



«МАТЕМАТИЧЕСКИЕ МЕТОДЫ РАСПОЗНАВАНИЯ ОБРАЗОВ»
18-я ВСЕРОССИЙСКАЯ КОНФЕРЕНЦИЯ С МЕЖДУНАРОДНЫМ УЧАСТИЕМ

26-29 ноября 2019 г.
г. Москва

Улучшение визуального качества изображений в авиационных системах улучшенного видения с использованием генеративных состязательных сетей

К.В. Доброходов, М.А. Лебедев, Ю.В. Визильтер,

Д.В. Комаров, О.В. Выголов

KDobrokhodov@gosniias.ru



ФГУП «Государственный научно-исследовательский институт авиационных систем»

Системы улучшенного видения (СУВ)

СУВ



Современные авиационные СУВ формируют изображения на основе данных, получаемых от различных датчиков технического зрения (чаще всего ТВ- и ИК- диапазонов).

**ТВ-диапазон
(0.4-0.9 мкм)**



**Ближний
ИК-диапазон
(0.9-1.7 мкм)**



**Дальний
ИК-диапазон
(8-14 мкм)**



Улучшение изображений в авиационных системах улучшенного видения

Одним из способов формирования улучшенного изображения является **выравнивание освещённости** с сохранением контраста в ярко и плохо освещенных областях для **улучшения визуального восприятия** летчиками закабинной обстановки.

Появление новых методов и алгоритмов обработки изображений позволяет предложить новые подходы к решению задачи формирования улучшенного изображения.



Исходное ИК



Улучшенное ИК



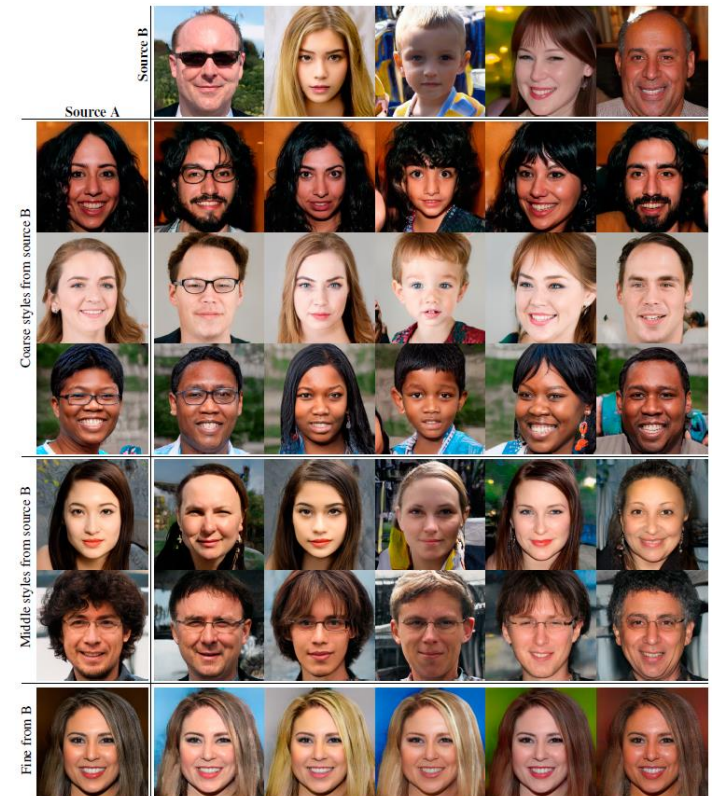
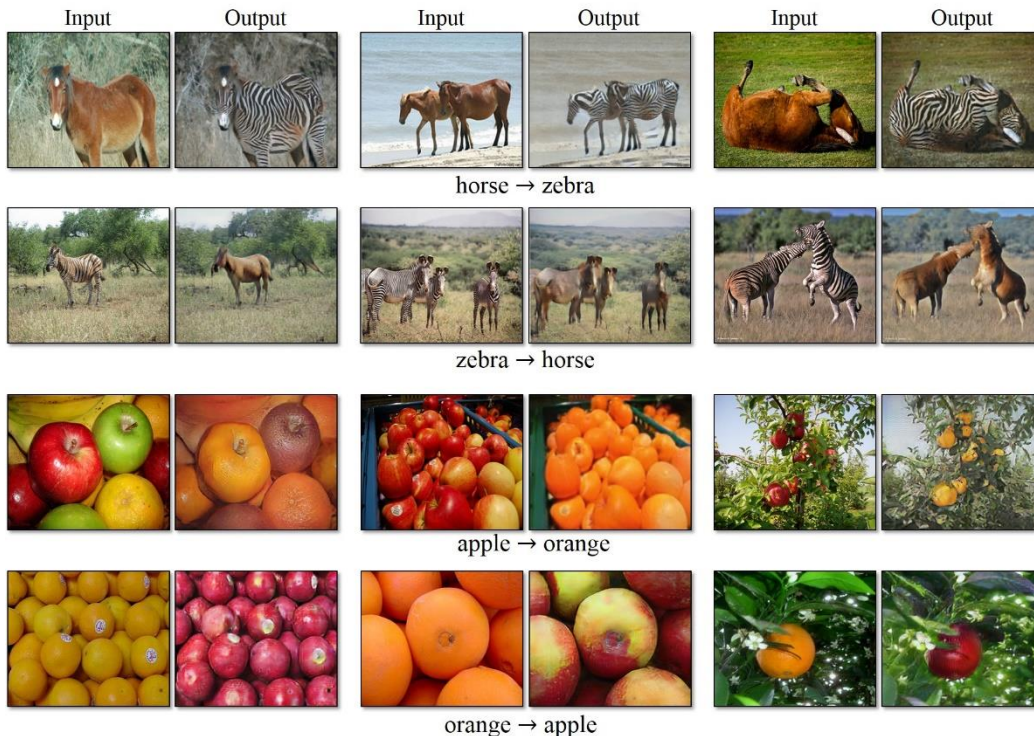
Исходное ТВ



