

Единый нейросетевой метод повышения разрешения для разнородных цифровых изображений дистанционного зондирования Земли

В.Ф. Коновалов¹, В.В. Мясников¹, В.В. Сергеев¹

¹ Самарский национальный исследовательский университет имени академика С.П. Королёва, 443086, Россия, г. Самара, Московское шоссе, д. 34

Аннотация

Работа посвящена поиску единого решения для практической задачи повышения разрешения разнородных цифровых изображений дистанционного зондирования Земли: панхроматических изображений, цветных изображений, мультиспектральных и гиперспектральных изображений, а также изображений, полученных радаром с синтезированной апертурой (РСА). Для получения такого решения из существующих наборов данных был собран агрегированный, включающий как пары изображений высокого и низкого разрешений, так и отдельные изображения, выступающие как изображения высокого разрешения. Для собранной базы данных и типовой модели искажений было проведено экспериментальное сравнение различных современных нейросетевых моделей повышения разрешения (включая преднастроенные и переобученные варианты), которые в обзорах фигурируют как state-of-the-art: сверточные, генеративно-состязательные и трансформенные (основанные на механизме внимания). Учитывая возможные ограничения моделей на число слоев космических снимков, рассматривались два варианта предварительной обработки изображений. Итоговый результат работы – единая нейросетевая модель повышения разрешения цифровых космических снимков, дополненная адаптерами для различных их типов, которая при обучении превосходит или не уступает современным специализированным state-of-the-art решениям.

Ключевые слова: цифровые изображения дистанционного зондирования Земли, повышение разрешения изображений, мультиспектральные и гиперспектральные изображения, радарные изображения.

Цитирование: Коновалов, В.Ф. Единый нейросетевой метод повышения разрешения для разнородных цифровых изображений дистанционного зондирования Земли / В.Ф. Коновалов, В.В. Мясников, В.В. Сергеев // Компьютерная оптика. – 2024. – Т. 48, № 6. – С. 944-955. – DOI: 10.18287/2412-6179-CO-1610.

Citation: Konovalov VF, Myasnikov VV, Sergeev VV. A unified neural network-based single super-resolution method for heterogeneous digital earth remote sensing images. Computer Optics 2024; 48(6): 944-955. DOI: 10.18287/2412-6179-CO-1610.

Введение

Повышение разрешения цифрового изображения – одна из популярных задач компьютерного зрения. По мере развития и все более широкого внедрения современных вычислительных средств она находит все большее применение: в телевидении, в цифровой фотографии, в задачах оцифровки архивных снимков и др. С определенного времени эта задача стала актуальной и востребованной и при наземной обработке цифровых изображений, полученных со спутников дистанционного зондирования Земли (ДЗЗ) [1]. Причинами такого, на первый взгляд парадоксального, направления использования являются чрезвычайно высокая стоимость и/или принципиальная невозможность замены аппаратуры съемки на космических аппаратах ДЗЗ по мере ее естественного устаревания, а также возможность использования нескольких изображений различного пространственного и спектрального разрешения для получения синтезированных снимков с одновременно лучшими спектральными и пространственными характеристиками, чем су-

ществующие. Последнее направление для оптических снимков обусловлено достаточно простым физическим законом сохранения энергии: количество энергии на единицу площади регистрации остается постоянным вне зависимости от числа каналов и разрешения изображения. В силу ограниченной чувствительности датчика в итоговом изображении может быть либо большее число спектральных каналов при меньшем пространственном разрешении, либо, наоборот, лучшее пространственное разрешение при меньшем числе спектральных каналов. Поэтому типичной задачей является, например, задача повышения разрешения цветного снимка низкого разрешения (LR – low resolution) по панхроматическому (полутоновому) изображению более высокого разрешения (HR – high resolution), именуемая в зарубежной литературе pansharpening [1], [2] и являющейся частным случаем класса задачи объединения/слияния (англ.: union) изображений космической съемки [3]. Хотя использование пар изображений является наиболее популярной задачей при обработке данных ДЗЗ, однако подходящие одновре-

менные пары снимков для решения задачи повышения разрешения далеко не всегда могут быть найдены. Это приводит нас к постановке задачи, когда повышение разрешения цифрового изображения ДЗЗ следует выполнять только по одному снимку. Такой тип задач при обработке космических снимков выделяется в отдельный класс, именуемый в англоязычной литературе Single Remote Sensing Image Super Resolution (SRSISR). *Настоящая работа направлена на решение именно этого класса задач.*

К настоящему времени существует целый ряд методов и алгоритмов решения задач SRSISR, базирующихся как на классических подходах компьютерного зрения (например, сплайновая интерполяция, интерполяция через спектр и др.), так и на нейросетевой парадигме. Решения, использующие глубокие нейронные сети (ГНС), в настоящем являются лидирующими, превосходя остальные по ключевым показателям качества [5]. При этом в рамках самой нейросетевой парадигмы есть ряд конкурирующих направлений решения задачи SRSISR [6], [7], [14]: сверточные нейронные сети (СНС) [8], генеративно-состязательные сети (GAN) [11] и сети-трансформеры, основанные на механизмах внимания [12]. При этом область применения таких решений дополнительно сужается классом космических снимков, для которых они могут быть использованы. Собственно класс изображения при этом может быть ограничен не только числом каналов и физическим типом сенсора (полноцветное, мультиспектральное, гиперспектральное и PCA), но даже итоговым разрешением изображения. Все это порождает огромное количество узкоспециализированных решений, которые оптимальны для некоторого конкретного класса изображений ДЗЗ, но плохи или неприменимы для изображений других классов. *Настоящая работа направлена на устранение обозначенной проблемы и имеет целью определение/нахождение единой нейросетевой модели для повышения разрешения разнородных цифровых изображений, которая по результатам экспериментальных исследований не будет уступать по качественным показателям узкоспециализированных решений SRSISR.*

Работа построена следующим образом. В первом параграфе представлен обзор существующих методов и алгоритмов SRSISR. Во втором параграфе представлена формальная постановка задачи SRSISR. Третий, основной параграф, содержит детальное описание всех этапов получения искомого результата: от деталей формирования обучающих наборов разнородных изображений ДЗЗ до результатов экспериментальных сравнений на них state-of-the-art моделей ГНС в различных постановках. Полученное в результате наилучшее решение также представлено в этом параграфе. Завершают работу заключение и список использованной литературы.

Основной вклад нашей работы следующий:

- объединенная база данных для задачи повышения разрешения цифровых изображений ДЗЗ SRSISR: RGB, мультиспектральные, панхроматические и PCA-изображения;
- результаты сравнения state-of-the-art моделей ГНС в задаче SRSISR для факторов «×2» и «×4» и различных типов изображений ДЗЗ;
- результаты сравнения способов предварительной обработки изображений ДЗЗ для приведения их к трехканальному виду;
- набор обученных алгоритмов (весов моделей ГНС) для задачи SRSISR для факторов «×2» и «×4» и различных типов изображений ДЗЗ;
- итоговое решение в виде наилучшей модели, результатов ее настройки для конкретных изображений ДЗЗ и факторов, способа предварительной обработки.

1. Современное состояние исследований

1.1. Классификация задач и методов ее решения

Задачи повышения разрешения цифровых изображений, как отмечалось во введении, делятся на задачу повышения разрешения LR-изображения по HR-изображению (IRS – Image Super Resolution) и задачу повышения разрешения одного изображения, интерпретируемого как LR (Single IRS – SIRS). Последняя, рассматриваемая в приложении к цифровым изображениям ДЗЗ и именуемая SRSISR, по сравнению с SIRS имеет ряд особенностей [14]:

- значительное расстояние от объекта регистрации до точки съемки, что позволяет рассматривать регистрируемую сцену как сцену малой глубины;
- большие размеры снимков и чрезвычайно большое число объектов различной физической природы;
- существенные различия снимков в различных географических поясах, существенно различного разрешения, полученные в различных погодных/временных условиях и при различной геометрии.

