VILNIUS UNIVERSITY
FACULTY OF MATHEMATICS AND INFORMATICS
INFORMATICS STUDY PROGRAMME

Spatial-temporal detection in satellite image monitoring

Erdvinis ir laikinis aptikimas palydovinių vaidų stebėjime

Master thesis

Author:	Žygimantas Augūnas
VU e-mail:	zygimantasaugunas@mif.stud.vu.lt
Supervisor:	dr. Linas Petkevičius
Reviewer:	Prof. dr. Gintautas Dzemyda

Vilnius – 2024

Contents

Acknowledgement

The author is thankful for the high performance computing resources provided by the Information Technology Research Center of Vilnius University.

Summary

Change detection is one of many topics in the remote sensing domain. Advancements in this topic aid various use cases spanning from infrastructure planning or urbanization tracking to environmental problem-solving and monitoring. Generally, remote sensing has gained a lot of attraction from researchers. Various models that have scored impressive results in the computer vision domain have also shown noteworthy accomplishments in the remote sensing area. However, popularity distribution seems to decay when coming to more concrete applications such as global warming change detection. Both, global and national attraction is a little scarce and one of the main problems could be challenges that come in the remote sensing domain: optical aberrations, data quality, seasonal changes interference and knowledge transfer difficulties. With all things in mind, this study has proposed medium analysis which aims to filter out noisy data and modified an existing neural network. The model was build to employ knowledge sharing by incorporating Prithvi's model feature maps into the detection pipeline. Its performance was slightly better across all 3 test cases: original OSCD dataset, MBSC dataset and manually gathered imagery of Lithuania. The objectives of this work include analysing relevant literature and datasets, investigating and proposing methods for working with multi-spectral images, developing new neural network by using existing algorithms and evaluating proposed solutions, providing recommendations. The hypothesis being tested is that it is possible to successfully track changes using existing satellite imagery resources.

Keywords: remote sensing, change detection, neural networks, machine learning, image segmentation, water bodies, global warming

Santrauka

Pokyčių nustatymas yra viena iš nuotolinio stebėjimo srities temų. Šios srities pažanga padeda spręsti įvairias problemas — nuo infrastruktūros planavimo, urbanizacijos stebėjimo iki aplinkosaugos problemų sprendimo ir stebėsenos. Per paskutinius metus nuotolinis stebėjimas sulaukė didelio tyrėjų dėmesio. Įvairūs modeliai, kurie pasiekė įspūdingų rezultatų kompiuterinės regos srityje, taip pat pasirodė esantys verti dėmesio nuotolinio stebėjimo srityje. Vis dėlto, matomas populiarumo sumažėjimas, kai pereinama prie konkretesnių taikymų, pavyzdžiui, globalinio atšilimo pokyčių nustatymo. Tiek pasauliniu, tiek nacionaliniu mastu susidomėjimas nėra didelis ir viena iš pagrindinių priežasčių galėtų būti nuotolinio stebėjimo srityje kylantys iššūkiai: optinės palydovinių nuotraukų aberacijos, duomenų kokybė, sezoniniai pokyčiai ir žinių perdavimo sunkumai. Atsižvelgiant į šiandieninę situaciją ir galimus iššūkius, šiuo tyrimu pasiūlyta medianos analizė, kuria siekiama išfiltruoti triukšmingus duomenis, patobulintas egzistuojantis neuroninis tinklas, kad būtų galima įgyvendinti dalijimasį žiniomis, į algoritmo veikimą įtraukiant ”Prithvi” modelio požymių žemėlapius. Šis modelis parodė nežymiai geresnius rezultatus 3-juose testinėse aplinkose: naudojant originalų OSCD duomenų rinkinį, MBSC duomenų rinkinį ir rankiniu būdu surinktus Lietuvos vaizdus. Uždaviniai: išanalizuoti susijusią literatūrą ir duomenų rinkinius, ištirti ir pasiūlyti darbo su daugiaspektriais vaizdais metodus, sukurti naują neuroninio tinklo architektūrą, patobulinant esamą ir įvertinti pasiūlytus sprendimus, pateikti rekomendacijas. Tikrinama hipotezė, kad pokyčius galima sėkmingai sekti naudojant esamus palydovinių vaizdų išteklius.

Raktiniai žodžiai: nuotolinis stebėjimas, pokyčių nustatymas, neuroniniai tinklai, nuotraukų segmentacija, vandens telkiniai, globalinis atšilimas

Introduction

Change detection is a process of identifying object variations by observing them at different times. Algorithms utilise various datasets in order to measure temporal effects on areas of interest. One of the domains where this is widely used in remote sensing analysis is change detection. Essentially, remote sensing change detection is applied by gathering data from Earth-orbiting satellites and identifying relief/structure permutations that have happened throughout a specific time track. This topic specifically has gained huge support recently and it has been applied in many fields, such as 1. Deforestation tracking [SPS+19]. 2. Waste identification [ZYD+22] [MDA18]. 3. Urban development tracking [MS21].

Widespread success has caused the development of new innovative methods. It is vital to note that change detection analysis started with traditional and classical methods. Such methods like algebraic analysis difference [KLZ+18] or regression [RL98] were implemented to identify significant changes in imagery. However, these methods are extremely sensitive to non-standard images and detect changes based on threshold values, which turned out to be ineffective in ever-changing conditions. To overcome these obstacles, neural networks were presented. Neural networks built for change detection could be split into two groups that differ by the data that they use. Some deep neural networks use traditional 3-channel images while others employ multi-spectral imagery which generally contains far more information. There are also methods that are capable of analysing both types of images.

Remote sensing change detection based on 3-channel imagery has widely adopted U-Net [RFB15] based networks [PZG19a]. Bent, split-connected architecture was especially effective in tandem with Siamese networks [DLB18a]. The combination managed to score state-of-the-art results.

On the flip side, the second group of change detection algorithms have also scored promising results. These networks utilise not only traditional images, but also bands that are not visible to the naked eye. OSCD [DLB+18] is one of several datasets that was tested with models that are pioneering remote sensing change detection domain. For example, some models used already mentioned siamese neural network architecture [DLB18a] while others have presented far more innovative ideas. ChangeStar [ZMZ+21] deep-neural network is one of them. It changed the way domain training is implemented and used single temporal rather than bi-temporal images. This allowed the algorithm to encompass much more semantic information which is extremely viable when identifying relief permutations. Transformer-based networks were also consolidated in the remote sensing domain. ChangeFormer [BP22] used transformer blocks in the encoder which shared weights with the other network in the Siamese unit and showed notable advancement. Thus, one can see that change detection in remote sensing is actively developed.

Besides having tremendous support in change detection, there are still several challenges in multi-band analysis that need to be overcome. Firstly, data quality and distribution. Remote sensing imagery applications usually require high-quality and high-resolution (HR) images for processing (an example of satellite imagery can be found in Figure 1).

Elements from the top-down view are very small, so in order to identify changes in the image,

Figure 1. Left: Image taken from an earth-orbiting satellite. Right: Same image with semantically segmented roads, buildings and vegetation. (photo taken from [NH18])

data has to be as good as possible [YLX+15]. Yet older satellites were not as advanced and since bi-temporal images make use of same place photography taken at different times, it is unsurprising that older images are more prone to lower quality and vulnerable to optical system aberration, atmospheric disturbance or noise. While on the subject, relief displacement is one of the notable problems in remote sensing no matter how new an Earth-orbiting satellite is [GHP+19]. Another worth mentioning interference is seasonal changes. Areas that experience all 4 seasons look very different throughout the year. The change detector's task is to ignore those changes and identify only significant relief or infrastructure modifications (example of extracted water body changes Figure 2).

However, even if a model is capable of extracting the correct change map from the image, there are some non-trivial obstacles. Transferring knowledge to other domains like building detection remains a problem. There were successful attempts to share extracted features by creating sophisticated model frameworks [LXG+22] but it appears to come with the efficiency and resourcefulness cost. This complex problem has to be overcome by utilising separate datasets that are populated with very different quality and band count images.

In order to maximize the algorithm's performance, one must choose the main problem of research. Global warming is a great fundamental challenge for remote sensing change detection. Satellite imagery has various types of bands that visualise water, vegetation and man-made structures distinctively. This complements the rather huge 60 meters recording area of satellite cameras[2]. Global warming has been addressed by remote sensing researchers. The topic has adopted neural network solutions [SWZ+21] although overall interest seems a little scarce. However, what appeals the most is water bodies change detection which seems to have gotten even less attraction among

[1]https://www.esri.com/arcgis-blog/products/arcgis-pro/imagery/change-detection-using-landsat-time-series

[2]Sentinel-2 satellite is capable of obtaining imagery at high spatial resolution (10 m to 60 m) over land and coastal waters. [22]

Figure 2. Image of an Amazon river from a blog page.[1] It's an output from the change raster. The pixel values indicate the period of time when the land cover was submerged in water. Red represents earlier years and yellow represents more recent years.

researchers. At the moment of writing there have only been several research papers that analysed coastal changes in collaboration with the general land site detection domain [EH11; KCR+17].

The lack of popularity of global warming change detection is also evident at the national level, as there has been limited research published on this topic specifically in the Lithuanian landscape. The closest finding was a study which conducted an analysis of different binary classification and regression algorithms for ice detection [SUG22]. To supplement existing Lithuanian studies [FKG+22; SUG22] this study will attempt to monitor climate change while using remote sensing images.

The aim of this thesis is to develop methods to monitor spacial changes under a given time track. The main context of this study contains the analysis of water-bodies satellite imagery of Lithuania.

To achieve the aim, several research objectives should be achieved:

1. Perform analysis of scientific papers about spatial-temporal urbanization and climate change detection by identifying SOTA models, data harvest methods and relevant datasets.

2. Investigate and propose methods that would allow working with multi-spectral images (e.g. segmentation).

3. Propose unique neural network that would help to identify relief changes

4. Evaluate proposed methods and solutions in spacial changes monitoring under given time track process, provide recommendations for practitioners.

Reaching these goals, the following hypothesis will be confirmed or disproved: Given existing satellite imagery resources, it is possible to successfully track spatial-temporal changes.

1. Related work

Data and amount of it is crucial in this domain. There are several available sources for remote sensing data gathering. Government institutions have created various missions that successfully collect huge amounts of satellite imagery data and make it available to the general public. This could be helpful while developing remote sensing change detection methods. Once created, something needs to identify how good those methods perform. This would include task formulation as tasks guides machine learning models towards higher performance.

Essentially, this section should provide some insights on required topics that would aid in the dataset analysis process. Section 1.1 presents valuable satellite imagery sources that are used as a resource for multi-spectral data harvesting. Section 1.2 walks through existing datasets. They play an essential role in dataset creation as it is an exact example of what it should look like. Section 1.3 go over two selected datasets that share similar geographic properties and are especially valuable as it helps to understand what are the upcoming challenges. To continue with challenges, Section 1.4 will dive deeper into an array of general problems that creators of such resources endure. Section 1.5 covers several loss functions to get a brief outline of what possible tasks could be adopted during dataset evaluation in order to get maximum results. Lastly, Section 1.6 will cover a couple of visualisation examples to better understand how certain properties of the GIS information can be analysed and presented.

