

АЛГОРИТМЫ КЛАССИФИКАЦИИ ОБЪЕКТОВ НА ИЗОБРАЖЕНИЯХ ПРИЕМНОЙ КАМЕРЫ СОВРЕМЕННОГО ФАНДОМАТА

© 2024 З. С. Гостев✉, А. А. Сирота

*Воронежский государственный университет
Университетская пл., 1, 394018 Воронеж, Российская Федерация*

Аннотация. В работе рассматривается задача разработки математического и программного обеспечения для анализа изображений в современных фандоматах. В рамках подхода, основанного на применении методов поверхностного машинного обучения, формулируется задача обработки изображений объектов, подлежащих сырьевой переработке в фандомате, как задача сегментации этих изображений с последующей классификацией по форме. Производится обзор и сравнение известных методов сегментации с целью выделения формы объектов и формирования признаков для классификации. В результате сравнения делается выбор в пользу метода активного контура. Предлагается и исследуется относительно простой алгоритм классификации сегментированных объектов на основе «случайного леса». Также предлагается два алгоритма выделения признаков классификации: алгоритм на основе анализа степени заполненности частей сегментированного изображения и алгоритм, вычисляющий свойства области объекта.

Как альтернативный подход также описывается метод классификации получаемых изображений в целом, без предварительного выделения признаков, основанный на глубоком обучении. Обсуждается проблема дефицита обучающих данных, приводятся возможные подходы к её решению. Описывается реализация сверточной нейронной сети — классификатора с архитектурой DenseNet, полученная с применением техники переноса обучения.

Представлены результаты экспериментов по оценке эффективности рассмотренных алгоритмов, проведенных на предоставленном обучающем наборе. Выбранный метод сегментации, а также оба рассмотренных классификатора продемонстрировали высокий уровень эффективности. Проведено сравнение результатов классификации для алгоритмов на основе поверхностного («случайный лес») и глубокого машинного обучения (сверточная нейронная сеть).

Ключевые слова: обработка изображений, фандоматы, сегментация изображений, машинное обучение, случайный лес, глубокое обучение, перенос обучения, аугментация данных, сверточные нейронные сети.

ВВЕДЕНИЕ

Одной из важнейших проблем на сегодняшний день является проблема загрязнения окружающей среды. Немаловажный вклад в этот процесс вносит накопление отходов потребления и производства. В России, согласно данным Росприроднадзора за 2016, накоплено 39.4 млрд тонн отходов производства и потребления [1].

Распространение практики повторной переработки отходов — важный шаг на пути к решению этой проблемы. Современные фандоматы являются одним из промежуточных звеньев в этой цепи. Они позволяют собирать пригодную для переработки тару и осуществлять ее сортировку в соответствии с типом материала изготовления (стекло, пластик, алюминий и т. п.).

Важным аспектом при конструировании фандоматов является обеспечение как можно более точной и практичной сортировки собираемой тары в виде пустых емкостей.

✉ Гостев Захар Сергеевич
e-mail: gostev_z_s@sc.vsu.ru



Контент доступен под лицензией Creative Commons Attribution 4.0 License.
The content is available under Creative Commons Attribution 4.0 License.

Одним из подходов является идентификация объекта по содержимому нанесённого на него штрихкода. Однако данный метод имеет ряд значительных ограничений: невозможно определить материал изготовления, если штрихкод отсутствует в базе данных, принять тару без этикетки или с повреждённым штрихкодом. Классификация емкостей при помощи технологий обработки изображений, машинного обучения и компьютерного зрения помогают преодолеть эти трудности.

Современные фандоматы устроены следующим образом: пользователь через приемное отверстие помещает в фандомат объект, предназначенный для переработки или повторного использования. Он попадает в приемную камеру, в которой производятся дальнейшие измерения с помощью различных установленных датчиков (весов, фотокамеры и т. д.). На основании измерений, проведённых в этом отсеке, производится классификация и сортировка принятого объекта. В результате он помещается в один из сортировочных контейнеров.

Подход сортировки на основе внешнего вида представлен, в частности, в публикации [2]. В работе исследуются применимость свёрточных нейронных сетей в условиях ограниченных вычислительных ресурсов для сортировки объектов на три класса: «бутылка», «пластиковая банка» и «прочее». Авторы сравнивают эффективность трёх предобученных сетей (AlexNet, SqueezeNet, MobileNet), обучая и проверяя их на скачанных из интернета 650 фотографиях бутылок, 650 фотографиях банок и 12000 фотографиях других объектов. Наилучшая точность, достигаемая в этой работе — 0.98.

Однако у данного исследования есть ряд ограничений:

- набор данных, собранных не на прототипе, может существенно отличаться от изображений, получаемых в фандомате;
- изображения классифицируемых объектов сделаны на разном фоне;
- модель не различает стеклянные и пластиковые бутылки;
- не производится сравнение эффективности свёрточных сетей с другими алгоритмами машинного обучения.

Исходя из этого, представляет интерес исследование, в котором на адекватных исходных данных проводится сравнение различных подходов к решению задачи классификации объектов.

Целью работы является рассмотрение и анализ алгоритмов предварительной сегментации изображений объектов, получаемых в приёмной камере фандомата, а также анализ эффективности методов классификации на основе алгоритмов поверхностного и глубокого обучения.

