

МАТЕМАТИКА . ИНФОРМАТИКА

DOI: 10.24412/2077-8481-2026-1-97-104

УДК 004.93

С. В. КУРАШОВ

И. И. МАКОВЕЦКИЙ, канд. физ.-мат. наук

Белорусско-Российский университет (Могилев, Беларусь)

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ИНФОРМАТИВНОСТИ ВИЗУАЛЬНЫХ ДЕСКРИПТОРОВ ДЛЯ КЛАССИФИКАЦИИ СОСТОЯНИЯ ДВЕРНЫХ ПРОЕМОВ В СИСТЕМАХ АССИСТИВНОГО ЗРЕНИЯ

Аннотация

Рассматривается задача автоматического определения состояния двери (открыта/закрыта) на изображениях со сложным фоном для систем навигации слабовидящих. Предложен гибридный подход, сочетающий нейросетевую локализацию объекта (YOLO) и классификацию его состояния методами машинного обучения. Сформирован вектор из семи визуальных признаков, описывающих текстурные, геометрические и яркостные характеристики изображения. На основе модели Random Forest проведен количественный анализ информативности признаков (Feature Importance). Экспериментально установлено, что наиболее значимым дескриптором является дисперсия яркости (Importance $\approx 0,19$), характеризующая однородность поверхности. Достигнута точность классификации (Accuracy) более 90 %, что подтверждает эффективность предложенного набора признаков.

Ключевые слова:

компьютерное зрение, классификация изображений, Random Forest, дескрипторы признаков, LBP, анализ текстур, ассистивные технологии.

Для цитирования:

Курашов, С. В. Сравнительный анализ информативности визуальных дескрипторов для классификации состояния дверных проемов в системах ассистивного зрения / С. В. Курашов, И. И. Маковецкий // Вестник Белорусско-Российского университета. – 2026. – № 1 (90). – С. 97–104.

Введение

Одна из ключевых задач в разработке ассистивных технологий для людей с нарушениями зрения – навигация внутри помещений. Критически важным аспектом безопасности является своевременное обнаружение дверей и определение их состояния (открыта или закрыта). Современные детекторы объектов на базе глубокого обучения (YOLO, SSD) успешно решают задачу локализации, выделяя объект «дверь»

ограничивающей рамкой [1]. Однако существует проблема «семантического разрыва»: стандартный детектор часто не способен надежно классифицировать состояние двери, т. к. визуально закрытая дверь и открытый проем могут иметь схожие очертания, особенно в условиях низкой освещенности или сложного фона.

Существующие решения часто опираются либо на дорогостоящие дальнометры (LiDAR, ультразвуковые сенсоры), которые эффективны лишь на ма-

лых дистанциях, либо на «тяжелые» нейросетевые модели, требующие мощных вычислительных ресурсов. Для носимых ассистивных устройств критически важен баланс между точностью и энергоэффективностью. Прямое использование сверточных нейронных сетей (CNN) для классификации состояния требует сбора огромных размеченных датасетов для каждого возможного ракурса и типа двери. В то же время гипотеза о том, что состояние объекта можно определить по набору интерпретируемых геометрических и текстурных признаков, позволяет существенно упростить вычислительный конвейер и повысить интерпретируемость принимаемых решений.

Цель работы – исследовать эффективность различных групп визуальных дескрипторов (признаков) и определить минимально необходимый набор, обеспечивающий высокую точность классификации состояния двери при низких вычислительных затратах.

Методика исследования

Для решения поставленной задачи разработан гибридный алгоритм, объединяющий методы глубокого обучения (Deep Learning) для локализации объектов и методы классического компьютерного зрения (Computer Vision) для анализа их состояния. Общая архитектура системы представляет собой двухэтапный конвейер. Подробнее об этапах конвейера рассказано в [2].

На первом этапе входной видеопоток подвергается предварительной обработке, после чего нейросетевой детектор выделяет область интереса (Region of Interest, ROI), содержащую объект класса «дверь». На втором этапе выполняются извлечение вектора признаков из ROI и его классификация. Рассмотрим математическое описание предлагаемых дескрипторов.

Поскольку цветовая информация (оттенок двери) не является инвариант-

ным признаком состояния (дверь может быть любого цвета), для снижения вычислительной сложности и размерности данных анализ выполняется в пространстве яркостей. Преобразование исходного RGB-изображения в полутоновое (Grayscale) осуществляется по следующей стандартной формуле взвешенной суммы каналов:

$$Y(x, y) = 0,299 \cdot R(x, y) + \\ + 0,587 \cdot G(x, y) + 0,114 \cdot B(x, y),$$

где $Y(x, y)$ – интенсивность пикселя с координатами (x, y) ; R, G, B – значения цветовых каналов.

