



**Сибирский государственный университет науки и технологий имени М.Ф. Решетнева**

# **Улучшение разрешения снимков ДЗЗ на основе глубоких генеративно- состязательных сетей**

**Профессор, д.т.н.  
Доцент, к.т.н.**

**Маргарита Николаевна Фаворская  
Андрей Иванович Пахирка**

**Бердск 2023**

# Области применения SR технологий

## ***Анализ снимков ДЗЗ***

- Обнаружение малоразмерных объектов интереса на снимках ДЗЗ.
- Поиск изменений на снимках ДЗЗ через длительные временные интервалы.

## ***Распознавание лиц***

- Распознавание лиц в условиях городского видеонаблюдения.
- Улучшение архивных фотографий.

## ***Анализ медицинских снимков***

- Улучшение снимков смазов крови.
- Анализ новообразований.

Снимки ДЗЗ высокого разрешения не всегда доступны, поэтому возникает задача алгоритмической реконструкции снимка низкого разрешения (LR, Low Resolution) в снимок так называемого сверхвысокого разрешения (SR, Super Resolution). Снимок SR является синтезированным и не совпадает с реальным снимком высокого разрешения (HR, High Resolution).

С математической точки зрения задача реконструкции изображения из LR снимка является некорректной.

# Нейросетевой подход

Различают методы SISR (Single Image Super-Resolution) и методы MISR (Multi Image Super-Resolution).

Традиционно методы SISR реализовывались на основе интерполяции и реконструкции. Методы на основе глубокого обучения используют интерполяцию в качестве предварительной обработки, а реконструкцию внедряют в саму модель глубокой нейронной сети.

В качестве моделей используют сверточные нейронные сети (СНС) и генеративно-состязательные сети (ГСС).

## ***Основные направления развития моделей ГСС:***

- Улучшение контуров.
- Улучшение текстуры.
- Повышение реалистичности снимков сверхвысокого разрешения.
- Реконструкция мультиспектральных и гиперспектральных снимков.
- Повышение производительности и стабильности обучения.

# Архитектура предлагаемой ГСС

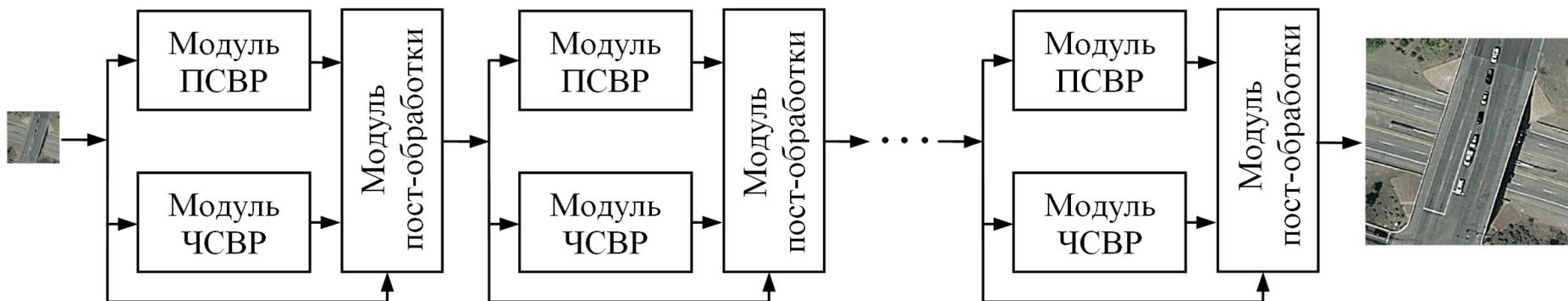
Предлагаемая модель глубокой сети включает три основных компонента:

- состязательное обучение,
- сохранение контуров,
- прогрессивное масштабирование.

