

Методы слияния гиперспектральных и мультиспектральных изображений на основе попиксельных спектральных преобразований и цикловой согласованности

© Р.Р. Ходырев, С.А. Баженов, Т.В. Кабанова, Д.А. Вражнов[✉], Ю.В. Кистенев

Национальный исследовательский Томский государственный университет,
Томск, Россия

[✉]e-mail: vda@iao.ru

Поступила в редакцию 05.12.2025 г.

В окончательной редакции 12.01.2026 г.

Принята к публикации 27.03.2026 г.

Представлены два подхода к слиянию гиперспектральных и мультиспектральных изображений. Первый подход основан на модуле WMatrixGenerator, генерирующем попиксельные матрицы спектрального преобразования с квадратичной коррекцией. Второй подход использует архитектуру HSIMSI-CycleGAN с цикловой согласованностью и физически мотивированными потерями. Экспериментальная оценка показала конкурентоспособность предложенных методов по метрикам PSNR, SSIM, SAM и ERGAS на наборе данных Real HSI/MSI/PAN. Методы работают на парных наблюдениях гиперспектральных и мультиспектральных изображениях одной сцены, но не требуют эталонных данных высокого пространственного разрешения.

Ключевые слова: слияние изображений, гиперспектральные изображения, мультиспектральные изображения, глубокое обучение, неотрицательные наименьшие квадраты, CycleGAN, дистанционное зондирование, нейронные сети.

DOI: 10.61011/OS.2026.05.63345.77-25

1. Введение

Гиперспектральные изображения (hyperspectral images, HSI) обладают высоким спектральным разрешением, предоставляя подробную информацию о составе материалов, но обычно имеют низкое пространственное разрешение [1]. Мультиспектральные изображения (multispectral images, MSI), наоборот, характеризуются высоким пространственным разрешением при ограниченном количестве спектральных каналов. Использование гиперспектральных камер на беспилотных летательных аппаратах (БПЛА) значительно расширило возможности регистрации гиперспектральных изображений в задачах агроэкологического мониторинга [2], например позволило увеличить пространственное разрешение с десятков метров (изображения со спутников) до десятков сантиметров [3]. Однако, принимая во внимание необходимость съемки больших по площади территорий с разрешением в несколько сантиметров, задача цифровой обработки гиперспектральных изображений с целью повышения их разрешения является актуальной. В настоящее время активно разрабатываются методы повышения пространственного разрешения гиперспектральных данных (HSI) либо спектрального разрешения мультиспектральных данных (MSI) за счет их совместной обработки [4]. Сравнительные характеристики MSI- и HSI-изображений даны в табл. 1.

За последние годы было предложено множество подходов для решения задачи слияния HSI-MSI. Их можно условно разделить на несколько категорий: методы на основе компонентного замещения, мультимасштабного

анализа, тензорного разложения, разреженного представления и в последнее время методы глубокого обучения. В табл. 2 представлена сводка основных методов с кратким описанием их принципов работы.

Эволюция методов слияния гиперспектральных и мультиспектральных данных проходит от линейных факторизационных подходов к тензорным, далее к модельно-ориентированным и основанным на физических принципах гибридным архитектурам. Базовые линейные методы, например неотрицательная факторизация матриц (nonnegative matrix factorization, NMF), метод главных компонент (principal component analysis, PCA), обеспечивают низкую вычислительную сложность и интерпретируемость, однако ограничены предпосылкой глобальной линейности и проявляют чувствительность к шуму, что приводит к умеренным значениям Peak Signal-to-Noise Ratio (PSNR) и недостаточно точному сохранению спектральной формы по метрике Spectral Angle Mapper (SAM).

Методы следующего уровня вводят структурные и статистические ограничения. Так, hyperspectral super-resolution via subspace-based regularization (HySure) использует подпространственную регуляризацию и совместную оценку функции рассеяния точки (point spread function, PSF) и спектральной функции отклика (spectral response function, SRF), что повышает устойчивость при корректной модели деградации. Nonnegative sparse representation (NSSR) опирается на разреженное представление с кластеризацией и неотрицательным словарем, улучшая оценку PSNR за счет роста вычислительных затрат.

Таблица 1. Сравнение HIS- и MSI-изображений

Характеристика	Мультиспектральные изображения	Гиперспектральные изображения
Количество каналов	3–10	>100
Ширина канала	Широкая (70–400 nm)	Узкая (5–10 nm)
Спектральное разрешение	Низкое	Высокое
Пространственное разрешение	До сантиметров	До десятков метров
Объем данных	Относительно небольшой	Очень большой („проклятие размерности“)
Ключевой принцип	Различить общие классы объектов	Идентифицировать материалы по их спектральной подписи

Таблица 2. Существующие методы слияния HIS- и MSI-изображений

Метод	Категория	Краткое описание
Nonnegative Matrix Factorization	Линейные	Неотрицательная матричная факторизация для разложения на базисные спектры
Principal Component Analysis	Линейные	Метод главных компонент для выделения основных направлений вариации
Coupled Sparse Tensor Factorization	Тензорные	Связанная разреженная тензорная факторизация с учетом сопряженности модальностей [5]
Hyperspectral Super-resolution via (Subspace-based) Regularization	Регуляризация	Подпространственная регуляризация с оценкой функции рассеяния точки и спектральной функции отклика [6]
Nonnegative Sparse Representation	Разреженные	Неотрицательное разреженное представление с кластеризацией и обучением словаря [7]
deep Tucker decomposition network	Тензорные и глубокое обучение	Разложение Таккера с нейросетевой аппроксимацией факторов [8]
Model Inspired Autoencoder	Модельные и глубокое обучение	Автоэнкодер с интегрированной NMF-структурой [9]
unsupervised Sparse Dirichlet-Net	Глубокое обучение	Неконтролируемая разреженная сеть Дирихле [10]
Coupled Convolutional Neural Network with Adaptive Response Function Learning	Глубокое обучение	Связанная сверточная сеть с адаптивной функцией отклика [11]
Coupled Unmixing Network with a cross-attention mechanism	Глубокое обучение	Сеть с перекрестным вниманием для разделения модальностей [12]
Unsupervised Deep Tensor Network	Тензорные и глубокое обучение	Неконтролируемая глубокая тензорная сеть [13]
Modality-Decoupled Spatial-Spectral Fusion	Глубокое обучение	Разложение на общие и комплементарные модальности [4]
Blind Estimation Generative Adversarial Network	Глубокое обучение	Двухэтапный подход с оценкой спектральной деградации [14]
Physics-Based Generative Adversarial Network	Глубокое обучение	Интеграция физической модели формирования изображений [15]
Registration And Fusion Nonconvex Low-Rank and Group-Sparse	Регуляризация и глубокое и групповой обучение	Регистрация с нелокальной регуляризацией разреженностью [16]
Efficient Hyperspectral Image Super-Resolution via Kolmogorov-Arnold Networks	Глубокое обучение	Сети на основе теоремы Колмогорова-Арнольда [14]

Тензорные методы, например Coupled Sparse Tensor Factorization (CSTF), deep Tucker decomposition network (Deep Tucker), Unsupervised Deep Tensor Network (UDTN), используют многомерную структуру данных. CSTF учитывает сопряженность модальностей, Deep Tucker добавляет нейросетевую аппроксимацию факторов, а UDTN достигает высоких метрик без парных эталонов, требуя внимательной настройки гиперпараметров. Модельные и сопряженно-сверточные архитектуры, такие как Model Inspired Autoencoder (MIAE), Coupled Convolutional Neural Network with Adaptive Response Function Learning (HyCoNet), Coupled Unmixing Network with a cross attention mechanism (CuCANet), unsupervised Sparse Dirichlet–Net (uSDN), Modality–Decoupled Spatial–Spectral Fusion (MossFuse), интегрируют априорные предположения в обучаемые структуры. Они сочетают интерпретируемость и нелинейность, однако чувствительны к размерности латентных пространств и точности калибровки спектрального отклика. Основанные на физических принципах и слепые подходы — Physics–Based Generative Adversarial Network (Physics GAN), Blind Estimation Generative Adversarial Network (BliEstGAN) — включают явную либо оцениваемую модель спектральной деградации. Первый достигает высоких значений метрики structural similarity index (SSIM) при зависимости от корректности модели, второй снижает требования к знаниям о сенсоре, но усложняет стабильность оптимизации. Методы с продвинутыми регуляризирующими механизмами (Registration And Fusion Nonconvex Low–Rank and Group–Sparse, RAF-NLRGS) и новые аппроксимирующие парадигмы (Efficient Hyperspectral Image Super–Resolution via Kolmogorov–Arnold Networks, HSR-KAN) обеспечивают минимальные значения метрики SAM за счет комплексной регуляризации и гибких нелинейных структур, сопровождаясь существенной ресурсоемкостью и незрелостью инструментов.

