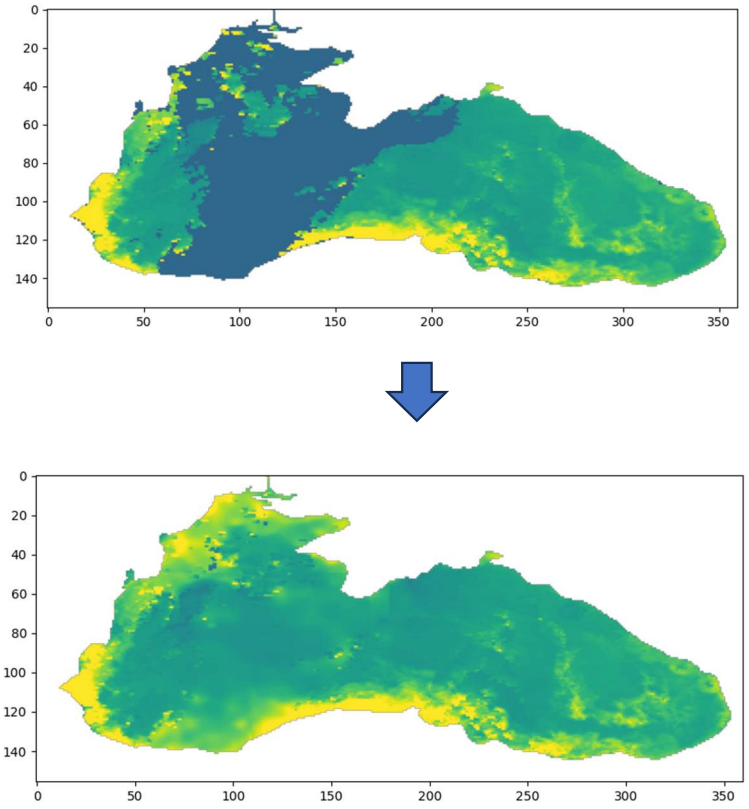
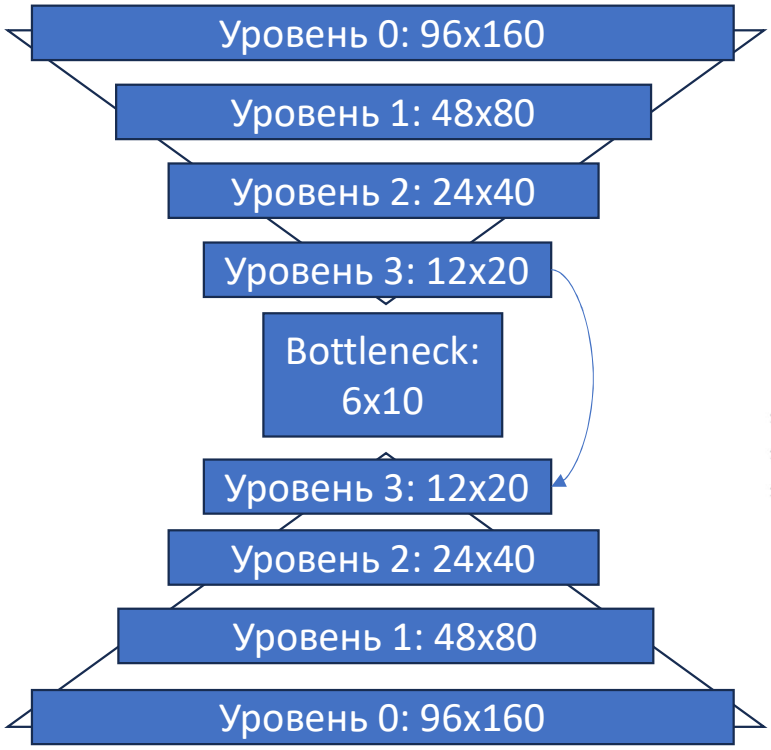
Выявление оптимального способа подготовки обучающих данных для U-Net при восстановлении суточных спутниковых карт концентрации хлорофилла «а» в Черном море