Каждый каскад состоит из двух ветвей:

- Для модуля повышения разрешения изображений в качестве базовой модели используется модель ESRGAN (Enhanced Super-Resolution Generative Adversarial Network). Модель ESRGAN имеет модифицированную структуру генератора, в котором слои пакетной нормализации заменены остаточным плотным блоком (RRDB, Residual-in-Residual Dense Block).
- Модуль дискретного вейвлет-преобразования построен на модели DWSR (Deep Wavelet SR). Модель DWSR вначале выполняет дискретное вейвлет-преобразование (ДВП) входного пространственного изображения, затем повышает разрешение изображения в ДВП-представлении и, наконец, выполняет обратное ДВП.

# Архитектура предлагаемой ГСС (продолжение)



Модель содержит несколько идентичных каскадов, повышающих разрешение  $\times 2$ . Каждый каскад обладает сложной структурой, состоящей из трех модулей:

- модуля пространственного сверхвысокого разрешения – модуля ПСВР,
- модуля частотного сверхвысокого разрешения – модуля ЧСВР (с помощью дискретного вейвлет-преобразования),
- модуля пост-обработки для объединения результатов работы двух предыдущих модулей.

# Постановка задачи

В простейшей форме генератор ГСС изучает отображение  $z \rightarrow y$ , где  $z$  – некоторый случайный скрытый вектор, который транслируется в пространство признаков, определяемое выходным вектором  $y$ .

Если ГСС используется для перевода одного изображения в другое, то задача генератора состоит в том, чтобы изучить отображение  $x \rightarrow y$  входного вектора  $x$  на выходной вектор  $y$ . Это делается путем преобразования  $x$  в скрытый вектор  $z$ ,  $x \rightarrow z$ , а затем в выходной вектор  $y$ ,  $z \rightarrow y$ . В данном случае изучается сопоставление входного LR снимка ДЗЗ реконструированному SR снимку ДЗЗ.

ГСС изучают сопоставления между входными и выходными данными с помощью состязательной функции  $\min \max L(G, D)$ , которую реализуют генератор  $G$  и дискриминатор  $D$  с функцией потерей  $L(G, D)$ :

$$\min \max L(G, D) = E_y \left[ \log(D_y) \right] + E_{x,z} \left[ \log(1 - D(x|G(x|z))) \right]. \quad (1)$$

Для оценки потерь генератора  $G$  рассчитывается среднеквадратическая ошибка (потери L2) для предотвращения размывания:

$$L2(G) = E_{x,y,z} \left[ \sqrt{(y - G(x|z))^2} \right]. \quad (2)$$

# Обучение и тестирование

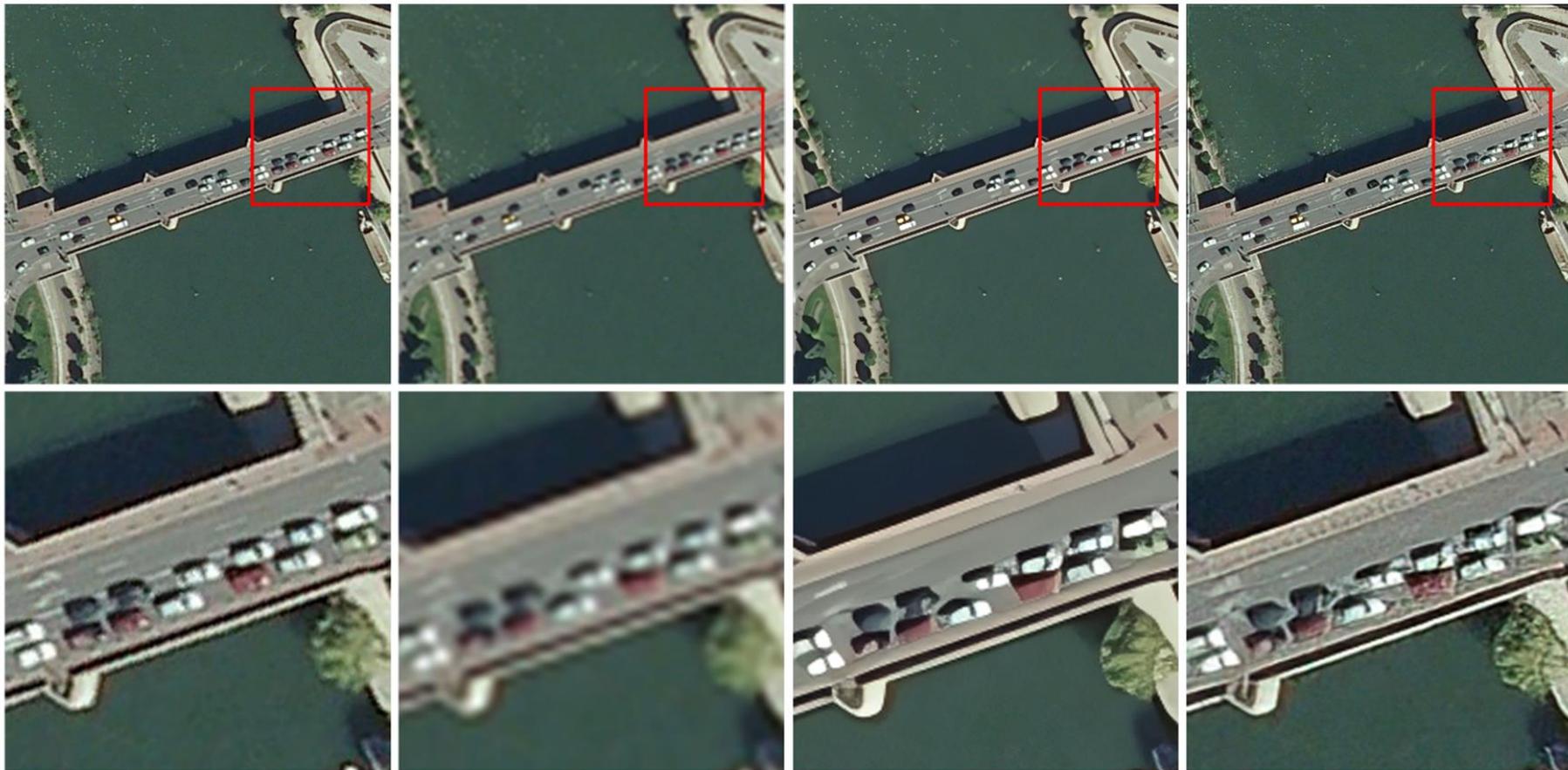
Для обучения и тестирования предложенной модели глубокой сети использовались два открытых набора данных Aerial Image Dataset (AID) и UC-Merced.

Набор данных AID содержит 10000 изображений разрешением 600×600 пикселей, принадлежащих 30 классам (аэропорты, стадионы, поля, пляжи, мосты, коммерческая застройка, пустыня, посевные площади, лес, горы, парки, стоянки, детские площадки, порты, железнодорожные станции, реки, школы, жилые массивы, площади, виадуки и т.д.).

Набор данных UC-Merced содержит более 2000 изображений разрешением 256×256 пикселей, принадлежащих 21 классу. Для аугментации данных все изображения подвергались только горизонтальному отражению.

Предлагаемая модель была реализована на языке Python с использованием репозитория Pytorch. В экспериментах применялась видеокарта NVIDIA Geforce RTX 2080 Ti (11 ГБ). Операционная система – MS Windows 10.

