

Классификация изображений в приемной камере фандомата

З. С. Гостев, email: gostev_z_s@sc.vsu.ru¹

А. А. Сирота, email: sir@cs.vsu.ru¹

Е. В. Попова^{1,2}

Д. А. Минаков²

¹ ФГБОУ ВО «Воронежский государственный университет»

² ООО «Формула Сортировки», г. Воронеж

***Аннотация.** Разработано математическое и программное обеспечение для сегментации и классификации изображений в приёмной камере фандомата.*

***Ключевые слова:** фандоматы, сегментация изображений, метод активного контура, случайный лес*

Введение

Одной из важнейших проблем общества на сегодняшний день является загрязнение окружающей среды. Немаловажный вклад в этот процесс вносит накопление мусора. К 2016 году объём образования отходов в мире достиг 22 млрд тонн, что на 5% больше чем в 2015. В России, согласно данным Росприроднадзора за этот же год, накоплено 39.4 млрд тонн отходов производства и потребления [1].

Распространение практики повторной переработки отходов - важный шаг на пути к решению этой проблемы. Фандоматы являются промежуточным звеном в этой цепи. Они позволяют собирать пригодную для переработки тару, которую не нужно будет впоследствии сортировать.

Важным аспектом при конструировании фандоматов является обеспечение как можно более точной и практичной сортировки собираемой тары. Одним из подходов является идентификация тары по содержимому нанесённого на неё штрихкода. Однако данный метод имеет ряд значительных ограничений: невозможно прочесть повреждённый или не сохранённый в базе данных штрихкод, принять тару без этикетки. Сортировка тары при помощи технологий обработки изображений, машинного обучения и компьютерного зрения помогают преодолеть эти трудности.

Целью данного исследования является разработка математического и программного обеспечения для сегментации и классификации изображений в приёмной камере современного фандомата.¹

1. Анализ предметной области

Фандомат – роботизированный агрегат, выменивающий возвратную тару, подлежащую переработке, у населения в обмен на вознаграждение. После того, как пользователь помещает тару в приемный отсек, аппарат делает серию фотоснимков принятого объекта. Полученные изображения используются для определения материала, из которого изготовлена тара.

Для переработки пригодно 3 типа тары (рис. 1):

- алюминиевая банка,
- стеклянная бутылка,
- пластиковая (ПЭТ) бутылка.

Необходимо классифицировать объект тары на полученном цифровом изображении, отнеся его к одному из вышеупомянутых классов.



Рис.1 Примеры изображений тары трёх классов: слева – алюминиевая банка, посередине – стеклянная бутылка, справа – пластиковая (ПЭТ) бутылка

2. Алгоритм проведения классификации тары

Общий алгоритм проведения классификации состоит из следующих шагов:

1. Сегментация исходного изображения тары.
2. Выделение признаков для классификации.
3. Классификации объекта на изображении.

3. Сегментация

Для того, чтобы классифицировать тару на полученном изображении, необходимо отделить область объекта от области фона, то есть провести сегментацию. Исходное цветное изображение

преобразуется к бинарному, в котором пиксели с максимальными значениями яркости соответствуют таре, а минимальными – фону.

В предлагаемом алгоритме сегментация проводилась с помощью метода активных контуров [2]. Данный метод минимизирует функционал энергии деформируемого контура (или *снейка*). Функционал составлен таким образом, чтобы его локальному минимуму соответствовал контур, выделяющий искомые границы объекта на изображении.

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

$$P_{line}(x, y) = -|\nabla I(x, y)|^2, \quad (5)$$

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

где $G\sigma$ – это двумерная функция Гаусса со стандартным отклонением σ , f – исходное изображение.

Для данной задачи используется коэффициент $w_{line} = 0$, поскольку требуется найти области перепадов, а не постоянной яркости. Коэффициенты α и β установлены экспериментальным путём в 0.5 и 1, соответственно. На рис. 2 приведены примеры сегментации методом активного контура с различными параметрами. В качестве начального приближения контура использовалась рамка, окаймляющая всё изображение.

Набор изображений, на которых тестировался алгоритм, был подготовлен в условиях, несколько отличающихся от условий внутри приёмочного отсека фандомата. В реальном фандомате можно обеспечить более мягкие условия для сегментации: зафиксировать фотокамеру, сохранять освещение постоянной яркости и тона. Изображения, получаемые таким образом, можно сегментировать более простыми методами: например, вычитанием снимка фона из изображения.



