

Военно-космическая академия имени А.Ф. Можайского
Кафедра оптико-электронных средств контроля

Подход к улучшению пространственной детальности оптико-электронных снимков с использованием субпиксельной сверточной нейронной сети

Дмитриков Г.Г.

Зуев Л.Г.

Голубков И.К.

Дементьев Д.С.

Санкт-Петербург

2021

Идея исследования

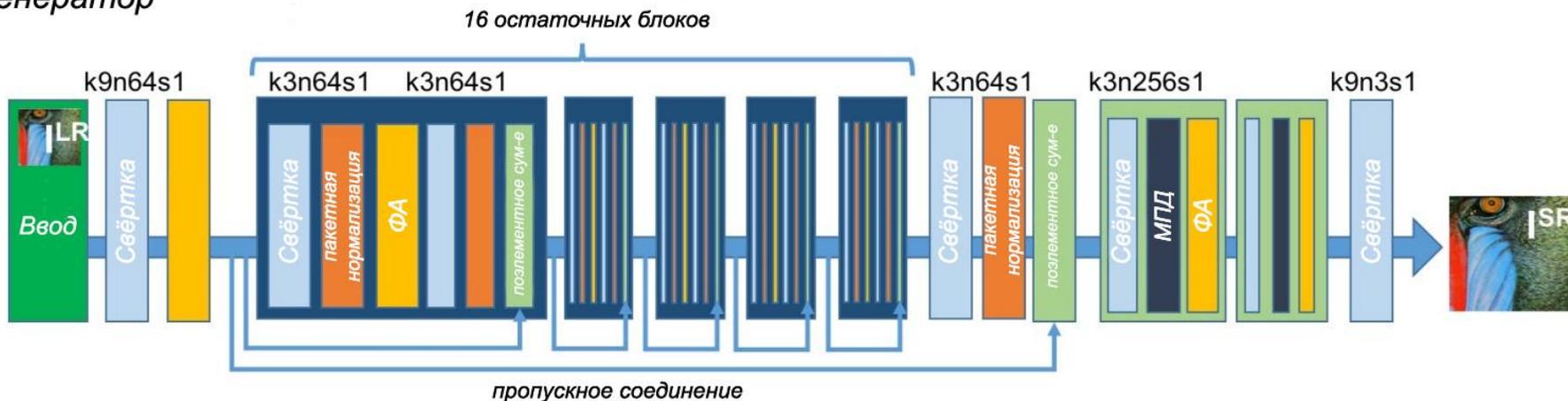
Идея исследования состоит в определении возможности использования архитектуры сети SRGAN для улучшения качества оптико-электронных снимков местности. Подобно архитектурам GAN сеть SRGAN также содержит две части: генератор и дискриминатор, где генератор выдает некоторые данные на основе распределения вероятностей, а дискриминатор пытается угадать данные, поступающие из входного набора данных или генератора. Генератор пытается оптимизировать сгенерированные данные, чтобы «обмануть» дискриминатор.



Рис. Подход к использованию сети SRGAN для улучшения пространственного разрешения оптико-электронных снимков

Генератор и дискриминатор сети SRGAN

Генератор



Дискриминатор

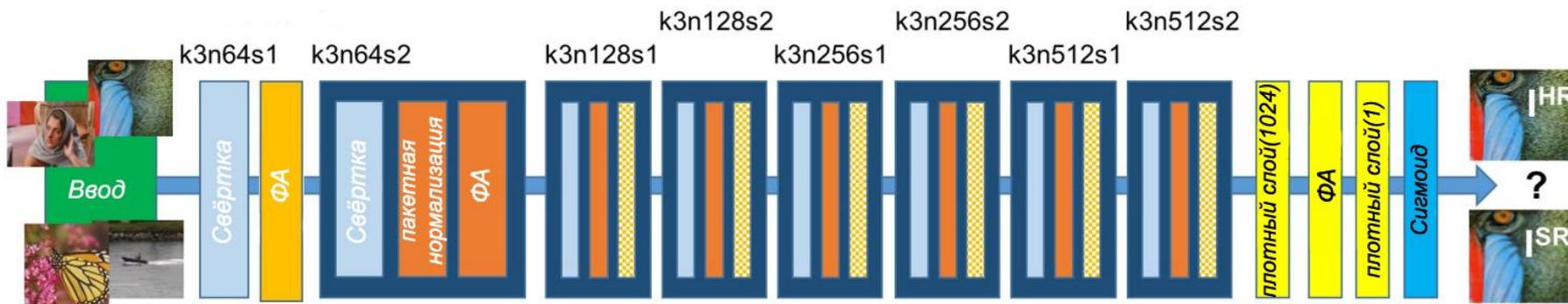


Рис. Архитектура сети генератора и дискриминатора с соответствующим размером ядра (k), количеством карт функций (n) и шаг (s) указаны для каждого сверточного слоя, где ФА – функция активации, МПД – метод повышающей дискретизации