# Примеры SR изображений



а

б

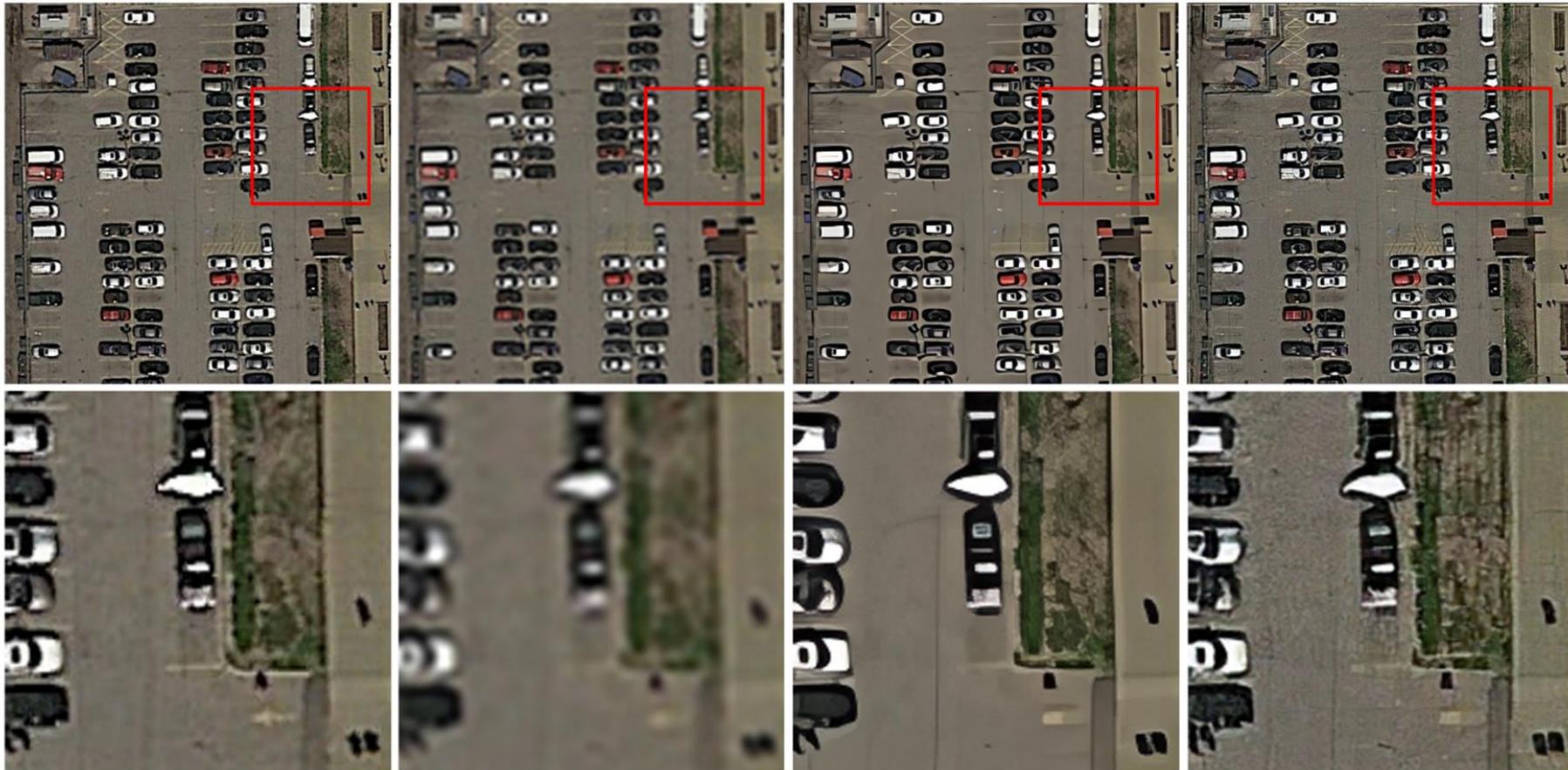
в

г

д

- а) входное изображение с разрешением  $150 \times 150$  пикселей,  
б) оригинальное изображение с разрешением  $600 \times 600$  пикселей,  
в) четырехкратное увеличение с использованием бикубической интерполяции,  
г) четырехкратное увеличение моделью ESRGAN,  
д) четырехкратное увеличение предложенной моделью.