1. МЕТОДЫ И МАТЕРИАЛЫ

1.1. Анализ предметной области

Автомат по приёму тары (или *фандомат*) — роботизированный агрегат, выменивающий возвратную тару (бутылки и банки из алюминия, стекла и других материалов) у населения в обмен на вознаграждение. Является одним из звеньев в цепи отдельного сбора и переработки отходов потребления.

Для осуществления качественной сортировки собираемой фандоматом тары необходимо корректно классифицировать её по типу (банка, бутылка) и материалу изготовления (пластик (ПЭТ), алюминий, стекло).

На рис. 1 приведены примеры емкостей различных классов. Из этих примеров видно, что внутри класса она может различаться по



Рис. 1. Примеры классифицируемых объектов различных классов

[Fig. 1. Examples of classified objects of various classes]

своему цвету, цвету крышки, размеру, форме, прозрачности, вытянутости горловины, размеру и положению этикетки, яркости освещения на фотографии.

Также пластиковая и алюминиевая емкость может быть деформирована. Пользователь фандомата может сдать наполненную тару, с остатками жидкости, а также без крышки или этикетки. В приемной камере могут содержаться фрагменты других принятых образцов. Все эти особенности необходимо учитывать при проектировании алгоритмов, анализирующих внешний вид тары.

1.2. Классификация на основе выделения сегментационных признаков

Предлагаемый общий алгоритм проведения классификации (рис. 2) состоит из следующих этапов.

1. Предварительная сегментация исходного изображения тары.
2. Выделение признаков для классификации.
3. Классификация объекта на изображении.

1.2.1. Методы сегментации изображений

Сегментация изображения — это процесс разделения цифрового изображения на однородные области. Её цель заключается в упрощении и/или изменении представления изображения, чтобы его было легче анализировать. Сегментация обычно используется для того, чтобы выделить объекты и границы (линии, кривые, и т. д.) на изображениях.

В данном случае сегментация нацелена на выделение маски объекта. Исходное цветное изображение преобразуется к бинарному, в котором пиксели с максимальными зна-

чениями яркости соответствуют объекту, а минимальными — фону. Сегментированное изображение используется как входное для алгоритмов выделения признаков классификации.

В ходе исследований предварительно рассматривался ряд алгоритмов, представлявших наиболее подходящими для данной задачи.

1. *Пороговый метод* сегментации [3, 4] основан на разграничении мод распределения яркостей на основе выбора глобального (в частном случае) порога T .

2. *Детектор контуров Кэнни* [3, 5] является методом выделения контуров объектов на изображении. Алгоритм состоит из следующих шагов.

- Сглаживание исходного изображения гауссовым фильтром.

- Формирование изображения модуля и направления градиента.

- Подавление немаксимальных точек изображения модуля градиента.

- Выполнение преобразования с двойным порогом и анализа связности для обнаружения и связывания контуров.

3. *Метод разрастания областей* [3, 6, 7] является одним из наиболее подходящих методов сегментации для данной задачи. В качестве входных данных этот метод принимает изображения и набор *семян (seeds)*. Семена отмечают объекты, которые нужно выделить. Области постепенно разрастаются, сравнивая все незанятые соседние пиксели с областью. Разность δ между яркостью пикселя и средней яркостью области используется как мера схожести. Пиксель с наименьшей такой разностью добавляется в соответствующую область. Процесс продолжается, пока все пиксели не будут добавлены в один из регионов.



Рис. 2. Общая схема алгоритма обработки изображений
 [Fig. 2. General diagram of the image processing algorithm]

Так как в данной задаче яркость объекта может сильно меняться, а фон — постоянный, семена выбираются в области фона. Соответственно, сегментируется не область объекта, а область фона. Полученная маска в последствии инвертируется.

4. *Метод разделения и слияния* [3], в отличие от метода выращивания областей, заключается в том, что изображение разбивается на множество произвольных непересекающихся областей для последующего их слияния и/или разделения.

5. *Методы разреза графа* [8] могут быть эффективно применены для сегментации изображений. В этих методах изображение представляется как взвешенный неориентированный граф. Обычно пиксель или группа пикселей ассоциируется вершиной, а веса рёбер определяют (не)похожесть соседних пикселей. Затем граф (изображение) разрезается согласно критерию, созданному для получения «хороших» кластеров. Каждая часть вершин (пикселей), получаемая этими алгоритмами, считается объектом на изображении.

6. *Метод активного контура*. Данный метод [9] минимизирует функционал энергии деформируемого контура (или *снейка*). Функционал составлен таким образом, чтобы его локальному минимуму соответствовал контур, выделяющий искомые границы объекта на изображении.

В ходе многочисленных экспериментов было установлено, что из числа перечисленных наиболее подходящим для предварительной сегментации является метод активного контура. Пороговый метод, метод разрастания областей имеют низкую эффективность работы для изображений с прозрачными объектами и объектами с цветом близким к фону. Метод разделения и слияния имеет слишком большую погрешность области сегментации. Наконец, для методов разреза графа и Кэнни слишком сложно подобрать универсальные значения параметров.

Стоит отметить, что набор изображений, на которых тестировался алгоритм, был подготовлен в более жестких условиях, по сравнению с условиями, которые потенциально можно обеспечить внутри приёмного отсека

фандомата. Изображения, получаемые таким образом, при необходимости, можно сегментировать более простыми методами: например, вычитанием снимка фона из изображения в сочетании с дальнейшей обработкой.

