

Двухэтапный подход к повышению качества спутниковых изображений на основе нейросетевых алгоритмов

Козуб В.А., Мурынин А.Б., Швед А.К.

Введение: Актуальность качества ДЗЗ

Высокое качество спутниковых снимков критично для детального анализа.

(Например, точное распознавание городских объектов требует высокого разрешения).

Разные сенсоры дают разный стиль изображений (разрешение, шум, контраст, битность).

Проблема: «Разрыв доменов» между снимками разных спутников мешает совместной обработке и снижает точность алгоритмов.

Задача и предлагаемое решение

Цель: Повысить разрешение и качество снимков, устранив расхождение между доменами сенсоров.

Идея: **ДВУХЭТАПНЫЙ ПОДХОД:** Перенос стиля + Суперразрешение (SR).

Этап 1 (Адаптация): Приводим LR-изображение (сенсор A) к стилю целевого датчика (сенсор B).

Этап 2 (SR): Повышаем разрешение и восстанавливаем детали на стилизованном изображении.

Последовательность методов даёт синергетический эффект, решая обе подзадачи поочерёдно.

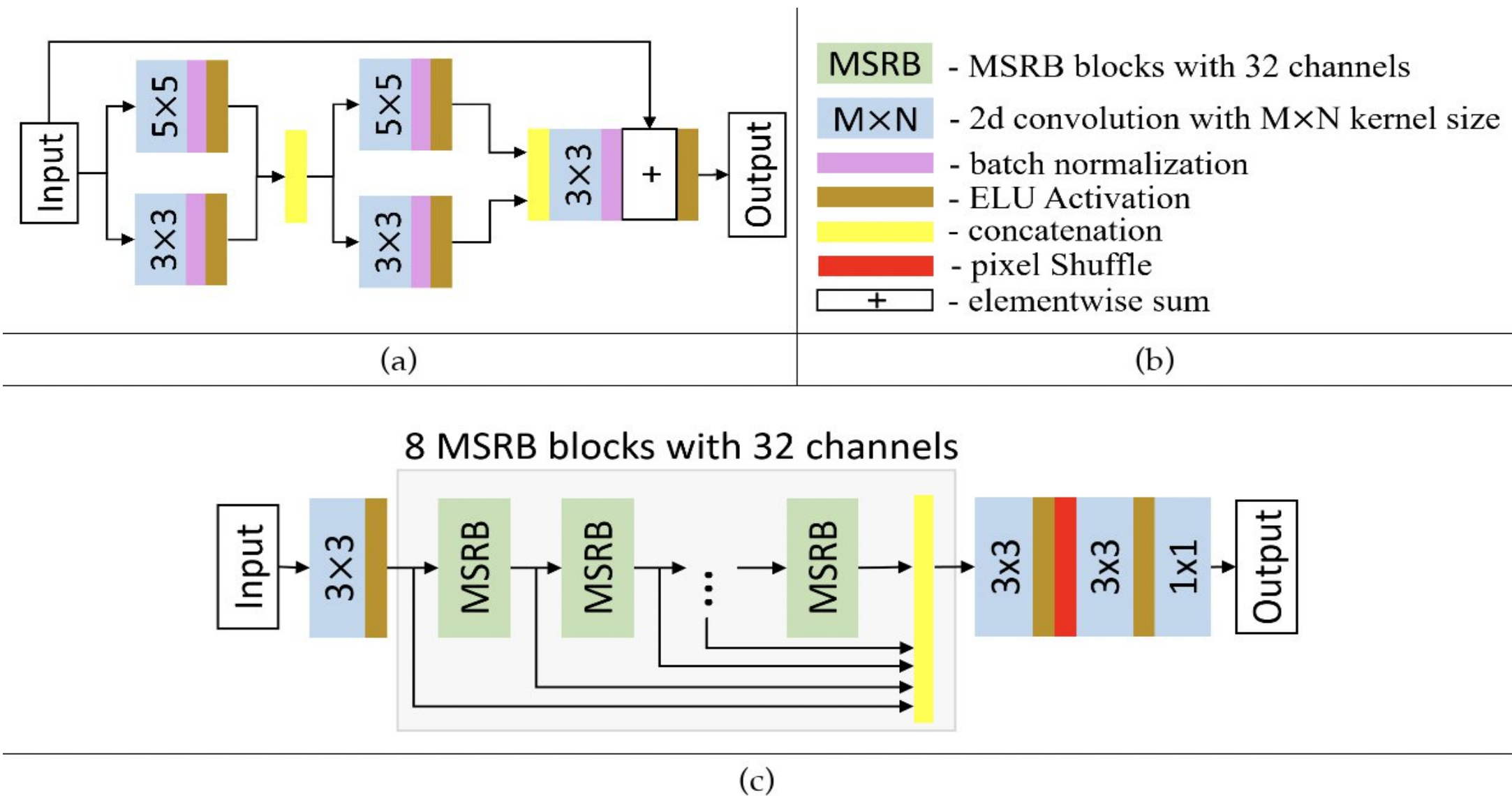


Рис. - Архитектура СНС сверхразрешения. (a) – структура Multi-Scale Residual Block, (b) – используемые условные обозначения (c) – архитектура ИНС сверхразрешения.