# Примеры SR изображений



а

б

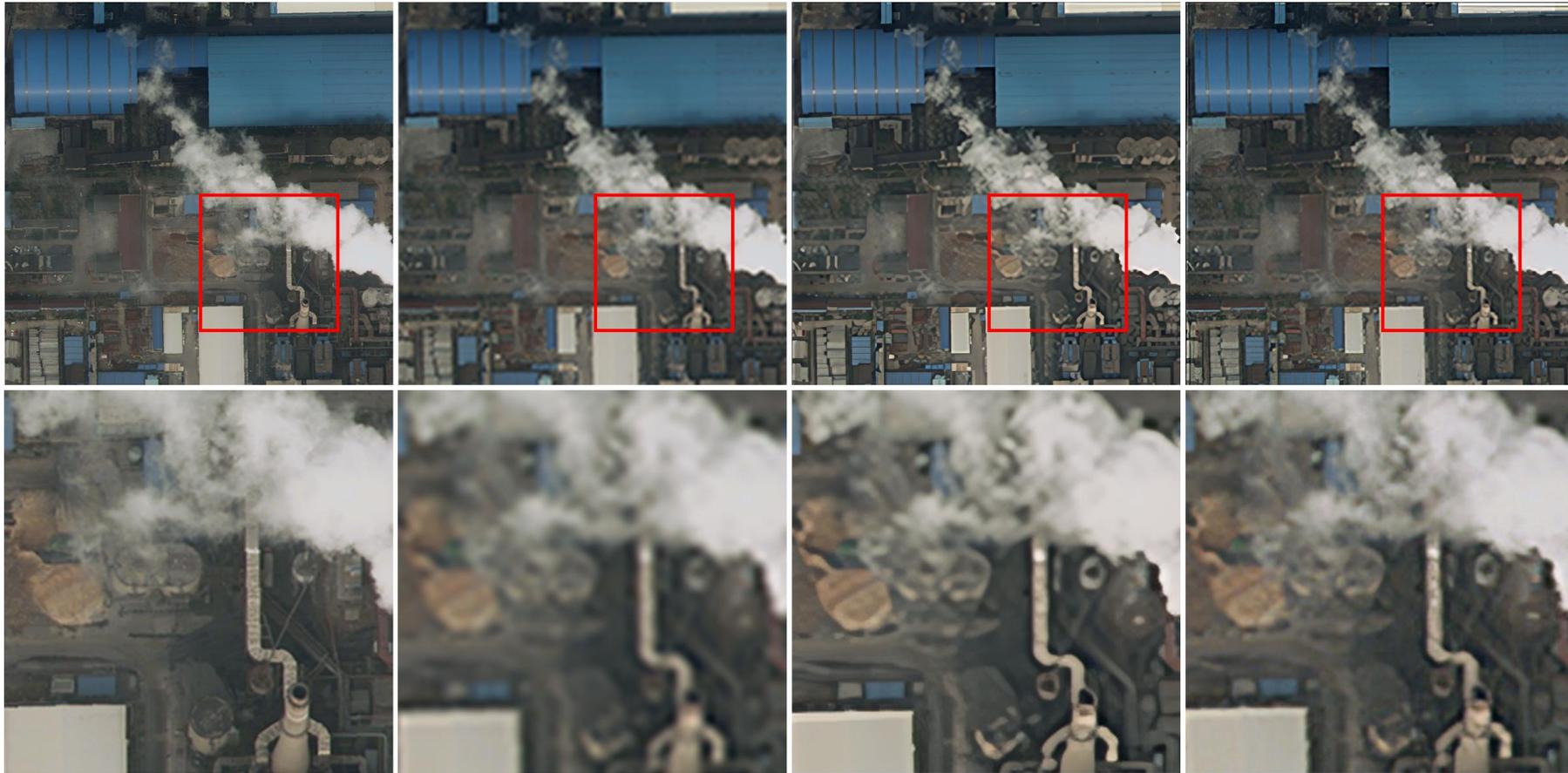
в

г

д

- а) входное изображение с разрешением  $150 \times 150$  пикселей,  
б) оригинальное изображение с разрешением  $600 \times 600$  пикселей,  
в) четырехкратное увеличение с использованием бикубической интерполяции,  
г) четырехкратное увеличение моделью ESRGAN,  
д) четырехкратное увеличение предложенной моделью.

