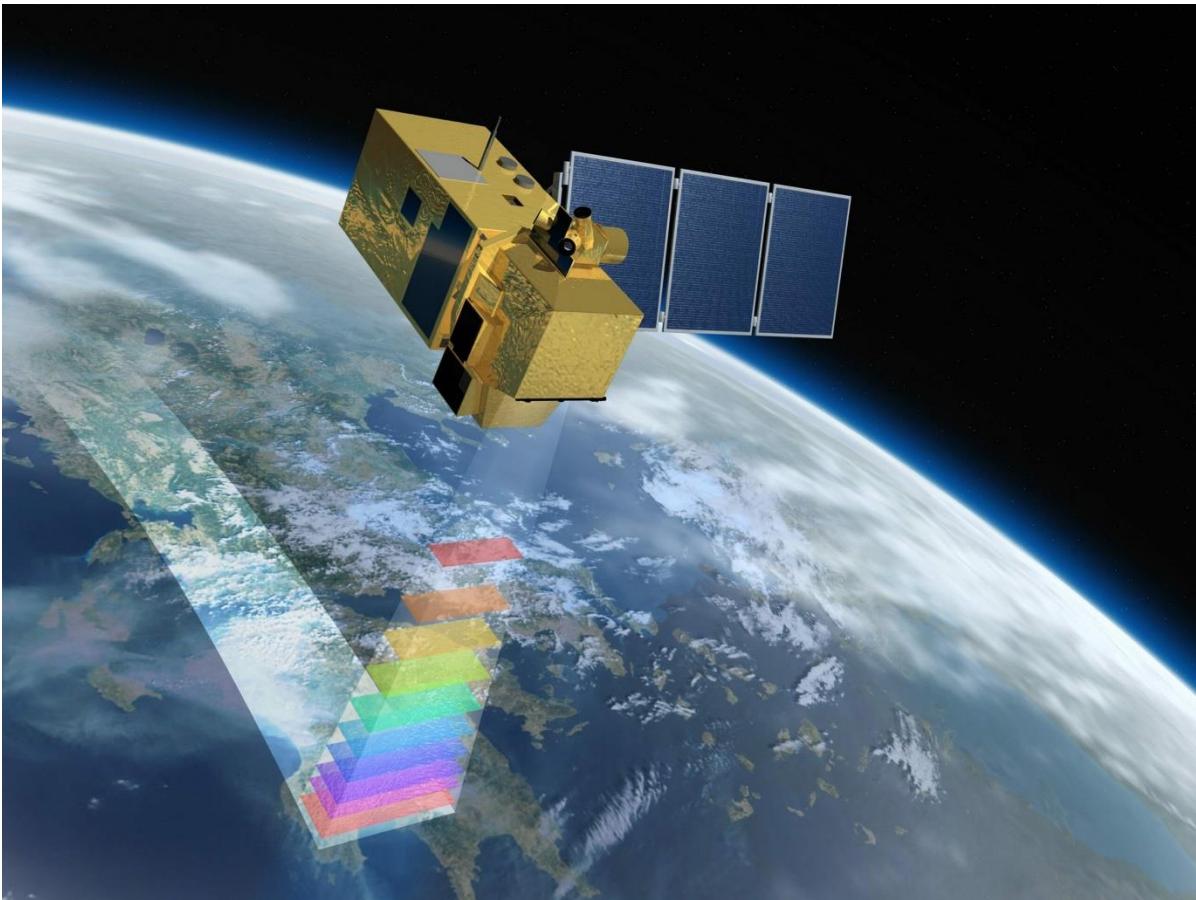


**Метод автоматического переноса тематических продуктов типа  
Scene Classification Land на данные прибора КМСС-2  
(Метеор-М) на основе текстурных признаков, перцентильной  
нормализации и методов машинного обучения**

*Колбудаев П.А., Плотников Д.Е.*

# Sentinel-2: общие характеристики



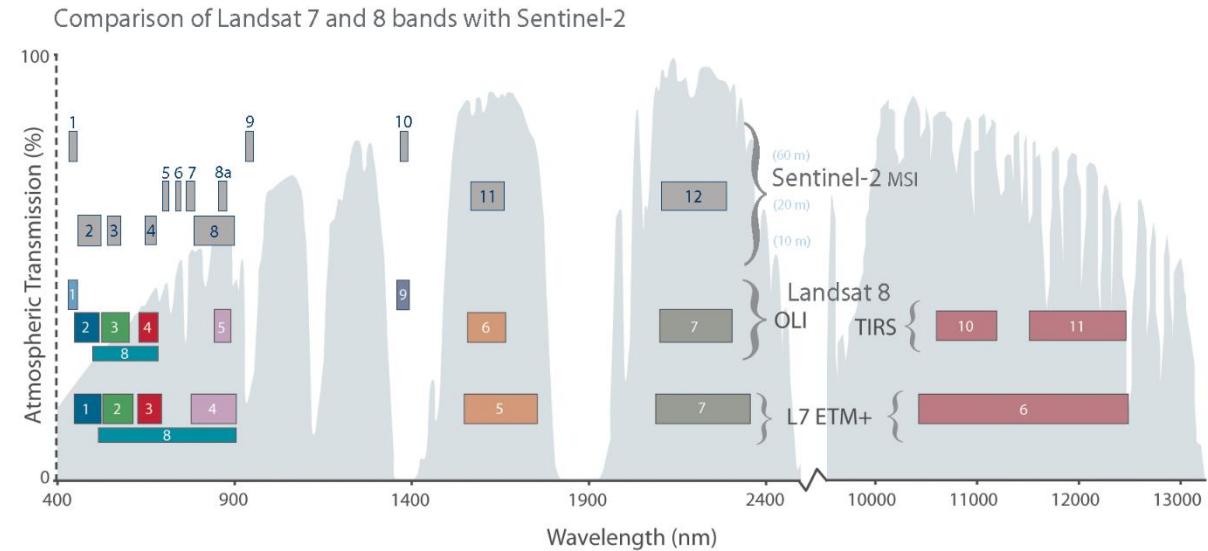
Sentinel-2	
	Модель спутника Sentinel-2
Технические характеристики	
Производитель	Airbus Defence and Space
Оператор	EKA
Задачи	Д33
Спутник	Земли
Стартовая площадка	Куру, ELV
Ракета-носитель	2А, 2В: Vega
Запуск	2А: 23 июня 2015 2В: 7 марта 2017
Длительность полёта	2А: 10 лет 2 месяца 8 дней 2В: 8 лет 5 месяцев 24 дня

Элементы орбиты	
Тип орбиты	солнечно-синхронная орбита
Наклонение	98,5°
Период обращения	100,5 мин.
Интервал повторения	5 дней (для 2 спутников)
Высота орбиты	786 км
Целевая аппаратура	
MSI	мультиспектральная камера
Пространственное разрешение	10, 20, 60 м
Полоса захвата	290 км
Спектральная полоса	443—2190 нм
Скорость передачи	2 Гбит/с
Бортовая память	2,4 ТБ

# Основные спектральные характеристики аппаратуры MSI

Band	Description	Central Wavelength (nm)	Bandwidth (nm)	Spatial Resolution (m)
Band 1	Coastal aerosol	443	21	60
Band 2	Blue	492	65	10
Band 3	Green	560	35	10
Band 4	Red	665	30	10
Band 5	Red edge1	704	15	20
Band 6	Red edge2	739	15	20
Band 7	Red edge3	783	20	20
Band 8	NIR	833	115	10
Band 9	Red edge4 (8a)	865	21	20
Band 10	Water Vapor	945	20	60
Band 11	Cirrus	1374	31	60
Band 12	SWIR1	1614	91	20
	SWIR2	2202	175	20

Спектральные каналы MSI и их характеристики



Сравнение каналов MSI с ETM+, OLI, TIRS

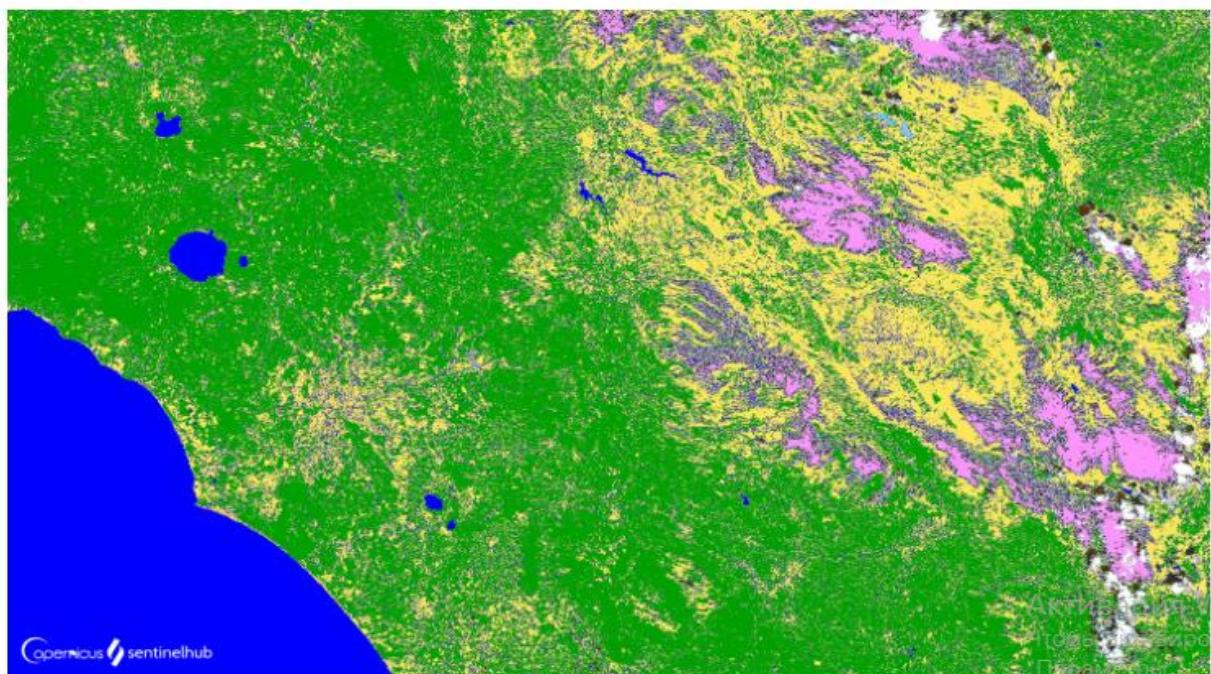
# Sentinel-2 L2A Scene Classification Land (SCL)

## General description

Scene classification was developed to distinguish between cloudy pixels, clear pixels and water pixels of Sentinel-2 data and is a result of the Scene classification algorithm run by ESA. Twelve different classifications are provided including classes of clouds, vegetation, soils/desert, water and snow. It does not constitute a land cover classification map in a strict sense.

## Description of representative images

Scene Classification of Rome.



## Color legend

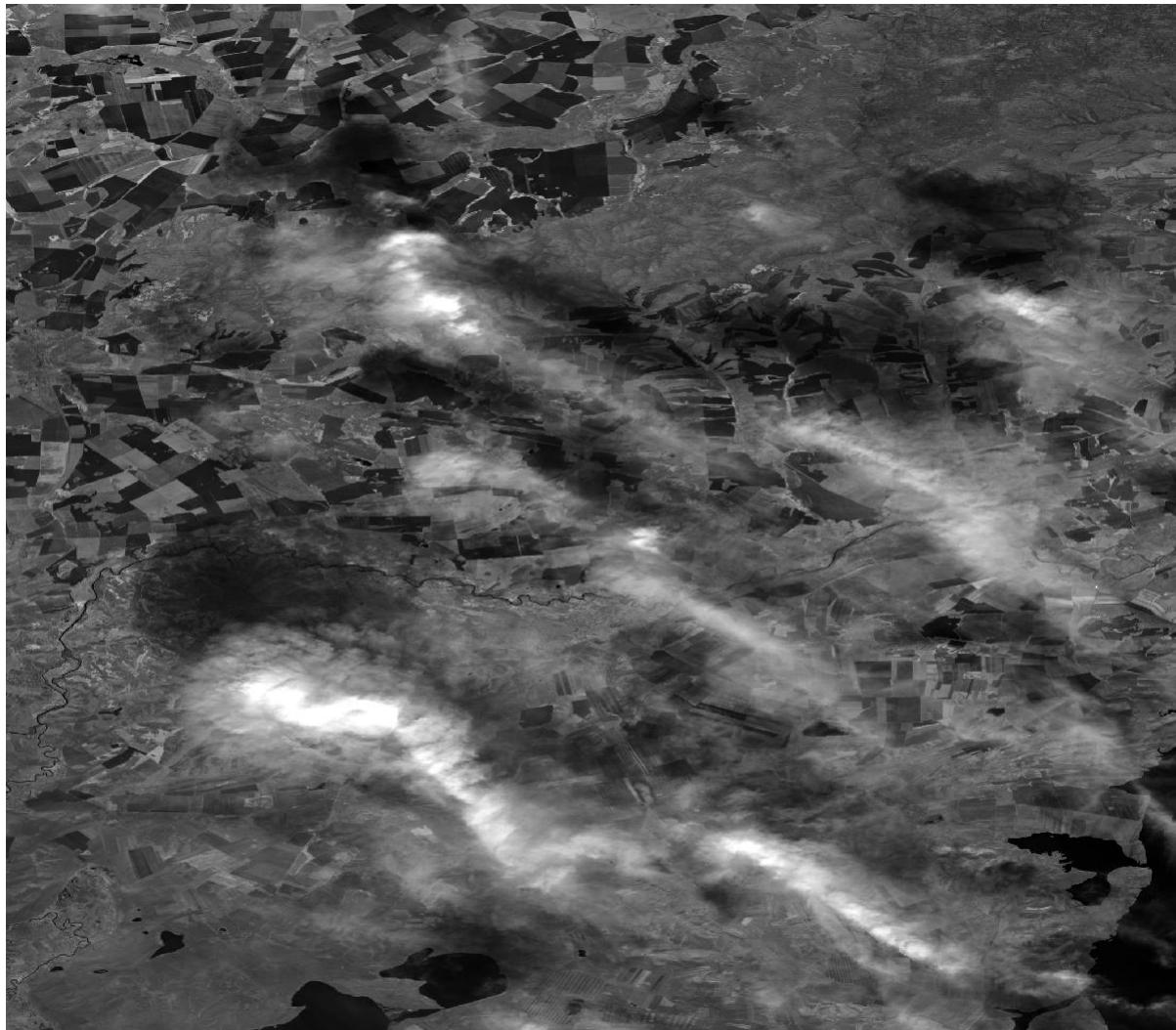
Value	Scene Classification	HTML color code	Color
0	No Data (Missing data)	#000000	
1	Saturated or defective pixel	#ff0000	
2	Topographic casted shadows (called "Dark features/Shadows" for data before 2022-01-25)	#2f2f2f	
3	Cloud shadows	#643200	
4	Vegetation	#00a000	
5	Not-vegetated	#ffe65a	
6	Water	#0000ff	
7	Unclassified	#808080	
8	Cloud medium probability	#c0c0c0	
9	Cloud high probability	#ffffff	
10	Thin cirrus	#64c8ff	
11	Snow or ice	#ff96ff	

