

Обнаружение клубней картофеля на фоне грунта

А. В. Левитин, Е. П. Мудров

Рязанский государственный радиотехнический университет имени В.Ф. Уткина, Рязань, Россия

Аннотация. В статье рассматривается задача сегментации изображений картофельных клубней на фоне грунта. Для классификации пикселей в качестве признаков применялись размытые цветовые каналы, а также два алгоритма машинного обучения: случайный лес (Random Forest) и линейный дискриминантный анализ (Linear Discriminant Analysis, LDA). Экспериментальные результаты подтвердили эффективность предложенного решения.

Ключевые слова: сегментация изображений, классификация, линейный дискриминантный анализ, случайный лес, обнаружение клубней.

Detection of potato tubers against the background of soil

A. V. Levitin, E. P. Mudrov

Ryazan State Radio Engineering University named after V.F.Utkin, Ryazan, Russia

Abstract. The article considers the problem of segmentation of images of potato tubers on the background of the soil. To classify pixels as features, blurred color channels were used, as well as two machine learning algorithms: Random Forest and Linear Discriminant Analysis (LDA). Experimental results have confirmed the effectiveness of the proposed solution.

Keywords: image segmentation, classification, linear discriminant analysis, random forest, tuber detection.

Введение

Задача обнаружения клубней картофеля на фоне грунта является ключевой при создании системы технического зрения для интеллектуального картофелеуборочного комбайна. Целью данной работы является разработка алгоритмов для СТЗ, способных решать эту задачу [1, 2].

Постановка задачи

Система технического зрения должна выполнять бинарную сегментацию получаемых изображений с разрешением 640×480 пикселей, классифицируя каждый элемент как «клубень» или «не клубень» на основе обученной модели. Сложность решения данной задачи заключается в том, что клубни часто бывают частично скрыты слоем грунта. Это требует анализа не отдельного пикселя, а его окрестности для формирования нелокальных признаков, что является критически важным для точной классификации.

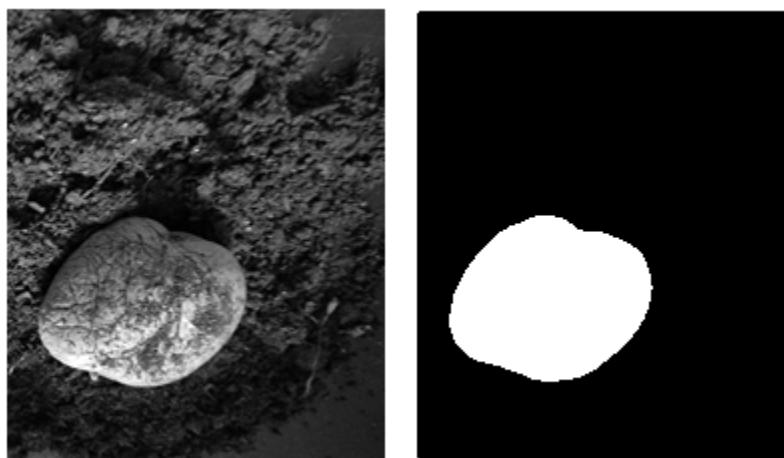


Рисунок 1. Пример исходного изображения и его маски

Теория

1. Признаки элементов изображения

Для классификации каждого элемента изображения (пикселя) использовался вектор признаков $\mathbf{z}_{ij} \in R^3$, компоненты которого вычисляются по формуле

$$z_{ij}^k = \sum_{n,m} a_{nm} q_{i+n, j+m}^k, \quad (1)$$

где q_{ij}^k – значение k -го канала ij -го элемента; a_{nm} – коэффициенты, подбираемые так, чтобы значение признака z_{ij}^k определялось значениями канала в некоторой окрестности ij -го элемента, обеспечивая тем самым нелокальность многомерного признака \mathbf{z}_{ij} .

Коэффициенты в составе соотношения (1) принимали значения

$$a_{nm} = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{n^2 + m^2}{2\sigma^2}\right), \quad (2)$$

где параметр σ , оптимальное значение которого подлежит экспериментальному определению, определяет глубину размытия изображения.

Размытие «скрывает» мелкие фрагменты грунта на поверхности клубня, сглаживая цветовые переходы и упрощая последующее разделение классов на основе обобщенной информации.

Под «каналами», используемыми для формирования вектора признаков, подразумеваются компоненты стандартных цветовых моделей. В различных экспериментах использовались: RGB или HSL. Выбор модели обусловлен необходимостью оценить, какое цветовое пространство наиболее эффективно для выделения клубней картофеля на фоне грунта, в условиях частичного перекрытия почвой.