Указанные особенности делают неэффективным применение настроенных/обученных методов/алгоритмов SIRS для решения задачи SRSISR.

Далее все множество методов решения задачи SRSISR делится на методы обучения с учителем и без учителя. Методы первого множества получают информацию для обучения в виде пар LR-HR-изображений, по которым они и конструируют искомое преобразование. Методы второго, называемые также непарными, используют только HR-изображения для «восприятия» их свойств и их «переноса» при преобразовании LR-изображения в HR [9]. В данной работе мы рассматриваем только методы обучения с учителем. Ниже будет дана классификация наиболее известных методов SRSISR, опубликованных за последние несколько лет, а также ряда методов SIRS, которые также будут участвовать в

сравнении на собранной в рамках настоящей работы базе данных изображений ДЗЗ.

Итак, современные глубокие нейросетевые модели повышения разрешения цифровых изображений подразделяются на следующие *типы*: *сверточные, генеративно-состязательные и трансформенные* (основанные на механизме внимания). *Сверточные сети* (англ.: convolutional neural network – CNN) в качестве основной операции преобразования используют операцию свертки или, если быть более точными, перекрестной корреляции входного двумерного поля изображения с набором ядер. Комбинируя указанную операцию с различными вариантами разделимых и неразделимых ядер, функций активаций, переброса соединений (ResNet), конкатенаций (DenseNet), слоями расширения и повышающей свертки и др., получают различные модели преобразований, настраиваемые на конкретные задачи.

Генеративно-состязательные сети (англ.: generative adversarial network – GAN) – подкласс методов машинного обучения, предложенный Ian Goodfellow с коллегами в 2014 [11] и предназначенный для генеративного синтеза изображений по базе данных с прототипами. Суть процесса обучения состоит в противостоянии двух ГНС – генератора и дискриминатора, в котором генератор пытается создавать изображения определенного класса, а дискриминатор указывает степень похожести изображения (например, созданного) на изображения этого класса. При наличии обучающего множества изображений «своего» и «иных» классов задача обучения дискриминатора представляется относительно простой. Гениальность идеи авторов GAN состоит в том, чтобы по мере обучения дискриминатора ему на вход подавать и те изображения, что созданы генератором, и выход первого использовать для обучения второго.

Сети-трансформеры (англ.: *Transormers*), или ГНС на основе механизма внимания – ГНС, которые в процессе последовательных пар преобразований используют, наряду с MLP (англ.: MLP – multi-layer perceptron), так называемый *механизм внимания* – нормированные посредством функции softmax выходы скалярных произведений линейно преобразованных векторов входной последовательности. Первоначально используемые в обработке последовательностей (текста) [12] модели трансформера были приспособлены и для обработки изображений путем переопределения понятия токена и способа позиционного кодирования [13].

1.2. Современные методы решения задачи SRSISR

Для задачи SRSISR распределение наиболее известных методов ее решения по типам и годам опубликования соответствующих решений представлены ниже в табл. 1. Более подробное изложение большинства из перечисленных в табл. 1 методов можно найти в обзорной работе [14].

Табл. 1. Распределение методов решения задачи SRSISR по типам и годам

	CNN	GAN	Transormers
2017	LGCNet, RS-DRL	-	-
2018	RS-CNN, DDRN, DMCN	TGAN	-
2019	CycleCNN, UDRB, WTCRR, MRNN, CSAE	CDGans, DRGAN, EEGAN	RSEB, RS-RCAN, MPSR
2020	NLASR, EDCNN, AMFFN, PRDNN	RDGAN, CCWGAN, RS-ESRGAN, SRPeruSatGAN	MHAN, CAFRN, PCRCAN, DSSR
2021	FGRDN, ERCNN	SCSE-GAN	DRDAN, RASAF, CARS
2022		NDSRGAN, MA-GAN, TE-SAGAN	RBAN-UNet, MA-GAN, TE-SAGAN
2023		SRADSGAN	
2024	SWCGAN	SWCGAN	Swin2 MoSE

Несмотря на существенные различия в качественных показателях, приводимых в оригинальных авторских работах (например, показатель PSNR мог варьироваться от 25 Дб до 45 дБ), как показывает проведенное в указанной обзорной работе [14] относительно объективное сравнение подмножества state-of-the art алгоритмов, показатели качества при повышении разрешения изображений у них различаются не так сильно и это различие становится все меньше по мере повышения разрешения с 2 до 8 раз. В частности, при коэффициенте «2» различия в группе CNN оказались не более 1 Дб, а в группе трансформеров – менее 0,5 Дб. При этом существенное различие на 4,2 Дб в группе GAN сопровождалось противоположными изменениями в показателях других критериев. При этом различие между наилучшими алгоритмами различных групп также оказывались незначительными и не превышали 1 Дб для коэффициента «2». Для определенности, в указанном исследовании сравнивались следующие методы:

CNN-методы

- SRCNN [19]: Super-Resolution Convolutional Neural Network,
- VDSR [20]: Very Deep Super-Resolution,
- LGCNet [21]: Local-Global Combined Network,
- PECNN [22]: Progressively Enhanced Convolutional Neural Network,
- RDN [23]: Residual Dense Network,
- DDRN [24]: Deep Distillation Recursive Network.

GAN-методы

- SRGAN [25]: Super-Resolution Generative Adversarial Networks,
- ESRGAN [26]: Enhanced Super-Resolution Generative Adversarial Networks,

EEGAN [27]: Edge-enhanced GAN.
Transformers-методы
 RCAN [28]: Residual Channel Attention Network,
 RSRCAN [29]: Remote Sensing Residual Channel Attention Network,
 HAN [30]: Holistic Attention Network,
 SAN [31]: Second order attention network,
 MHAN [32]: Mixed High order Attention Network,
 CARS [33]: Channel Attention Remote Sensing.

Для определенности, исследование проводилось на панхроматических изображениях с пространственным разрешением 0,31–1 м и спектральных с разрешением 1,24–4 м, в нем не участвовали мультиспектральные и PCA-изображения.

1.3. Современные методы решения задачи SISR

SISR-решения для повышения разрешения цифровых изображений появились в компьютерном зрении достаточно давно [37]. Первые и наиболее простые решения использовали сплайн-интерполяцию, например, билинейную, доступную во многих современных редакторах изображений GIMP, Adobe Photoshop. Классические подходы, называемые в зарубежной литературе “upsampling”, также включают интерполяцию изображения с помощью преобразования Фурье (оптимально, если исходный сигнал имел дискретизацию в соответствии с теоремой отсчетов Котельникова, но использование на обычных изображениях может приводить к нарушению неотрицательности функции яркости) и различные способы фильтрации «расширенного» изображения.

Последнюю декаду классические методы уступили место подходам с использованием ГНС. Для них классификация методов решения полностью соответствует приведенному в пп. 1.1 выше: СНС, GAN и сети-трансформеры.

Подробный обзор методов можно найти в работе 2024 года [38]. В таблице ниже представлены названия методов, которые сравнивались в указанном обзоре.