1.1. Used resources

Creating a dataset takes research on already existing resources. This applies to imagery resources as well as already built datasets review. A few satellites will be presented here. Most of them will help in this research and will also be mentioned in existing datasets analysis:

- **Sentinel 1** - is Copernicus initiative mission of two satellites orbiting around the globe. The Sentinel-1 mission includes C-band imaging with resolution up to 400 km. Most used satellite imagery is gathered from its Synthetic Aperture Radar (SAR). This imagery stands out because of its unique ability to be unaffected by clouds or bad weather. Satellites orbit in near-polar, sun-synchronous setting and gather imagery all around the world. With both satellites operating, the repeat cycle of the same point in the globe is 6 days.[3]

- **Sentinel 2** - is a continuation mission of its predecessor. Another two satellites placed in the same sun-synchronous orbit. However, sentinel 2 satellites are phased 180 degrees from each other. This minimizes revisit frequency of equator to 5 days. Image quality is also much better, satellites can produce wide-swath, high-resolution (up to 60 m. spatial resolution), multi-spectral imagery and capture data across the world.[4]

- **Moderate Resolution Imaging Spectroradiometer (MODIS)** - a system which joins two satellites: Terra and Aqua. Both satellites are also orbiting in near-pole way but each in

[3]Link to information about Sentinel 1
[4]Link to information about Sentinel 2

opposite direction: Terra orbit north to south , while Aqua flies south to north. Terra MODIS and Aqua MODIS are viewing the entire Earth's surface every 1 to 2 days and has a viewing swath width of 2,330 km. Captured data spatial resolution reaches up to 1,000m.[5]

- **Landsat** - satellite imagery collection program which have been active since 1972. Mission offers researchers a high-quality multispectral data (up to 100 m. spatial resolution) taken from all across the world. Satellites from the most recent missions (Landsat 9 at the time of writing) completes about 14 full orbits each day and crosses the entire Earth in 16 days. [6]

Researchers have used these resources to build accurate data for further analysis [SHQ+19b; SHZ18]. Some of them even includes labels for indicating changes in relief or land use. Further section will cover these solutions.

1.2. Existing datasets

Datasets are the primary source of knowledge when training the model. Their unique characteristics allow scientists to select the best-suiting one. It also forces researchers to carefully weigh each dataset from a wide variety of properties.

The dataset creation process also benefits from the research of relevant examples. To gather insight on dataset creation progress we need to formulate criteria. These measures will allow emphasizing datasets that have had similar challenges and in general are related in terms of geographic information, and data gathering processes. The following criteria will be taken into account when choosing datasets to analyse:

- **Dataset popularity** - while not the most important aspect, dataset must be used. Dataset usage is a great indicator of a good resource. Datasets' popularity will be determined by the citation count.

- **Accessibility** - even though this criterion could look like an obvious indicator of a proper dataset, there are some datasets that could have been an example but were not chosen because of broken mirror links [LJ20]. It is vital to have access to the dataset in order to make use of it, run example scripts and understand implementation details.

- **Relatedness** - when choosing references, the most relatable resources will be the ones that use the same resources. This thesis will prioritise data gathering from Sentinel satellites. Similar dataset that incorporates these resources will be covered in this section.

Change detection datasets that are mentioned in the following section can be analysed in the table Table 1. Note that several properties do not depict dataset fully. For example revisit frequency of a dataset might be a little smaller (less frequent) due to various interferences. Some pictures could have been deleted because of cloudy weather or bad quality.

[5]Link to information about MODIS
[6]Link to information about Landsat missions
[7]Multispectral images with varying resolutions

Table 1. Overview of analysed datasets. N/A indicates that no information was found about that property regarding the dataset, LC - Land usage mapping, LULC - Land usage and land cover mapping, C - Classification task, SS - Semantic segmentation, TS - Time series analysis, OD - Obejct detection. Data was taken in 2023-05-01. Accuracy was taken from the best performing model found.

	Freely available	**Popularity (citation count)**	**Data format**	**Nr. of images**	**Use cases**	**Task**	**Reference**
SEN1-2	Yes	153	RGB (in Geo-TIFF)	564,768	LC	SS, C	[SHZ18]
SEN12-MS	Yes	173	GeoTIFF	541,986	LC	SS, C	[SHQ+19b]
OSCD	Yes	247	GeoTIFF	24^7	LC	C	[DLB+18]
Hi-UCD	No	29	N/A	1293	LC	C	[TZZ+22]
HRSCD	Yes	291	GeoTIFF	291	LC	C, SS	[CLB+19]
Dynamic Earth Net	Yes	13	GeoTIFF	54,750	LULC	SS, TS	[TKW+22]
FloodNet	Yes	71	GeoTIFF	3200	Post disaster images	OD, SS	[RCS+21]
LoveDA	Yes	60	RGB	5987	LC	SS	[WZM+21]

1.3. Selected datasets

Creating a dataset is a challenging task. Data must be standardized because satellite imagery packs a diverse palette of information [AA19]. Different types of hardware in satellites can visualize imagery and are prone to interference differently. Not to mention, researchers have to decide whether to use visible imagery or work with multi-spectral data.

This section will present two example datasets and challenges that scientists have encountered. Both of them pack data from the same Copernicus/Sentinel satellites. In essence, understanding the characteristics of the imagery and having comprehensive information about the dataset's creation will be particularly relevant and valuable since this thesis will also utilize these satellites as a crucial source of data.

SEN1-2 and SEN12MS datasets are created by the same authors [SHQ+19b; SHZ18]. The latter is an improvement of the former. As mentioned, they both are comprised of Sentinel satellite imagery. The main differences between these two datasets:

1. Limited polarization - SEN1-2 is comprised of vertically polarized imagery. This lowers dataset complexity at the cost of possible information.

2. Limited image bands - Sentinel 2 is capable of capturing multi-spectral data which can be useful in situations when plain RGB bands fail to represent relief correctly. SEN12MS dataset already packs this piece of information while SEN1-2 does not.

3. Geo-location data - SEN1-2 does not come with any geographical location information.

This section will describe both datasets and will emphasize the differences between them in terms of dataset structure even further. The general challenges of remote sensing dataset creation will be covered here as well.

SEN1-2 is the first iteration dataset which had imagery from both versions of Copernicus Sentinel satellites: Sentinel 1 and Sentinel 2. Authors have extracted Synthetic-aperture radar (SAR) images from the first satellite and RGB channels from Sentinel 2. These restrictions were said to make dataset creation easier with the cost of information loss from the discarded portion of available bands. SEN1-2 used Google Earth Engine (GEE) as the source of images. The environment let authors easily access freely available satellite imagery. An image processing pipeline was created in order to make images clouds and shadows free. The pipeline consists of random ranges of interest (ROI) sampling, image mosaicking, exporting to GeoTiffs and multiple manual inspections after additional transformations.

Both datasets are split into 4 seasons. This method allows controlling imagery pre-processing for temporal groups that differ extensively. However, due to extreme cloudiness grouping did not fully resolve the issue and the whole process yield was significantly reduced even more after manual inspection.

During the processing step, the authors noted a few challenges. First of all, due to the implementation details of GEE image mosaicking, some images had large blank areas. The process failed to successfully find a replacement for cloudy areas thus these regions became unoccupied in the final image. Secondly, Sentinel-2 granule clouds information sometimes appeared inaccurate. The data acquired by the satellite is assigned to the whole granule rather than to a separate pixel. Essentially, this means that cloud distribution property in filtered area might not affect interested part of the relief at all or affect the chosen ROI event though the property's value is pretty low. This uncertainty resulted in much data being discarded after the manual verification.

Authors have also discovered quite an unusual challenge when comparing other datasets. They found out that some images had severely distorted colouration. It was decided to remove these images. Since they did not collect other than RGB bands information, there was nothing to compensate for such distortion. With these limitations the first dataset has been released with a promise to work on an improved version.

A year after SEN12MS [SHQ+19b], the second dataset was published. Authors have collected SAR and multi-spectral information from Sentinel 1 and 2 respectively. This time they also used MODIS (Moderate Resolution Imaging Spectroradiometer) satellite for landcover insights. Cloudiness was not an issue with the former and Sentinel 1 instruments since they solely relied on SAR (Synthetic Aperture Radar) data. However, Sentinel 2 imagery suffered from cloudy images and required further data transformation. To remove cloudy regions from data, authors used methods described in [SHQ+19a]. This paper will be used later when talking about the methodology of cloud removal. For now, it is valuable to note that pre-processing consisted of three parts: data query, quality score computation and image merging. The first part was responsible for data loading from selected ROIs and time periods. After data extraction authors needed to assign additional

information to pixel data which would later be beneficial in cloud removal. Last but not the least, the merging module used previously added data in mosaicking to merge the least cloudy images.

This pipeline helped resolving cloudiness issue. Nevertheless, there are still more challenges that other authors have bumped into.

1.4. Challenges

Dataset creation is a complex exercise which involves various aspects. A couple of them will be presented here:

First of all, a dataset must have exceptional image quality. As mentioned before, this directly influences algorithms' performance. Authors have reported such inconveniences regarding poor data quality. Images from the satellite were just enough to segment significant building changes. However, since data segmentation was performed manually, it was tough to recognize smaller relief transformations resulting in lower confidence in segmented data quality [DLB+18]. In addition to small changes, [TZZ+22] reported that certain small objects like cars highly contributed to already great interference problem. Thus, when creating a dataset smaller changes in satellite imagery should be taken with care to reduce penalty for dataset quality.

Secondly, data frequency and inconsistency between multiple sources have also put a spoke in wheels [DLB+18]. For example, Urban Atlas[8] information and maps from BD ORTHO appeared to be contradicting. This essentially lowered dataset quality. A problem was encountered when creating [CLB+19] dataset. The dataset uses European Environment Agency's (EEA) Copernicus Land Monitoring Service - Urban Atlas project resources that provide land use labels. However, EEA only guarantees 80-85% of label accuracy. On top of that, not all objects were present at the time of capturing which means that there were some images that did not depict the real use of the land. In addition to that, older satellite images were captured less frequently which makes creating a dataset a challenge [DLB+18; TZZ+22]. For example, when OpenStreetMap was used as a change reference, there were times when a change in a map was not present in the satellite image at the same time of capturing. Despite these inconveniences, there are also datasets that especially stand out for their pixel-wise annotation and frequency of captured data [TKW+22]. Examples like this should be noted, shared knowledge could help get great accuracy results.

Moving back to challenges, multiple papers also complained about data class imbalance when doing change detection in semantic segmentation setting [CLB+19; RCS+21; SHQ+19b; TKW+22; TZZ+22; WZM+21]. Some have claimed that 9 dataset labels are simply not enough to represent all areas in multi-class challenges [TZZ+22]. Others noted that the vast majority of change map content is considered as „no change“ which should be taken into account when modelling the task. Also, land use labels experienced class disproportion. For [CLB+19] most of the area labelled was from agricultural area.

Nevertheless, if a dataset is for supervised learning, it all boils down to labels. Manual label creation is a tedious job. Inaccurate labels as well as inexact borders are one of the problems that

[8] https://land.copernicus.eu/local/urban-atlas

supervised datasets encounter [CLB+19]. Not to mention the time it takes to carefully label so much data is already very expensive.

Lastly, there were also some other challenges that authors had to deal with. Challenges like shadows and clouds [KKS+21; SHQ+19a], increased saturation [TZZ+22] and distorted colours [SHQ+19b] are all problems that should be taken care. All things considered, the details pointed out in this section can be a great disclaimer of what to expect before satellite imagery analysis.

1.5. Loss functions for satellite imagery analysis

Satellite imagery change detection is a complex task which involves multiple stages of data processing. [AA19] Authors of referenced paper have referred to these stages as steps in a change detection framework. The presented pipeline consisted of data acquisition, pre-processing, running through a change detection algorithm and finally segmentation (Figure 3 for visual reference).

