



Министерство науки и высшего образования Российской Федерации
Федеральное государственное бюджетное образовательное учреждение
высшего образования
«Московский государственный технический университет
имени Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)

ФАКУЛЬТЕТ ИУ «Информатика и системы управления»

КАФЕДРА ИУ-7 «Программное обеспечение ЭВМ и информационные технологии»

**РАСЧЕТНО-ПОЯСНИТЕЛЬНАЯ ЗАПИСКА
К НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ
НА ТЕМУ:**

***«Классификация методов повышения разрешающей
способности изображений»***

Студент ИУ7-73Б _____ Светличная А. А.

Руководитель НИР _____ Филиппов М. В.

Рекомендуемая руководителем НИР оценка _____

РЕФЕРАТ

Научно-исследовательская работа содержит 31 страницу, 15 рисунков, 3 таблицы, 29 формул, 14 источников, 1 приложение.

ЦИФРОВАЯ ОБРАБОТКА ИЗОБРАЖЕНИЙ, РАЗРЕШАЮЩАЯ СПОСОБНОСТЬ ИЗОБРАЖЕНИЙ, ИНТЕРПОЛЯЦИЯ, НЕЙРОННЫЕ СЕТИ

Цель — провести обзор методов повышения разрешающей способности изображений.

В процессе работы проводилась классификация существующих методов повышения разрешающей способности изображений, а также сравнительный анализ данных методов по сформулированным критериям оценки.

Результаты — нейронные методы показывают визуально более качественные результаты, однако являются ресурсозатратными (вычислительные, человеческие, по памяти и т. д.), в то время как классические методы демонстрируют приемлемое для отдельных задач качество и являются в разы более легкими. Из названных типов методов можно выделить как наиболее хорошо показавших себя методы бикубической интерполяции и Ланцоша из классических и сверточные нейронные сети из соответственно нейронных методов.

СОДЕРЖАНИЕ

РЕФЕРАТ	3
ВВЕДЕНИЕ	5
1 Анализ предметной области	6
1.1 Основные направления цифровой обработки изображений . . .	6
1.2 Основные определения	6
2 Классификация методов повышения разрешающей способности	11
2.1 Классические методы	11
2.1.1 Метод ближайшего соседа	12
2.1.2 Билинейная интерполяция	13
2.1.3 Бикубическая интерполяция	15
2.1.4 Фильтр Ланцоша	20
2.2 Нейронные методы	22
2.2.1 Сверточные нейронные сети	22
2.3 Генеративно-состязательные нейронные сети	24
2.4 Сравнение методов повышения разрешающей способности изображения	26
ЗАКЛЮЧЕНИЕ	30
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	31
ПРИЛОЖЕНИЕ А Презентация к работе	33

ВВЕДЕНИЕ

Цифровая обработка изображений — это востребованная и быстроразвивающаяся область, результаты которой применяются в медицине, науке, промышленности, торговле и многих других отраслях. Например, в 2021 году приложения для обработки изображений были самыми загружаемыми в магазинах приложений, таких как App Store и Google Play [1]. А это лишь малая часть применения цифровой обработки изображений.

В цифровой обработке изображений огромное количество направлений, посвященных как улучшению, так и ухудшению качества изображения, а также анализу изображений на наличие тех или иных признаков. Для данной работы была выбрана такая характеристика, как разрешение изображения, а именно повышение данной характеристики. Возможность увеличения или уменьшения разрешения изображения встречается не только в приложениях, непосредственно посвященных обработке изображений, но даже и в тех, которые, казалось бы, не имеют отношения к данной сфере, а значит, тема является актуальной.

Цель: классифицировать методы повышения разрешающей способности изображений.

Для достижения поставленной цели необходимо решить ряд задач.

Задачи:

- провести анализ предметной области;
- описать существующие методы повышения разрешения изображений;
- сформулировать критерии сравнения методов повышения разрешающей способности изображений;
- сравнить описанные методы на основании выделенных критериев;
- описать результаты сравнения рассмотренных методов.

1 Анализ предметной области

1.1 Основные направления цифровой обработки изображений

Цифровая обработка изображений предоставляет инструменты и методы для анализа, улучшения и интерпретации изображений и видео, полученных с помощью цифровых устройств. Эту область трудно описать полностью, однако очень легко привести некоторые примеры цифровой обработки изображений:

- фильтрация изображений: уменьшение резкости, усиление краев, устранение шумов медианным замещением, сглаживание с использованием функции Гаусса;
- бинаризация: перевод в черно-белый формат, разделение изображения на два класса (черный и белый);
- улучшение контраста: расширение диапазона интенсивности, выравнивание гистограммы;
- морфологические операции: изменение размера объектов;
- сегментация изображений: разделение изображения на регионы, группировка пикселей схожих характеристик;
- изменение разрешения: увеличение или уменьшение разрешения изображений.

В данной работе будут рассмотрены методы последнего пункта представленного списка, а именно методы повышения разрешающей способности изображения.

1.2 Основные определения

Разрешение изображения — количество пикселей (точек) на единицу площади изображения [2].

Цифровые методы повышения разрешающей способности можно разделить на классические и нейронные.

Классические методы используют интерполяцию, и их можно разделить на нелинейные адаптивные и линейные неадаптивные [3].

Адаптивные методы учитывают локальную структуру и содержание изображения при принятии решений о том, как интерполировать значения пикселей. Они могут анализировать окружающие пиксели и динамически выбирать метод интерполяции в зависимости от характеристик окружающих данных [3].

Неадаптивные методы не учитывают контекст или структуру изображения, а применяют один и тот же метод интерполяции ко всем пикселям [3].

Интерполяция — процесс определения значений между известными точками данных [4].

На рисунке 1.1 показан оригинал изображения в низком разрешении, использованного для интерполяции, а на 1.2 и 1.3 — изображения одного разрешения, но до и после интерполяции соответственно.

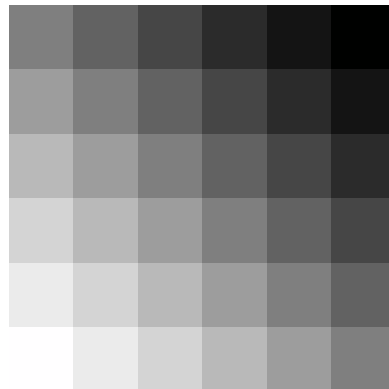


Рисунок 1.1 – Оригинал изображения

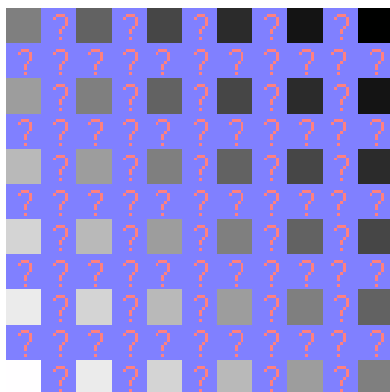


Рисунок 1.2 – До интерполяции

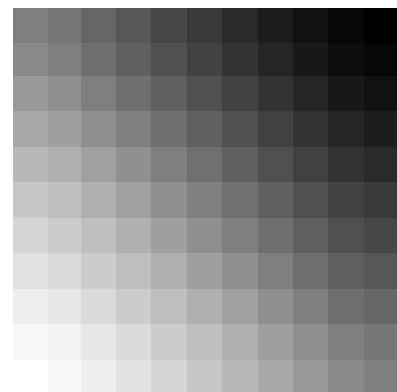


Рисунок 1.3 – После интерполяции

Однако в результате интерполяции могут возникнуть следующие де-

фекты: ступенчатость, размытие и граничные гало [5]. Пример данных дефектов показан на рисунке 1.4.

Ступенчатость — видимые лестничные ступеньки или резкие перепады яркости вдоль контуров объектов [3].

Размытие — потеря деталей и четкости изображения при сглаживании краев [3].

Граничное гало (иногда данное явление называют эффектом Гиббса) — светлые или темные области, окружающие объекты или контуры на изображении [3].