Куимов А.П.
(1) Институт океанологии им. П.П. Ширшова РАН, Москва, Россия
(2) Московский физико-технический институт (государственный университет), Москва, Росси

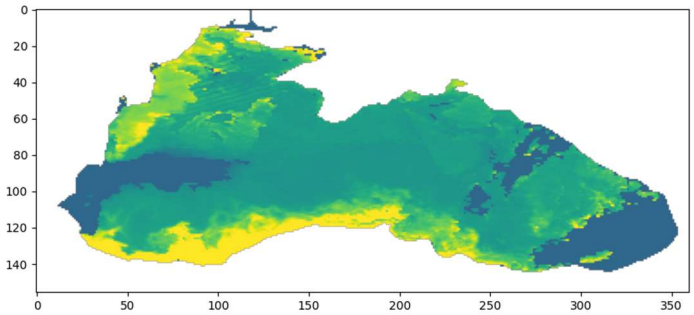


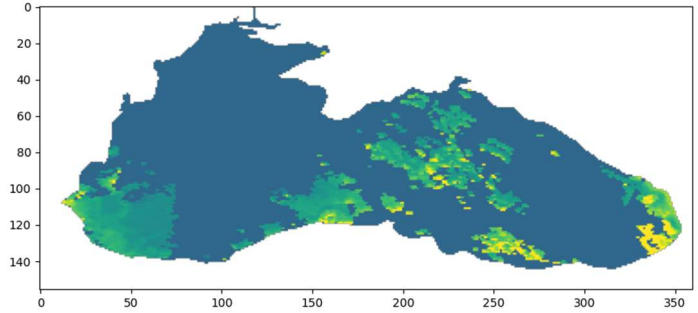
Суточные карты концентрации хлорофилла «а», получаемые со спутников, содержат пропуски, возникающие из-за облачности. В областях пропусков скрывается информация о концентрации хлорофилла «а». Одним из естественных решений является применение нейронных сетей архитектуры U-Net, предназначенных для генерации незашумлённых изображений. Однако возникает вопрос: как формировать обучающие выборки для обучения нейронной сети?

Архитектура и данные.



Спроектированы U-Net состоящие из четырёх свёрточных блоков с понижением пространственной размерности и четырёх последующих свёрточных блоков с повышением пространственной размерности. Для повышения разрешающей способности добавлены связи между нижними уровнями. Функциями активации выступали Mish, функцией ошибок L2.