# Загрубление Sentinel-2 L2A Scene Classification Land

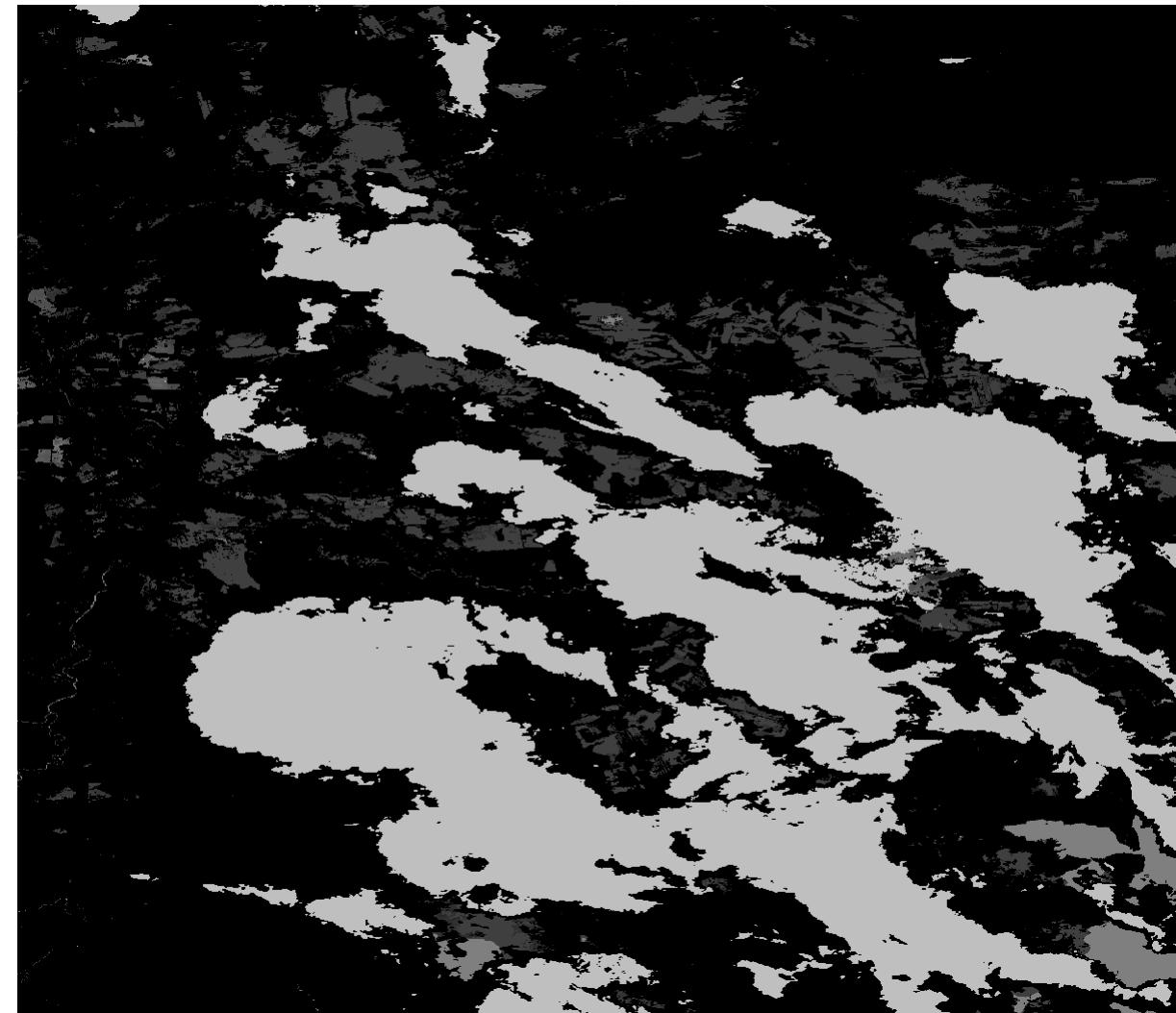
## Color legend

Value	Scene Classification	HTML color code	Color	
0	No Data (Missing data)	#000000		<b>нет данных</b>
1	Saturated or defective pixel	#ff0000		
2	Topographic casted shadows (called "Dark features/Shadows" for data before 2022-01-25)	#2f2f2f		<b>тени</b>
3	Cloud shadows	#643200		
4	Vegetation	#00a000		<b>чистые данные</b>
5	Not-vegetated	#ffe65a		
6	Water	#0000ff		<b>вода</b>
7	Unclassified	#808080		<b>нет данных</b>
8	Cloud medium probability	#c0c0c0		
9	Cloud high probability	#ffffff		<b>облачность</b>
10	Thin cirrus	#64c8ff		
11	Snow or ice	#ff96ff		<b>снег</b>