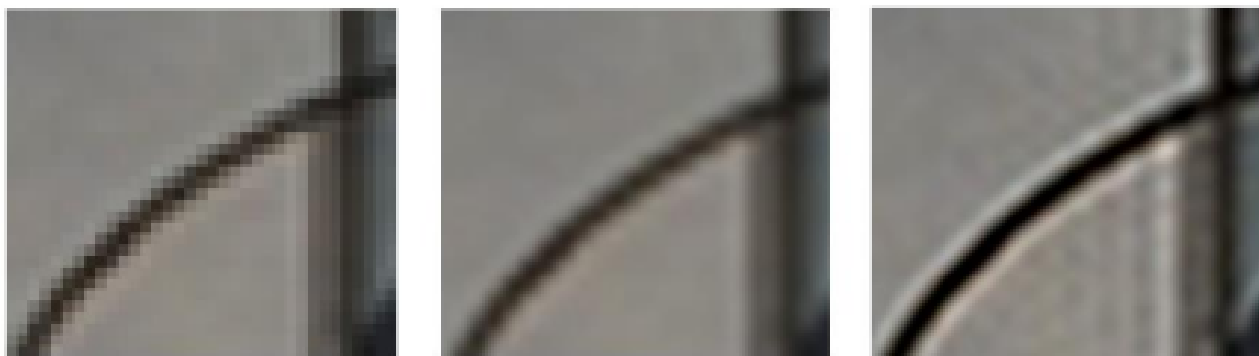


Рисунок 1.4 – Пример дефектов неадаптивных методов(слева направо ступенчатость, размытие, граничные гало)

Нейронные методы используют искусственные нейронные сети, которые принадлежат сфере машинного обучения.

Машинное обучение — направление искусственного интеллекта, сосредоточенное на создании систем, которые обучаются и развиваются на основе получаемых ими данных [6].

Машинное обучение без учителя (неконтролируемое обучение) — метод машинного обучения, при котором модель обучается выявлять закономерности и скрытые взаимосвязи на наборах неразмеченных данных без контроля со стороны пользователя. При неконтролируемом обучении алгоритму не сообщается конечная цель или шаблоны, а только предоставляются массивы данных, общие признаки распознаются автоматически [6].

Машинное обучение с учителем — метод машинного обучения построения моделей на основе множества примеров, содержащих пары «известный вход – известный выход». Иными словами, алгоритм должен работать с примерами, которые содержат не только вектор независимых переменных (ат-

рибутов, признаков), но и значение, которое должна выдавать модель после обучения [7].

Искусственные нейронные сети — системы, преобразующие информацию подобно процессам, происходящим в человеческом мозге [6].

Одной из главных функциональных особенностей нейронных сетей является способность к обучению. Обучение нейронных сетей заключается в итерационном процессе подбора весовых коэффициентов, обеспечивающих ее требуемую реакцию сети на входные сигналы [8].

Структурной единицей нейронной сети является нейрон представляющий из себя аналог биологического нейрона. С точки зрения математики искусственный нейрон это нелинейная функция, обрабатывающая все входные сигналы и посылающая результат на единственный выход нейрона. Эта нелинейная функция называется активационной функцией или функцией активации [8].

Искусственный нейрон включает в себя группу синапсов, которые представляют собой однонаправленные связи, соединенные с предыдущими нейронами. Каждый синапс можно охарактеризовать с помощью весов W . Это аналог электрической проводимости биологического нейрона. Для того чтобы параллельно обрабатывать сигналы нейроны объединяются в слои в зависимости от их функционирования. Первый слой называется входным. Нейроны этого слоя принимают входные данные и передают их следующему слою. Последний слой называется выходным, т.к. в качестве выходных данных он выдает результат работы нейронной сети. Слои, находящиеся между входным и выходным слоем, называются скрытыми или внутренними. В этих слоях происходит большая часть обработки данных. Когда выходные сигналы нейронов одного слоя передаются всем нейронам следующего слоя, слой, получающий эти сигналы, называется полносвязным [8].

При обучении для определения корректности результата работы применяется функция потерь. В задаче обучения с учителем функция потерь показывает разницу между реальными данными и данными полученными в результате работы нейронной сети [8].

Наиболее используемыми методами повышения разрешающей способности изображения с использованием нейронных сетей являются сверточные и генеративно-состязательные нейронные сети.

Вывод

В данной главе была проанализирована область цифровой обработки изображений, а также введены все термины и понятия, необходимые для дальнейшего изучения методов повышения разрешения изображений.

2 Классификация методов повышения разрешающей способности

2.1 Классические методы

Адаптивные нелинейные методы не получили большого распространения в силу их узкой направленности, однако в лицензированных программах иногда встречаются коммерческие алгоритмы на основе адаптивных методов. По этим причинам далее эти методы рассмотрены не будут.

Известны следующие неадаптивные методы повышения разрешения изображений:

- метод ближайшего соседа;
- билинейный;
- бикубический;
- метод Ланцоша (фильтр Ланцоша) [5].

На рисунках 2.1, 2.2, 2.3, 2.5 показаны примеры работы данных методов соответственно списку.

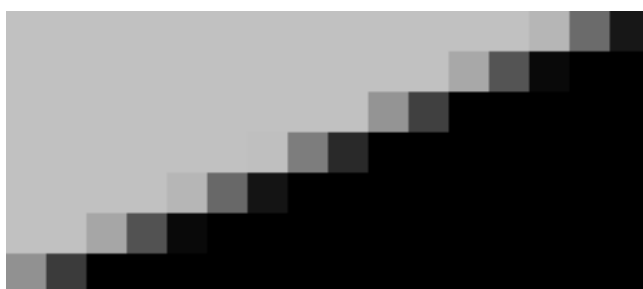


Рисунок 2.1 – Пример работы метода ближайшего соседа



Рисунок 2.2 – Пример работы билинейного метода



Рисунок 2.3 – Пример работы бикубического метода



Рисунок 2.4 – Пример работы метода Ланцоша

2.1.1 Метод ближайшего соседа

Наиболее базовый из всех алгоритмов интерполяции, который требует наименьшего времени обработки, поскольку учитывает только один пиксель — ближайший к точке интерполяции [4].

$$\begin{aligned}
 C_1[i][j] &= C_0[k_1 \cdot i][k_2 \cdot j], \\
 i &= 0..(H_0 - 1), \\
 j &= 0..(W_0 - 1), \\
 k_1 &= \frac{H_0}{H_1}, \\
 k_2 &= \frac{W_0}{W_1},
 \end{aligned} \tag{2.1}$$

где

- H_0 и W_0 — высота и ширина исходного изображения соответственно;
- H_1 и W_1 — высота и ширина нового изображения;
- k_1 и k_2 — масштабные коэффициенты.

2.1.2 Билинейная интерполяция

Метод рассматривает квадрат 2×2 известных пикселей, окружающих неизвестный. В качестве интерполированного значения используется взвешенное усреднение этих четырех пикселей. В результате изображения выглядят значительно более гладко, чем результат работы метода ближайшего соседа [4].



Рисунок 2.5 – Диаграмма билинейной интерполяции

В вычислительной математике *билинейная интерполяция* — расширение линейной интерполяции для функций в двух переменных, то есть это линейная интерполяция первого порядка в двумерном случае. Ключевая идея заключается в том, чтобы провести обычную линейную интерполяцию сначала в одном направлении, затем в другом [9].

Рассмотрим пример, график которого изображен на рисунке 2.6. Четыре красные точки представляют собой известные значения функции. Значение в зеленой точке должно быть интерполировано.

