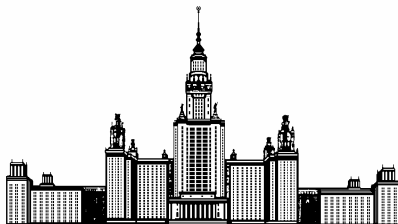


Московский государственный университет имени М. В. Ломоносова



Факультет Вычислительной Математики и Кибернетики

Кафедра Математических Методов Прогнозирования

## **КУРСОВАЯ РАБОТА**

**«Генерация признаков формы по изображениям листьев  
для биоиндикации»**

**«Shape feature generation for bioindication using leaf images»**

Выполнил студент 5 курса  
517 группы:

*Сенин Александр Николаевич*

Научный руководитель:  
д.т.н., профессор

*Местецкий Леонид Моисеевич*

Москва, 2022

# Содержание

<b>1.</b>	<b>Введение .....</b>	<b>3</b>
<b>2.</b>	<b>Постановка задачи .....</b>	<b>4</b>
<b>3.</b>	<b>Описание данных .....</b>	<b>4</b>
<b>4.</b>	<b>Извлечение признаков формы .....</b>	<b>5</b>
4.1.	Удаление черенка.....	5
4.2.	Построение аппроксимирующего многоугольника .....	6
4.2.1.	Вычисление площади и периметра .....	6
<b>5.</b>	<b>Связь признаков формы с экологической обстановкой .....</b>	<b>6</b>
5.1.	Форм-фактор.....	6
5.2.	Влияние экологической обстановки .....	7
5.2.1.	Корректность наблюдений.....	7
5.3.	Качественные признаки формы .....	9
<b>6.</b>	<b>Классификация экологической обстановки .....</b>	<b>10</b>
6.1.	Классификация по одному листу .....	10
6.2.	Классификация по выборке листьев .....	11
6.2.1.	Оценка числа листьев для надежности классификации.....	11
6.3.	Состязательная классификация двух районов .....	12
6.4.	Интерпретация результата .....	13
<b>7.</b>	<b>Выводы.....</b>	<b>13</b>
<b>8.</b>	<b>Список источников.....</b>	<b>14</b>

### **Аннотация**

Простые маркеры, отражающие состояние окружающей среды в регионе, ценны в задачах экологии. Такими маркерами могут быть листья деревьев, например, липы. В работе показано, что форма листьев зависит от экологии в районе, где растет дерево, причем это наблюдается и для простых признаков формы, например, для периметра и площади. Удаление черенка перед извлечением признаков снижает зашумленность в данных, предложен алгоритм удаления. Полученные признаки можно использовать для косвенной оценки экологического состояния по собранным листьям. Построены классификаторы состояния по одному листу, по выборке листьев, предложен подход для сравнения качества экологии в двух районах. Проведено сравнение с принятым в экспертном обществе подходом к оцениванию состояния по извлеченным признакам, показано преимущество предложенных методов классификации.

## 1. Введение

Развитие современных естественных наук во-многом определяется прогрессом в области информационных технологий. Технология цифровой фотографии позволяет получать высококачественные изображения в большом разрешении для самых разных объектов исследования. В свою очередь современные подходы к обработке и распознаванию изображений дают возможность напрямую проверять гипотезы, поставленные биологами и экологами. В случае выполнения некоторых гипотез возможно построить алгоритмы, которые будут решать экологическую задачу методами компьютерного зрения и машинного обучения.

Экологи давно ищут простые маркеры, отражающие состояние окружающей среды в регионе. Такими маркерами могут быть листья деревьев, в частности липы. Результатом многолетних наблюдений стала гипотеза о зависимости формы листа от экологической обстановки в месте, где растет дерево. Было высказано предположение, что выпуклость зубчиков листа коррелирует с экологическим состоянием региона.

Для работы с листьями, проверки гипотез о их форме удобно использовать сканирование. В результате получается цифровое изображение в большом разрешении, к которому можно применять алгоритмы обработки и распознавания изображений, а затем получать статистики и проверять гипотезы.

В обществе экспертов, использующих листья в качестве маркеров окружающей среды, существуют устоявшиеся численные показатели, описывающие форму листа. Общее название таких показателей – форм-фактор. Одним из таких показателей является нормированное отношение площади к квадрату периметра. Эмпирический опыт экспертов экологов подсказывал, что такой форм-фактор достаточен для определения экологического состояния региона.

Естественно предположить, что выпуклость зубчиков существенно влияет на площадь листа и несущественно влияет на его периметр. Поэтому для проверки гипотезы о выпуклости зубчиков было решено начать с оценивания площади и периметра для листьев из разных по экологическому состоянию регионов.

Было обнаружено, что черенки листьев вносят существенную шумовую компоненту в оцениваемые параметры периметра и площади. Варьируя длину черенка, можно значительно менять периметр листа, при этом почти не меняя его площадь. Для избавления от шумовой компоненты был построен алгоритм удаления черенка.