Табл. 2. Распределение методов решения задачи SISR по типам и годам

	<i>CNN</i>	<i>GAN</i>	<i>Transormers</i>
2016	SRCNN, VDSR, FSRCNN, DRCN, ESPCNN		
2017	DRRN, MDSR SR-DenseNet, MixNet	SRGAN	SelNet
2018	CARN, MSRN SRMDNF, RDN, D-DBPN	ZSSR, ESRGAN	RCAN, SENet, SESR, ScSe, RAM
2019	IMDN CBPN	RaGAN Rank-SRGAN	SAN RNAN
2020	FDC	Unpaired GAN	CSFM RFANet
2021	MCSN KARN		UAN
2022			DRLN

Как можно заметить, некоторые методы присутствуют в обеих таблицах. Более того, выводы относительно близости результатов, представленные по методам SRSISR, можно отнести и к методам SISR из табл. 2, а именно: в обзорной работе рассматривалась задача повышения разрешения в 4 раза, и подавляющее число методов ГНС показывает удивительно близкие результаты, отличающиеся менее чем в 1 Дб.

На основании результатов сравнения, проведенных в двух указанных обзорных работах, можно сделать вывод об относительном паритете современных передовых методов повышения разрешения цифровых изображений ДЗЗ. Это, в свою очередь, показывает *целесообразность поиска единого решения, качественные показатели которого были бы близки (или превосходили) показатели современных state-of-the art методов SRSISR*. Очевидно, такое решение может основываться либо на моделях SRSISR, либо на более универсальных моделях SISR.

1.4. Базы данных и соревнования

Базы данных для задач SR делятся на два типа – с искусственно синтезированными изображениями низкого разрешения и с реальными парами изображений низкого–высокого разрешений. Получение реальных пар изображений на практике затруднительно, поэтому существует лишь небольшое количество подобных наборов данных, например, DIV2K для SISR и SEN2Venus для SRSISR. В качестве баз данных с искусственно синтезированными изображениями низкого разрешения может использоваться любой набор данных изображений, например для SISR – Set5 и Set14, для SRSISR – UCMerced.

В области повышения разрешения цифровых изображений также проводятся соревнования. Например, соревнование NTIRE (англ.: New Trends in Image Restoration and Enhancement workshop and associated challenges) [39], которое в 2024 прошло в девятый раз. По его результатам можно судить о состоянии области и получать количественные оценки для сравнения эффективности предлагаемых решений. Соревнования нацелены на выявление наиболее эффективных решений/алгоритмов по нескольким направлениям: числу параметров, качественным характеристикам, наибольшей скорости и т.д.

2. Постановка задачи SRSISR

Обозначим $I_{N_1N_2}$ изображение высокого разрешения, а $I_{M_1M_2}(M_1 < N_1, M_2 < N_2)$ – изображение низкого разрешения, связанное с $I_{N_1N_2}$ через некоторое неизвестное нам преобразование Φ . Это преобразование обычно интерпретируется как *преобразование деградации*, понижающее разрешение изображения высокого разрешения:

$$I_{M_1M_2} = (I_{N_1N_2}). \tag{1}$$

Задача SRSISR в целом заключается в построении «обратного» к Φ преобразования, обозначаемого далее ϑ , такого, чтобы выполнялось точное или приближенное равенство:

$$I_{N_1 N_2} \cong_{\vartheta} (I_{M_1 M_2}).$$

Здесь ϑ – параметры конструируемого преобразования ϑ . Учитывая, что подобная задача в самой общей постановке не является корректной, решение (то есть вектор/тензор параметров ϑ) в классе методов обучения с учителем ищется как решение оптимизационной задачи следующего вида:

$$E \left\{ L \left(I_{N_1 N_2}, \vartheta \left(I_{M_1 M_2} \right) \right) \right\}_{\vartheta} \rightarrow \min. \quad (2)$$

Здесь $L(\dots, \dots)$ – функция, определяющая величину рассогласования между двумя изображениями (в машинном обучении называемая функцией потерь или loss-функцией), а $E\{\dots\}$ – оператор математического ожидания, связанный с распределением класса рассматриваемых изображений.

На практике математическое ожидание в задаче (2) заменяется на эмпирическое среднее по множеству пар изображений высокого и низкого разрешений:

$$\left\{ I_{N_1 N_2}^j, I_{M_1 M_2}^j \right\}_{j=1}^J.$$

Оценка параметров преобразования тогда определяется следующим выражением:

$$\hat{\vartheta} = \arg \min_{\vartheta} \left(\frac{1}{J} \sum_{j=1}^J L \left(I_{N_1 N_2}^j, \vartheta \left(I_{M_1 M_2}^j \right) \right) \right).$$

Дополнительно следует отметить, что не всегда на практике удается получить/наблюдать согласованную пару $I_{N_1 N_2}^j, I_{M_1 M_2}^j$ изображений. Довольно часто доступно только одно из них. Практическим приемом решения этой проблемы является интерпретация наблюдаемого изображения как изображения высокого разрешения и введение модели деградиационного преобразования $\vartheta(\dots)$, которая настраивается на выполнение преобразования (1), то есть на преобразование изображений высокого разрешения в изображения низкого разрешения. После настройки ϑ параметров преобразования $\vartheta(\dots)$ по априорным данным или по парам наблюдаемых изображений высокого и низкого разрешения оно может быть использовано для получения согласованных пар изображений. То есть для получения оценки изображения \tilde{I} низкого разрешения по изображению высокого разрешения:

$$\tilde{I}_{M_1 M_2} =_{\vartheta} \left(I_{N_1 N_2}^j \right). \quad (3)$$

Хотя полученная оценка \tilde{I} изображения низкого разрешения может и не совпадать с истинным абсолютно, оно, наряду с реальными парами LR- и HR-изображений, может быть использовано для оценки параметров ϑ искомого преобразования ϑ .

Самым простым способом получения варианта изображения низкого разрешения является обычная бикубическая интерполяция [15]. Физически более корректной моделью $\Phi_{\vartheta}(\dots)$ является линейная модель наблюдения в виде [16]:

$$\tilde{I}_{M_1 M_2}^j = \left[I_{N_1 N_2}^j * h_{\vartheta} \right] \downarrow_{M_1 M_2}^{N_1 N_2} + n_{\vartheta}, \quad (4)$$

где h_{ϑ} – ядро линейного преобразования, зависящее некоторым образом от параметров ϑ , а n_{ϑ} – шум, параметры которого также определяются вектором/тензором ϑ . Оператор $\downarrow_{M_1 M_2}^{N_1 N_2}$ – оператор передискретизации. В практически наиболее часто рассматриваемом случае $N_i = M_i k (i=1, 2)$, а целочисленный коэффициент k принимает значения из множества $\{2, 3, 4, 8\}$. Тогда сам оператор $\downarrow_{M_1 M_2}^{N_1 N_2}$ оказывается обычным прореживанием отсчетов с коэффициентом k . В общем случае построение преобразования (3) для данных ДЗЗ должно учитывать характеристики оптико-электронного тракта регистрации, используемого на конкретном космическом аппарате, а также параметры искажения атмосферы.

3. Построение единого нейросетевого метода повышения разрешения

Процесс поиска и построения единого нейросетевого метода повышения разрешения выполняется в несколько этапов, включающих следующие.

- 1) *Формирование базы данных цифровых изображений ДЗЗ различного типа:* панхроматических, полноцветных, мульти- или гиперспектральных и РСА. Сформированная база данных априори должна содержать пары изображений высокого и низкого разрешения для нескольких сенсоров КА.
- 2) *Расширение базы данных за счет синтеза изображений* низкого разрешения по существующим изображениям (для тех изображений, у которых в оригинальных базах данных отсутствуют парные LR- или HR-изображения).
- 3) *Определение вариантов предварительной обработки изображений ДЗЗ* к «типовому» формату алгоритмов повышения разрешения. В качестве типового формата в настоящей работе приняты трехканальные изображения.
- 4) *Определение критериев сравнения.*
- 5) *Определение состава SRSISR и SISR нейросетевых моделей*, участвующих в качестве претендентов в исследовании. Выбор алгоритма, результаты которого могут быть приняты за т.н. минимально допустимые, или, в английских терминах, за «baseline».
- 6) *Проведение экспериментальных исследований* по сравнению эффективности отобранных моделей в совокупности с предложенными в п. 3 вариантами адаптации.
- 7) *Анализ результатов и выбор оптимального решения.*

Ниже все эти этапы рассмотрены подробнее.