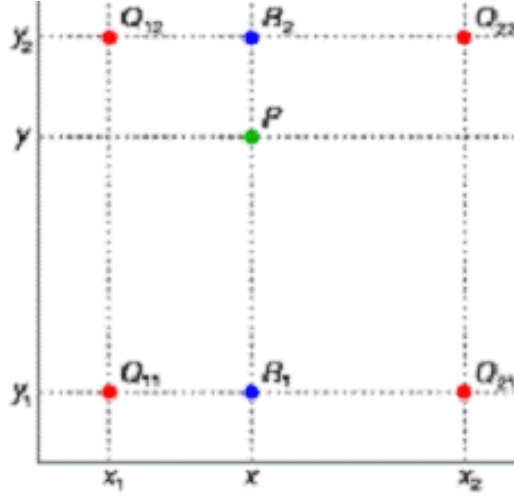


Рисунок 2.6 – Пример билинейной интерполяции

Допустим, что необходимо интерполировать значение функции f в точке $P = (x, y)$. Для этого необходимо знать значения функций в (окружающих P) точках $Q_{11} = (x_1, y_1)$, $Q_{12} = (x_1, y_2)$, $Q_{21} = (x_2, y_1)$, $Q_{22} = (x_2, y_2)$. Первым шагом интерполируется (линейно) значение вспомогательных точек R_1 и R_2 вдоль оси абсцисс, где $R_1 = (x, y_1)$, $R_2 = (x, y_2)$ [9].

$$f(R_1) \approx \frac{x_2 - x}{x_2 - x_1} f(Q_{11}) + \frac{x - x_1}{x_2 - x_1} f(Q_{21}). \quad (2.2)$$

$$f(R_2) \approx \frac{x_2 - x}{x_2 - x_1} f(Q_{12}) + \frac{x - x_1}{x_2 - x_1} f(Q_{22}). \quad (2.3)$$

$$f(P) \approx \frac{y_2 - y}{y_2 - y_1} f(R_1) + \frac{y - y_1}{y_2 - y_1} f(R_2). \quad (2.4)$$

Это и есть приблизительное значение функции в точке $P = f(x, y)$.

$$\begin{aligned} f(x, y) = & \frac{f(Q_{11})}{(x_2 - x_1)(y_2 - y_1)}(x_2 - x)(y_2 - y) \\ & + \frac{f(Q_{21})}{(x_2 - x_1)(y_2 - y_1)}(x - x_1)(y_2 - y) \\ & + \frac{f(Q_{12})}{(x_2 - x_1)(y_2 - y_1)}(x_2 - x)(y - y_1) \\ & + \frac{f(Q_{22})}{(x_2 - x_1)(y_2 - y_1)}(x - x_1)(y - y_1). \end{aligned} \quad (2.5)$$

В особом случае, когда известные точки находятся на вершинах единичного квадрата, т. е. имеют координаты $(0, 0)$, $(0, 1)$, $(1, 0)$, и $(1, 1)$, формула

билинейной интерполяции упрощается до:

$$f(x, y) = (1 - x) \cdot (1 - y) \cdot f(0, 0) + x \cdot (1 - y) \cdot f(1, 0) + (1 - x) \cdot y \cdot f(0, 1) + x \cdot y \cdot f(1, 1). \quad (2.6)$$

Или же с помощью умножения векторов с матрицей:

$$f(x, y) \approx \begin{bmatrix} 1 - x & x \end{bmatrix} \begin{bmatrix} f(0, 0) & f(0, 1) \\ f(1, 0) & f(1, 1) \end{bmatrix} \begin{bmatrix} 1 - y \\ y \end{bmatrix}. \quad (2.7)$$

Результат билинейной интерполяции не зависит от порядка шагов. Возможно сначала интерполировать между известными точками вдоль оси ординат и затем, получив два вспомогательных значения, интерполировать между ними вдоль оси абсцисс [9].

2.1.3 Бикубическая интерполяция

Бикубическая интерполяция идет на один шаг дальше билинейной, рассматривая массив из 4x4 окружающих пикселей — всего 16. Поскольку они находятся на разных расстояниях от неизвестного пикселя, ближайшие пиксели получают при расчете больший вес [4].

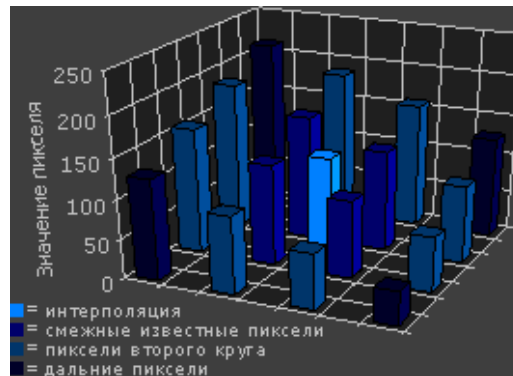


Рисунок 2.7 – Диаграмма бикубической интерполяции

В вычислительной математике бикубическая интерполяция — расширение кубической интерполяции на случай функции двух переменных, значения которой заданы на двумерной регулярной сетке. Поверхность, полученная в результате бикубической интерполяции, является гладкой функцией, в отличие от поверхностей, полученных в результате билинейной интерполяции или интерполяции методом ближайшего соседа [10].

Чтобы лучше понять принцип бикубической интерполяции, необходимо рассмотреть принцип кубической интерполяции.

Если значения функции $f(x)$ и ее производной известны в точках $x = 0$ и $x = 1$, тогда функция может быть интерполирована на интервале $[0, 1]$ используя полином третьего порядка. Полином третьего порядка и его производная:

$$f(x) = ax^3 + bx^2 + cx + d, \quad (2.8)$$

$$f'(x) = 3ax^2 + 2bx + c. \quad (2.9)$$

Значения полинома и его производной в точках $x = 0$ и $x = 1$:

$$\begin{aligned} f(0) &= d, \\ f'(0) &= c, \\ f(1) &= a + b + c + d, \\ f'(1) &= 3a + 2b + c. \end{aligned} \quad (2.10)$$

Эти четыре тождества могут быть записаны как:

$$\begin{aligned} a &= 2f(0) - 2f(1) + f'(0) + f'(1), \\ b &= -3f(0) + 3f(1) - 2f'(0) - f'(1), \\ c &= f'(0), \\ d &= f(0). \end{aligned} \quad (2.11)$$

Итак, интерполяционная формула получена, однако на практике алгоритм используется для интерполяции функции, имея некие известные значения в заданных точках. В данном случае невозможно знать производную функции. Принятие производной в заданных точках, как 0, могло бы быть вариантом. Но для получения более гладких и правдоподобных графиков функций принимается за производную уклон линии между предыдущей и следующей точкой. Таким образом, для расчетов необходимо использовать 4 точки. Предположим, есть 4 значения функции в точках p_0 , p_1 , p_2 , p_3 , расположенных соответственно на $x = -1$, $x = 0$, $x = 1$, $x = 2$.

Подставив полученные значения $f(0)$, $f(1)$, $f(2)$, $f(3)$, получается:

$$\begin{aligned} f(0) &= p_1, \\ f'(0) &= \frac{p_2 - p_0}{2}, \\ f(1) &= p_2, \\ f'(1) &= \frac{p_3 - p_1}{2}. \end{aligned} \tag{2.12}$$

После сопоставления этих данных с полученными ранее формулами имеются:

$$\begin{aligned} a &= -\frac{1}{2}p_0 + \frac{3}{2}p_1 - \frac{3}{2}p_2 + \frac{1}{2}p_3, \\ b &= p_0 - \frac{5}{2}p_1 + 2p_2 - \frac{1}{2}p_3, \\ c &= -\frac{1}{2}p_0 + \frac{1}{2}p_2, \\ d &= p_1. \end{aligned} \tag{2.13}$$

Результат:

$$\begin{aligned} f(x, a, b, c, d) &= \left(-\frac{1}{2}a + \frac{3}{2}b - \frac{3}{2}c + \frac{1}{2}d\right)x^3 \\ &+ \left(a - \frac{5}{2}b + 2c - \frac{1}{2}d\right)x^2 \\ &+ \left(-\frac{1}{2}a + \frac{1}{2}c\right)x + b. \end{aligned} \tag{2.14}$$

Бикубическая интерполяция представляет собой кубическую интерполяцию в двух измерениях.

