

УДК 004.93

DOI: [10.26102/2310-6018/2026.52.1.002](https://doi.org/10.26102/2310-6018/2026.52.1.002)

Экспериментальные исследования дескриптора формы бинарной фигуры на основе профиля вращения

О.С. Середин, Н.А. Ломов, Д.В. Ляхов, Н.С. Митюгов✉, О.А. Кушнир,
А.В. Копылов

Тулский государственный университет, Тула, Российская Федерация

Резюме. В работе представлены результаты экспериментальных исследований дескриптора формы, основанного на профиле вращения, для задач классификации листьев растений. Дескриптор представляет собой последовательность значений, полученных путем поворота фигуры относительно самой себя с заданным угловым шагом в интервале от 0 до 180 градусов. Для каждого угла поворота вычисляется мера Жаккара, отражающая сходство между исходной и повернутой фигурами. Предложенный дескриптор обладает инвариантностью к преобразованиям подобия, что обеспечивает его эффективность при анализе объектов с разной формой. Эксперименты проводились на четырех задачах классификации с применением трех типов классификаторов: метода опорных векторов (SVM), градиентного бустинга (XGBoost) и нейронной сети (NN Simple). Эффективность дескриптора сравнивалась с традиционными подходами, включая моменты Зернике, геометрические моменты и моменты Ху. Кроме того, для сравнения использовались распознавания на растровых изображениях с помощью сверточных нейронных сетей (ResNet50, VGG16, CNN Simple). Результаты показали высокую точность и стабильность предложенного дескриптора формы в различных контекстах классификации, а также подтвердили его большой потенциал для задач анализа формы в компьютерном зрении.

Ключевые слова: компьютерное зрение, бинарное растровое изображение, анализ формы, мера Жаккара, профиль вращения.

Благодарности: Работа выполнена при финансовой поддержке Министерства науки и высшего образования РФ в рамках государственного задания FEWG-2024-0001.

Для цитирования: Середин О.С., Ломов Н.А., Ляхов Д.В., Митюгов Н.С., Кушнир О.А., Копылов А.В. Экспериментальные исследования дескриптора формы бинарной фигуры на основе профиля вращения. *Моделирование, оптимизация и информационные технологии*. 2026;14(1). URL: <https://moitvvt.ru/ru/journal/pdf?id=2043> DOI: 10.26102/2310-6018/2026.52.1.002

Experimental study of the rotation profile based binary shape descriptor

O.S. Seredin, N.A. Lomov, D.V. Liakhov, N.S. Mityugov✉, O.A. Kushnir, A.V. Kopylov

Tula State University, Tula, the Russian Federation

Abstract. This paper presents the results of an experimental study of a shape descriptor based on a Rotation Profile for tasks of leaf classification. The descriptor is a sequence of values obtained by rotating the shape around itself with a fixed angular step within the range of 0 to 180 degrees. For each rotation angle, the Jaccard measure, reflecting the similarity between the original and rotated shapes, is calculated. The proposed descriptor is invariant to similarity transformations, ensuring its effectiveness in analyzing objects with varying shapes. Experiments were conducted on four classification tasks using three types of classifiers: Support Vector Machine (SVM), Gradient Boosting (XGBoost), and a simple neural network (NN Simple). The descriptor's performance was compared with traditional approaches, including Zernike moments, geometric moments, and Hu moments. Additionally, recognition was performed directly on raster images using convolutional neural networks (ResNet50, VGG16, CNN Simple). The results demonstrated high accuracy and stability of the proposed shape descriptor across

different classification contexts and confirmed its strong potential for shape analysis tasks in computer vision.

Keywords: computer vision, binary raster image, shape analysis, Jaccard measure, rotation profile.

Acknowledgements: The work was carried out with the financial support of the Ministry of Science and Higher Education of the Russian Federation within the framework of the state assignment FEWG-2024-0001.

For citation: Seredin O.S., Lomov N.A., Liakhov D.V., Mityugov N.S., Kushnir O.A., Kopylov A.V. Experimental study of the rotation profile based binary shape descriptor. *Modeling, Optimization and Information Technology*. 2026;14(1). (In Russ.). URL: <https://moitvvt.ru/ru/journal/pdf?id=2043> DOI: 10.26102/2310-6018/2026.52.1.002

Введение

В компьютерном зрении и обработке изображений анализ формы является одним из ключевых аспектов при распознавании объектов. Форма объекта несет важную визуальную информацию, отражающую его структуру и свойства. Для эффективного и точного анализа формы применяются специальные методы ее описания – дескрипторы формы. Они должны не только точно извлекать и компактно представлять форму, но и сохранять инвариантность к различным преобразованиям, таким как масштабирование, смещение и вращение, что значительно усложняет их разработку.

Современные дескрипторы формы принято разделять на две группы, а именно дескрипторы на основе контуров и на основе областей.

Дескрипторы формы на основе контуров используют только информацию о границах и не отражают внутреннее содержание изображения объекта. Кроме того, они не могут работать с формами, структура которых сильно затрудняет получение точной границы, что ограничивает их практическое применение. Также контурные дескрипторы сильно подвержены влиянию шумов. К этой группе можно отнести Фурье-дескрипторы [1], эллиптические Фурье-дескрипторы [2] и многомасштабные дескрипторы на основе кривизны [3]. Фурье-дескрипторы обладают такими характеристиками, как простое с вычислительной точки зрения получение и простая нормализация, что сделало их очень популярными в широком спектре приложений. Однако Фурье-дескрипторы обладают следующими недостатками: во-первых, они используют информацию о границе формы целиком и не могут извлекать локальные характеристики, а во-вторых, они не всегда обеспечивают инвариантность к различным преобразованиям, таким как повороты, масштабирование и сдвиг. В некоторых случаях это может затруднить распознавание объектов при изменении их положения или размера.

Дескрипторы формы на основе областей используют моменты для представления информации об области в целом. В эту группу входят геометрические моменты, моменты Зернике [4], псевдо-моменты Зернике, моменты Лежандра, моменты Ху [5], а также новый дескриптор формы, основанный на профиле вращения, предложенный в [6].

Помимо этого, в работе [7] описаны классические ортогональные моменты, которые делятся на несколько групп, исходя из использованных ортогональных функций: полиномы Якоби, гармонические функции и собственные функции. В первую группу входят ранее названные моменты Зернике и псевдо-моменты Зернике, а также ортогональные моменты Фурье-Меллина, моменты Чебышева-Фурье и другие. Ортогональные моменты с использованием полиномов Якоби часто используются в оптической физике, цифровой обработке изображений и распознавании образов. Однако определение их базисных функций приводит к высокой вычислительной сложности из-за использования факториалов и гамма-функций. Во вторую группу входят радиальные гармонические моменты Фурье и моменты экспоненты-Фурье. По сравнению с

полиномами Якоби, ортогональные моменты, использующие только гармонические функции, не требуют сложных вычислений, а также имеют меньшую временную сложность и лучшую численную стабильность. Ортогональных моментов, основанных на собственных функциях на данный момент относительно мало, одним из представителей этой группы являются моменты Бесселя-Фурье [8]. Используемая для их формирования радиальная базисная функция опирается на бесконечные ряды и факториалы. Поэтому по сравнению с полиномами Якоби и гармоническими функциями ортогональные моменты, основанные на собственных функциях, теоретически имеют значительно более высокую сложность.