Реализация метода активного контура.

Рассмотрим более подробно реализованную в конечном варианте общего алгоритма обработки версию этого алгоритма. Энергия активного контура состоит из внешней и внутренней. Внутренняя энергия зависит только от формы кривой и определяет ее упругость и жесткость. Внешняя энергия зависит от интенсивности изображения, на котором определен активный контур.

Определим энергию активного контура:

$$E_{snake}(C) = \int_0^1 (E_{int}(C(s)) + E_{ext}(C(s))) ds, \quad (1)$$

где $C(s)$ — контур, $E_{int}(C(s))$ — внутренняя энергия контура, определённая как

$$E_{int}(C(s)) = \alpha(s)|C'(s)|^2 + \beta(s)|C''(s)|^2, \quad (2)$$

где $\alpha(s)$ регулирует упругость контура и $\beta(s)$ регулирует жесткость контура, а $E_{ext}(C(s))$ — внешняя энергия, определённая как

$$E_{ext}(C(s)) = P(C(s)). \quad (3)$$

Функция P называется потенциалом и определяет внешнее воздействие на активный контур, зависящее от свойств изображения. Потенциал определяется следующим образом:

$$P(x, y) = -w_{line}P_{line} - w_{edge}P_{edge}, \quad (4)$$

где w_{line} и w_{edge} — весовые коэффициенты соответствующих потенциалов. P_{line} и P_{edge} определяются как

$$P_{line}(x, y) = -G_{\sigma}(x, y) * f(x, y), \quad (5)$$

$$P_{edge}(x, y) = -|\nabla(G_{\sigma}(x, y) * f(x, y))|^2, \quad (6)$$

где G_{σ} — это двумерная функция Гаусса со стандартным отклонением σ , f — исходное изображение.

Для решаемой задачи использовался коэффициент $w_{line} = 0$, поскольку требуется найти области перепадов, а не постоянного уровня яркости. Значения коэффициентов α и β подобраны экспериментальным путём как 0.5 и 1, соответственно. На рис. 3 приведены примеры сегментации методом активного контура с различными параметрами. В каче-

стве начального приближения контура использовалась рамка, окаймляющая всё изображение.

1.2.2. Предлагаемые алгоритмы извлечения признаков и используемый классификатор

Алгоритм извлечения признаков с помощью «слайсов». Алгоритм разделяет сегментированное изображение на заданное количество прямоугольников с одинаковой площадью, а потом вычисляет степень заполненности (от 0 до 1) каждого прямоугольника. Полученные значения являются признаками для классификации.

Последовательность шагов алгоритма следующая.

1. Выравнивание пространственной ориентации объекта на сегментированном изображении.

2. Кадрирование, в ходе которого вырезается минимальный ограничивающий прямо-

угольник, содержащий область объекта.

3. Разделение изображения на внутренние сегменты прямоугольной формы («слайсы»).

4. Вычисление степени заполненности сегментов (отношение количества пикселей объекта внутри сегмента к общей площади сегмента).

«Нарезку» можно проводить по одной или обоим осям изображения с различным количеством сегментов (см. рис. 3).

Алгоритм извлечения признаков на основе свойств сегментированного изображения. Данный алгоритм заключается в вычислении свойств области объекта на сегментированном изображении, распределения значений которых должны отличаться для каждого класса изображений. В работе использовались следующие свойства области объекта (рис. 3): длина, ширина минимального охватывающего прямоугольника, относительное смещение центра масс по каждой оси изображения, дефект выпуклости, «дефект минимального

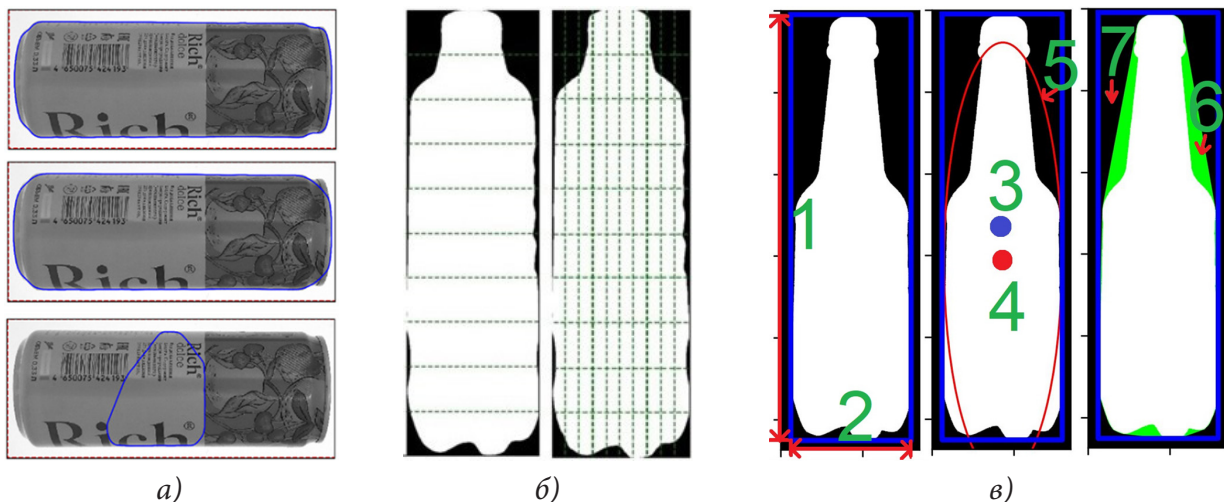


Рис. 3. Примеры различных преобразований изображений.

