

ВВЕДЕНИЕ

Актуальность темы исследования обусловлена стремительным развитием технологий дистанционного зондирования Земли (ДЗЗ) и расширением спектра задач, решаемых с их помощью. Спутниковые снимки высокого разрешения востребованы в картографии, мониторинге экологической обстановки, сельском хозяйстве, градостроительстве и задачах обеспечения безопасности. Однако получение изображений сверхвысокого разрешения аппаратными средствами сопряжено с существенными техническими и экономическими ограничениями, связанными со стоимостью оптики, ограничениями каналов передачи данных и физическими параметрами орбитальных аппаратов.

В связи с этим возникает необходимость применения программных методов повышения пространственного разрешения, которые позволяют восстанавливать детализацию изображений на основе алгоритмической обработки. На сегодняшний день существует широкий спектр подходов к решению данной задачи: от классических методов интерполяции до современных алгоритмов, основанных на глубоком машинном обучении.

Несмотря на активное развитие нейросетевых технологий, вопрос выбора оптимального метода для обработки именно спутниковых данных остается открытым. Спутниковые снимки обладают специфическими характеристиками, такими как сложная текстура, наличие шумов атмосферного происхождения и разнообразие масштабов объектов, что отличает их от обычных фотографий. Поэтому проведение сравнительного анализа классических и нейросетевых подходов является важной исследовательской задачей, позволяющей определить границы применимости каждого из методов.

Объектом исследования является процесс цифровой обработки изображений дистанционного зондирования Земли.

Предметом исследования выступают методы и алгоритмы повышения пространственного разрешения изображений.

Целью курсовой работы является выявление наиболее эффективных подходов к повышению разрешения спутниковых снимков на основе сравнительного анализа классических методов интерполяции и алгоритмов, основанных на глубоких нейронных сетях.

Для достижения поставленной цели необходимо решить следующие задачи:

1. Провести теоретический анализ классических и нейросетевых методов повышения разрешения (Super-Resolution) изображений.
2. Разработать алгоритмы для подготовки данных на основе спутниковых снимков, их последующей обработки с применением классических и нейросетевых методов и визуализации полученных результатов.
3. Провести сравнительный эксперимент по оценке эффективности методов на основе количественных метрик и визуального анализа.
4. Сформулировать выводы о преимуществах, недостатках и границах применимости рассмотренных подходов для задач обработки спутниковых снимков Земли.

TODO доделать методы исследования пересмотреть

Методы исследования. В работе используются методы теории цифровой обработки сигналов и изображений, методы машинного обучения, в частности глубокие нейронные сети (CNN, GAN), а также методы статистического анализа данных для оценки качества работы алгоритмов. Реализация практической части выполняется с использованием языка программирования Python и специализированных библиотек компьютерного зрения.

Практическая значимость работы заключается в том, что полученные результаты и разработанные программные модули могут быть использованы для улучшения качества архивных спутниковых снимков, подготовки данных для систем автоматического распознавания объектов.

МАТЕМАТИЧЕСКАЯ ПОСТАНОВКА ЗАДАЧИ

Задача повышения разрешения (Super Resolution, SR) изображений относится к классу некорректно поставленных обратных задач (ill-posed inverse problems). Основная цель заключается в восстановлении высокочастотных деталей изображения высокого разрешения (High Resolution, HR) на основе имеющегося изображения низкого разрешения (Low Resolution, LR).

Пусть $I_{HR} \in \mathbb{R}^{H \times W \times C}$ — исходное изображение высокого разрешения (High Resolution), где H и W — высота и ширина изображения, а C — количество цветовых каналов.

Процесс формирования наблюдаемого изображения низкого разрешения $I_{LR} \in \mathbb{R}^{h \times w \times C}$ (где $h = H/s$, $w = W/s$, а s — коэффициент масштабирования) описывается моделью деградации. Данная модель математически имитирует потерю качества, происходящую при съемке аппаратурой спутника из-за ограничений оптики и сенсоров. В общем виде связь между I_{LR} и I_{HR} выражается следующим образом:

$$I_{LR} = D(I_{HR}; \delta)$$

где D — оператор деградации, зависящий от параметров δ (масштабирование, размытие, шум).

В рамках данной работы для проведения экспериментов используется стандартная модель деградации, включающая операцию бикубического понижения дискретизации (downsampling). Это позволяет сформировать обучающую выборку из пар изображений, используя имеющиеся высококачественные спутниковые снимки в качестве эталонов:

$$I_{LR} = (I_{HR}) \downarrow_s$$

где \downarrow_s — операция уменьшения пространственного разрешения с коэффициентом s (обычно $s = 2, 4$ или 8).