Геометрические моменты [9] – это способ описания формы, основанный на моментных инвариантах для представления формы и измерения их схожести. Моментные инварианты являются результатом вычисления моментов формы и остаются инвариантными относительно двумерных геометрических преобразований формы. Преимущество использования методики геометрических моментных дескрипторов заключается в том, что это очень компактное представление формы, и вычисления несложны, однако сложно получить моментные инварианты более высокого порядка.

Моменты Зернике позволяют создавать независимые моментные инварианты произвольно высокого порядка. Для получения комплексных моментов Зернике используются полиномы Зернике. Теория моментов Зернике аналогична теории преобразования Фурье, которая заключается в разложении сигнала в ряд ортогональных базисных функций. Однако вычисление дескрипторов моментов Зернике не требует знания информации о границе, что делает его подходящим для представления более сложных форм. Вычисление моментов Зернике имеет множество проблем, таких как преобразование пространства координат изображения в область, где определен ортогональный полином, непрерывные интегралы должны быть приближены дискретными суммами. Это приближение не только приводит к численным ошибкам в вычисленных моментах, но также существенно влияет на аналитические свойства, такие как вращательная инвариантность и ортогональность.

В контексте описания форм также можно рассмотреть подходы, основанные на алгебраической инвариантности. В работе [10] представлено минимальное множество однородных порождающих элементов алгебры инвариантов бинарной формы степени 7. Используя метод полуинвариантов и семитрансвектантов, автор вычисляет явную систему из 30 фундаментальных инвариантов, что подтверждает результаты классической теории.

В работе [11] были предложены метод кодирования скелета с помощью цепочки примитивов, включающих длину ребра и угол между данным ребром и соседним, и функция различия пары скелетов, основанная на сравнении таких цепочек. Однако данный способ кодирования не учитывает радиальную функцию скелета, которая является важным дискриминантным признаком в решении многих задач. Эту проблему устранили посредством включения радиальной функции скелета через коэффициенты полинома Лежандра.

Важность надежного и точного дескриптора формы выходит за рамки простого описания различных форм. Исследование, проведенное в работе [12], показало высокую эффективность использования простых дескрипторов формы в функциях потерь для сетей глубокой сегментации. Это открытие потенциально может определить направление будущих исследований.

В работе [6] был предложен новый дескриптор формы, основанный на профиле вращения, представляющий собой последовательность чисел, получаемых при повороте фигуры относительно самой себя с фиксированным угловым шагом в диапазоне от 0 до 180 градусов. Для каждого угла в этой последовательности вычисляется мера Жаккара,

измеряющая сходство между исходной и повернутой фигурами. Данный дескриптор обладает инвариантностью к сдвигам, вращениям, отражениям и масштабированию, что обеспечивает его высокую эффективность при анализе объектов с различными формами. Он отличается понятной концепцией, очевидной интерпретацией и простой алгоритмической реализацией.

В настоящей работе проведены экспериментальные исследования этого нового дескриптора формы с точки зрения точности классификации объектов на четырех различных задачах. Мы сравнили дескриптор формы, основанный на профиле вращения, с представлениями объектов на изображениях с использованием таких дескрипторов, как моменты Зернике, геометрические моменты и моменты Ху. Для каждого из рассмотренных дескрипторов были проведены распознавания с использованием трех типов классификаторов: метода опорных векторов (SVM), градиентного бустинга (XGBoost) и глубоких нейронных сетей (NN Simple). Дополнительно для оценки описательных свойств формы были выполнены эксперименты с бинарными и цветными растровыми изображениями, где в роли классификаторов выступали сверточные нейронные сети (ResNet50, VGG16, CNN Simple).

Результаты экспериментальных исследований позволяют сделать вывод о высокой эффективности и устойчивости дескриптора формы на основе профиля вращения в различных контекстах классификации, а также демонстрируют высокий потенциал данного дескриптора для задач анализа формы в области компьютерного зрения.

Материалы и методы

Профиль вращения для бинарного растрового изображения. Профиль вращения представляет собой последовательность фиксированной длины из некоторого набора чисел, которые были получены путем поворота фигуры относительно самой себя на угол 180 градусов с некоторым шагом. В данной работе мы использовали шаг в один угловой градус. Для каждого угла поворота от 1 до 180 градусов этой последовательности рассчитывается мера Жаккара [13], которая отражает степень сходства между исходной и повернутой фигурами:

$$J(c, I, I_c^\varphi) = \frac{|I \cap I_c^\varphi|}{|I \cup I_c^\varphi|} = \frac{|I \cap I_c^\varphi|}{2|I| - |I \cap I_c^\varphi|}, \quad (1)$$

где I – множество пикселей изображения, принадлежащих фигуре (далее – фигура), I_c^φ – фигура, полученная поворотом фигуры I на угол φ относительно некоторой точки c с координатами (c_x, c_y) .

Поскольку мера Жаккара для изображений, полученных поворотами вокруг одной и той же точки на углы φ_1 и φ_2 , зависит лишь от угла поворота между этими изображениями, $|\varphi_2 - \varphi_1|$ формула (1) может быть записана в следующем виде:

$$J(c, \varphi) = J(c, I_c^{\varphi_1}, I_c^{\varphi_2}), \varphi = |\varphi_2 - \varphi_1|. \quad (2)$$

Ряд значений $J(c, \varphi)$ при повороте фигуры вокруг точки c на все возможные углы φ из диапазона $[0, 2\pi]$ представляет собой профиль вращения. Он определяет степень сходства фигуры с самой собой относительно меры Жаккара при различных углах поворота. В качестве точки поворота выбран фокус центральной симметрии [14]. Эксперименты показывают, что шага поворота в один градус достаточно для подавляющего большинства задач. Таким образом дескриптор на основе профиля вращения содержит только 180 чисел в замкнутом интервале от 0 до 1. Дополнительным свойством дескриптора является то, что мы можем получить информацию о том,

насколько симметрична фигура при повороте на разные углы относительно определенной точки [14, 15]. Данный дескриптор обладает инвариантностью к смещению, повороту, отражению и масштабированию, что делает его эффективным в анализе объектов различных форм. Идея данного дескриптора достаточно прозрачна, он имеет явную интерпретацию и простую алгоритмическую реализацию. Кроме того, профиль вращения – это дескриптор без параметров, в отличие от, например, моментов Зернике (радиус и степень).

Классификаторы, используемые в экспериментальных исследованиях. Экспериментальные исследования дескриптора формы проводились на четырех задачах классификации изображений листьев растений. В рамках каждой из задач сравнивались классификаторы, обученные на дескрипторе формы, основанном на профиле вращения [6], моментах Зернике [4], геометрических моментах [9], моментах Ху [5], а также классификаторы, обученные непосредственно на бинарных и исходных цветных изображениях.