В результате выяснилось, что форм-фактор отношения площади к квадрату периметра мало коррелирует с экологическим положением. Тем не менее, сами по себе признаки площади и периметра оказались достаточно сильными. Была обнаружена зависимость между этими показателями формы листьев и экологической обстановкой мест, где выросли деревья липы. Более того, обнаруженная зависимость позволяет построить классификаторы с привлечением алгоритмов машинного обучения. Например, простейший линейный классификатор позволяет решать задачу определения экологической обстановки в заданном регионе по одному листу с качеством выше 0.93 ROC AUC на тестовой выборке.

Возникла идея оценивать экологическую обстановку не по одному листу, а по выборке. Построен алгоритм, определяющий состояние окружающей по выборке листьев. Сценарий использования такого алгоритма выглядит перспективно – для оценки экологического состояния территории достаточно собрать несколько листьев с дерева липы, отсканировать их и применить алгоритм. В таком сценарии не потребуются прибегать к помощи экспертов в области экологии, специально замерять загрязненность воздуха, оценивать близость к автомобильным трассам и вредным производствам, учитывать розы ветров и т.п.

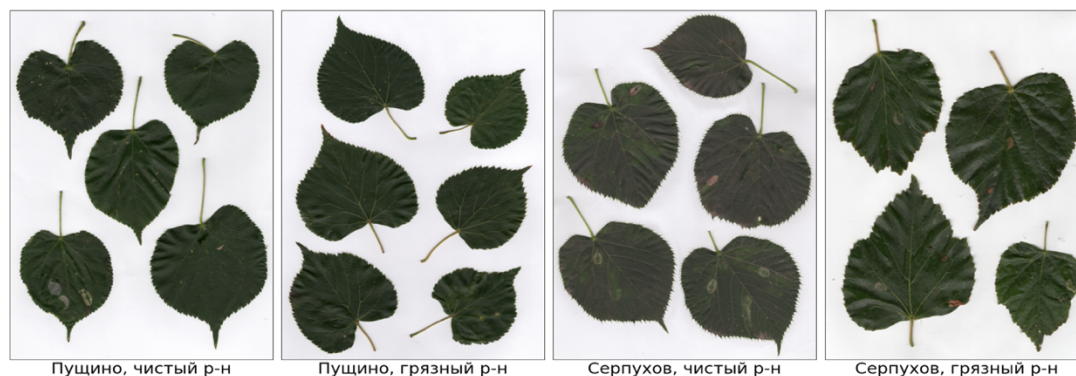
## 2. Постановка задачи

Будем работать с листьями липы. Экспертами экологами было высказано предположение о существовании зависимости между выпуклостью зубчиков на листе и экологической обстановкой в районе, где растет дерево. Было решено начать с более простой гипотезы о существовании связи между простейшими признаками формы (площадь и периметр) и экологическим состоянием региона. Задача строится из следующих шагов:

1. Построить многоугольную аппроксимацию листа, удалить черенок, оценить простейшие признаки формы.
2. Проверить предположение о связи между признаками формы и экологической ситуацией в регионе.
3. Предложить классификатор, определяющий экологическую обстановку по одному листу. Предложить аналогичный классификатор по выборке листьев.
4. Сравнить предложенный подход с принятым в экспертном обществе подходом, использующим форм-фактор отношения площади к квадрату периметра.

## 3. Описание данных

Экспертами экологами были собраны листья липы в двух городах Московской области: в Пушкино и Серпухове. Деревья выбраны одного биологического вида, но в разных по экологическому состоянию районах. Будем называть районы экологически чистыми и экологически грязными соответственно. В Пушкино листья собраны с 5 деревьев: 2 дерева в чистом районе, 3 в грязном. В Серпухове листья собраны с 8 деревьев: 3 в чистом районе и 5 в грязном. В среднем с одного дерева собрано около 75 листьев, после фильтрации выбросов имеем в совокупности 628 листьев из грязного района и 294 листа из чистого района.



**Рисунок 1** – Примеры сканов листьев из чистого и грязного районов в Пушкино и Серпухове.

## 4. Извлечение признаков формы

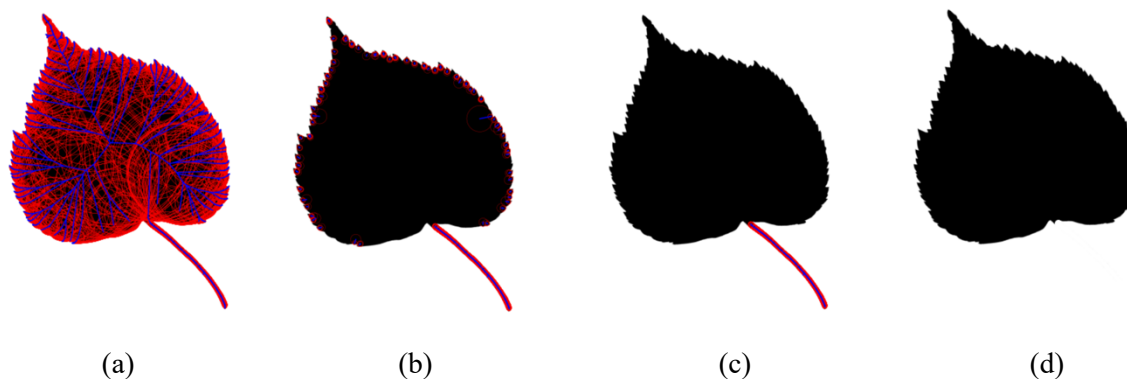
Будем вычислять простейшие признаки формы листа – площадь и периметр (длина границы). На рисунке 1 приведены примеры сканов листьев. Черенки листьев вносят существенную шумовую компоненту в вычисляемые признаки формы. Варьируя длину черенка, можно значительно менять периметр листа, при этом почти не меняя его площадь. Поэтому после бинаризации необходимо удалить черенки, а после вычислить площадь и периметр фигуры каждого листа.