## Пример сцены Sentinel-2

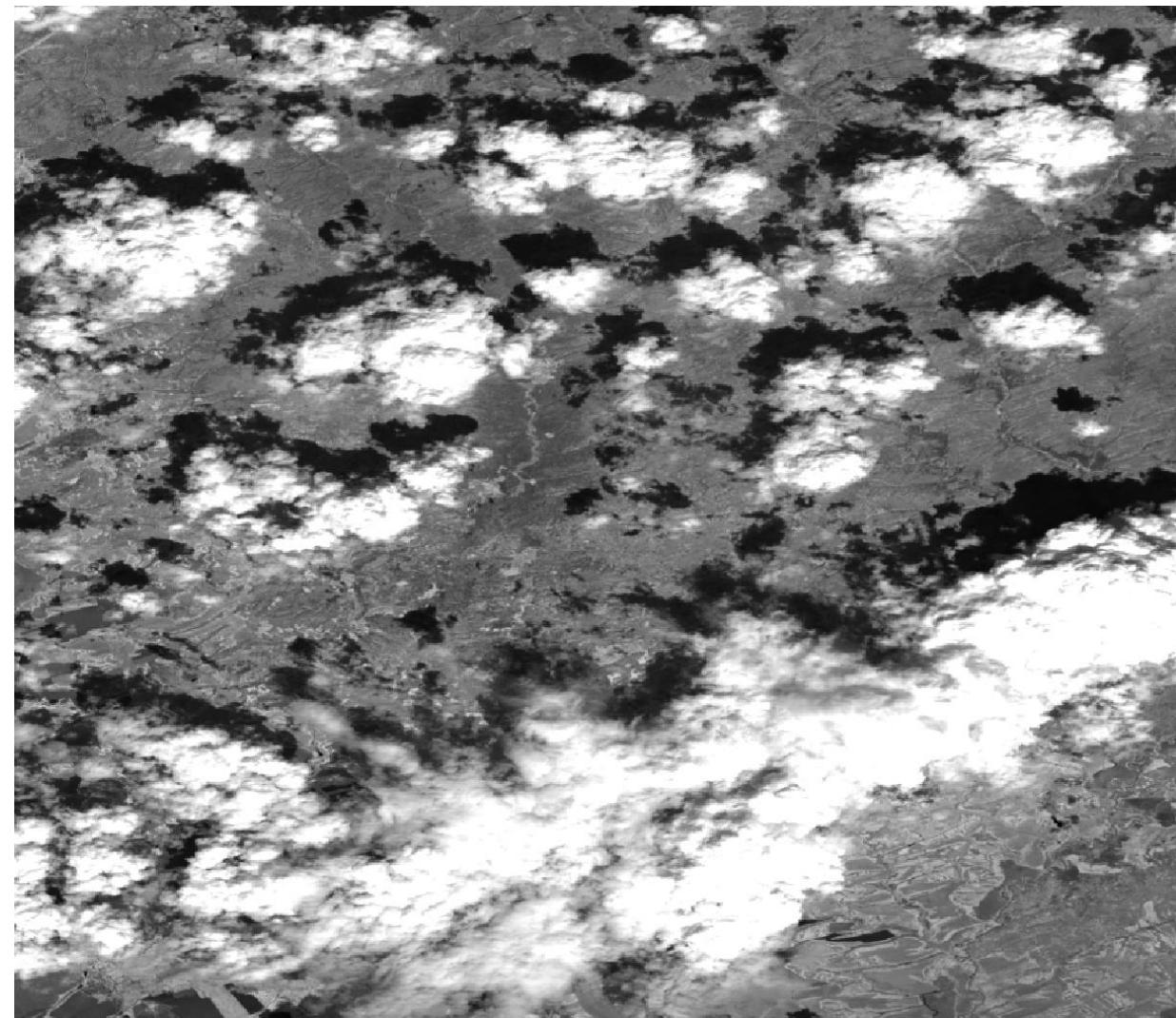


часть сцены msi в ИК канале

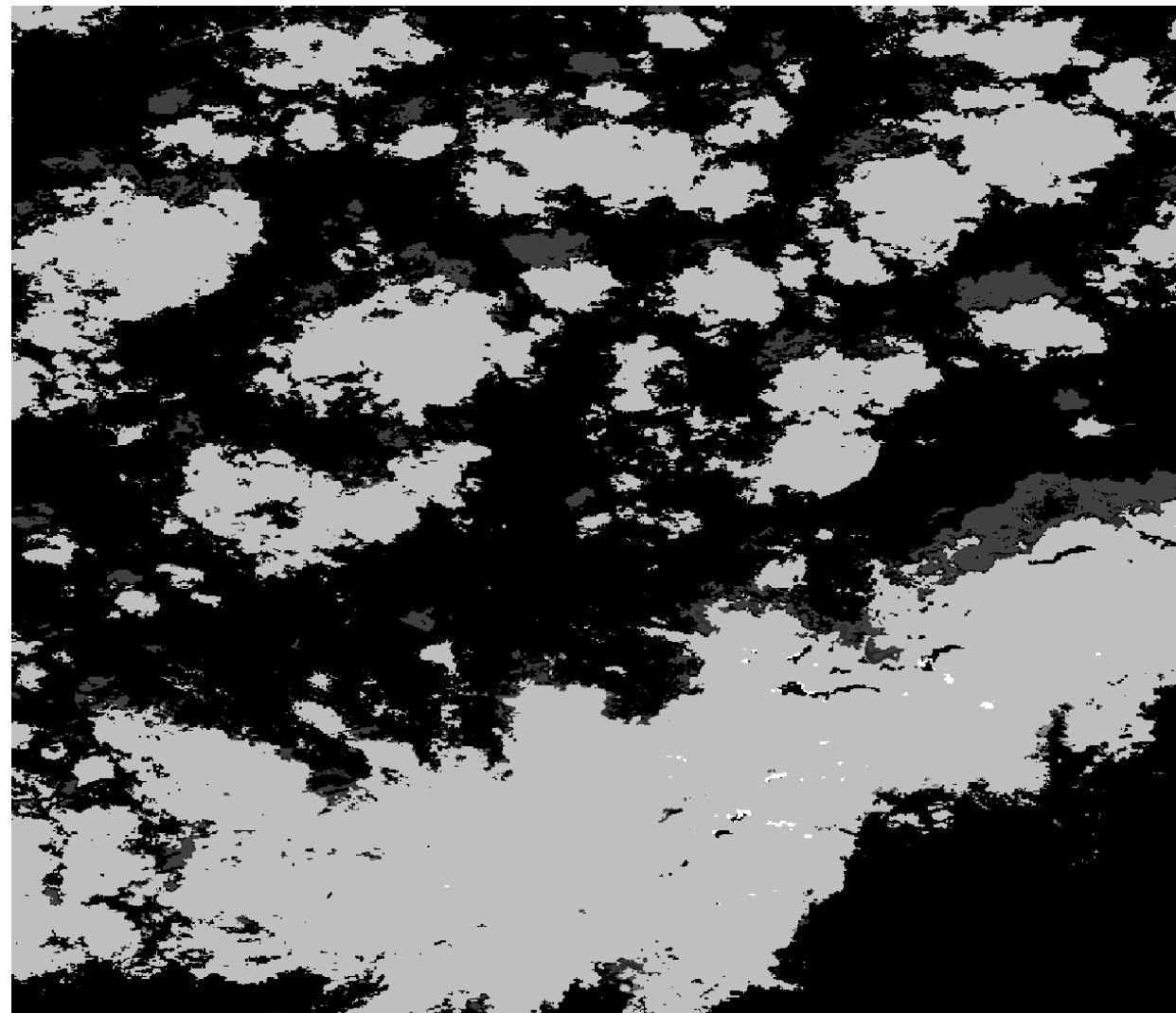


часть загруженной маски scl

## Пример сцены Sentinel-2

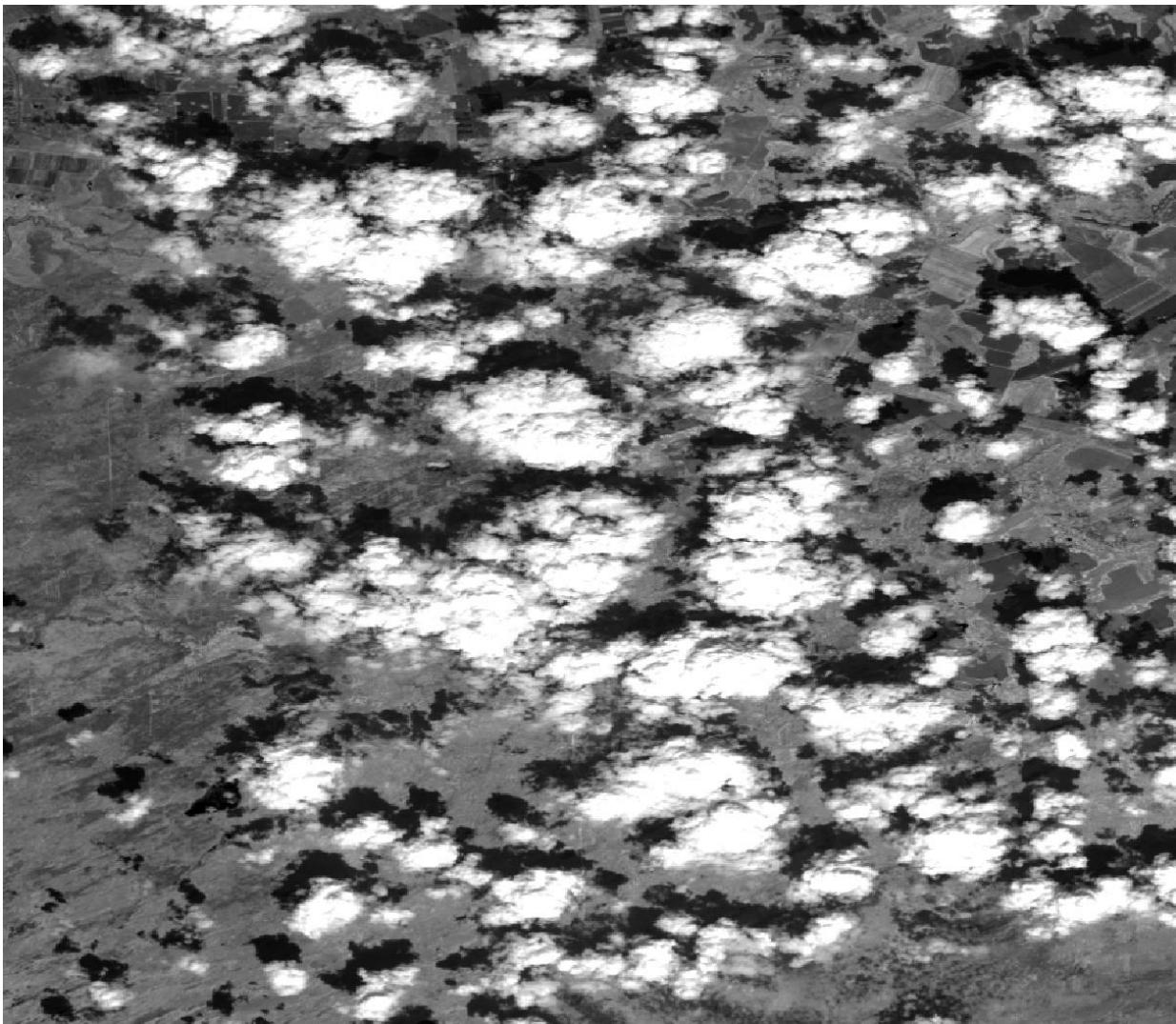


часть сцены msi в ИК канале

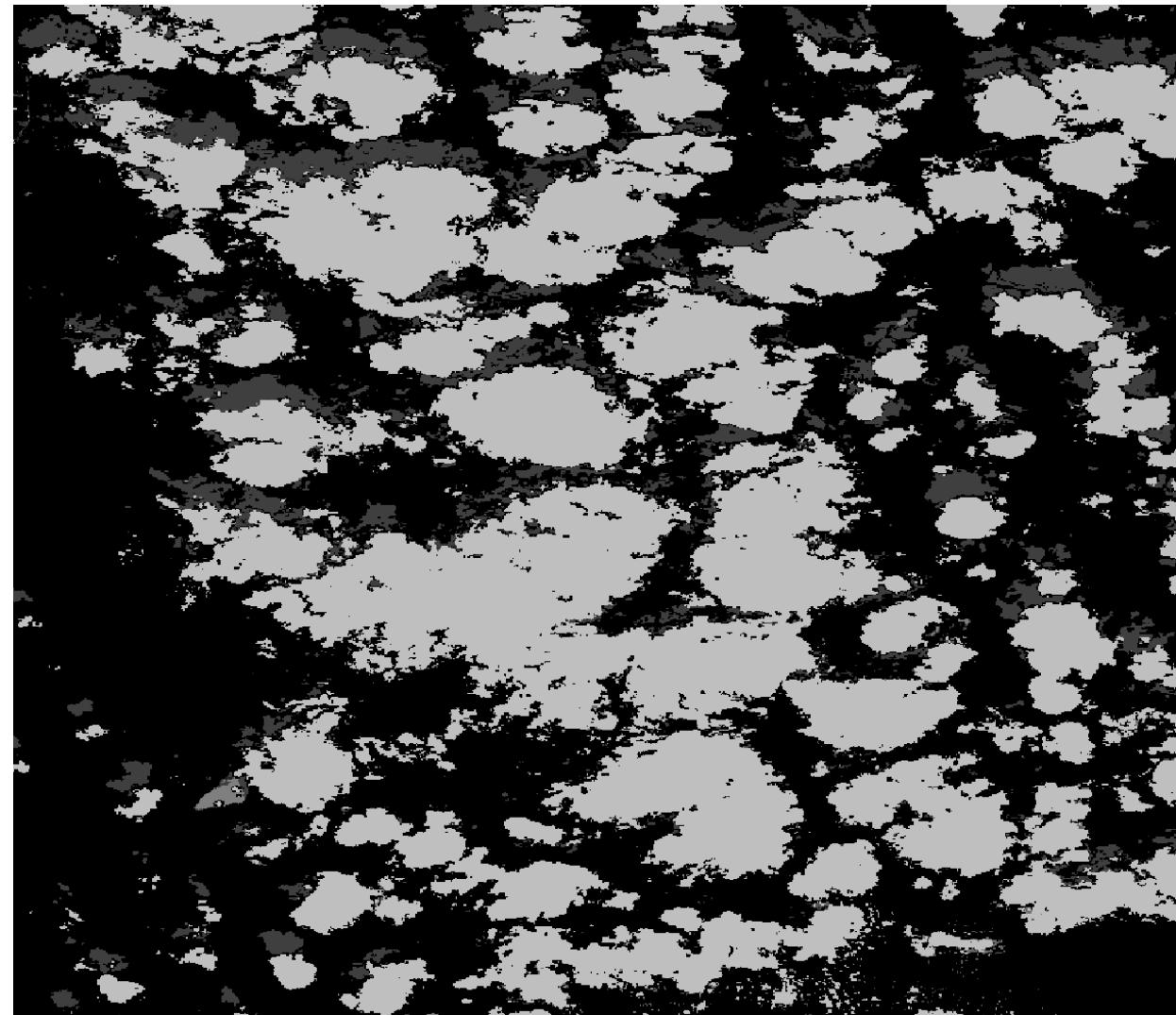


часть загруженной маски scl

## Пример сцены Sentinel-2

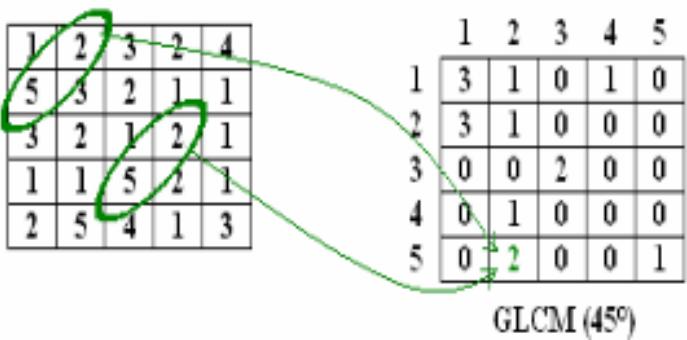
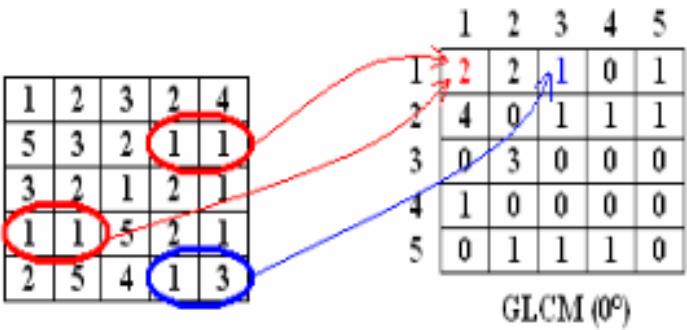


часть сцены msi в ИК канале



часть загруженной маски scl

# Gray-Level Co-Occurrence Matrix (GLCM)



P

Input array. `P` is the gray-level co-occurrence histogram for which to compute the specified property. The value `P[i,j,d,theta]` is the number of times that gray-level j occurs at a distance d and at an angle theta from gray-level i.

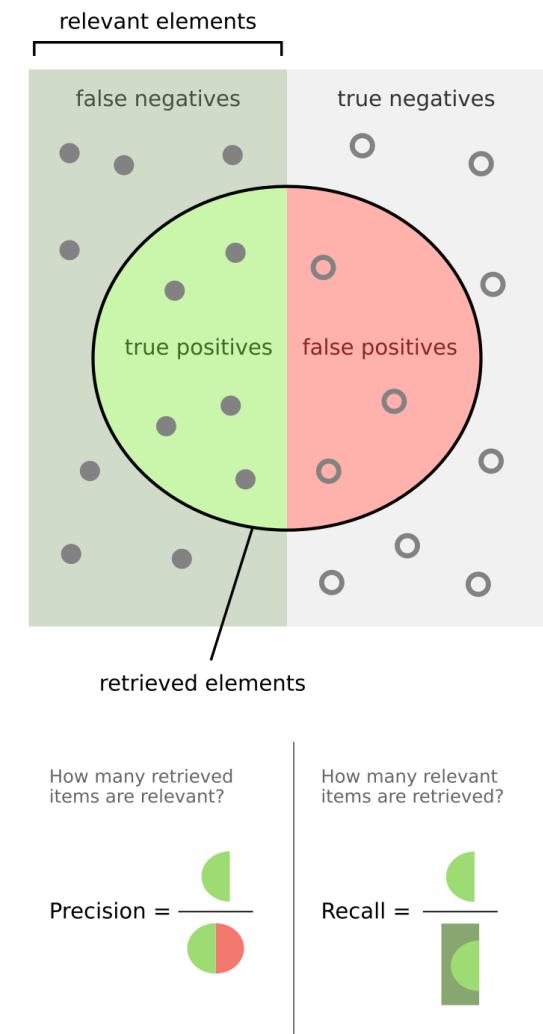
- 'contrast':  $\sum_{i,j=0}^{levels-1} P_{i,j}(i - j)^2$
- 'dissimilarity':  $\sum_{i,j=0}^{levels-1} P_{i,j}|i - j|$
- 'homogeneity':  $\sum_{i,j=0}^{levels-1} \frac{P_{i,j}}{1+(i-j)^2}$
- 'ASM':  $\sum_{i,j=0}^{levels-1} P_{i,j}^2$
- 'energy':  $\sqrt{ASM}$
- '**correlation**':

$$\sum_{i,j=0}^{levels-1} P_{i,j} \left[ \frac{(i - \mu_i)(j - \mu_j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}} \right]$$

- 'mean':  $\sum_{i=0}^{levels-1} i * P_i$
- 'variance':  $\sum_{i=0}^{levels-1} P_i * (i - mean)^2$
- 'std':  $\sqrt{variance}$
- 'entropy':  $\sum_{i,j=0}^{levels-1} -P_{i,j} * \log(P_{i,j})$

# Подбор параметров patch\_size, distance в GLCM

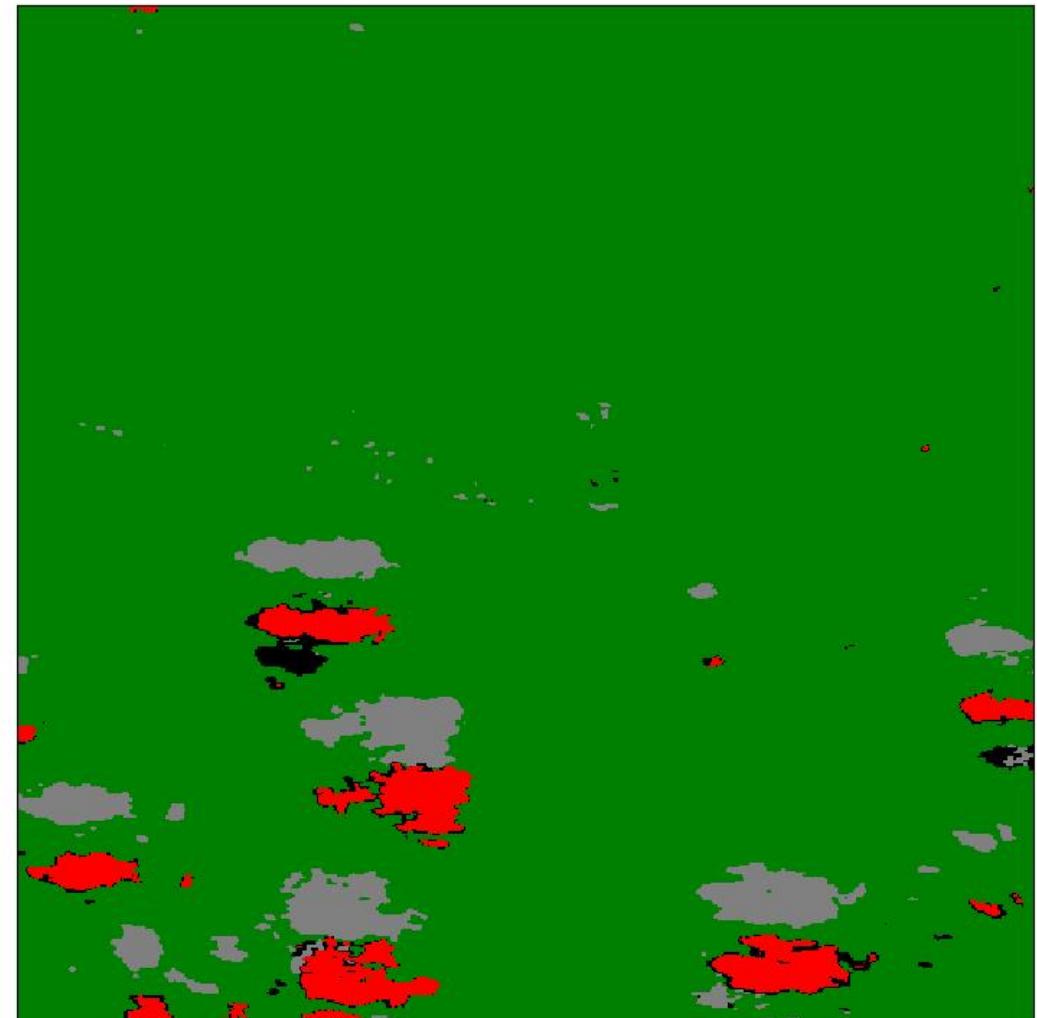
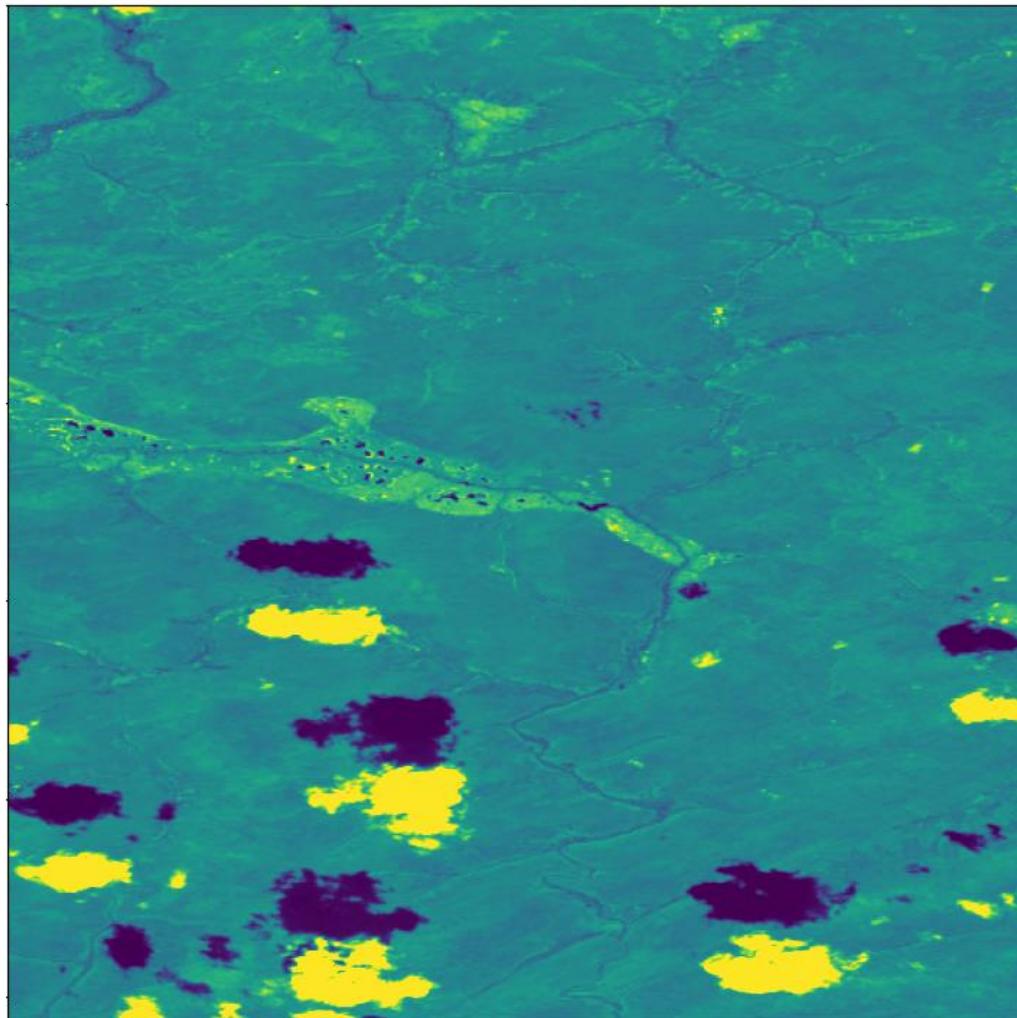
	<b>Distance=1</b>	<b>Distance=3</b>	<b>Distance=5</b>	<b>Distance=7</b>	<b>Distance=10</b>
Patch=5	0.280813	0.213445	0.000000	0.000000	0.000000
Patch=10	<b>0.332336</b>	0.321765	0.320212	0.284687	0.000000
Patch=15	0.327761	0.328312	0.329729	0.319577	0.293593
Patch=20	0.327395	0.327530	0.328723	0.327731	0.319656
Patch=25	0.320297	0.327619	0.324680	0.319776	0.315179
Patch=30	0.322501	0.323773	0.318983	0.315839	0.309315
Patch=35	0.307943	0.313917	0.315437	0.313621	0.313452
Patch=40	0.289545	0.296772	0.294594	0.289374	0.292138



Средняя разница метрики f-score на основе сцен с “качественной” и “некачественной” масками scl после обучения классификатора случайный лес и оценке качества на тестовых данных