Алгоритм вычисления профиля вращения был реализован на C++, с использованием библиотеки компьютерного зрения OpenCV [16]. Геометрические моменты и моменты Ху также были взяты из библиотеки OpenCV. Для вычисления моментов Зернике была использована библиотека компьютерного зрения Mahotas [17].

Для дескрипторов формы были применены классификаторы на основе метода опорных векторов (SVM), градиентного бустинга (XGBoost) [18] и глубоких нейронных сетей (NN Simple). В случае цветных изображений и их бинарных силуэтов в качестве классификаторов применялись сверточные нейронные сети (CNN Simple, ResNet50, VGG16). Такой подход позволил оценить влияние предложенного дескриптора на качество классификации по сравнению с классификаторами, обученными напрямую на цветных и бинарных изображениях.

В качестве потенциальной функции (kernel) классификатора на основе метода опорных векторов использовалась радиальная базисная функция Гаусса (RBF). Параметры классификатора выбирались перебором по сетке при помощи процедуры кросс-проверки.

NN Simple и CNN Simple – простейшие нейронные сети схожей структуры. Первая из них обучалась на профиле вращения. Она включала в себя слой Batch Normalization для входных данных размерности 180, за которым следует полносвязный слой с 180 нейронами и функцией активации ReLU. Затем применяется слой Dropout с коэффициентом отсечения 0,5 для регуляризации, после чего еще один полносвязный с числом нейронов, соответствующим количеству классов, и функцией активации Softmax. Вторая – обучалась непосредственно на изображениях, как на цветных, так и бинарных. Состоит она из четырех сверточных слоев с функцией активации ReLU, последовательно уменьшающих пространственные размеры входных данных через слои максимального пулинга. Затем следуют слой выравнивания и два полносвязных слоя: первый с 128 нейронами и функцией активации ReLU, а второй с числом нейронов, соответствующим количеству классов, и функцией активации Softmax.

Следует подчеркнуть, что оценка точности распознавания проводилась с использованием метода кросс-валидации. для этого обучающая выборка была равномерно разделена на десять непересекающихся частей (фолдов) одинакового размера (10-fold Cross Validation). В каждой из десяти итераций одна часть использовалась в качестве тестовой выборки, а оставшиеся девять частей – для обучения классификатора. Такой подход позволяет максимально эффективно использовать данные, обеспечивая объективную оценку модели на различных подмножествах выборки и снижая риск переобучения. Результаты каждой итерации затем усреднялись для получения итоговой оценки качества распознавания.

Результаты

Задача классификации сканированных изображений листьев растений для экологического мониторинга. В исследовании с использованием базы растений, содержащей сканированные изображения листьев двух широко распространенных видов – *Tilia cordata* (липа) и *Betula pendula* (береза) [19], решались две задачи: отделение листьев березы от листьев липы, а также классификация листьев одного вида, собранных в местах с различным экологическим состоянием. Листья березы имеют более узкую форму, из-за чего мера Жаккара уменьшается при повороте угла. В отличие от них, листья липы характеризуются более округлой формой, и поворот влияет на меру незначительно, что продемонстрировано на Рисунке 1. Такие особенности позволяют оценить эффективность описательных свойств дескриптора на объектах с принципиально разной формой.

Из-за того, что профиль вращения вычисляется на основе бинарных изображений, было необходимо сегментировать и бинаризовать базу растений. Для этой цели был применен метод, основанный на одноклассовой цветовой сегментации с использованием вероятностной гамма-нормальной модели, описанный в работе [20].

Чтобы исключить влияние черенков на форму листьев, черенки были отрезаны следующим образом: сначала были найдены две точки на контуре, евклидово расстояние между которыми меньше первого порогового значения (толщина черенка), и при этом минимальное из двух геодезических расстояний вдоль контура больше второго порогового значения (длина черенка). Затем, с использованием этих двух точек, часть изображения с черенком закрашивалась цветом фона.

Полученные таким образом бинарные изображения в дальнейшем использовались и для обучения сверточных нейронных сетей на силуэтах. Изображения в базе листьев растений имеют разные размеры, в связи с тем, что для обучения нейронных сетей данные должны быть унифицированы, изображения были приведены к размеру 256 на 256 пикселей.

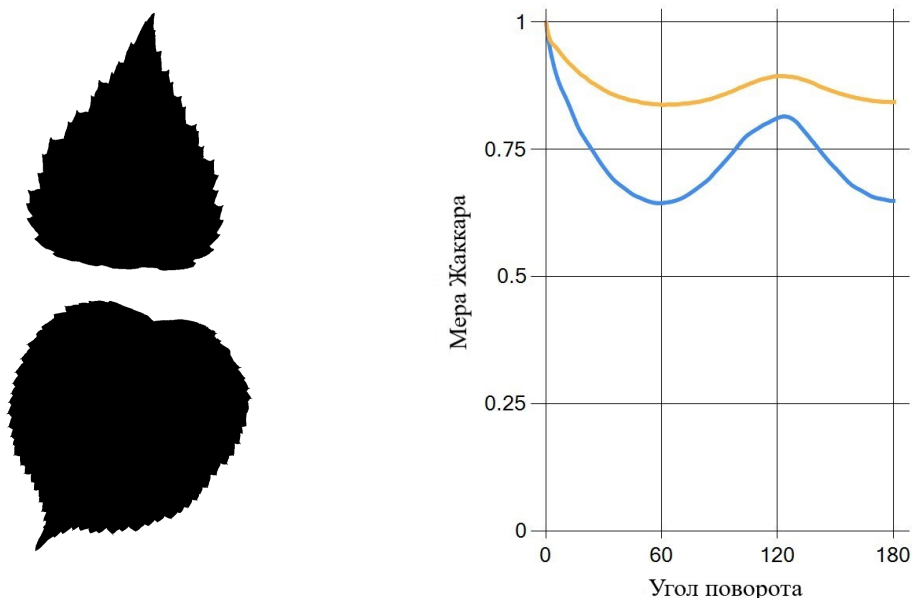


Рисунок 1 – Сравнение профилей, принадлежащих листам березы (слева) и липы (справа)
(синяя линия – лист березы, оранжевая линия – лист липы)

Figure 1 – Comparison of birch (left) and linden (right) leaf profiles (blue curve – birch leaves,
orange curve – linden leaves)

В Таблице 1 представлены результаты работы классификаторов SVM, XGBoost и NN Simple, обученных на профиле вращения (Profile), а также ResNet50, VGG16 и CNN Simple, обученных непосредственно на цветных (Color) и бинарных (Binary) изображениях листьев растений. Параметры TP (True Positive), FP (False Positive), FN (False Negative) и TN (True negative) в Таблице 1 – это стандартные метрики для оценки работы классификаторов, особенно в задачах бинарной классификации, где мы пытаемся разделить данные на два класса: положительный (Positive) и отрицательный (Negative).

Для оценивания качества классификаторов в задачах бинарной классификации использовалась недавно предложенная мера P_4 [21, 22]:

$$P_4 = \frac{4}{\frac{FP+FN}{TP} + \frac{FP+FN}{TN} + 4}. \quad (3)$$

Для задачи многоклассовой классификации использовалась модификация формулы (3), описанная в [23].