Пусть известно 16 точек p_{ij} , с точкой начала координат в $(i - 1, j - 1)$, где i, j изменяются от 0 до 3. В этом случае возможно интерполировать поверхность на участке $[0, 0] \times [1, 1]$, для этого интерполируют 4 колонки и потом интерполируют полученные результаты в горизонтальном направлении:

$$\begin{aligned} g(x, y) &= f(x, f(y, p_{00}, p_{01}, p_{02}, p_{03}), f(y, p_{10}, p_{11}, p_{12}, p_{13}), \\ &f(y, p_{20}, p_{21}, p_{22}, p_{23}), f(y, p_{30}, p_{31}, p_{32}, p_{33})). \end{aligned} \tag{2.15}$$

Можно переписать формулу для бикубической интерполяции как мно-

гомерный полином:

$$g(x, y) = \sum_{i=0}^3 \sum_{j=0}^3 a_{ij} x^i y^j. \quad (2.16)$$

Теперь необходимо вычислить коэффициенты a_{ij} :

$$\begin{aligned}
a_{00} &= p_{11}, \\
a_{01} &= -\frac{1}{2}p_{10} + \frac{1}{2}p_{12}, \\
a_{02} &= p_{10} - \frac{5}{2}p_{11} + 2p_{12} - \frac{1}{2}p_{13}, \\
a_{03} &= -\frac{1}{2}p_{10} + \frac{3}{2}p_{11} - \frac{3}{2}p_{12} + \frac{1}{2}p_{13}, \\
a_{10} &= -\frac{1}{2}p_{01} + \frac{1}{2}p_{21}, \\
a_{11} &= \frac{1}{4}p_{00} - \frac{1}{4}p_{02} - \frac{1}{4}p_{20} + \frac{1}{4}p_{22}, \\
a_{12} &= -\frac{1}{2}p_{00} + \frac{5}{4}p_{01} - p_{02} + \frac{1}{4}p_{03} + \frac{1}{2}p_{20} - \frac{5}{4}p_{21} + p_{22} - \frac{1}{4}p_{23}, \\
a_{13} &= \frac{1}{4}p_{00} - \frac{3}{4}p_{01} + \frac{3}{4}p_{02} - \frac{1}{4}p_{03} - \frac{1}{4}p_{20} + \frac{3}{4}p_{21} - \frac{3}{4}p_{22} + \frac{1}{4}p_{23}, \\
a_{20} &= p_{01} - \frac{5}{2}p_{11} + 2p_{21} - \frac{1}{2}p_{31}, \\
a_{21} &= -\frac{1}{2}p_{00} + \frac{1}{2}p_{02} + \frac{5}{4}p_{10} - \frac{5}{4}p_{12} - p_{20} + p_{22} + \frac{1}{4}p_{30} - \frac{1}{4}p_{32}, \\
a_{22} &= p_{00} - \frac{5}{2}p_{01} + 2p_{02} - \frac{1}{2}p_{03} - \frac{5}{2}p_{10} + \frac{25}{4}p_{11} - 5p_{12} + \frac{5}{4}p_{13} + 2p_{20} \\
&\quad - 5p_{21} + 4p_{22} - p_{23} - \frac{1}{2}p_{30} + \frac{5}{4}p_{31} - p_{32} + \frac{1}{4}p_{33}, \\
a_{23} &= -\frac{1}{2}p_{00} + \frac{3}{2}p_{01} - \frac{3}{2}p_{02} + \frac{1}{2}p_{03} + \frac{5}{4}p_{10} - \frac{15}{4}p_{11} + \frac{15}{4}p_{12} - \frac{5}{4}p_{13} - p_{20} \\
&\quad + 3p_{21} - 3p_{22} + p_{23} + \frac{1}{4}p_{30} - \frac{3}{4}p_{31} + \frac{3}{4}p_{32} - \frac{1}{4}p_{33}, \\
a_{30} &= -\frac{1}{2}p_{01} + \frac{3}{2}p_{11} - \frac{3}{2}p_{21} + \frac{1}{2}p_{31}, \\
a_{31} &= \frac{1}{4}p_{00} - \frac{1}{4}p_{02} - \frac{3}{4}p_{10} + \frac{3}{4}p_{12} + \frac{3}{4}p_{20} - \frac{3}{4}p_{22} - \frac{1}{4}p_{30} + \frac{1}{4}p_{32}, \\
a_{32} &= -\frac{1}{2}p_{00} + \frac{5}{4}p_{01} - p_{02} + \frac{1}{4}p_{03} + \frac{3}{2}p_{10} - \frac{15}{4}p_{11} + 3p_{12} - \frac{3}{4}p_{13} - \frac{3}{2}p_{20} \\
&\quad + \frac{15}{4}p_{21} - 3p_{22} + \frac{3}{4}p_{23} + \frac{1}{2}p_{30} - \frac{5}{4}p_{31} + p_{32} - \frac{1}{4}p_{33}, \\
a_{33} &= \frac{1}{4}p_{00} - \frac{3}{4}p_{01} + \frac{3}{4}p_{02} - \frac{1}{4}p_{03} - \frac{3}{4}p_{10} + \frac{9}{4}p_{11} - \frac{9}{4}p_{12} + \frac{3}{4}p_{13} + \frac{3}{4}p_{20} \\
&\quad - \frac{9}{4}p_{21} + \frac{9}{4}p_{22} - \frac{3}{4}p_{23} - \frac{1}{4}p_{30} + \frac{3}{4}p_{31} - \frac{3}{4}p_{32} + \frac{1}{4}p_{33}.
\end{aligned} \tag{2.17}$$

2.1.4 Фильтр Ланцоша

Идея фильтра Ланцоша основана на применении нормированной функции $\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$ (график функции изображен на рисунке 2.8) с растянутым по оси x главным лепестком и равной нулю вне заданного параметром ширины a интервала [11].

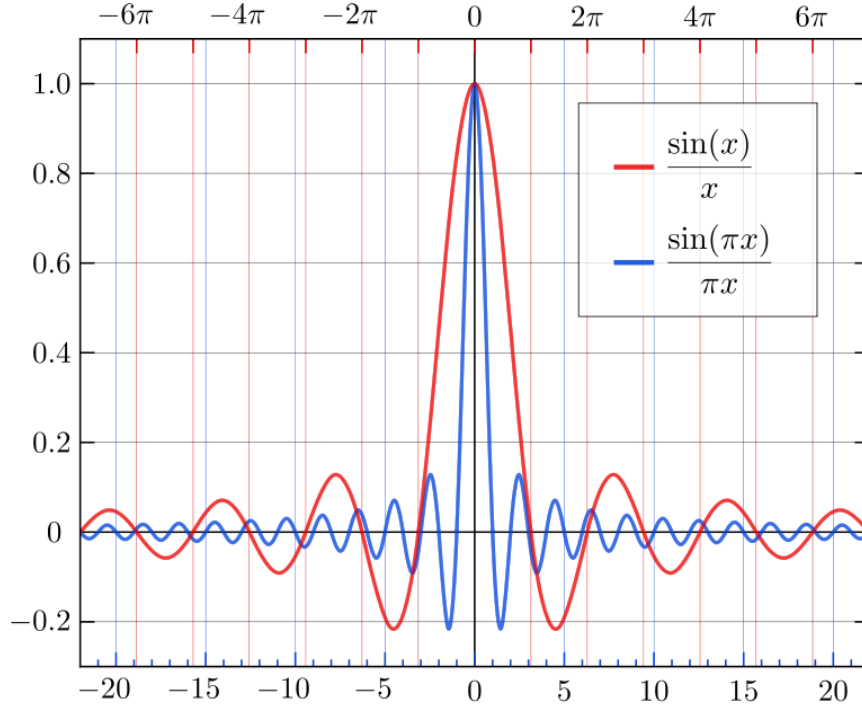


Рисунок 2.8 – Графики нормированной и ненормированной функций $\text{sinc}(x)$ в диапазоне $-7\pi \leq x \leq 7\pi$