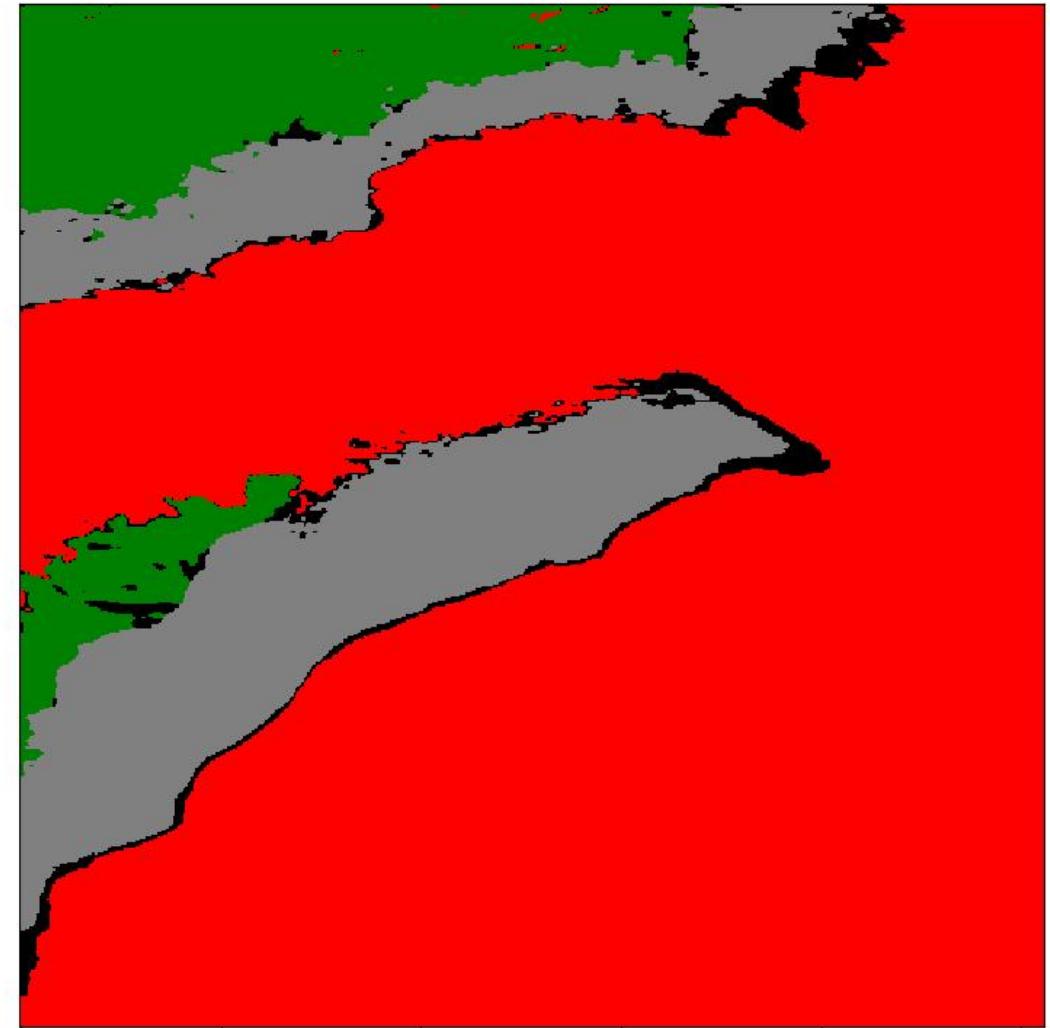
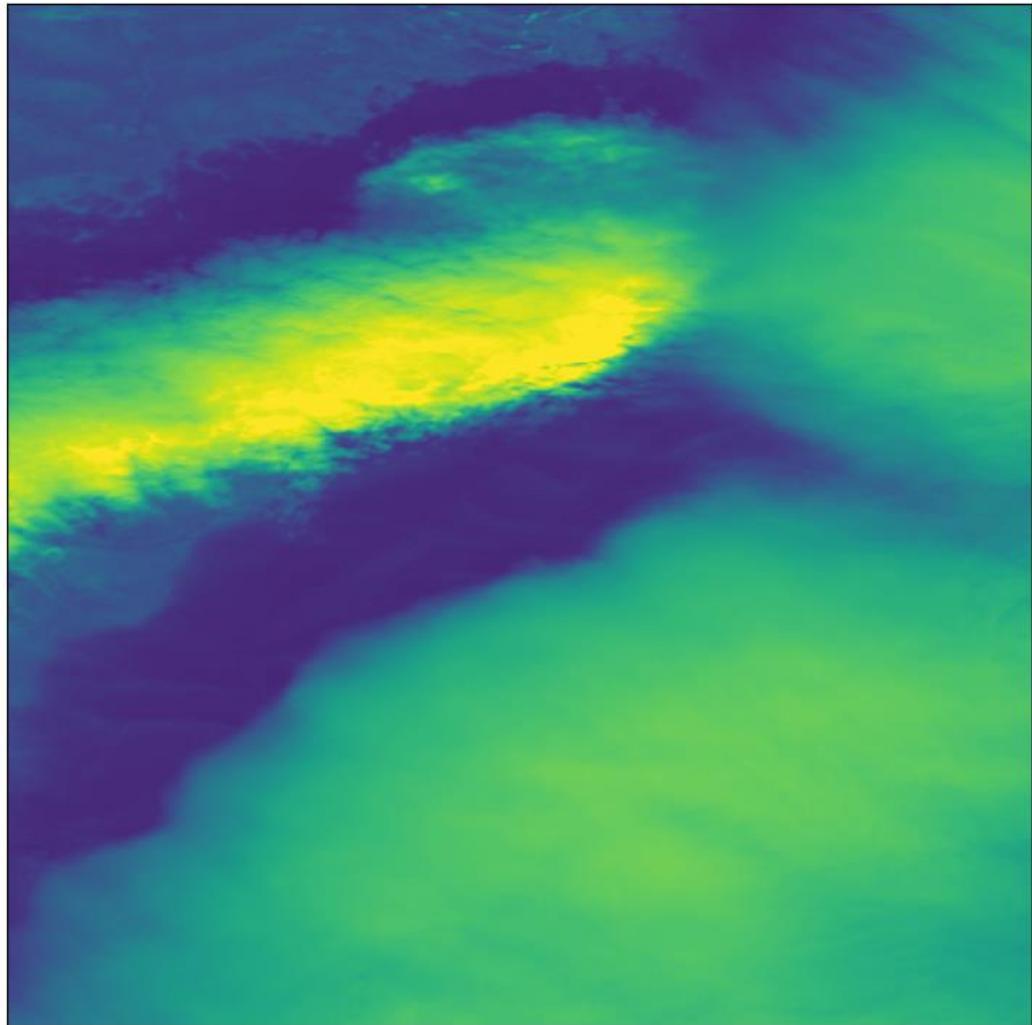
**F-score**

## Локальная оценка качества маски SCL на основе GLCM



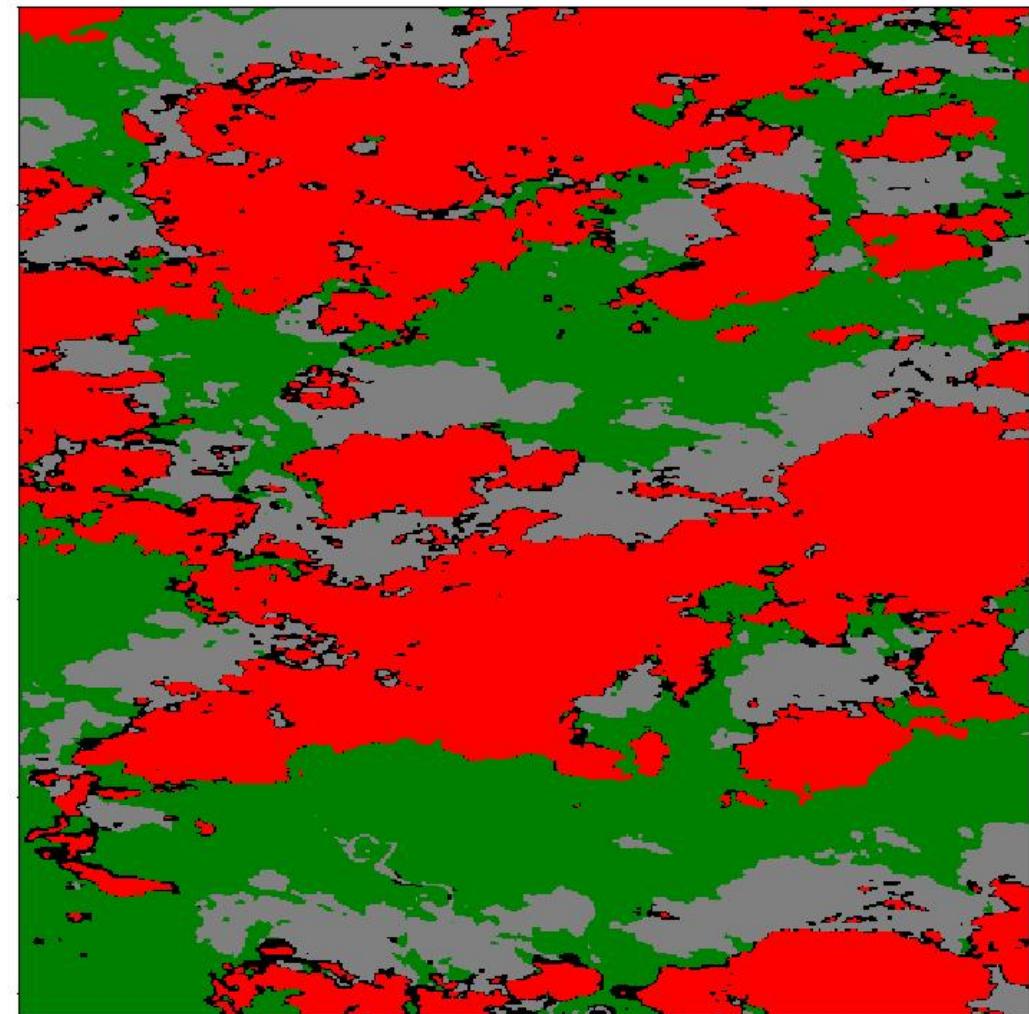
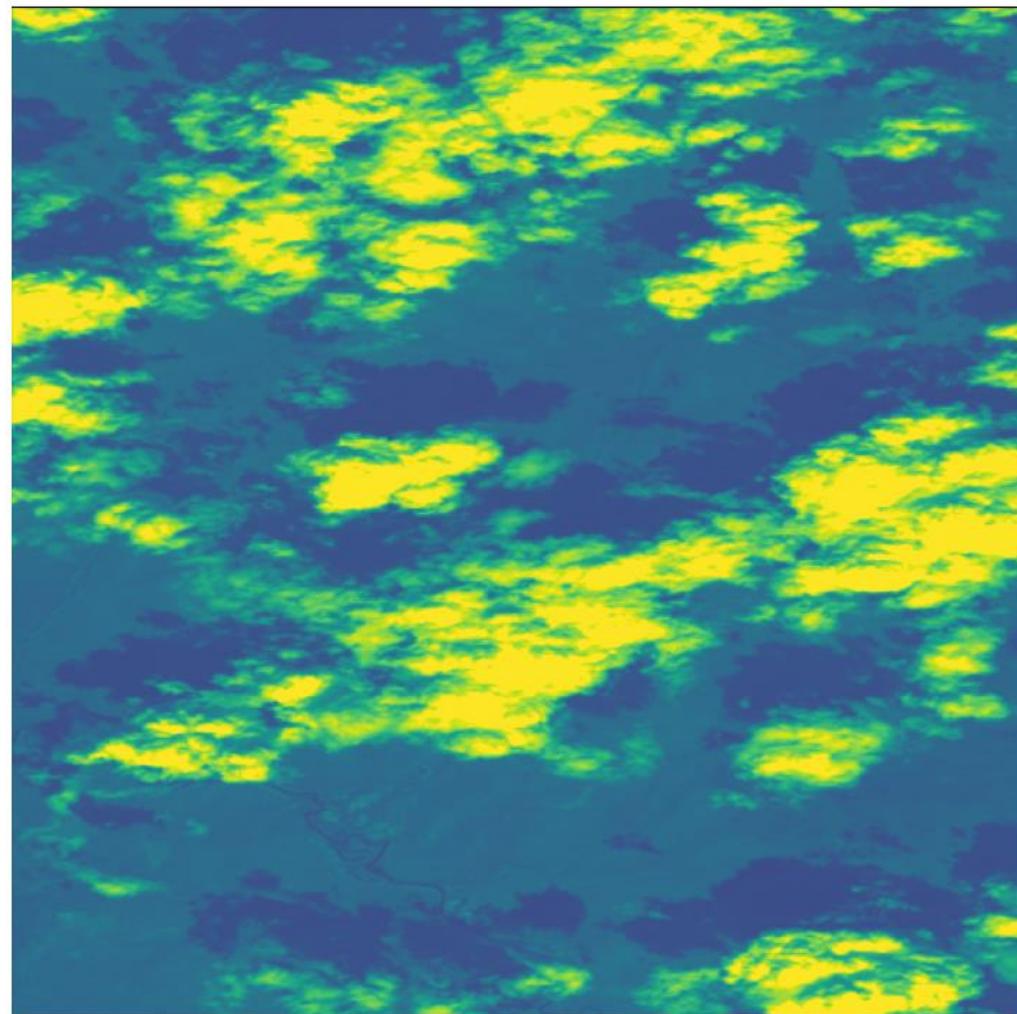
часть сцены MSI в ИК канале (слева) и соответствующая часть маски scl. зеленый цвет-чистая территория (f-score 99%), серый-тени (f-score 79%), красный-облачность (f-score 91%), черный-unclassified.