Ключевые преимущества двухэтапного подхода

Без парных данных: Не нужны пары «LR→HR» одной сцены для обучения — используем **несопряжённые выборки**.

Адаптация к сенсору B: После Этапа 1 изображение статистически соответствует целевому домену (контраст, шум), поэтому SR работает в «правильном» пространстве.

Устранение Domain Shift при SR: Обычная SR-модель, обученная на одном сенсоре, теряет качество на другом. Наш подход **полностью устраняет** этот сдвиг.

Этап 1: Доменная адаптация (Перенос стиля)

Используем **GAN (Generative Adversarial Network)** для перевода стиля.

Архитектура: Генератор — ResNet, Дискриминатор — PatchGAN.

Обучение: Без учителя, на несопряжённых наборах (LR A vs. деградированные HR B).

Результат: Изображение с сохранением содержимого исходного LR, но **по стилю — как данные сенсора B**.

Схема GAN для переноса стиля

Вход LR (A) → Генератор (GAN)
→ Выход LR (Стиль B)

Дискриминатор сравнивает выход с эталонным распределением LR (B).

Сохранение структуры при переносе стиля

Проблема: GAN может исказить структуру объектов, «фантазируя».

Решение (1): Добавляем **структурный лосс F-LSeSim** (Spatially-Correlative Loss).

Идея F-LSeSim: Сохранять внутренние корреляции между частями изображения, чтобы дорога осталась дорогой. (Контрастивный лосс на самоподобие: меняем только низкоуровневый стиль).

Решение (2): L1-регуляризация по интенсивности. Штрафует большие изменения яркости/цвета, предотвращая ложные артефакты.

Этап 2: Сверхразрешение (Super-Resolution, x3)

Отдельная нейросеть для увеличения разрешения (фактор **x3**).

Архитектура: Компактная CNN с несколькими мультискейл-Residual блоками.

Декодер: Слой **PixelShuffle** для повышения размера без блочных артефактов (sub-pixel convolution).

Обучение: На парах вход LR_B (стилизованное) \rightarrow цель HR_B (оригинал).

Функция потерь: L1 (MAE) между выходом и HR (для стабильности и точности).

Данные для обучения (Сенсоры А и В)

Сенсоры

Сенсор А (Низкокачественный): ~2 м/пикс, 8-bit, заметный шум. (Аналог «Канопус-В»)

Сенсор В (Высококачественный): ~0.7 м/пикс, 10-bit, высокий контраст. (Аналог «Ресурс-П»)

Подготовка и обучение

Собраны **несопряжённые выборки** (разные регионы) для А и В.

Создание LR_B: HR изображения В деградированы (downscale + шум) до качества А.

Обучение GAN (Этап 1): LR_A сравнение с распределением LR_B.

Обучение SR (Этап 2): LR_B в HR_B.

Метрики качества (безэталонные)

PSNR/SSIM: Используются только для справки на синтетических данных (нет настоящего HR для теста).

FID (Fréchet Inception Distance): Снижен после переноса стиля — **домены сближены**.

NIQE/BRISQUE (No-Ref): Изображения более «естественные» (меньше искажений).

Резкость (BREN): Выросла относительно исходных LR и SR-без-стиля.

Контраст (VAR, GLLV): Вырос — изображения стали менее блеклыми, лучше распределение градиентов.

Вывод: Полный метод (Стиль + SR) > Только SR > Исходные LR.

Визуальные результаты (Фрагменты сцен)