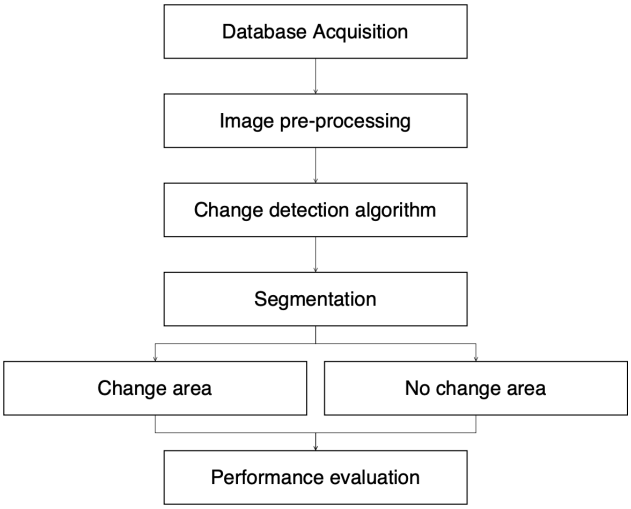


Figure 3. A framework of change detection method from [AA19]

This diagram visualizes the high-level design of the change detection segmentation process. Some tasks are a composite of multiple. For example, change detection and segmentation tasks have theoretical sub-tasks formulated that are used in machine learning as a tool to indicate the most important feature the model should learn. They also could be used to measure the model's performance.

These tasks are carefully designed to emphasize different features of provided input, such as edges, colours or shapes. There are a plethora of possible loss functions that can be used in model training. This section will focus on several use cases and loss functions that are used to accomplish acceptable results in corresponding applications.

The classification task is the first one that will be covered in this section. It is a fundamental task in machine learning where the goal is to predict a categorical label for a given input. The input could be an image, text, audio, or any other form of data, and the output could be a binary (two classes) or multi-class (more than two classes) label. The classification process involves training a model on a labelled dataset, where the model learns to identify features like patterns or colours that are relevant for making accurate predictions. Probably the most popular loss function used in machine learning

for various classification tasks - Cross entropy loss. This loss measures the difference between the predicted probability distribution of a model and the true probability distribution of the data. In the case of binary classification, this loss function can be expressed as:

$$\mathcal{L}_{\text{BCE}}(y,p) = \begin{cases} -\log(p), & y = 1 \\ -\log(1-p), & \text{otherwise} \end{cases} \tag{1}$$

Here Equation 1 $y \in \{0,1\}$ - the ground truth label, $p \in (0,1)$ - estimated label value. Binary cross-entropy loss ($\mathcal{L}_{BCE}$) can be used to express change map [BIS+23; PZG19b; RAD+20], because binary maps indicate whether that pixel has changed (binary 1) or not (binary 0) which is what BCE function compares. There is also a categorical cross-entropy loss which is used for multi-class classification problems where the output variable is categorical. It calculates the cross-entropy loss for each class and then takes the sum of these losses to obtain a single overall loss value (see below, Equation 2):

$$-\sum_{c=1}^{M} y_{o,c} \log(p_{o,c}) \tag{2}$$

here $y_{o,c}$ and $p_{o,c}$ are true output and a prediction of a class label c ($c \in M$) for observation o ($o \in N$, N - all samples).

Another cross-entropy variation that is used in classification tasks such as remote sensing change detection is weighted cross-entropy loss. This variant takes each class's importance into account when evaluating predicted labels. Each class has assigned a weight based on their importance or frequency. It is commonly used when dealing with imbalanced datasets where one class may be significantly underrepresented compared to others [BIS+23; PZG19b; ZS20]. For example, change maps are imbalanced by nature. There are much more pixels labelled as „No change“ than „Changed“. The weighted cross-entropy loss gives more weight to the minority class to ensure that the model pays equal attention to all classes during training.

Cross-entropy variations have different use cases where they can be used. The extendability and versatility aspects of such loss function make it very appealing for various tasks including remote sensing change detection. However, these are not the only losses that are used against class imbalance (Separable loss [XWJ+21], focal loss [LGG+20]).

Another loss function used for a classification task that takes class imbalance into consideration is focal loss ($\mathcal{L}_{FL}$) [LGG+20]. The key idea behind focal loss is to down-weight the loss assigned to well-classified examples and give more attention to misclassified examples. This aspect is implemented as a modulating factor $(1-p)^{\gamma}$ in the following formula (Equation 3):

$$\mathcal{L}_{FL} = -y(1-p)^{\gamma}\log(p) + (1-y)p^{\gamma}\log(1-p) \tag{3}$$

where γ is the *focusing* parameter ($\gamma \geq 0$) defined as a hyper-parameter to tune during the training or a constant to control the effect of the loss. It increases the importance of wrongly classified examples. This helps the model to focus more on hard examples, which are usually the minority class, and improve its ability to correctly classify them.

To conclude, classification involves predicting a categorical label for an entire input data or a region of interest in it. It is vital for a model to correctly pay attention to critical data features in order to decide *what* is in the data.

Nevertheless, there is a task which is concerned with identifying *where* each object is in the data space. This application is called - Semantic segmentation. Semantic segmentation involves assigning a class label to each pixel in an image. Its goal is to enable machines to understand the meaning of images at a pixel level and to identify the boundaries between different objects in the scene. Such a complex task has a unique set of losses as well as shares several from the classification domain.

For example, focal loss which was already mentioned is also used in segmentation tasks. It could be paired with a dice loss [CPY+22; SLV+17] to successfully segment changed areas in the satellite data. Even though most of the SOT models are pretty good at pinpointing major changes, it is still a challenge to identify changes in small roads or other minor variations. A model simply lacks the ability to detect minimal fluctuations in the earth's surface. As contrary as it could sound, according to the authors of RDP-Net [CPY+22] the underlying cause for this fact is that most of the loss functions are more focused on the class imbalance problems rather than boundary regions. Dice loss ($\mathcal{L}_{DL}$) is commonly used in segmentation problems. It measures the overlap between the predicted output and the actual output. This function especially helps to identify the boundaries of objects. Here is a formula for the dice loss (Equation 4):

$$\mathcal{L}_{DL} = 1 - \frac{2yp + \epsilon}{y + p + \epsilon} \tag{4}$$

where ϵ - is a small constant value that prevents division by zero on sets that do not overlap. These 2 loss functions are especially useful for segmentation tasks in general computer vision machine learning [IKW+18; SYL+23; YLX+15] as well as specific remote sensing [RAD+20; STL+20; WSH+22] domains.

One more loss function used for boundary detection is edge loss. [CPY+22] has designed a new type of loss function which was supposed to fix poor change detection performance when detecting small objects or bigger object's boundaries. The mathematical presentation of this loss function can be observed below (Equation 5):

$$\begin{aligned} w_{\text{edge}}\left(c\right) &= \alpha \left| \frac{1}{n} \sum_{k \in N} L(k) - L(c) \right| \\ \mathcal{L}_{EL}\left(p\right) &= -w_{\text{edge}} \log\left(p\right) \end{aligned} \tag{5}$$

Here $w_{\text{edge}}\left(c\right)$ is a weight of a point c. Each neighbouring label value L_k of a point k in the range N is subtracted from the real label value L_c of a regional centre point c. The sum of all subtractions is divided by the n points. Essentially, having binary label values makes this loss function act like an average point label of its neighbourhood multiplied by defined coefficient $\alpha, \alpha > 0$.

Generally, this weight function wrapped in the cross-entropy context can be used as an ideal loss function for region boundary differentiation. The loss function can also be presented visually (Figure 4).

Figure 4. Ground truth image (left) and a map of w_{edge} (right). White pixels represents the changed area. Images taken from [CPY+22]

Authors [CPY+22] have combined (Equation 6) all three mentioned loss functions (focal, dice, edge losses) to train their model and achieved state-of-the-art performance in [LVV+18] dataset:

$$\mathcal{L} = \mathcal{L}_{EL} + \mathcal{L}_{FL} + \mathcal{L}_{DL} \tag{6}$$

Both of these losses can be as helper tools to correctly segment small object boundaries. Another loss function that is used to train models and almost in every segmentation performance table as a score of precision is Jaccard loss [BSB+22]. Jaccard Loss [DVD+21] is a popular loss function used in machine learning for semantic segmentation tasks. It is also known as the intersection-over-union loss or IOU loss. Both Jaccard and dice losses are derived from corresponding similarity indexes. The jaccard index measures the similarity between two sets of data by comparing their overlapping area to their total area. Its corresponding loss function could be used to train the model to predict a segmentation map that is as close as possible to the ground truth segmentation map. The Jaccard Loss is calculated as follows:

$$\mathcal{L}_{\text{JAC}}(y, p) = 1 - \frac{(y \cdot p) + \epsilon}{(y + p - y \cdot p) + \epsilon} \tag{7}$$

In addition to focal loss, gradient profile loss is specifically designed to help segment images [BSB+22]. The main selling point of gradient profile loss (GP) is its ability to take spatial context into account. Unlike Jaccard loss, GP loss would not treat two rotated change maps with the same similarity to the original image (Figure 5)

A formulation of task can be examined below (Equation 8):

$$\begin{aligned} \mathcal{S}(y, p) &= \sum_{c} \left(\frac{1}{H} \operatorname{tr}\left(y_c \cdot p_c^{\tau}\right) + \frac{1}{W} \operatorname{tr}\left(y_c^{\tau} \cdot p_c\right) \right) \\ \mathcal{L}_{\text{GP}}(y, p) &= -\mathcal{S}(\nabla y, \nabla p) \end{aligned} \tag{8}$$

Here y_c is the ground truth and p_c - predicted value of channel c of image with dimensions $H \times W$. $tr(.)$ represents trace of a matrix and $(.)\tau$ represents transpose of a matrix. The loss function is computed by running image gradients ∇y and ∇p through a function $\mathcal{S}(y, p)$.

GP loss has shown surprisingly promising results in the context of deep learning with SAR

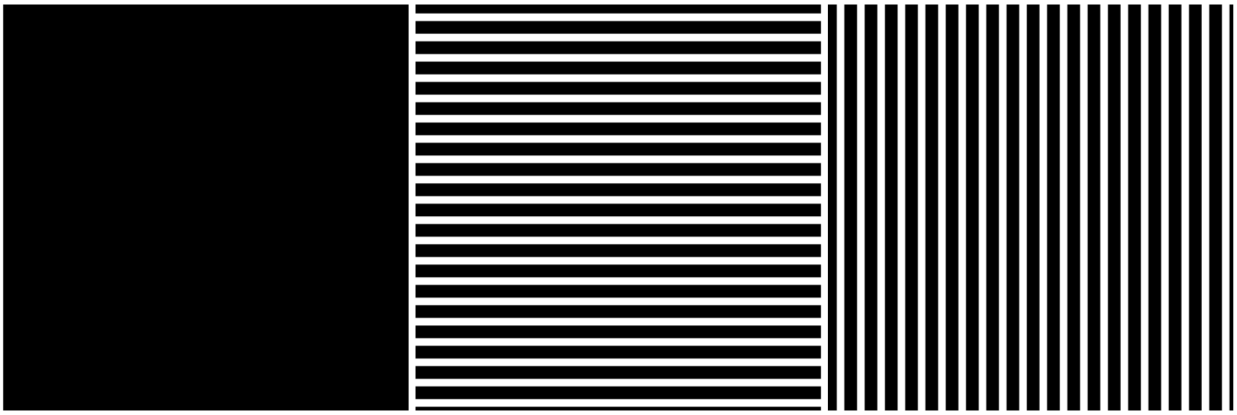


Figure 5. Visualization showing how important spatial information is when the number of pixels remains the same. Left image - source A; center image - target Bl right image - target C

images [BSB+22] and scored SOT scores in oceanography-related datasets [ZMH+22].