## Локальная оценка качества маски SCL на основе GLCM



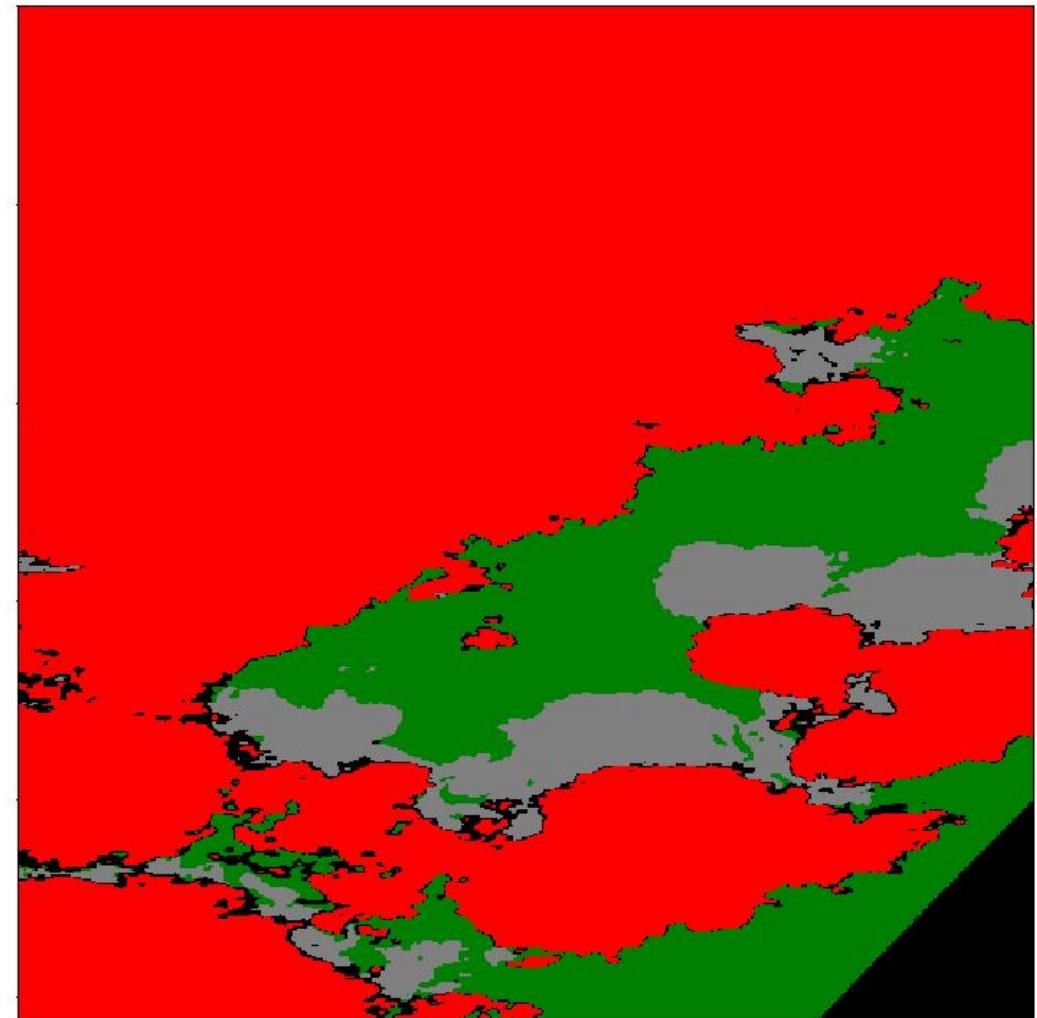
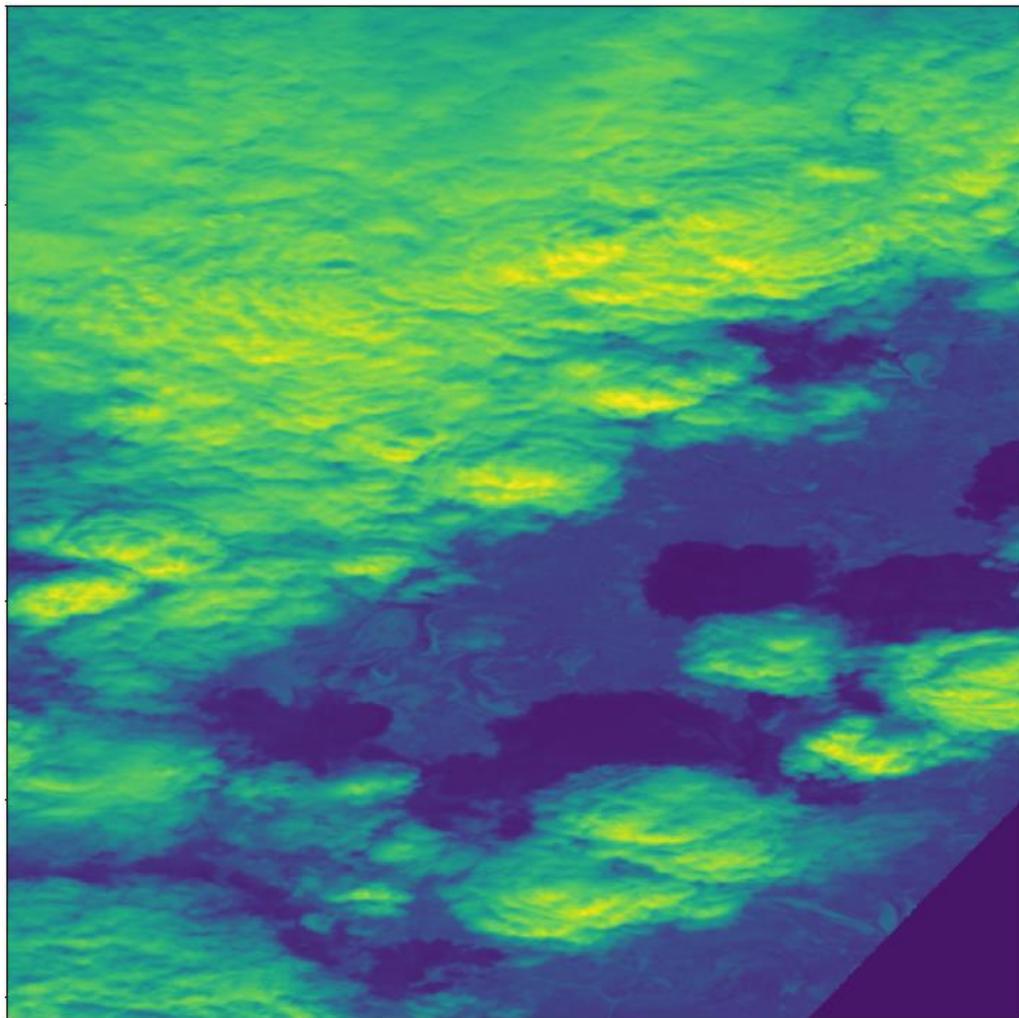
часть сцены MSI в ИК канале (слева) и соответствующая часть маски scl. зеленый цвет-чистая территория (f-score 91%), серый-тени (f-score 92%), красный-облачность (f-score 99%), черный-unclassified.

## Локальная оценка качества маски SCL на основе GLCM



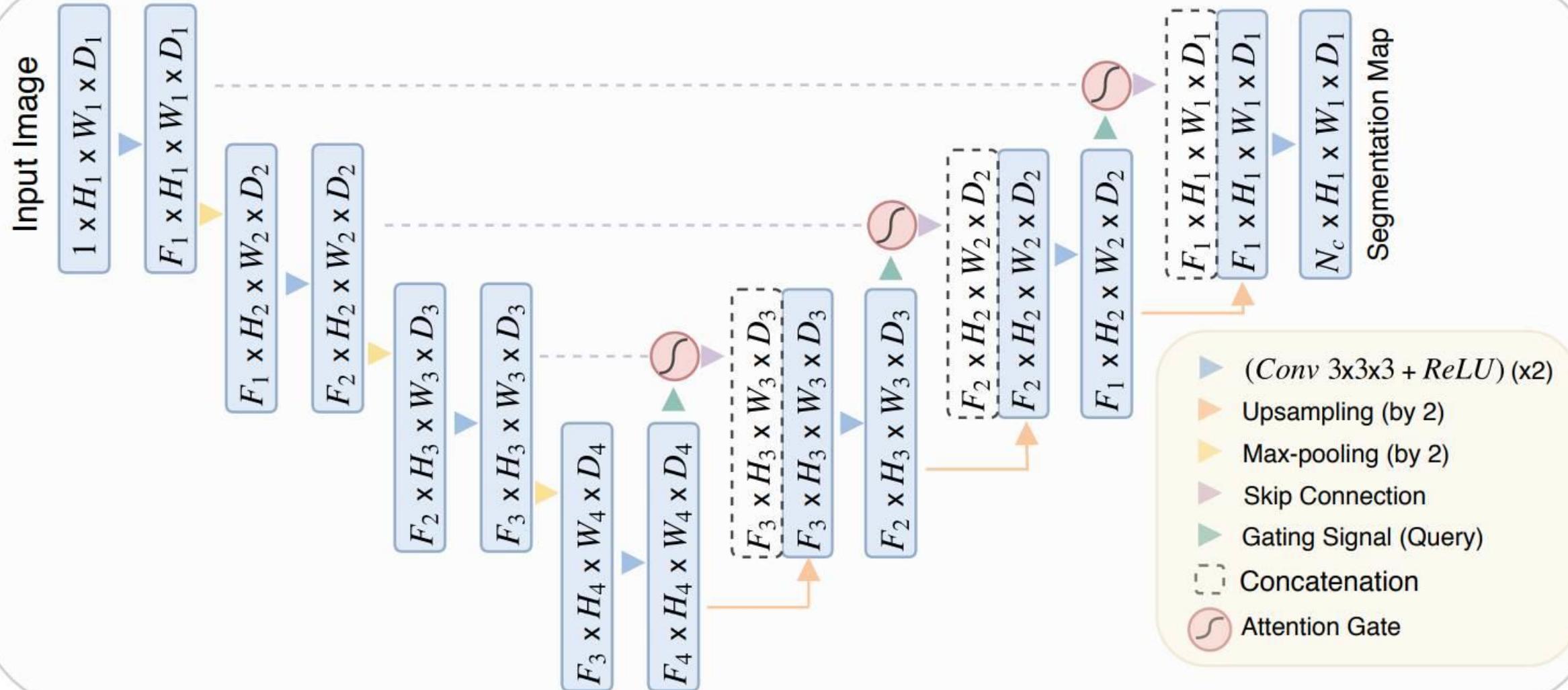
часть сцены MSI в ИК канале (слева) и соответствующая часть маски scl. зеленый цвет-чистая территория (f-score 85%), серый-тени (f-score 76%), красный-облачность (f-score 92%), черный-unclassified.

## Локальная оценка качества маски SCL на основе GLCM

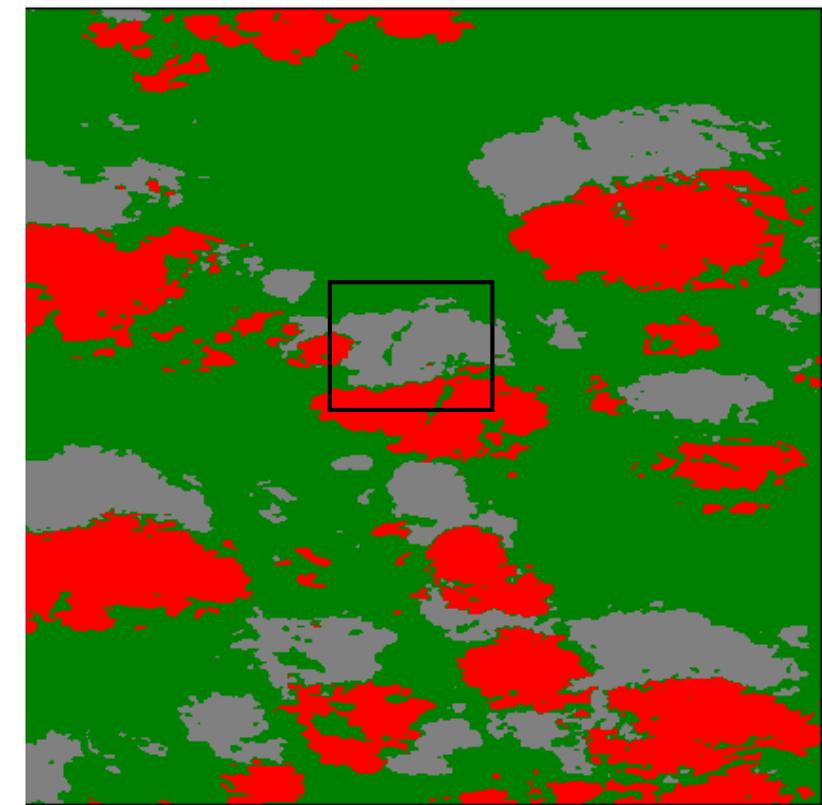
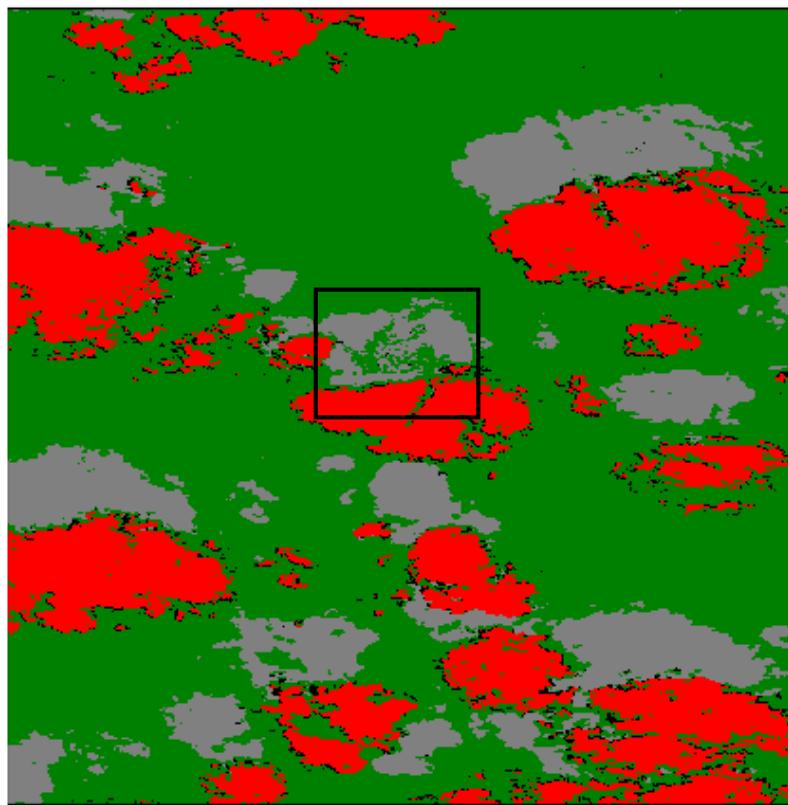
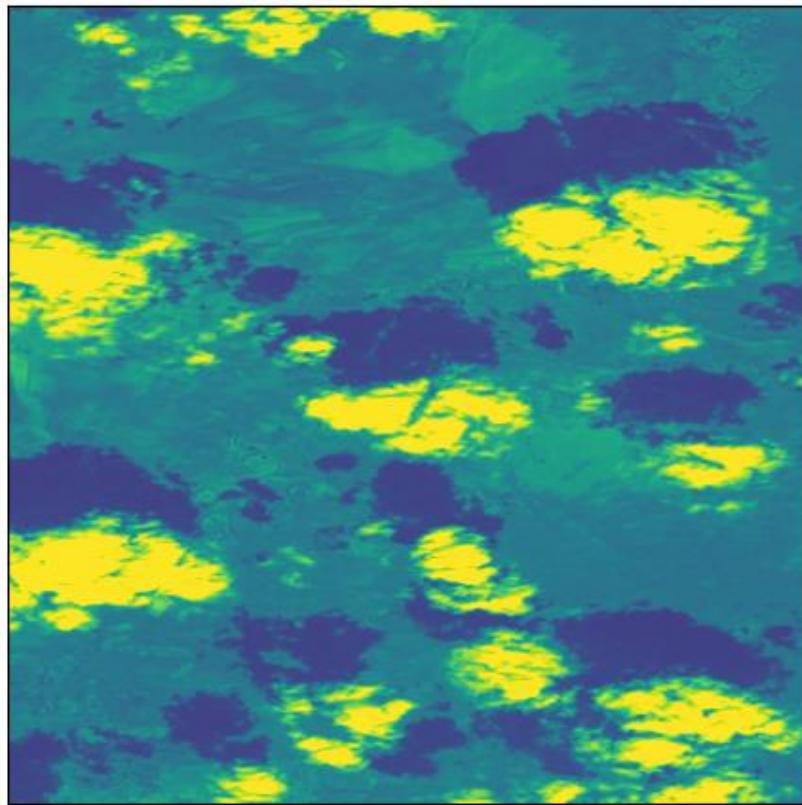


часть сцены MSI в ИК канале (слева) и соответствующая часть маски scl. зеленый цвет-чистая территория (f-score 87%), серый-тени (f-score 68%), красный-облачность (f-score 98%), черный-unclassified.