Улучшенное ТВ

Применение нейронных сетей в задачах технического зрения

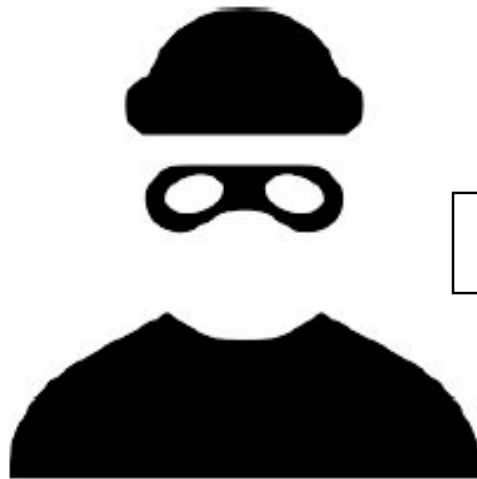
На сегодняшний день наиболее показательные результаты по работе с изображениями демонстрируют глубокие конволюционные нейронные сети (ГКНС), которые превзошли результаты алгоритмов прошлых поколений и существенно увеличили область применения компьютерного зрения. Точность, получаемая ГКНС близка к возможностям человека, а в ряде случаев даже превосходит.



Генеративно-сопоставительная сеть

Для формирования улучшенного изображения предлагается модель, основанная на **генеративно-сопоставительных сетях (GAN)**. Главными составляющими генеративно-сопоставительной сети являются две конкурирующие нейронные сети: генератор G и дискриминатор D .

Для наглядности рассмотрим принцип обучения GAN на примере фальшивомонетчика и машинки для проверки денег.



5000



5000



Схема генератора

В основе генератора **используется сеть «ResNet»**.