Несмотря на значительные успехи, достигнутые в области слияния HSI и MSI, многие существующие методы сталкиваются с рядом ограничений, таких как высокая вычислительная сложность, недостаточная робастность к шумам и вариативности сенсоров, а также ограниченная способность к обобщению на новые данные. Современные подходы требуют обширных размеченных данных или точного знания характеристик датчиков, что затрудняет их практическое применение. Настоящая работа предлагает решения этих проблем через разработку двух инновационных подходов: метода попиксельного спектрального преобразования на основе нейросетевой генерации матриц и архитектуры цикловой согласованности, использующей парные наблюдения HSI/MSI без привлечения эталонного гиперспектрального изображения высокого разрешения. Оба подхода демонстрируют баланс между точностью восстановления и практической применимостью в условиях реальных задач дистанционного зондирования.

Целью настоящей работы является разработка метода слияния HSI/MSI на основе архитектуры цикловой согласованности с физически мотивированными потерями. Для этого были решены следующие задачи.

- Проведен комплексный анализ современных методов слияния HSI и MSI, выявлены их преимущества и недостатки.
- Разработаны и реализованы метод слияния на основе попиксельного спектрального преобразования с использованием нейросетевой генерации матриц, оценена его эффективность на реальных данных.
- Разработан и реализован метод слияния на основе архитектуры цикловой согласованности с физически мотивированными потерями, исследовано влияние различных компонентов функции потерь на метрики качества.
- Описаны процедуры подготовки данных и обучения для предложенных методов, обеспечивающие воспроизводимость экспериментов.
- Проведена экспериментальная качественная и количественная оценка предложенных подходов на стандартных наборах данных, выполнен сравнительный анализ с существующими решениями по метрикам PSNR, SSIM, SAM и relative dimensionless global error in synthesis (ERGAS).

2. Материалы и методы

2.1. Математическая постановка задачи

Формально задача слияния формулируется следующим образом. Дано гиперспектральное изображение низкого пространственного разрешения $H \in \mathbb{R}^{h \times w \times L_h}$, где $h \times w$ — пространственные размеры (высота и ширина в пикселях), L_h — количество спектральных каналов (обычно десятки или сотни). Также дано мультиспектральное изображение высокого разрешения $M \in \mathbb{R}^{H \times W \times L_m}$, где $h_{\text{MSI}} \times w_{\text{MSI}}$ — пространственные размеры ($h_{\text{MSI}} \gg h$, $w_{\text{MSI}} \gg w$), L_m — количество спектральных каналов ($L_m \ll L_h$, обычно единицы). Требуется восстановить изображение высокого качества $\hat{Y} \in \mathbb{R}^{h_{\text{MSI}} \times w_{\text{MSI}} \times L_h}$ (high-quality hyperspectral image, HQHSI), обладающее одновременно высоким пространственным разрешением (как у MSI) и высоким спектральным разрешением (как у HSI).

Процесс формирования наблюдаемых данных можно описать следующими моделями деградации [17,18]:

$$H = D_s(Y) + n_h = E(Y) \downarrow_r + n_h,$$

$$M = D_\lambda(Y) + n_m = RY + n_m,$$

где $Y \in \mathbb{R}^{h_{\text{MSI}} \times w_{\text{MSI}} \times L_h}$ — искомое HQHSI высокого пространственного и спектрального разрешения; $D_s(Y)$ — оператор пространственной деградации, включающий размытие $E(Y)$ и понижение дискретизации (downsampling) \downarrow_r с коэффициентом $r = h_{\text{MSI}}/h = w_{\text{MSI}}/w$; D_λ — оператор спектральной

деградации, аппроксимируемый линейным преобразованием через матрицу спектрального отклика сенсора $R \in \mathbb{R}^{L_m \times L_h}$; n_h и n_m — аддитивные шумы.

При разработке методов глубокого обучения для задачи слияния HSI–MSI важным понятием является рецептивное поле (receptive field) — область входных данных, влияющая на вычисление конкретного выходного элемента. Увеличение рецептивного поля позволяет модели учитывать более широкий пространственный и спектральный контекст при формировании предсказания, что критично для точного восстановления спектральных характеристик и сохранения пространственных структур.

Ограничения съёмочной аппаратуры приводят к компромиссу между пространственным и спектральным разрешением. Получение HSI высокого пространственного разрешения напрямую сопряжено со значительными техническими и стоимостными трудностями. Методы слияния призваны преодолеть этот барьер путем математической обработки доступных HSI низкого разрешения (LR–HSI) и MSI высокого разрешения (HR–MSI) одной и той же сцены (рис. 1).

2.2. Описание используемых данных и их предобработка

В настоящем исследовании использовались синтетические данные, сгенерированные на основе реальных гиперспектральных изображений из набора CAVE (<http://www.cs.columbia.edu/CAVE/databases/multispectral/>). Обучение моделей осуществлялось на реальных сценах Washington DC Mall, Pavia Center и Botswana, а также набора Real HSI/MSI/PAN (<https://www.kaggle.com/datasets/shuangliangli123/real-hsimsipan-image-dataset>), тогда как тестирование выполнялось на синтетически сформированных данных на основе CAVE. В рассматриваемой конфигурации размеры низкоразрешенных гиперспектральных кубов составляют $64 \times 64 \times 76$ (пространственные размеры $h \times w$ и количество спектральных каналов L_h), а мультиспектральных изображений высокого разрешения — $192 \times 192 \times 8$ (пространственные размеры и количество каналов L_m), что соответствует коэффициенту масштабирования по пространству 3 и используется для контролируемых экспериментов с известным эталоном высокого разрешения.

Процесс генерации синтетических тестовых данных из набора CAVE включал следующие этапы деградации для имитации реалистичных условий наблюдения.

1. Пространственное размытие: к эталонным HQHSI применялся гауссовский фильтр по пространственным измерениям с параметром $\sigma_{\text{blur}} = 0.8$ для моделирования ограничений оптической системы и атмосферных эффектов.

2. Добавление шума: к размытым изображениям добавлялся аддитивный гауссовский шум с нулевым средним и стандартным отклонением $\sigma_{\text{noise}} = 0.8$ для имита-

ции инструментальных шумов сенсора. Итоговые значения обрезались к диапазону $[0,1]$.

3. Понижение пространственного разрешения: выполнялось понижение дискретизации с коэффициентом $r = 3$ методом усреднения по области (area interpolation) для получения LR–HSI размером $64 \times 64 \times 76$.

4. Спектральная деградация: для формирования HR–MSI применялось линейное спектральное преобразование через матрицу $R \in \mathbb{R}^{8 \times 76}$, моделирующую спектральный отклик мультиспектрального сенсора.

Предобработка данных включала следующие этапы, реализованные в предоставленном программном обеспечении.

1. Загрузка и нормализация: исходные данные HSI и MSI загружались и нормализовались методом min-max-нормализации к диапазону $[0,1]$ для обеспечения сопоставимости значений и стабильности последующих вычислений:

$$x_{\text{norm}} = \frac{(x - x_{\min})}{x_{\max} - x_{\min}}.$$

2. Обрезка изображений: для обеспечения совместности размеров и возможности нарезки на патчи (небольшие фрагменты изображения фиксированного размера, используемые для обучения нейронных сетей) исходные изображения HSI и MSI обрезались так, чтобы их пространственные размеры (высота и ширина в пикселях) были кратны размеру патча. Для MSI учитывался коэффициент пространственного масштабирования относительно HSI (64×64).

3. Нарезка на патчи: из подготовленных изображений формировались наборы обучающих патчей размером 64×64 для HSI и 192×192 для MSI без перекрытия, что способствует увеличению объема обучающей выборки и повышению устойчивости моделей.

4. Изменение размера HSI: для сопоставления с MSI, HSI низкого разрешения увеличивалось до размеров MSI с использованием бикубической интерполяции, что обеспечивает корректное сопоставление пространственных структур.

Данные этапы обеспечивают качественную подготовку исходных данных для подачи в модели слияния и способствуют повышению достоверности получаемых результатов.

2.3. Метрики оценки качества

Для количественной оценки качества восстановленных изображений используется набор метрик, покрывающих различные аспекты схожести с эталоном.