Рис.2 Примеры сегментации изображения алюминиевой банки методом активного контура с различными α и β : слева - $\alpha = 0.1, \beta = 1$, посередине - $\alpha = 0.4, \beta = 1$, справа - $\alpha = 0.7, \beta = 1$. Красный контур – начальное приближение, синий – итоговый контур

4. Пороговый алгоритм классификации

Наиболее простым и естественным для данной задачи является использование *порогового алгоритма* классификации, который позволяет отличать изображения банок от изображений стеклянных и пластиковых бутылок. Данный метод основан на внешней несхожести этих двух классов: бутылка имеет сужение на одном из концов, а банка не имеет.

Выделение признаков для порогового алгоритма состоит из следующих шагов:

1. Выравнивание пространственной ориентации тары на сегментированном изображении.
2. Кадрирование изображения. Вырезается минимальный прямоугольник, содержащий область тары.
3. Кадрирование изображения отступом от края на заданную относительную величину.
4. Разделение изображения на прямоугольные сегменты (*слайсы*) с шагом по длиннейшей стороне.
5. Поиск модуля разницы относительной доли белых пикселей (пикселей тары) первого и последнего слайсов.

Промежуточные изображения, получаемые на шагах 1-4, изображены на рис. 3.

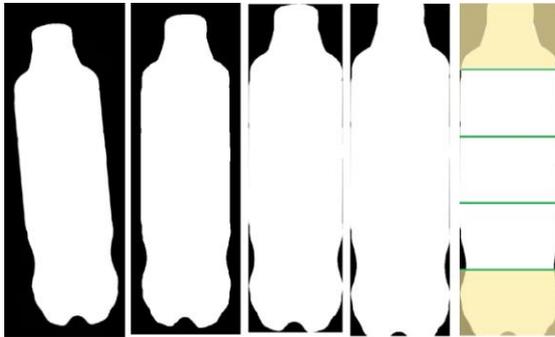


Рис.3. Примеры промежуточных изображений на этапе извлечения признаков для порогового алгоритма классификации: слева – исходное сегментированное изображение бутылки, далее слева направо – шаги 1-4 алгоритма извлечения признаков

Окончательное значение признака (шаг 5) вычисляется по следующей формуле:

$$feature = \left| \frac{n_{first}}{area_{first}} - \frac{n_{last}}{area_{last}} \right|, \quad (7)$$

где $feature$ – итоговое значение признака, n_{first} и n_{last} – количество пикселей объекта на первом и последнем слайсе, $area_{first}$ и $area_{last}$ – общее количество пикселей первого и последнего слайса, соответственно.

Изображение классифицируется как объект первого класса, если значение признака превышает заданный порог, иначе – как объект второго класса.

5. Классификация методом случайного леса

Более общий, но в то же время и более сложный подход – классификация методом *случайного леса*. В отличие от порогового классификатора, случайный лес позволяет классифицировать изображения на все три имеющихся класса. Данный метод 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.

Мерой загрязнённости узла является индекс Джини:

$$I_G(t) = \sum_{i=1}^c p(i|t)(1 - p(i|t)) = 1 - \sum_{i=1}^c p(i|t)^2, \quad (8)$$

где $p(i|t)$ – доля образцов, которые принадлежат классу i для индивидуального узла t , c – количество классов.

В качестве признаков для классификации используются величины заполненности слайсов, полученные при выделении признаков для порогового метода. В данном случае изображение допустимо разделять на прямоугольники, как по одной, так и сразу по обеим осям (рис. 4). Из набора признаков исключаются те, значимость которых ниже заданного значения. Количество слайсов и способ их извлечения, порог влияния признаков для отсеивания, количество решающих деревьев в ансамбле являются гиперпараметрами.

6. Результаты экспериментальных исследований

Для оценки эффективности реализованных методов был собран набор изображений различных тар каждого класса, сфотографированных на 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 Смарт Рэй в рамках проекта. Также вручную была размечена маска тары каждого изображения для проверки точности сегментации.

В качестве меры правильности сегментации была использована метрика IoU (intersection over union), которая задаётся следующим выражением:

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

где TP - положительные пиксели (пиксели объекта), FP - ложноположительные пиксели, FN - ложноотрицательные пиксели.