3.1. Формирование базы данных цифровых изображений ДЗЗ различного типа

После изучения существующих наборов данных ДЗЗ, представленных в пп. 1.3, было отмечено достаточно малое количество специализированных наборов данных ДЗЗ именно для задачи *повышения разрешения по одному изображению в рамках методов обучения с учителем*. Это связано с объективной сложностью получения пар изображений космического ДЗЗ низкого и высокого разрешения. Основной проблемой является требование одновременности съемки или достаточно малого интервала времени между съемками. Далее требуется полное (координатное) согласование изображений, полученных с разных камер, что на практике приводит к необходимости геометрической коррекции одного из них (и снижения разрешающей способности). Также нужны компенсация различных искажений и отличий между камерами и т. д. Эта проблема неспецифична для ДЗЗ и решается путем создания синтетических изображений низкого разрешения (см. пп. 3.2). Описание использованных в исследованиях наборов данных мы представили ниже.

Набор данных SEN2Venus – мультиспектральный набор данных, содержащий пары HR- и LR-изображений. Данный набор содержит реальные пары HR- и LR-изображений, полученные совмещением данных ДЗЗ со спутников Sentinel-2 и Venus. Sentinel-2 имеет 4 спектральные полосы с разрешением 10 метров/пиксель (R, G, B, Wide NIR) и 4 спектральные полосы с разрешением 20 метров/пиксель (Red 700 nm, Red 740 nm, Red 780 nm, Narrow NIR), Venus – 7 спектральных полос с разрешением 5 метров/пиксель (R, G, B, Red 700 nm, Red 740 nm, Red 780 nm, Narrow NIR). Из них сформированы пары четырехканальных мультиспектральных изображений низкого – высокого разрешения.

Нужно отметить, что все изображения этого набора были получены с временным разбегом не более чем в 30 минут. Из всего набора данных было отобрано 1000 изображений в случайных локациях и временных отрезках. Разрешение набора данных: 256×256.

Набор UC Merced Land Use содержит 2100 полноцветных RGB-изображений формата .tif с диапазоном [0, 255], разделенных на 21 категорию (сельское хозяйство, самолеты, жилая застройка,...). Набор предполагался как унифицированный по размеру – 256×256 пикселей, но размеры некоторых изображений отличались и были нами удалены. Оставшиеся изображения были разделены на тренировочную (90 %) и тестовую (10 %) выборки. Так как набор данных изначально не предоставляет изображения в низком разрешении, они были получены синтетически путем бикубической интерполяции (см. пп.3.2).

Набор Grayscale оптических панхроматических (полутонных) изображений ДЗЗ был получен из набора RGB-данных (см. выше) путем перевода изображений в оттенки серого. Для совместимости со всеми исследуемыми методами изображения были дополнены до 3 каналов путем копирования канала.

Набор PCA-изображений, полученных радаром с синтезированной апертурой, представляет собой поднабор набора данных SEN2MS. Изображения этого набора одноканальные, размера 256×256, формата PNG.

Изображения LR для последних двух наборов были получены так, как описано ниже в пп. 3.2.

Наборы данных для гиперспектральных изображений ДЗЗ не формировались по причине отсутствия в свободном доступе достаточных данных.

3.2. Расширение базы данных за счет синтеза изображений

Изображения низкого разрешения были получены двумя способами, для каждого из которых проводились эксперименты:

1. *Бикубическая интерполяция* [15]. Этот метод использовался для оптических изображений и PCA.
2. *Логарифмирование с добавлением шума (для SAR)*. Поскольку PCA-изображения отличаются по физическим принципам их получения от оптических, они очень зашумлены спекл-шумом. И при бикубической интерполяции этот шум будет «рассеян» на соседние пиксели. Поэтому как альтернативу бикубической интерполяции мы использовали следующую модель получения LR PCA-изображений, основанную на модели (4):

$$\tilde{I}_{M_1 M_2} = \exp\left(\ln(I_{N_1 N_2}) * h_9 + n_9 \cdot 0.05\right) \downarrow_{M_1 M_2}^{N_1 N_2},$$

где $n_9 \sim N(0, 1)$; h_9 – ядро усреднения размером $k \times k$.

3.3. Определение вариантов предварительной обработки изображений ДЗЗ

Практически все существующие методы повышения разрешения в качестве входных данных принимают трехканальные изображения. Это делает необходимым адаптацию (преобразование) некоторых изображений так, чтобы они могли быть использованы этими методами. Мы использовали два варианта адаптации, которые сравнивались между собой:

- 1) *Деление многоканальных изображений на трехканальные с пересечением*.

В частности, четырехканальное изображение разделялось на два трехканальных, где выборка каналов происходила последовательно. Такая схема выборки представляется наиболее корректной, поскольку она наиболее соответствует последовательности полноцветных RGB-изображений. В данном случае изображение получается «в ложных RGB-цветах».

2) Деление многоканальных изображений на одно-канальные с последующим превращением их в трехканальные копированием.

Для набора Grayscale, где изображения одноканальные, они преобразовывались в трехканальные копированием каналов.

3.4. Определение критериев сравнения

Выбор показателей/мер качества решения задачи SRSISR/SISR является важной задачей. Как правило, в задачах повышения разрешения применяется несколько метрик либо их комбинация вместе с визуальным контролем качества. Мы использовали три показателя.

PSNR (англ.: Peak Signal to Noise Ratio), или *пиковое отношение сигнала к погрешности*, является одним из наиболее часто используемых показателей качества изображения (при сравнении с образцом):

$$PSNR = 10 \cdot \log_{10} \left(\frac{(\max(I_{N_1 N_2}))^2}{\frac{1}{N_1 N_2} \sum_{n=1}^{N_1} \sum_{m=1}^{N_2} (I(n, m) - \tilde{I}(n, m))^2} \right).$$

Эта мера является легко интерпретируемой, позволяет легко сравнивать результаты с полученными ранее и в других исследованиях. Однако она обладает известным недостатком – она склонна принимать размытые изображения как более качественные по сравнению с изображениями с точечными, но большими ошибками.

Другой используемой мерой качества является *мера структурного сходства* SSIM (Structural Similarity Index Measure) [17]. Она изменяется от «-1» (противоположная корреляция) до «1» (абсолютное сходство). Данная мера делится на три составляющие: сходство яркостей l , контраста c и структуры s . В результате эта мера дает более точную картину качества изображения. Особенно в задаче сверхразрешения, где восстановление четких границ изображения важно:

$$SSIM(I, \tilde{I}) = l^\alpha(I, \tilde{I}) \cdot c^\beta(I, \tilde{I}) \cdot s^\gamma(I, \tilde{I}).$$

Здесь

$$l(I, \tilde{I}) = \frac{2\mu_I \mu_{\tilde{I}} + c_1}{\mu_I^2 + \mu_{\tilde{I}}^2 + c_1}, \quad c(I, \tilde{I}) = \frac{2\sigma_I \sigma_{\tilde{I}} + c_2}{\sigma_I^2 + \sigma_{\tilde{I}}^2 + c_2},$$

$$s(I, \tilde{I}) = \frac{2\sigma_{\tilde{I}\tilde{I}} + c_3}{\sigma_I + \sigma_{\tilde{I}} + c_3}.$$

В указанных выражениях величины μ_I и $\mu_{\tilde{I}}$ – средние значения яркостей для соответствующих изображений (или их фрагментов), σ_I^2 и $\sigma_{\tilde{I}}^2$ – их дисперсии, а $\sigma_{\tilde{I}\tilde{I}}$ – ковариация между ними. Константы $c_3 = c_2/2$, $c_1 = (0,01L)^2$, $c_2 = (0,03L)^2$ предотвращают появления нуля в знаменателях, а L – величина, характеризующая динамический диапазон изображения (для

8-битового представления отсчетов она равна $2^8 = 256$). Показатели α, β, γ обычно полагаются равными «1».