PSNR — метрика фотометрической точности, оценивающая среднеквадратичную ошибку амплитуд пикселей [19]:

$$\text{PSNR} = 10 \log_{10} \left(\frac{\text{MAX}_I^2}{1/(L_h N) \sum_b \sum_i (H_{b,i} - \hat{H}_{b,i})^2} \right),$$

где L_h — количество спектральных каналов, N — общее число пикселей, MAX_I — максимальное возможное

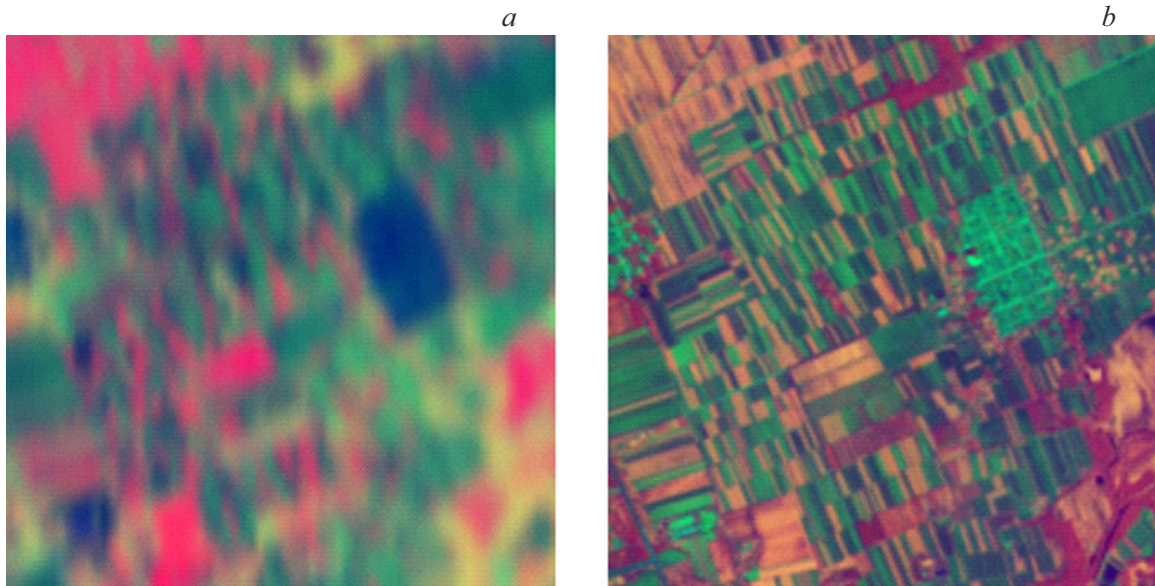


Рис. 1. Пример HSI низкого пространственного разрешения (a) и MSI высокого пространственного разрешения (b) одной и той же сцены (из датасета <https://www.kaggle.com/datasets/shuangliangli123/real-hsimsipan-image-dataset>). Для визуализации многоканальных изображений применен PCA: используются первые три главные компоненты, содержащие основную долю полезного сигнала, которые нормируются и отображаются в соответствующие каналы цветовой модели RGB с проекцией на первые три главные компоненты и отображением в соответствующее цветовое пространство.

значение пикселя, $H_{b,i}$ — эталонное значение пикселя i в канале b , $\hat{H}_{b,i}$ — восстановленное значение пикселя.

SSIM — метрика структурного сходства, оценивающая схожесть локальных патчей по яркости, контрасту и структуре [20]:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)},$$

где x, y — сравниваемые локальные фрагменты (патчи); μ_x, μ_y — средние значения яркости фрагментов x, y ; σ_x, σ_y — дисперсии (контраст) фрагментов x, y ; σ_{xy} — ковариация фрагментов x, y (мера структурного сходства); C_1, C_2 — малые константы для стабилизации деления.

SAM — метрика спектральной формы, измеряющая угол между спектральными векторами пикселей (в градусах) [21]:

$$SAM = \frac{180}{\pi} \frac{1}{N} \sum_i \arccos\left(\frac{\langle h_i, \hat{h}_i \rangle}{\|h_i\|_2 \|\hat{h}_i\|_2 + \varepsilon}\right),$$

где N — общее число пикселей в изображении, h_i — спектр эталонного пикселя i , \hat{h}_i — спектр восстановленного пикселя i , $\langle h_i, \hat{h}_i \rangle$ — скалярное произведение, $\|h_i\|_2$ — евклидова норма h_i , ε — малая константа для численной стабильности (предотвращения деления на ноль).

ERGAS — метрика относительной интегральной ошибки по каналам с учетом масштаба сигнала [22]:

$$ERGAS = 100 \frac{1}{d} \sqrt{\frac{1}{L_h} \sum_b \left(\frac{\sqrt{\frac{1}{N} \sum_i (H_{b,i} - \hat{H}_{b,i})^2}}{\mu_b} \right)^2},$$

где d — коэффициент понижения пространственного разрешения (коэффициент масштабирования); μ_b — среднее значение эталонного изображения в канале b .

Совместное использование (PSNR, SSIM, SAM, ERGAS) покрывает фотометрическое, структурное и спектральное качество восстановления.

2.4. Описание предлагаемых подходов

На основе анализа существующих решений и проведенных экспериментов были разработаны и исследованы два основных подхода к слиянию HSI и MSI.

Подход 1: Спектральное преобразование на основе NMF

Основы метода NMF. NMF представляет собой метод разложения неотрицательной матрицы $X \in \mathbb{R}_+^{m \times n}$ на произведение двух неотрицательных матриц меньшей размерности: $X \approx WH$, где $W \in \mathbb{R}_+^{m \times r}$ и $H \in \mathbb{R}_+^{r \times n}$, при этом ранг факторизации $r \ll \min(m, n)$. В контексте слияния изображений NMF интерпретируется следующим образом: матрица W представляет базисные спектральные профили (эндмемберы), а матрица H — коэффициенты

смешения этих профилей для каждого пикселя. Задача NMF формулируется как оптимизационная проблема [23,24]:

$$\min_{W, H \geq 0} \|X - WH\|_F^2,$$

где $\|\cdot\|_F$ обозначает норму Фробениуса. Неотрицательность факторов обеспечивает физическую интерпретируемость, соответствующую аддитивной природе спектральных измерений. Для задачи слияния HSI–MSI классический NMF находит глобальную матрицу спектрального преобразования W , связывающую пространство MSI и HSI.

Предлагаемое расширение. Этот подход направлен на поиск нелинейного преобразования, отображающего MSI в пространство HSI. Базовая линейная часть предполагает, что каждый спектральный канал HSI может быть аппроксимирован линейной комбинацией каналов MSI: $H_R \approx WM$, где $H_R \in \mathbb{R}^{L_h \times N}$ — гиперспектральное изображение, увеличенное до размеров MSI методом интерполяции и развернутое в матрицу (развертывание тензора $\mathbb{R}^{h_{\text{MSI}} \times w_{\text{MSI}} \times L_h}$ в матрицу $\mathbb{R}^{L_h \times N}$, где $N = h_{\text{MSI}} \times w_{\text{MSI}}$ — общее число пикселей, для данной конфигурации 76×36864), $M \in \mathbb{R}^{L_m \times N}$ — мультиспектральное изображение (развертывание тензора $\mathbb{R}^{h_{\text{MSI}} \times w_{\text{MSI}} \times L_m}$ в матрицу $\mathbb{R}^{L_m \times N}$, для данной конфигурации 8×36864) и $W \in \mathbb{R}^{L_h \times L_m}$ — искомая матрица спектрального преобразования (в данной реализации 76×8). Однако для учета нелинейных эффектов в спектральном отклике сенсоров и атмосферных искажений к линейному преобразованию добавляется квадратичная коррекция.

Для реализации подхода используется модуль WMatrixGenerator, генерирующий попиксельные (тензорные) матрицы спектрального преобразования W' и коэффициенты квадратичной коррекции, что представляет собой нелинейное обобщение глобального NMF на локально-адаптивное преобразование.

Архитектура модуля WMatrixGenerator. Данный модуль реализует генерацию пространственно-вариантной (попиксельной) матрицы спектрального преобразования W' и коэффициентов квадратичного полинома для отображения MSI в пространство HSI (рис. 2). Входом служит конкатенация интерполированного HSI низкого разрешения H_R и нормализованного MSI, M — количество каналов.

Спецификация модуля включает следующие элементы.

• **Вход:** $H_R \in \mathbb{R}^{\text{Batch} \times L_h \times h_{\text{MSI}} \times w_{\text{MSI}}}$ и $M \in \mathbb{R}^{\text{Batch} \times L_m \times h_{\text{MSI}} \times w_{\text{MSI}}}$ (приведенные к одному пространственному разрешению тензоры), где *Batch* — количество батчей.

