



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

**26-29 ноября 2019 г.**  
**г. Москва**

# **Автоматическое совмещение изображений в задачах улучшенного и комбинированного видения с использованием генеративных состязательных сетей**

*С.А. Неклюдов, М.А. Лебедев, Ю.В. Визильтер,  
О.В. Выголов, К.В. Доброходов*  
**[sneklyudov@gosniias.ru](mailto:sneklyudov@gosniias.ru)**

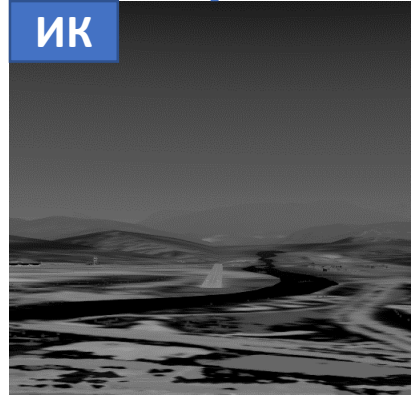
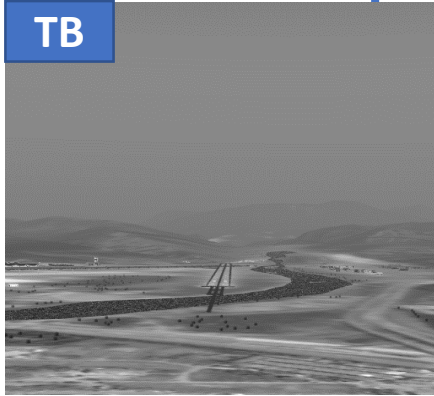


ФГУП «Государственный научно-исследовательский  
институт авиационных систем»

# Системы повышения ситуационной информированности экипажа

Система улучшенного видения (EVS)

Комплексирование



Система синтезированного видения (SVS)

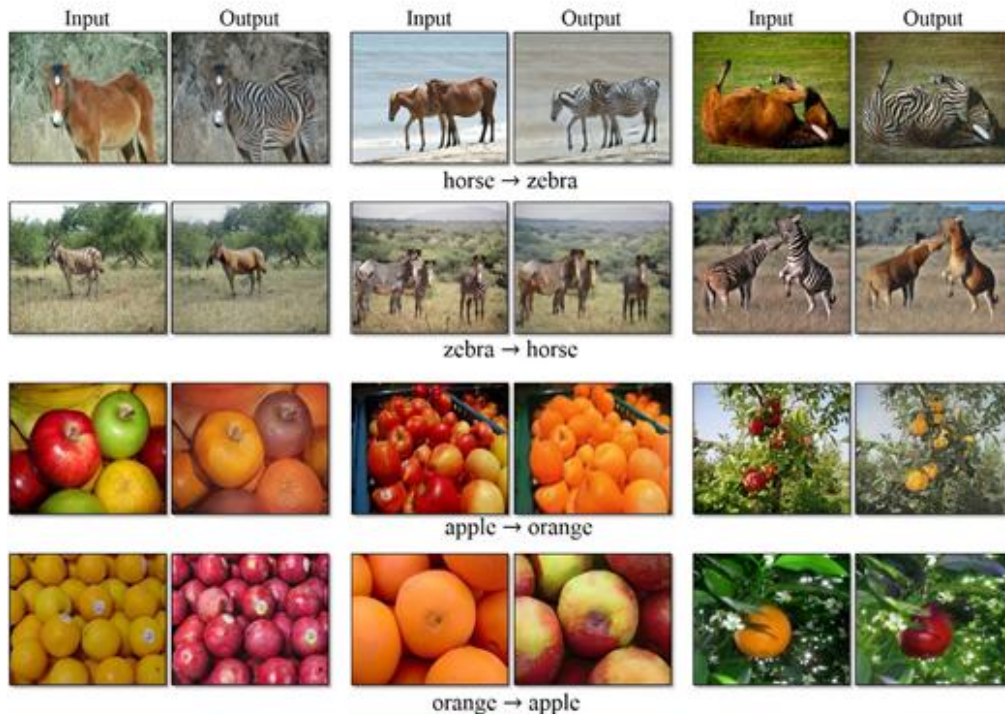


Система комбинированного видения (CVS)



# Применение нейронных сетей в задачах технического зрения

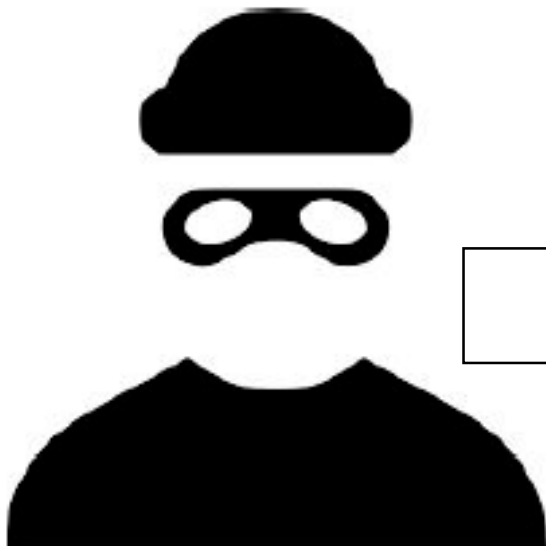
На сегодняшний день использование глубоких конволюционных нейронных сетей (ГКНС) стало важнейшим инструментом для обработки изображений. ГКНС показывают себя более эффективными относительно классических алгоритмов обработки изображений. Результаты, получаемый при помощи ГКНС близки к возможностям человека, а в ряде случаев даже превосходит.