# Примеры SR изображений



а

б

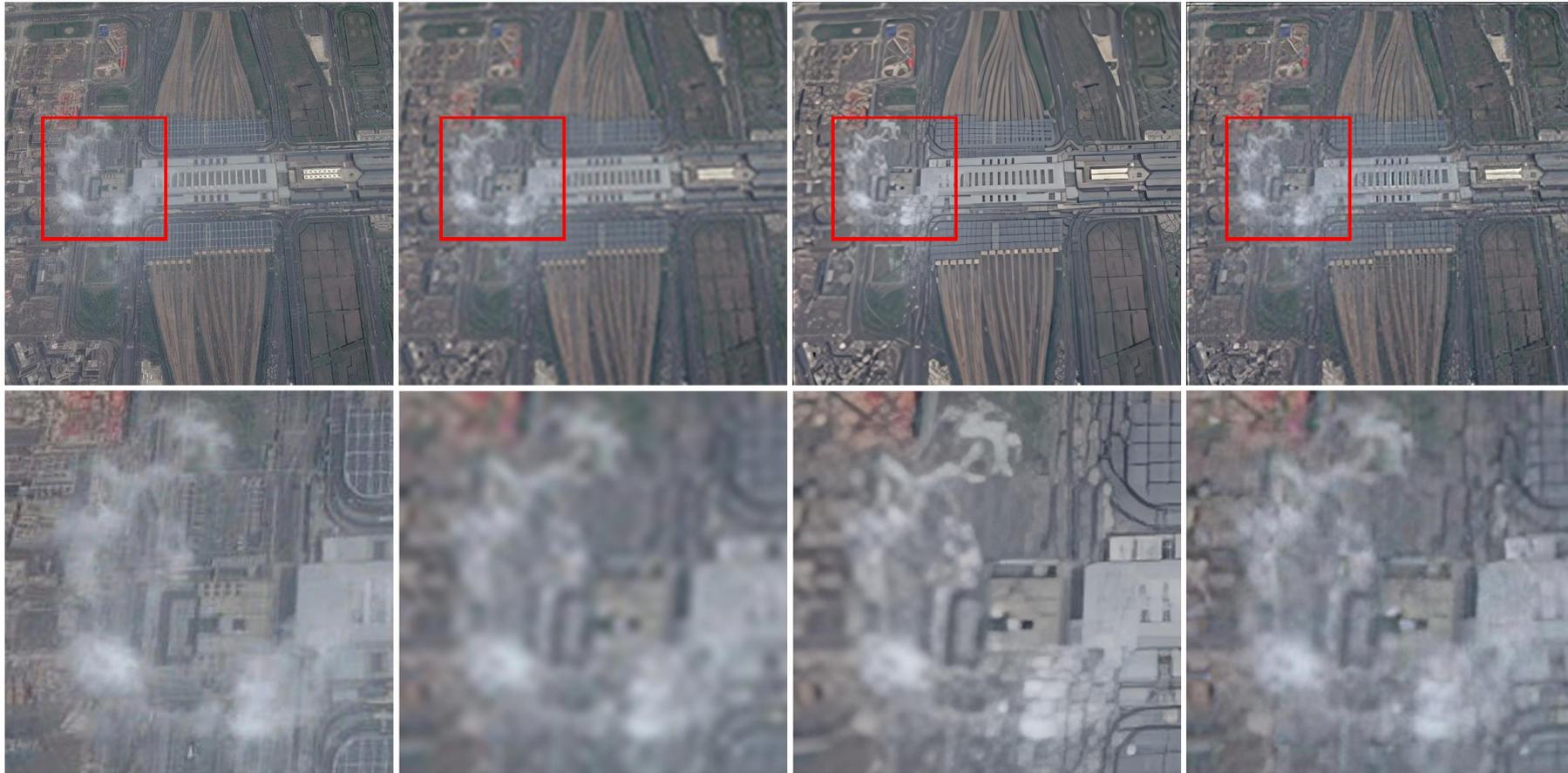
в

г

д

- а) входное изображение с разрешением  $150 \times 150$  пикселей,  
б) оригинальное изображение с разрешением  $600 \times 600$  пикселей,  
в) четырехкратное увеличение с использованием бикубической интерполяции,  
г) четырехкратное увеличение моделью ESRGAN,  
д) четырехкратное увеличение предложенной моделью.

# Примеры SR изображений



а

б

в

г

д

- а) входное изображение с разрешением  $150 \times 150$  пикселей,  
б) оригинальное изображение с разрешением  $600 \times 600$  пикселей,  
в) четырехкратное увеличение с использованием бикубической интерполяции,  
г) четырехкратное увеличение моделью ESRGAN,  
д) четырехкратное увеличение предложенной моделью.

# Метрики оценки изображений

Для оценки качества полученных SR изображений применялись различные метрики:

- Пиковое отношение сигнала к шуму PSNR (Peak Signal-to Noise Ratio).
- Индекс структурного сходства SSIM (Structural SIMilarity).
- Метрика LPIPS (Learned Perceptual Image Patch Similarity) для оценки визуального качества изображений.

Средние значения показателей реконструированных снимков ДЗЗ

Метод	PSNR, дБ	SSIM	LPIPS
Бикубическая интерполяция	28,95	0,89	0,175
ESRGAN	27,76	0,88	0,119
Предложенная модель	28,45	0,88	0,096

# Заключение

- Полученные объективные оценки свидетельствуют о том, что разработанная модель способна обеспечить качественную генерацию изображений сверхвысокого разрешения, сохраняя при этом их текстурные особенности.
- Было выяснено, что небольшая облачность, присутствующая на исходных снимках ДЗЗ, принципиально не влияет на работу предложенной нейросетевой модели.
- Предложенная нейросетевая модель на основе ГСС может генерировать более реалистичные снимки ДЗЗ. Данный подход характеризуется повышенной чувствительностью к деталям изображения.