Описание используемых функций потерь

SRGAN использует функцию перспективных потерь (L_{SR}), которая представляет собой взвешенную сумму двух компонентов потерь: потери контента и состязательной потери. Эта потеря очень важна для производительности архитектуры генератора.

Пиксельная потеря MSE: функция потери для архитектуры SRResnet, которая является наиболее распространенной потерей MSE для изображения со сверхвысоким разрешением. Однако потеря MSE не может справиться с высокочастотным содержанием изображения, которое привело к получению чрезмерно гладких изображений.

$$l_{MSE}^{SR} = \frac{1}{r^2WH} \sum_{x=1}^{rW} \sum_{y=1}^{rH} (I_{x,y}^{HR} - G_{\theta_G}(I^{LR})_{x,y})^2$$

Описание используемых функций потерь

Потеря контента: потеря VGG основана на уровнях активации ReLU предварительно обученной 19-уровневой сети VGG. Эта потеря определяется как

$$l_{VGG/i,j}^{SR} = \frac{1}{W_{i,j}H_{i,j}} \sum_{x=1}^{W_{i,j}} \sum_{y=1}^{H_{i,j}} (\phi_{i,j}(I^{HR})_{x,y} - \phi_{i,j}(G_{\theta_G}(I^{LR}))_{x,y})^2$$

Состязательная потеря: функция потерь, которая заставляет генератор создавать изображение, более похожее на изображение с высоким разрешением, с помощью дискриминатора, который обучен различать изображения с высоким и сверхвысоким разрешением

$$l_{Gen}^{SR} = \sum_{n=1}^N -\log D_{\theta_D}(G_{\theta_G}(I^{LR}))$$

Общая потеря контента в архитектуре:

$$l^{SR} = \underbrace{l_X^{SR}}_{\text{потеря контента}} + \underbrace{10^{-3}l_{Gen}^{SR}}_{\text{состязательная потеря}}$$

Общая потеря контента(для потери VGG)

Результаты эксперимента

В приведенной таблице MSE означает, что среднеквадратичная ошибка значения пикселя принята как потеря контента. VGG22 указывает карту характеристик, полученную посредством 2-й свертки (после активации) перед 2-м уровнем максимального объединения в сети VGG19. Потеря VGG вычислена с использованием формулы, описанной выше. Таким образом, эта потеря является потерей функций нижнего уровня. Аналогичным образом VGG 54 использует потери, рассчитанные на карте характеристик, полученной в результате 4-й свертки (после активации) перед 5-м слоем maxpooling в сети VGG19. Это представляет собой потерю функций более высокого уровня из более глубоких сетевых слоев, что позволяет с большим потенциалом анализировать содержание изображений.

Set5	SRResNet-		SRGAN-		
	MSE	VGG22	MSE	VGG22	VGG54
PSNR	32.05	30.51	30.64	29.84	29.40
SSIM	0.9019	0.8803	0.8701	0.8468	0.8472
MOS	3.37	3.46	3.77	3.78	3.58
Set14					
PSNR	28.49	27.19	26.92	26.44	26.02
SSIM	0.8184	0.7807	0.7611	0.7518	0.7397
MOS	2.98	3.15*	3.43	3.57	3.72*