# Генеративная состязательная сеть

Для автоматического совмещения изображений предлагается модель, основанная на **генеративно-состязательных сетях (GAN)**. Основными составляющими генеративно-состязательной сети являются две конкурирующие нейронные сети: генератор  $G$  и дискриминатор  $D$ .

Для наглядности рассмотрим принцип обучения GAN на примере фальшивомонетчика и кассира, проверяющего купюры.



5000

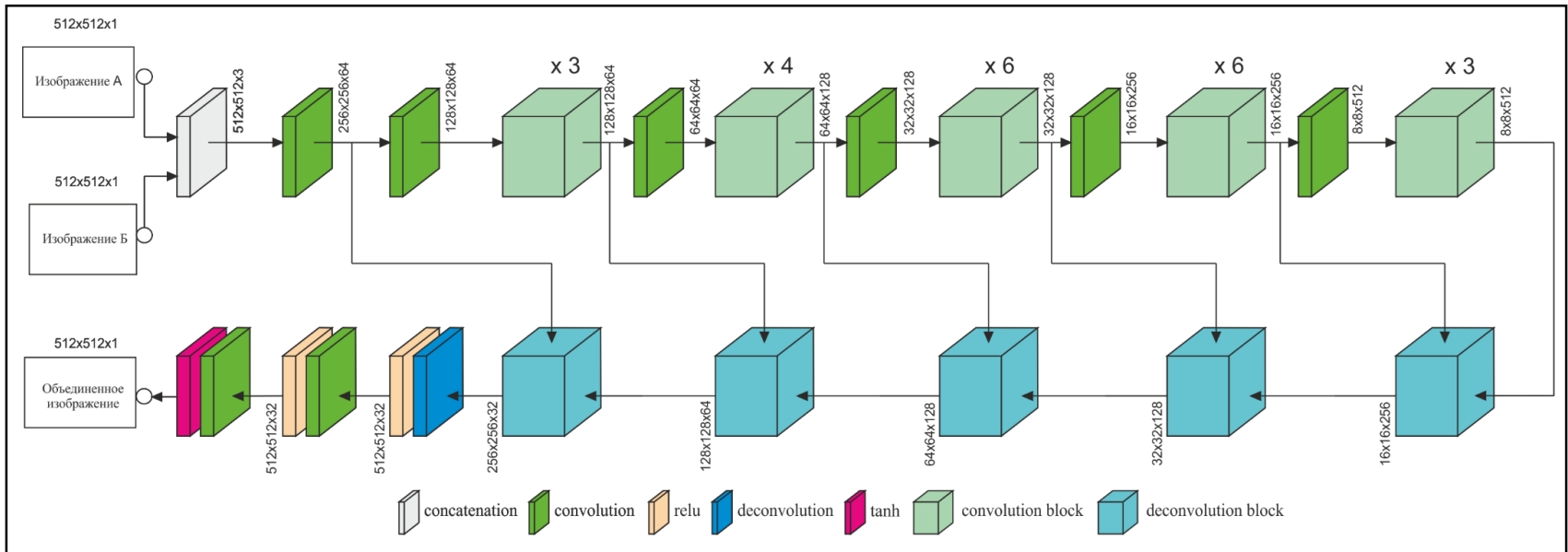


5000

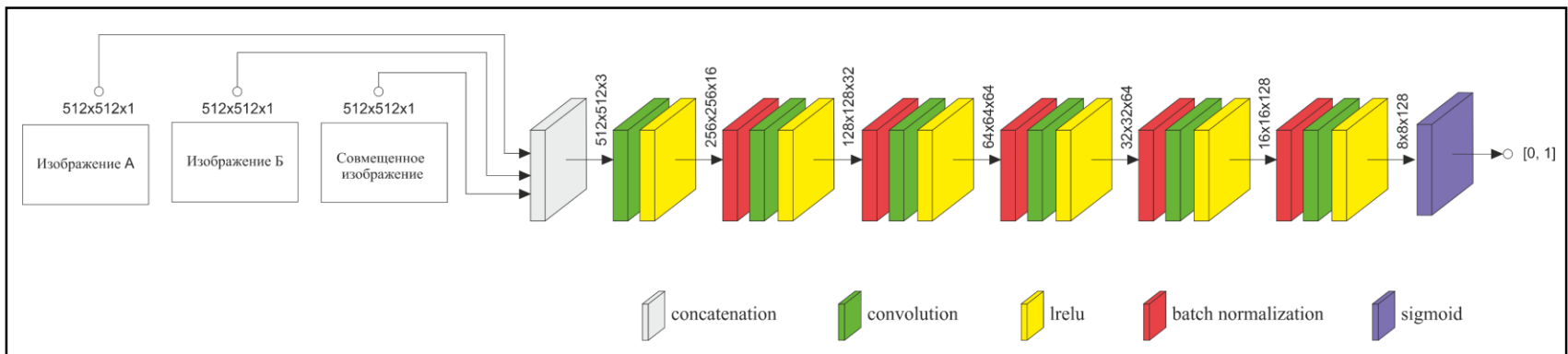


# Схема генератора и дискриминатора

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



Дискриминатор – это другая отдельная сеть, архитектура которой называется **«PatchGAN».**



# Схема конволюционного и деконволюционного блока

Схема конволюционного блока:

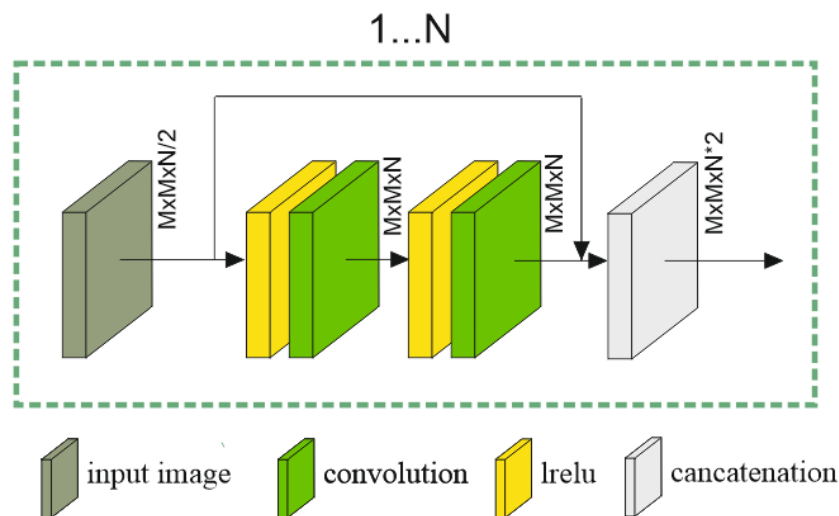
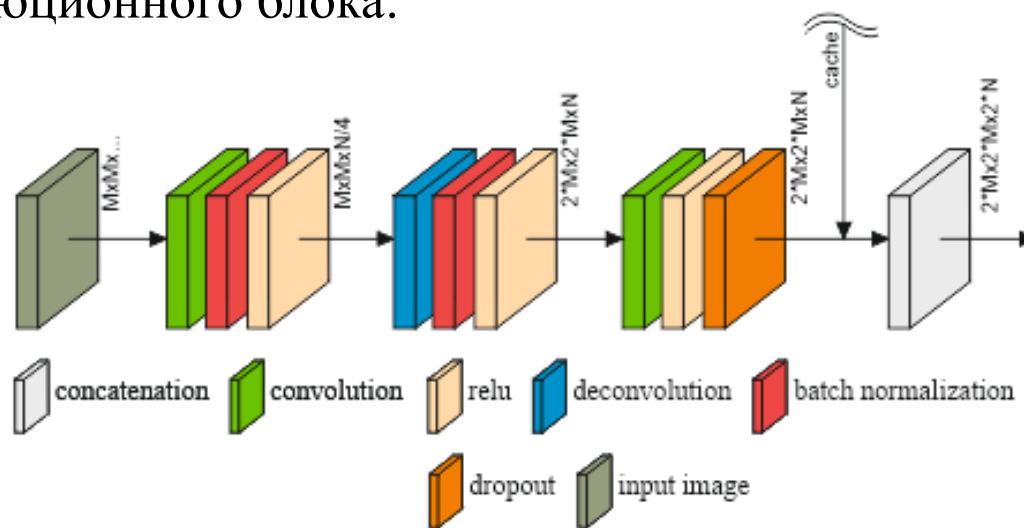


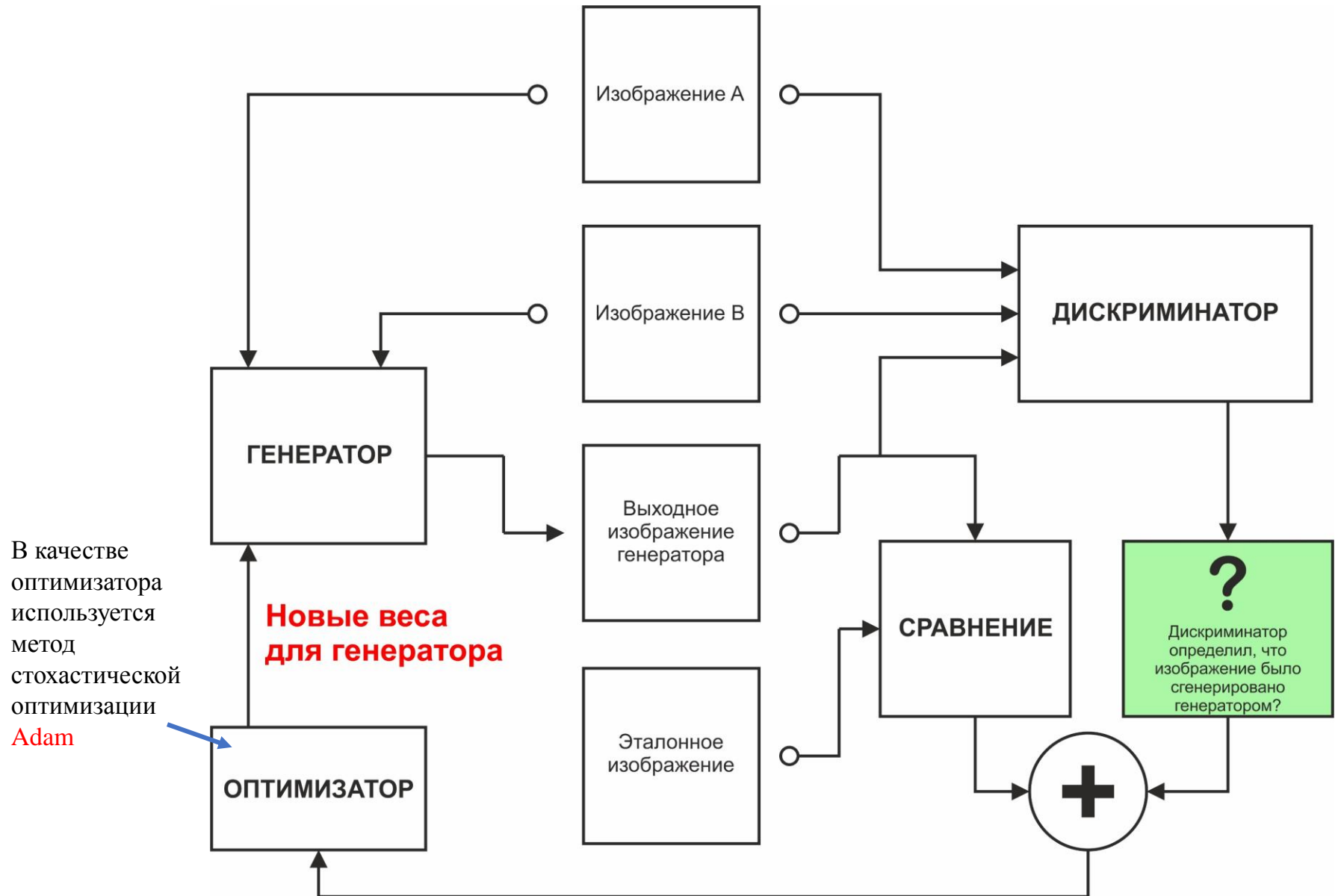
Схема деконволюционного блока:



# Блок-схема обучения дискриминатора



# Блок-схема обучения генератора





# Программная реализация

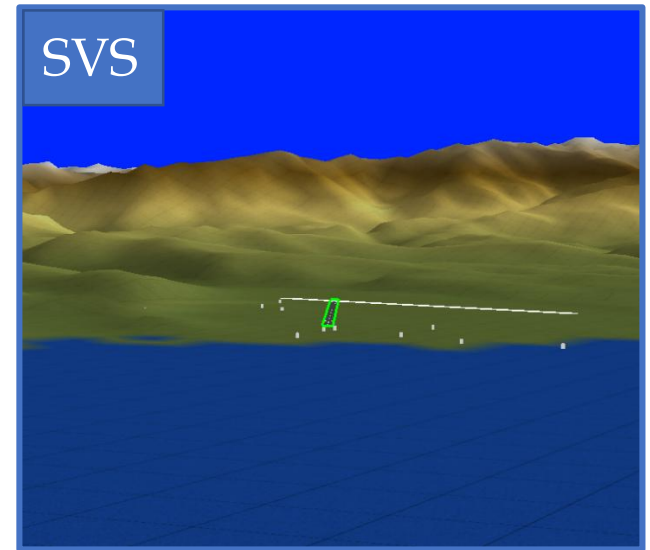
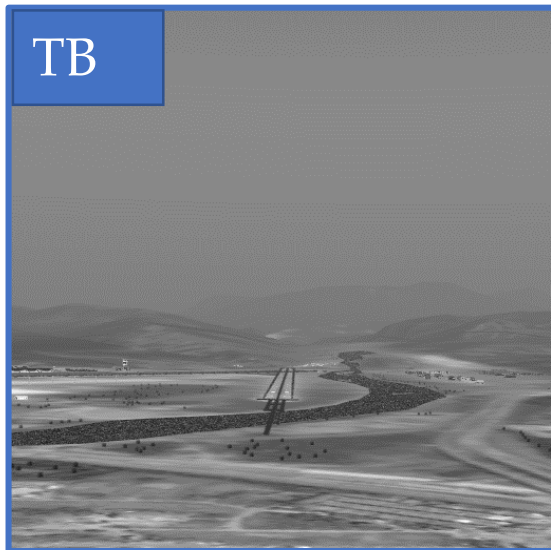
Предложенная архитектура сети была реализована при помощи открытой программной библиотеки для глубокого обучения **TensorFlow** на языке программирования **Python**.

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



# Исходные данные

Создание обучающих выборок осуществлялось с использованием стендовой базы ФГУП «ГосНИИАС». Выборка состояла из изображений телевизионного (ТВ) и инфракрасного (ИК) каналов разрешением 512x512 пикселей. Синтезированные изображения формировались на основе имеющейся базы данных подстилающей поверхности, препятствий, объектов аэродромной инфраструктуры и пилотажно-навигационной информации без погрешности и с добавлением погрешности ( $\pm 1.5$  градуса по осям ориентации).