• **Выход:** матрицы спектральных весов $W' \in \mathbb{R}^{\text{Batch} \times L_m \times h_{\text{MSI}} \times w_{\text{MSI}}}$ и коэффициенты квадратичного полинома $A, B, C \in \mathbb{R}^{\text{Batch} \times L_m \times h_{\text{MSI}} \times w_{\text{MSI}}}$.

• **Применение:** линейное преобразование $y_{\text{lin}} = W'M$ с последующей квадратичной коррекцией $\hat{y} = Ay_{\text{lin}}^2 + By_{\text{lin}} + C$.

Архитектура (обозначения: *Chan* — базовые каналы, d_{emb} — размерность позиционного кодирования) реализует следующую последовательность этапов.

• **Позиционное кодирование:** формируются нормализованные координатные сетки, которые через небольшую MLP преобразуются в позиционные представления $E \in \mathbb{R}^{\text{Batch} \times d_{\text{emb}} \times h_{\text{MSI}} \times w_{\text{MSI}}}$, добавляя модели информацию о местоположении пикселя.

• **Входная свертка:** конкатенация $[H_R, M]$ проходит Conv2d с ядром 3×3 и ReLU, формируя базовый тензор признаков F_0 .

• **Понижение дискретизации и расширение рецептивного поля:** к признакам добавляются позиционные представления, далее выполняются блоки с понижением разрешения и residual-обработкой (stride2 для переходов $\text{Chan} \rightarrow 2\text{Chan}$ и $2\text{Chan} \rightarrow 4\text{Chan}$) для аккумуляции контекста [25].

• **Повышение дискретизации:** два транспонированных сверточных слоя ($k = 4, s = 2, p = 1$) возвращают пространственное разрешение и уменьшают количество каналов ($4\text{Chan} \rightarrow 2\text{Chan} \rightarrow \text{Chan}$) без появления шахматных артефактов.

• **Заключительная обработка:** residual-блок и свертка 3×3 формируют сглаженное представление F_{ref} , после чего одномерная свертка позволяет найти параметры W' и (A, B, C) .

Нелинейное преобразование: для каждого пикселя выполняется линейная свертка спектра MSI:

$$y_{\text{lin}}[\text{batch}, l_h, h, w] = \sum_j W'[\text{batch}, l_h, j, h, w] \times M[\text{batch}, l_h, h, w], \quad (1)$$

где *batch* — номер батча, l_h — спектральный канал, h и w — высота и ширина в пикселях соответственно. Затем к y_{lin} применяется ограниченный квадратичный полином:

$$\hat{y}[\text{batch}, l_h, h, w] = A[\text{batch}, l_h, h, w]y_{\text{lin}}^2[\text{batch}, l_h, h, w] + B[\text{batch}, l_h, h, w]y_{\text{lin}}[\text{batch}, l_h, h, w] + C[\text{batch}, l_h, h, w], \quad (2)$$

где параметры A, B проходят масштабирование функцией гиперболического тангенса для численной устойчивости:

$$A \leftarrow 0.1 \tanh(\cdot), \\ B \leftarrow 1 + 0.5 \tanh(\cdot).$$

Итоговое изображение высокого разрешения формируется с остаточным соединением для повышения стабильности обучения:

$$Y_{\text{fake}} = y_{\text{lin}} + \hat{y}. \quad (3)$$

Свойства и связь с NMF: Попиксельные матрицы W' обеспечивают локальную адаптивность, квадратичная коррекция добавляет умеренную нелинейность, понижение дискретизации расширяет рецептивное поле. Линейная часть y_{lin} обобщает NMF на локальные пиксельные преобразования, где глобальная матрица W

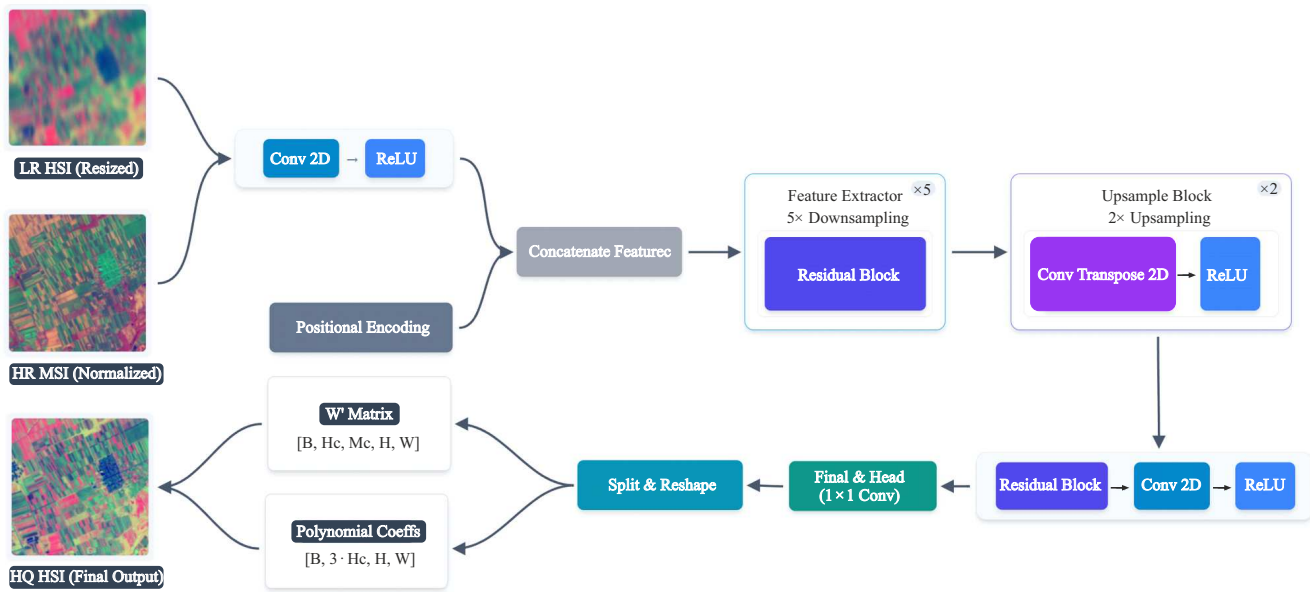


Рис. 2. Визуализация схемы WMatrixGenerator. Здесь Conv2D — фильтр, извлекающий пространственные и спектральные признаки из изображений путем свертки, ResidualBlock — блок с обходным соединением, позволяющий обучать глубокую сеть без потери информации об исходном сигнале, Positional Encoding — механизм добавления координат пикселя к данным для локальной адаптации спектрального преобразования, Concatenate Features — операция конкатенации признаков от HSI и MSI для их совместной обработки, Feature Extractor — модуль сжатия данных для анализа контекста и выделения главных характеристик сцены, Upsample Block (ConvTranspose2D) — блок восстановления высокого пространственного разрешения из сжатых признаков, Final Head (1 × 1 Conv) — слой, преобразующий накопленные признаки в конкретные значения параметров матриц, Split & Reshape — техническая операция разделения данных и изменения их размерности для корректного перемножения, W' Matrix/Polynomial Coeffs — выходные параметры сети (матрица и коэффициенты) для математического преобразования спектра MSI в HSI.

заменена адаптивными $W'_{h,w}$. Модуль может обучаться как в режиме принудительного обучения (с эталонными целями W_{NMF}), так и в режиме сквозного обучения совместно с цикловыми потерями.

Псевдообратный генератор W^+ (PseudoInverseGenerator). Симметричный модуль строит попиксельные псевдообратные матрицы для перехода из пространства HSI обратно в MSI и обучается совместно с W' по цикловой схеме ниже. Термин псевдообратность в описании модуля PseudoInverseGenerator относится к восстановлению линейной компоненты спектрального преобразования.

- **Входная свертка:** Conv2d($L_H \rightarrow C, 3 \times 3$) с ReLU.
- **Экстракция и апсэмплинг:** каскад residual-блоков с понижением размерности до $4C$ и последующим апсэмплингом ConvTranspose2d к исходному разрешению.
- **Финальный блок:** residual-блок и свертка 3×3 формируют сглаженное представление.
- **Выход:** Conv2d($C \rightarrow L_M L_H, 1 \times 1$) с масштабированием к тензору $W^+ \in \mathbb{R}^{B \times h_c \times L_M \times L_H \times h_{MSI} \times w_{MSI}}$.