Все дальнейшие вычисления признаков производятся над матрицей Y .

Базовая гипотеза исследования заключается в том, что закрытая дверь представляет собой относительно однородную поверхность (полотно), тогда как в открытом дверном проеме наблюдается сложная сцена (интерьер комнаты, мебель, перепады освещения). Для количественной оценки этой неоднородности используются статистические моменты распределения яркости.

Дисперсия яркости (Simple Variance). Данный признак характеризует глобальный разброс значений интенсивности пикселей относительно среднего. Высокое значение дисперсии свидетельствует о наличии контрастных перепадов, характерных для открытого проема.

Дисперсию σ^2 можно найти по формуле

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (I_i - \mu)^2,$$

где N – общее количество пикселей в выделенной области интереса (ROI); I_i – значение яркости i -го пикселя; μ – математическое ожидание (средняя

яркость) ROI, определяемое как

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i .$$

Разностная региональная дисперсия (Regional Variance Difference).

Для детектирования ситуаций, когда дверь приоткрыта лишь частично (например, видны щель или дверной косяк), глобальной дисперсии может быть недостаточно. Предложен метод зонального анализа: ROI разбивается на три вертикальные полосы равной ширины. Для каждой полосы $k \in \{1, 2, 3\}$ вычисляется локальная дисперсия σ_k^2 . Итоговый дескриптор F_{var_diff} рассчитывается как размах дисперсий:

$$F_{var_diff} = \max(\sigma_1^2, \sigma_2^2, \sigma_3^2) - \min(\sigma_1^2, \sigma_2^2, \sigma_3^2) .$$

Большое значение данного признака указывает на неравномерность визуальной сложности по ширине объекта, что свойственно полуоткрытым дверям.

Для анализа микроструктуры изображения используется оператор локальных бинарных шаблонов (Local Binary Patterns, LBP), обладающий устойчивостью к монотонным изменениям освещения. Код LBP для пикселя (x_c, y_c) вычисляется путем порогового сравнения яркости центрального пикселя с его соседями по окружности радиусом R :

$$LBP_{P,R}(x_c, y_c) = \sum_{p=0}^{P-1} s(g_s - g_p) 2^p ;$$

$$s(x) = \begin{cases} 1, & x \geq 0; \\ 0, & x < 0, \end{cases}$$

где g_s – яркость центрального пикселя; g_p – яркость p -го соседа; P – количе-

ство точек-соседей.

Использовались два производных признака:

1) дисперсия LBP ($R = 1, P = 8$). Описывает мелкомасштабную текстуру поверхности;

2) энтропия LBP ($R = 3, P = 8$). Описывает информационную сложность текстуры на большем масштабе.

Энтропия Шеннона H вычисляется по гистограмме распределения LBP-кодов как

$$H = - \sum_j p_j \log_2 p_j ,$$

где p_j – вероятность появления j -го бинарного шаблона в ROI.

Высокая энтропия коррелирует с хаотичностью текстуры, наблюдаемой в открытом проеме.

Геометрия закрытой двери характеризуется наличием четких контуров дверной коробки и полотна. Для формализации этих свойств применяются методы детектирования границ.

Плотность границ (Canny Edges).

Используется детектор границ Кэнни, включающий этапы гауссовского размытия, вычисления градиента, подавления немаксимумов и гистерезисной пороговой фильтрации. Дескриптор F_{edges} представляет собой нормированное количество пикселей, отнесенных к границам:

$$F_{edges} = \frac{1}{W \cdot H} \sum_{x,y} B(x, y) ,$$

где $B(x, y) = 1$, если пиксель является границей, и 0 – в противном случае; W, H – ширина и высота ROI.

Метрика вертикальности (Verticality). Для различения вертикальных линий (косяки двери) и хаотичных линий фона разработан специальный алгоритм на основе анализа гистограммы ориентированных градиентов (Histogram of

Oriented Gradients). Вычисляются магнитуда $M(x, y)$ и ориентация $\theta(x, y)$ градиента в каждой точке:

$$M(x, y) = \sqrt{G_x^2 + G_y^2};$$

$$\theta(x, y) = \arctg\left(\frac{G_y}{G_x}\right).$$

Строится взвешенная гистограмма ориентаций. Вычисляется суммарная энергия вертикальных градиентов E_V (углы $\theta \in [90^\circ \pm \delta]$) и горизонтальных градиентов E_H ($\theta \in [0^\circ \pm \delta] \cup [180^\circ \pm \delta]$). Итоговая метрика вертикальности определяется как

$$F_{vert} = \min\left(\frac{E_V}{E_H + \varepsilon}, 100\right),$$

где ε – малая константа для исключения деления на ноль.