Целью алгоритма повышения разрешения является нахождение функции отображения F , которая позволяет получить восстановленное изображение I_{SR} (Super-Resolution) из I_{LR} , максимально приближенное к исходному I_{HR} :

$$I_{SR} = F(I_{LR}) \approx I_{HR}$$

TODO пересмотреть потери и метрики

Задача сводится к минимизации функции потерь \mathcal{L} между восстановленным и эталонным изображениями. В случае использования методов глубокого обучения (нейронных сетей) происходит поиск оптимальных весовых коэффициентов θ сети:

$$\hat{\theta} = \arg \min_{\theta} \frac{1}{N} \sum_{i=1}^N \mathcal{L}(F(I_{LR}^{(i)}; \theta), I_{HR}^{(i)})$$

Где N — количество изображений в обучающем наборе данных. В качестве функции потерь \mathcal{L} могут выступать попиксельные метрики (MSE, MAE) для обеспечения точного цветового соответствия, либо более сложные перцептивные метрики (Perceptual Loss, Adversarial Loss) для улучшения визуального восприятия текстур ландшафта.

КЛАССИЧЕСКИЕ МЕТОДЫ ПОВЫШЕНИЯ РАЗРЕШЕНИЯ

Классические подходы к задаче Super-Resolution базируются на методах интерполяции. Их основной принцип заключается в оценке значений пикселей изображения высокого разрешения на основе локальной окрестности пикселей исходного изображения низкого разрешения. В рамках данной работы рассматриваются и программно реализуются два наиболее распространенных метода: бикубическая интерполяция и интерполяция Ланцоша.

2.1 Бикубическая интерполяция

Бикубическая интерполяция является усовершенствованием билинейного метода и обеспечивает более гладкие переходы яркости, сохраняя при этом больше деталей. Математически метод аппроксимирует значение пикселя с использованием полиномов третьей степени.

Для вычисления значения нового пикселя рассматривается окрестность размером 4×4 пикселя исходного изображения. Вклад каждого соседа определяется весовой функцией $W(d)$, зависящей от расстояния d до искомой точки.

В работе реализован алгоритм на основе стандартного кубического ядра свертки (weighting kernel). Функция весов $W(d)$ задается следующей кусочно-заданной функцией:

$$W(d) = \begin{cases} (a + 2)|d|^3 - (a + 3)|d|^2 + 1, & |d| \leq 1 \\ a|d|^3 - 5a|d|^2 + 8a|d| - 4a, & 1 < |d| < 2 \\ 0, & \text{else} \end{cases}$$

где a — свободный параметр, определяющий характер интерполяции (резкость). Обычно принимается значение $a = -0.5$.

С вычислительной точки зрения процесс интерполяции реализуется как последовательная (сепарабельная) свертка: сначала вычисляются промежуточные значения вдоль одной оси (например, по строкам), а затем полученный результат интерполируется по другой оси (по столбцам). Это позволяет существенно сократить количество операций по сравнению с двумерной сверткой.

2.2 Интерполяция Ланцоша

Метод Ланцоша (Lanczos resampling) считается одним из лучших классических алгоритмов для изменения размера изображений, обеспечивая высокий уровень четкости ("резкости") границ объектов. В его основе лежит математический аппарат оконных функций и использования кардинального синуса (Sinc).

Ядро фильтра Ланцоша $L(x)$ представляет собой функцию Sinc, ограниченную оконной функцией того же типа. Формула для вычисления весовых коэффициентов имеет вид:

$$L(x) = \begin{cases} \text{sinc}(x) \cdot \text{sinc}\left(\frac{x}{a}\right), & -a < x < a \\ 0, & \text{else} \end{cases}$$

где параметр a определяет размер окна фильтра (радиус влияния). Функция $\text{sinc}(x)$ определяется как нормированный кардинальный синус:

$$\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$$

В рамках данной работы используется конфигурация Lanczos-3 (параметр $a = 3$). Это означает, что для вычисления значения одного пикселя результирующего изображения используется информация из окрестности размером $2a \times 2a$, то есть 6×6 пикселей исходного изображения (диапазон x от -3 до 3).