В контексте первой задачи класс «Липа» считается положительным классом (Positive), а класс «Береза» – отрицательным классом (Negative). В таблице также приведено количество объектов в базе. Обучение нейронных сетей проводилось на видеокарте Nvidia Tesla A100 40G, а классификаторов SVM и XGBoost – на вычислителе с процессором Intel(R) Xeon(R) Platinum 8255C, 2.50GHz. Среднее время распознавания классификаторами одного объекта приведено в последнем столбце. В этой таблице и далее наилучший результат среди бинарных объектов (все изображение целиком или профиль вращения) выделен жирным шрифтом, в то время как наилучший результат, полученный для дескрипторов на основе формы, выделен курсивом. При этом результаты для цветных растровых изображений приведены только для справки и не рассматриваются как конкурирующие, поскольку они выходят за рамки анализа формы.

Таблица 1 – Качество классификаторов на задаче с отделением листьев липы (1709) от листьев березы (1445)

Table 1 – Classifier performance for distinguishing linden leaves (1709) from birch leaves (1445)

Дескриптор	Классификатор	TP	FP	FN	TN	P_4	Ср. время обработки объекта (мс)
Профиль вращения	SVM	1662	47	34	1411	0,97415	0,008
	XGBoost	1651	58	49	1396	0,96584	0,01
	NN Simple	1679	30	23	1422	0,98308	0,52
Моменты Зернике	SVM	1671	38	42	1403	0,97444	0,004
	XGBoost	1674	35	31	1414	0,97892	0,01
	NN Simple	1684	25	16	1429	0,98691	0,88
Геом. моменты	SVM	1490	219	226	1219	0,85768	0,02
	XGBoost	1633	76	88	1357	0,94757	0,01
	NN Simple	1650	59	84	1361	0,95425	0,27
Моменты Ху	SVM	1608	101	136	1309	0,92411	0,01
	XGBoost	1634	75	101	1344	0,94369	0,04
	NN Simple	1641	68	87	1358	0,95043	0,27
Бинарные изображения	ResNet50	1518	191	19	1426	0,93335	4,9
	VGG16	1687	22	11	1434	0,98946	6,19
	CNN Simple	1666	43	40	1405	0,97349	2,7
Цветные изображения	ResNet50	1512	197	135	1310	0,89425	8,17
	VGG16	1690	19	13	1432	0,98978	2,97
	CNN Simple	1682	27	41	1404	0,97826	2,82

Вторая задача, выполняемая с использованием базы сканированных изображений, заключается в идентификации листьев деревьев одного вида, произрастающих в различных экологических условиях. Листья растений являются одним из наиболее доступных и информативных объектов для изучения и анализа [19], и служат важным интегральным показателем состояния экосистемы. Будем условно называть районы произрастания растений, как и сами листья, «чистыми» и «грязными» в зависимости от экологического состояния окружающей среды. В качестве вида растения для анализа была взята *Tilia cordata* (липа) [19]. На Рисунке 2 представлены профили, принадлежащие чистым (лист слева) и грязным (лист справа) листьям липы. Можно заметить, что форма самих листьев схожа и обнаруживает различия практически только в величине и форме зубчиков, что приводит к некоторым трудностям при классификации (Таблица 2). Класс «Чистая липа» отнесен к категории Positive, а класс «Грязная липа» – к категории Negative. Как видно в этой задаче описатель формы на основе профиля вращения показал наилучшие результаты, даже в сравнении с полными бинарными изображениями листьев.

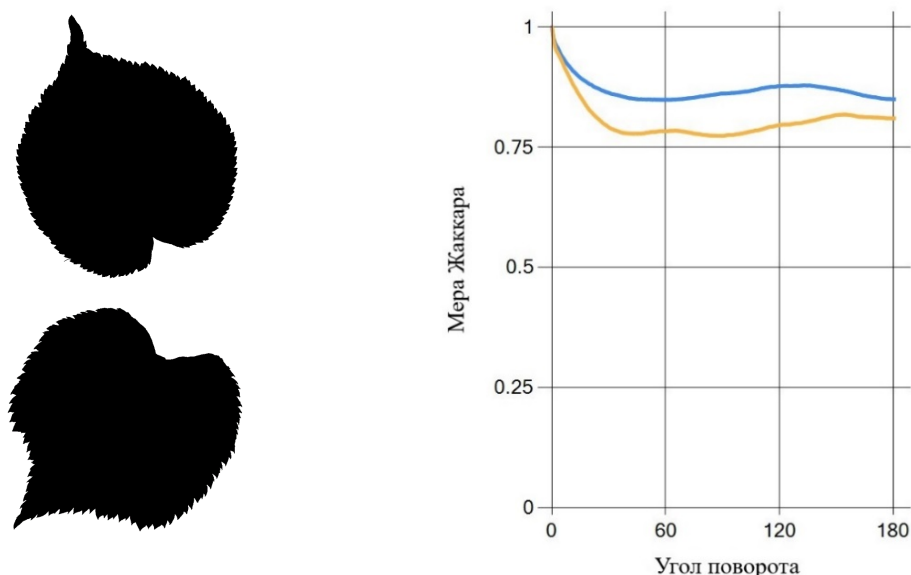


Рисунок 2 – Сравнение профилей, принадлежащих чистому (слева) и грязному (справа) листам липы (синяя линия – чистый лист, оранжевая линия – грязный лист)

Figure 2 – Comparison of the profiles of clean (left) and contaminated (right) linden leaves (blue curve – clean leaf, orange curve – contaminated leaf)

Таблица 2 – Качество классификаторов на задаче с отделением грязных листьев липы (948) от чистых (761)

Table 2 – Classifier performance for distinguishing clean (761) and contaminated (948) linden leaves

Дескриптор	Классификатор	TP	FP	FN	TN	P_4	Ср. время обработки объекта (мс)
Профиль вращения	SVM	526	235	198	750	0,74066	0,15
	XGBoost	530	231	198	750	0,74329	0,03
	NN Simple	582	179	184	764	0,78449	0,35
Моменты Зернике	SVM	530	231	175	773	0,75596	0,04
	XGBoost	524	237	208	740	0,73386	0,01
	NN Simple	505	256	168	780	0,74305	0,49

Таблица 2 (продолжение)
Table 2 (continued)

Геом. моменты	SVM	438	323	242	706	0,65678	0,04
	XGBoost	492	269	231	717	0,70008	0,08
	NN Simple	493	268	216	732	0,70884	0,49
Моменты Ху	SVM	477	284	332	616	0,63578	0,05
	XGBoost	452	309	247	701	0,66409	0,13
	NN Simple	435	326	166	782	0,69442	0,49
Бинарные изображения	ResNet50	479	282	559	389	0,50519	6,14
	VGG16	591	170	220	728	0,76988	5,22
	CNN Simple	536	225	212	736	0,7395	3,33
Цветные изображения	ResNet50	573	188	472	476	0,61177	7,28
	VGG16	582	179	173	775	0,79066	4,16
	CNN Simple	561	200	163	785	0,78285	4,87

Многоклассовая задача классификации изображений листьев растений из базы FLAVIA. В другой задаче классификации применялась известная коллекция бинарных растровых изображений листьев растений FLAVIA [24] (Рисунок 3). В состав базы входит 1907 изображений листьев с разрешением 800 на 600 пикселей, представляющих тридцать два различных класса. В каждом классе количество образцов находится в диапазоне от 50 до 70 штук. С целью снижения требований к вычислительным ресурсам для обучения нейронных сетей изображения были сжаты до размеров 200 на 150 пикселей.

