

Горохов Олег Андреевич,
аспирант, Ангарский государственный технический университет,
Истомин Андрей Леонидович,
д.т.н., профессор, Ангарский государственный технический университет,
e-mail: a.l.istomin@mail.ru

РАСПОЗНАВАНИЕ ОБРАЗОВ: МЕТОДЫ И ПРОГРАММНЫЕ СРЕДСТВА РЕАЛИЗАЦИИ

Gorokhov O.A., Istomin A.L.

PATTERN RECOGNITION: METHODS AND SOFTWARE IMPLEMENTATION TOOLS

Аннотация. Детальный анализ методов распознавания образов и программных средств для его использования.

Ключевые слова: распознавание образов, метод, программное обеспечение.

Abstract. Detailed analysis of pattern recognition methods and software for its use.

Keywords: pattern recognition, method, software.

В современном мире контроль состояния технологического оборудования на производственных площадках и способность упреждать возникновение аварий является одной из важных задач. Зачастую процесс контроля осуществляется в ручном режиме – оператор, ответственный за вверенный ему участок, следит за показаниями измерительных датчиков через видеорегистрирующее оборудование и на основе имеющейся информации о критических значениях показателей определяет состояние технологического оборудования – аварийное или предаварийное – и предпринимает соответствующие действия по решению возникшей проблемы.

Однако в ходе работы может сказаться влияние человеческого фактора, что может привести к неверному результату анализа показаний датчиков. В данном случае может понадобиться использование средств автоматизированного распознавания образов, их анализа и принятия соответствующих решений.

Распознавание образов – это отнесение исходных данных к определённого класса с помощью выделения существенных признаков, характеризующих эти данные, из общей массы данных [1]. Данный раздел информатики находит своё применение в различных областях, как, например, медицина, транспорт, криминалистика и другие.

В зависимости от условий задачи распознавания образов, могут применяться различные методы, которые будут рассмотрены далее.

Метод сравнения с эталоном, который применим в том случае, когда каждому классу – подмножеству множества объектов идентификации – можно сопоставить конечный набор эталонных образов [2]. Образ – это совокупность данных об объекте, позволяющих выделить его из окружающей среды и объединить его с другими объектами для принятия решения. Принцип сравнения с эталоном

также называют принципом перечисления. В таком случае процесс распознавания сводится к сопоставлению эталонных образов, поступающих на вход распознающего устройства или алгоритма, с эталонами классов, на основе выбранной меры сходства. Принцип сравнения с эталоном – один из первых подходов, возникших при построении технических систем распознавания при ограниченности возможностей вычислительных систем. Этот метод применяется и в настоящее время в аналоговых и аналогово-цифровых системах распознавания.

Распространенным примером применения такого подхода является распознавание печатных шрифтов. В тематической обработке изображений он может применяться при распознавании текстур и выделении объектов определенной формы. Недостатком метода является необходимость создания большого количества эталонов, для каждого из которых необходимо вычислить расстояние. На эту операцию затрачивается значительное количество времени и вычислительных ресурсов.

Метод общности свойств существенно использует связи между элементами образа. Как правило, он применяется в тех случаях, когда множество образов каждого класса слишком велико, чтобы получить надежное описание конечного числа эталонов. Выявляется достаточное количество отличительных особенностей классов по конечным выборкам образов. Выявленные свойства кодируются на основе математической модели и хранятся в памяти в виде информационных структур, функций или отношений. В процессе распознавания производится анализ образа по схеме, позволяющей выявить необходимые свойства образа и затем они сопоставляются со свойствами классов. Обобщающим свойством может быть сам алгоритм порождения образов. В этом случае классы образов задаются алгоритмами порождения структур определённого вида.

Если признаки представляют собой наборы измерений (параметров) без каких-либо явным образом заданных взаимосвязей, то возможно использование методов кластеризации и образ представляется как n -мерный вектор в признаковом пространстве.