Преобладание вертикальной энергии характерно для закрытого состояния двери.

Необходимость введения специализированной метрики вертикальности обусловлена особенностями конструкции дверных проемов. Стандартная закрытая дверь характеризуется наличием ярко выраженных вертикальных границ (дверная коробка, полотно, филенки), которые доминируют в гистограмме градиентов. В случае открытого проема, даже при наличии вертикальных косяков, внутреннее пространство (фон) вносит хаотичные градиенты различных направлений, что «размывает» пики в гистограмме ориентаций. Использование нормализованной гистограммы позволяет сделать этот признак инвариантным к масштабу изображения (расстоянию до двери).

Сформированный вектор признаков $X = \{\sigma^2, F_{var_diff}, \mu, \sigma_{LBP}^2, H_{LBP}, F_{edges}, F_{vert}\}$

нормализуется и подается на вход классификатора. В качестве алгоритма классификации выбран метод случайного леса (Random Forest), представляющий собой ансамбль из K решающих деревьев. Финальное решение о классе $C \in \{\text{"Открыто"}, \text{"Закрыто"}\}$ принимается путем мажоритарного голосования деревьев:

$$C = \text{mode}\{C_1(X), C_2(X), \dots, C_K(X)\}.$$

Выбор Random Forest обусловлен его высокой обобщающей способностью, устойчивостью к выбросам и возможностью оценки важности признаков (Feature Importance), что необходимо для анализа результатов исследования.

Экспериментальная часть

Для обучения и тестирования разработанной модели был сформирован репрезентативный набор данных, объединяющий изображения из общедоступного датасета Open Images V7 и специализированного набора, собранного в реальных условиях эксплуатации (общезития, учебные корпуса, офисные помещения).

Общий объем выборки составил 6110 изображений. Особенностью набора данных является естественный дисбаланс классов, характерный для задачи навигации: количество изображений с закрытыми дверями (класс «Closed») превышает количество открытых проемов (класс «Open») в соотношении 3,5:1.

Для обеспечения корректности оценки обобщающей способности модели данные были разделены на обучающую (Train) и тестовую (Test) выборки в соотношении 75 % к 25 % с использованием стратифицированного разбиения (Stratified Split). Это позволило сохранить исходное распределение классов в обеих подвыборках и избежать статистических искажений при обучении.

Перед подачей на вход классификатора векторы признаков подвергались стандартизации (*Z-score normalization*) для приведения к нулевому среднему и единичной дисперсии.

В качестве классификатора использовался алгоритм *Random Forest* со следующими гиперпараметрами, подобранными эмпирически для достижения баланса между точностью и скоростью инференса.

1. Количество деревьев решений (*n_estimators*): 100.

2. Критерий расщепления: Джини (*Gini impurity*).

3. Балансировка весов классов (*class_weight*): 'balanced' (автоматическая корректировка весов обратно пропорциональна частоте классов во входных данных).

Использование параметра балансировки позволило компенсировать дисбаланс выборки и повысить чувствительность модели к минорному, но критически важному классу «Открыто».

В задачах ассистивных технологий

стоимость ошибок первого и второго рода неравнозначна. Ложное срабатывание (*False Positive*), когда закрытая дверь классифицируется как открытая, может привести к травмированию пользователя. Поэтому, помимо общей точности (*Accuracy*), ключевыми метриками являлись:

1) *Precision* (Точность) для класса «Открыто» – доля действительно открытых дверей среди всех объектов, классифицированных как открытые;

2) *Recall* (Полнота) – доля найденных открытых дверей от их общего числа;

3) *F1-score* – гармоническое среднее точности и полноты;

4) *ROC AUC* – площадь под кривой ошибок, оценивающая качество ранжирования объектов.

Результаты и обсуждение

Результаты тестирования обученной модели на отложенной выборке представлены в табл. 1.

Табл. 1. Показатели эффективности классификации состояния двери

Метрика	Значение
Accuracy	0,904
Precision	0,918
Recall	0,629
F1-score	0,747
ROC AUC	0,932

Анализ полученных данных показывает, что предложенный подход обеспечивает высокую общую точность (90,4 %). Критически важным результатом является высокое значение метрики *Precision* (0,918). Это означает, что в 91,8 % случаев, когда система сообщает пользователю «Дверь открыта», путь действительно свободен. Такой результат минимизирует риски столкновения с препятствием, что является

приоритетом для систем безопасности незрячих.