# Attention U-Net

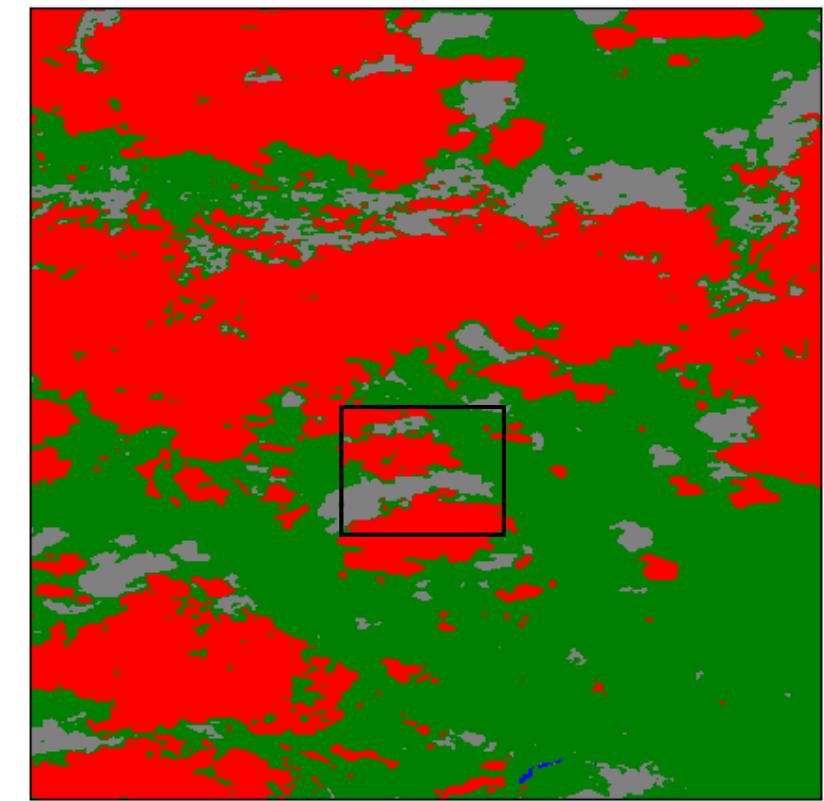
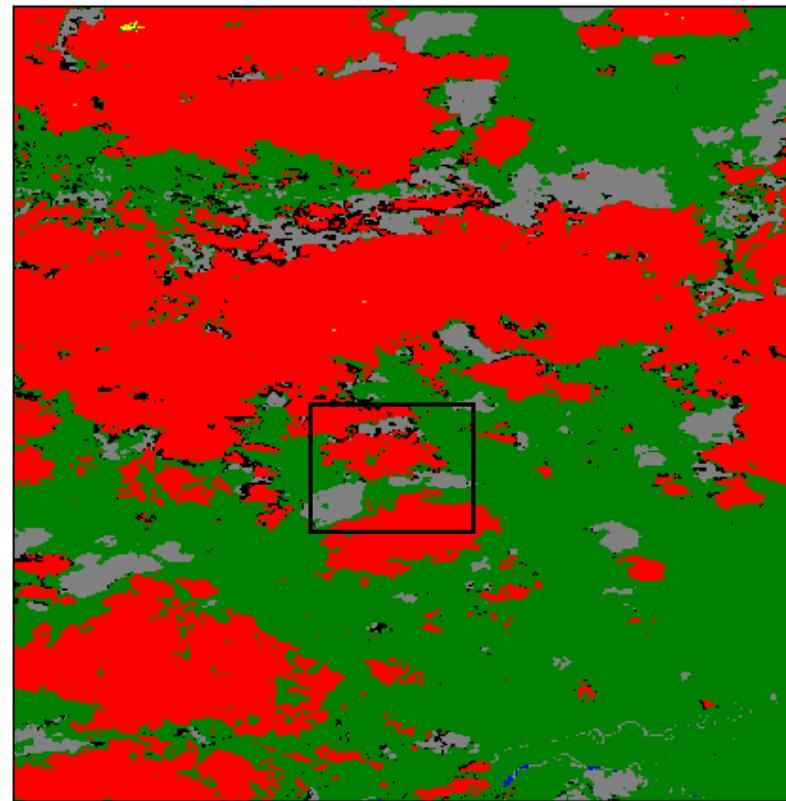
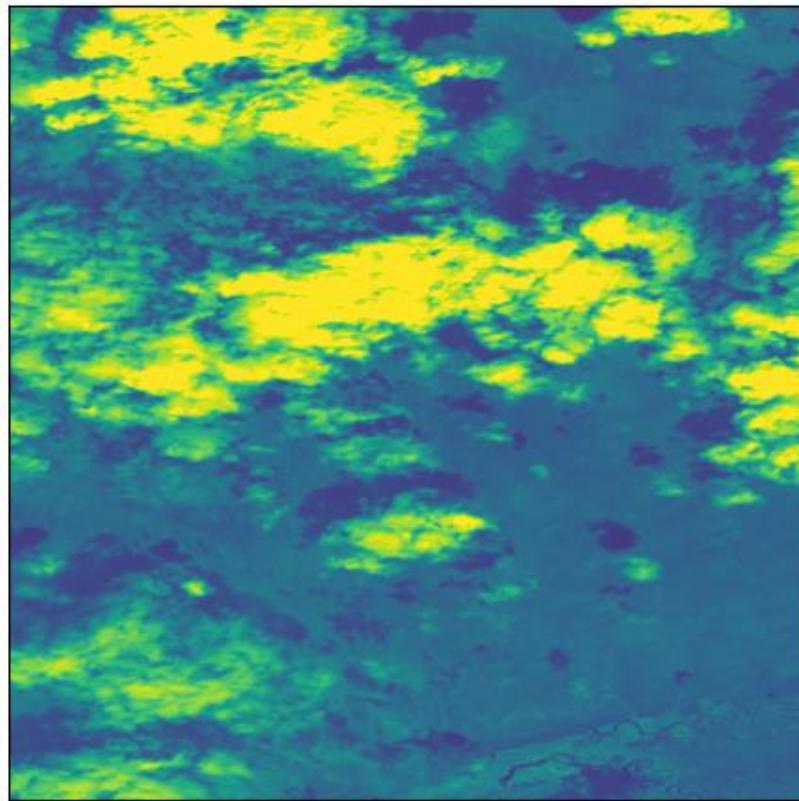


## Улучшение загруженной маски SCL



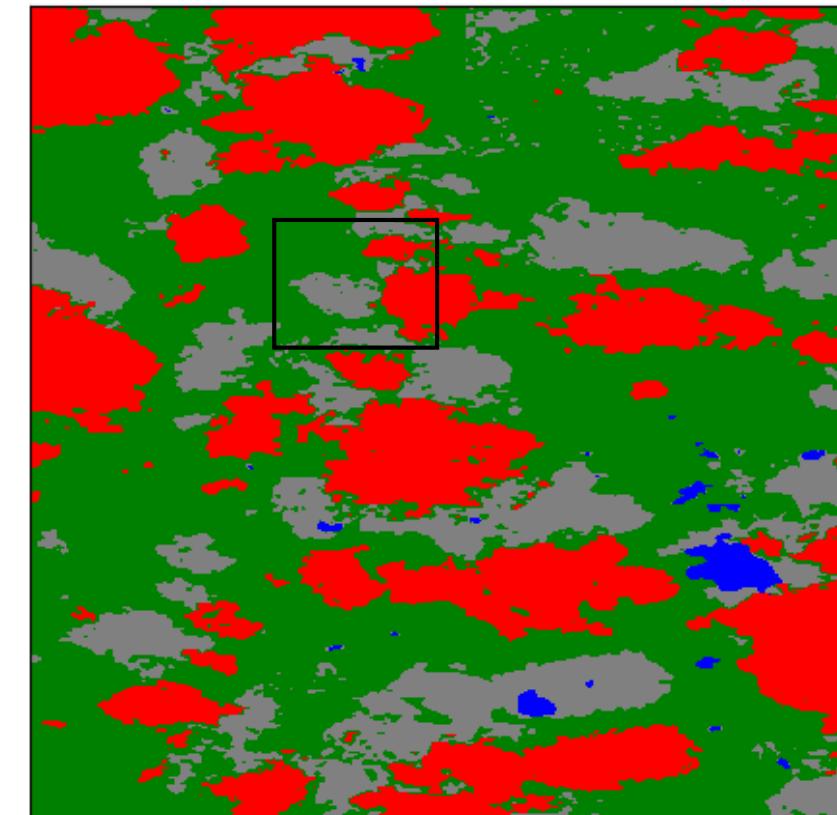
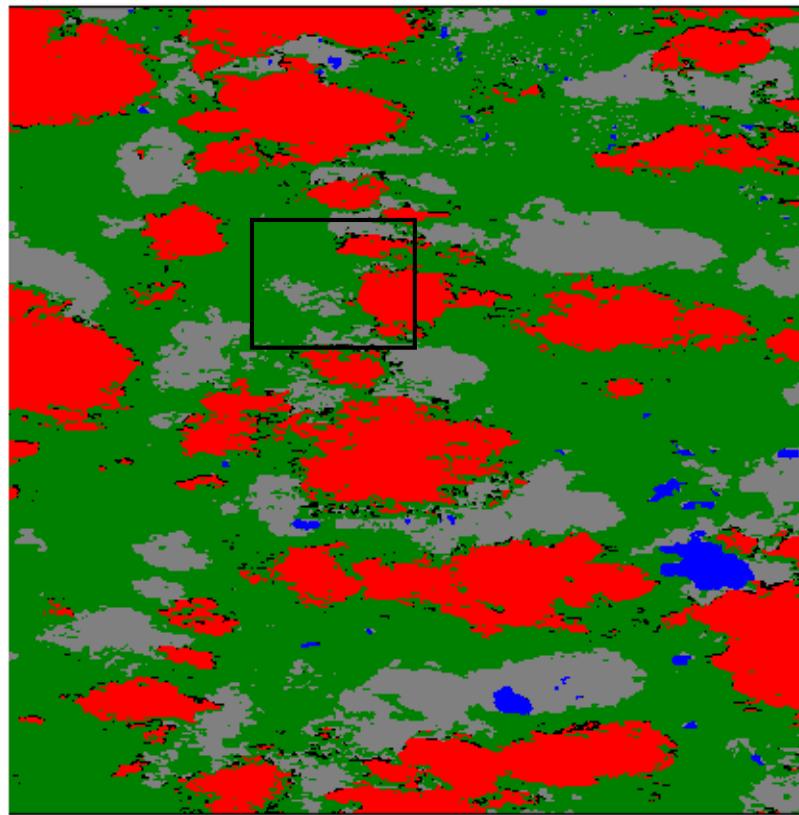
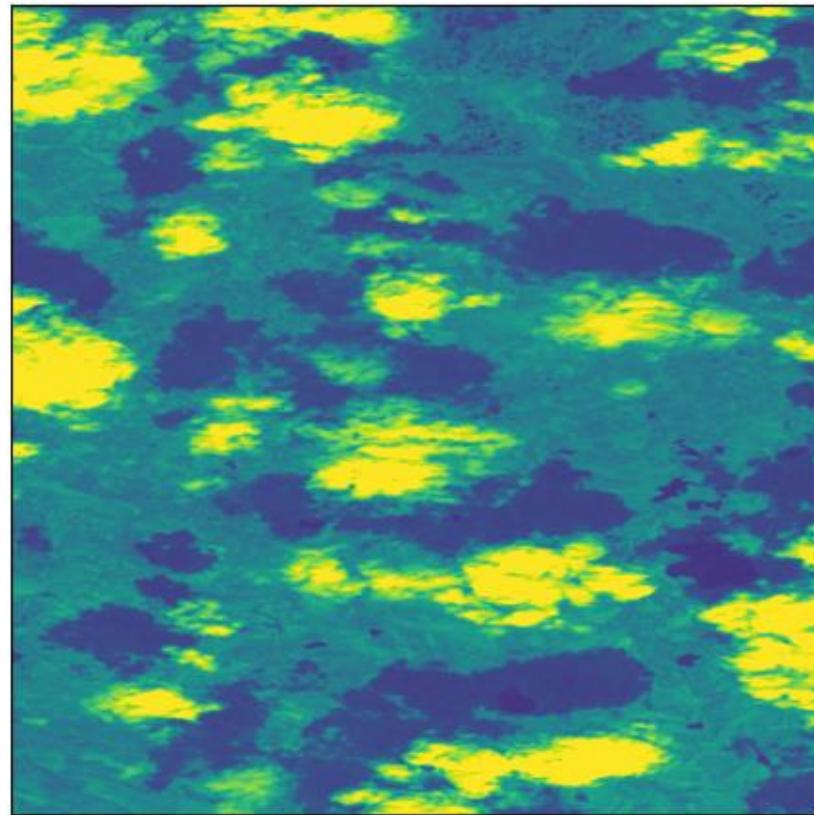
часть сцены msi в ик канале (слева), загруженная маска scl (центр), маска на основе attention u-net (справа).

## Улучшение загруженной маски SCL



часть сцены msi в ик канале (слева), загруженная маска scl (центр), маска на основе attention u-net (справа).

## Улучшение загруженной маски SCL



часть сцены msi в ик канале (слева), загруженная маска scl (центр), маска на основе attention u-net (справа).

# Метеор-М: общие характеристики



**Тип спутника —  
метеорологический**

**Орбита КА —  
круговая, солнечно-  
синхронная**

**высота: 814-843 км  
наклонение: 98,57°**

# Основные технические характеристики аппаратуры КМСС-2/МСУ-100ТМ



Рисунок 1.1 - Внешний вид прибора МСУ-100ТМ

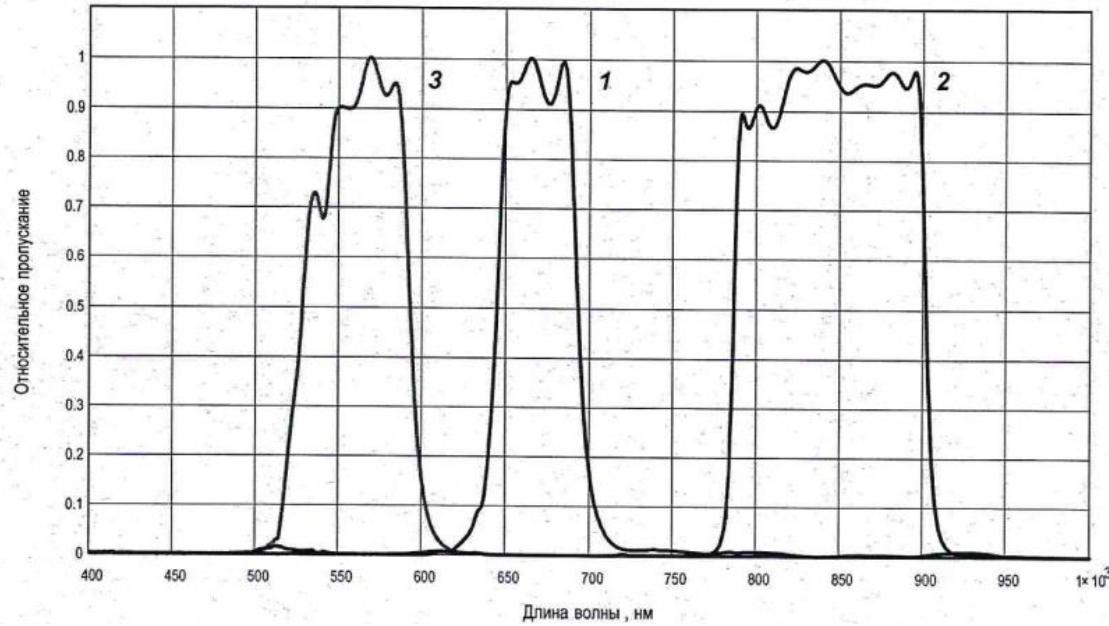
**Формируемая полоса обзора – не менее 1000 км  
(2 камеры)**