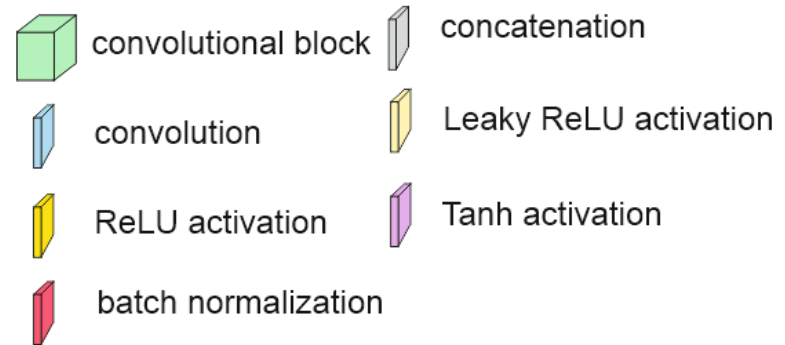
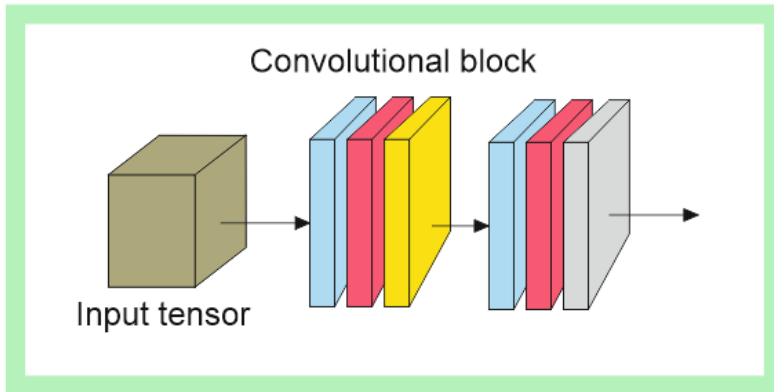
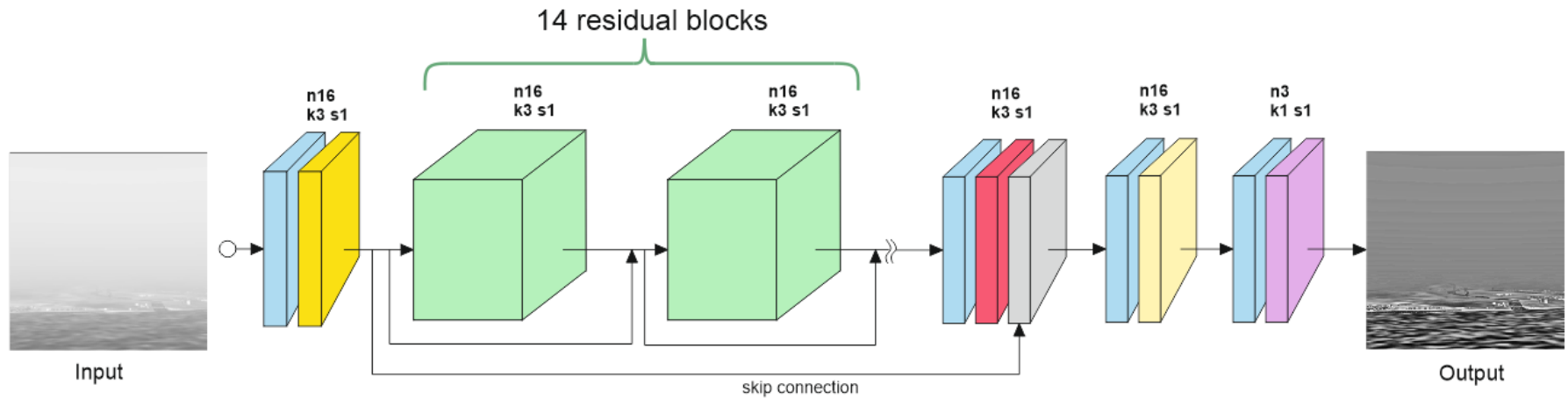
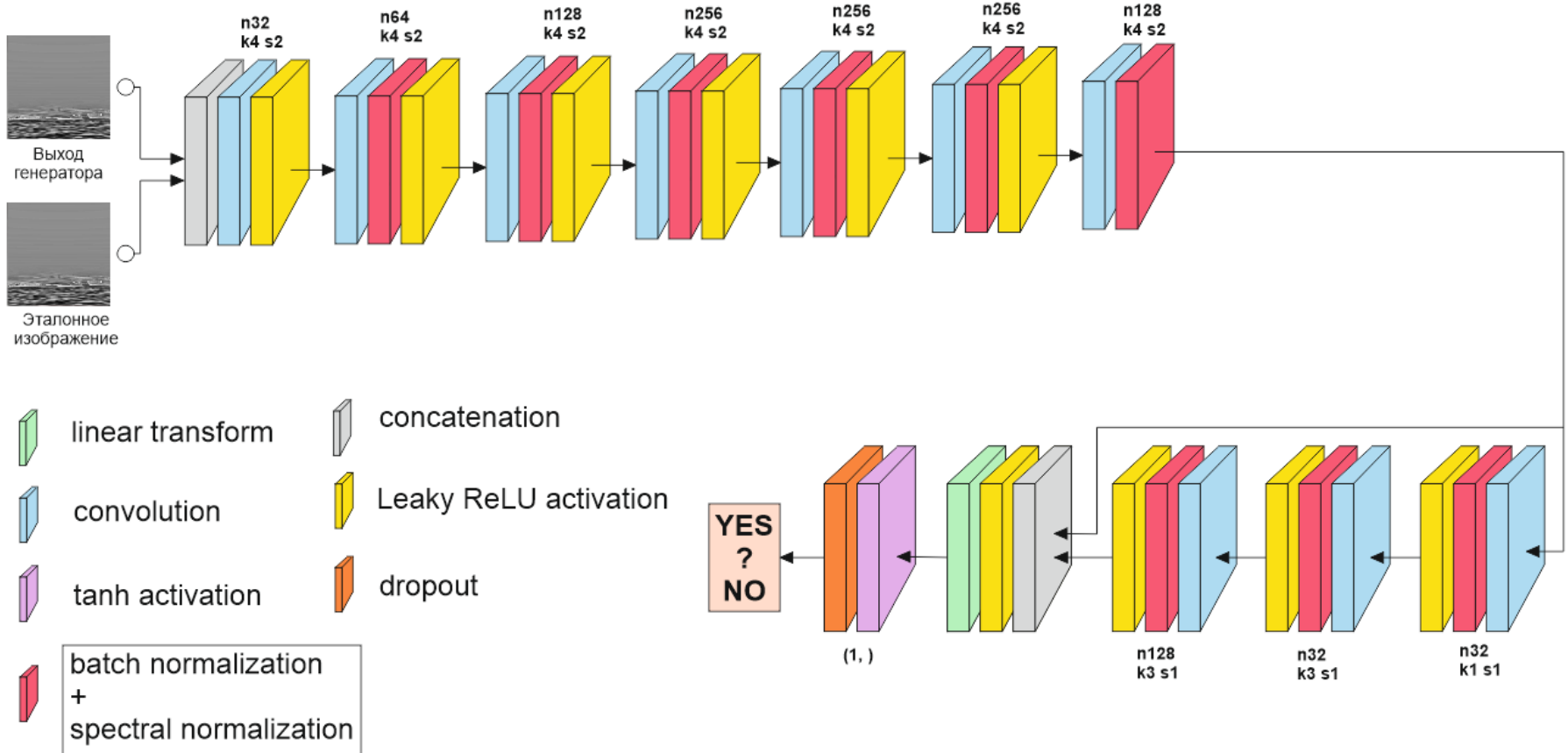