Существуют меры, целью которых является наиболее близкое к человеческому сравнению качества полученных изображений. Одной из таких мер является обученная мера LPIPS (англ.: Learned Perceptual Image Patch Similarity) [18]. Ее авторы установили, что последние слои нейросетей, обученных на достаточно широком наборе данных, показывают схожее с мозгом восприятие изображений. Это позволило им использовать выходы этих слоев в качестве некоего скрытого представления изображения и рассчитывать меру (расогласование) уже по нему. Чем меньше значение этой меры, тем лучше качество изображения.

3.5. Определение состава ГНС-моделей

Для исследования было выбрано несколько лучших моделей ГНС общего назначения (SISR):

- SRCNN [19],
- RealESRGAN [26],
- SwinIR Lightweight и Classical [34],

представляющих различные категории нейронных сетей СНС, GAN и ИНС-трансформер соответственно. Также в исследовании участвовало несколько специализированных моделей SRSISR: сверточная сеть LGCNet [21] и одна из самых современных ГНС-трансформер Swin2mose [35], которая является доработкой сети-трансформера SwinIR для задачи SRSISR, то есть для работы с ДЗЗ.

Предобученные модели в случае их существования были взяты из официальных репозиториях моделей. Сеть LGCNet, в силу отсутствия официальной реализации и весов модели, была создана «с нуля» с использованием PyTorch.

В качестве «baseline»-варианта повышения разрешения мы использовали результаты бикубической интерполяции (в таблицах – «BICUBIC»).

3.6. Проведение экспериментальных исследований

Целью эксперимента было:

- определение наилучшей модели ГНС-решения задачи SRSISR для цифровых изображений ДЗЗ различной природы (оптических и мультиспектральных) при повышении разрешения в 2 и 4 раза,
- получение конкретных обученных моделей для всех типов изображений - оптических (панхроматических, RGB, мультиспектральных) и PCA
- выбор способа предварительной обработки изображений ДЗЗ для их использования соответствующим методом повышения разрешения.

Второстепенной задачей исследования являлась проверка следующих предположений:

- а) алгоритмы SISR общего назначения после обучения на изображениях ДЗЗ могут быть эффективно использованы и для решения задач SRSIRS;

б) алгоритмы, обученные на оптических RGB-изображениях ДЗЗ, могут быть использованы для повышения разрешения оптических изображений ДЗЗ другого типа (панхроматических и мультиспектральных).

Возможность использования алгоритмов, обученных для оптических данных, для изображений PCA требовало отдельной проверки.

Чтобы сравнить эффективность и генерализуемость моделей ГНС в задаче SRSISR, было проведено обучение моделей на тренировочной части набора данных UC Merced Land Use. Поскольку полное обучение моделей с нуля, особенно моделей на основе трансформера, является вычислительно затратным и требует большого количества данных, было решено использовать дообучение существующих базовых моделей общего назначения. Это уменьшает количество необходимых данных, а также время для адаптации под конкретную задачу. Также в тех случаях, когда это было возможно, применялось обучение модели сначала для случая 2×, а затем полученная модель обучалась для случая 4×. Это позволяет в ряде случаев дополнительно уменьшить время обучения.

Для каждого вида нейросетей использовался свой набор параметров, основой для которого являлся файл конфигурации обучения нейронной сети из репозитория. Общим также являлось количество изображений в тренировочном наборе данных и его состав.

Ниже приведены уточняющие комментарии к процессам обучения выбранных моделей ГНС:

- LGCNet: конфигурация согласно [21];
- SRCNN: обучение для «×2» в течение 10 000 эпох, затем для «×4» в течение 5 000 эпох;
- Real-ESRGAN: дообучение в течение 200 000 шагов для «×2», затем в течение 100 000 шагов для «×4», batch size изменен на «6»;
- SwinIR-Lightweight и SwinIR-Classical: дообучение в течение 100 000 шагов для «×2», затем в течение 100 000 шагов для «×4»;
- Swin2MoSE не дообучалась и использовалась только для сравнения с RGB результатами.

3.7. Анализ результатов и выбор оптимального решения

Непосредственным результатом проведенных экспериментов является массив таблиц с показателями эффективности по каждой из обученных и базовой моделей для каждого из типов данных. В силу ограниченного объема настоящей публикации мы включили в ее текст только отдельные материалы, демонстрирующие либо типовые зависимости, либо значимые особенности.

Из табл. 3 мы можем видеть, что обучение ГНС общего назначения на RGB-изображениях ДЗЗ позволяет улучшить или оставить прежним показатель качества. Наибольшее отличие от базового метода BICUBIC наблюдается для моделей ГНС

RealESRGAN и SwinIR по метрике LPIPS. На Рис. 1 представлены примеры полученных изображений.

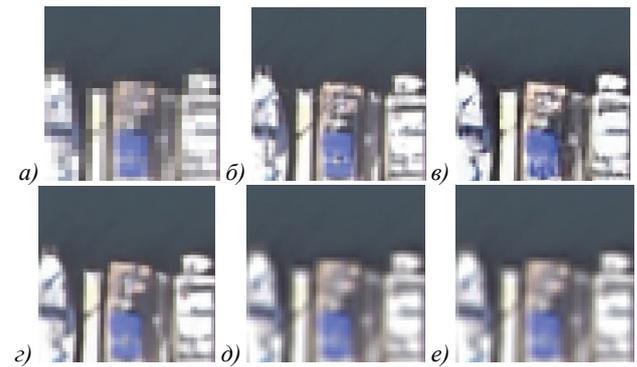


Рис. 1. Примеры полученных изображений (×2):
а) LR-изображение; б) HR-изображение;
в) модель RealESRGAN; з) модель SwinIR;
д) модель SRCNN; е) модель BICUBIC

Табл. 3. Показатели качества для задачи SRSISR, RGB-изображения

×2						
Нейросеть	До переобучения			После переобучения		
	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS
LGCNet_x2	нет	нет	нет	31,385	0,890	0,192
SRCNN_x2	нет	нет	нет	30,285	0,874	0,277
Swin-Lightweight_64_8_x2	34,771	0,935	0,084	34,907	0,935	0,087
Swin-Classical_48_8_x2	34,990	0,937	0,081	35,251	0,938	0,079
Swin2mose_sen2venusavg	24,065	0,864	0,200	нет	нет	нет
Swin2Mose_sen2venusred	24,160	0,865	0,199	нет	нет	нет
Swin2mose_sen2venusgray	25,424	0,888	0,295	нет	нет	нет
BICUBIC	31,429	0,890	0,193	31,429	0,890	0,193
×4						
Нейросеть	До переобучения			После переобучения		
	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS
LGCNet_x4	нет	нет	нет	25,897	0,684	0,474
SRCNN_x4	нет	нет	нет	26,173	0,712	0,399
SwinLightweight_64_8_x4	27,991	0,769	0,282	28,743	0,784	0,273
SwinClassical_48_8_x4	28,340	0,786	0,264	28,337	0,798	0,248
Swin2mose_sen2venusavg	20,695	0,672	0,405	нет	нет	нет
Swin2Mose_sen2venusred	20,760	0,670	0,407	нет	нет	нет
Swin2mose_sen2venusgray	19,895	0,669	0,476	нет	нет	нет
BICUBIC	25,910	0,684	0,474	25,910	0,684	0,474

Табл. 4 показывает эффективность полученной обучением на RGB-изображениях модели на тестовых изображениях в оттенках серого. Эффективность модели не снижается для этих изображений, что мо-

жет говорить о возможной применимости этих моделей для изображений с произвольным числом каналов путем их поканальной обработки, например, PCA и мультиспектральных изображений, что и будет проведено далее.