**Угол установки относительно местной вертикали  
–  $\pm 15.2$  град.**

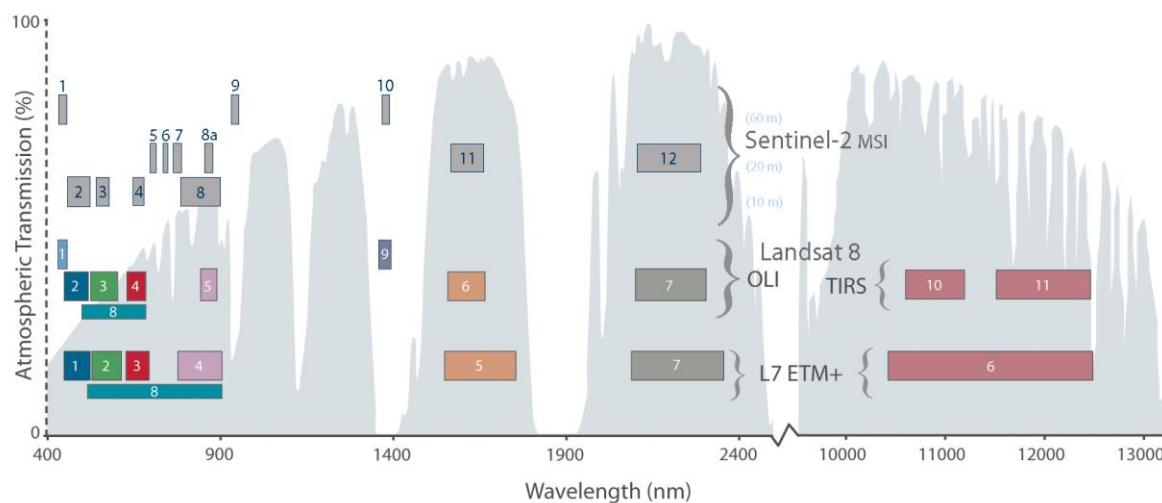
**Размер проекции элемента разрешения на  
земную поверхность (в направлении оптической  
оси прибора) – 60 метров.**

**Спектральные зоны – 520-590 нм, 640-690 нм,  
785-900 нм.**

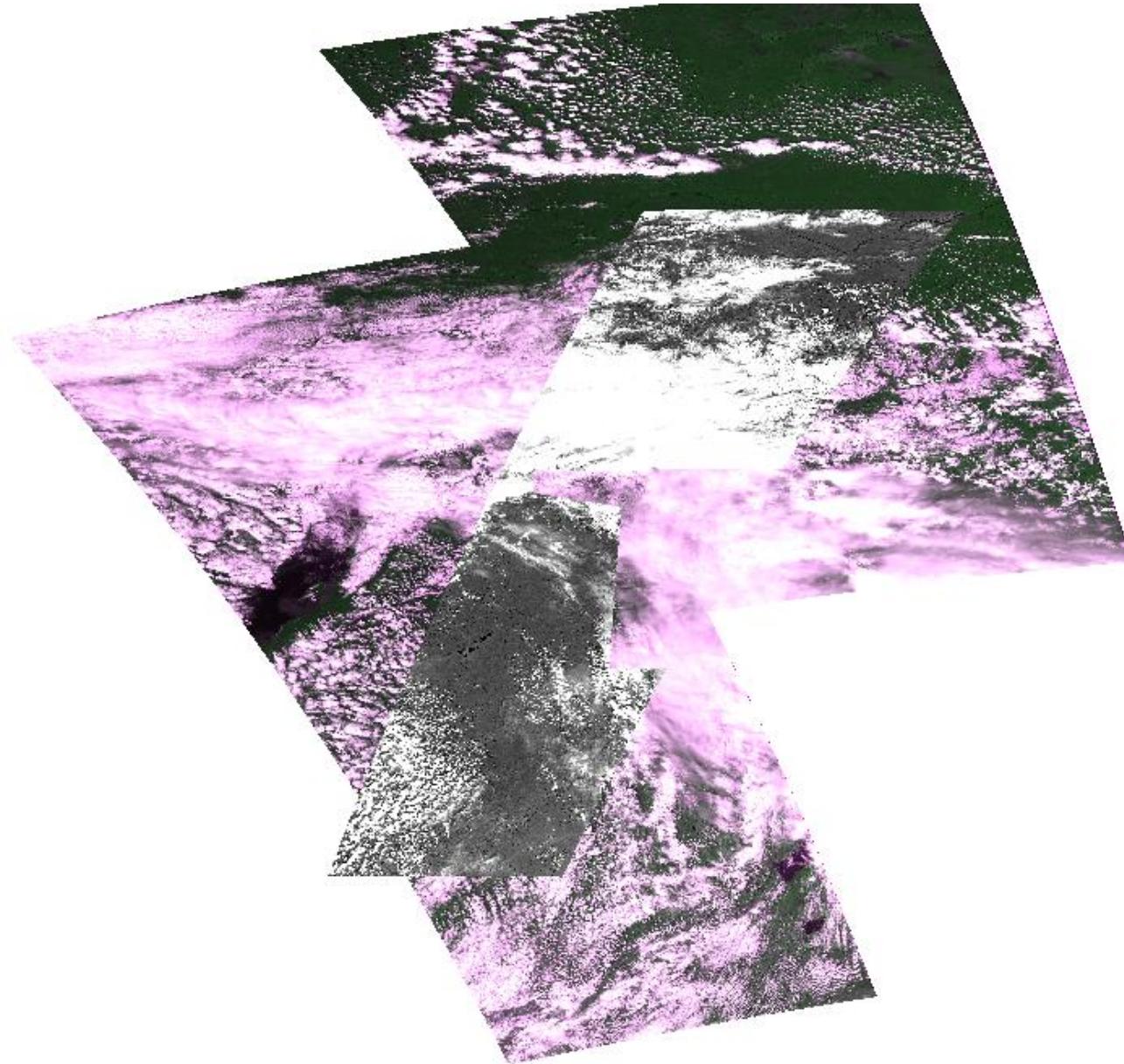
# ФСЧ КМСС и MSI



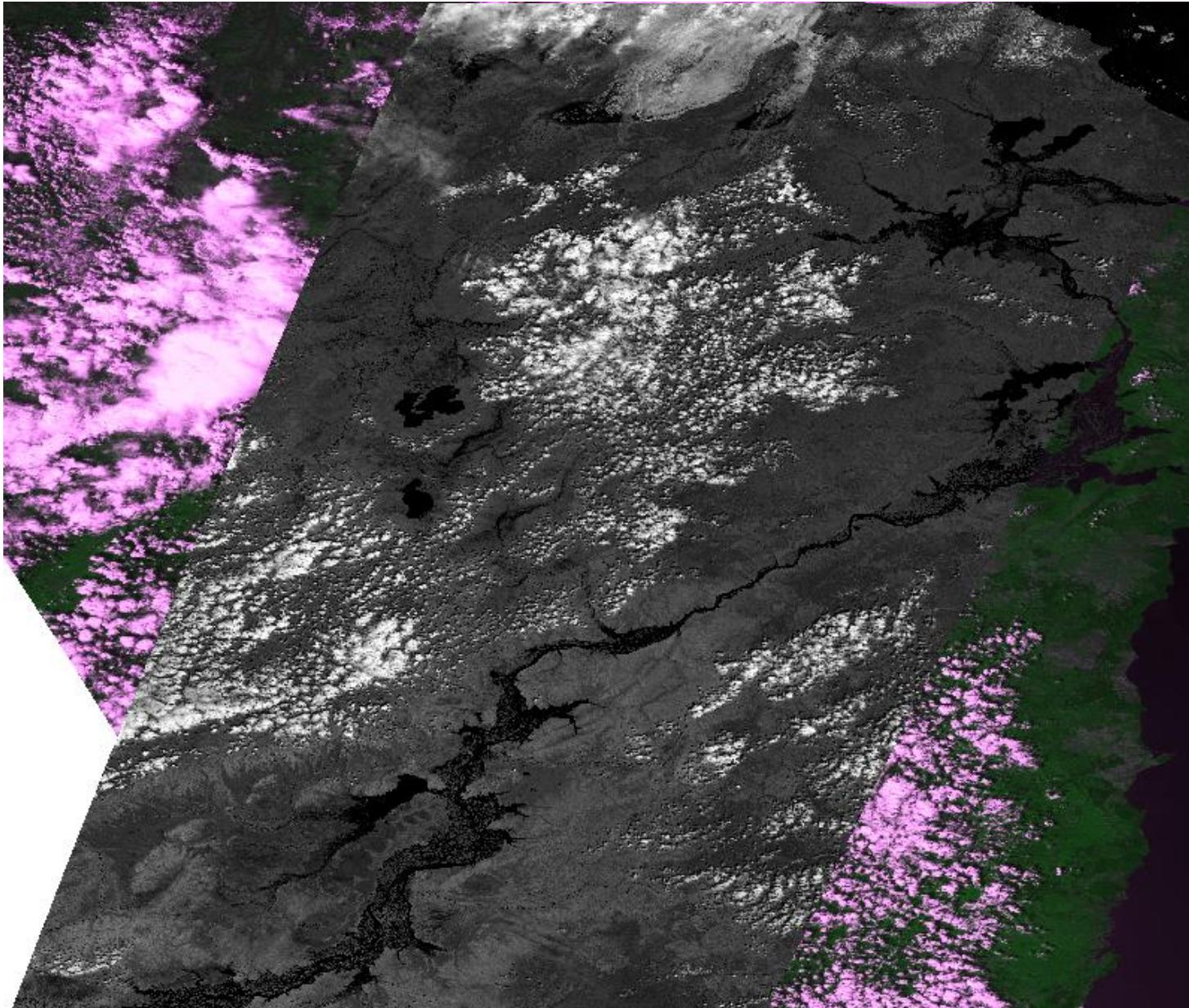
Comparison of Landsat 7 and 8 bands with Sentinel-2



## Общее покрытие данными приборов КМСС и MSI



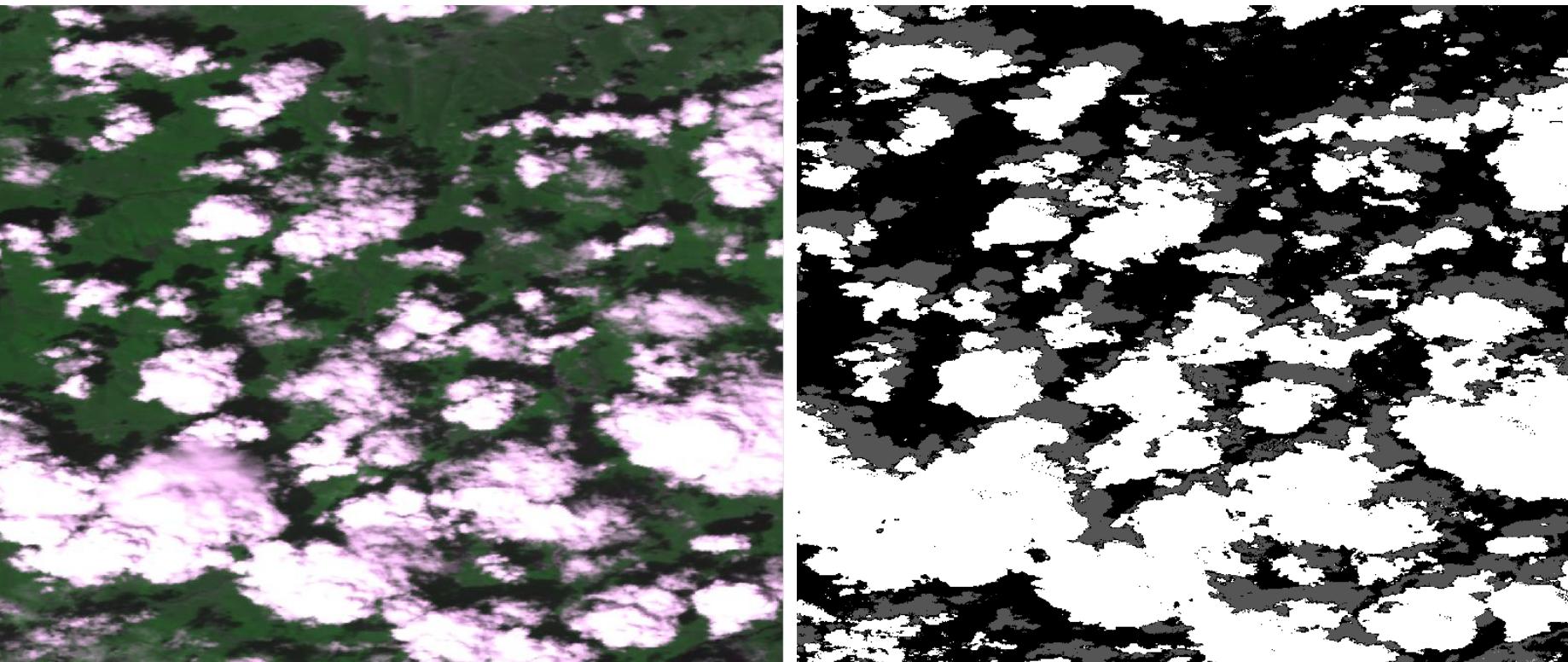
## Общее покрытие данными приборов КМСС и MSI



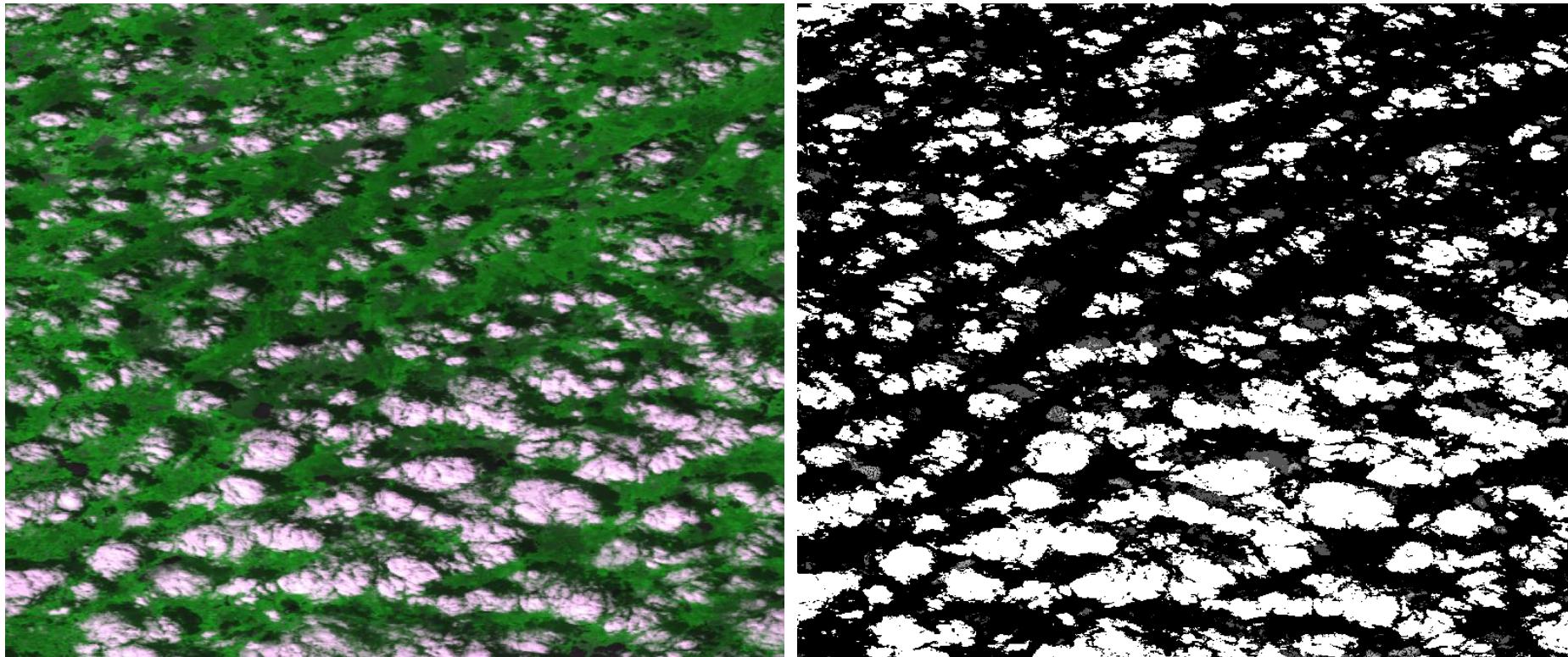
## Автоматический метод переноса масок scl на данные КМСС