Схема дискриминатора



Исходные данные

Создание обучающей выборки осуществлялось с использованием стендовой базы ФГУП «ГосНИИАС». Выборка состояла из изображений телевизионного (ТВ) и инфракрасного (ИК) каналов разрешением 512x512 пикселей.



ИК



ТВ



Формирование обучающей выборки

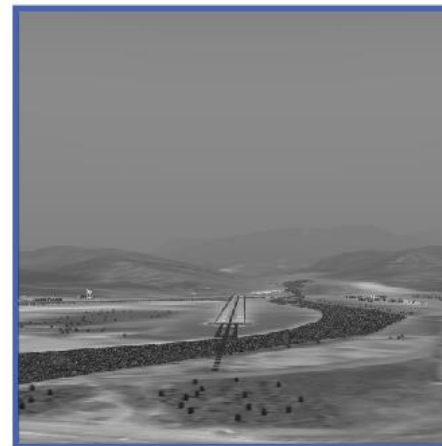
Для обучающей выборки были отобраны 9600 изображений, синтезированных на демонстраторе информационно управляющего поля (ИУП) кабины, на основе которых формируются эталонные улучшенные изображения, получаемые при помощи классического подхода (обработка входных изображений алгоритмом MultiScale Retinex¹).



Исходное ИК



Улучшенное ИК



Исходное ТВ



Улучшенное ТВ

¹ Ana Belén Petro, Catalina Sbert, and Jean-Michel Morel, Multiscale Retinex // Image Processing On Line, pp. 71–88.