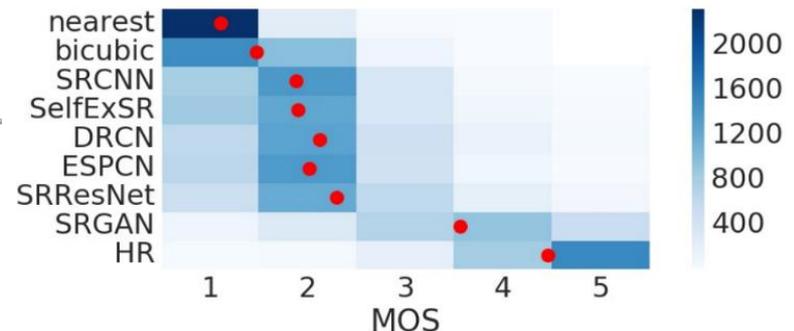


Рис. Сравнение подходов к искусственному улучшению разрешения изображения

Примеры результатов обработки по улучшению пространственного разрешения снимков

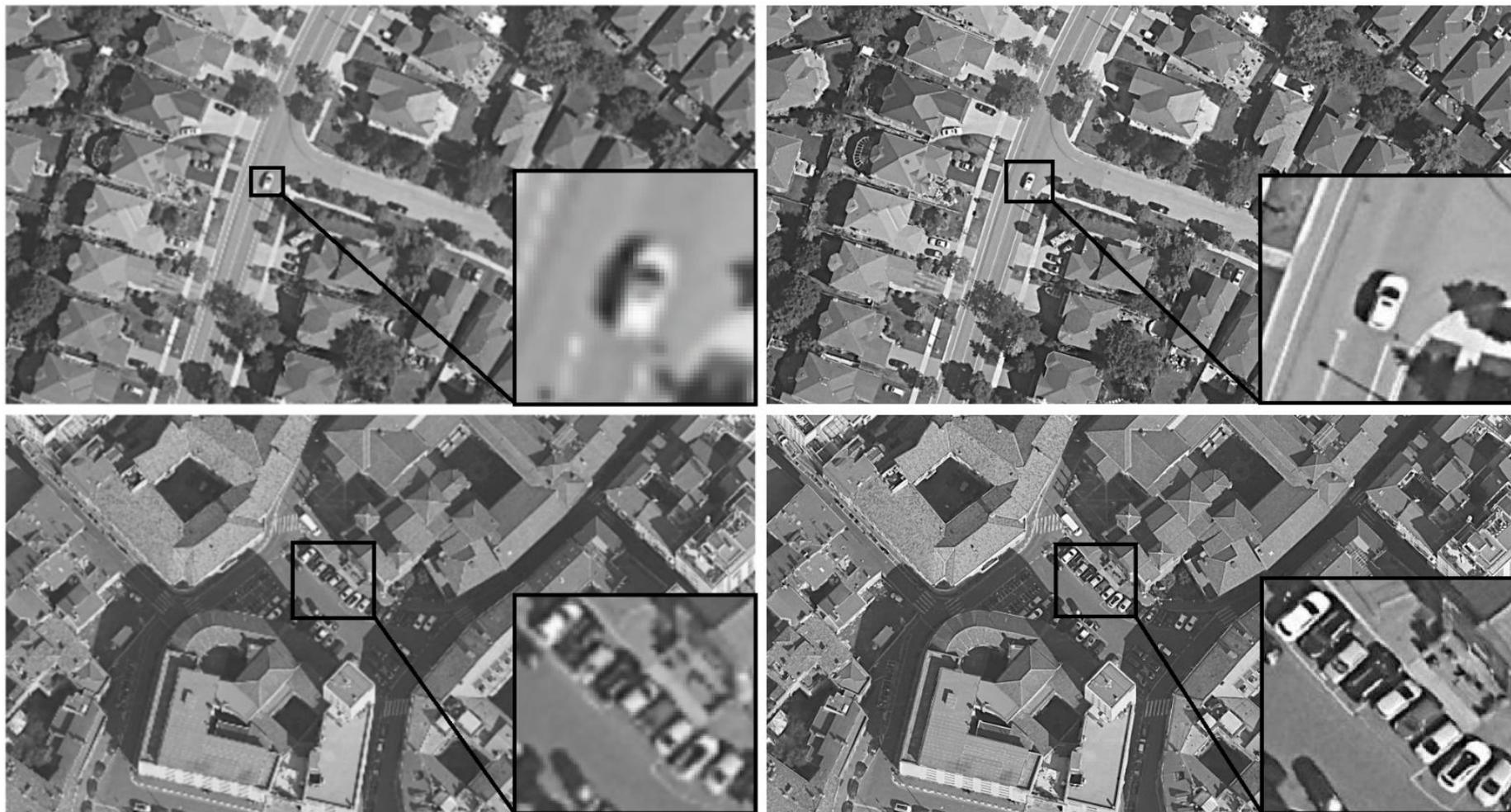


Рис. Исходное изображение с низким разрешением (слева), изображение с улучшенным разрешением, полученное путем использования сети SRGAN (справа)