- Выбор пар на основании минимальной разницы во пространстве и времени
- Маскирование облачности по данным Sentinel-2 и КМСС на основе соответствующих ранее разработанных нейросетевых моделей
- По каждому из источников делается перцентильная нормализация, то есть значения яркости безоблачной поверхности переводятся в значения соответствующих перцентиляй. После этого оба источника становятся сравнимыми и пригодными для переноса модели
- Допривязывалось изображение КМСС по данным в ИК канале методом на основе sift с годовым композитным изображением Sentinel-2 в качестве эталона
- На основе кропов scl с высокими значениями метрик f-score по всем классам обучалась модель Random Forest для Sentinel-2 которая переносилась на соответствующую территорию КМСС
- Локализованная проверка качества перенесенной маски на основе текстурных признаков по данным КМСС

## Результат переноса маски SCL на данные КМСС



# Проверка результатов переноса маски SCL на данные КМСС на основе текстурных признаков GLCM (Gray Level Co-Occurrence Matrix)



Значение f-score на валидационных данных по классу теней на основе  
Random Forest по текстурным признакам составляет 0.10

# **Результаты обучения нейросетей на тестовых данных КМСС**

## **Unet**

**N=32 accuracy: 0.8534, f1-score: 0.6888**  
**N=64 accuracy: 0.8857, f1-score: 0.6911**  
**N=96 accuracy: 0.8794, f1-score: 0.6791**

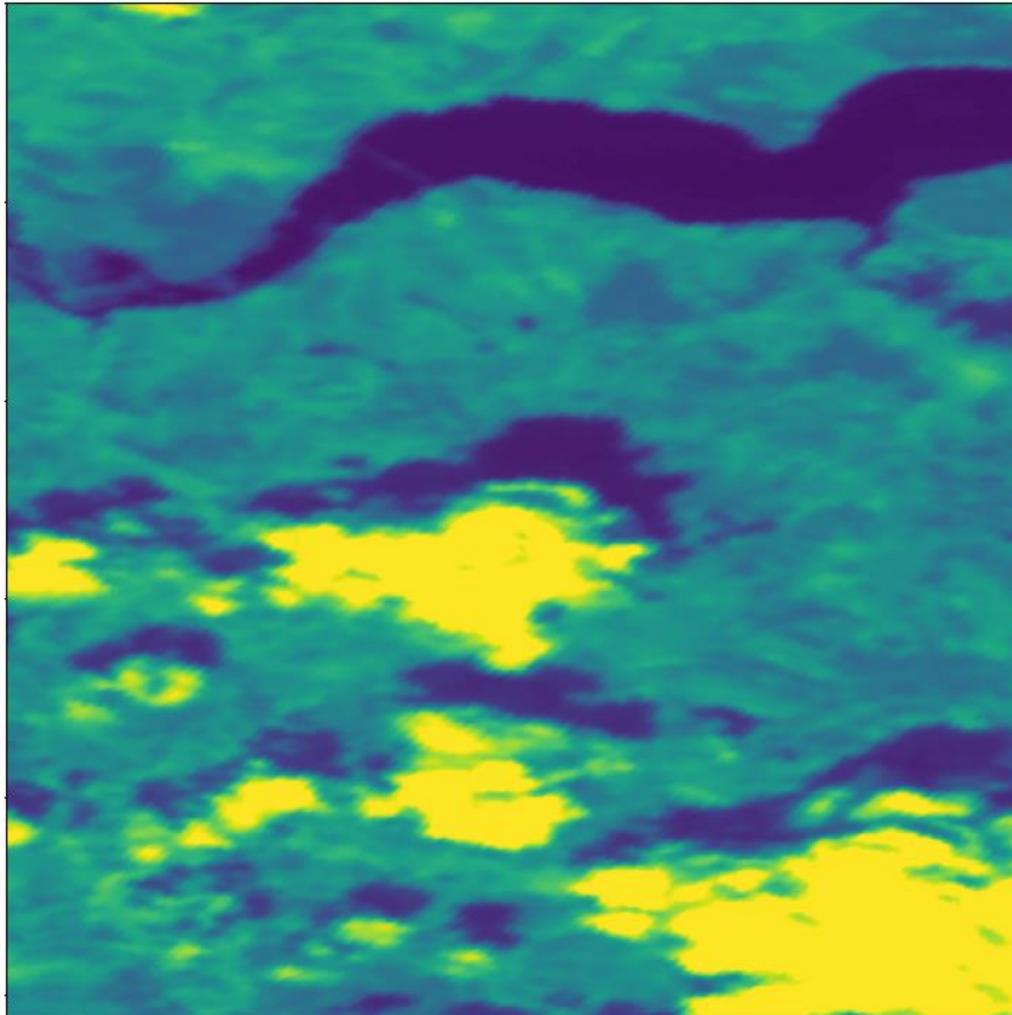
## **Attention Unet**

**N=32 accuracy: 0.8513, f1-score: 0.6825**  
**N=64 accuracy: 0.9509, f1-score: 0.7462**  
**N=96 accuracy: 0.9508, f1-score: 0.7483**

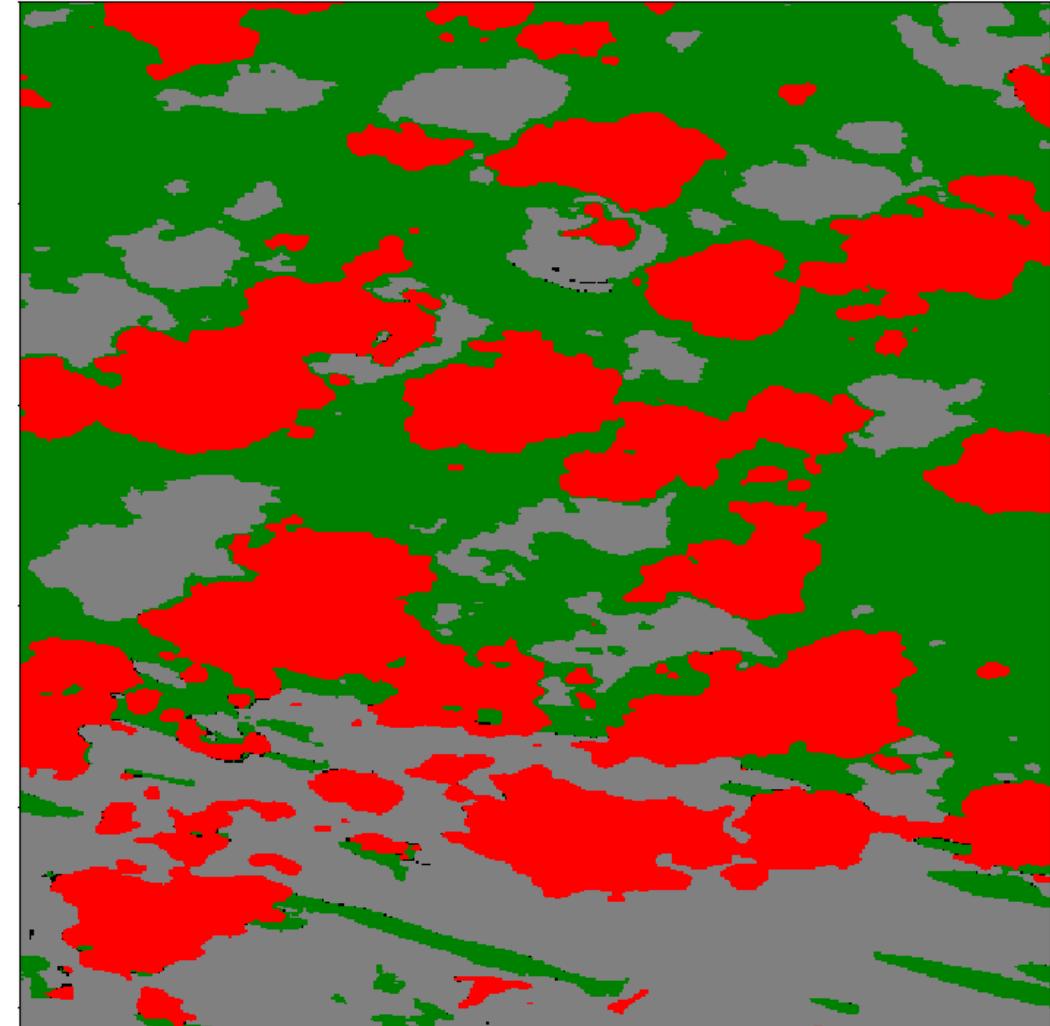
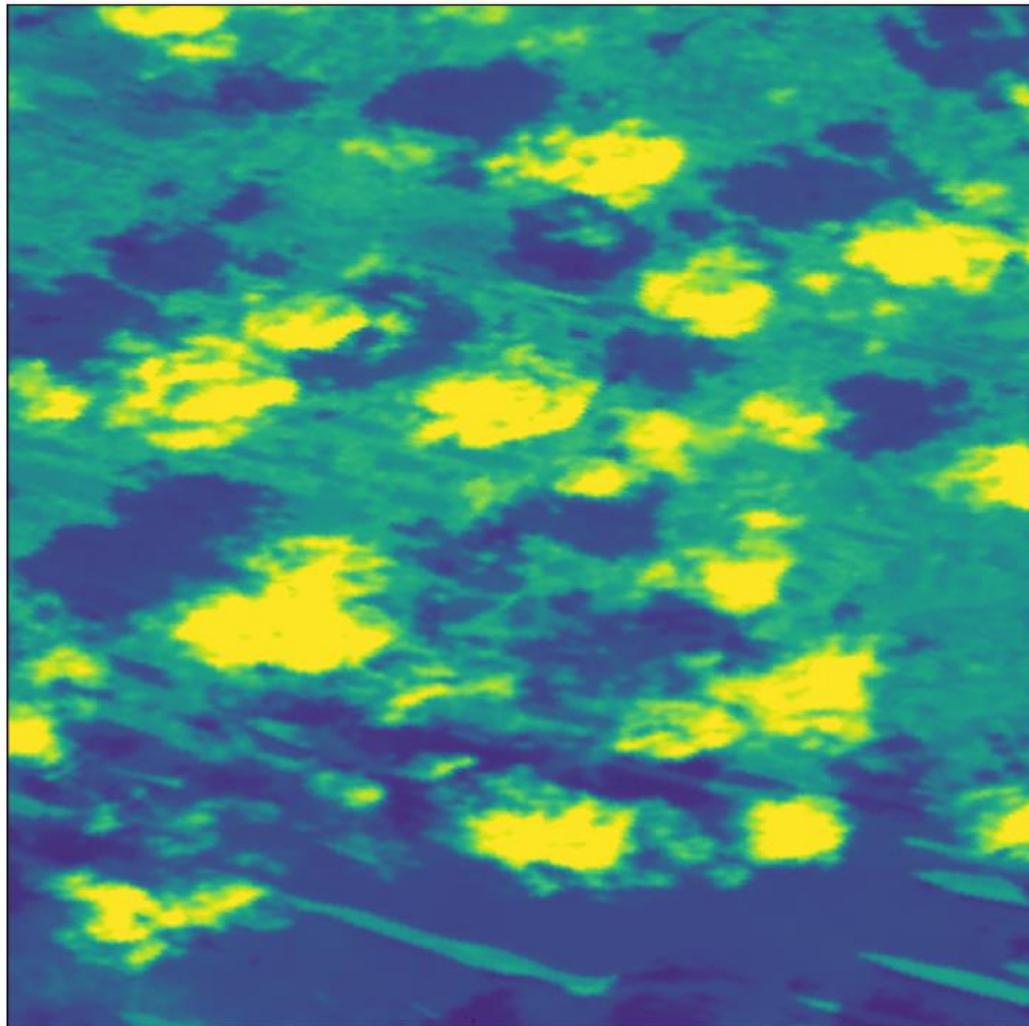
## **Attention Res. Unet**

**N=32 accuracy: 0.9339, f1-score: 0.7499**  
**N=64 accuracy: 0.9308, f1-score: 0.7321**  
**N=96 accuracy: 0.9555, f1-score: 0.7564**

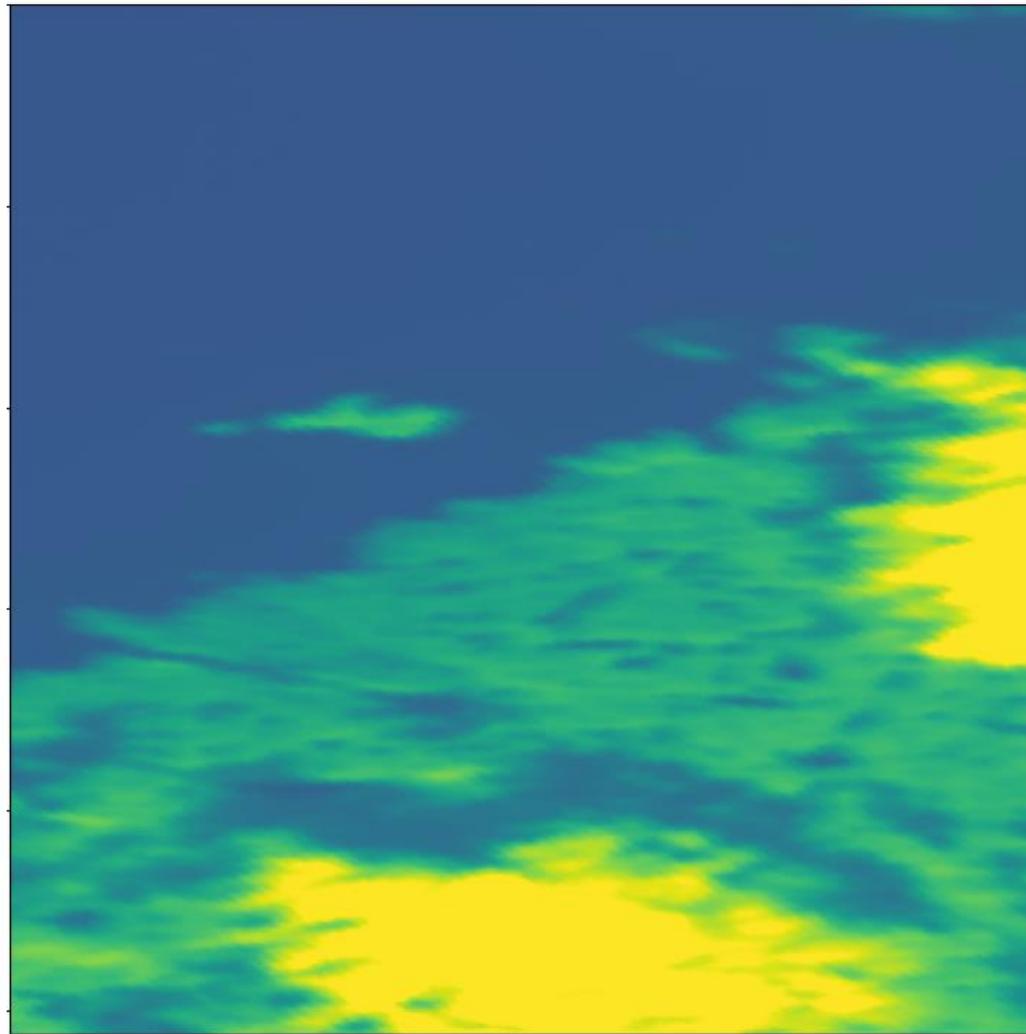
## Результат обучения нейросети на тестовых данных КМСС



## Результат обучения нейросети на тестовых данных КМСС



## Результат обучения нейросети на тестовых данных КМСС



## Выводы и перспективы

- Проведен эксперимент по сравнительному анализу информативности спектральных каналов и их соответствующих glcm метрик для решения задачи независимой оценки качества маски scl
- Разработан метод независимой оценки качества масок облачности, теней на основе текстурных признаков glcm. Преимуществом метода является использование данных всего-лишь двух спектральных каналов
- Создан набор данных Sentinel-2 с масками scl высокого качества для обучений нейросети с целью улучшения тематического продукта
- Разработан автоматический метод переноса масок scl на данные КМСС на основе перцентильных признаков
- Создан набор данных КМСС с перенесенными масками высокого качества для обучения нейросетей

**Исследование выполнено при поддержке Министерства науки и высшего образования Российской Федерации (соглашение № 075-15-2024-663, уникальный идентификатор 13.2251.21.0264).**