**1. Исходный LR (Сенсор
A)**

**2. Только SR (Ошибка
домена)**

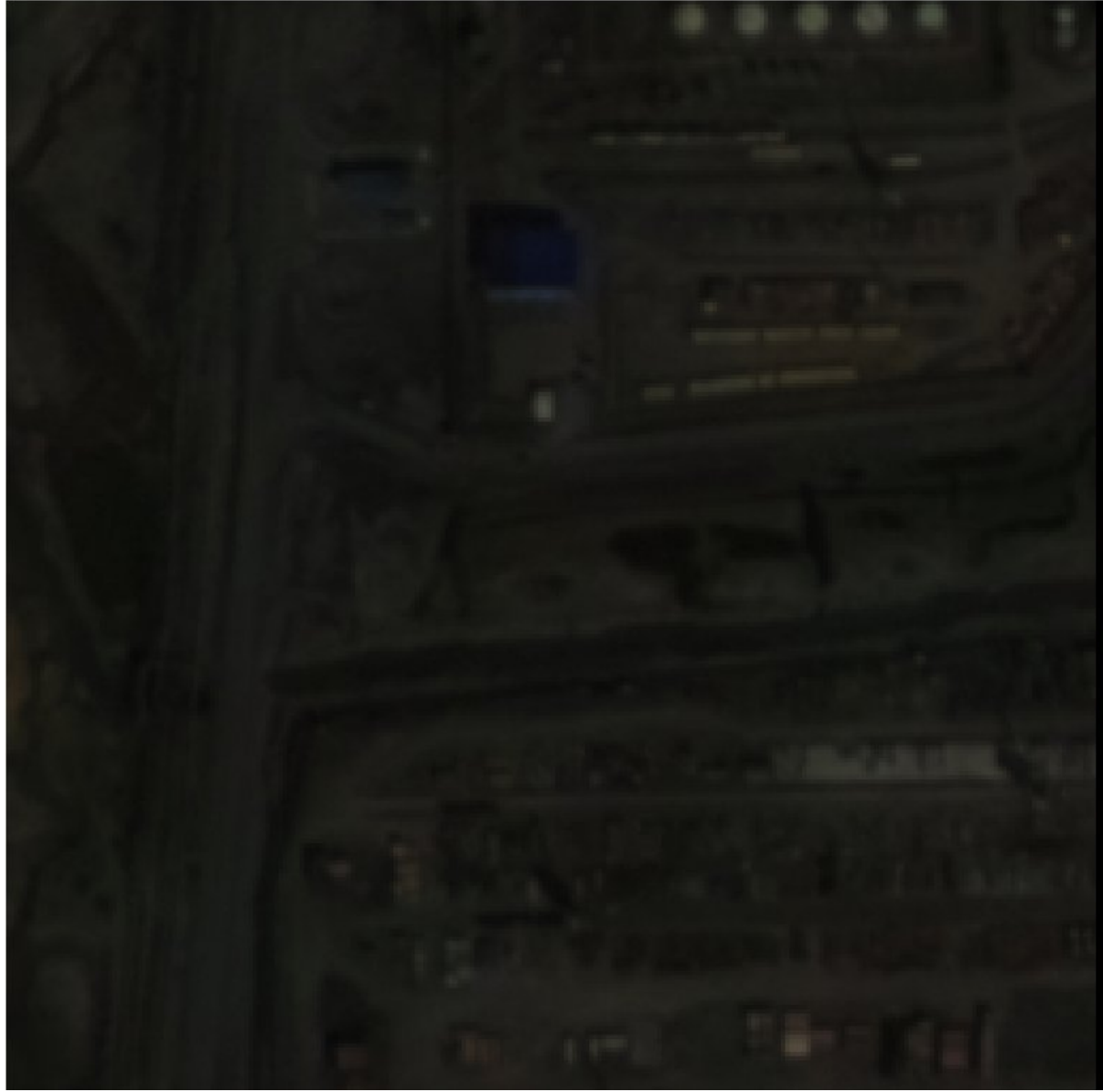
**3. Наш метод (Стиль +
SR)**

Примеры работ алгоритма



Можно заметить, что получаемые изображения обладают намного более высокой контрастностью и высоким семантическим соответствием (т.е. объекты одного класса не переходят в объекты другого класса).





Ограничения и риски

Перенос нежелательных особенностей: Генератор может перенести и недостатки целевого сенсора В (например, его собственный шум).

Локальные искажения от GAN: В сложных сценах возможно небольшое искажение мелких объектов или текстур. (Снижено с помощью структурного лосса, но полностью исключить артефакты нельзя).

Не универсальность модели: Обучена на конкретные домены. При смене сенсора требуется перенастройка.

Пути улучшения (Roadmap)

End-to-end модель: Объединить Этап 1 и 2 в единую сквозную сеть для ускорения применения.

Диффузионные модели для SR: Заменить текущую CNN на диффузионный декодер (этап 2). Ожидается более стабильное добавление деталей без артефактов при высоком увеличении (x4, x8).

Мульти-доменная адаптация: Расширить подход для одновременной нормализации **нескольких сенсоров** в единый целевой стиль, повышая универсальность системы.

Выводы

Разработан двухэтапный метод: **Перенос стиля + Сверхразрешение.**

Метод **устраняет межсенсорный domain shift** и повышает детализацию без необходимости в парных данных.

Эксперименты подтвердили, что наш подход превосходит одиночное SR, подчеркивая важность стилевой адаптации.

Решение масштабируется: Полезно для интеграции разнородных данных, выравнивая изображения разных спутников под единый стандарт.

Спасибо за внимание! Вопросы?