Метод связан с оконной функцией Ланцоша $L_w(x)$, представляющей собой главный лепесток функции $\text{sinc}(x)$, вне этого лепестка оконная функция равна нулю:

$$L_w(x) = \text{sinc}\left(\frac{x}{a}\right). \quad (2.18)$$

Отфильтрованная функция $S(x)$ является дискретной сверткой исходной дискретной функции заданной упорядоченным массивом выборок s_i с функцией называемой ядром Ланцоша:

$$S(x) = \sum_{i=\lfloor x \rfloor - a + 1}^{\lfloor x \rfloor + a} s_i L(x - i). \quad (2.19)$$

Ядро Ланцоша представляет собой произведение функции $\text{sinc}(x)$ на

оконную функцию Ланцоша, по определению равную нулю вне заданного параметром носителя функции a :

$$L(x) = \begin{cases} \text{sinc}(x) \text{sinc}(\frac{x}{a}), & \text{если } -a < x < a, \\ 0, & \text{иначе,} \end{cases} \quad (2.20)$$

или

$$L(x) = \begin{cases} 1, & \text{если } x = 0, \\ \frac{a \sin(\pi x) \sin(\frac{\pi x}{a})}{\pi^2 x^2}, & \text{если } -a \leq x < a \text{ и } x \neq a, \\ 0, & \text{иначе.} \end{cases} \quad (2.21)$$

При обработке изображений, так как они являются двумерными функциями для свертки применяется двумерное ядро Ланцоша $L(x, y) = L(x)L(y)$.

При обработке возможно появление нежелательных артефактов типа граничного гало. Возникновение ореолов обусловлено тем, что при значении параметра $a > 1$ ядро Ланцоша принимает отрицательные значения при некоторых значениях аргумента. Поэтому обработанный сигнал может принимать даже отрицательные значения при положительных значениях выборок. При практической обработке изображений удовлетворительное качество достигается при значении параметра $a = 2$ или $a = 3$ [11]. Графики функции ядра Ланцоша для $a = 2$ и $a = 3$ показаны на рисунке 2.9.

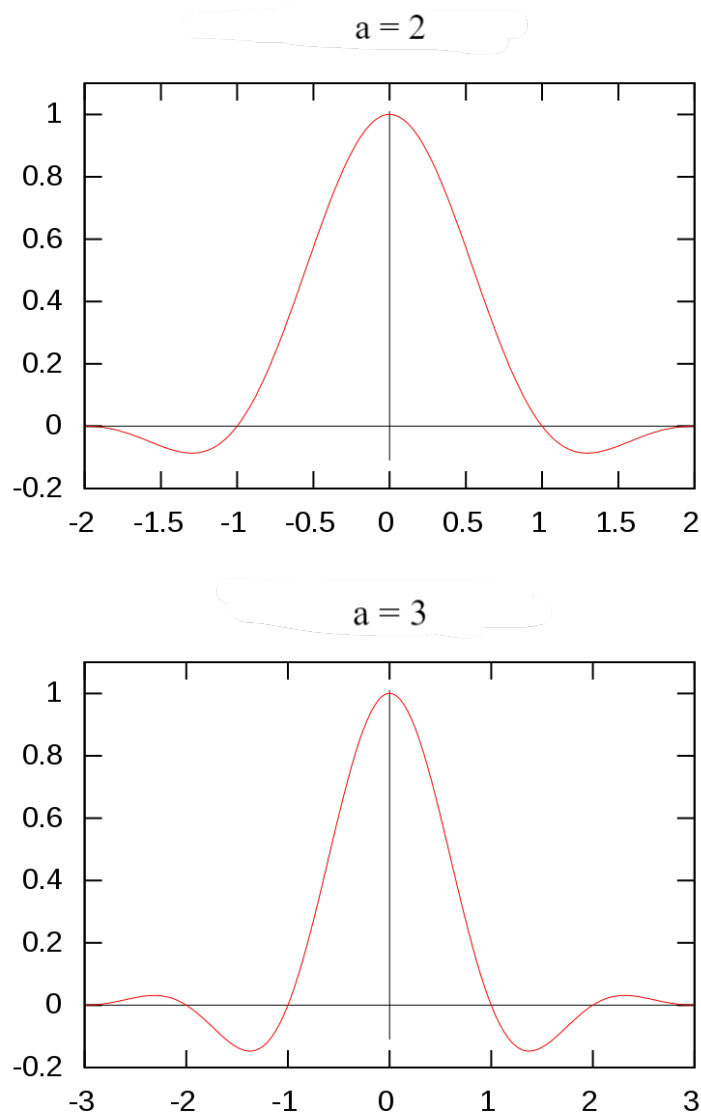


Рисунок 2.9 – Графики функции ядра Ланцоша для $a = 2$ и $a = 3$

2.2 Нейронные методы

2.2.1 Сверточные нейронные сети

Сверточные нейронные сети представляют собой класс глубоких нейронных сетей, разработанных специально для обработки и анализа структурированных данных, таких как изображения. Основная идея заключается в автоматическом извлечении иерархии признаков из входных данных.

В начальной стадии, сверточная нейронная сеть принимает на вход тензор, размерности которого есть высота, ширина и число цветовых каналов. Процесс анализа начинается с применения сверточных слоев, осуществляющих операции свертки с целью обнаружения локальных признаков [12].

После свертки применяются слои подвыборки, которые служат для уменьшения размерности данных и создания инвариантности к масштабу и местоположению объектов. Последующие полносвязные слои объединяют выделенные признаки [12].

Сети обучаются с использованием метода обратного распространения ошибки, который корректирует веса сети, минимизируя ошибку прогнозов.

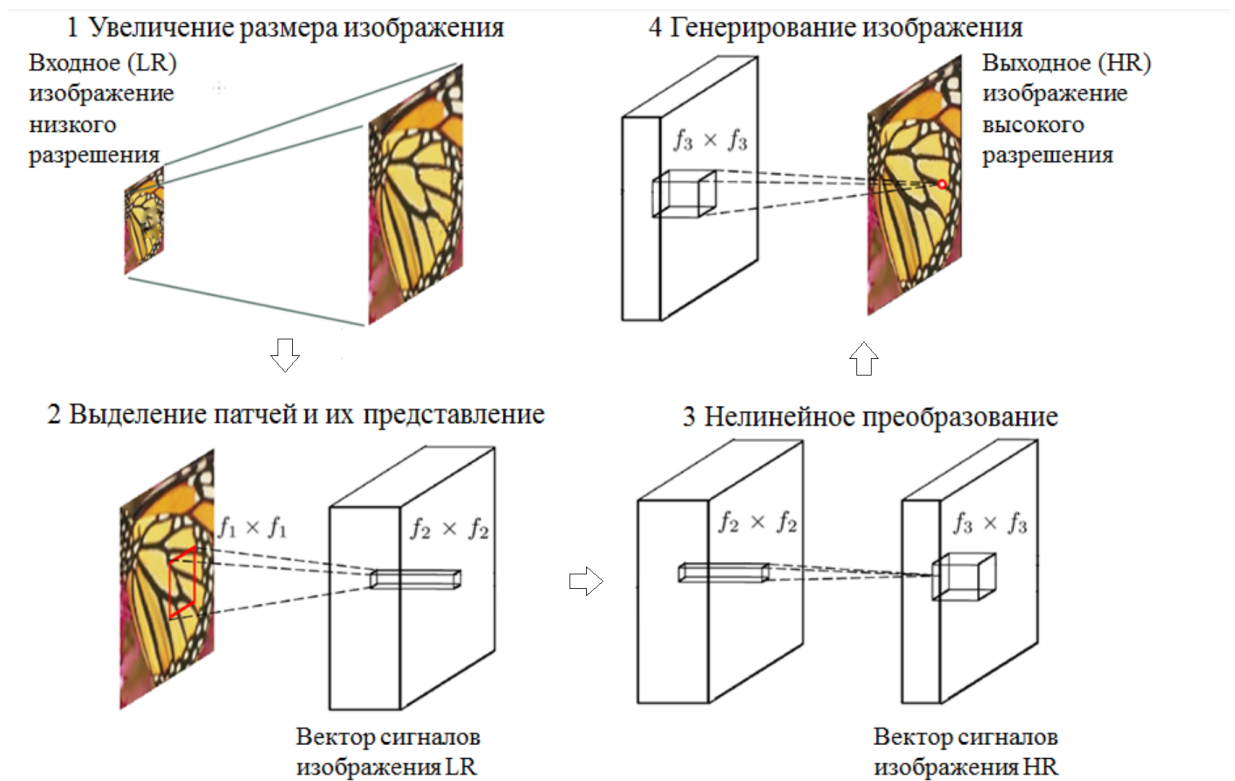


Рисунок 2.10 – Схема сверточной нейронной сети для решения задачи увеличения разрешения

Сверточными нейронными сетями изображения обрабатываются в несколько этапов.