# Формирование обучающей выборки для задачи комплексирования изображений

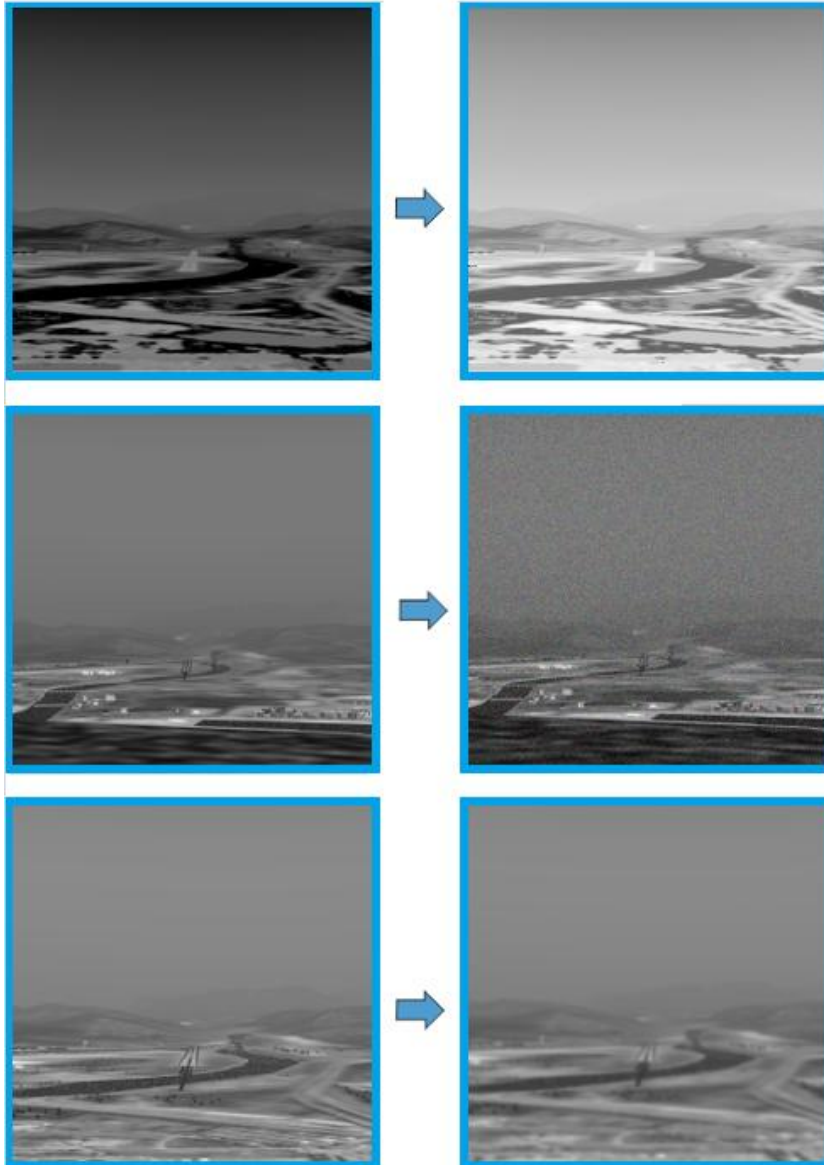
Для обучающей выборки были отобраны 9600 изображений, синтезированных на демонстраторе информационно управляющего поля (ИУП) кабины, на основе которых формируются эталонные комплексированные изображения, получаемые при помощи классического подхода (улучшение входных изображений алгоритмом MultiScale Retinex<sup>1</sup> с последующим комплексированием на основе пирамиды Лапласианов<sup>2</sup>).



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

<sup>2</sup> В. В. Инсаров, С. В. Тихонова, И. И. Михайлов. Проблемы построения систем технического зрения, использующих комплексирование информационных каналов различных спектральных диапазонов // М.: Новые технологии, приложение к журналу "Информационные технологии", № 3, 2014.

# Аугментация данных

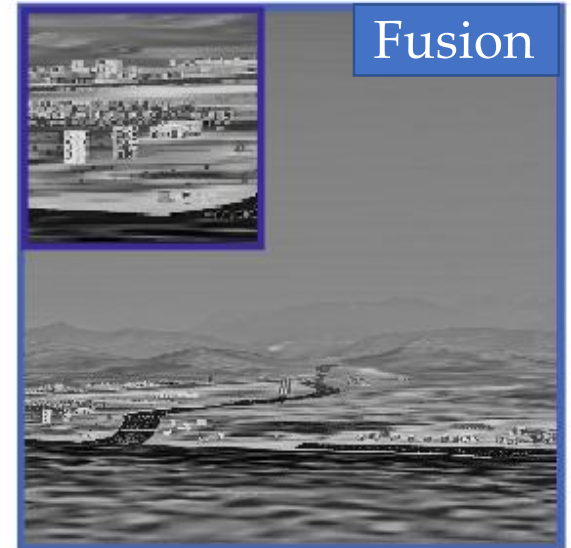
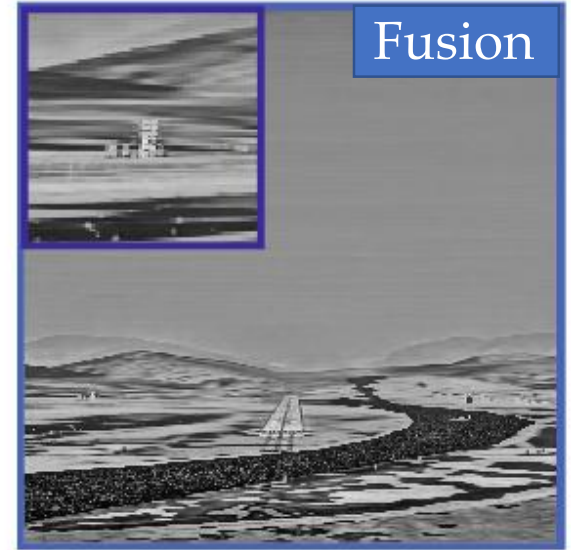


Имитация времени суток осуществлялась простым случайным **изменением параметров яркости** для изображений.

