

S. N. Nesterenkov
*PhD Associate professor of department of
the software of information technologies*

*Center for Informatization and Innovative Developments of the Belarusian State University of Informatics and Radio Electronics, Republic of Belarus
Belarusian State University of computer science and Radio Electronics, Republic of Belarus
E-mail: mariya.gm.4@mail.ru, s.nesterenkov@bsuir.by.*

Abstract. This article deals with the search for images by visual pattern. A technique for developing an algorithm for estimating the visual proximity of a pair of images based on machine learning is proposed. Image features that can be used in search algorithms are analyzed. A solution to the problem of choosing among them the key ones is presented.

Keywords: Image search by visual pattern, machine learning, neural networks.