Lastly, there is also worth noting less adopted loss functions that are used in the remote sensing domain. Contrastive loss [NH20] is often used in the context of image retrieval, where the goal is to retrieve similar images from a large database given a query image. In this context, contrastive loss is used to train a neural network to learn a feature representation of the images that captures similarities and differences between them. The contrastive loss function encourages similar images to be mapped close to each other in the feature space, while dissimilar images are mapped further apart. This is achieved by minimizing the distance between the features of similar images and maximizing the distance between the features of dissimilar images. By using contrastive loss to train a neural network, it is possible to learn a feature representation that is well-suited for image retrieval in remote sensing applications, allowing to quickly and accurately retrieve images that are similar to a given query image.

In conclusion, this section delved into the realm of loss functions that could be used for satellite imagery analysis. By exploring various loss functions designed for both classification and segmentation tasks, one could have gained valuable insights into their respective strengths and limitations. The selection of an appropriate loss function plays a crucial role in optimizing model performance and enhancing the accuracy of change map classification. Essentially, this work will be valuable in proposed methodologies evaluation.

1.6. Visualisation techniques for different types of analysis

Selected loss functions depict tasks that a model should train for. As mentioned before, a loss could be specifically used in road infrastructure changes detection. However, these changes should also be visualised in a clear and intuitive way. This is the underlying reason why this section will glimpse on various visualisation types.

The following paragraphs will cover change detection map visualisations for the following tasks:

1. Binary change detection (climate or building)

(a) Rio in 24/04/2016. (b) Rio in 11/10/2017. (c) Ground truth.

Figure 6. (a), (b) - bi-temporal images of Rio de Janeiro; (c) - Ground truth, manually labelled binary change map between first two images. Source: [LVV+18]

2. Land cover land use (LULC) visualisations for several use cases

3. Intensity of certain properties (changes of pH scale in water bodies)

When the most important part of analysis is change, researchers usually use binary change detection maps. As covered in Section 1.5 these maps are comprised of pixels that can have only one of two values: 0 - indicating no change and 1 - change detected. Binary maps may visualize infrastructure change Figure 6.

Here we can see Rio de Janeiro city in different time spans. The difference between these images is depicted in the third image. Third image acts as a label that supervised model is trying to correctly replicate on other data. When just the change is not enough, there are other, more information packing methods - semantic segmentation.

Semantic segmentation in remote sensing refers to the identification and delineation of specific objects or areas of interest in satellite images. This technique plays a crucial role in numerous applications, including urban planning [TZZ+22], environmental monitoring [KKS+21] and disaster management [WSH+22]. The following image could help visualizing the output image model must correctly generate Figure 7.

Different colours denote different land use type. Models could potentially indicate not only the type itself but changes to each class as well. There is a third use case of a map where the feature indicates the existence of a certain object's property in the image.

This is a perfect example of a scenario where one is trying to determine whether a water body is blooming. This is solved by managing chlorophyll concentration [GP]. The image below demonstrates how specific properties of the object become more pronounced as certain data bands are increased.

Overall, it is possible to visualise data in different methods. Some methods indicate a change by lighting up the pixel in the area where the change is and showing pitch black for pixels that have not changed. Other methods classify areas into blobs that mean different classes of land use. Lastly, by increasing the intensity of certain bands, it is possible to identify objects' properties. This helps in environmental issues, urban infrastructure or disaster analysis.

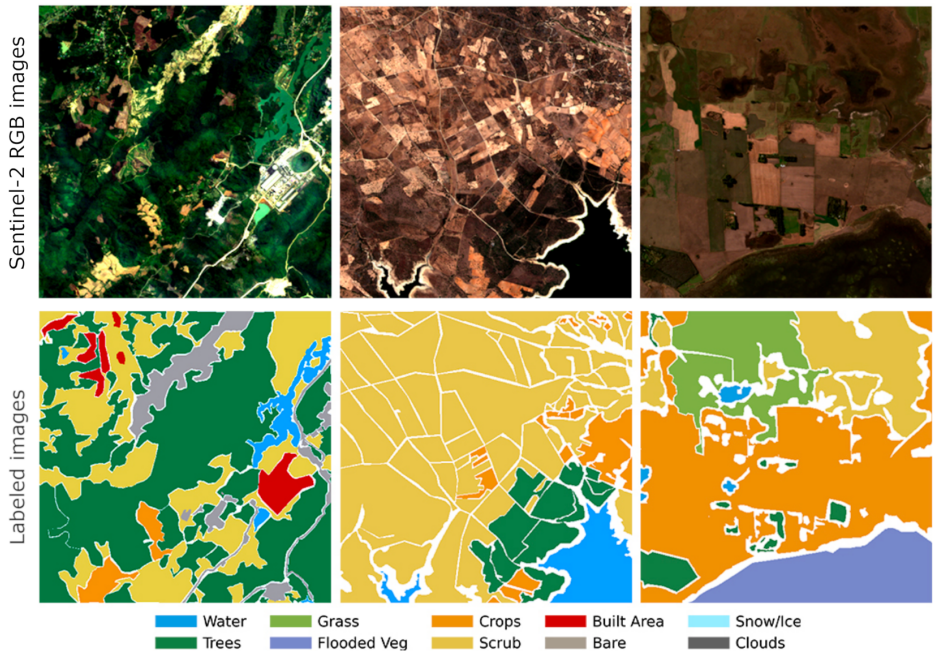


Figure 7. Image (top) and target (bottom) pairs of Land cover land use dataset [KKS+21].

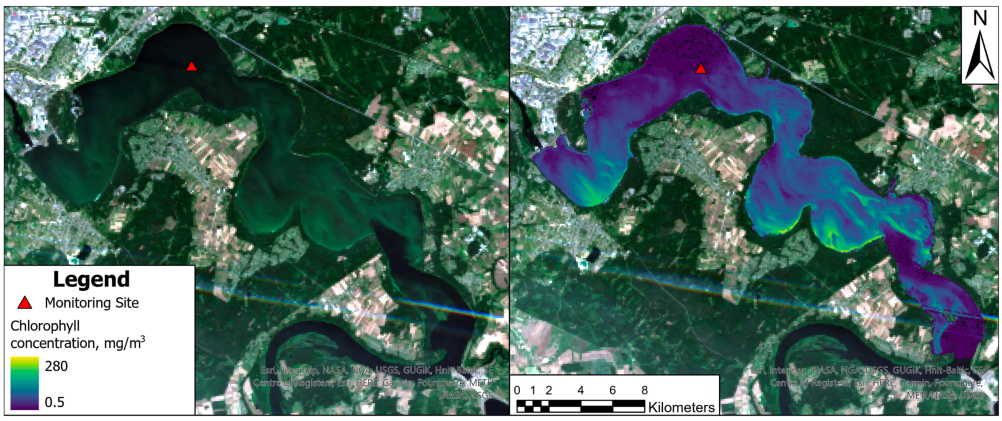


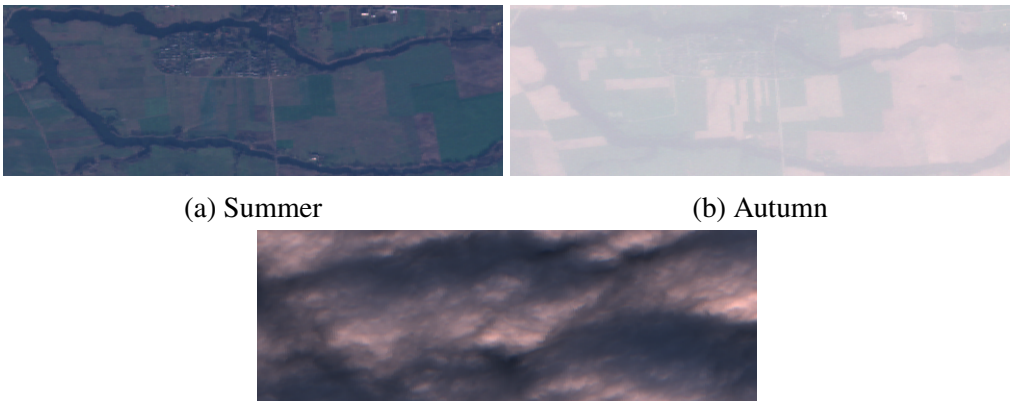
Figure 8. A pair of images taken from [GP]. Sample image of a water body on the left and chlorophyll concentration visualisation made with the help of other bands on the right.

2. Data gathering methods

To extract data that could be used as a dataset, it was decided to conduct several analyses that would give an insight into data extraction process implementation. This section will describe these in detail. Firstly, there will be introduced a way to extract valuable information from images that are affected by common obstacles. After that extracted data will be tested to decide whether it is usable in any training setting and how well it performs with a selected state-of-art model in a change detection setting without any pretraining. This will also suggest whether the results of inferred change maps are useable in manual labelling procedure.

2.1. Satellite imagery quality evaluation

As already mentioned, in previous sections the creation of a dataset requires deep analysis of the gathered data. Satellite imagery tends to be contaminated by various obstacles like clouds and reflections. These occlusions worsen the quality or even make it unusable. Data quality is a crucial point when creating a new dataset, so it needs to be examined before use. For example, Figure 9 shows a couple of lake Voveriai photos. These photos are an example of what an image could look like. However, it is rather unusual to think that an image in Figure 9c could provide any useful information. In this section, a data extraction method will be presented. Extracted images will vary in quality but with the help of further analysis, it will be hopefully clear whether occlusion-affected images weigh any valuable information from other bands.



(a) Summer (b) Autumn

(c) Winter

Figure 9. Lake Voveriai in different times of the year. These images were exported raw to show exactly what problems are active in geospatial data harvesting. Images do not depict the whole season they're assigned to.

Even though cloudiness is a general problem in satellite imagery analysis, it is especially active in regions that have all 4 seasons like Lithuania - the country in which data is gathered. To gather as much usable information as possible 5 lakes in Lithuania were selected for deeper analysis Figure 10.

These 5 water bodies were selected to encompass 5 possibly different weather areas. Even though, Lithuania is a rather small country, Baltic Sea influence might correlate with some weather conditions.

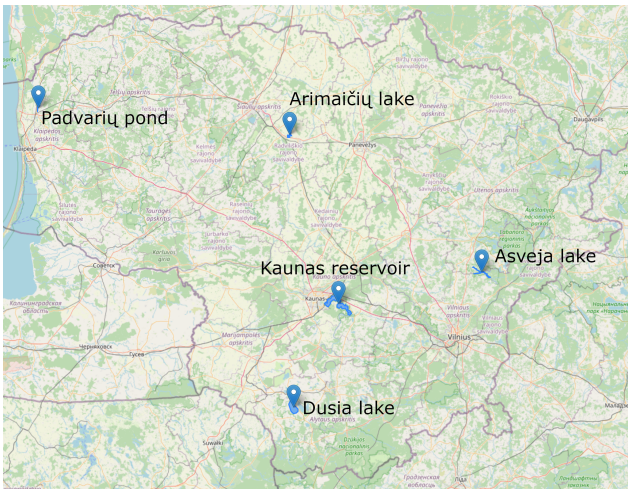


Figure 10. Lithuanian water bodies picked for analysis. 5 water bodies encompassing 5 varying in weather areas: pond Padvariai - north-west side, lake Arimaičiai - north side, Kaunas reservoir - centre side, lake Asveja - west side, lake Dusia - south side.