Табл. 5 и 6 отражают количественные показатели эффективности полученной обучением на RGB-изображениях модели на тестовых изображениях PCA для двух вариантов синтеза SAR-изображений низкого разрешения – с добавлением speckle шума и без его добавления. Ожидаемо, вариант с добавлением шума оказался более сложным для восстановления, однако в обоих случаях как обученные заново, так и базовые методы показали себя значительно лучше, чем базовый BICUBIC и специализированный LGCNet, за исключением SRCNN.

Табл. 4. Показатели качества для задачи SRSISR, grayscale изображения

×2						
Нейросеть	До переобучения			После переобучения		
	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS
LGCNet_x2	нет	нет	нет	31,470	0,894	0,178
Swin-Classical_48_8_x2	34,917	0,938	0,080	34,912	0,938	0,085
BICUBIC	31,516	0,894	0,178	31,516	0,895	0,178
×4						
Нейросеть	До переобучения			После переобучения		
	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS
LGCNet_x4	нет	нет	нет	25,879	0,688	0,460
SwinClassical_48_8_x4		0,780	0,273	28,121	0,780	0,273
BICUBIC	25,892	0,688	0,460	25,892	0,688	0,460

Табл. 5. Показатели качества для задачи SRSISR, SAR-noise изображения

×2						
Нейросеть	До переобучения			После переобучения		
	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS
LGCNet_x2	нет	нет	нет	25,320	0,814	0,500
Swin-Classical_48_8_x2	24,879	0,820	0,374	24,480	0,807	0,350
BICUBIC	25,335	0,814	0,500	25,336	0,814	0,500
×4						
Нейросеть	До переобучения			После переобучения		
	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS
LGCNet_x4	нет	нет	нет	20,941	0,450	0,879
SwinClassical_48_8_x4	20,666	0,497	0,819	19,774	0,447	0,762
BICUBIC	20,946	0,450	0,879	20,946	0,450	0,880

Табл. 7 и 8 содержат результаты для двух случаев обработки мультиспектральных изображений. Можно видеть, что метод деления на два RGB-подобных изображения в ложных цветах оказывается значи-

тельно более эффективным, чем поканальное деление на изображения в оттенках серого.

Табл. 6. Показатели качества для задачи SRSISR, SAR-bicubic изображения

×2						
Нейросеть	До переобучения			После переобучения		
	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS
LGCNet_x2	нет	нет	нет	26,472	0,849	0,311
Swin-Classical_48_8_x2	27,949	0,902	0,245	27,495	0,893	0,195
BICUBIC	26,482	0,849	0,312	26,482	0,849	0,311
×4						
Нейросеть	До переобучения			После переобучения		
	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS
LGCNet_x4	нет	нет	нет	21,352	0,467	0,579
SwinClassical_48_8_x4	21,696	0,547	0,546	21,172	0,524	0,498
BICUBIC	21,355	0,467	0,579	21,355	0,467	0,579

Табл. 7. Показатели качества для задачи SRSISR, Multispectral – 3channel изображения

×2						
Нейросеть	До переобучения			После переобучения		
	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS
LGCNet_x2	нет	нет	нет	39,723	0,963	0,056
Swin-Classical_48_8_x2	39,714	0,962	0,047	39,714	0,962	0,047
BICUBIC	39,723	0,963	0,075	39,723	0,963	0,075
×4						
Нейросеть	До переобучения			После переобучения		
	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS
LGCNet_x4	нет	нет	нет	35,297	0,896	0,321
SwinClassical_48_8_x4	35,525	0,902	0,256	35,525	0,903	0,256
BICUBIC	35,298	0,610	0,380	35,298	0,610	0,380

Табл. 8. Показатели качества для задачи SRSISR, Multispectral – single channel изображения

×2						
Нейросеть	До переобучения			После переобучения		
	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS
LGCNet_x2	нет	нет	нет	22,670	0,636	0,434
SwinClassical_48_8_x2	22,651	0,630	0,436	22,650	0,630	0,435
BICUBIC	39,723	0,963	0,075	39,723	0,963	0,075
×4						
Нейросеть	До переобучения			После переобучения		
	PSNR	SSIM	LPIPS	PSNR	SSIM	LPIPS
LGCNet_x4	нет	нет	нет	20,689	0,518	0,630
SwinClassical_48_8_x4	20,645	0,508	0,591	20,645	0,508	0,591
BICUBIC	35,298	0,610	0,3801	35,298	0,610	0,380

В целом по серии экспериментов можно сделать вывод, что:

- дообучение SISR-моделей на изображениях ДЗЗ позволяет улучшить их качество для задачи SRSIRS. То есть модели SISR могут быть использованы и для решения задач SRSIRS посредством обычного переобучения;
- модели ГНС, (до)обученные на одном типе ДЗЗ данных (в работе – на оптических RGB-изображениях), могут достаточно успешно применяться для повышения разрешения других типов оптических данных ДЗЗ. При этом они показывают качественные показатели лучше, чем базовый алгоритм (BICUBIC);
- качество повышения разрешения PCA-изображений для ГНС-моделей не имеет преимущества перед наиболее простым базовым подходом (BICUBIC), а также значительно отстает от качества повышения разрешения для оптических изображений. В качестве модели синтеза LR-изображений по итоговым характеристикам более предпочтительным является вариант «без шума» – SAR-bicubic.

Результатом достижения основных целей исследования являются следующие положения:

- вариант предварительной подготовки типа «Деление многоканальных изображений на трехканальные с пересечением» является предпочтительным;
- получены конкретные обученные модели ГНС для всех типов изображений – оптических (панхроматических, RGB, мультиспектральных) и PCA;
- *наилучшей среди исследованных ГНС-моделей для задачи SRSIRS является модель SwinIR, а конкретно SwinIR-Classical/Lightweight. Она показывает наиболее стабильные, а в большинстве случаев и наилучшие результаты.*

Заключение

В работе представлен единый нейросетевой метод повышения разрешения для разнородных цифровых изображений ДЗЗ. В качестве каркаса метода выступает ГНС-модель трансформера *SwinIR-Classical или Lightweight* (в зависимости от числа параметров), дополненная процедурой модификации входных изображений ДЗЗ, число слоев которых отлично от трех. Полученное решение показывает наилучшие или близкие к ним показатели качества повышения разрешения на всем спектре изображений космического ДЗЗ: панхроматических, полноцветных RGB, мультиспектральных и PCA-изображений, полученных радаром с синтезированной апертурой. Для проведения такого сравнения в работе собраны четыре независимых набора данных (RGB, мультиспектральные, панхроматические и PCA-изображения), часть из которых имела пары LR- и HR-изображений, а для другой формировались LR-изображения путем моделирования процедуры де-

градации. Сравнение моделей производилось как с преднастроенными, так и с переобученными ГНС-претендентами, которые в обзорах фигурируют как state-of-the-art: сверточные, генеративно-состязательные и трансформенные (основанные на механизме внимания).