Первый этап: увеличение размера изображения. Прежде всего, необходимо увеличить размер исходного изображения до нужного, это чаще всего делают с помощью бикубической интерполяции.

Второй этап: выделение патчей и их представление. На этом этапе проводится получение отдельных участков из изображения и представление их в виде многомерных векторов содержащих карты признаков размером равным размерности вектора по формуле:

$$F_1(y) = \max(0, W_1 * y + B_1), \quad (2.22)$$

где W_1 и B_1 соответствуют фильтру и смещению, а знак « $*$ » означает операцию свертки. W_1 представляет собой набор фильтров в количестве n_1 с размерностью $c \times f_1 \times f_1$, где c — количество каналов изображения, а f_1 — размер фильтра. Результат представляет собой набор карт свойств в количестве n_1 . Взятие поэлементного максимума осуществляется с использованием функции активации ReLU [12].

Третий этап: нелинейное преобразование. Здесь под преобразованием понимается отображение каждого многомерного вектора на другие векторы. Каждый отображенный вектор представляет патч высокого разрешения, а вектора содержат новый набор карт признаков.

Преобразование одного n_1 -мерного вектора в другой n_2 -мерный вектор осуществляется по формуле:

$$F_2(y) = \max(0, W_2 * F_1(y) + B_2), \quad (2.23)$$

где W_2 представляет собой набор фильтров в количестве n_2 с размерностью $n_1 \times f_2 \times f_2$ [12].

Каждый новый вектор размерности n_2 концептуально представляет собой участок высокого разрешения, связанный с аналогичным участком низкого разрешения.

Четвертый этап: генерирование изображения высокого разрешения. Отображения высокого разрешения каждого патча из предыдущего шага используются для генерации изображения высокого разрешения. На этом шаге из n_2 -мерного вектора строится искомое изображение высокого разрешения по формуле:

$$F(y) = W_3 * F_2(y) + B_3, \quad (2.24)$$

где W_3 означает набор из линейных фильтров с размерностью $n_2 \times f_3 \times f_3$ в количестве c , а B_3 является вектором размерностью c [12].

2.3 Генеративно-состязательные нейронные сети

Генеративно-состязательные сети представляют собой систему из двух нейронных сетей, выполняющих роли генератора и дискриминатора. Модель-

генератор G преобразует переменные латентного множества z с использованием параметров $\Theta^{(G)}$. В свою очередь, модель-дискриминатор D использует наблюдаемые переменные x и использует $\Theta^{(D)}$ в качестве набора параметров.

В оригинальной формулировке задачи целью дискриминатора является минимизация функции потерь $J^{(D)}$:

$$J^{(D)} = -\frac{1}{2}\mathbb{E}_{x \sim p_{\text{data}}} [\log D(x)] - \frac{1}{2}\mathbb{E}_z [\log (1 - D(G(z)))], \quad (2.25)$$

где p_{data} — это распределение обучающей выборки

В свою очередь, целью генератора является минимизация функции $J^{(D)}$:

$$J^{(G)} = -\frac{1}{2}\mathbb{E}_z [\log D(G(z))], \quad (2.26)$$

Можно определить следующее интуитивное объяснение принципа работы модели: сеть дискриминатор обучается с целью отличать реальные данные от результатов работы сети-генератора. Одновременно, генератор учится «обманывать» дискриминатор, генерируя все более реалистичные изображения с каждым шагом обучения. Результатом такого итеративного обучения является сеть-генератор, которая способна производить изображения, находящиеся в латентном множестве реальных изображений, согласно представлению дискриминатора.

В своей оригинальной форме архитектура GAN используется для генерации убедительных изображений из векторов случайного шума. Такой подход неприменим к задаче повышения разрешения изображений, так как целью является генерация изображений, близких по содержанию к входным изображениям низкого разрешения. Для использования генеративного подхода в данной задаче, необходимо модифицировать функцию потерь сети-генератора. Сеть-генератор будет использовать функцию, основанную на по-пиксельной средней квадратичной ошибке (MSE – Mean Squared Error).

$$L_{MSE} = \frac{1}{n} \sum_n (Y_{Gen} - Y_{HR})^2. \quad (2.27)$$

Предложенная функция потерь определена как комбинация ошибки генератора и MSE. Это позволяет сети-генератору частично отклоняться от оптимальных с точки зрения MSE результатов, генерируя натуральные детали для изображенных объектов:

$$L_{SR} = L_{MSE} + 10^{-3} J^{(G)}. \quad (2.28)$$

2.4 Сравнение методов повышения разрешающей способности изображения

Задача сравнения классических и нейронных методов повышения разрешающей способности не является адекватной, так как данные типы методов предоставляют кардинально разное качество результата, однако и требуют различных вычислительных мощностей. Именно поэтому в данной работе будет проведено несколько сравнений, а именно классических и нейронных методов между собой и классических и нейронных методов как два типа друг с другом.

Однако все критерии будут оценены по следующей схеме: 1 — лучший результат по данному критерию, n — худший результат, где n — количество сравниваемых методов. Если критерии имеют одинаковый результат, то они получают баллы по формуле $\frac{n_1+n_2}{2}$, это необходимо для правильного масштабирования весов оценки, если далее будет найдено среднее арифметическое всех оценок, то данная формула даст более корректный результат.

Критерии сравнения **классических** методов:

- K_1 — ступенчатость;
- K_2 — размытие;
- K_3 — граничное гало;
- K_4 — вычислительная сложность.

Результаты сравнения классических методов увеличения разрешающей способности изображений представлены в таблице 2.3.

Как было сказано ранее, классические методы интерполяции всегда являются компромиссом между тремя дефектами, соответствующими критериям K_1, K_2, K_3 (графическое сравнение представлено на рисунке 2.11). Таким

Таблица 2.1 – Сравнение классических методов повышения разрешающей способности

Метод	K_1	K_2	K_3	K_4
Ближайшего соседа	4	1	1.5	1
Билинейный	3	3	1.5	2
Бикубический	1	4	3.5	3
Ланцоша	2	2	3.5	4

образом, методами показывающими наилучший результат по качеству являюмя методы бикубической интерполяции и метод Ланцоша, это подтверждается их частотой использования. Однако из представленных они же являются наиболее затратными, что позволяет судить об их неприменимости в некоторых ситуациях, которые требуют минимальной вычислительной сложности. Здесь лучшие результаты показывает метод ближайшего соседа. И как компромисс между качеством и сложностью — метод билинейной интерполяции.