Although there is no quick and easy way to delete a cloud from an image and it would be more beneficial to remove those images entirely, this scenario raises other problems. Data in one region is limited and removing part of it is extremely costly. The five water bodies analysis could potentially answer whether cloudy images bring any possibly useful information in other bands. They may not seem useful with a naked eye (Reg-Gree-Blue or RGB band images) but it is still information that models can use to determine surfaces. It is also worth noting that different bands react to different reliefs which could aid in water distinction.

This experiment continued by selecting two points on the map: one from the area in the lake and another one outside of it (see Appendix Nr. 1). Geospatial data from these pixels were collected and plotted in pixel intensity graphs. Two pixels' intensities from each of the lakes were visualised to see the distribution of it throughout the years. Even though the pixel does not represent the whole picture, it can still suggest the alteration frequency and magnitude of each band at that point during different times.

Since there are too many images, this analysis will only emphasize a few valuable mentions with appropriate examples. These examples will use B12, a short-wave infrared band only. For example, analysing the closest and furthest water body from the Baltic Sea Appendix Nr. 2 shows that there is no meaningful difference whether the object has a wetter climate or dryer. This means every decision from this point, can be applied to every selected water body.

What is also worth mentioning is that since there is a slight difference between intensities in the lake and outside the lake, band data should be pre-processed individually by setting proportional properties. Also, the difference between in and outside lake data could be because of the natural band sensitivities towards different relief types. For example B12 band is used in combination with other bands to emphasise vegetation which would explain the increase of intensity in outside water body analysis.

Another point worth noting is that the most monotonic intensities were taken during the warm seasons. Sadly, this finding might indicate the harsh truth that off-season data have too much noise

to be valuable.

Considering all factors, the results raise the question: how to use this information effectively? Mostly all of the band's data has times when pixel intensities outlied the average. These values might be an indication of an unstable signal or weather obstacles. However, the data between such periods is perfectly fine. To separate the wheat from the chaff, a running median SMM was introduced as a threshold. The formula below was used:

$$M = \begin{cases} X'_{\frac{n+1}{2}} & \text{if } n \text{ is odd} \\ \frac{X'_{\frac{n}{2}} + X'_{\frac{n}{2}+1}}{2} & \text{if } n \text{ is even} \end{cases}, \text{where } \{X'_1 \leq X'_2 \leq \ldots \leq X'_n\}$$

$$SMM_i = M(\{X_{i-k+1}, X_{i-k+2}, \ldots, X_i\})$$

Every value X_n where n is the nth sequence value $X_1, X_2, \ldots, X_N$ and $n \in N$ that is above the median at a specific time window $k = 20$ is deleted. This way unusable data can be filtered while retaining more off-season data. The distribution of such data and moving medians can be observed in Figure 11 and Appendix Nr. 3

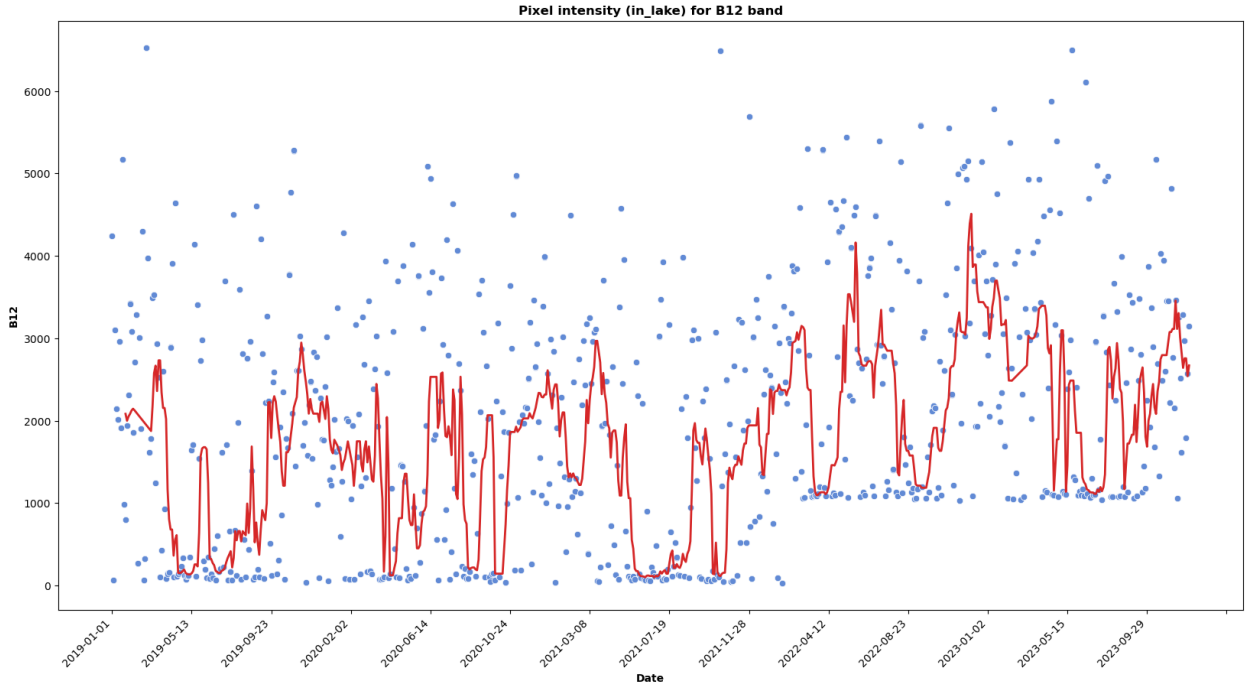


Figure 11. Running median distribution of pond Padvariai (pixel inside the lake), 2019-2013

However, individual plots do not contain much meaningful information. It was decided to accumulate different band findings by counting how many of them were outside the error bounds. The error bounds were defined by taking the moving median as the middle line and 20% from each side of the median as the upper and lower bounds. Each band had its error range and outliers were counted. For example, an image in the collection that has 8 bands that are outside the error range was assigned 8 to the count property. This way a histogram was created (refer to Figure 12) that allows guessing which images bring more information.

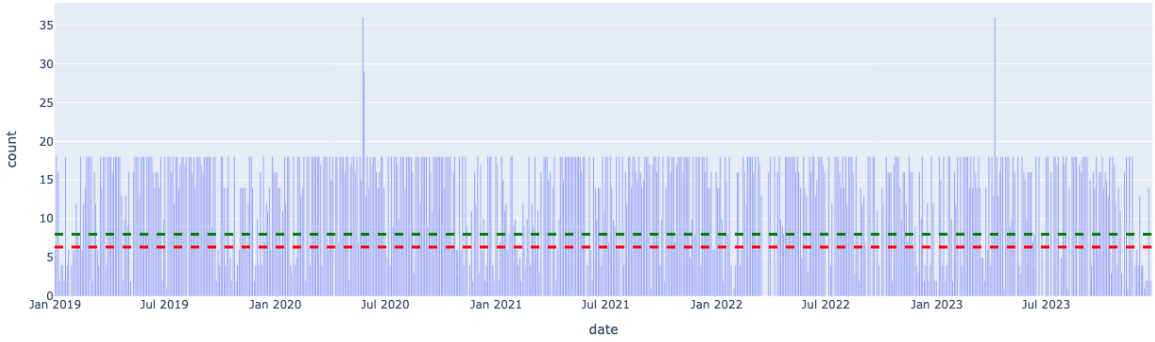


Figure 12. Outlier bands count distribution throughout the whole image collection (2019-2023). Red line - average, green - median

The calculated median was used as a threshold value that helps to determine whether the image is good for consumption. Finally, these calculations were automated to make the process easily integrated when used in other location analyses.

2.2. Data extraction and applicability

Once data harvesting methods are acquired, data extraction is the next step. It is vital to test how well previously described methodologies work with real data. With the help of further experiments, it will be possible to weigh several hypotheses:

1. Data quality - whether downloaded images can be used with existing methods or models.
2. Automatic labelling possibilities - even though manual labelling is said to be the standard way to get ground truth images, automatic labelling would aid this process by providing at least the base of labels. It could also possibly label images whose RGB bands are affected by clouds.

To test these hypotheses, 3 different areas were selected:

1. Visoriai/Bajorai - quickly emerging Vilnius district is suitable for urban change detection.
2. Construction site/ruins of the National stadium in Vilnius - recently resumed site constructions are a perfect opportunity to detect changes.
3. The annual floods of river Šyša - the river is commonly known to flood during the winter. This is a perfect opportunity to identify changes in the water body.

Sentinel hub[9] was used for image exploration. Sentinel Hub is a satellite data processing engine, that facilitates convenient access to Sentinel and Landsat Earth observations. All 4 areas were explored throughout the time interval 2019-2023. For easier insights, only the cloudless images

[9]https://www.sentinel-hub.com/

were taken. It was easier to determine whether these methodologies work at all. Only then they could be tested with lower-quality data. Selected images can be observed in Appendix Nr. 4.

All images were downloaded as GeoTiff files in EPSG:3857 projection, each having 13 bands. They were run through the Fully Convolutional Early Fusion (FC-EF) [DLB18b] model to acquire change maps. The original authors' code change map generation was too strict for such testing. In the original code, the model returns two-channel (change and not changed) tensor with logits in it. The code creates a one-channel matrix in which each pixel is assigned a value of either 0 or 1. This value indicates whether the highest value across two channels occurred in channel 0 or channel 1, respectively. In most cases, the examples returned empty change maps. To address such instances, the output channel was passed through a sigmoid function. Since the output became probability for a pixel to be marked as „changed", it was easy to apply thresholds to make change maps that vary in sensitivity.

Running data through the model yielded quite interesting results. Change maps can be seen in Figure 13.

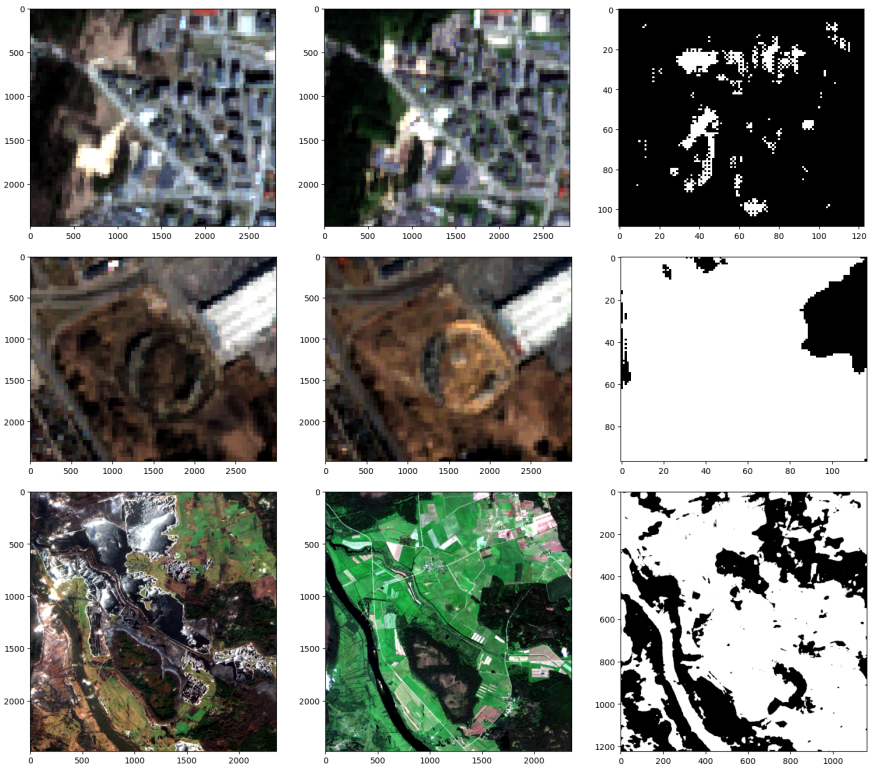


Figure 13. Spatial temporal image pairs with predicted change maps of Bajorai district (1), National stadium construction site (2) and river Šyša (3).