Конечным результатом работы являются: рекомендованная наилучшая модель ГНС *SwinIR-Classical или Lightweight* для задачи SRSIRS, набор обученных алгоритмов (весов указанной модели) для задач повышения разрешения цифровых изображений ДЗЗ (RGB, мультиспектральные, панхроматические и PCA-изображения) для факторов «×2» и «×4», а также наилучший по результатам сравнений способ предварительной обработки изображений ДЗЗ для приведения их к трехканальному виду.

Результаты работы могут быть использованы для построения наземных комплексов автоматического формирования покрытий/мозаик спутниковыми изображениями территорий Земли.

Благодарности

Результаты исследования были получены при финансовой поддержке Российской Федерации в лице Минобрнауки России (Соглашение № 075-15-2024-558).

References

- [1] Wang Y, Bashir SMA, Khan M, Ullah Q, Wang R, Song Y, Guo Z, Niu Y. Remote sensing image super-resolution and object detection: Benchmark and state of the art. *Expert Syst Appl* 2022; 197: 116793. DOI: 10.1016/j.eswa.2022.116793.
- [2] Meng X, Shen H, Li H, Zhang L, Fu R. Review of the pansharpening methods for remote sensing images based on the idea of meta-analysis: Practical discussion and challenges. *Inf Fusion* 2018; 46: 102-113. DOI: 10.1016/j.inffus.2018.05.006.
- [3] Schowengerdt RA. Remote sensing, models, and methods for image processing. San Diego: Academic Press; 1997. ISBN: 0-12-628981-6.
- [4] Lia J, Hong D, Gao L, Yao J, Zheng K, Zhang B, Chanussot J. Deep learning in multimodal remote sensing data fusion: A comprehensive review. *Int J Appl Earth Obs Geoinf* 2022; 112: 102926. DOI: 10.1016/j.jag.2022.102926.
- [5] Malhotra R, Singh P. Recent advances in deep learning models: a systematic literature review. *Multimed Tools Appl* 2023; 82(29): 44977-45060. DOI: 10.1007/s11042-023-15295-z.
- [6] Wang X, Yi J, Guo J, Song Y, Lyu J, Xu J, Yan W, Zhao J, Cai Q, Min H. A review of image super-resolution approaches based on deep learning and applications in remote sensing. *Remote Sens* 2022; 14(21): 5423. DOI: 10.3390/rs14215423.
- [7] Reshad Md, Burks R, Kwan C, Li J. Deep learning for remote sensing image super-resolution. 2019 IEEE 10th Annual Ubiquitous Computing, Electronics & Mobile Communication Conf (UEMCON) 2019: 286-292. DOI: 10.1109/UEMCON47517.2019.8993047.
- [8] Zhao X, Wang L, Zhang Y. A review of convolutional neural networks in computer vision. *Artif Intell Rev* 2024; 57(4): 99. DOI: 10.1007/s10462-024-10721-6.

- [9] Wu J, Cong R, Fang L. Unpaired remote sensing image super-resolution with content-preserving weak supervision neural network. *Sci China Inf Sci* 2023; 66: 119105. DOI: 10.1007/s11432-021-3575-1.
- [10] Wang P, Zhang H, Zhou F, Jiang Z. Unsupervised remote sensing image super-resolution using cycle CNN. *IGARSS 2019 – 2019 IEEE Int Geoscience and Remote Sensing Symposium* 2019; 3117-3120. DOI: 10.1109/IGARSS.2019.8898648.
- [11] Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville AC, Bengio Y. Generative adversarial nets. *Neural Information Processing Systems* 2014: 1-9.
- [12] Vaswani A, et al. Attention is all you need. *NIPS'17: Proc 31st Int Conf on Neural Information Processing Systems* 2017: 6000-6010.
- [13] Kolesnikov A, Dosovitskiy A, Weissenborn D, Heigold G, Uszkoreit J, Beyer L, Minderer M, Dehghani M, Houselby N, Gelly S, Unterthiner T, Zhai X. An image is worth 16x16 words: Transformers for image recognition at scale. *Int Conf on Learning Representations* 2021: 1-21.
- [14] Wang P, Bayram B, Sertel E. A comprehensive review on deep learning based remote sensing image super-resolution methods. *Earth-Sci Rev* 2022; 232: 104110. DOI: 10.1016/j.earscirev.2022.104110.
- [15] Tuna C, Unal G, Sertel E. Single-frame super resolution of remote-sensing images by convolutional neural networks. *Int J Remote Sens* 2018; 39(8): 2463-2479. DOI: 10.1080/01431161.2018.1425561.
- [16] Zhou S, Zhang J, Zuo W, Loy CC. Cross-scale internal graph neural network for image super-resolution. *34th Conf on Neural Information Processing Systems (NeurIPS 2020)* 2020; 1-13.
- [17] Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. *IEEE Trans on Image Process* 2004; 13(4): 600-612. DOI: 10.1109/TIP.2003.819861.
- [18] Zhang R, Isola P, Efros AA, Shechtman E, Wang O. The unreasonable effectiveness of deep features as a perceptual metric. *2018 IEEE/CVF Conf on Computer Vision and Pattern Recognition* 2018; 586-595. DOI: 10.1109/CVPR.2018.00068.
- [19] Dong C, Loy CC, He K, Tang X. Image super-resolution using deep convolutional networks. *IEEE Trans Pattern Anal Machine Intell* 2016; 38(2): 295-307. DOI: 10.1109/TPAMI.2015.2439281.
- [20] Kim J, Lee JK, Lee KM. Accurate image super-resolution using very deep convolutional networks. *Proc 2016 IEEE Conf on Computer Vision and Pattern Recognition* 2016: 1646-1654. DOI: 10.1109/CVPR.2016.182.
- [21] Ji Y, Zhang H, Gao F, Sun H, Wei H, Wang N, Yang B. LGCNet: A local-to-global context-aware feature augmentation network for salient object detection. *Inf Sci* 2017; 584: 399-416. DOI: 10.1016/j.ins.2021.10.055.
- [22] Kui J, Zhongyuan W, Pen Y, Junjun J. A progressively enhanced network for video satellite imagery superresolution. *IEEE Signal Process Lett* 2018; 25(11): 1630-1634. DOI: 10.1109/LSP.2018.2870536.
- [23] Rika S, Bayu SA, Donny D, Ketut W. Generative adversarial network with residual dense generator for remote sensing image super resolution. *ICRAMET* 2020: 34-39. DOI: 10.1109/ICRAMET51080.2020.9298648.
- [24] Jiang K, Wang Z, Yi P, Jiang J, Xiao J, Yao Y. Deep distillation recursive network for remote sensing imagery super-resolution. *Remote Sens* 2018; 10(11): 1700. DOI: 10.3390/rs10111700.
- [25] Ledig C, Theis L, Huszár F, Caballero J, Cunningham A, Acosta A, Aitken A, Tejani A, Totz J, Wang Z, Shi W. Photo-realistic single image super-resolution using a generative adversarial network. *Proc IEEE Conf on Computer Vision and Pattern Recognition* 2017: 4681-4690. DOI: 10.1109/CVPR.2017.19.
- [26] Wang X, Xie L, Dong C, Shan Y. Real-ESRGAN: Training real-world blind super-resolution with pure synthetic data. *2021 IEEE/CVF Int Conf on Computer Vision Workshops (ICCVW)* 2021: 1905-1914. DOI: 10.1109/ICCVW54120.2021.00217.
- [27] Kui J, Wang Z, Yi P, Wang G, Lu T, Jiang J. Edge-enhanced GAN for remote sensing image superresolution. *IEEE Trans Geosci Remote Sens* 2019; 57(8): 5799-5812. DOI: 10.1109/TGRS.2019.2902431.
- [28] Zhang, Yulun, Kungpeng Li, Kai Li, Lichen Wang, Bineng Zhong, Yun Fu. Image super-resolution using very deep residual channel attention networks. In Book: Ferrari V, Hebert M, Sminchisescu C, Weiss Y, eds. *Computer Vision – ECCV 2018*. 15th European Conference, Munich, Germany, September 8-14, 2018, Proceedings, Part VII. Cham, Switzerland: Springer Nature Switzerland AG; 2018: 286-301. DOI: 10.1007/978-3-030-01234-2_18.
- [29] Haut JM, Fernandez-Beltran R, Paoletti ME, Plaza J, Plaza A. Remote sensing image superresolution using deep residual channel attention. *IEEE Trans Geosci Remote Sens* 2019; 57(11): 9277-9289. DOI: 10.1109/TGRS.2019.2924818.
- [30] Niu B, Wen W, Ren W, Zhang X, Yang L, Wang S, Zhang K, Cao X, Shen H. Single image super-resolution via a holistic attention network. In Book: Vedaldi A, Bischof H, Brox T, Frahm J-M, eds. *Computer Vision – ECCV 2020*. 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XII. Cham, Switzerland: Springer Nature Switzerland AG; 2020: 191-207. DOI: 10.1007/978-3-030-58610-2_12.
- [31] Dai T, Cai J, Zhang Y, Xia S-T, Zhang L. Second-order attention network for single image super-resolution. *2019 IEEE/CVF Conf on Computer Vision and Pattern Recognition (CVPR)* 2019: 11057-11066. DOI: 10.1109/CVPR.2019.01132.
- [32] Zhang D, Shao J, Li X, Shen HT. Remote sensing image super-resolution via mixed high-order attention network. *IEEE Trans Geosci Remote Sens* 2021; 59(6): 5183-5196. DOI: 10.1109/TGRS.2020.3009918.
- [33] Wang P, Bayram B, Sertel E. Super-resolution of remotely sensed data using channel attention based deep learning approach. *Int J Remote Sens* 2021; 42(16): 6048-6065. DOI: 10.1080/01431161.2021.1934598.
- [34] Liang J, Cao J, Sun G, Zhang K, Van Gool L, Timofte R. SwinIR: Image restoration using swin transformer. *2021 IEEE/CVF Int Conf on Computer Vision Workshops (ICCVW)* 2021: 1833-1844. DOI: 10.1109/ICCVW54120.2021.00210.
- [35] Rossi L, Bernuzzi V, Fontanini T, Bertozzi M, Prati A. Swin2-MoSE: A new single image super-resolution model for remote sensing. *arXiv Preprint*. 2024. Source: <<https://arxiv.org/abs/2404.18924>>. DOI: 10.48550/arXiv.2404.18924.
- [36] Michel J, Vinasco-Salinas J, Inglada J, Hagolle O. SEN2VEN μ S, a dataset for the training of Sentinel-2 super-resolution algorithms. *Data* 2022; 7(7): 96. DOI: 10.3390/data7070096.
- [37] Chen H, He X, Qing L, Wu Y, Ren C, Sheriff RE, Zhu C. Real-world single image super-resolution: A brief review. *Inf Fusion* 2022; 79: 124-145. DOI: 10.1016/j.inffus.2021.09.005.