а) Примеры сегментации изображения алюминиевой банки методом активного контура с различными коэффициентами α и β . б) Примеры «нарезки» изображения бутылки на различное количество слайсов. в) Иллюстрация свойств сегментированной области объекта, используемых для извлечения признаков классификации. 1, 2 и 3 — длина, ширина и центр минимального прямоугольника, 5, 4 — эллипс с площадью, равной площади объекта и центр эллипса (он же центр масс), 6 — выпуклая оболочка объекта, 7 — иллюстрация «дефекта минимального прямоугольника»

[Fig. 3. Examples of various image transformations. a) Examples of segmentation of an aluminum can using the active contour method with various coefficients α and β . b) Examples of “slicing” a bottle image into different numbers of slices. c) Illustration of segmented object region properties used to extract classification features. 1, 2 and 3 — length, width and center of the minimal rectangle, 5, 4 — ellipse with an area equal to the area of the object and the center of the ellipse (also known as the center of mass), 6 — convex hull of the object, 7 — illustration of the “minimum rectangle defect”]

прямоугольника», эксцентриситет эллипса с площадью, равной площади объекта.

Алгоритм классификации на основе «случайного леса». Алгоритм «случайный лес» [10, 11] является ансамблевым: он объединяет в один классификатор несколько деревьев принятия решений. Это позволяет построить более надёжную модель, которая обладает большей эффективностью обобщения и меньшей восприимчивостью к переобучению. Данный метод предлагается использовать для классификации.

Общий набор признаков корректируется: из него исключаются признаки, значимость которых ниже заданного значения. Количество признаков и способ их извлечения, порог влияния признаков для отсеивания, количество решающих деревьев в ансамбле являются гиперпараметрами предлагаемого алгоритма.

1.3. Классификация с использованием методов глубокого обучения

Классификация с использованием глубоких нейронных сетей имеет ряд положительных сторон, позволяющих решить некоторые проблемы, возникающих при использовании других алгоритмов. Однако данный подход имеет и некоторые ограничения.

Достоинства:

- отпадает необходимость предварительной сегментации;
- признаки для классификации выделяются автоматически.

Недостатки:

- необходим более объёмный набор данных;
- усложняется интерпретация признаков классификации и самой модели.

1.3.1. Проблема дефицита обучающих данных и методы её решения

Набор данных, собранный в рамках исследования, состоит всего из 164 фотографий. Классы набора не сбалансированы: пластиковых объектов примерно в 3 раза больше, чем стеклянных или алюминиевых. В связи с этим возникла проблема недостаточного количества данных.

Основные подходы к решению описанной проблемы основаны на генерации и аугментации синтетических данных, использовании техники переноса обучения.

Аугментация данных. Аугментация данных [12] — это процесс дополнения набора данных сходными, созданными из уже имеющихся. Аугментация позволяет «размножить» набор данных до нужного объёма, но не превышая некоторого максимального порога.

Рассматривались следующие аугментирующие преобразования изображений: обреза, поворот по часовой и против часовой стрелки, отражение по вертикали и горизонтали, смещение, масштабирование, пластическая деформация.

В результате из описанных преобразований было выбрано одно преобразование отражения по вертикали. Изображения тренировочной и тестовой выборки были обработаны по-отдельности, чтобы доли добавленных образцов были одинаковы в них обеих. В итоге исходный набор данных был увеличен примерно на 30 %.

Перенос обучения. Техника переноса обучения [13] позволяет использовать внутренние признаки глубокой нейронной сети, полученные в ходе ее обучения для одной задачи, при решении другой, схожей задачи классификации. С помощью этого метода можно обучать модель, даже когда данных для её обучения с нуля недостаточно.

Перенос обучения состоит из следующих шагов:

1. Обучение базовой модели на достаточном количестве данных другой задачи. Зачастую можно использовать уже обученную нейронную сеть.

2. Создание новой модели с помощью замены полносвязных слоёв исходной модели на новые и её обучение на текущих данных. Слои базовой модели при этом «замораживаются», то есть не обучаются.

3. Обучение итоговой модели целиком (все слои активируются, замороженных нет) на текущих данных с очень низкой скоростью обучения. Этот шаг является необязательным и в данной работе он не выполняется.

1.3.2. Архитектура глубокой свёрточной нейронной сети

В качестве базовой модели для переноса обучения использовалась свёрточная нейронная сеть архитектуры *DenseNet* [14]. В то время как традиционные свёрточные сети с L слоями имеют L соединений — по одному между каждым слоем и его последующим слоем — данная сеть имеет $L(L + 1)/2$ прямых соединений. Для каждого слоя карты признаков всех предыдущих слоев используются в качестве входных данных, а его собственные карты признаков используются в качестве входных данных для всех последующих слоев. У *DenseNet* есть несколько неоспоримых преимуществ: эта архитектура облегчает проблему исчезающего градиента, улучшает распространение признаков, увеличивает повторное использование признаков и существенно сокращает количество параметров.