Совместное обучение W' и W^+ в цикловой архитектуре уменьшает накопление ошибок при последовательных переходах MSI–HSI–MSI, обеспечивает взаимную регуляризацию и связывает метод с классической NMF-

парадигмой через локальные адаптивные матрицы преобразования.

В предложенной модели полное преобразование является композицией линейного отображения через прямую матрицу и аддитивной нелинейной коррекции полиномиальными членами. Обратное преобразование построено симметрично.

Обеспечение псевдообратимости реализуется на двух уровнях. На линейном уровне ортогональная функция потерь обеспечивает псевдообратность матрицы обратного преобразования по отношению к матрице прямого преобразования. В свою очередь это гарантирует корректное восстановление основной линейной структуры сигнала. На нелинейном уровне полная псевдообратимость отображения и компенсация квадратичных членов достигаются за счет обучения параметров обратного полинома под действием функции потерь цикловой согласованности.

Таким образом, генерируемая обратная матрица обеспечивает псевдообратность линейного базиса, в то время как нелинейные остатки инвертируются обучаемыми полиномиальными коэффициентами псевдообратного генератора в рамках сквозной оптимизации сети.

Подход 2: Архитектура HSIMSICycleGAN (цикловая модель с попиксельными спектральными ядрами)

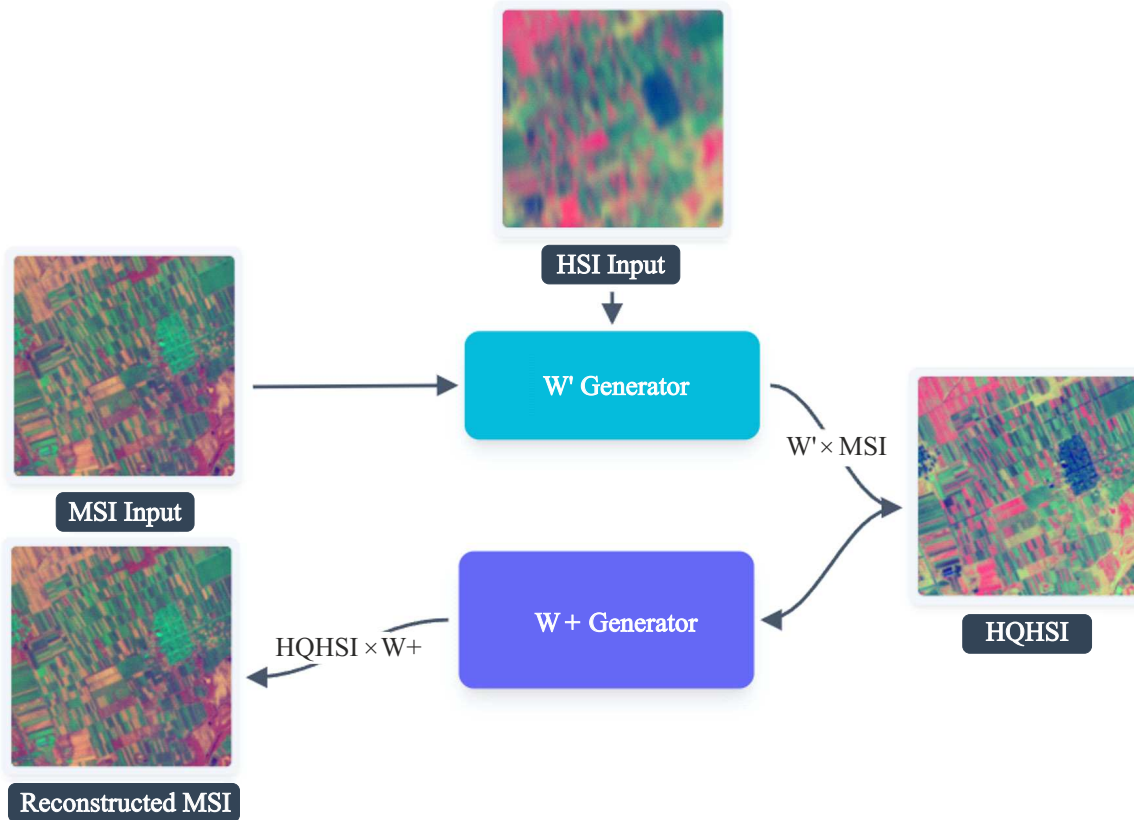


Рис. 3. Визуализация схемы архитектуры HSIMSiCycleGAN. Входные MSI и HIS подаются в блок W' Generator для генерации матрицы W' , далее осуществляется смешивание W' и MSI в HQHSI. Следующим шагом блок W^+ Generator создает матрицу W^+ и происходит ее смешивание с HQHSI для генерации итогового реконструированного MSI.

Второй подход реализуется в виде цикловой архитектуры HSIMSiCycleGAN (рис. 3), расширяющей классическую идею цикловой согласованности [26] применением (i) попиксельных спектральных матриц прямого и обратного преобразования, (ii) параметризованных квадратичных нелинейностей, (iii) позиционных представлений для повышения локальной адаптивности. Модель состоит из двух взаимно обратных отображений: $G_M \rightarrow H$, $G_H \rightarrow M$.

Прямое отображение. На вход подаются интерполированное HSI H_R и мультиспектральное изображение M . Генератор WMatrixGenerator формирует попиксельные матрицы W' и коэффициенты A, B, C , после чего выполняется линейное смешение спектров и квадратичная коррекция по формулам (1)–(3).

Обратное отображение. Модуль PseudoInverseGenerator принимает Y_{fake} и генерирует псевдообратные матрицы W^+ вместе с коэффициентами нелинейной поправки, что позволяет восстановить MSI и замкнуть цикл:

$$x_{\text{lin}}[\text{batch}, l_h, h, w] = \sum_{i=1}^{L_H} W^+[\text{batch}, l_h, i, h, w] \times Y_{\text{fake}}[\text{batch}, i, h, w],$$

$$\begin{aligned} \dot{x}[\text{batch}, l_h, h, w] &= A'[\text{batch}, l_h, h, w] \\ &\times (x_{\text{lin}}[\text{batch}, l_h, h, w])^2 + B'[\text{batch}, l_h, h, w] \\ &\times x_{\text{lin}}[\text{batch}, l_h, h, w] + C'[\text{batch}, l_h, h, w], \\ X_{\text{rec}} &= x_{\text{lin}} + \dot{x}. \end{aligned}$$

В обоих генераторах используются позиционные кодировки, конкатенируемые с признаками перед блоками с понижением и повышением дискретизации, что увеличивает эффективное рецептивное поле при сохранении локальной адаптивности.

Ортогональность прямого и обратного преобразований обеспечивается дополнительной регуляризацией, минимизирующей разность между произведением W^+W' и единичной матрицей для каждого пикселя:

$$L_{\text{ortho}} = \|W^+W' - I\|_1.$$

Функция потерь включает цикловую согласованность, спектральную и пространственную компоненты деградации, градиентное согласование, SAM-потерю, SSIM-потерю, реконструкционную составляющую и регуляризацию норм матриц, что отражено в табл. 3.

Схема обучения. Обучение проводится в три стадии по 75 эпох каждая. Предварительно WMatrixGenerator

Таблица 3. Компоненты функции потерь и веса на трех стадиях обучения

Компонента	Краткая формула	Назначение	Стадия 1	Стадия 2	Стадия 3
Цикл	$\ M - X_{\text{rec}}\ _1$	Согласованность MSI→HSI→MSI	1.0	1.0	1.0
Спектральная (HSI→MSI)	$\ Conv_{1 \times 1}(Y_{\text{fake}}) - M\ _1$	SRF-аппроксимация	1.0	1.0	1.0
Пространственная (HR→LR)	$\ down_{area}(blur(Y_{\text{fake}})) - H_R\ _1$	Согласование деградации	1.0	1.0	1.0
Градиентная	$\sum_d \ \partial_d(\tilde{Y}_{\text{fake}}) - \partial_d(\tilde{M})\ _1$	Сохранение границ	1.0	1.0	1.0
SAM	$SAM(Y_{\text{fake}}, up_{\text{bicubic}}(H_R))$	Спектральная форма	1.0	5.0	5.0
SSIM	$1 - SSIM(Conv_{1 \times 1}(Y_{\text{fake}}), M)$	Локальная структурная близость	2.0	2.0	2.0
Recon (HR–HSI)	$\ Y_{\text{fake}} - up(H_R)\ _1$	Фотометрическая реконструкция	0.0	5.0	5.0
Спектральная гладкость	$\ Y_{\text{fake}}[:, 1, :] - Y_{\text{fake}}[:, :, -1]\ _1$	Сглаживание спектров	0.0	0.0	0.0
Распределительная	$W_1(hist(Y_{\text{fake}}), hist(up(H_R)))$	Согласование гистограмм каналов	1.0	1.0	1.0
Ортогональность	$\ W^+W' - I\ _1$	Псевдообратность локальных матриц	0.1	0.1	0.1
Регуляризация	$\ W'\ _2^2 + \ W^+\ _2^2$	Контроль норм веса	0.01	0.01	0.01

инициализируется из линейной аппроксимации, после чего все блоки обучаются совместно.