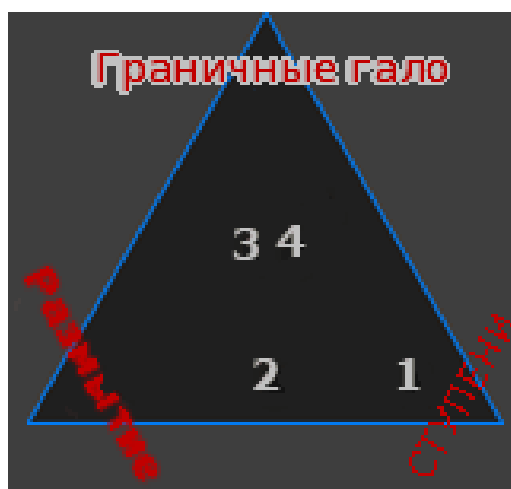


Рисунок 2.11 – Диаграмма сравнения методов интерполяции на наличие трех дефектов

Критерии сравнения **нейронных** методов:

- K_1 — время обучения;
- K_2 — вычислительные ресурсы;
- K_3 — память;
- K_4 — сложность реализации;

- K_5 — пиковый сигнально-шумовой коэффициент;
- K_6 — коэффициент структурной схожести;

Пиковое отношение сигнала к шуму — это метрика, используемая для измерения качества восстановленных изображений путем сравнения их с оригиналами. Она основана на отношении максимально возможного значения сигнала (максимальной яркости пикселей) к среднеквадратичной ошибке между оригинальным и восстановленным изображением.

Индекс структурного сходства — это метрика, оценивающая структурное сходство между двумя изображениями. Она учитывает яркость, контраст и структурное подобие пикселей в оригинальном и восстановленном изображении.

Таблица 2.2 – Сравнение нейронных методов повышения разрешающей способности

Метод	K_1	K_2	K_3	K_4	K_5	K_6
Сверточный	1	1	1	1	1	2
Генеративно-состязательный	2	2	2	2	2	1

Подводя итог сравнения нейронных методов, можно отметить, что генеративно-состязательные нейронные сети требуют большее количество ресурсов, однако они предоставляют высокое визуальное качество и реализм, в то время как сверточные нейронные сети больше сосредоточены на точности восстановления и минимизации ошибок.

Критерии сравнения **классических и нейронных** методов:

- K_1 — вычислительная сложность;
- K_2 — сложность реализации;
- K_3 — качество.

Так как критерии не являются объективными, оценить их можно только посредством экспертного мнения. Данные типы методов были классифицированы по выделенным критериям согласно источникам [9] и [13].

Таблица 2.3 – Сравнение нейронных методов повышения разрешающей способности

Тип метода	K_1	K_2	K_3
Классические	1	1	2
Нейронные	2	2	1

Получены в некотором роде очевидные результаты: нейронные методы более ресурсоемкие, однако они обеспечивают визуально качественные изображения, в то время как классические методы являются более «быстрыми» и «легкими», однако их даже лучшие результаты нельзя назвать удовлетворительными по сравнению с нейронными сетями.

Как итог, использование того или иного типа метода и самого метода зачислит от постановки задачи.

Вывод

В данной главе были описаны классические и нейронные методы повышения разрешения изображения, проведена их классификация, по введенным ранее критериям.

ЗАКЛЮЧЕНИЕ

В ходе выполнения данной работы были классифицированы методы повышения разрешающей способности изображений.

На основании результатов сравнения методов можно сделать вывод о том, что в общем случае предпочтительными методами являются нейронные, так как они предоставляют более качественный результат, но являются более ресурсозатратными, чем классические, которые тоже имеют право на использование в силу своей легкости и удовлетворительного качества таких, как методы Ланцоша и бикубический, другие классические методы оказались еще более простыми, но их качество результата слишком сильно подвержено дефектам, присущим данному типу. Однако конечный выбор того или иного метода должен зависеть от постановки задачи.

Цель, поставленная в начале работы, была достигнута — классифицированы методы повышения разрешающей способности изображений. В ходе ее выполнения были решены все задачи:

- проведен анализ предметной области;
- описаны существующие методы повышения разрешения изображений;
- сформулированы критерии сравнения методов повышения разрешающей способности изображений;
- проведено сравнение методов на основании выделенных критериев;
- описаны результаты сравнения рассмотренных методов.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. Отчет App Annie из электронной статьи «Мобильные приложения мировой рынок» [Электронный ресурс]. URL: [https://www.tadviser.ru/index.php/%D0%A1%D1%82%D0%B0%D1%82%D1%8C%D1%8F:%D0%9C%D0%BE%D0%B1%D0%B8%D0%BB%D1%8C%D0%BD%D1%8B%D0%B5_%D0%BF%D1%80%D0%B8%D0%BB%D0%BE%D0%B6%D0%B5%D0%BD%D0%B8%D1%8F_\(%D0%BC%D0%B8%D1%80%D0%BE%D0%B2%D0%BE%D0%B9_%D1%80%D1%8B%D0%BD%D0%BE%D0%BA](https://www.tadviser.ru/index.php/%D0%A1%D1%82%D0%B0%D1%82%D1%8C%D1%8F:%D0%9C%D0%BE%D0%B1%D0%B8%D0%BB%D1%8C%D0%BD%D1%8B%D0%B5_%D0%BF%D1%80%D0%B8%D0%BB%D0%BE%D0%B6%D0%B5%D0%BD%D0%B8%D1%8F_(%D0%BC%D0%B8%D1%80%D0%BE%D0%B2%D0%BE%D0%B9_%D1%80%D1%8B%D0%BD%D0%BE%D0%BA) (дата обращения: 23.11.2023).
2. Основы компьютерной графики: учебное пособие / А.Ю. Дёмин; Томский политехнический университет. – Томск: Изд-во Томского политехнического университета, 2011. – 191 с.
3. Захаренко А. В. Алгоритмы повышения разрешения изображений // Белорусский государственный университет информатики и радиоэлектроники. 2020.
4. Электронная статья на тему «Интерполяция цифрового изображения» [Электронный ресурс]. URL: <https://www.cambridgeincolour.com/ru/tutorials-ru/image-interpolation.htm> (дата обращения: 16.11.2023).
5. Электронная статья на тему «Увеличение цифровых фотографий» [Электронный ресурс]. URL: <https://www.cambridgeincolour.com/ru/tutorials-ru/digital-photo-enlargement.htm> (дата обращения: 16.11.2023).
6. Основы машинного обучения : учебное пособие / О.В. Лимановская, Т.И. Алферьева ; Мин-во науки и высш. образования РФ.— Екатеринбург : Изд-во Урал. ун-та, 2020. — 88 с.
7. Т.Р. Назаров, Н.А. Мамедова Автоматизированное решение задачи детектирования промышленных объектов на ортофотоплане с помощью нейронной сети // Программные продукты и системы. 2023. №1.
8. Галетка А. В. Анализ случайных процессов с помощью нейронных сетей // Московский государственный университет им. М. В. Ломоносова. 2008.

9. A. Kumar, C. Kumar. Estimation of motion among shifted images at coarse and fine levels // International Journal of Computational Intelligence and Information Security. 2011.
10. Веселов Ю. Г., Островский А. С. К вопросу оценки разрешающей способности при масштабировании цифровых изображений // Наука и образование. 2010.
11. Duchon Claude E. Lanczos Filtering in One and Two Dimensions // Journal of applied meteorology. 1979.
12. Гафаров Ф.М Г12 Искусственные нейронные сети и приложения: учеб. пособие / Ф.М. Гафаров, А.Ф. Галимянов. – Казань: Изд-во Казан. ун-та, 2018. – 121 с.
13. Захаренко А. В. Алгоритмы повышения пространственного разрешения изображений земной поверхности // Белорусский государственный университет информатики и радиоэлектроники. 2021.

ПРИЛОЖЕНИЕ А

Презентация к работе



**Федеральное государственное бюджетное образовательное учреждение высшего образования
«Московский государственный технический университет имени Н.Э. Баумана
(национальный исследовательский университет)»
(МГТУ им. Н.Э. Баумана)**

Классификация методов повышения разрешающей способности изображений

Студент: Светличная Алина Алексеевна (группа – ИУ7-73Б)

Научный руководитель: Филиппов Михаил Владимирович

Цель и задачи

Цель: классифицировать методы повышения разрешающей способности изображений.