Оценка результатов правильности сегментации представлена в табл. 1.

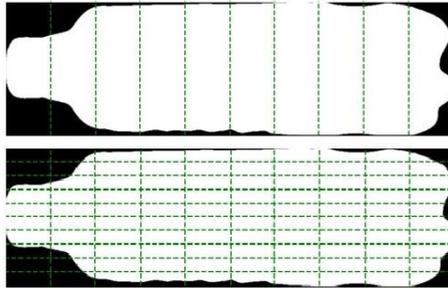


Рис.4. Примеры нарезки изображения бутылки на различное количество слайсов: сверху – изображение бутылки, нарезанное на 10 слайсов по ширине, снизу – то же изображение, нарезанное на 10 слайсов по ширине и 10 по длине

Таблица 1

Результаты сегментации

	Среднее	Минимум	Максимум	Количество
Алюминиевая банка	0.98	0.96	0.99	35
Стеклоная бутылка	0.98	0.97	0.99	32
Пластиковая бутылка	0.96	0.70	0.99	97
Все классы	0.97	0.70	0.99	164

Наилучшие результаты классификации методом случайного леса были достигнуты при следующих значениях гиперпараметров: 10 решающих деревьев в ансамбле, 10 слайсов, нарезанных с шагом по длиннейшей оси, порог отсечения значимости признаков – 0.05. Результаты классификации пороговым методом, а также результаты перекрестной проверки по десяти блокам с размером тестовой выборки, равным 0.2, для случайного леса представлены в табл. 2.

Результаты классификации

	Результат
Пороговый метод	1.00
Метод случайного леса	0.95

7. Обсуждение результатов

Как видно из результатов, рассмотренные методы позволяют достичь достаточно высокой эффективности сегментации и классификации для данной задачи. Наихудшее значение правильности сегментации было получено для изображений пластиковых бутылок. Это связано с тем, что пустая тара из этого материала обладает слишком высокой прозрачностью, что уменьшает контрастность у её границ на изображении.

Несмотря на то, что пороговый метод классификации не допустил ни одной ошибки, метод случайного леса является более полезным, поскольку позволяет разделять все три класса объектов. По результатам перекрестной проверки ансамблевый алгоритм допускает от нуля до двух ошибок. Чаще всего данный классификатор путает стеклянные бутылки и пластиковые, которые имеют вытянутую, как у стеклянных, горловину (рис. 5).



Рис.5. Сравнение форм стеклянной и пластиковой бутылок: слева – пластиковая бутылка, ошибочно классифицированная как стеклянная, в середине – стеклянная бутылка, справа – пластиковая бутылка

В практическом плане результаты исследований по созданию математического и программного обеспечения в полном объеме могут

быть реализованы в организациях промышленности, в том числе, на предприятии - промышленном партнере (ООО «Смарт Рэй») в разрабатываемых средствах экспресс-анализа и системах машинного зрения работающих в режиме реального времени, как в плане определения основных направлений развития производства, так и в конкретных конструкторских решениях и работах по модернизации выпускаемого оборудования.

Заключение

В ходе работы реализовано математическое и программное обеспечение для сегментации изображений в приёмной камере фандомата. При сегментации методом активного контура достигнута точность 0.97. Точность классификации алгоритмом случайного леса – 0.95.

Работа выполнена при финансовой поддержке Фонда содействия развитию малых форм предприятий в научно-технической сфере в рамках программы Старт-21-1 (2 очередь) / Н4. Новые приборы и интеллектуальные производственные технологии. Договор №4230ГС1/70439 от 27.10.2021 – «Разработка мультиспектрального фандомата для приёма вторичного сырья».

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. Snakes: Active Contour Models / M. Kass, A. Witkin, D. Terzopoulos; Schlumberger Palo Alto Research // Internation Journal Of Computer Vision. – Boston : Kluwer Academic Publicshers, 1987. – P. 321-331.
3. Classification and Regression Trees : book / L. Breiman [et al.]. – 1st Edition. – New York : Chapman & Hall/CRC, 1984. – 368 p.
4. Understanding Random Forests : From Theory to Practice : PhD dissertation / G. Louppe. – Belgium: University of Liège, 2014. – 223 p.
5. Generalized Intersection over Union: A Metric and A Loss for Bounding Box Regression / H. Rezatofighi [et al.]. – USA : Stanford University, 2019. – 9 p.