Суть метода состоит в том, что каждому классу сопоставляется некоторое множество векторов в пространстве. В результате признаковое пространство разбивается на области, соответствующие классам, которые называют кластерами или таксонами. Одному классу может соответствовать любое количество областей пространства. Эти области могут перекрываться, но методология принятия решения в конечном итоге обеспечивает однозначное отнесение точки пространства к определенному классу.

Принцип кластеризации (таксономии) широко применяется в различных прикладных направлениях при обработке количественных данных, в том числе и в системах компьютерного анализа многозональных и спектральнозональных аэрокосмических изображений (классификация по спектральным признакам).

Классификация образов при помощи функций расстояния – один из базовых методов распознавания. Одной из функций расстояния является функция расстояния Евклида. Для классификатора принимается, что эталонные образы определяются как векторы математического ожидания образов из этого класса.

Определяется, что объект относится к тому классу, эталон которого является ближайшим. Одной из распространенных мер расстояния является евклидова норма. Таким образом, объект относится к тому классу, для которого евклидова норма оказывается наименьшей.

Данный метод является эффективным для классов, которые характеризуются степенью изменчивости, ограниченной в заданных пределах.

Корреляционное сопоставление предполагает наличие эталона, определенного на множестве, при этом носителем изображения является область изображения, совпадающая по форме с областью, соответствующей изображению реального объекта [3]. Носителем изображения называется множество точек с координатами, в которых его значение не равно нулю. Это означает, что с помощью смещения вдоль координат множества область изображения можно привести в область реального объекта.

Суть метода состоит в нахождении экстремума корреляционной функции.

Преимуществом данного метода является отсутствие необходимости нахождения признаков, а также предварительного выделения области объекта. Однако для каждого эталона необходимо выполнить большое количество вычислений корреляционной функции. Это требует значительных временных затрат.

Аппарат нечетких множеств и нечеткой логики применяется к задачам классификации изображения с конца восьмидесятых годов прошлого века [4]. Эти методы хорошо показали себя при работе с объектами регулярной формы, однако, несмотря на множество исследований, не решены проблемы с множеством объектов нерегулярной формы, низкого качества и контрастности.

Для работы байесовского классификатора необходимо наличие функции плотности распределения каждого из классов и априорная вероятность появления каждого класса [5].

Для этого метода вводится предположение о том, что существует некая вероятность соответствия изображения определённому классу.

Байесовский классификатор сводится к вычислению дискриминантной функции, и изображение соотносится с тем классом, у которого вычисленное значение дискриминантной функции оказывается наибольшим.

Как было указано выше, необходимо наличие функции плотности распределения каждого из классов и априорная вероятность появления каждого класса. Однако определение функции плотности распределения является трудоемкой задачей, так как является многомерной функцией и требует использования методов теории вероятности, что сложно применить на практике.

Нейронные сети начали использовать для обработки изображений в шестидесятые года двадцатого века. Впервые нейронную сеть, имитирующую работу мозга, – персептрон – создали в 1958 году. Однослойный персептрон по аналогии с мозгом состоит из нейронов – простейших вычислительных элементов, дендритов – входных сигналов и аксона – выходного сигнала.

Нейрон имеет несколько входов с весами – дендритами – и одним выходом – аксоном.

На выходе персептрон выдает результат, основанный на нескольких вещественных входных объектах, путем формирования линейной комбинации с использованием весовых коэффициентов.

Наибольшее применение получила архитектура многослойного персептрона. В многослойном персептроне создаются несколько слоев нейронов разного размера, при этом выходы предыдущего слоя являются входами для последующего [6].

Сигналы распространяются как в одном направлении, так и существуют модели с обратным распространением для обновления весов.

Несмотря на широкое распространение искусственных нейронных сетей, у них есть недостатки и ограничения:

- требования к наличию большого количества эталонных изображений, соотнесенных с классами,
- наличие проблемы с переобучением сети при снижении точности полученных результатов,
- сложность понимания поведения сети и предсказуемо повлиять на результаты её работы.