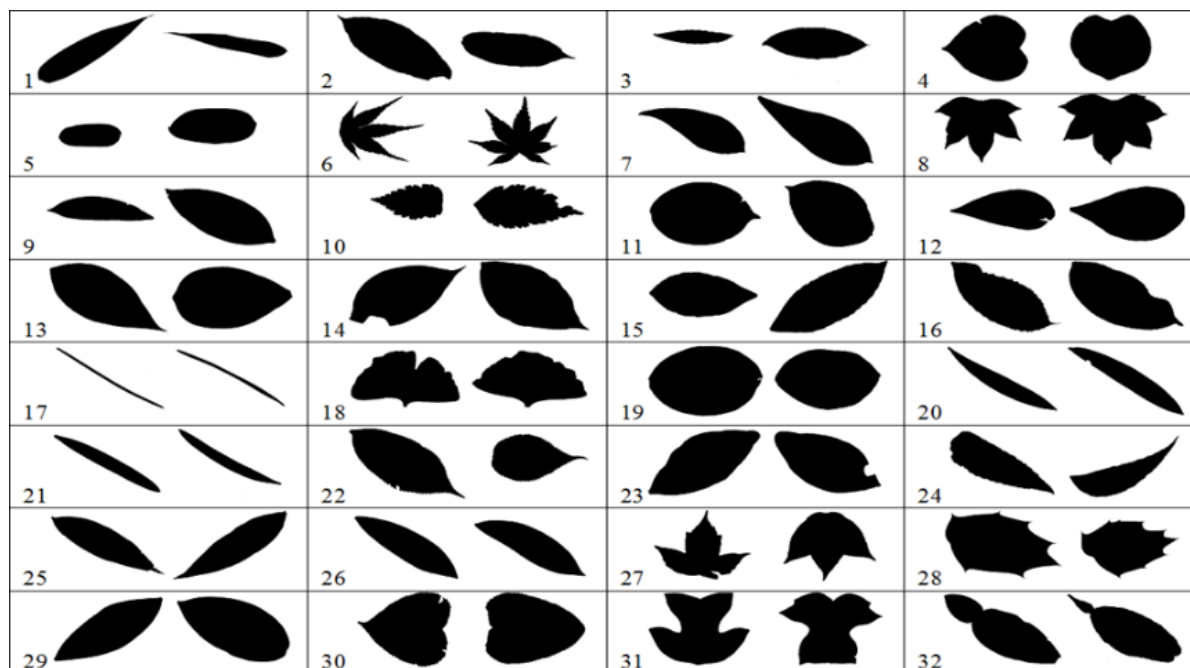


Рисунок 3 – Примеры бинаризованных изображений различных классов из базы FLAVIA
Figure 3 – Examples of binarized images from various classes of the FLAVIA database

В Таблице 3 представлено качество работы различных классификаторов по распознаванию 32 классов изображений листьев из базы FLAVIA. Согласно Таблице 3, наилучшие результаты показали нейронные сети VGG16 и CNN Simple, которые обучались на бинарных изображениях. Следующие два места заняли классификаторы, обученные на профиле вращения (NN Simple, SVM).

Таблица 3 – Качество классификаторов на задаче с распознаванием изображений листьев из базы FLAVIA

Table 3 – Classifier performance for distinguishing leaf images from the FLAVIA database

Дескриптор	Классификатор	P_4	Ср. время обработки объекта (мс)
Профиль вращения	SVM	0,90320	0,15
	XGBoost	0,83931	0,02
	NN Simple	0,89615	0,34
Моменты Зернике	SVM	0,84472	0,1
	XGBoost	0,78693	0,02
	NN Simple	0,85384	0,3
Геом. моменты	SVM	0,82189	0,09
	XGBoost	0,85941	0,02
	NN Simple	0,86482	0,4
Моменты Ху	SVM	0,71214	0,08
	XGBoost	0,64761	0,06
	NN Simple	0,73127	0,43
Бинарные изображения	ResNet50	0,84134	8,09
	VGG16	0,91626	3,95
	CNN Simple	0,90817	1,49

Матрицы ошибок результата классификации для бинарного изображения (классификатор на основе сверточной нейронной сети VVG16) и дескриптора формы на основе профиля вращения (классификатор нейросетью NN Simple) представлены на Рисунках 4 и 5 соответственно. Метки классов, находящиеся на рисунках слева, – истинные, а сверху – предсказанные.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32			
1	51		3																		3			1	1								5		
2		46	3				1	3	1						2								2	1	2				2				6		
3	1	1	57						6																									6	
4				68																									1		3			7	
5					67					5					1																			7	
6						56																												5	
7							58	2																1										1	
8								52																										5	
9	2	7		1		1		32	2							1								2	4	2			1					5	
10				4	1					48	3	1	1										1											5	
11			2								34	4	2					2				3							3					5	
12										1		61																		1				6	
13										1	4		41	2									4											5	
14										1	3		2	50		1							3	1					3				1	6	
15	2			1					1				1		53									1						1				6	
16		4							1					1	42								2	5					1					5	
17																	77																	7	
18																			62															6	
19				1							3									56									1					6	
20	1																						64	1										6	
21	2																						1	57										6	
22										2	5		3											43	2									5	
23		2		1							1			2	2	6		1					2	36					1	1				5	
24							2																		55	6	2							6	
25			1							8															5	36	4							5	
26	1								1													2			1	1	46							5	
27						1						1																51						5	
28				1						5				4																				5	
29		2													4	2																		5	
30				1																		1												6	
31																																	52		5
32											2												1											53	5
	56	59	71	74	73	58	62	52	54	61	55	67	52	62	62	52	77	62	63	67	61	59	47	66	50	54	52	51	56	65	52	55			

Рисунок 4 – Матрица ошибок результата классификации базы изображений FLAVIA, полученной через кросс-проверку нейронной сетью VGG16, обученной на бинарных изображениях

Figure 4 – Confusion matrix for the FLAVIA image database classification obtained through cross-validation by the VGG16 neural network trained on binary images

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32					
1	56																			1	2												59				
2		55	1						2							3						1							1					63			
3	1	1	44												1					6					4	8								65			
4				71																1						4	8							72			
5					72															1														73			
6						56																												56			
7			1				53		1																	5								2	62		
8								52																											52		
9		3	1				4		21							8	1							1	1	12	1			2				55			
10										56				1																2				59			
11				1							28		3	2						15			1											50			
12												62							1															63			
13											3	2	32	12		1													2					52			
14											5		4	53	1			1						1										65			
15							1		3				1		42	1								5						7				60			
16		5			1								1	6	2	38														3				56			
17																	77																		77		
18												1							61																62		
19				1							2									58															61		
20																					64	2													66		
21	5																					55													60		
22											1			2		2							47	1					2						55		
23		1												1	8				1	5				39											55		
24							1		1																59	4									65		
25		1	7				6		5						2										1	28	3			1					54		
26			5																							2	45								52		
27																													52				1		53		
28											4																			51					55		
29			1		1		2		1							6	3									1	1			41					57		
30				5																												59			64		
31																																1	51			53	
32			1								1				2														1							52	56
	62	66	61	77	75	56	67	52	34	57	43	65	42	78	70	49	77	64	80	71	59	48	48	61	56	58	53	57	55	60	52	54					