Стадия 1 (эпохи 1–75). Используются веса $\lambda_{\text{spectral}} = \lambda_{\text{spatial}} = \lambda_{\text{grad}} = \lambda_{\text{SAM}} = 1$, $\lambda_{\text{SSIM}} = 2$ при отсутствующей реконструкционной компоненте ($\lambda_{\text{recon}} = 0$). Скорость обучения равна $1e-4$ и обеспечивает стабилизацию линейного отображения.

Стадия 2 (эпохи 76–150). Добавляется реконструкционная потеря по HSI ($\lambda_{\text{recon}} = 5$) и усиливается спектрально-угловой штраф ($\lambda_{\text{SAM}} = 5$). Остальные веса сохраняются, шаг обучения уменьшается до $1e-5$.

Стадия 3 (эпохи 151–225). Весовые коэффициенты совпадают со второй стадией, однако шаг оптимизации увеличивается до $3e-5$ для деликатной донастройки и сохранения величины SAM.

На всех стадиях используется разбиение данных на обучающую, проверочную и тестовую выборки без пересечений по сценам. Ранняя остановка контролируется по проверочным SAM и SSIM, а малые веса спектральной гладкости и распределительной потери поддерживают стабильность обучения.

3. Результаты и обсуждение

3.1. Сравнительный анализ качества работы существующих методов

Для оценки эффективности предложенных методов проведено количественное сравнение с современными подходами к слиянию HSI–MSI. В табл. 4 представлены значения метрик PSNR, SSIM, SAM и ERGAS для различных методов, включая предложенные в настоящей работе подходы.

По численным метрикам наблюдается последовательное повышение PSNR и SSIM и снижение ERGAS при

переходе от линейных моделей к гибридным и GAN-/attention-архитектурам. Средние значения по группам методов можно обобщить следующим образом.

- Линейные методы (NMF, PCA): средний PSNR ≈ 18.2 dB, SSIM ≈ 0.40 , SAM $\approx 12.6^\circ$, ERGAS ≈ 3.34 — базовый уровень качества с высокой интерпретируемостью.

- Регуляризационные и разреженные методы (HySure, NSSR): средний PSNR ≈ 29.1 dB, SSIM ≈ 0.87 , SAM $\approx 15.4^\circ$, ERGAS ≈ 2.77 — заметное улучшение структурной согласованности при росте вычислительной сложности.

- Тензорные модели (CSTF, Deep Tucker, UDTN): средний PSNR ≈ 34.5 dB, SSIM ≈ 0.94 , SAM $\approx 7.8^\circ$, ERGAS ≈ 1.30 — значительный прирост за счет использования многомерной фактурности.

- Современные подходы на основе глубокого обучения (MIAE, HSR–KAN): средний PSNR ≈ 40.3 dB, SSIM ≈ 0.98 , SAM $\approx 5.5^\circ$, ERGAS ≈ 1.20 — максимальные показатели при высокой ресурсоемкости и требованиях к данным.

Критически малые значения SAM достигаются у моделей с явной структурной регуляризацией (RAF–NLRGS: 1.063°) и гибридных архитектур с тензорными компонентами (Deep Tucker: 4.23° , HSR–KAN: 3.17°). Предложенные в работе подходы дополняют существующий спектр решений. HSI-MSCycleGAN демонстрирует промежуточные результаты (PSNR = 25.30 dB, SSIM = 0.8133, SAM = 13.51°), превосходя базовые методы и уступая передовым гибридным архитектурам. Базовый WMatrixGenerator пока ограничен (PSNR = 11.72 dB, SSIM = 0.2185, SAM = 31.76°), что указывает на необходимость дальнейшей регуляризации и доработки архитектуры. Современные методы глубо-

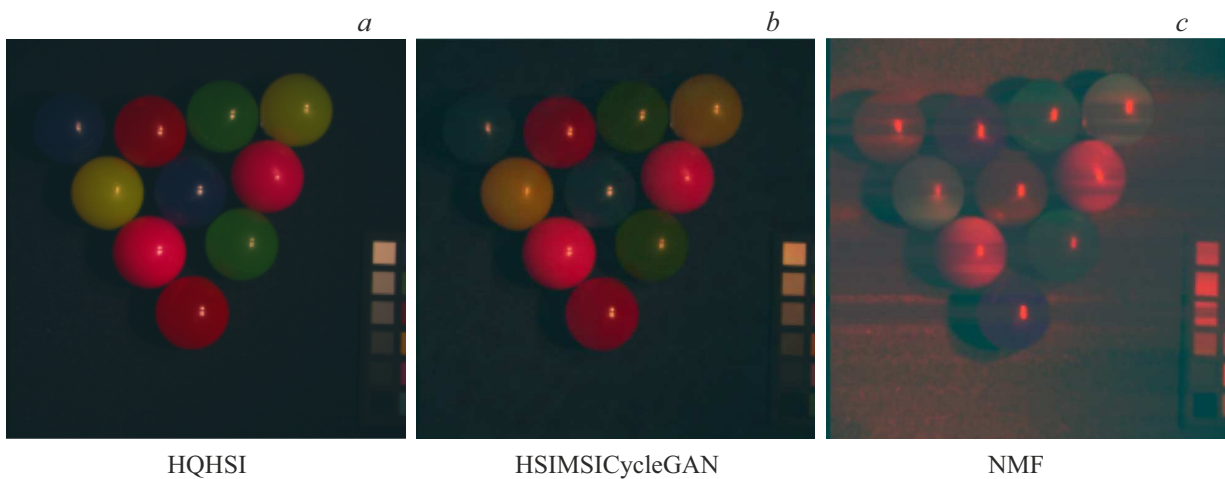


Рис. 4. Пример сцены из синтетически сгенерированного набора данных на основе CAVE: (a) HQHSI (высококачественное гиперспектральное изображение), (b) восстановленное высоко-разрешённое HSI, полученное архитектурой HSIMSiCycleGAN, (c) результат метода на основе NMF. Для визуализации гиперспектральных изображений (a,b,c) применен метод главных компонент (PCA) с проекцией на первые три главные компоненты и последующим отображением в RGB.

Таблица 4. Сравнительные метрики качества для различных методов повышения разрешения HSI–MSI (\uparrow означает, что большее значение лучше, \downarrow — меньшее значение лучше)

Метод	PSNR \uparrow	SSIM \uparrow	SAM \downarrow	ERGAS \downarrow
NMF	17.17	0.3865	13.32	3.56
PCA	19.23	0.4217	11.87	3.12
CSTF	28.55	0.904	12.26	1.94
HySure	27.63	0.841	16.63	4.01
NSSR	30.59	0.899	14.17	1.53
Deep Tucker	35.18	0.938	4.23	1.40
MIAE	35.58	0.927	10.38	0.98
uSDN	37.74	0.974	9.28	0.70
HyCoNet	38.72	0.966	6.90	1.94
Physics-Based GAN	38.97	0.990	5.53	1.20
UDTN	39.8	0.982	6.78	0.55
MossFuse	42.05	0.990	6.57	0.42
BliEstGAN	43.9	0.990	6.89	—
RAF-NLRGS	43.57	0.991	1.063	2.075
HSR-KAN	44.89	0.991	3.17	2.04
Предложенные методы				
NMF	11.72	0.2185	31.76	—
HSIMSiCycleGAN	25.30	0.8133	13.51	—

кого обучения (MossFuse, HSR–KAN, RAF–NLRGS) обеспечивают наилучшие количественные показате-

ли и подтверждают эффективность локально адаптивных нелинейных преобразований. Предложенный HSIMSiCycleGAN обеспечивает разумный компромисс между качеством и доступностью данных, полагаясь на парные наблюдения HSI/MSI без эталонного HQHSI (рис. 4, 5).

В целом прослеживается устойчивая тенденция повышения качества при переходе от глобальных линейных моделей к локально адаптивным нелинейным архитектурам, сочетающим физические знания и обучаемые компоненты.