Новая модель создается на основе *DenseNet* с помощью замены классификационного слоя на набор слоёв: *flatten layer*, *dropout layer* и *dense layer* с активацией *softmax*, рассчитанной на три рассматриваемых класса объектов. *Dropout layer* используется для уменьшения переобучения нейронной сети. Итоговая модель дана на рис. 4.

1.3.3. Обучение нейронной сети

Полученная нейросетевая модель обучалась в течение 30 эпох на тренировочных данных, описанных в разделе 1.1. Для отыскания минимума оптимизируемой функции проводилось сравнение различных оптимизаторов (*SGD*, *Adam*, *RMSProp*) из библиотеки *Keras* с различными скоростями обучения. Использовалась методика создания *чекпоинтов* (*checkpoints*) для сохранения лучшей мо-

дели. Если нейронная сеть в течение эпохи не улучшается, то текущие веса не сохраняются. Кроме того, использовалась практика ступенчатого уменьшения скорости обучения с выходом на «плато», при которой, если точность модели изменяется незначительно в течение некоторого заранее заданного числа эпох, то скорость обучения временно уменьшается для отыскания локальных минимумов.

2. РЕЗУЛЬТАТЫ И ИХ ОБСУЖДЕНИЕ

Для оценки эффективности реализованных методов был собран набор изображений различных объектов каждого класса, сфотографированных на одинаковом равномерном фоне. Примеры таких изображений показаны на рис. 1. Этот датасет содержит 35 фотографий алюминиевых банок, 32 фотографии стеклянных и 97 фотографий пластиковых бутылок. Фотографии получены на прототипе, разработанном компанией ООО «Формула Сортировки» и предоставлены авторам для проведения исследований.

Для проверки качества сегментации вручную были размечены маски объектов на всех изображениях.

2.1. Результаты сегментации методом активного контура

В качестве меры правильности сегментации была использована метрика *IoU*, которая задаётся следующим выражением:

$$IoU = \frac{TP}{TP + FP + FN}, \quad (7)$$

где *TP* — положительные пиксели (пиксели объекта), *FP* — ложноположительные пиксели, *FN* — ложноотрицательные. Оценка результатов правильности сегментации представлена в табл. 1.

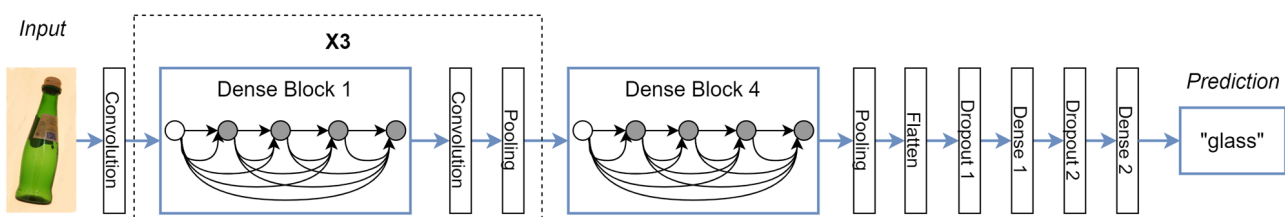


Рис. 4. Схема итоговой архитектуры свёрточной нейронной сети
 [Fig. 4. Diagram of the final architecture of the convolutional neural network]

Таблица 1. Результаты сегментации
[Table 1. Segmentation results]

	IoU	Количество
Алюминиевая банка	0.98	35
Стеклоанная бутылка	0.98	32
Пластиковая бутылка	0.96	97
Взвешенное среднее	0.97	164

Здесь взвешенное среднее определяется как

$$\frac{1}{\sum_{l \in L} |y_l|} \sum_{l \in L} |y_l| M(y_l, \hat{y}_l), \quad (8)$$

где y — множество истинных значений пар образца и метки, \hat{y} — множество предсказанных пар значений образца и метки, L — множество меток, M — выбранная метрика.

Наихудшее значение метрики *IoU* сегментации было получено для изображений пластиковых бутылок. Это связано с тем, что пустая тара из этого материала обладает высокой прозрачностью, что уменьшает контрастность у её границ на изображении.

2.2. Результаты классификации методом на основе случайного леса

Были проведены эксперименты классификации методом случайного леса для наборов данных, полученных с помощью обоих предложенных алгоритмов выделения признаков. Ансамбль состоял из 100 решающих деревьев, порог значимости признаков для отсека — 0.05. Наилучшие результаты для метода на основе слайсов были достигнуты при нарезке изображения по длиннейшей на 8 слайсов, из набора признаков по порогу значимости были исключены 2 средних слайса (№ 4 и № 5). Для данных, полученных с помощью алгоритма на основе свойств сегментированного изображения, был отброшен признак смещения центра масс по короткой оси изображения. Результаты классификации представлены в табл. 2 и 3. Выбранные метрики определяются согласно формулам:

$$accuracy = \frac{TP + TN}{TP + FP + TN + FN}, \quad (9)$$

$$precision = \frac{TP}{TP + FP}, \quad recall = \frac{TP}{TP + FN}, \quad (10)$$

$$f1_score = 2 \frac{precision * recall}{precision + recall}, \quad (11)$$

где *TP* — истинно положительные, *FP* — ложноположительные, *TN* — истинно отрицательные, *FN* — ложноотрицательные образцы. Макро-среднее — среднее среди всех классов.