Задачи:

- провести анализ предметной области;
- описать существующие методы повышения разрешения изображений;
- сформулировать критерии сравнения методов повышения разрешающей способности изображений;
- сравнить описанные методы на основании выделенных критериев;
- описать результаты сравнения рассмотренных методов.

Введение в классические методы

Адаптивные

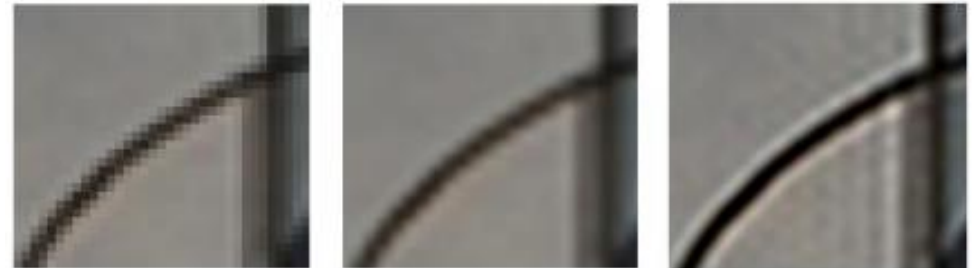
учитывают локальную структуру и содержание изображения при принятии решений о том, как интерполировать

Неадаптивные

не учитывают контекст или структуру изображения, а применяют один и тот же метод интерполяции ко всем пикселям


Интерполяция — процесс определения значений между известными точками данных.

- **Ступенчатость** — видимые лестничные ступеньки или резкие перепады яркости вдоль контуров объектов.
- **Размытие** — потеря деталей и четкости изображения при сглаживании краев.
- **Граничное гало** — светлые или темные области, окружающие объекты или контуры на изображении.



Введение в нейронные методы

Машинное обучение — направление искусственного интеллекта, сосредоточенное на создании систем, которые обучаются и развиваются на основе получаемых ими данных.



Без учителя

обучение выявлять закономерности и скрытые взаимосвязи на наборах неразмеченных данных без контроля со стороны пользователя

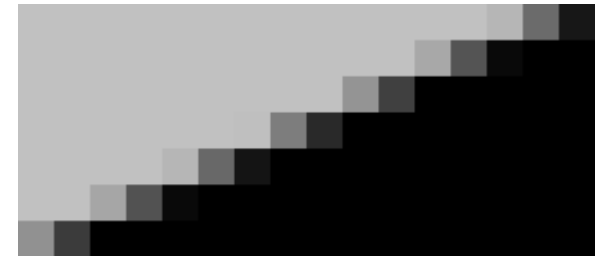
С учителем

построение моделей на основе множества примеров, содержащих пары «известный вход — известный выход»

Искусственные нейронные сети — системы, преобразующие информацию подобно процессам, происходящим в человеческом мозге.

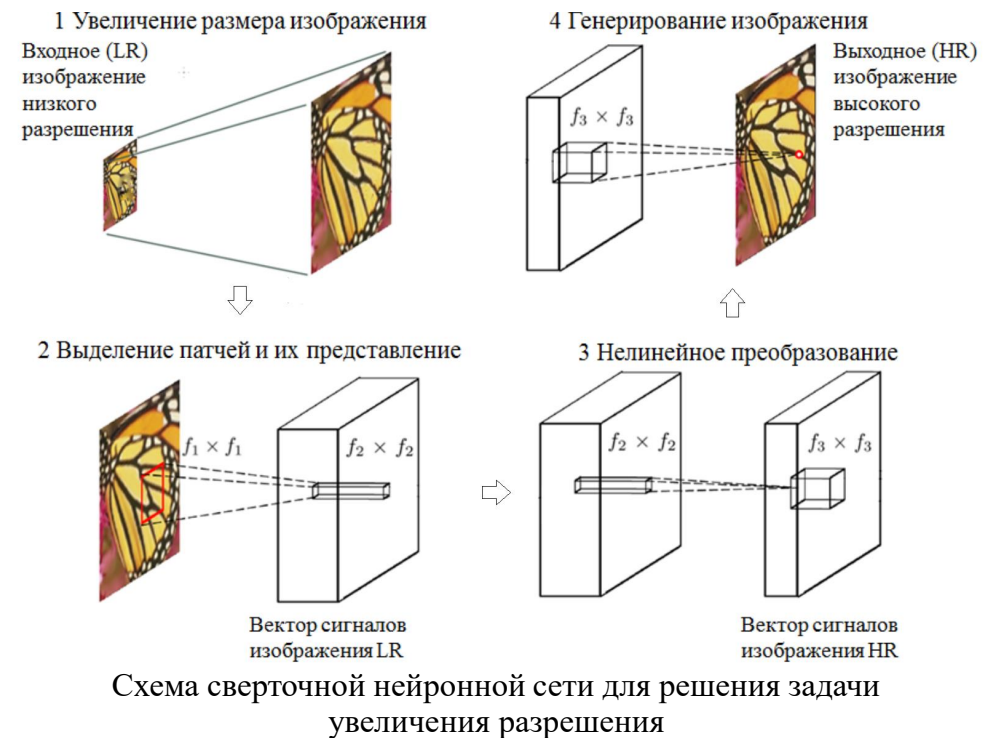
Классические методы

- **Метод ближайшего соседа** учитывает только один пиксель ближайший к точке интерполяции.
- **Билинейная интерполяция** рассматривает квадрат 2×2 известных пикселей, окружающих неизвестный.
- **Бикубическая интерполяция** рассматривает массив из 4×4 окружающих пикселей.
- **Метод Ланцоша** основан на применении нормированной функции $\text{sinc}()$.



Нейронные методы

- **Сверточные нейронные сети** – автоматическое извлечение иерархии признаков из входных данных.
- **Генеративно-состязательные нейронные сети** – система из двух нейронных сетей, выполняющих роли генератора (генерирует изображения) и дискриминатора (оценивает реалистичность сгенерированных изображений).



Классификация методов

Классические методы:

- K_1 – ступенчатость
- K_2 – размытие
- K_3 – граничное гало
- K_4 – вычислительная сложность

Метод	K_1	K_2	K_3	K_4
Ближайшего соседа	4	1	1.5	1
Билинейный	3	3	1.5	2
Бикубический	1	4	3.5	3
Ланцоша	2	2	3.5	4

Нейронные методы:

- K_1 – время обучения
- K_2 – вычислительная сложность
- K_3 – память
- K_4 – сложность реализации
- K_5 – пиковый сигнальный шум
- K_6 – коэффициент структурной схожести

Метод	K_1	K_2	K_3	K_4	K_5	K_6
Сверточный	1	1	1	1	1	2
Генеративно-состязательный	2	2	2	2	2	1

Классификация методов

Классические и нейронные типы методов:

- K_1 – вычислительная сложность
- K_2 – сложность реализации
- K_3 – качество

Методы	K_1	K_2	K_3
Классические	1	1	2
Нейронные	2	2	1

Вывод: нейронные методы показывают визуально более качественные результаты, однако являются ресурсозатратными (вычислительные, человеческие, по памяти и т.д.), в то время как классические методы демонстрируют приемлемое для отдельных задач качество и являются в разы более легкими. Из названных типов методов можно выделить как наиболее хорошо показавших себя методы бикубической интерполяции и Ланцоша из классических и сверточные нейронные сети из соответственно нейронных методов.

Заключение

В ходе выполнения научно-исследовательской работы поставленная **цель** была достигнута: классифицированы методы повышения разрешающей способности изображений.

Все **задачи** выполнены:

- проведен анализ предметной области;
- описаны существующие методы повышения разрешения изображений;
- сформулированы критерии сравнения методов повышения разрешающей способности изображений;
- проведено сравнение описанных методов на основании выделенных критериев;
- описаны результаты сравнения рассмотренных методов.