3.2. Обсуждение результатов

Количественная оценка предложенных методов показала различия в их эффективности. HSIMSiCycleGAN достигает умеренного качества реконструкции (PSNR = 25.30 dB, SSIM = 0.7546) при относительно хорошем сохранении спектральной характеристики (SAM = 14°), что позиционирует метод как перспективный для задач, где доступны синхронные наблюдения HSI/MSI, но отсутствует эталонное HQHSI. Модуль WMatrixGenerator в базовой конфигурации демонстрирует ограниченную эффективность (PSNR = 11.72 dB, SSIM = 0.2185, SAM = 31.76°), указывая на необходимость интеграции более сложных регуляризационных механизмов.

Используемая архитектура сочетает локально-адаптивные матрицы, умеренную нелинейность и цикловую согласованность, обеспечивая интерпретируемость преобразований W' и W^+ и улучшая баланс между фотометрическими, структурными и спектральными метриками.

В задачах анализа изображений интерпретируемость выступает ключевым фактором, позволяющим верифицировать физическую правдоподобность восстановлен-

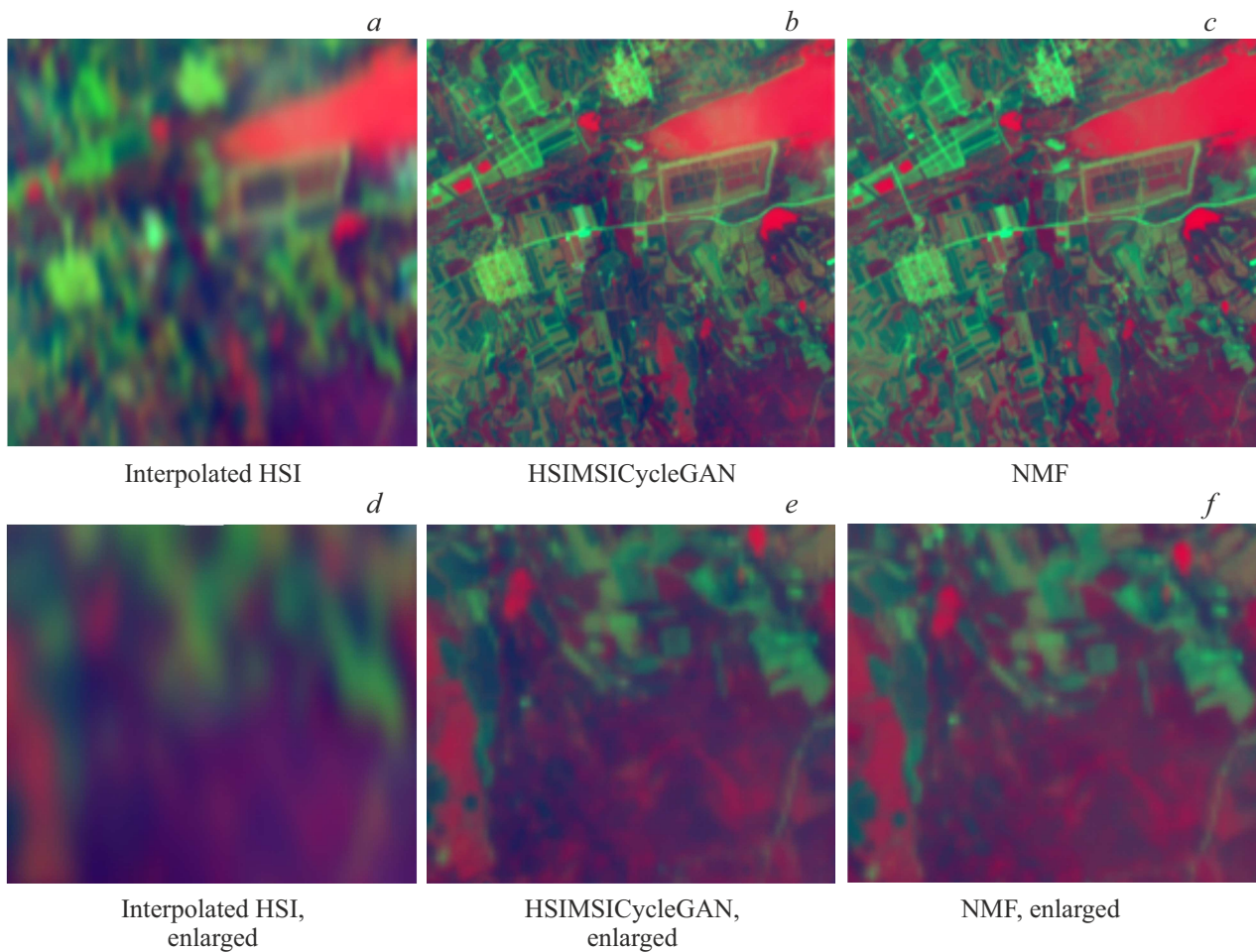


Рис. 5. Визуальное сравнение результатов слияния (*a,b,c*) и соответствующих увеличенных фрагментов (*d,f,g*). Слева направо: интерполированное HSI (192×192 , из Real HSI/MSI/PAN), результат HSIMSiCycleGAN (192×192) и результат метода на основе NMF (192×192).

ных данных. Она гарантирует, что результат является следствием математически обоснованного преобразования измерений, а не артефактом работы непрозрачных вычислительных процедур [27].

В подходе на основе NMF данная интерпретируемость обеспечивается строгим соответствием модели линейного смешения с ограничениями неотрицательности ($W, H \geq 0$), где компоненты несут прямой физический смысл спектральных сигнатур и их концентраций.

В рамках предложенной архитектуры HSIMSiCycleGAN жесткое ограничение неотрицательности для генерируемых матриц W' ослабляется. Это позволяет моделировать сложные зависимости и эффекты спектральной коррекции, выходящие за рамки простейшей аддитивной модели. Преимущество данного подхода заключается в реализации механизма спектрального вычитания (spectral subtraction): использование отрицательных весов допускается для подавления паразитного сигнала от перекрывающихся широких каналов MSI при восстановлении узких полос HSI. Такой механизм обеспечивает более

высокую фотометрическую точность спектральной реконструкции, потенциально недостижимую при строгих ограничениях неотрицательности.

Матрица W' в данном случае интерпретируется не как набор статичных базисных спектральных профилей, а как динамический оператор спектральной реконструкции, адаптирующийся к локальному контексту сцены. Сохранение матричной структуры преобразования позволяет напрямую отслеживать вклад каждого канала MSI в формирование итогового спектра, что обеспечивает прозрачность модели для анализа, в отличие от полностью скрытых слоев глубоких нейронных сетей.

Достоинства проведенного исследования

- Проведен обзор современных методов слияния HSI–MSI.
- Предложены, реализованы и предварительно исследованы два новых подхода: попиксельное спектральное преобразование на основе NMF и архитектура HSIMSiCycleGAN.

- Описаны практические процедуры подготовки данных и обучения, что облегчает воспроизводимость результатов.

Недостатки и ограничения

- Полученные оценки носят предварительный характер и требуют расширенного сравнения с большим числом передовых методов на различных наборах данных.
- Обучение нейросетевых моделей, особенно HSIMSI-CycleGAN, сопровождается высокой вычислительной стоимостью.
- Необходим тщательный подбор гиперпараметров и архитектурных компонентов для устойчивой сходимости.
- Следует отметить, что обучение на патчах может приводить к артефактам при итоговом слиянии, однако в настоящей работе этот вопрос не исследовался.

Перспективы развития

Планируется масштабная экспериментальная проверка на разнообразных наборах данных с различными пространственными и спектральными характеристиками, а также исследование альтернативных механизмов внимания и стратегий регуляризации. Интерес представляют интеграция физически обоснованных моделей формирования изображений, разработка легких архитектур для встроенных систем и использование трансферного обучения для быстрой адаптации к новым сенсорным платформам.

Область применения данной технологии включает как повышение разрешения гипер- и мультиспектральных изображений, полученных со спутников, БПЛА, так и широкое применение в биофотонике [28]. Так, в работе [29] нейросетевые методы реконструкции суперразрешения используются для построения предиктивных моделей диагностики заболеваний растений на основе гиперспектральных изображений. Кроме того, гиперспектральная визуализация кожи человека является потенциальным инструментом диагностики неоплазм [30–32]. Увеличение разрешения таких изображений при помощи RGB-изображений [33] также позволит существенно улучшить разрешение HSI.