Суть алгоритма заключается в вычислении взвешенной суммы интенсивностей соседних пикселей, где веса зависят от расстояния до центра интерполяции и рассчитываются по приведенной формуле ядра $L(x)$. Метод Ланцоша позволяет минимизировать эффект алиасинга (ступенчатости наклонных линий), однако требует больших вычислительных ресурсов по сравнению с бикубической интерполяцией из-за более широкого окна выборки и сложности вычисления тригонометрических функций.

НЕЙРОСЕТЕВЫЕ МЕТОДЫ ПОВЫШЕНИЯ РАЗРЕШЕНИЯ

В отличие от классических методов интерполяции, которые вычисляют недостающие пиксели по фиксированной аналитической формуле, методы на основе глубокого обучения (Deep Learning) восстанавливают высокочастотные детали, опираясь на закономерности, выявленные в ходе обучения на больших наборах данных.

3.1 Сверточные нейронные сети (SRCNN и FSRCNN)

Пионерским подходом в данной области является метод SRCNN (Super-Resolution Convolutional Neural Network)[1]. Идея метода заключается в прямом обучении отображения между изображением низкого разрешения (LR) и высокого разрешения (HR) в сквозном (end-to-end) режиме.

Архитектура SRCNN состоит из трех основных этапов, которые интерпретируются как слои сверточной нейронной сети:

1. Извлечение признаков (Patch extraction and representation): извлечение перекрывающихся участков изображения и представление их в виде векторов признаков.
2. Нелинейное отображение (Non-linear mapping): проецирование векторов признаков из пространства низкой размерности в пространство высокой размерности.
3. Реконструкция (Reconstruction): восстановление итогового изображения высокого разрешения.

Важной особенностью SRCNN является то, что на вход сети подается не исходное маленькое изображение, а уже увеличенное до целевого размера (с помощью бикубической интерполяции) изображение.

Математически процесс обработки описывается композицией операций свертки:

$$F(Y) = W_3 * \phi(W_2 * \phi(W_1 * Y + B_1) + B_2) + B_3$$

Где Y — входное интерполированное изображение, W_i и B_i — веса и смещения слоев, $*$ — операция свертки, ϕ — функция активации (обычно ReLU).

Рис. 1: Архитектура сети SRCNN

Несмотря на высокое качество работы по сравнению с классическими методами, SRCNN имеет существенный недостаток: высокая вычислительная сложность. Поскольку сеть обрабатывает изображение уже в высоком разрешении, количество операций свертки значительно возрастает пропорционально квадрату коэффициента масштабирования s^2 .

Для решения этой проблемы и ускорения обработки в данной курсовой работе выбран и реализован метод FSRCNN (Fast Super-Resolution Convolutional Neural Network) [2].

Ключевое отличие FSRCNN заключается в том, что сеть принимает на вход оригинальное изображение низкого разрешения без предварительной интерполяции. Увеличение размерности происходит только на самом последнем слое с помощью операции деконволюции (transposed convolution). Это позволяет проводить все нелинейные преобразования в пространстве низкой размерности, что существенно снижает вычислительные затраты.

Архитектура FSRCNN включает пять этапов:

1. Извлечение признаков: первая свертка выполняется непосредственно над сырым LR-изображением.
2. Сжатие (Shrinking): уменьшение размерности карт признаков с помощью фильтров 1×1 для сокращения количества параметров.
3. Нелинейное отображение (Mapping): последовательность сверточных слоев размером 3×3 , выполняющих основную обработку.
4. Расширение (Expanding): обратное увеличение размерности карт признаков перед восстановлением.
5. Деконволюция (Deconvolution): слой, который выполняет функцию `upsampling` (повышение разрешения) и формирует итоговое HR-изображение.

Слой деконволюции можно рассматривать как операцию, обратную свертке. Если свертка с шагом k уменьшает размер изображения, то деконволюция с шагом k увеличивает его. Математически это позволяет обучаемой модели самой находить оптимальные фильтры для повышения разрешения, превосходящие фиксированные ядра бикубической интерполяции.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

TODO сделать по госту

- [1] Dong C., Loy C. C., He K., Tang X. Image super-resolution using deep convolutional networks // IEEE transactions on pattern analysis and machine intelligence. – 2015. – Vol. 38. – №. 2. – P. 295-307.
- [2] TODO: fsrnn work
- [3] Wang X. et al. Esrgan: Enhanced super-resolution generative adversarial networks // Proceedings of the European conference on computer vision (ECCV) workshops. – 2018.