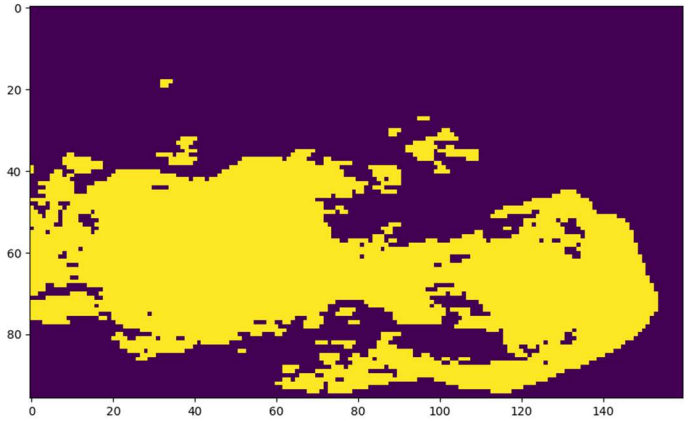




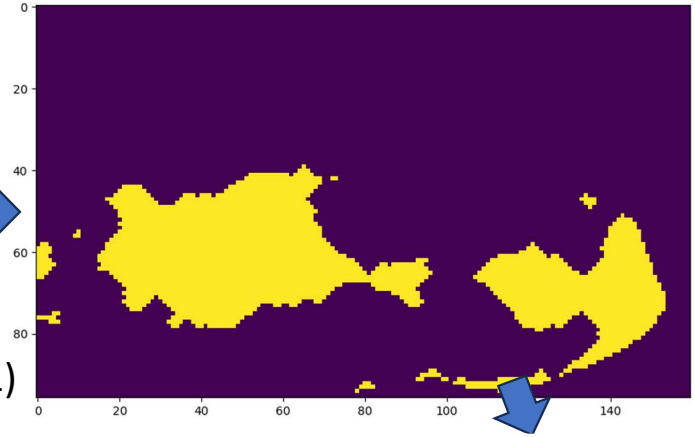
Суточные карты 3 уровня Aqua-MODIS, Terra-MODIS, NOAA20-VIIRS, SNPP-VIIRS, Sentinel-3A и Sentinel-3B скачивались с сайта Nasa. Данные карты усреднялись, и бралась область чёрного моря ((40.5, 47.0) (27.0, 42.0)). Данные логарифмировались для сглаживания В результате получались изображения выше.
После они нарезались на более малые изображения 96x160, которые подавались на вход. Обучение проводилось на выбранных участках.

Выбор обучающей области

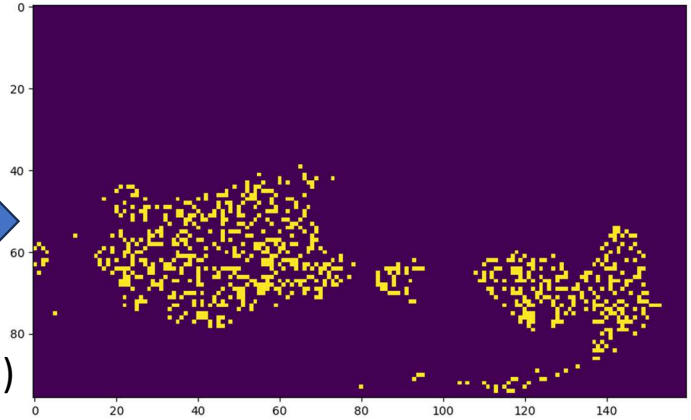
На изображение 1 представлена маска, обозначающая наличие данных (жёлтый) или пропуск(тёмный). Необходимо выбрать участки, на которых нейросеть будет настраивать свои веса. Будем выбирать только те точки, которые в радиусе имеют известных соседей (изображение 2)



(1)



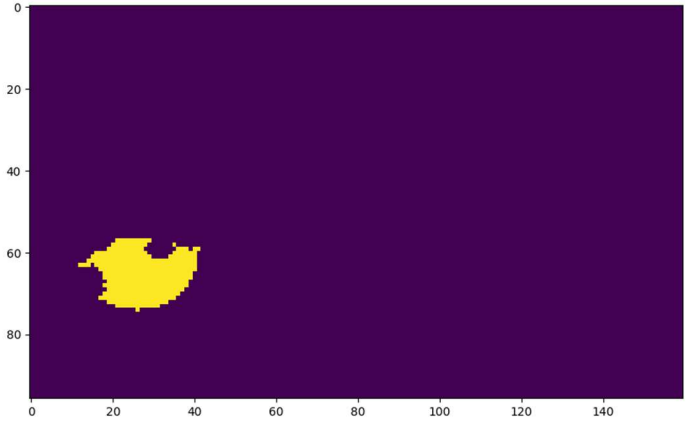
(2)



(3)

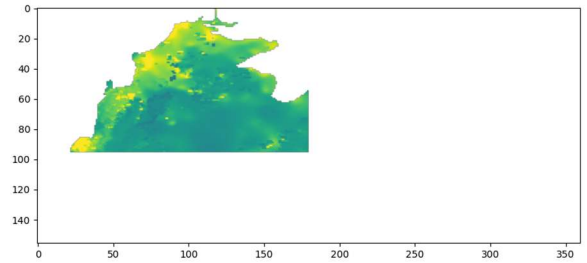
Можно выбрать случайные точки(изображение 3) или целую область(изображение 4).