Таблица 2. Результаты классификации для «слайсового» алгоритма выделения признаков
[Table 2. Classification results for the “slicing” feature extraction algorithm]

	precision	recall	f1-score
Алюминиевая банка	1	1	1
Стеклоанная бутылка	0,86	1	0,93
Пластиковая бутылка	1	0.95	0.97
Макро-среднее	0.95	0.98	0.97
Взвешенное среднее	0.97	0.97	0.97
accuracy	0.97		

Таблица 3. Результаты классификации для алгоритма выделения признаков на основе свойств сегментированного изображения
[Table 3. Classification results for the feature extraction algorithm based on the properties of the segmented image]

	precision	recall	f1-score
Алюминиевая банка	1.00	1.00	1.00
Стеклоанная бутылка	1.00	1.00	1.00
Пластиковая бутылка	1.00	1.00	1.00
Макро-среднее	1.00	1.00	1.00
Взвешенное среднее	1.00	1.00	1.00
accuracy	1.00		

2.3. Результаты классификации с использованием нейронной сети

Результаты оценки модели на тестовой выборке представлены в табл. 4.

Таблица 4. Результаты классификации с помощью сверточной нейронной сети
 [Table 4. Classification results using convolutional neural network]

	precision	recall	f1-score	мера
Алюминиевая банка	1.00	1.00	1.00	14
Стеклоянная бутылка	0.91	0.83	0.87	12
Пластиковая бутылка	0.95	0.97	0.96	38
Макро-среднее	0.95	0.95	0.94	64
Взвешенное среднее	0.95	0.95	0.95	64
accuracy	0.95			64

2.4. Обсуждение результатов классификации и сравнение классификаторов

Как видно из результатов, рассмотренные методы позволяют достичь достаточно высокой эффективности сегментации и классификации для данной задачи. Средневзвешенная f1-мера для алгоритма классификации с использованием сегментационных признаков составляет 0.97...1.00 (в зависимости от выбранного алгоритма выделения признаков), для сверточной нейронной сети — 0.95.

По результатам перекрестной проверки алгоритм на основе случайного леса допустил максимум две ошибки. Чаще всего данный классификатор путает стеклянные бутылки и пластиковые, которые имеют вытянутую, как у стеклянных, горловину. Эффективность классификации при использовании разных методов выделения признаков несколько отличается в пользу метода на основе свойств сегментированного изображения.

Из полученных результатов можно сделать вывод, что оба классификатора хорошо справляются со своей задачей. Классификатор на основе случайного леса является существенно менее требовательным с точки зрения вычислительных ресурсов и объема обучающей выборки. Однако он учитывает лишь форму объекта на изображении и игнорирует цветовой канал. Кроме того, для использования этого алгоритма необходимо

проведение предварительной сегментации и выделение признаков. Поэтому, в условиях достаточной производительности компьютера, установленного в сортирующей системе, а также при наличии более объемной обучающей выборки, нейросетевой классификатор, все-таки, в перспективе является предпочтительным.

Стоит отметить, что при подготовке набора данных для обучения и тестирования нейросетевого классификатора все изображения были масштабированы к одному размеру и, в связи с особенностями процесса сбора фотографий на прототипе, была потеряна информация об относительном размере объектов. Присутствие этой информации могло бы улучшить результаты.

ЗАКЛЮЧЕНИЕ

В данной работе была рассмотрена задача разработки математического и программного обеспечения для процесса обработки изображений, получаемых в приёмной камере фандомата.

В ходе исследования проведено рассмотрение известных методов для решения задачи сегментации изображений. В результате экспериментов было выяснено, что в реализованных на прототипе условиях фотосъёмки наиболее подходящим методом предварительной сегментации является метод активного контура.

Предложен поэтапный алгоритм классификации, включающий выделение признаков классификации и проведение классификации с использованием алгоритма поверхностного машинного обучения — алгоритма «случайный лес». Предложены два собственных алгоритма выделения признаков.

В качестве альтернативы рассмотрена сверточная нейронная сеть, построенная с помощью переноса обучения на основе модели DenseNet. Обучение проводилось с использованием технологии аугментации данных.

Экспериментальные исследования проведены на собранном на прототипе фандомата наборе данных, состоящем из образцов трёх классов. Классификатор на основе слу-

чайного леса и нейросетевой классификатор обеспечивают достаточно высокую точность классификации, что говорит о целесообразности применения методов поверхностного и глубокого машинного обучения в задаче сортировки тары в фандомате. Полученный результат сопоставим с лучшими результатами известных работ. Несмотря на несколько лучшие результаты классификатора на основе случайного леса, в перспективе нейросетевой классификатор является предпочтительным, поскольку при увеличении обучающей выборки способен продемонстрировать большую обобщающую способность.

В ходе дальнейшего развития представляется полезным расширить набор рассмотренных методов сегментации (например, методом водоразделов), расширить обучающий набор данных, а также провести исследования для образцов с возможными при эксплуатации системы дефектами и искажениями формы.

Авторы выражают благодарность сотрудникам компании ООО «Формула Сортировки» за предоставленные для обработки реальные данные.