### 4.1. Удаление черенка

Для удаления черенка будем строить скелетное представление для листа. Скелетом будем называть множество центров всех максимальных пустых кругов фигуры, т.е. всех пустых кругов, каждый из которых не содержится ни в каком другом пустом круге. Пустым кругом фигуры будем называть такое замкнутое, целиком лежащее в фигуре множество точек, что расстояние между любыми двумя точками множества не превосходит радиуса круга. Будем представлять скелет в виде связного плоского графа, вершины будут отвечать центрам максимальных пустых кругов, каждой вершине приписан радиус соответствующего пустого круга. Подробное описание алгоритма построения скелета можно найти в книге [1].

Алгоритм удаления черенка формулируется следующим образом:

- Строим скелетный граф для каждого листа, каждой вершине приписан радиус максимального пустого круга. См. шаг (a) на рисунке 2.
- Определяем диапазон, в который попадают радиусы всех вершин в черенке. Этот диапазон будет универсальным для всех листьев при условии одинакового масштаба всех изображений.
- Оставляем только те ребра, для которых радиус хотя бы одной вершины попадает в найденный диапазон. См. шаг (b) на рисунке 2.
- Теперь граф потерял связность. Находим нужную связную компоненту. Черенку будет отвечать наибольшая по числу вершин связная компонента. См. шаг (c) на рисунке 2.
- Удаляем черенок с фигуры листа. См. шаг (d) на рисунке 2.



**Рисунок 2** – Последовательность шагов по удалению черенка. Ребра выделены синим, максимальные пустые круги выделены красным.

## 4.2. Построение аппроксимирующего многоугольника

Современные библиотеки предлагают готовые решения, вычисляющие площадь и периметр, однако часто эти показатели вычисляются попиксельно. Попиксельное вычисление вполне может использоваться для изображений с большим разрешением. Однако, несмотря на большое разрешение сканов листьев, в данной задаче важно более точно приблизить границу, особенно в участках, требующих тщательной детализации, например, около зубцов листьев. Поэтому будем строить аппроксимирующий многоугольник.

Используем алгоритм обхода "подвижным мостом", получаем за линейное время граничный коридор. Строим минимальный разделяющий многоугольник, лежащий внутри граничного коридора, т.е. многоугольник минимальной длины среди всех замкнутых линий, лежащих внутри коридора. Физической интерпретацией построенного многоугольника может быть резиновая нить, проложенная и натянутая внутри граничного коридора. Подробное описание алгоритма построения разделяющего многоугольника можно найти в книге [1].

### 4.2.1. Вычисление площади и периметра

Для найденного минимального разделяющего многоугольника необходимо вычислить длину границы (периметр) и площадь. Длина границы вычисляется при естественном последовательном обходе по всем вершинам многоугольника. Для вычисления площади разобьем многоугольник на треугольники с общей вершиной, общую вершину выбираем из вершин многоугольника. Вновь последовательно обходим по всем вершинам многоугольника, каждый раз вычисляя площадь соответствующего прилежащего треугольника. Искомая площадь будет суммой площадей данных треугольников.

## 5. Связь признаков формы с экологической обстановкой

Оценим, существует ли связь между простейшими показателями формы листа (площадь и периметр) и экологическим состоянием района. Эксперты в предметной области часто используют устоявшиеся показатели формы. Введем определение такого показателя.

### 5.1. Форм-фактор

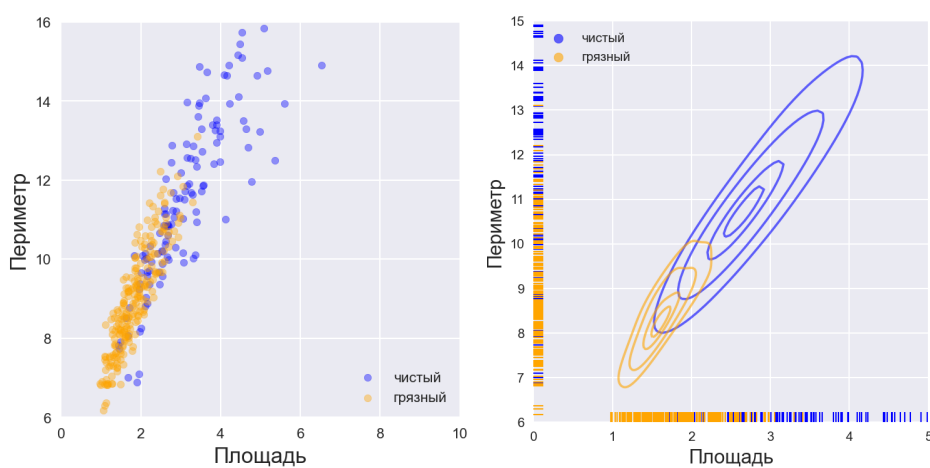
Определим форм-фактор  $F$  как нормированное отношение площади фигуры  $S$  к квадрату длины границы  $l$

$$F = 4 \pi S / l^2 \quad (1)$$

Нормировочная константа подобрана так, что для круга форм-фактор будет равен 1. В этом и заключается интуиция использования форм-фактора: косвенное сравнение, насколько близка фигура к кругу. Форм-фактор в виде (1) часто применяется для описания формы листьев в биологических и экологических исследованиях.