Значение *Recall* (0,629) указывает на то, что модель проявляет определенный консерватизм, пропуская часть открытых дверей (например, в условиях плохой видимости или при малом угле открытия). В контексте задачи навигации это является приемлемым компромиссом: лучше не обнаружить проход, чем направить пользователя в закры-

тую дверь.

Помимо точностных характеристик, для ассистивных систем критически важным параметром является латентность (задержка обработки). Была проведена оценка производительности разработанного программного комплекса на тестовом стенде CPU Intel Core i5 без использования дискретного GPU-ускорителя. Замерялось полное время прохождения кадра по конвейеру.

Средняя скорость обработки составила 27...29 кадров в секунду (FPS), что соответствует времени обработки одного кадра порядка 35...37 мс. Из этого времени детекция занимает около 60 % времени, расчет текстурных признаков и фильтрация Кэнни – около 35 %, классификация – менее 5 %.

Такая производительность позволяет использовать систему в режиме реального времени на мобильных устройствах, обеспечивая мгновенную обратную связь для пользователя при движении со средней скоростью пешехода. Отсутствие необходимости в тяжелых вычислениях на этапе классификации подтверждает эффективность гибридного подхода для встраиваемых систем.

Для верификации гипотез, выдвинутых в разделе «Методика исследования», был проведен анализ важности признаков (Feature Importance) на основе среднего уменьшения неопределенности (Mean Decrease in Impurity). Результаты ранжирования дескрипторов представлены на рис. 1.

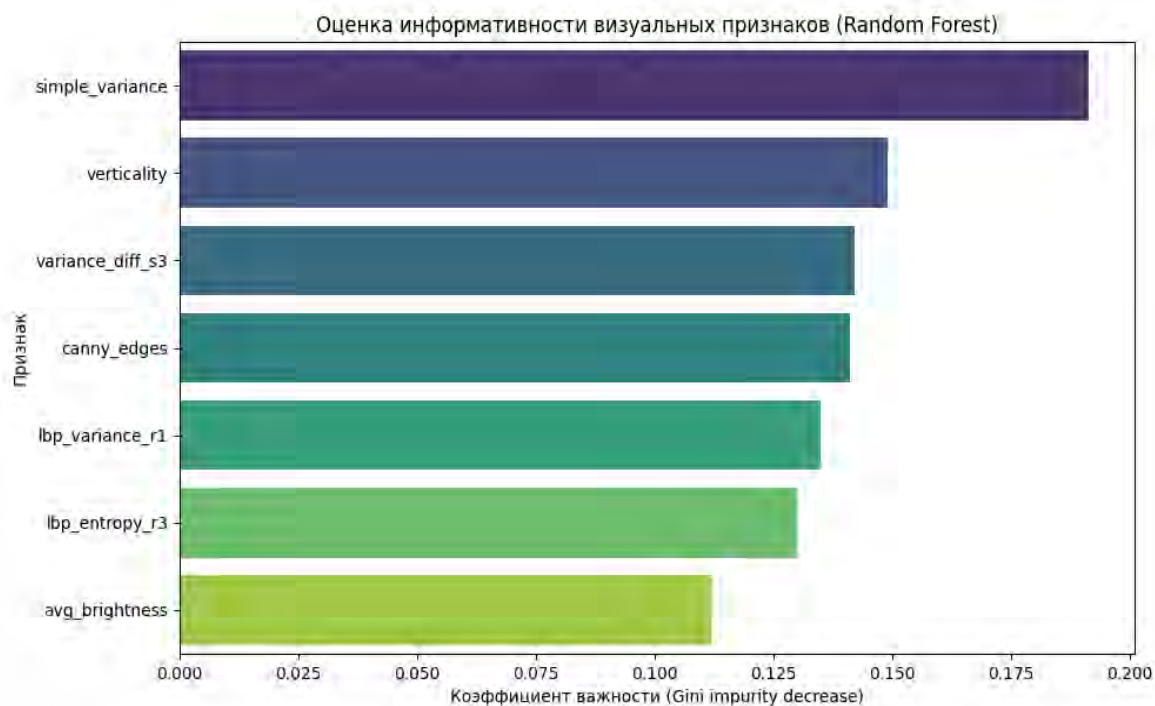


Рис. 1. Диаграмма оценки информативности визуальных признаков (Random Forest)

Количественный анализ важности признаков позволяет сделать следующие выводы о физике процесса классификации.