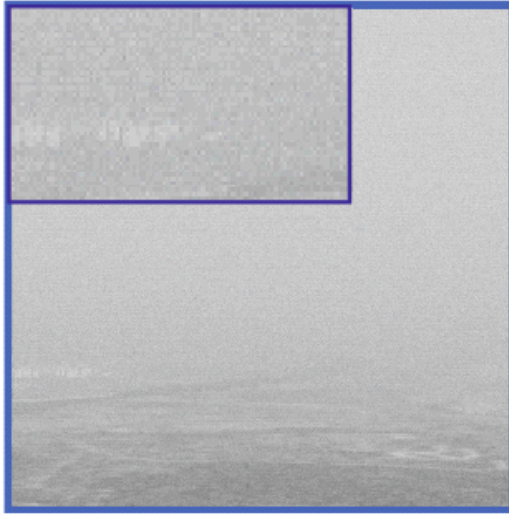
Для имитации характерных для реальных сенсоров помех **добавлялся случайный аддитивный гауссовский шум** с нулевым математическим ожиданием и максимальной дисперсией  $\sigma = 20$ , а также **геометрические несоответствия** до 5 пикселей.

Имитация изображений не в фокусе или тумана для ТВ- канала осуществлялась **добавлением случайного гауссовского размытия** с максимальной сигмой  $\sigma = 2$ .

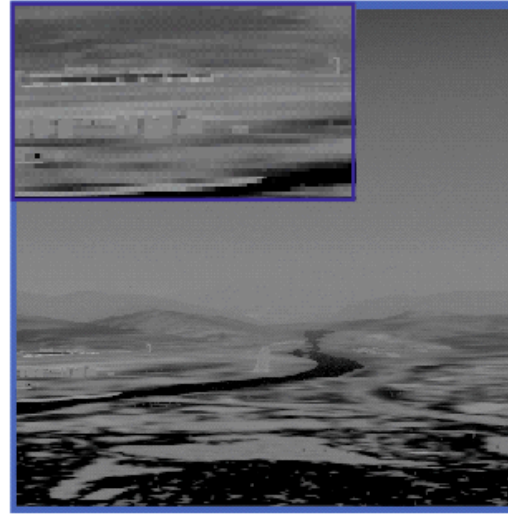
# Результаты работы сети



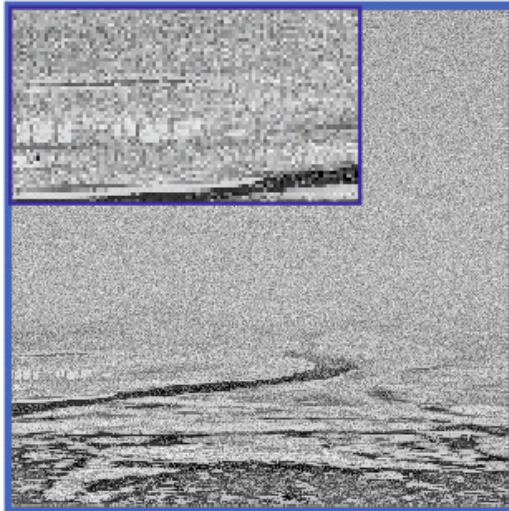
# Сравнение результата работы сети с классическим алгоритмом при плохих условиях видимости



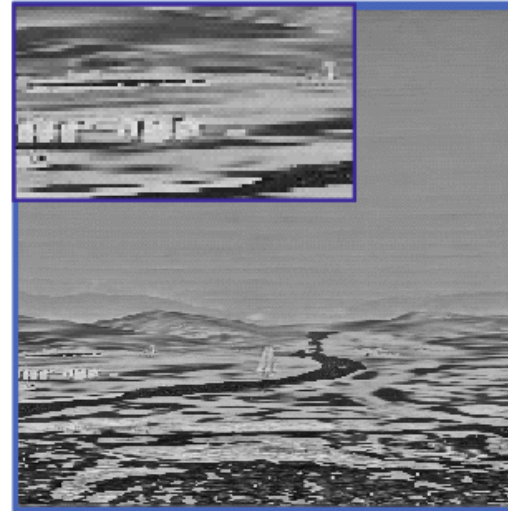
Исходное ТВ



Исходное ИК

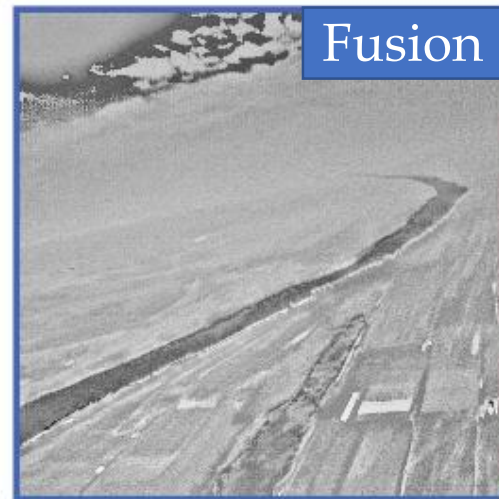
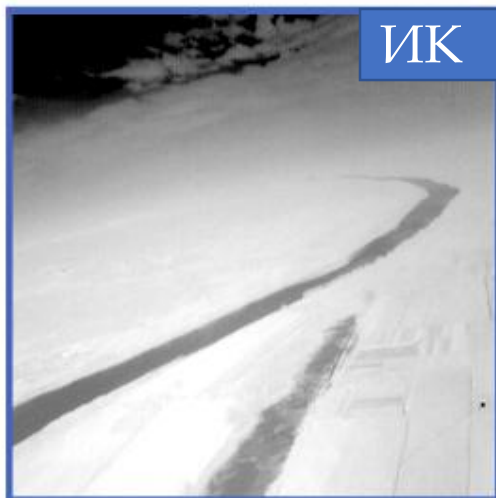
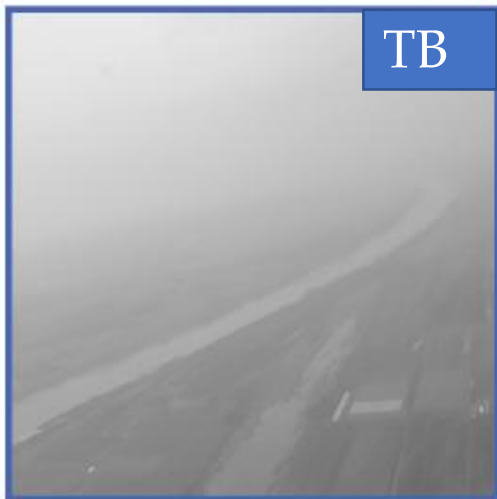