## 5.2. Влияние экологической обстановки

Применим построенные алгоритмы и рассчитаем признаки периметра и площади для всех листьев в двух городах. Отобразим листья точками в двумерном признаковом пространстве (см. пример для Пущино на Рисунке 3). Точки разбиваются на кластеры в зависимости от экологического состояния, есть корреляция между признаками площади и периметра – в Пущино выборочная корреляция Пирсона в чистом районе 0.89, в грязном районе 0.88. Несмотря на некоторое пересечение кластеров, существует связь между распределениями признаков и чистотой района. Чистому району отвечает кластер с большей дисперсией, средней площадью и бóльшим средним периметром.



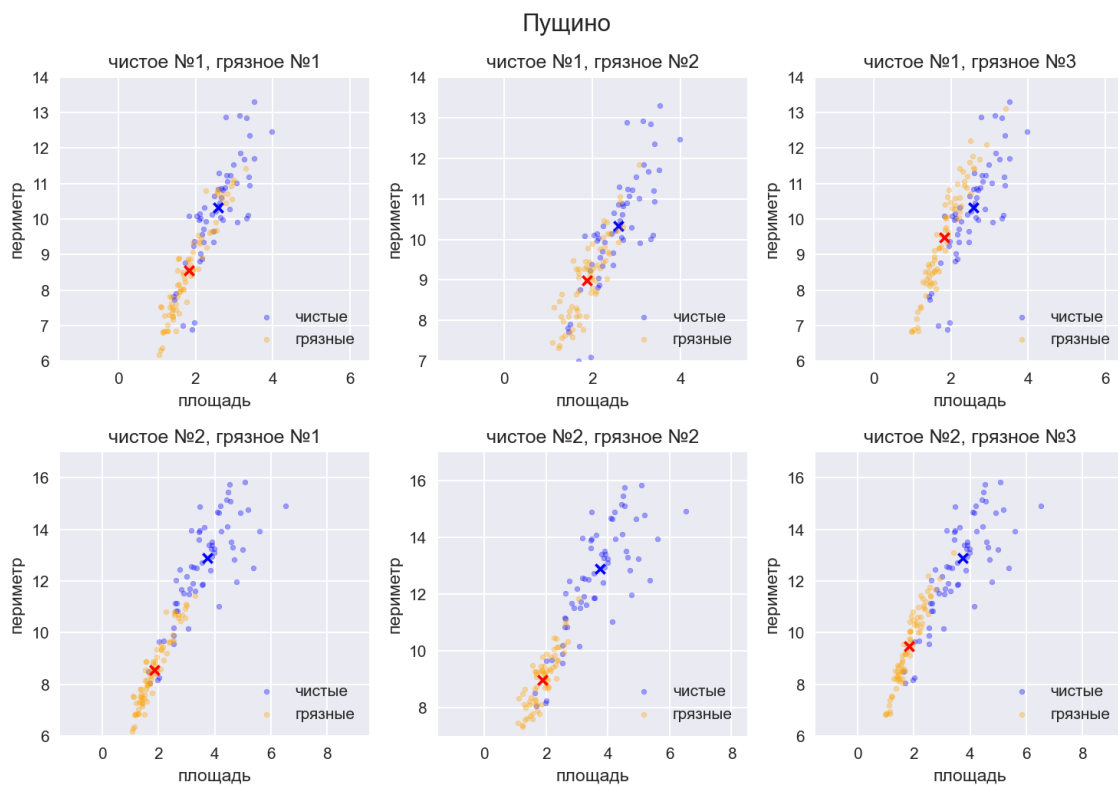
**Рисунок 3** – Визуализация распределений признаков формы в зависимости от чистоты района в Пущино. На диаграмме рассеяния (слева) точка отвечает одному листу. Синие точки соответствуют чистым районам, оранжевые – грязным.