2. Методы классификации элементов изображения

Для задачи бинарной классификации пикселей были применены и сравниены два алгоритма: линейный дискриминантный анализ [3] и случайный лес [4].

LDA предполагает, что векторы признаков \mathbf{z}_{ij} для обоих классов («клубень» и «грунт») имеют многомерное нормальное распределение с разными средними значениями, но с одинаковой ковариационной матрицей. Введем обозначения: $x_{ij} \in T_0$ – элемент изображения x_{ij} расположен в области T_0 с грунтом; $x_{ij} \in T_1$ – элемент x_{ij} расположен в области T_1 с клубнем. Тогда, в соответствии с выше сделанными предположениями LDA, справедливо соотношение

$$\ln p(x_{ij} \in T_k | \mathbf{z}_{ij}) = \mathbf{v}_k^T \mathbf{z}_{ij} + \eta_k + c, \quad (3)$$

где $p(x_{ij} \in T_k | \mathbf{z}_{ij})$ – оценка апостериорной вероятности того, что $x_{ij} \in T_k$; $\mathbf{v}_k = \Sigma^{-1} \mathbf{m}_k$;

$\eta_k = -\frac{1}{2} \mathbf{m}_k^T \Sigma^{-1} \mathbf{m}_k + \ln p(x_{ij} \in T_k)$; \mathbf{m}_k – выборочная оценка математического ожидания \mathbf{z}_{ij} при

$x_{ij} \in T_k$; Σ – выборочная оценка ковариационной матрицы \mathbf{z}_{ij} ; $p(x_{ij} \in T_k)$ – оценка априорной вероятности того, что $x_{ij} \in T_k$; c – нормирующая величина, обеспечивающая равенство

$$\sum_{k=0}^1 p(x_{ij} \in T_k | \mathbf{z}_{ij}) = 1.$$

Случайный лес — это ансамблевый метод, который строит множество решающих деревьев и объединяет их результаты для повышения общей точности и устойчивости модели. Каждое дерево обучается на своей собственной бутстрэп-выборке (случайной подвыборке данных с повторениями). При построении каждого узла дерева рассматривается не весь набор признаков, а лишь их случайное подмножество. Это предотвращает переобучение. После обучения каждое дерево в ансамбле «голосует» за отнесение объекта к тому или иному классу. Итоговый прогноз модели определяется по результату большинства голосов.

Результаты экспериментов

На рисунках 2 и 3 представлены результаты сегментации тестового RGB-изображения для Случайного леса и LDA соответственно. Вероятности принадлежности пикселей к классу "клубни" получены с помощью метода `predict_proba()`. Значения метрик составляют: $\text{precision} = 0.94$; $\text{recall} = 0.95$ при $\sigma = 14$ для Случайного леса с параметрами максимальной глубины равной 5 и числа деревьев равного 15 (данные параметры были подобраны с помощью поиска по сетке) и $\text{precision} = 0.84$; $\text{recall} = 1.0$ при $\sigma = 23$ для LDA.



Рисунок 2. Результат сегментации тестового RGB-изображения с использованием Random Forest

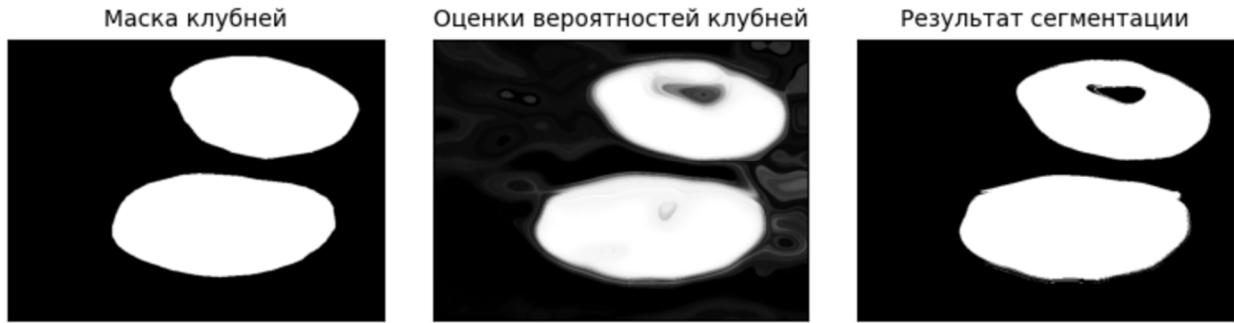


Рисунок 3. Результат сегментации тестового RGB-изображения с использованием LDA

На рисунке 4 показаны зависимости точности (precision) и полноты (recall) обнаружения пикселей клубней от параметра глубины размытия для методов LDA и RF.