Классическое  
комплексирование



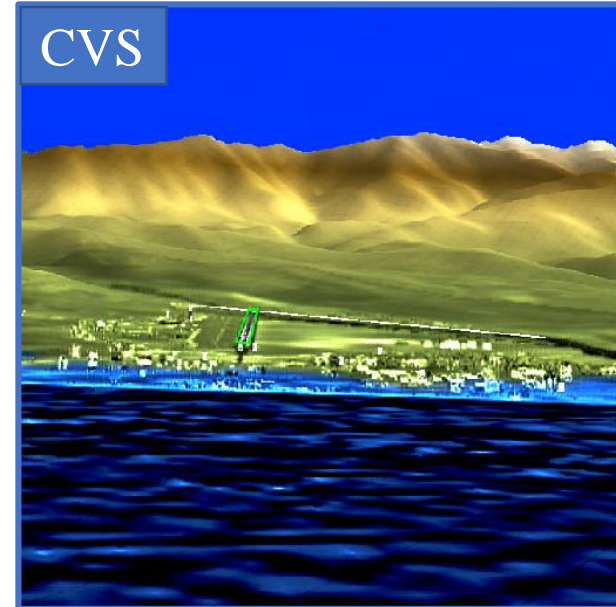
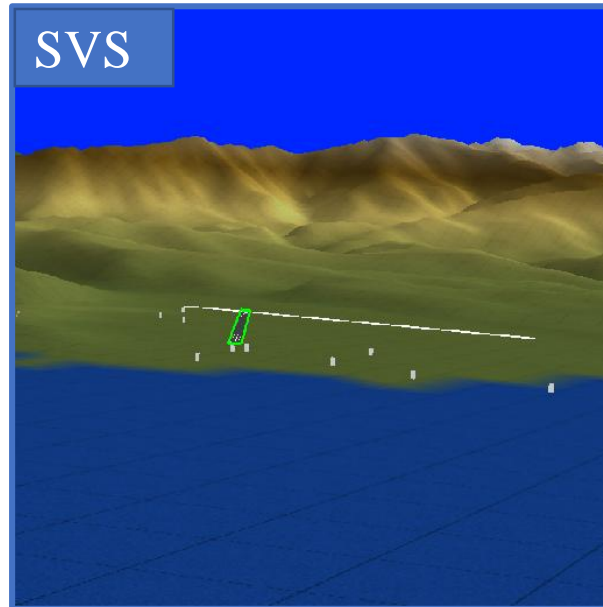
Нейросетевое  
комплексирование

# Результаты работы сети на реальных изображениях с бортовой камеры ЛА



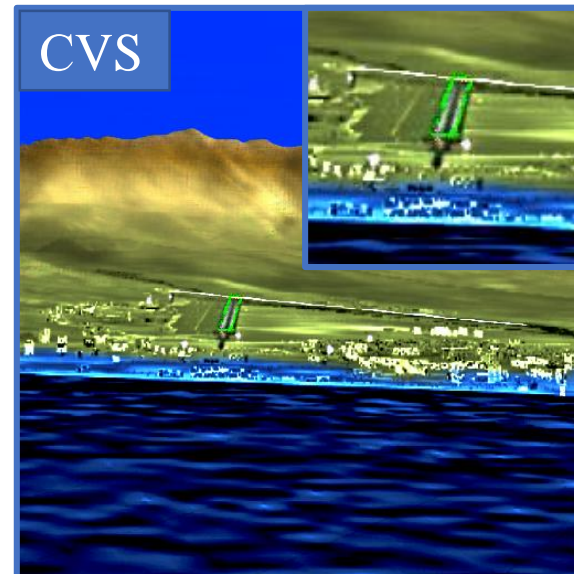
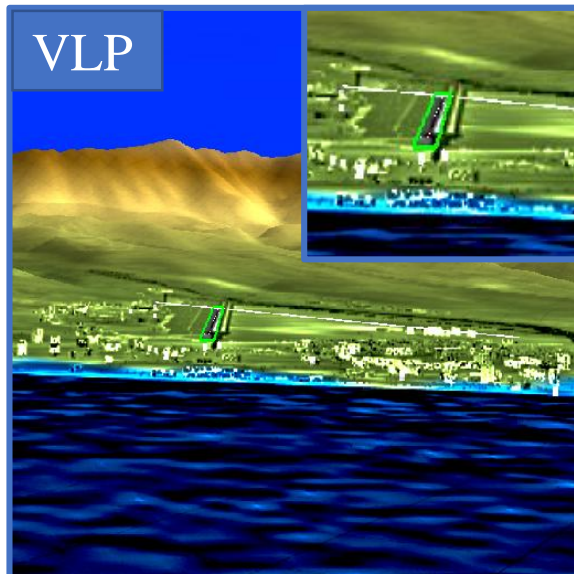
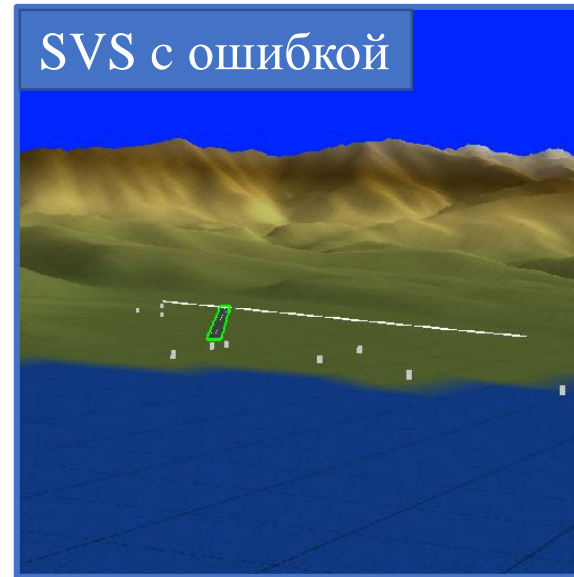
# Формирование обучающей выборки для задачи построения комбинированного изображения