### 5.2.1. Корректность наблюдений

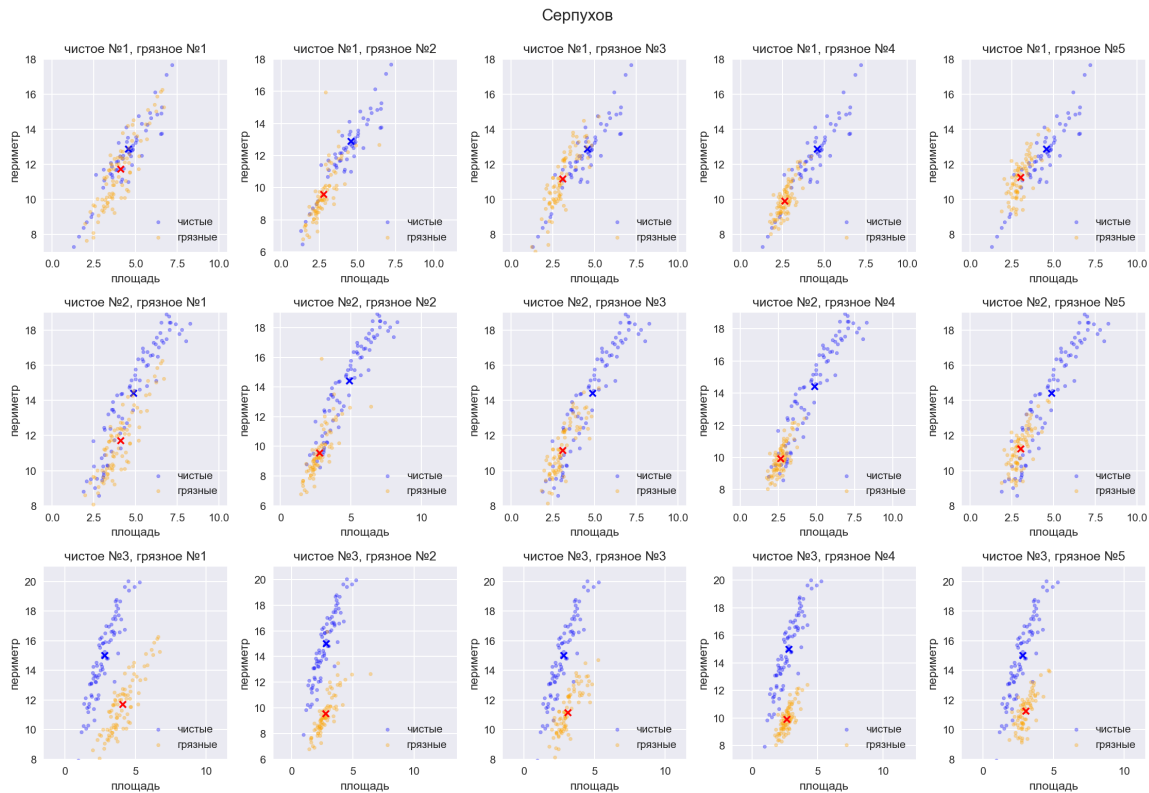
Сперва зависимость признаков формы от экологии в районе была обнаружена на изображениях листьев с 3 деревьев в Пущино. Важно, чтобы эта зависимость не оказалась следствием ложной корреляции, возможной ошибки в определении вида дерева или разницы во времени сканирования. Листья были собраны под контролем эксперта в предметной области, на территории одного города (Пущино и Серпухов соответственно), при равных температурных и погодных условиях. Кроме того, результат должен сохраняться при попарном сравнении деревьев из чистого и грязного районов. Поэтому для дополнительной проверки существования зависимости проведем попарное сравнение на 5 деревьях в Пущино и 8 деревьях в Серпухове.

В Пущино имеем попарное сравнение 2 деревьев из чистого района с 3 деревьями из грязного (Рисунок 4), в Серпухове сравнение 3 деревьев из чистого района с 6 деревьями из грязного (Рисунок 5). Во всех случаях сохраняется обнаруженная зависимость: листья разбиваются на кластеры по чистоте района, центры масс кластеров разнесены, центр масс чистого района лежит правее и выше центра масс грязного района. Под центром масс понимаем точку, координаты которой равны выборочному среднему соответствующих признаков.





**Рисунок 4** – Попарное сравнение деревьев в Пушино. Строки отвечают 2 деревьям в чистом районе, столбцы – 3 деревьям в грязном районе. Крестиком обозначен центр масс кластера.



**Рисунок 5** – Попарное сравнение деревьев в Серпухове. Строки отвечают 3 деревьям в чистом районе, столбцы – 5 деревьям в грязном районе. Крестиком обозначен центр масс кластера.

### 5.3. Качественные признаки формы

Подберем одномерный показатель, агрегирующий площадь и периметр, который при этом будет статистически значимо отличать выборки из разных по экологическому состоянию районов. Для описания удаленности центра масс от начала координат подойдут простейшие функции, например, среднее  $\frac{S+l}{2}$  или норма точки в двумерном признаковом пространстве  $\sqrt{S^2 + l^2}$ . Выберем в качестве показателя норму и сравним (Рисунок 6) распределения нормы и форм-фактора (1).