Div2K

Для создания большей обучающей выборки, помимо набора данных, сформированных на стенде ИУП были добавлены 2000 изображений из датасета Div2K². Особенность данной выборки заключается в высоком разрешении (от 1500x1500 пикселей), а также вариативности изображений (размеры объекта/фона, количество маленьких деталей), что позволяет улучшить результат работы сети.



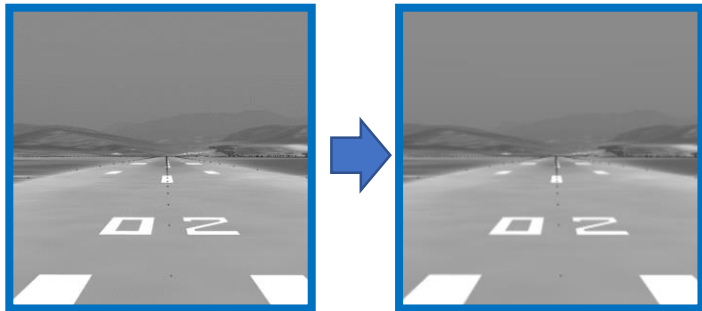
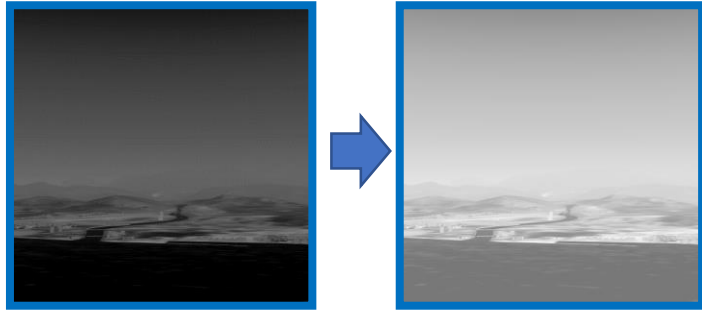
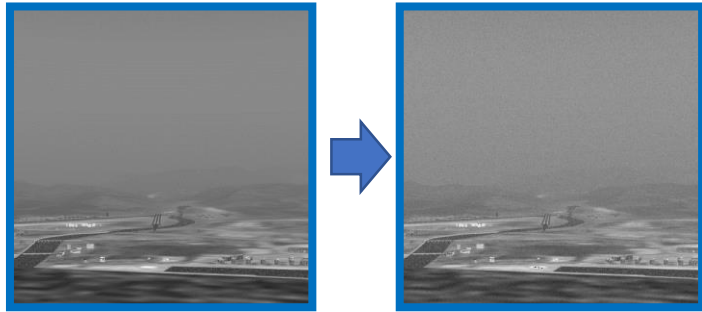
Исходное изображение



PowerNet

²DIV2K Dataset: <https://data.vision.ee.ethz.ch/cvl/DIV2K/>

Аугментация данных



Для приближения данных в обучающей выборке к реальным данным была произведена аугментация данных.

- ❖ Для имитации времени суток были **изменены параметры яркости.**
- ❖ Для имитации свойств реальных сенсоров добавлен **случайный аддитивный шум.**
- ❖ Для имитации тумана, либо нахождения изображений не в фокусе добавлено **случайное гауссовское размытие.**
- ❖ Для имитации увеличения детализации часть изображений подвергались **2-х кратному сжатию с последующим билинейным увеличением** до исходного размера.

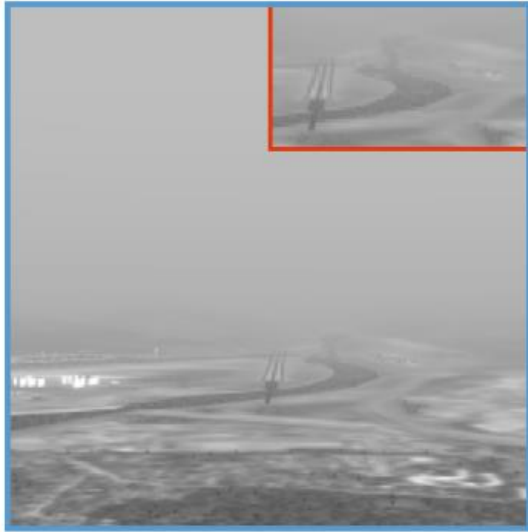
Программная реализация

Предложенная архитектура сети была реализована при помощи открытой программной библиотеки для машинного обучения **TensorFlow** на языке программирования **Python**.