Рисунок 5 – Матрица ошибок результата классификации базы изображений FLAVIA, полученной через кросс-проверку нейронной сетью NN Simple, обученной на профиле вращения
Figure 5 – Confusion matrix for the FLAVIA image database classification obtained through cross-validation by the NN Simple neural network trained on the Rotation Profile

Многоклассовая задача классификации изображений листьев растений из базы Swedish leaf. Для последней задачи оценки классификации с помощью профиля вращения является многоклассовая задача распознавания листьев растений из базы Swedish leaf¹ (Рисунок 6). В базе содержится 15 видов (классов) изображений листьев растений. В каждом классе содержится по 75 изображений.

Так как база Swedish leaf имеет цветные изображения, а профиль вращения вычисляется на основе бинарных изображений, то ее необходимо было сегментировать и бинаризовать. Для этой цели, как и в задаче с листьями березы и липы, был применен метод, описанный в работе [20]. Однако этот метод не учитывает дыры в объектах на изображениях и закрашивает их. Поэтому для двух классов растений (*Sorbus aucuparia*, *Sorbus intermedia*) была применена модель сегментации Background Erase Network (BEN) [25]. Помимо этого, для исключения влияния черенков на форму листьев они были удалены тем же способом, что и в задаче классификации сканированных изображений листьев растений для экологического мониторинга.

В Таблице 4 представлено качество работы различных классификаторов по распознаванию 32 классов изображений листьев из базы Swedish leaf. Согласно Таблице 4, наилучшие результаты показала нейронная сеть VGG16, которая обучалась на бинарных изображениях. Следующие два места заняли классификаторы, обученные на профиле вращения (NN Simple, SVM).

¹ Söderkvist O.J.O. Swedish Leaf Dataset. Linköping universitet. URL: <https://www.cvl.isy.liu.se/en/research/datasets/swedish-leaf/> (дата обращения: 21.05.2025).

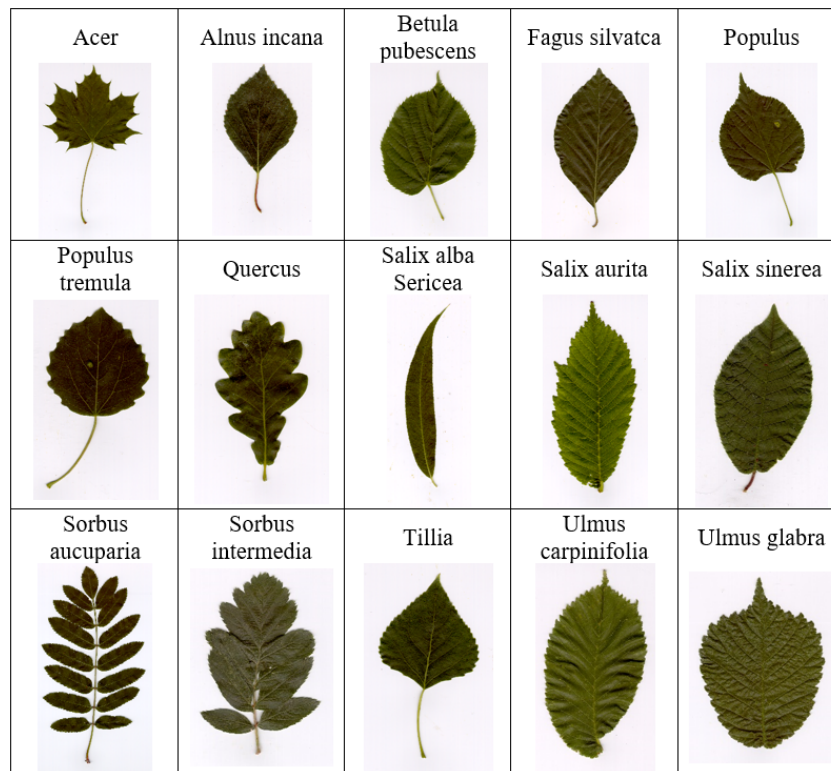


Рисунок 6 – Примеры изображений различных классов из базы Swedish leaf
Figure 6 – Examples of images from various classes of the Swedish leaf database

Таблица 4 – Качество классификаторов на задаче с распознаванием изображений листьев из базы Swedish leaf

Table 4 – Classifier performance for distinguishing leaf images from the Swedish leaf database

Дескриптор	Классификатор	P_4	Ср. время обработки объекта (мс)
Профиль вращения	SVM	0,94855	0,06
	XGBoost	0,93068	0,02
	NN Simple	0,96856	0,7
Моменты Зернике	SVM	0,83855	0,04
	XGBoost	0,81215	0,02
	NN Simple	0,84530	0,6
Геом. моменты	SVM	0,90962	0,02
	XGBoost	0,90063	0,02
	NN Simple	0,91851	0,6
Моменты Ху	SVM	0,73929	0,03
	XGBoost	0,78511	0,02
	NN Simple	0,81693	0,6
Бинарные изображения	ResNet50	0,85113	8,58
	VGG16	0,97152	5,11
	CNN Simple	0,94366	2,09
Цветные изображения	ResNet50	0,84642	8,52
	VGG16	0,96659	4,11
	CNN Simple	0,949	2,13

Матрицы ошибок результата классификации для бинарного изображения (классификатор на основе сверточной нейронной сети VVG16) и дескриптора формы на основе профиля вращения (классификатор нейросетью NN Simple) представлены на Рисунках 7 и 8 соответственно. Метки классов, находящиеся на рисунках слева, – истинные, а сверху – предсказанные.