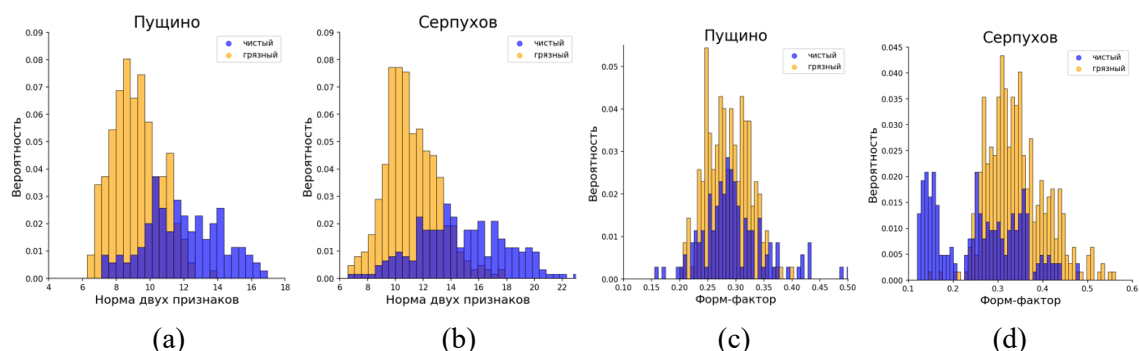


Рисунок 6 – Сравнение нормы площади и периметра с форм-фактором для Пушино и Серпухова.

Распределение нормы существенно отличается в зависимости от экологического состояния, чего не наблюдается для распределения форм-фактора (1). Оценим статистическую значимость отличия средних между выборкой листьев из чистого района и выборкой из грязного.

Сформулируем нулевую гипотезу о равенстве матожиданий нормы в двух выборках для чистого и грязного района. О равенстве дисперсий нам неизвестно, имеем выборки разного размера, поэтому будем предполагать нормальность обеих выборок и использовать t-критерий Уэлча [2]. Для Пушино (Рисунок 6а) получаем p-value  $10^{-21}$ , для Серпухова (Рисунок 6б) получаем p-value  $10^{-35}$ . Для большинства общепринятых уровней значимости  $\alpha$  отвергаем нулевую гипотезу о равенстве матожиданий в двух выборках. Другими словами, различие в средних для нормы двух признаков между экологически чистым и грязным районами обладает статистической значимостью, причем в обоих рассматриваемых городах.

Теперь сформулируем аналогичную гипотезу о равенстве форм-факторов (1) в двух различных по экологическому состоянию выборках. Предположение о нормальности будет некорректно для Серпухова, см. Рисунок 6д. Для Пушино (Рисунок 6с) получаем p-value 0.079, а значит на уровне значимости  $\alpha = 0.05$  мы не можем отвергать нулевую гипотезу. Другими словами, различие в средних для форм-фактора (1) между чистым и грязным районами в Пушино не обладает статистической значимостью. Это означает, что выборочная оценка среднего для форм-фактора (1) не позволит определить экологическое состояние района. И напротив, подсчет средней нормы для признаков площади и периметра будет корректно использовать для оценки экологического состояния.

## 6. Классификация экологической обстановки

Будем считать выполненным предположение о существовании зависимости между простейшими признаками формы и экологической обстановкой. Теперь оценим, насколько "сильны" признаки площади и периметра с точки зрения классического машинного обучения.

### 6.1. Классификация по одному листу

Работаем с малым объемом данных, поэтому существует большой риск переобучения, если будем использовать сложные модели машинного обучения. Поэтому построим простой линейный классификатор – логистическую регрессию. Будем по площади и периметру одного листа предсказывать (2) экологическое состояние района (класс), где растет дерево этого листа. Дополнительно обучим логистическую регрессию предсказывать класс района по всего одному признаку. В качестве такого признака переберем выбранные одномерные показатели – среднее (3) и норма (4) площади и периметра, а также форм-фактор (5).

$$a_1(S, l, \vec{w}) = \sigma(w_1 S + w_2 l + w_0) \quad (2)$$

$$a_2(S, l, \vec{w}) = \sigma(w_1 (S + l)/2 + w_0) \quad (3)$$

$$a_3(S, l, \vec{w}) = \sigma(w_1 \sqrt{S^2 + l^2} + w_0) \quad (4)$$

$$a_4(S, l, \vec{w}) = \sigma(w_1 4 \pi S/l^2 + w_0) \quad (5)$$

Классификатор будет возвращать оценку принадлежности к классу листьев из чистого района. Имеем 4 классификатора, веса  $\vec{w}$  для каждого подбираем при обучении. Обучим классификаторы на данных из Пущино. Для корректного оценивания листа в обучении и тесте должны быть с разных деревьев. В обучении будет 138 листьев с 1 чистого и 1 грязного дерева. В тесте будет 211 листьев с 1 чистого и 2 грязных деревьев, множества деревьев между обучением и тестом не пересекаются.

В качестве оценки качества классификации используем метрику ROC AUC на тестовой выборке. Величину этой метрики можно интерпретировать как вероятность того, что оценка алгоритма для случайно выбранного листа из чистого района будет больше оценки алгоритма для случайно выбранного листа из грязного района.

Таблица 1 – ROC AUC на тестовой выборке

Оба признака $a_1$	Среднее $a_2$	Норма $a_3$	Форм-фактор $a_4$
0.938	0.928	0.916	0.42

Результаты приведены в Таблице 1. Пару признаков площади и периметра можно считать достаточно сильной, классификатор дает на этой паре больше 0.93 ROC AUC. Отметим, что одномерная агрегация через среднее и норму сохраняет потенциал к качественной классификации, модель показывает высокое качество даже на одном таком признаке. В свою очередь классификатор на форм-факторе (1) показывает очень слабую разделяющую способность.