All results considered, it appears that the model has different sensitivity towards change detection in different areas. For example, the Bajorai district change prediction threshold is 0.499, national stadium - 0.488, and river Šyša - 0.475. While these thresholds do not vary too much, their change percentage distribution makes it impossible to determine one size fit for all procedures. Mostly all of the 3 change maps do not have a higher prediction than 0.5 for a changed pixel but according Figure 14 Bajorai district and National stadium construction distribution is tighter than Šyša river's).

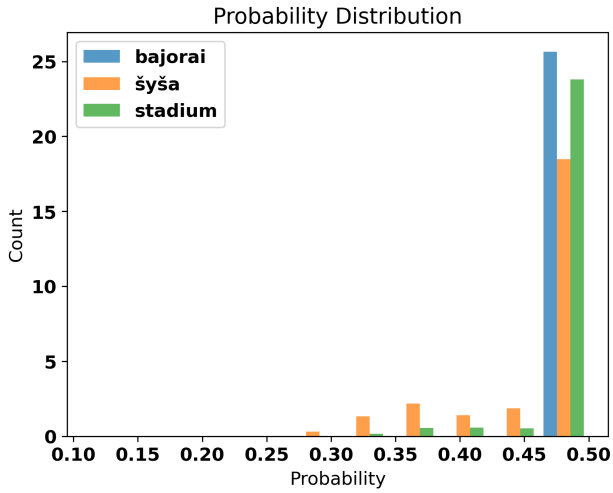


Figure 14. Probability distribution of 3 images

Worth noting that this histogram is only to visualise pixel value distribution. Lower probability does not infer the quality of prediction. However, if the distribution is wider like for Šyša river, threshold setting requires manual interference. This scenario is not ideal when making a huge dataset and does not make work easier.

To sum up, the results of analysed images suggest that dataset creation will require more manual work than one would want. To conclude properly, conclusions will be split into two parts, each explaining how well each hypothesis stood against the results.

Regarding data quality evaluations, images that cannot be analysed from their RGB values due to bad weather could only be used for unsupervised or semi-supervised training. While this analysis did not conduct any performance testing to find out whether such approach works, the written code can conveniently select such images that hypothetically could enhance the model's performance. Regarding supervised training, it will not benefit from such data addition mainly because those images lack labels and labelling solutions that were tested do not yield anything robust. Proper change maps require additional tinkering with configuration which is very tedious if not impossible in images that do not have visible RGB bands.

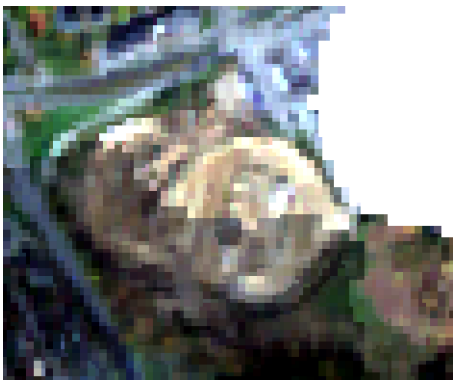
Speaking of automatic labelling, the second hypothesis could be half-proved. While the methods do not make the whole process automatic, output change maps could be a foundation for manual work. If integrated with any labelling tool, such platform could be highly efficient.

2.3. Mosaicking - worth mentioning data gathering method

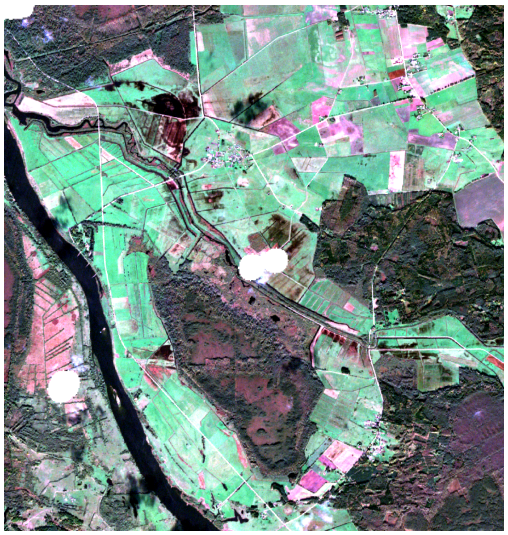
During times when satellite images are of very poor quality, using them one by one is not practical. However, having multiple of such images most of the time means that it is possible to get a single image from all of them. It is done by using the best values across all images for each pixel. This process is called mosaicking. Google Earth Engine provides functionality to convert image collection into a single image. However, with the help of a little custom logic, it is possible to get even better results. A couple of examples are depicted in Figure 15.

This method can also be used in data extraction. While it is true that mosaicking reduces the

(a) Bajorai district, 2023-11-01 - 2024-01-01 60% CPF, 40% CPT



(b) National stadium, 2023-11-01 - 2024-01-02, 80% CPF, 40 % CPT



(c) Šyša river, 2023-10-01 - 2024-01-01, 30% CPF, 40% CPT

Figure 15. Mosaicked image examples. CPF - cloud probability filter, CPT - cloud probability threshold. White spots are areas that have got excluded by the cloud probability threshold.

amount of data, a better quality image that is gathered from a reasonable time interval might have a bigger impact while training models. Carefully created mosaicked images will have both: a good quality in all bands and spatial-temporal changes - a model objective.

3. Using feature maps as model's input

Previous sections reasoned about different ways to get new data that could be used in remote sensing model development. Another possible way to increase model performance is to use another model's output and adapt its knowledge. The main hypothesis of this idea states that existing model's output contains concentrated features about the same data. Naturally it raises the question whether such interpretation of the same input could be used in tandem with original data to increase problem solving performance.

At the time of writing, FC-Siam-Diff was considered as the SOT model in OSCD dataset benchmark. The model achieves 57.92% F1 score and 51.84% precision [10]. Even the second-place model is from the same Fully convolutional model family. This thesis aims to enhance the SOT model's performance by adding additional information for the training process. Recently, there was a model introduced by IBM and NASA - Prithvi [JCF+23]. The new vision transformer based foundational model was built using first-of-its-kind segmentation training framework for multispectral satellite imagery. More than 1TB of imagery was used in order to achieve promising results in spatial-temporal footage segmentation.

This new foundational model was used to extract useful information from existing OSCD images. This strategy works only if the following hypothesis stands correct: latent information from images has valuable information that can be compared and their difference evaluated. An analysis was conducted to prove this. Latent features were compared using multiple comparison metrics:

1. **L1** similarity, also known as Manhattan distance, calculates the absolute difference between corresponding elements of two vectors (X_1 and X_2): $\sum_{i=1}^{n} |X_{1_i} - X_{2_i}|$.
2. **L2** similarity, also called Euclidean distance, computes the square root of the sum of squared differences between corresponding elements of two vectors (X_1 and X_2): $\sqrt{\sum_{i=1}^{n} (X_{1_i} - X_{2_i})^2}$.
3. **L-Inf** similarity calculates the maximum absolute difference between corresponding elements of two vectors (X_1 and X_2): $\max_i |X_{1_i} - X_{2_i}|$.
4. **Cosine similarity** measures the cosine of the angle between two vectors (X_1 and X_2), indicating their similarity regardless of their magnitude: $\frac{\sum_{i=1}^{n} X_{1_i} \cdot X_{2_i}}{\sqrt{\sum_{i=1}^{n} (X_{1_i})^2} \cdot \sqrt{\sum_{i=1}^{n} (X_{2_i})^2}}$.
5. **SSIM (Structural Similarity Index Measure)** assesses the similarity between two images (X and Y) by considering luminance, contrast, and structure: $\text{SSIM}(X, Y) = \frac{(2\mu_X\mu_Y + c_1)(2\sigma_{XY} + c_2)}{(\mu_X^2 + \mu_Y^2 + c_1)(\sigma_X^2 + \sigma_Y^2 + c_2)}$, where μ_X and μ_Y are the means of X and Y, σ_X^2 and σ_Y^2 are the variances of X and Y, σ_{XY} is the covariance of X and Y, and c_1 and c_2 are constants to stabilize the division.

[10] Leaderboards can be found here

6. **Multi-scale SSIM (MS-SSIM)** extends SSIM to multiple scales for improved performance: $\text{MS-SSIM}(X, Y) = \frac{1}{L} \sum_{i=1}^{L} \text{SSIM}(X_i, Y_i)$, where X_i and Y_i are the decomposed images at scale i, and L is the total number of scales.

A random 900 samples were picked to have a reasonable sample size and each image pair was run through the model. The outputs were then compared with all metrics and written to file. Images below (Figure 16) show normalised data points.

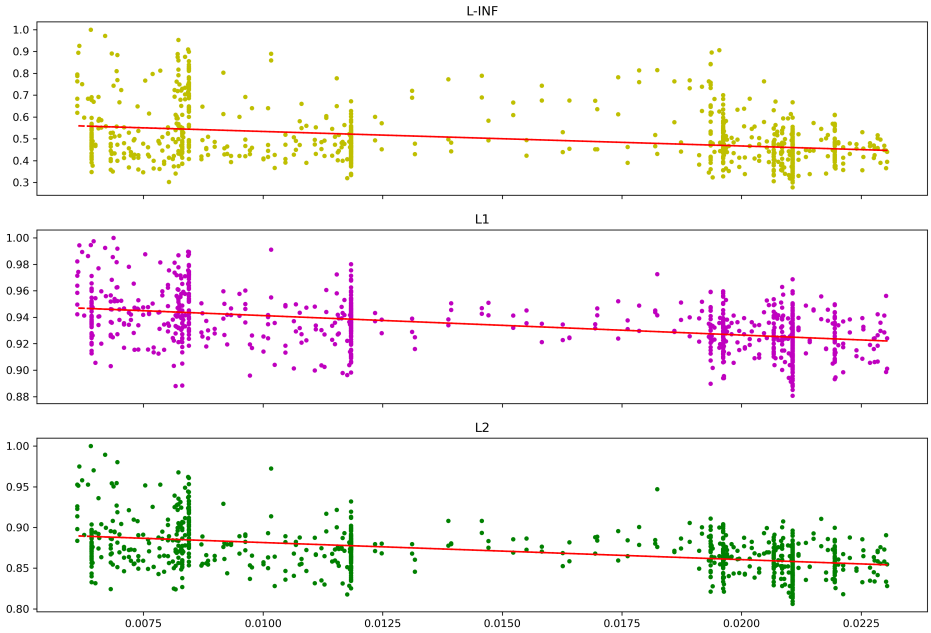


Figure 16. Changed/Unchanged pixels ration - X axis, similarity score - Y axis. Red line - scatter plot trend line.

The X axis shows the ratio between changed and unchanged pixels and the Y axis shows similarity. If the metric correctly depicts the similarity between feature maps, data pattern should follow a descending trend. This would indicate that the more change there is (bigger ration between changed/unchangend pixels), the smaller similarity score this pair scores.