Для задачи объединения изображений была сформирована выборка из комплексированных (EVS) и синтезированных (SVS) изображений. Эталонные комбинированные (CVS) изображения формировались путем объединения изображений EVS и SVS без ошибки при помощи алгоритма Vivid Light Photoshop (VLP).

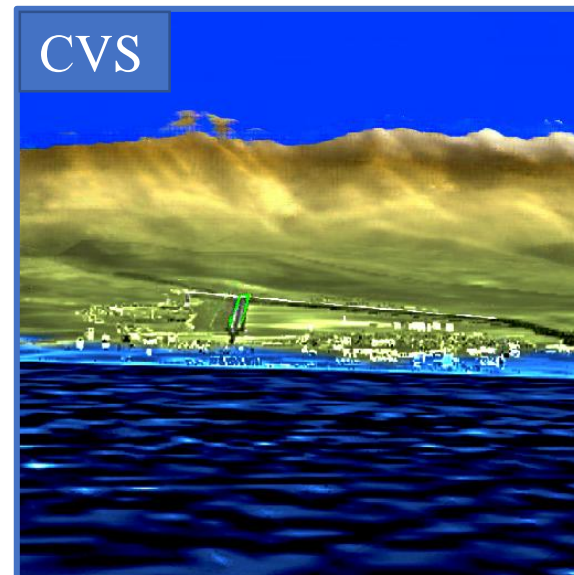
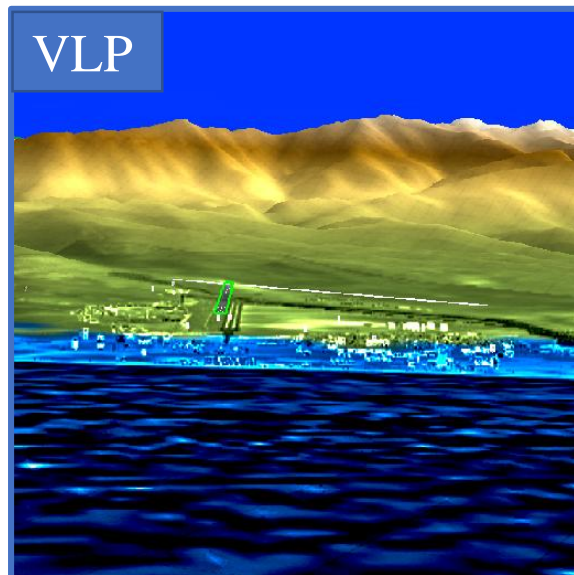
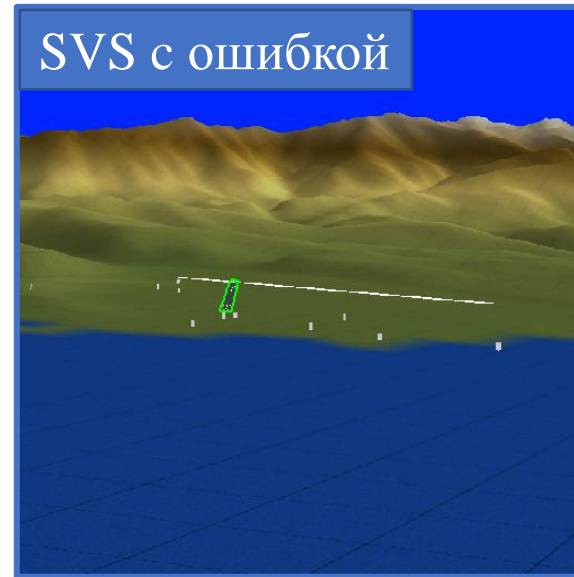




# Результаты работы сети при объединении изображений



# Результаты работы сети при объединении изображений



## Заключение

Предложена оригинальная архитектура генеративной нейронной сети для автоматического объединения изображений разной физической природы разрешением 512x512 пикселей с сохранением качества обработки мелких объектов. Продемонстрирована работоспособность применительно к таким практическим задачам как комплексирование разноспектральных изображений и формировании комбинированных изображений.

Комплексирование изображений телевизионного и инфракрасного диапазонов в совокупности с аугментацией данных при обучении характеризуется низким уровнем шума и сохранением информативных областей со всех входных изображений различных спектров.

Предложенная методика обучения сети для объединения улучшенного и синтезированного изображений позволяет на этапе работы алгоритма не осуществлять процедуру уточнения пилотажно-навигационных данных, а сразу осуществлять комбинирование изображений.

*Спасибо за внимание!*