Отдельно обратим внимание, что удаление черенка повышает метрику в среднем на 0.03 пункта ROC AUC.

## 6.2. Классификация по выборке листьев

Добьемся наилучшего качества классификации. На практике кажется более разумным оценивать экологическую обстановку не по одному листу, а по выборке листьев.

Пусть  $X_c^i$  – выборка норм площади и периметра для листьев с дерева  $i$  в чистом районе, а  $X_d^j$  – выборка норм для листьев с дерева  $j$  в грязном районе. В геометрической интерпретации выборка норм – это набор радиусов окружностей, на которых лежат точки в признаковом пространстве. Найдем порог, однозначно разделяющий любые два кластера из противоположных классов

$$t = (\min_i \bar{X}_c^i + \max_j \bar{X}_d^j) / 2 \quad (6)$$

Порогу (6) отвечает окружность, разделяющая два ближайших кластера из противоположных классов. Класс выборки норм для листьев  $X$  будет определяться положением кластера относительно этой окружности

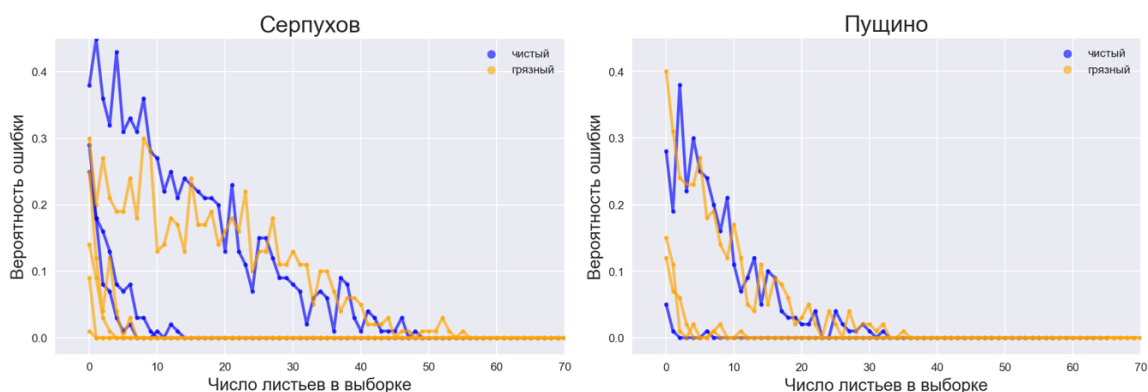
$$a_5(X, t) = I[\bar{X} > t] \quad (7)$$

Этап обучения в такой постановке заключается в нахождении порога (6) по выборке. Такой классификатор (7) достигает 100% точности при классификации любого дерева в Пушино и Серпухове по всем предоставленным листьям.

### 6.2.1. Оценка числа листьев для надежности классификации

Построенный классификатор идеально разделяет кластеры деревьев по чистоте района, но это выполняется при классификации на всех доступных листьях с дерева. Оценим, сколько нужно на практике собрать листьев, чтобы быть уверенным в прогнозе на заданном уровне вероятности ошибки.

Будем использовать бутстрэп [3] для оценки вероятности ошибки для каждого дерева. Фиксируем дерево, последовательно перебираем число листьев  $k$  в выборке, для каждого  $k$  повторяем  $n$  раз сэмплирование, при этом на каждой сэмплированной подвыборке проверяем качество классификации. Оцениваем вероятность ошибки для данного числа  $k$  как долю ошибочных классификаций среди  $n$  испытаний. Кривые вероятностей в зависимости от числа листьев для каждого дерева отложим на Рисунке 7. Заметим, что для некоторых деревьев вероятность ошибки выше при прочих равных, это можно объяснить неоднородностью в экологическом состоянии районов, эти кривые отвечают ближайшим двум кластерам из противоположных классов.



**Рисунок 7** – Бутстрэп оценка вероятности ошибки при классификации по выборке листьев ограниченного объема. Каждая кривая соответствует отдельному дереву.

Проинтерпретируем полученные кривые: если мы хотим иметь вероятность ошибки для всех деревьев ниже 0.05, то в Серпухове будет достаточно собрать с дерева 40 листьев, а в Пущино 20 листьев.

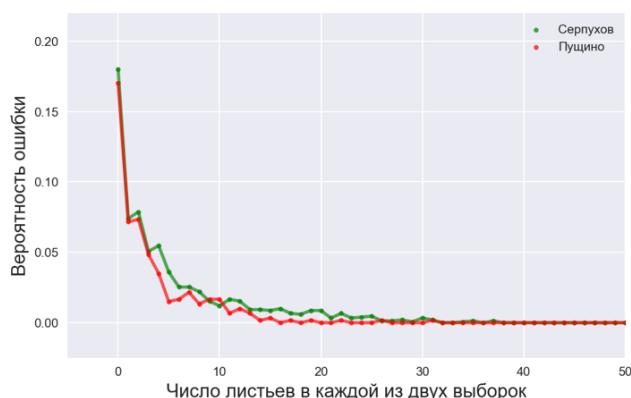
### 6.3. Состязательная классификация двух районов

Справедливо будет заметить, что 100% точность классификации по выборке является оптимистичной оценкой качества подхода, так как фактически мы оцениваем классификатор на тех же данных, на которых он обучался. Подход состязательной классификации лишен такой проблемы.