При работе с нейронными сетями возникает необходимость в разработке представления о том, что будет являться для них входными сигналами, так как при передаче на вход цифровое изображение приведёт к значительным временным затратам на работу и потребует больших вычислительных мощностей

В методологии принятия решений при распознавании образов выделяют три основных направления: эвристические методы; математические методы; лингвистические (синтаксические) методы.

Эвристические методы основаны на опыте специалиста, разрабатывающего систему распознавания. Такие методы ориентированы на решение конкретного типа задач распознавания и непосредственно привязаны к способу синтеза образов. Наиболее часто применяются при использовании принципа сравнения с эталоном и принципа общности свойств.

Математические методы опираются на использование классического математического аппарата: методов линейного программирования, корреляционного анализа, теории статистических решений и тому подобное. Применяются в тех случаях, когда признаки представлены измерениями (параметрами), а их связи могут быть описаны в виде аналитических зависимостей. В той или иной

степени, математические методы используются во всех трёх подходах к распознаванию.

Математические методы обычно подразделяют на детерминированные и статистические. В задачах анализа данных наиболее часто применяют статистические методы. Это обусловлено высокой степенью изменчивости условий наблюдения и самих объектов исследования, а также потерей информации на разных этапах проведения исследования. Для оценки и анализа изображений используется его статическое свойство – гистограмма яркости.

Лингвистические (синтаксические, структурные) методы применяются в тех случаях, когда образ представляет собой некоторую структуру, состоящую из так называемых непримитивных (первичных) элементов и признаков, описывающих связи между ними. В этих методах широко используется аппарат алгебры логики и теории формальных языков. При обработке данных и в геоинформационных технологиях эти методы могут использоваться при анализе пространственных структур, некоторые из них применяются также в экспертных системах. Современный уровень развития вычислительных средств позволяет комбинировать в системах распознавания как различные подходы к описанию образов, так и методы, используемые в процессе распознавания.

Задачи распознавания образов могут решаться с применением компьютерного (машинного, технического) зрения. Под техническим зрением понимается теория и технология создания машин, которые могут производить обнаружение, отслеживание и классификацию объектов. В контексте распознавания образов подразумевается возможность вычислительного устройства извлекать необходимую для конкретной задачи информацию в целях её дальнейшей обработки.

За всё время существования распознавания образов было выпущено немало программных решений для работы с компьютерным зрением, в том числе с применением систем искусственного интеллекта и нейросетей.

Одним из популярных инструментов является OpenCV (Open Source Computer Vision Library, библиотека компьютерного зрения с открытым исходным кодом), библиотека программного обеспечения для компьютерного зрения и машинного обучения с открытым исходным кодом. OpenCV был создан для обеспечения общей инфраструктуры для приложений компьютерного зрения и ускорения использования машинного восприятия в коммерческих продуктах [7].

Библиотека содержит более 2500 оптимизированных алгоритмов, которые включают в себя полный набор как классических, так и самых современных алгоритмов компьютерного зрения и машинного обучения. Эти алгоритмы могут быть использованы для обнаружения и распознавания лиц, идентификации объектов, отслеживания движений камеры, извлечения 3D-моделей объектов, поиска похожих изображений из базы данных изображений и так далее.

Компания Mathworks имеет свой инструмент для работы с машинным зрением под названием Computer Vision Toolbox, предлагающий пользователю визуальный осмотр, обнаружение объектов и их отслеживание, обнаружение, извлечение и сопоставление характеристик; автоматизацию калибровки различных камер; стереовидение, обработку облаков точек, построение структуры по движению; предварительно обученные детекторы объектов или возможность их обучить с помощью алгоритмов глубокого обучения и машинной обучаемости; классификацию изображений с помощью визуальных преобразователей [8].

Помимо наборов библиотек для языков программирования, предполагающих ручную правку написанного кода при изменении условий работы программы, имеется возможность применения инструментов машинного обучения, среди которых можно отметить TensorFlow.