КОНФЛИКТ ИНТЕРЕСОВ

Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи

СПИСОК ЛИТЕРАТУРЫ

1. Волкова А. В. Рынок утилизации отходов – 2018 [Электронный ресурс]: экономическое исследование / А. В. Волкова. – ВШЭ, 2018. – Режим доступа: <https://dcenter.hse.ru/data/2018/07/11/1151608260/Рынок%20утилизации%20отходов%202018.pdf>
2. Ахметзянов К. Р. Сравнение сверточных нейронных сетей для задач сортировки мусорных отходов / К. Р. Ахметзянов, А. А. Южаков // Известия СПбГЭТУ ЛЭТИ. – 2018. – № 6. – С. 27–32. – EDN XVVOPR.
3. Гонсалес Р. С. Цифровая обработка изображений / Р. С. Гонсалес, Р. Е. Вудс. – 3-е издание, исправленное и дополненное. – Москва: Техносфера, 2012. – 1104 с.
4. Chow C. K, Kaneko T. Automatic boundary detection of the left ventricle from cineangiograms / C. K. Chow, T. Kaneko // *Comput Biomed Res.* – 1972 Aug. – P. 388–410.
5. Canny J. A Computational Approach to Edge Detection / J. Canny // *IEEE Trans. Pattern Analysis Machine Intelligence.* – Vol. 8, no 6 – P. 679–698.
6. Fu K. A survey on image segmentation / K. Fu, J. K. Mui // *Pattern Recognit.* – 1981. – P. 3–16.
7. Hojjatoleslami S. A. Region growing: a new approach / S. A. Hojjatoleslami, J. Kittler. // *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society.* – 1998. – P. 1079–1084.
8. Shi J. Normalized Cuts and Image Segmentation / Jianbo Shi and Jitendra Malik // Архивная копия от 6 июня 2011 на Wayback Machine, *IEEE Transactions on pattern analysis and machine intelligence.* – 2000. – P. 888–905.
9. Snakes: Active Contour Models / M. Kass, A. Witkin, D. Terzopoulos; Schlumberger Palo Alto Research // *International Journal Of Computer Vision.* – Boston : Kluwer Academic Publishers, 1987. – P. 321–331.
10. Classification and Regression Trees : book / L. Breiman [et al.]. – 1st Edition. – New York : Chapman & Hall/CRC, 1984. – 368 p.
11. Understanding Random Forests : From Theory to Practice : PhD dissertation / G. Louppe. – Belgium : University of Liège, 2014. – 223 p.
12. Shorten C. A survey on Image Data Augmentation for Deep Learning / C. Shorten, T. M. Khoshgoftaar // *J Big Data* 6. – 60. – 2019.
13. Weiss Karl. A survey of transfer learning / Karl Weiss, Taghi Khoshgoftaar, DingDing Wang // *Journal of Big Data.* 3. 10.1186/s40537-016-0043-6. – 2016.
14. Huang G. Densely Connected Convolutional Networks / Gao Huang, Zhuang Liu, Laurens van der Maaten, Kilian Q. Weinberger // *Computer vision foundation.* – 2016
15. Uijlings J. Selective Search for Object Recognition / Uijlings Jasper, Sande K., Gevers T., Smeulders A.W.M. // *International Journal of Computer Vision.* 104. 154-171. 10.1007/s11263-013-0620-5, 2013.

16. Hosseini S. Histograms of Oriented Gradients for Human Detection / S. Hosseini, E. Fazl // 10.13140/RG.2.2.23788.85122. – 2021.

17. Гонсалес Р. С. Обработка изображений в среде MATLAB / Р. С. Гонсалес, Р. Е. Вудс, С. Эддинс. – Перевод с английского В. В. Чепыжова. – Москва : Техносфера, 2006. – 616 с.

18. Felzenszwalb P. F. Object detection with discriminatively trained part-based models / P. F. Felzenszwalb [и др.] // *IEEE Trans. Pattern Anal. Mach. Intell.* – 2010. – № 32. – С. 1627–1645.

19. OpenCV documentation index [Электронный ресурс]. – Режим доступа: <https://docs.opencv.org>

Гостев Захар Сергеевич — аспирант кафедры технологий обработки и защиты информации, факультета компьютерных наук, Воронежского государственного университета.

E-mail: gostev_z_s@sc.vsu.ru

ORCID iD: <https://orcid.org/0009-0003-5104-2278>

Сирота Александр Анатольевич — д-р техн. наук, проф., заведующий кафедрой технологий обработки и защиты информации, факультета компьютерных наук, Воронежского государственного университета.

E-mail: sir@cs.vsu.ru

ORCID iD: <https://orcid.org/0000-0002-5785-8513>

DOI: <https://doi.org/10.17308/sait/1995-5499/2024/1/114-126>

ISSN 1995-5499

Received 28.10.2023

Accepted 01.04.2024

ALGORITHMS FOR CLASSIFICATION OF OBJECTS IN IMAGES OF THE RECEIVING CAMERA OF A MODERN REVERSE VENDING MACHINES

© 2024 Z. S. Gostev✉, A. A. Sirota

Voronezh State University

1, Universitetskaya Square, 394018 Voronezh, Russian Federation

Annotation. The work examines the problem of developing mathematical and software for image analysis in modern reverse vending machines. As part of the approach based on the use of shallow machine learning methods, the problem of processing images of objects subject to raw material processing in reverse vending machines is formulated as a task of segmenting these images, followed by subsequent classification by shape. A review and comparison of known segmentation methods is carried out in order to extract the shape of objects and generate features for classification. As a result of the comparison, a choice is made in favor of the active contours method. A relatively simple algorithm for classifying segmented objects based on “random forest” is proposed and studied. We also propose two algorithms for extracting classification features: an algorithm based on the analysis of the degree of fullness of parts of a segmented image and an algorithm that calculates the properties of the object area.