As seen in the figure, scores from all three metrics follow the same trend and can be used to prove feature maps similarity. Aside these three, other metrics were also calculated. To find other methods that correctly interpret given feature maps, Cosine similarity, Structural similarity index and Multi-scale structural similarity index measures were tested. The result can be found below (Figure 17)

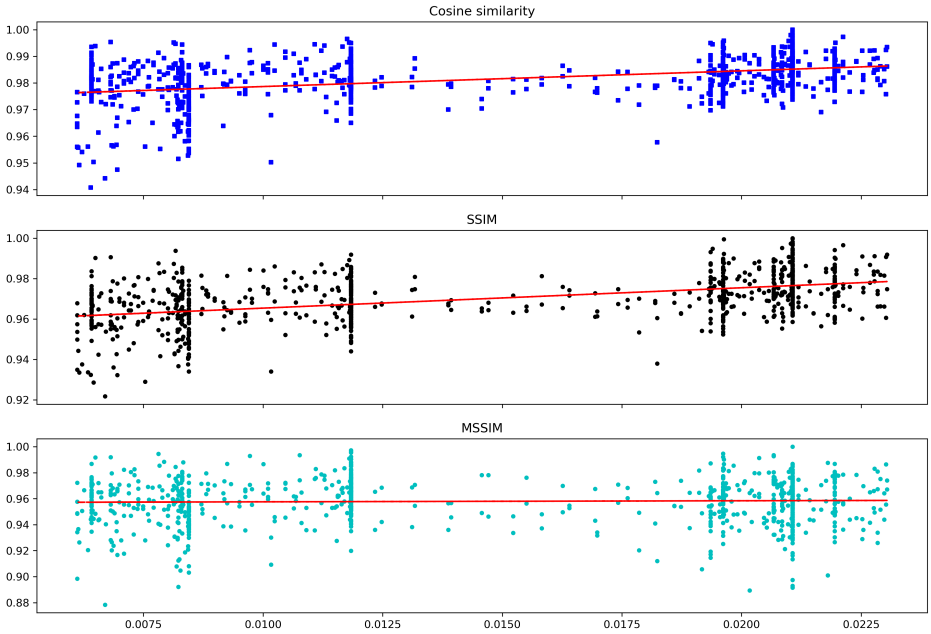


Figure 17. Changed/Unchanged pixels ration - X axis, similarity score - Y axis. Red line - scatter plot trend line.

However, the latter three metrics did not convey any useful information about feature map similarity.

Even though some metrics did not visualise correct correlation between feature map pair closeness and actual changes, first three metrics can be confidently used to prove the hypothesis. In conclusion, feature maps can be used to extract useful information about relation between two images. This data from foundational model Prithvi can be used as additional input for FC-Siam Diff model.

4. Model development

This work presents a new model with customised architecture and a new input. The input consists of tensor of shape $B \times 38400$. These feature maps are taken from the middle of Prithvi model (Figure 18)

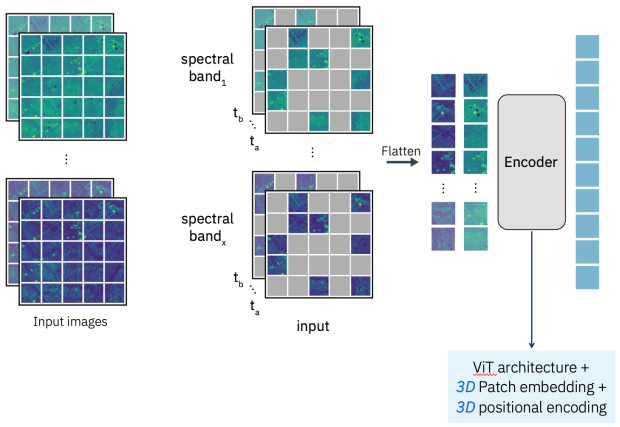


Figure 18. Prithvi model architecture encoder part that was integrated into the custom model architecture

Essentially, this output from the model is latent representation of the image. In the original pipeline this representation was used in decoder to recostruct masked images. This means that model was trained to make exclusively informational latent representations to make decoder work possible. The similarity information between bi-temporal images was reasoned in Section 3 and this section will present further details regarding model development.

4.1. Model structure changes

As seen in previous section (Section 3), feature maps are proven to contain change information. The next step is to integrate this new kind of input into the existing pipeline. An existing model FC-Siam-Diff (Figure 19) was used as a base model. Aside from existing base inputs a new 2 feature maps were injected into the model.

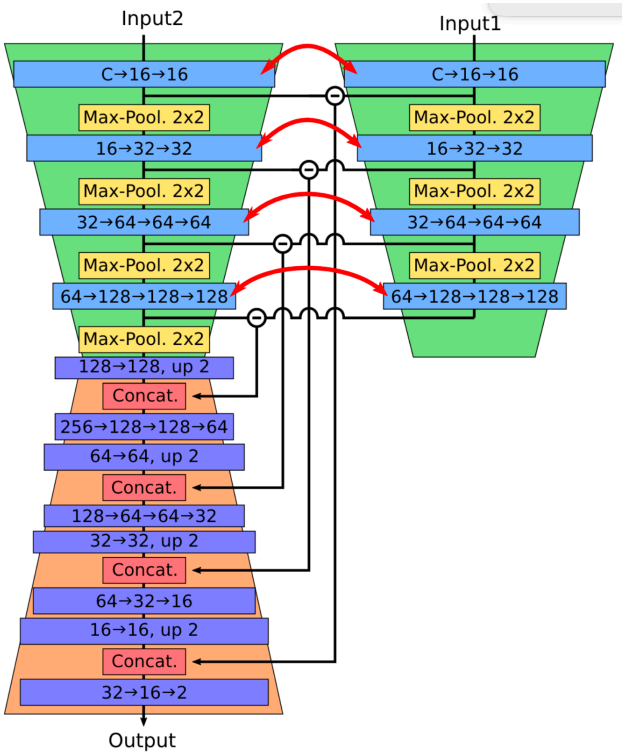


Figure 19. Fully Convolutional Siamese - Difference (FC-Siam-diff) model. This was taken as a base architecture of an improved model.

Outputs of Prithvi model are already in encoded form so it is safe to include them right into the second stage of the model - the decoder part. However, the decoder had to adapt to accept more data (Figure 20).

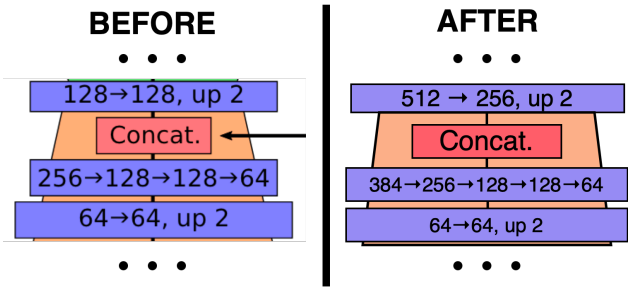


Figure 20. Part of the updated model architecture to allow to accept additional inputs.

To match existing dimensions inputs were padded with zeros to the shape of $B \times 128 \times 28 \times 28$. Features then were reshaped into the correct shape of the inputs from the first stage. Then both of them were concatenated with output from encoder. Resulted tensor was run through the first layer of decoder. The only difference was to accept tensor of size 512 and reduce it to 256. The last layer followed the same strategy as in the original architecture - gradually reduce the shape of the data.

The original FC-Siam-Diff model was customised to accept additional 2 inputs - Prithvi model's output features maps of the same bi-temporal images. This could potentially increase model's performance. Next sections will analyse customised model's performance.

5. Performance evaluation

A developed model has to be tested against various data to identify the quality of development. Later subsections will present 3 test cases:

1. Testing against original dataset (OSCD) (Section 5.1).
2. Testing against new external dataset (MSBC) (Section 5.2).
3. Testing against custom Lithuanian data (Section 5.3).

5.1. OSCD dataset testing

Firstly, the model was tested with the original dataset test subset. OSCD dataset has predefined test case consisting of 7 cities: Montpellier, Norcia, Saclay, Valencia, Dubai, Milano and Chongqing. These multispectral images were cropped to appropriate size that the new custom model accepts (224×224) and run through the model in evaluation mode. The following results were outputted (Table 2):

	Accuracy	F1 score	IoU	Precision	Recall
Custom model	0.95	0.31	0.19	0.53	0.31

Table 2. Model's performance on OSCD test data

As can be seen from the table, results are not high. Model performance does not reach FC-Siam-Diff model F1 score of 57.92 (or 0.5792 if on the same scale). However, custom model's precision surpassed original predecessor's 51.84 (0.5184) score. This model also showed interesting ROC data (Figure 21).

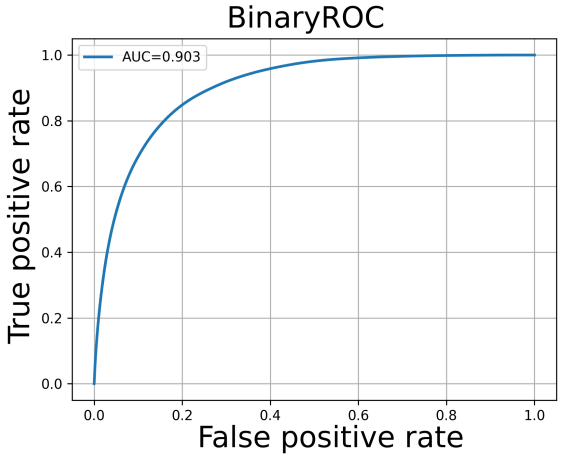


Figure 21. Model's ROC on test data

The Receiver Operating Characteristic (ROC) curve rises sharply towards the top left corner of the plot. This and high Area Under the Curve (AUC) score indicates low false positives rate and high

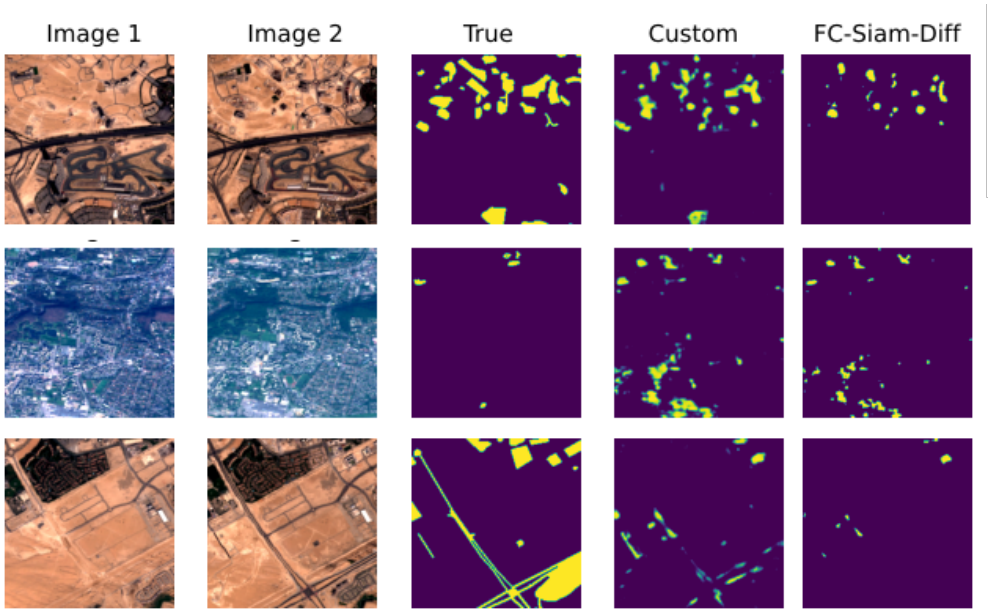


Figure 22. Sample model output against OSCD dataset

true positives count. However, low F1 and Intersection over Union (IoU) scores suggest the contrary. AUC metric may be misleading due to a likely class imbalance or the model's overemphasis on the majority class. This is a perfect example, how crucial it is to choose correct evaluation tools to be able to make correct conclusions.