	Acer	Alnus incana	Betula pubescens	Fagus sylvatica	Populus	Populus tremula	Quercus	Salix alba Sericea	Salix aurita	Salix sinerea	Sorbus aucuparia	Sorbus intermedia	Tilia	Ulmus carpinifolia	Ulmus glabra
Acer	75														75
Alnus incana		73	1	1											75
Betula pubescens		1	69		3	1							1		75
Fagus sylvatica				73					2						75
Populus			2		72	1									75
Populus tremula			1			73						1			75
Quercus							75								75
Salix alba Sericea								75							75
Salix aurita						1		66	3				3	2	75
Salix sinerea				3		1		1	69		1				75
Sorbus aucuparia										75					75
Sorbus intermedia											75				75
Tilia												75			75
Ulmus carpinifolia				2				2					64	7	75
Ulmus glabra								6	2				6	61	75
	75	74	73	79	75	75	77	75	75	76	75	76	76	74	70

Рисунок 7 – Матрица ошибок результата классификации базы изображений Swedish leaf, полученной через кросс-проверку нейронной сетью VGG16, обученной на бинарных изображениях
Figure 7 – Confusion matrix for the Swedish leaf image database classification obtained through cross-validation by the VGG16 neural network trained on binary images

	Acer	Alnus incana	Betula pubescens	Fagus sylvatica	Populus	Populus tremula	Quercus	Salix alba Sericea	Salix aurita	Salix sinerea	Sorbus aucuparia	Sorbus intermedia	Tilia	Ulmus carpinifolia	Ulmus glabra
Acer	75														75
Alnus incana		72	2	1											75
Betula pubescens			68		5						1	1			75
Fagus sylvatica				74				1							75
Populus			3		72										75
Populus tremula						73						2			75
Quercus							75								75
Salix alba Sericea								75							75
Salix aurita									57				1	17	75
Salix sinerea				1					74						75
Sorbus aucuparia										75					75
Sorbus intermedia											75				75
Tilia					1							74			75
Ulmus carpinifolia			1					1					67	6	75
Ulmus glabra			1					12					1	61	75
	75	72	75	76	77	74	75	75	70	75	75	77	70	84	

Рисунок 8 – Матрица ошибок результата классификации базы изображений Swedish leaf, полученной через кросс-проверку нейронной сетью NN Simple, обученной на профиле вращения
Figure 8 – Confusion matrix for the Swedish leaf image database classification obtained through cross-validation by the NN Simple neural network trained on the Rotation Profile

Заключение

Исследование предложенного дескриптора было произведено на нескольких задачах классификации изображений. Полученные результаты обобщены на Рисунке 9. Наилучшая средняя мера P_4 по всем четырем прикладным задачам для дескриптора формы на основе профиля вращения (классификатор NN Simple) превзошла все остальные дескрипторы (моменты X_y , геометрические моменты и моменты Зернике). В то же время в сравнении с анализом полного бинарного растрового изображения (классификатор VGG16) ухудшение составляет всего 0,003 процента. Отметим, что профиль вращения дает более компактное и гибкое описание, инвариантное к различным преобразованиям фигуры и обеспечивает существенно быстрое распознавание по сравнению с анализом полного изображения. Из Таблиц 1–4 следует, что такое ускорение может быть оценено как 3–10 раз.

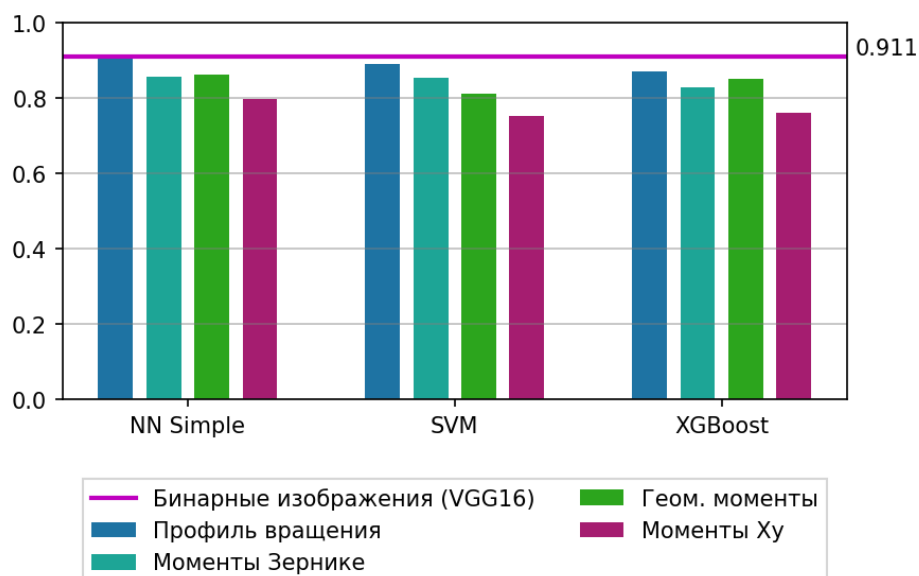


Рисунок 9 – Мера P_4 классификаторов для различных дескрипторов формы, усредненная по всем четырем прикладным задачам, в сравнении с наилучшей усредненной мерой P_4 классификатора на основе CNN (VGG16) для бинарных растровых изображений
Figure 9 – Measure P_4 of classifiers for different shape descriptors, averaged on all four applied problems in comparison with the best averaged measure P_4 of CNN-based classifier (VGG16) on binary bitmaps

Классификаторы, обученные на профиле вращения, хорошо решают не только такую тривиальную задачу, как разделение листьев березы от липы, но и более сложную задачу разделения листьев липы, произрастающих в благоприятных условиях и подверженных факторам экологического загрязнения. Можно сделать вывод, что профиль вращения является надежным индикатором для классификации растительной флоры, несмотря на практически идентичную форму листьев, собранных в разных экологических точках. Кроме того, они успешно справились с задачами многоклассовой классификация листьев различных видов растений.

СПИСОК ИСТОЧНИКОВ / REFERENCES

1. Zhang D., Lu G. Shape-Based Image Retrieval Using Generic Fourier Descriptor. *Signal Processing: Image Communication*. 2002;17(10):825–848. [https://doi.org/10.1016/S0923-5965\(02\)00084-X](https://doi.org/10.1016/S0923-5965(02)00084-X)