TensorFlow – открытая программная библиотека для машинного обучения, разработанная компанией Google для решения задач построения и тренировки нейронной сети с целью автоматического нахождения и классификации образов, достигая качества человеческого восприятия [8]. Применяется как для исследований, так и для разработки собственных продуктов Google. Основной API для работы с библиотекой реализован для Python, также существуют реализации для C#, C++, Java, JavaScript и других языков программирования.

TensorFlow предоставляет ряд инструментов компьютерного зрения и классификации изображений через библиотеки Keras более высокого уровня и модуль tf.image более низкого уровня. В большинстве случаев библиотеки Keras считаются авторами проекта более удобными, чем встроенные альтернативы TensorFlow. Но, если параметры Keras не подходят для пользовательского варианта использования или ему нужен контроль нижнего уровня над предварительной обработкой изображений, то могут понадобиться инструменты TensorFlow нижнего уровня [9].

KerasCV – это библиотека модульных компонентов CV, построенная на Keras Core. KerasCV включает модели, слои, метрики, обратные вызовы и другие инструменты, расширяющие высокоуровневый API Keras для задач CV. API-интерфейсы KerasCV могут помочь с увеличением данных, классификацией, обнаружением объектов, сегментацией, генерацией изображений и другими пространственными рабочими процессами CV. Пользователь может использовать KerasCV для быстрой сборки современных конвейеров обучения и вывода промышленного уровня.

Если KerasCV не подходит для пользовательского варианта использования, он может использовать tf.image и tf.data для написания собственных конвейеров или слоев увеличения данных.

Модуль tf.image содержит различные функции для обработки изображений, такие как настройка яркости, конвертация цветовой палитры RGB в чёрно-

белый спектр, а API tf.data позволяет создавать сложные конвейеры ввода из простых, многократно используемых частей.

В заключение можно сделать вывод – задачи распознавания образов предполагают несколько методов работы и инструментов для программной реализации алгоритмов. Учитывая актуальность данного направления, разобранные в статье инструменты будут продолжать развиваться, предлагая пользователю необходимый для решения его задач функционал.

ЛИТЕРАТУРА

1. **Ту, Д.** Принципы распознавания образов / Д. Ту, Р. Гонсалес. – Москва : Мир, 1978.
2. **Гонсалес, Р.** Цифровая обработка изображений / Р. Гонсалес, Р. Вудс. – М.: Техносфера, 2005. – 1072 с.
3. **Wang, F.** Fuzzy supervised classification of remote sensing images / F.Wang // IEEE transactions of geoscience and remote sensing. – 1990. – Vol.28(2). – P. 194-201.
4. **Cannon, R. L.** Segmentation of a semantic mapper image using fuzzy C-means clustering algorithm / R. L. Cannon // IEEE Trans. Geosci remote sensing. V.10. – 1986. – GE-24. – P. 400-408.
5. **Фомин, Я. А.** Статистическая теория распознавания образов / Я. А. Фомин, Г. Р. Тарловский. – М.: Радио и связь. – 1986. – 624 с.
6. **Щеголькова, Д. В.** О подходах к выбору структуры нейронной сети на примере многослойного персептрона / Д. В. Щеголькова, Е. И. Орешкина, Л.В. Липинский // Актуальные проблемы авиации и космонавтики. – 2014. – Т. 1, № 10. – С. 305-306.
7. OpenCV. Официальный сайт. About – OpenCV.. URL: <https://opencv.org/about/> (дата обращения 23.04.2024).
8. MathWorks. Официальный сайт. Computer Vision Toolbox. URL: <https://www.mathworks.com/products/computer-vision.html> (дата обращения 23.04.2024).
9. Википедия. Свободная энциклопедия. TensorFlow. 2024. <https://ru.wikipedia.org/wiki/TensorFlow> (дата обращения 23.04.2024).
10. TensorFlow Core. Компьютерное зрение с TensorFlow. 2024. <https://www.tensorflow.org/tutorials/images?hl=ru> (дата обращения 23.04.2024).