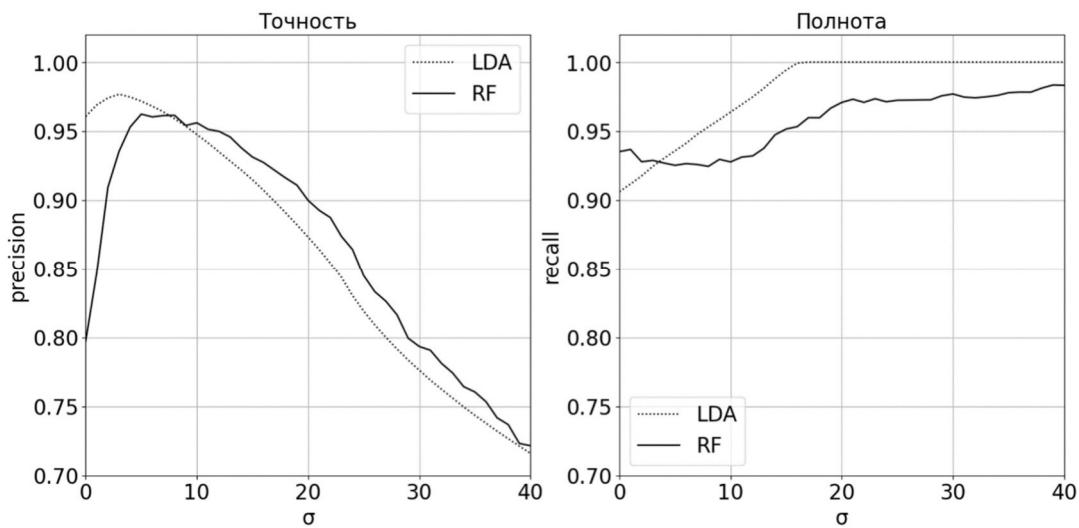


Рисунок 4. Зависимости характеристик обнаружения клубней от параметра глубины размытия

На рисунке 5 представлены кривые precision-recall и зависимости площади под этими кривыми (AUC-ROC) от параметра глубины размытия.

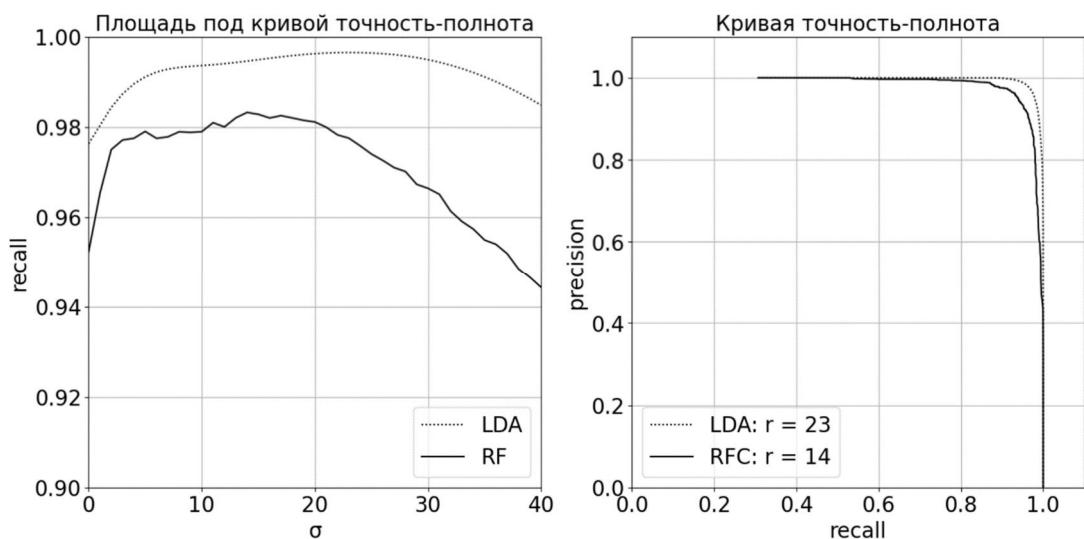


Рисунок 5. Кривые precision–recall и зависимость площади под кривой от параметра глубины размытия

На рисунках 6 и 7 представлены результаты сегментации тестового HSL-изображения для Случайного леса и LDA соответственно. Значения метрик составляют: precision = 0.93; recall = 0.86 при $\sigma = 0$ для Случайного леса с параметрами максимальной глубины равной 9 и числа деревьев равного 20 и precision = 0.92; recall = 0.85 при $\sigma = 3$ для LDA.



Рисунок 6. Результат сегментации тестового HSL-изображения с использованием Random Forest



Рисунок 7. Результат сегментации тестового HSL-изображения с использованием LDA

На рисунке 8 показаны зависимости точности (precision) и полноты (recall) обнаружения пикселей клубней от параметра глубины размытия для методов LDA и RF.