- Применение вейвлет-преобразования позволяет сохранить контуры и повысить четкость изображения. Будущие модификации разработанной модели направлены на повышение ее производительности.

# Библиографические ссылки

1. *Courtrai L., Pham, M. T., Lefevre S.* Small object detection in remote sensing images based on super-resolution with auxiliary generative adversarial networks // *Remote Sensing*. 2020. V. 12. N. 19. P. 3152.1-3152.19.
2. *Bashir S.M.A., Wang Y.* Small object detection in remote sensing images with residual feature aggregation-based super-resolution and object detector network // *Remote Sensing*. 2021. V. 13. N. 9. P. 1854.1-1854.21.
3. *Parelius E.J.* A review of deep-learning methods for change detection in multispectral remote sensing images // *Remote Sensing*. 2023, V. 15. N. 8. P. 2092.1-2092.30.
4. *Dong C., Loy C.C., He K., Tang X.* Image super-resolution using deep convolutional networks // *IEEE Trans Pattern Anal Mach Intell*. 2016. V. 38. N. 2. P. 295–307.
5. *Kim J., Lee J.K., Lee K.M.* Accurate image super-resolution using very deep convolutional networks // 2016 IEEE Conference on Computer Vision and Pattern Recognition. 2016. IEEE: Las Vegas, NV, USA. P. 1646-1654.
6. *Mao X.J., Shen C., Yang Y.B.* Image restoration using convolutional auto-encoders with symmetric skip connections // In: Lee D.D., Sugiyama M., Luxburg U.V., Guyon I., Garnett R. (eds) *Advances in Neural Information Processing Systems: Annual Conference on Neural Information Processing Systems 2016, Barcelona, Spain*. 2016. ANIPS, V. 29. P. 2802-2810.
7. *Zhang Y., Li K., Li K., Wang L., Zhong B., Fu Y.* Image super-resolution using very deep residual channel attention networks // In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds) *Computer Vision – ECCV 2018*. 2018. Springer: Cham. LNCS, V. 11211. P. 294–310.