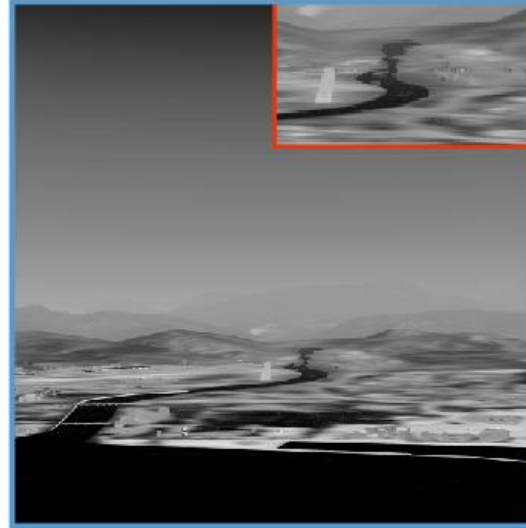
Обучение и тестирование сети осуществлялось на видеокарте **NVIDIA GeForce GTX 1080** с фреймбуфером объемом 8 GB GDDR5X.



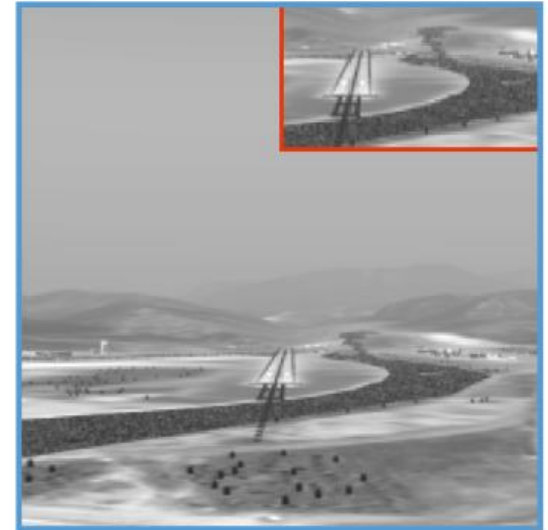
Результаты работы сети PowerNet



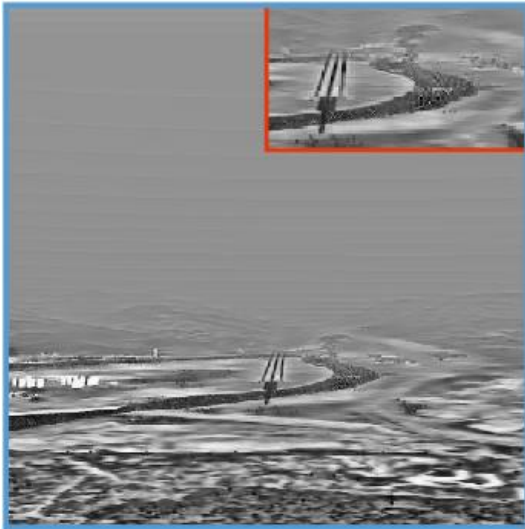
Исходное изображение



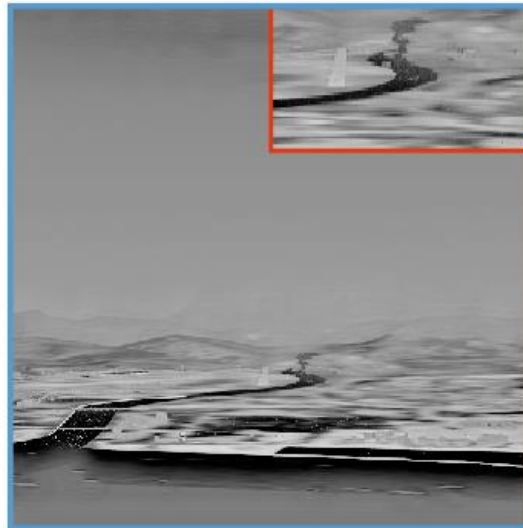
Исходное изображение



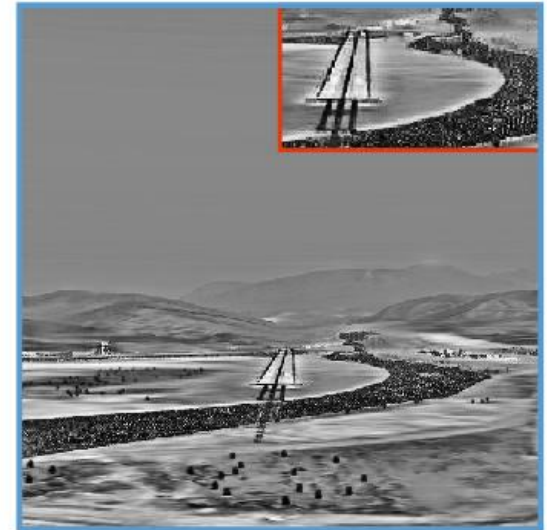
Исходное изображение



PowerNet

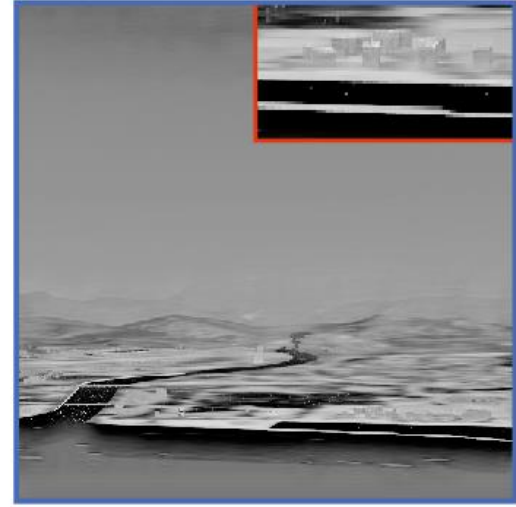
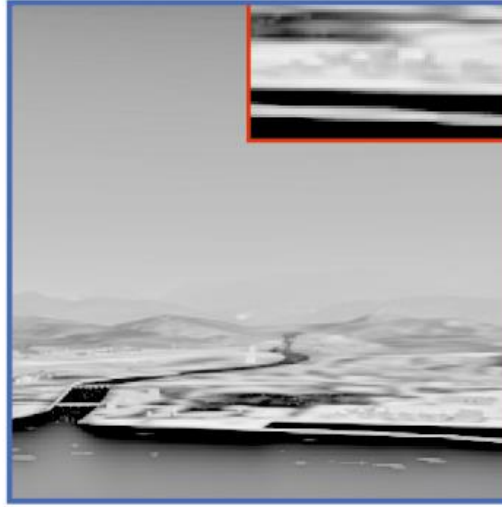
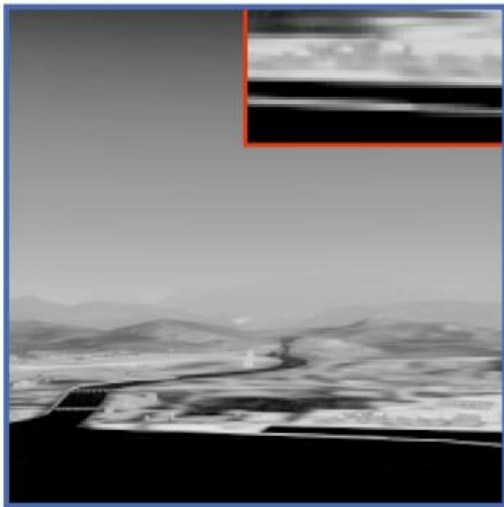


PowerNet



PowerNet

Сравнение результата работы сети PowerNet с классическим алгоритмом при плохих условиях видимости



Исходные изображения

Улучшенное

Результат PowerNet

Результаты работы сети PowerNet на реальных изображениях с бортовой камеры ЛА



Исходное изображение



Исходное изображение



Исходное изображение



PowerNet



PowerNet



PowerNet

Работа сети PowerNet на видеоряде, полученном с бортовой камеры ЛА



Заключение

Предложена оригинальная архитектура генеративной состязательной сети PowerNet для улучшения телевизионного и инфракрасного изображений разрешением 512x512 пикселей. В основе предложенной архитектуры сети генератора использована сеть типа ResNet, которая в совокупности с аугментацией данных при обучении позволяет улучшить изображение за счет повышения детализации изображения, устранения шумов и размытия в виде тумана.

Работоспособность предложенного алгоритма улучшения изображений продемонстрирована как на модельных данных с демонстратора ИУП кабины, так и на реальных изображениях с бортовой камеры ЛА.

Спасибо за внимание!