- [38] Kehtarnavaz N, Kim N. Digital signal processing system-level design using LabVIEW. Newnes; 2005: 43-53. ISBN: 978-0750679145.
- [39] Al-Mekhlafi H, Liu S. Single image super-resolution: a comprehensive review and recent insight. Front Comput Sci 2024; 18: 181702. 2024. DOI: 10.1007/s11704-023-2588-9.
- [40] New trend in image restoration and enhancement. 2024. Source: <<https://cvi.lai.net/ntire/2024/>>.
- [41] Bavrina AYu, Myasnikov VV, Sergeev AV. Method of parametric estimation of optoelectronic tract of remote sensed optical image formation. Computer Optics 2011; 35(4): 500-507.
- [42] Bavrina AYu, Myasnikov VV, Sergeyev VV, Tresheva EV, Chupshev NV. Modelling of videoinformational tract of optoelectronic remote sensing systems of Earth: solutions, problems and tasks. Computer Optics 2012; 36(4): 572-585.

Сведения об авторах

Коновалов Виталий Федорович, 1999 года рождения, в 2022 г. окончил Самарский национальный исследовательский университет имени академика С.П. Королева (Самарский университет). В настоящее время работает инженером АО «Самара-Информспутник». Круг научных интересов включает компьютерное зрение, генеративные нейронные сети, вычисления общего назначения на графических процессорах, обработку и восстановление изображений. E-mail: vifvk@gmail.com

Информацию об авторах **Мясников Владислав Валерьевич** и **Сергеев Владислав Викторович** см. стр. 943 этого номера.

ГРНТИ: 89.57.35

Поступила в редакцию 02 сентября 2024 г. Окончательный вариант – 16 сентября 2024 г.

A unified neural network-based single super-resolution method for heterogeneous digital earth remote sensing images

V.F. Konovalov¹, V.V. Myasnikov¹, V.V. Sergeev¹

¹ Samara National Research University, 443086, Samara, Russia, Moskovskoye Shosse 34

Abstract

This paper is devoted to finding a unified solution for the practical problem of increasing the resolution of heterogeneous digital images of remote sensing of the Earth: panchromatic images, color images, multispectral and hyperspectral images, as well as images obtained by synthetic aperture radar (SAR). To obtain such a solution, an aggregated dataset was collected from the existing datasets, including both pairs of high- and low-resolution images and individual images acting as high-resolution images. For the collected database and a typical distortion model, an experimental comparison of various modern neural network models of resolution enhancement (including pre-tuned and retrained options) was carried out, which are listed in reviews as state-of-the-art: convolutional, generative adversarial and transform (based on the attention mechanism). Taking into account possible limitations of the models on the number of layers of space images, two options for image preprocessing were considered. The final result of the work is a single neural network model for increasing the resolution of digital space images, supplemented by adapters for their various types, which, when trained, surpasses or is not inferior to modern specialized state-of-the-art solutions.

Keywords: digital remote sensing images, image enhancement, multispectral and hyperspectral images, radar images.

Citation: Konovalov VF, Myasnikov VV, Sergeev VV. A unified neural network-based single super-resolution method for heterogeneous digital earth remote sensing images. *Computer Optics* 2024; 48(6): 944-955. DOI: 10.18287/2412-6179-CO-1610.

Acknowledgements: This work was supported by the Ministry of Science and Higher Education, Russia (Agreement No.075-15-2024-558).

Author's information

Vitaly Fedorovich Konovalov, (b. 1999), graduated from Samara National Research University (Samara University) in 2022. Currently works as an engineer at Samara-InformSputnik. Research interests: generative neural networks, computer vision, GPGPU computing, generative neural networks, image processing and image restoration.
E-mail: vitfvk@gmail.com

Vladislav Valerievich Myasnikov, received his DrSc degree in Physics & Maths (2008). Currently he works as a professor at the Geoinformatics and Information Security department in Samara National Research University. The range of scientific interests: computer vision, pattern recognition and artificial intelligence, machine learning and geoinformatics. He has about 200 publications, including more than 100 articles and three monographs.
E-mail: vmyas@geosamara.ru

Vladislav Viktorovich Sergeev, (b. 1951) graduated from Samara National Research University (Samara University) in 1974. In 1978 he received the Candidates of Technical Sciences degree and in 1993 – the Doctor of Technical Sciences degree. Currently he works as the head of Geoinformatics and Information Security department at Samara University. Research interests are digital signals and image processing, GIS and pattern recognition.
E-mail: vserg@geosamara.ru

Received September 02, 2024. The final version – September 16, 2024.