На рисунке 9 представлены кривые precision–recall и зависимости площади под этими кривыми (AUC-ROC) от параметра глубины размытия.

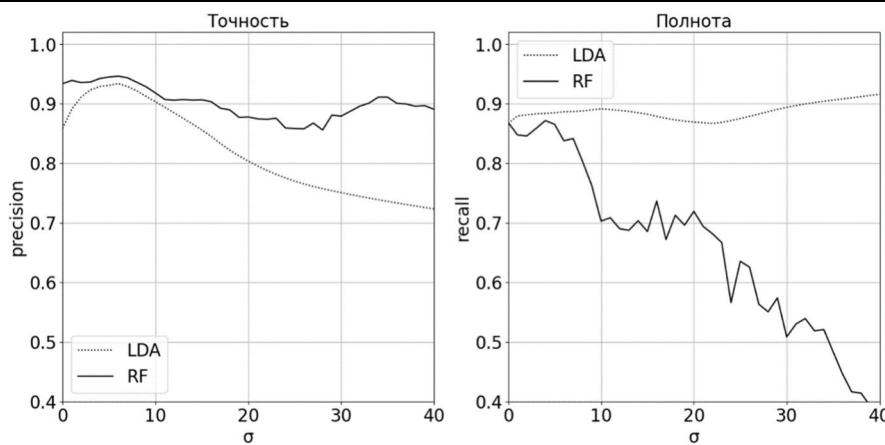


Рисунок 8. Зависимости характеристик обнаружения клубней от параметра глубины размытия

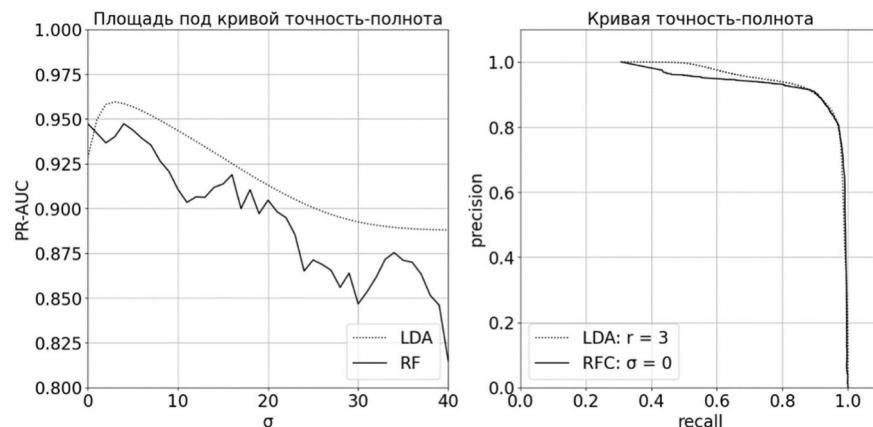


Рисунок 9. Кривые precision–recall и зависимость площади под кривой от параметра глубины размытия

Выводы

Результаты исследования показали, что предложенные подходы демонстрируют разную эффективность в зависимости от используемого цветового пространства.

В пространстве RGB оба метода показали хорошие результаты, при этом метод LDA продемонстрировал значительно более высокое качество классификации. Также преимуществом LDA стала значительно меньшая вычислительная сложность по сравнению с Random Forest.

В отличие от RGB, в пространстве HSL оба метода показали неудовлетворительные результаты классификации, что свидетельствует о меньшей пригодности этого цветового пространства для решения данной конкретной задачи сегментации.

Оптимальное значение параметра глубины размытия σ может быть найдено путем максимизации площади под кривой точность–полнота для обучающего изображения.

Список литературы

1. Левитин А. В., Мудров Е. П. Обнаружение клубней картофеля на фоне грунта // Новые информационные технологии в научных исследованиях: материалы XXIX Всероссийской научно-технической конференции студентов, молодых ученых и специалистов, Рязань, 27–29 ноября 2024 года / Рязанский государственный радиотехнический университет им. В.Ф. Уткина. 2024. Т. 2. С. 118.
2. Левитин А. В., Мудров Е. П. Обнаружение клубней картофеля на фоне грунта // Современные технологии в науке и образовании – СТНО-2025: сборник трудов VIII международного научно-технического форума, Рязань, 4–6 марта 2025 года / Рязанский государственный радиотехнический университет им. В.Ф. Уткина, 2025 Т. 5. С. 142–147.
3. Duda R. O., Hart P. E., Stork D. G. Pattern Classification. 2nd ed. New York: John Wiley & Sons, 2001. Section 2.6.2.
4. Breiman L. Random Forests // Machine Learning. 2001. Vol. 45, no. 1. P. 5–32.