4. Заключение

В работе представлен анализ современных методов слияния гиперспектральных и мультиспектральных изображений и два новых подхода к решению данной задачи. Экспериментальная оценка показала, что архитектура HSIMSI-CycleGAN демонстрирует умеренные результаты ($PSNR = 25.30$ dB, $SSIM = 0.7546$, $SAM = 14^\circ$), превосходя классические методы, но уступая современным передовым решениям. Ключевым преимуществом является возможность обучения на парных наблюдениях HSI/MSI без использования эталонных данных высокого

пространственного разрешения, что важно для практических приложений дистанционного зондирования.

Результаты подтверждают тенденцию превосходства методов глубокого обучения над классическими подходами при важности баланса между точностью и вычислительной эффективностью. Работа создает основу для дальнейших исследований в области неконтролируемого слияния многомодальных спутниковых данных.

Выбор метода должен основываться на специфике задачи и доступных ресурсах. Для быстрых приложений с умеренными требованиями к качеству рекомендуются классические методы (NMF, PCA). При наличии парных наблюдений HSI низкого разрешения и MSI высокого разрешения, но без эталонного HQHSI, целесообразно применение цикловых архитектур типа HSIMSI-CycleGAN. Для достижения максимального качества при наличии обширных размеченных данных предпочтительны современные глубокие методы с физическими ограничениями или гибридные архитектуры.

Финансирование работы

Исследование выполнено при финансовой поддержке Министерства науки и высшего образования Российской Федерации (грант № 075-15-2024-557 от 25.04.2024 г.).

Конфликт интересов

Авторы заявляют, что у них нет конфликта интересов.

Список литературы

- [1] *Hyperspectral Data Exploitation: Theory and Applications*, ed. by C.I. Chang (John Wiley & Sons, 2007). DOI: 10.1002/9780470124628.ch3
- [2] T. Adão, J. Hruška, L. Pádua, J. Bessa, E. Peres, R. Morais, J.J. Sousa. *Remote Sensing*, **9** (11), 1110 (2017). DOI: 10.3390/rs9111110
- [3] G. Avola, A. Matese, E. Riggi. *Remote Sensing*, **15** (7), 1917 (2023). DOI: 10.3390/rs15071917
- [4] S. Du, Y. Zou, Z. Wang, X. Li, Y. Li, C. Shang, Q. Shen. *arXiv:2412.04802*, (2024). DOI: 10.48550/arXiv.2412.04802
- [5] J. Yang, L. Xiao, Y.-Q. Zhao, J.C.-W. Chan. *IEEE Transactions on Neural Networks and Learning Systems*, **35** (9), 13017–13031 (2024). DOI: 10.1109/TNNLS.2023.3266038
- [6] J. Liu, Z. Wu, L. Xiao, X.-J. Wu. *IEEE Transactions on Geoscience and Remote Sensing*, **60**, 1–12 (2021). DOI: 10.1109/TGRS.2022.3143156
- [7] K. Zheng, L. Gao, W. Liao, D. Hong, B. Zhang, X. Cui. *IEEE Transactions on Neural Networks and Learning Systems*, **59** (3), 2487–2502 (2021). DOI: 10.1109/TGRS.2020.3006534
- [8] J. Yao, D. Hong, J. Chanussot, D. Meng, X. Zhu, Z. Xu. *European Conference on Computer Vision* (Cham: Springer International Publishing, 2020), p. 208–224. DOI: 10.1007/978-3-030-58526-6_13
- [9] Y. Qu, H. Qi, C. Kwan. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2511–2520 (2018). DOI: 10.48550/arXiv.1904.12175

- [10] S. Li, R. Dian, L. Fang, J.M. Bioucas-Dias. IEEE Transactions on Image Processing, 2018. DOI: 10.1109/TIP.2018.2836307
- [11] M. Simões, J. Bioucas-Dias, L.B. Almeida, J. Chanussot. IEEE Transactions on Geoscience and Remote Sensing, 2015. DOI: 10.1109/TGRS.2014.2375320
- [12] W. Dong, F. Fu, G. Shi, X. Cao, J. Wu, G. Li, X. Li. IEEE Transactions on Image Processing, 2016. DOI: 10.1109/TIP.2016.2542360
- [13] J. Yang, L. Xiao, Y. Zhao, J.C.-W. Chan. IEEE Transactions on Neural Networks and Learning Systems. **35** (9), 13017–13031 (2023). DOI: 10.1109/TNNLS.2023.3266038
- [14] B. Li, X. Wang, H. Xu. arXiv:2409.06705, 2024. DOI: 10.48550/arXiv.2409.06705
- [15] H. Wang, Y. Xu, Z. Wu, Z. Wei. IEEE Transactions on Neural Networks and Learning Systems, 2024. DOI: 10.1109/TNNLS.2024.3457781
- [16] K. Yang, M. Bai, T. Lu. arXiv preprint arXiv:2407.05279, 2024. DOI: 10.48550/arXiv.2407.05279
- [17] L. Wald. *Data Fusion: Definitions and Architectures: Fusion of Images of Different Spatial Resolutions* (Presses des MINES, 2002).
- [18] R.A. Schowengerdt. *Remote Sensing: Models and Methods for Image Processing*, 3rd ed. (Academic Press, London, 2006).
- [19] A. Hore, D. Ziou. 20th International Conference on Pattern Recognition. IEEE, 2366–2369 (2010). DOI: 10.1109/ICPR.2010.579
- [20] Z. Wang, A.C. Bovik, H.R. Sheikh, E.P. Simoncelli. IEEE Transactions on Image Processing, **13** (4), 600–612 (2004). DOI: 10.1109/TIP.2003.819861
- [21] R.H. Yuhas, A.F.H. Goetz, J.W. Boardman. JPL, Summaries of the Third Annual JPL Airborne Geoscience Workshop, vol. 1: AVIRIS Workshop (1992).
- [22] K. Kolikov, G. Krastev, Y. Epitropov, D. Hristozov. Chemometrics and Intelligent Laboratory Systems, **102** (1), 15–19 (2010). DOI: 10.1016/j.chemolab.2010.03.001
- [23] G.H. Golub, C.F. Van Loan. *Matrix Computations*, 4th ed. (Baltimore, MD, 2013). DOI: 10.56021/9781421407944
- [24] G. Strang. *Introduction to Linear Algebra* (Wellesley-Cambridge Press, 2022). DOI: 10.1137/1.9781733146678
- [25] K. He, X. Zhang, S. Ren, J. Sun. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, p. 770–778 (2016). DOI: 10.1109/CVPR.2016.90
- [26] J.-Y. Zhu, T. Park, P. Isola, A. Efros. Proceedings of the IEEE International Conference on Computer Vision, 2223–2232 (2017). DOI: 10.1109/ICCV.2017.244
- [27] Z. Salahuddin, H.C. Woodruff, A. Chatterjee, P. Lambin. Computers in Biology and Medicine, **140**, 105111 (2022). DOI: 10.1016/j.combiomed.2021.105111
- [28] M. Yako. Optical Review, 1–13 (2025). DOI: 10.1007/s10043-025-01001-x
- [29] M.M. Antony, C.S.S. Sandeep, H.-T. Lim, M.V. Matham. J. Biomedical Photonics & Engineering, **9** (3), 030304 (2023). DOI: 10.18287/JBPE23.09.030304
- [30] I.A. Bratchenko, V.P. Sherendak, O.O. Myakinin, D.N. Artemyev, A.A. Moryatov, E. Borisova, L. Avramov, L.A. Zherdeva, A.E. Orlov, S.V. Kozlov, V.P. Zakharov. J. Biomedical Photonics & Engineering, **4** (1), 5–12 (2018). DOI: 10.18287/JBPE17.04.010301
- [31] B.V. Grechkin, V.O. Vinokurov, Y.A. Khristoforova, I.A. Matveeva. J. Biomedical Photonics & Engineering, **9** (4), 50–55 (2023). DOI: 10.18287/JBPE23.09.040304
- [32] I.A. Bratchenko, O.O. Myakinin, V.P. Sherendak, P.N. Volkhin, Y.A. Khristoforova, L.A. Bratchenko, D.N. Artemyev, A.A. Moryatov, O.V. Polschikova, A.S. Machikhin, V.E. Pozhar, S.V. Kozlov, V.P. Zakharov. J. Biomedical Photonics & Engineering, **4** (4), 040301 (2018). DOI: 10.18287/JBPE18.04.040301
- [33] Z. Lai, Y. Fu, J. Zhang. IEEE Transactions on Neural Networks and Learning Systems, 2024. DOI: 10.1109/TNNLS.2023.3340561