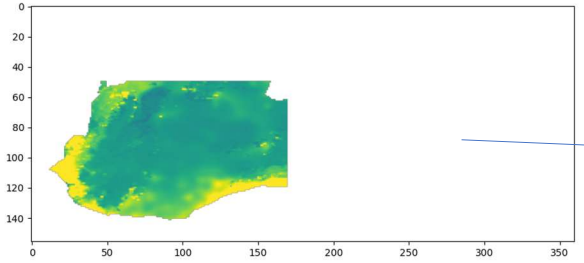
Как показали эксперименты, нейросеть, обученная на целой область даёт большую выразительную способность, чем случайные точки. Точки в обучающей области лучше выбрать нулевыми.

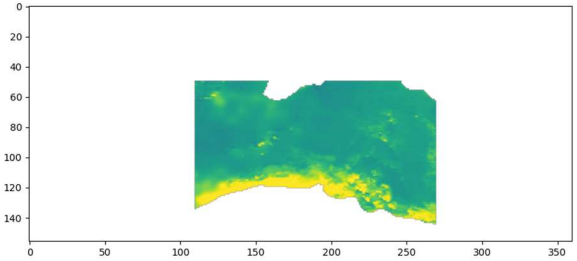


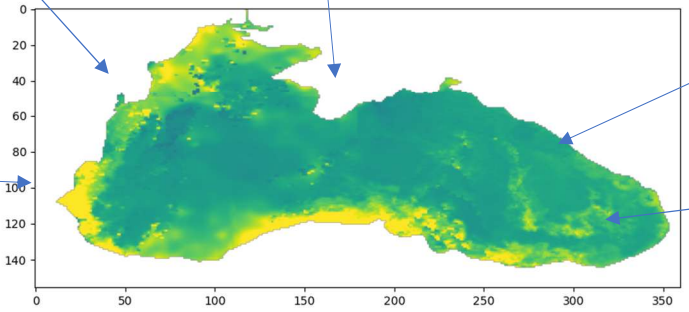
(4)

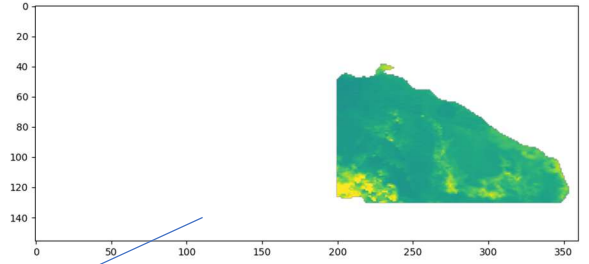
Сшивание картинок

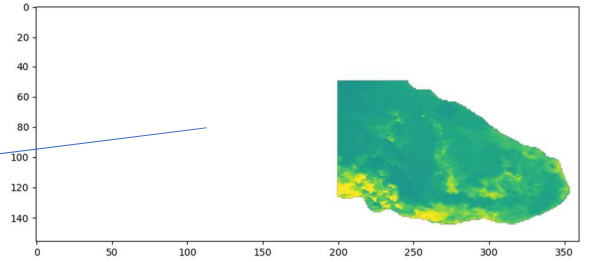




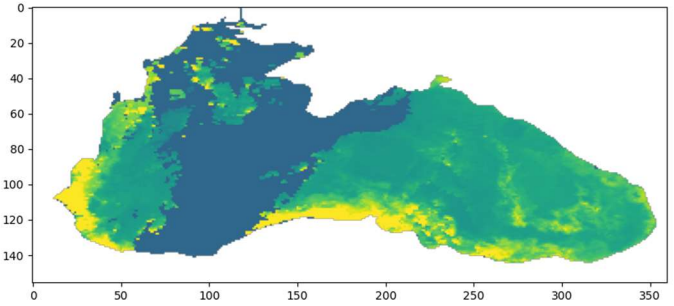








Показано действие нейросети, обученной на односвязных областях. Результаты были вставлены только в местах пропусков.



Изначальное изображение

Данная модель сильно смазывает структуру, но даёт понимание «Есть возрастание хлорофилла или нет в данной области»