1. Признак *simple_variance* показал

наивысшую информативность (коэффициент важности $\approx 0,19$). Это подтверждает предположение о том, что глобальная неоднородность сцены является самым надежным индикатором откры-

того проема. Закрытая дверь, как правило, представляет собой плоский объект с равномерным отражением света, тогда как открытый проем содержит глубину сцены и множественные перепады яркости.

2. Вторым по важности признаком ($\approx 0,15$) стала метрика *verticality*, основанная на гистограмме градиентов. Это свидетельствует о том, что сохранение строгой вертикальной структуры линий является сильным признаком закрытого состояния (дверное полотно, косяки). В открытом проеме вертикальные линии часто «размываются» фоном.

3. Признаки, основанные на детекторе границ (*canny_edges*) и зональной дисперсии (*variance_diff_s3*), показали высокую значимость ($\sim 0,14$), эффективно дополняя глобальные дескрипторы. В частности, *variance_diff_s3* позволяет успешно детектировать полуоткрытые двери, создающие асимметрию в изображении.

4. Признак средней яркости (*avg_brightness*) оказался наименее значимым ($\approx 0,11$). Это позитивный результат, указывающий на то, что модель опирается на структурные и текстурные особенности объекта, а не на уровень

освещенности, что обеспечивает робастность системы при работе в разное время суток.

Заключение

Предложен и исследован метод классификации состояния дверных проемов, основанный на комбинации нейросетевой локализации и анализа вектора визуальных дескрипторов.

Экспериментально доказано, что использование ограниченного набора из семи интерпретируемых признаков (текстурных, геометрических и яркостных) в сочетании с алгоритмом Random Forest позволяет достичь точности классификации более 90 %. Установлено, что наиболее дискриминативными признаками являются дисперсия яркости и метрика вертикальности, вносящие более 34 % вклада в принятие решения.

Полученные результаты внедрены в программный комплекс DoorSight, что позволило реализовать функционал определения состояния двери в режиме реального времени с высокой надежностью и низкими требованиями к вычислительным ресурсам [3].

СПИСОК ИСПОЛЬЗОВАННОЙ ЛИТЕРАТУРЫ

1. **Redmon, J.** You only look once: Unified, real-time object detection / J. Redmon // Proceedings of the IEEE conference on computer vision and pattern recognition. – 2016. – P. 779–788.
2. **Курашов, С. В.** Использование комбинации глубокого обучения и классических методов компьютерного зрения для определения состояния двери / С. В. Курашов // 61 студен. науч.-техн. конф., Могилев, 15 мая 2025 г. – Могилев : Бел.-Рос. ун-т, 2025. – С. 83.
3. **Курашов, С. В.** Архитектура и реализация полностекового веб-приложения в сфере ассистивных технологий / С. В. Курашов // Новые материалы, оборудование и технологии в промышленности : материалы Междунар. науч.-техн. конф. молодых ученых, Могилев, 23–24 окт. 2025 г. – Могилев : Бел.-Рос. ун-т, 2025. – С. 204–205.

Статья сдана в редакцию 19 января 2026 года

Контакты:

kurashov2015@yandex.by (Курашов Савелий Валерьевич);
fiu@exes.bgu.by (Маковецкий Илья Иванович).

S. V. KURASHOV, I. I. MAKOVETSKY

COMPARATIVE ANALYSIS OF VISUAL DESCRIPTOR INFORMATIVENESS FOR DOORWAY STATE CLASSIFICATION IN ASSISTIVE VISION SYSTEMS

Abstract

The article addresses the problem of automatic door state detection (open/closed) in images with complex backgrounds for navigation systems designed for the visually impaired. A hybrid approach is proposed combining neural network-based object localization (YOLO) and state classification using machine learning methods. A vector of seven visual features was formed, which describe textural, geometric, and brightness characteristics of the image. Based on the Random Forest model, a quantitative analysis of feature informativeness (Feature Importance) was conducted. Experimental results established that the most significant descriptor is brightness variance (Importance $\approx 0,19$), characterizing surface homogeneity. A classification accuracy (Accuracy) of over 90 % was achieved, confirming the effectiveness of the proposed feature set.

Keywords:

computer vision, image classification, Random Forest, feature descriptors, LBP, texture analysis, assistive technologies.

For citation:

Kurashov, S. V. Comparative analysis of visual descriptor informativeness for doorway state classification in assistive vision systems / S. V. Kurashov, I. I. Makovetsky // Belarusian-Russian University Bulletin. – 2026. – № 1 (90). – P. 97–104.