2. Kuhl F.P., Giardina Ch.R. Elliptic Fourier Features of a Closed Contour. *Computer Graphics and Image Processing*. 1982;18(3):236–258. [https://doi.org/10.1016/0146-664X\(82\)90034-X](https://doi.org/10.1016/0146-664X(82)90034-X)
3. Dudek G., Tsotsos J.K. Shape Representation and Recognition from Multiscale Curvature. *Computer Vision and Image Understanding*. 1997;68(2):170–189. <https://doi.org/10.1006/cviu.1997.0533>
4. Zhang D., Lu G. A Comparative Study of Three Region Shape Descriptors. In: *DICTA2002: Digital Image Computing Techniques and Applications, 21–22 January 2002, Melbourne, Australia*. 2002. P. 86–91.
5. Li E., Li H. Reflection Invariant and Symmetry Detection. arXiv. URL: <https://arxiv.org/abs/1705.10768> [Accessed 21st May 2025].
6. Seredin O., Lomov N., Liakhov D., et al. Rotation Profile-Based Binary Shape Descriptor. *The Visual Computer*. 2025;41:8911–8933. <https://doi.org/10.1007/s00371-025-03906-9>
7. Qi Sh., Zhang Yu., Wang Ch., Zhou J., Cao X. A Survey of Orthogonal Moments for Image Representation: Theory, Implementation, and Evaluation. *ACM Computing Surveys*. 2021;55(1). <https://doi.org/10.1145/3479428>
8. Xiao B., Ma J.-F., Wang X. Image Analysis by Bessel-Fourier Moments. *Pattern Recognition*. 2010;43(8):2620–2629. <https://doi.org/10.1016/j.patcog.2010.03.013>
9. Teague M.R. Image Analysis via the General Theory of Moments. *Journal of the Optical Society of America*. 1980;70(8):920–930. <https://doi.org/10.1364/JOSA.70.000920>
10. Bedratyuk L. On Complete System of Invariants for the Binary Form of Degree 7. arXiv. URL: <https://arxiv.org/abs/math/0611122> [Accessed 21st May 2025].
11. Рогов А.А., Быстров М.Ю. Структурное распознавание бинарных изображений с использованием скелетов. *Математические методы распознавания образов*. 2011;15(1):420–423.
12. Kervadec H., Bahig H., Létourneau-Guillon L., Dolz J., Ayed I.B. Beyond Pixel-Wise Supervision for Segmentation: A Few Global Shape Descriptors Might Be Surprisingly Good! In: *Medical Imaging with Deep Learning, 07–09 July 2021, Lübeck, Germany*. PMLR; 2021. P. 354–368.
13. Jaccard P. Étude comparative de la distribution florale dans une portion des Alpes et du Jura. *Bulletin de la Société Vaudoise des Sciences Naturelles*. 1901;37(142):547–579. <https://doi.org/10.5169/seals-266450>
14. Seredin O., Liakhov D., Kushnir O., Lomov N. Jaccard Index-Based Detection of Order 2 Rotational Quasi-Symmetry Focus for Binary Images. *Pattern Recognition and Image Analysis*. 2022;32(3):672–681. <https://doi.org/10.1134/S1054661822030403>
15. Lomov N., Seredin O., Kushnir O., Liakhov D. Search for Rotational Symmetry of Binary Images via Radon Transform and Fourier Analysis. In: *Proceedings of the 18th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications: Volume 4, 19–21 February 2023, Lisbon, Portugal*. SCITEPRESS; 2023. P. 280–289. <https://doi.org/10.5220/0011679900003417>
16. Bradski G. The OpenCV Library. *Dr. Dobb's Journal: Software Tools for the Professional Programmer*. 2000;25(11):120–123.
17. Coelho L.P. Mahotas: Open Source Software for Scriptable Computer Vision. arXiv. URL: <https://arxiv.org/abs/1211.4907> [Accessed 21st May 2025].
18. Chen T., Guestrin C. XGBoost: A Scalable Tree Boosting System. In: *KDD '16: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 13–17 August 2016, San Francisco, CA, USA*. New York: Association for Computing Machinery; 2016. P. 785–794. <https://doi.org/10.1145/2939672.2939785>

19. Сенин А.Н., Тирас Х.П., Местецкий Л.М. Генерация признаков формы по изображениям листьев для биоиндикации. *Экология. Экономика. Информатика. Серия: Системный анализ и моделирование экономических и экологических систем*. 2022;1(7):66–75. <https://doi.org/10.23885/2500-395X-2022-1-7-66-75>
Senin A.N., Tiras K.P., Mestetskiy L.M. Shape Feature Generation for Bioindication Using Leaf Images. *Ecology. Economy. Informatics. System Analysis and Mathematical Modeling of Ecological and Economic Systems*. 2022;1(7):66–75. (In Russ.). <https://doi.org/10.23885/2500-395X-2022-1-7-66-75>
20. Liakhov D.V., Mityugov N.S., Gracheva I.A., Kopylov A.V., Seredin O.S., Tiras Kh.P. Scanned Plant Leaves Boundary Detection in the Presence of a Colored Shadow. *Pattern Recognition and Image Analysis*. 2022;32(3):575–585. <https://doi.org/10.1134/S1054661822030221>
21. Sitarz M. Extending F1 Metric, Probabilistic Approach. *Advances in Artificial Intelligence and Machine Learning*. 2023;3(2):1025–1038. <https://doi.org/10.54364/aaiml.2023.1161>
22. Sitarz M. Extending F1 Metric, Probabilistic Approach. arXiv. URL: <https://arxiv.org/abs/2210.11997> [Accessed 21st May 2025].
23. Seredin O.S., Kopylov A.V. Harmonic Averaging in Classifier Quality Assessment. *Pattern Recognition and Image Analysis*. 2024;34(4):1160–1171. <https://doi.org/10.1134/S1054661824701220>
24. Wu S.G., Bao F.Sh., Xu E.Y., Wang Yu-X., Chang Y.-F., Xiang Q.-L. A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network. In: *2007 IEEE International Symposium on Signal Processing and Information Technology, 15–18 December 2007, Giza, Egypt*. IEEE; 2007. P. 11–16. <https://doi.org/10.1109/ISSPIT.2007.4458016>
25. Meyer M., Spruyt J. BEN: Using Confidence-Guided Matting for Dichotomous Image Segmentation. arXiv. URL: <https://arxiv.org/abs/2501.06230> [Accessed 21st May 2025].

ИНФОРМАЦИЯ ОБ АВТОРАХ / INFORMATION ABOUT THE AUTHORS

Середин Олег Сергеевич, кандидат физико-математических наук, доцент, Тульский государственный университет, Тула, Российская Федерация. **Oleg S. Seredin**, Candidate of Physico-Mathematical Sciences, Docent, Tula State University, Tula, the Russian Federation.
e-mail: oseredin@yandex.ru
ORCID: [0000-0003-0410-7705](https://orcid.org/0000-0003-0410-7705)

Ломов Никита Александрович, кандидат физико-математических наук, Тульский государственный университет, Тула, Российская Федерация. **Nikita A. Lomov**, Candidate of Physico-Mathematical Sciences, Tula State University, Tula, the Russian Federation.
e-mail: nikita-lomov@mail.ru
ORCID: [0000-0003-4286-1768](https://orcid.org/0000-0003-4286-1768)

Ляхов Даниил Викторович, аспирант, Тульский государственный университет, Тула, Российская Федерация. **Daniil V. Liakhov**, Postgraduate, Tula State University, Tula, the Russian Federation.
e-mail: liakhov.daniil@mail.ru
ORCID: [0000-0003-1105-9780](https://orcid.org/0000-0003-1105-9780)

Митюгов Никита Сергеевич, аспирант, **Nikita S. Mityugov**, Postgraduate, Tula State University, Tula, the Russian Federation.

e-mail: nikita.mityugov.2001@mail.ru

ORCID: [0000-0001-7319-7068](https://orcid.org/0000-0001-7319-7068)

Кушнир Олеся Александровна, кандидат технических наук, доцент, Тульский государственный университет, Тула, Российская Федерация. **Olesia A. Kushnir**, Candidate of Engineering Sciences, Docent, Tula State University, Tula, the Russian Federation.

e-mail: kushnir-olesya@rambler.ru

ORCID: [0000-0001-7879-9463](https://orcid.org/0000-0001-7879-9463)

Копылов Андрей Валериевич, кандидат технических наук, доцент, Тульский государственный университет, Тула, Российская Федерация. **Andrei V. Kopylov**, Candidate of Engineering Sciences, Docent, Tula State University, Tula, the Russian Federation.

e-mail: and.kopylov@gmail.com

ORCID: [0000-0003-3193-583X](https://orcid.org/0000-0003-3193-583X)

Статья поступила в редакцию 14.10.2025; одобрена после рецензирования 14.11.2025; принята к публикации 30.12.2025.

The article was submitted 14.10.2025; approved after reviewing 14.11.2025; accepted for publication 30.12.2025.