Будем классифицировать состояние не в одном регионе, а сравнивать состояния в двух регионах. Здесь не будет обучаемых параметров, решение будем принимать, сравнивая средние нормы площадей и периметров

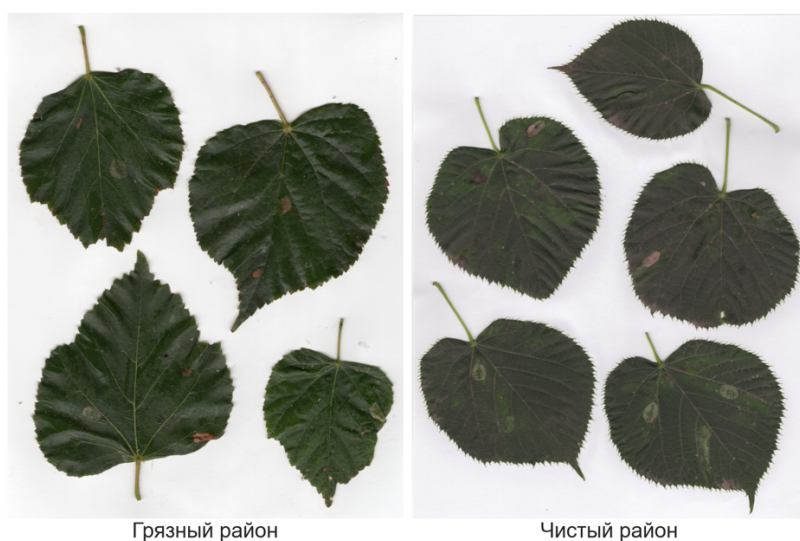
$$a_6(X_c, X_d) = I[\overline{X_c} > \overline{X_d}] \quad (8)$$

Классификацию (8) будем называть состязательной. Вновь оценим вероятность ошибки в такого рода классификации. Ошибкой будем считать неверное определение классов на хотя бы одной паре деревьев, где одно дерево в паре будет из чистого района, а другое – из грязного. Аналогично, используем бутстрэп [3], последовательно перебираем число листьев  $k$ , сэмплируем  $k$  листьев для каждого дерева, сравниваем все возможные пары деревьев разных классов по тому же принципу, что на Рисунках 4 и 5. Для каждого  $k$  повторяем процедуру  $n$  раз, оцениваем вероятность ошибки как долю ошибок среди  $n$  испытаний. Визуализируем кривые на Рисунке 8. Для требования вероятности ошибки ниже 0.05 достаточно будет собрать по 10 листьев в каждом районе.



**Рисунок 8** – Бутстрэп оценка вероятности ошибки при состязательной классификации по выборке листьев ограниченного объема.

#### 6.4. Интерпретация результата



**Рисунок 9** – Сравнение примеров листьев из разных по экологическому состоянию районов. Отличия в основании у черенка и в выпуклости зубчиков.

Было показано, что форма листа зависит от экологической обстановки в районе, где растет дерево, причем эта зависимость видна даже на примитивных признаках формы, таких как площадь и периметр. С точки зрения прикладной области это можно объяснить следующими наблюдениями:

- В чистых районах чаще вырастают крупные листья.
- В чистых районах основания у черенка более выпуклые и закругленные, а в грязных более плоские (см. Рисунок 9).
- В чистых районах чаще встречаются заостренные зубчики (см. Рисунок 9).

## 7. Выводы

Листья липы могут служить качественными маркерами экологического состояния среды. Форма листьев зависит от экологии в районе, где растет дерево, причем это наблюдается и для

простых признаков формы, например, для периметра и площади. Черенок вносит шумовую компоненту при вычислении площади и периметра, его удаление приводит к более качественным результатам. Принятый в обществе экологов форм-фактор не может статистически значимо определить состояние в районе, тем не менее это возможно для достаточно простых функций, например, для среднего или нормы. Можно использовать признаки площади и периметра для построения классификаторов состояния окружающей среды по собранным листьям. Например, простейший линейный классификатор достигает выше 0.93 ROC AUC при предсказании состояния среды по одному листу. Аналогично, можно оценивать состояние среды не по одному листу, а по выборке. Такие подходы достигают высокого качества, причем можно оценить надежность их прогноза в зависимости от числа собранных листьев.

Подход с оцениванием состояния окружающей среды по признакам формы листьев выглядит перспективным, его можно применять без привлечения экспертов в области биологии и экологии, он не требует ресурсоемких и сложных расчетов, а самое главное – для его применения достаточно лишь собрать несколько листьев, отсканировать и применить алгоритм.

## **8. Список источников**

- [1] Местецкий Л.М. Непрерывная морфология бинарных изображений. // Москва, Физматлит, 2009.
- [2] Welch B.L. The generalization of "Student's" problem when several different population variances are involved. // *Biometrika*. 1947. № 34 (1–2). С. 28–35.
- [3] Efron B. Bootstrap methods: Another look at the jackknife. // *The Annals of Statistics*. 1979. №7 (1). С. 1–26.