## Библиографические ссылки (продолжение)

8. *Ledig C., Theis L., Huszar F., Caballero J., Aitken A., Tejani A., Totz J., Wang Z., Shi W.* Photo-realistic single image super-resolution using a generative adversarial network // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017. IEEE: Honolulu, HI, USA. P. 105-114.

9. *Hou B., Liu Q., Wang H., Wang Y.* From W-Net to CDGAN: Bi-temporal change detection via deep learning techniques // IEEE Transactions on Geoscience and Remote Sensing. 2020. V. 58. N. 3. P. 1790-1802.

10. *Salgueiro L., Marcello J., Vilaplana V.* SEG-ESRGAN: A multi-task network for super-resolution and semantic segmentation of remote sensing images // Remote Sensing. 2022. V. 14. N. 22. P. 5862.1-5862.28.

11. *Zhang Y., Yu X., Lu X., Liu P.* Pro-UIGAN: Progressive face hallucination from occluded thumbnails // IEEE Transactions on Image Processing. 2022. V.31. P. 3236-3250.

12. *Muzahid A.A.M., Wanggen W., Sohel F., Bennamoun M., Hou L., Ullah H.* Progressive conditional GAN-based augmentation for 3D object recognition // Neurocomputing. 2021. V. 460. P. 20-30.

13. *Zhou, Q., Yin, H.* A U-Net based progressive GAN for microscopic image augmentation // In: Yang, G., Aviles-Rivero, A., Roberts, M., Schönlieb, CB. (eds) Medical Image Understanding and Analysis. 2022. Springer: Cham. LNCS, V. 13413. P. 458-468.

14. *Wang X., Yu K., Wu S., Gu J., Liu Y., Dong C., Qiao Y., Loy C.C.* ESRGAN: Enhanced super-resolution generative adversarial networks // In: Leal-Taixé, L., Roth, S. (eds) Computer Vision – ECCV 2018 Workshops. 2019. Springer, Cham. LNCS, V. 11133. P. 63-79.

## Библиографические ссылки (продолжение)

15. *Guo T., Mousavi H.S., Vu T.H., Monga V.* Deep wavelet prediction for image super-resolution // 2017 IEEE Conf. Computer Vision and Pattern Recognit. Workshops (CVPRW). 2017. IEEE: Honolulu, HI, USA. P. 104-113.
16. *Xia G.-S., Hu J., Hu F., Shi B., Bai X., Zhong Y., Zhang L.* AID: A benchmark dataset for performance evaluation of aerial scene classification // IEEE Transactions on Geoscience and Remote Sensing. 2017. V. 55. N. 7. P. 3965-3981.
17. UC Merced Land Use Dataset. <http://weegee.vision.ucmerced.edu/datasets/landuse.html> (дата обращения 06.06.2023).
18. *Zhang R., Isola P., Efros A.A., Shechtman E., Wang O.* The unreasonable effectiveness of deep features as a perceptual metric // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2018. IEEE: Salt Lake City, UT, USA. P. 586-595.

*Спасибо за внимание*  
*Ваши вопросы*