As an alternative approach, a method for classifying the resulting images as a whole, without preliminary feature selection, based on deep learning, is also described. The problem of shortage of training data is discussed, and possible approaches to solving it are given. We describe the implementation of a convolutional neural network — a classifier with the DenseNet architecture, obtained using the transfer learning technique.

The results of experiments to evaluate the effectiveness of the considered algorithms, conducted on the provided training set, are presented. The selected segmentation method, as well as both considered classifiers, demonstrated a high level of efficiency. When comparing the results of classification algorithms based on shallow (“ran-

✉ Гостев Захар Сергеевич
e-mail: gostev_z_s@sc.vsu.ru

dom forest”) and deep machine learning (convolutional neural network), a choice was made in favor of the neural network approach when certain conditions were met for the training data.

Keywords: image processing, reverse vending machines, image segmentation, machine learning, random forest, deep learning, transfer learning, data augmentation, convolutional neural networks.

CONFLICT OF INTEREST

The authors declare the absence of obvious and potential conflicts of interest related to the publication of this article.

REFERENCES

1. Volkova A. V. (2018) Rynok utilizatsii otkhodov – 2018 [Waste disposal market <https://dcenter.hse.ru/data/2018/07/11/1151608260/Рынок%20утилизации%20отходов%202018.pdf> (in Russian)]
2. Akhmetzyanov K. R. and Yuzhakov A. A. (2018) Sravnenie svertochnykh neironnykh setei dlya zadach sortirovki musornykh otkhodov [Comparison of convolutional neural networks for waste sorting tasks]. *Izvestiya SPbGETU LETI*. (6). P. 27–32. (in Russian)
3. González R. C., Woods R. E. and Masters B. R. (2009). Digital Image Processing, Third Edition. *Journal of Biomedical Optics*. 14. 029901.
4. Chow C. K. and Kaneko T. (1972). Automatic boundary detection of the left ventricle from cineangiograms. *Computers and biomedical research, an international journal*. 5(4). P. 388–410. [https://doi.org/10.1016/0010-4809\(72\)90070-5](https://doi.org/10.1016/0010-4809(72)90070-5)
5. Canny J. (1986). A computational approach to edge detection. *IEEE transactions on pattern analysis and machine intelligence*. 8(6). P. 679–698.
6. Fu K. and Mui J. K. A survey on image segmentation. *Pattern Recognit.* 1981. P. 3–16.
7. Hojjatoleslami S. A. and Kittler J. (1998). Region growing: a new approach. *IEEE transactions on image processing: a publication of the IEEE Signal Processing Society*. 7(7). P. 1079–1084. <https://doi.org/10.1109/83.701170>
8. Jianbo Shi and Malik J. (2000). Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 22(8). P. 888–905. doi:10.1109/34.868688
9. Kass M., Witkin A. and Terzopoulos D. (1988) Snakes: Active contour models. *Int J Computer Vision* 1, P. 321–331. <https://doi.org/10.1007/BF00133570>
10. Breiman L. (1984). Classification and Regression Trees (1st ed.). *Routledge*. <https://doi.org/10.1201/9781315139470>
11. Louppe G. (2014). Understanding Random Forests: From Theory to Practice. arXiv: Machine Learning.
12. Shorten C. and Khoshgoftaar T. M. (2019). A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*. 6. P. 1–48.
13. Weiss K. R., Khoshgoftaar T. M. and Wang D. (2016). A survey of transfer learning. *Journal of Big Data*, 3.
14. Huang G., Liu Z., Van Der Maaten L. and Weinberger K. Q. (2017) Densely Connected Convolutional Networks. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA. P. 2261–2269. doi: 10.1109/CVPR.2017.243.
15. Uijlings J. R., Sande K. E., Gevers T. and Smeulders A. W. (2013). Selective Search for Object Recognition. *International Journal of Computer Vision*. 104. P. 154–171.
16. Dalal N. and Triggs B. (2005) Histograms of oriented gradients for human detection (2005). *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, San Diego, CA, USA. vol. 1. P. 886–893. doi: 10.1109/CVPR.2005.177.
17. González R. C., Woods R. E. and Eddins S. L. (2006). Digital image processing using MATLAB.
18. Felzenszwalb P. F., Girshick R. B., McAllester D. and Ramanan D. (2010). Object detection with discriminatively trained part-based models. *IEEE transactions on pattern analysis and machine intelligence*. 32(9). P. 1627–1645. <https://doi.org/10.1109/TPAMI.2009.167>
19. OpenCV command. (2015) OpenCV documentation index <https://docs.opencv.org>

Gostev Zakhar S. — PhD student, Department of Information Security and Processing Technologies, Faculty of Computer Sciences, Voronezh State University.

E-mail: gostev_z_s@sc.vsu.ru

ORCID iD: <https://orcid.org/0009-0003-5104-2278>

Sirota Alexander A. — DSc in Technical Sciences, Head of the Department of Information Security and Processing Technologies, Faculty of Computer Sciences, Voronezh State University.

E-mail: sir@cs.vsu.ru

ORCID iD: <https://orcid.org/0000-0002-5785-8513>