To correctly conclude model performance evaluation against OSCD dataset a couple of examples are provided in the Figure 22. As one can observe, the first example could be considered as the best from all of 3. Model correctly acknowledges most notable boxes as changes, although failing to segment regions fully. Other examples are worse. The second example shows how model creates non-existing changes. The third one suggests that model also struggles to identify narrow lines (roads).

The performance testing has shown model's flaws and spots where model could be improved. Nonetheless, this performance is not extraordinary or not expected taking into account a slightly worse than original model's metrics. In fact, original model is also failing to identify many different regions and performing worse in those 3 provided examples. Thus, custom model performance results might be considered as better.

5.2. External dataset testing

To test the model correctly, another evaluation was conducted. An external MSBC (Multisource Built-up change) [LZZ+22] multispectral imagery dataset was picked to test image on the same domain but different location. This dataset consists of Sentinel-1 and Sentinel-2 images of Guigang City, China. Images from the former satellite have two bands (VV and VH), from latter - 7 bands (B2, B3, B4, B8, B8A, B11, B12). Time interval spans 2 years and reaches spatial resolution of 2 m. Sample from original paper is provided below in Figure 23.

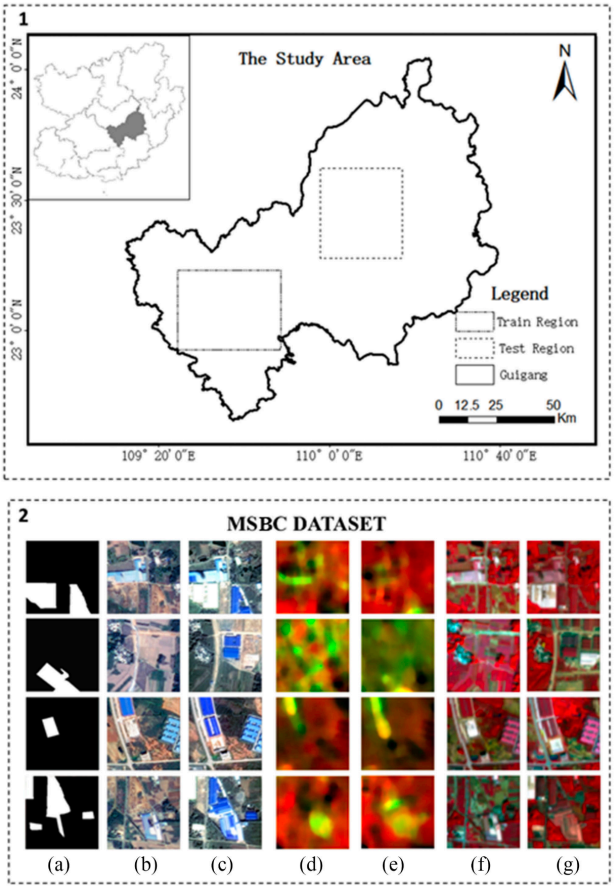


Figure 23. MSBC dataset. (a) Ground truth. (b) Images in time1. (c) Images in time2. (d) SAR in time1. (e) SAR in time2. (f) Multispectral data in time1. (g) Multispectral data in time2.

Even though, images from MSBC dataset are multispectral, the number of bands is not sufficient. Both, original and custom models take input of 13 channels. MSBC dataset lacks infrared (B5, B6, B7), aerosols (B1), water vapor (B9) and SWIR bands (B11, B12). Substitution for missing bands does not exist however, it is possible to assign zeros to the missing bands. This approach does not interfere with inference and provide correct results.

To test the dataset, a new Pytorch Dataset module was developed and run through both FC-Siam-Diff and Custom models. Results can be seen in Table 3

	FC-Siam-Diff	Custom
Accuracy	0.93	0.94
F1 score	0.35	0.37
IoU	0.22	0.24
Precision	0.43	0.43
Recall	0.35	0.41

Table 3. Model performance with MSBC dataset

The Custom model outperforms FC-Siam-Diff in most metrics, demonstrating a higher accuracy. It also shows an improved F1 score (0.37 vs. 0.35), better IoU (0.24 vs 0.22) and recall (0.41

vs 0.35). Precision remains the same for both models at 0.43.

Overall, the Custom model provides a minute but notable enhancement in performance over FC-Siam-Diff. This dataset is one of a few performance metrics where custom model shows positive improvement indications.

5.3. Manual testing

To test the model even further, another test was conducted. Model was tested with Lithuanian satellite imagery.

50 random points from Lithuania were selected (Figure 24).

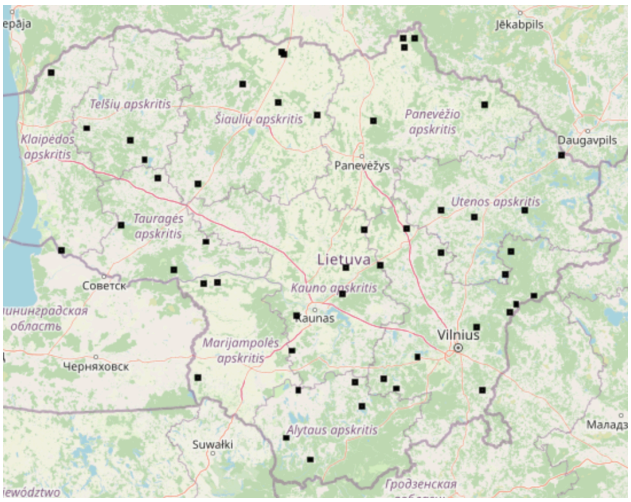


Figure 24. Random selected points from Lithuania

Firstly, Google Earth Engine was used to select the all images below 10% cloudiness. To ensure the best quality of downloaded images, the collection was filtered to only collect images from 2020 to 2023 and only from May till October. All downloaded points were $1km^2$ squares in 224×224 dimensions. To ensure correct projection, *EPSG:3857* system was used.

These bi-temporal images were used to understand what changes a model sees throughout the time. An individualised threshold then filtered change map from predictions below a certain probability. Filtered values were then averaged to get so-called „confidence“ score. This score allowed to sort change maps and get images with the highest predictions. Top 20 of these were selected and then manually reviewed for proper performance examples.

A few best examples will be presented here to reason about model's performance. All examples are named as Points A to D (Figure 32, Figure 33, Figure 34, Figure 35, condensed view - Figure 25).

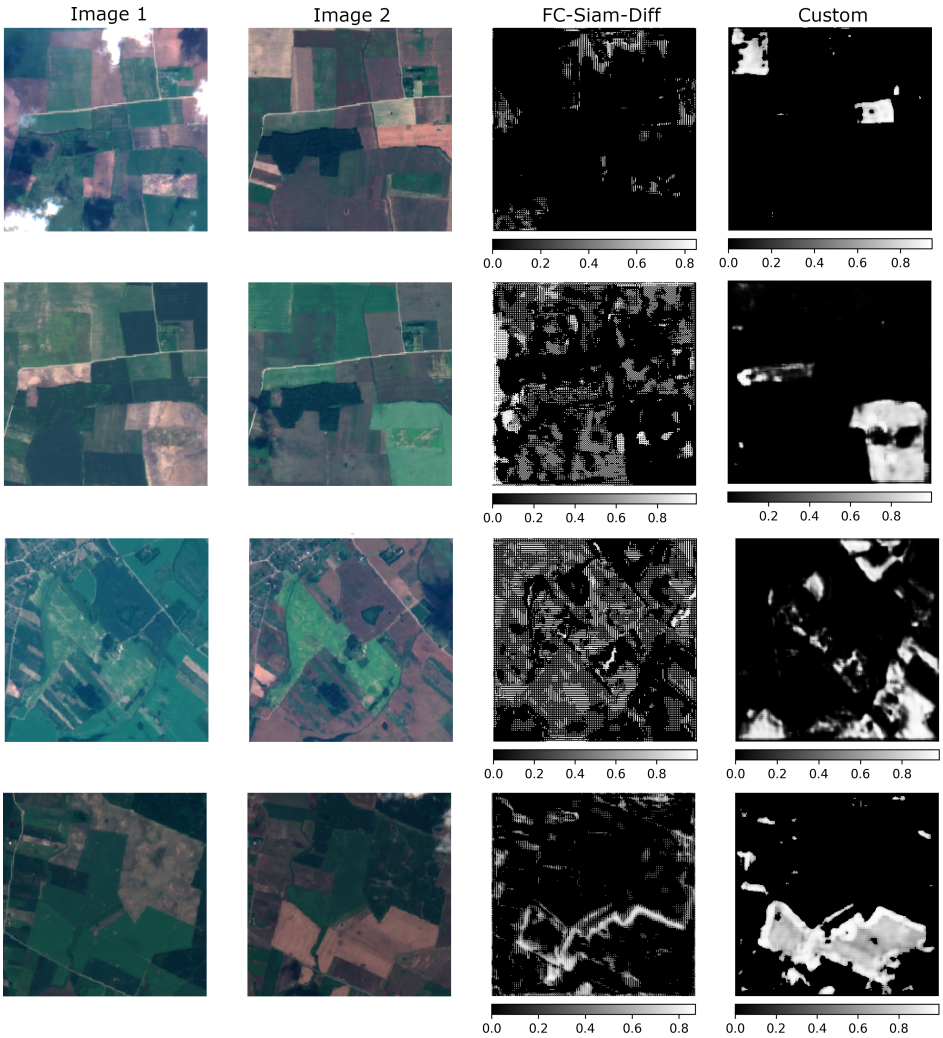


Figure 25. Example points. These locations are referred as points from A to D, top to bottom respectively

Looking at the first four examples (Figure 32, Figure 33, Figure 34, Figure 35), one can see that images contain agricultural areas. Both models fail to see every change. However, it is clearly visible that FC-Siam-Diff model fails to indicate any meaningful change. No matter how high the threshold is set, FC-Siam-Diff model shows most of the image as ”slight” change: the whole image is gray and does not have regions where probability is high enough to be filtered out by the threshold.

Another point worth mentioning, first image has clouds in the image 1. What is interesting, the custom model sees this as interference and does not visualise it as change. This does not apply to all images that have clouds. Both models are vulnerable to bigger cloud concentration in the image.

Speaking of the last triplet (Figure 35) example, both models indicate a clear change. A big spot in the bottom of the image region has changed and became an agricultural area. However, FC-Siam-Diff does not fill that region and highlights the edges only.

This and 3 other examples show that there are images where the custom model outperforms FC-Siam-Diff. It by no means allows for making generalised conclusions or proving that the custom model is superior across the entire Lithuanian region. Results from manual performance testing indicates potential improvements of land-use, land-cover change detection.

Results and conclusions

The thesis concludes with the results:

1. Developed median analysis methodology and tested against spatial-temporal data from 3 different areas. Results has shown that the new data quality is only for training in unsupervised or semi-supervised settings.
2. Proposed idea to use the output from Prithvi model as another model's input. This could compensate domain's data deficit.
3. Developed a neural network model that accepts 4 inputs: 2 bi-temporal images and 2 feature maps
4. Tested proposed model against OSCD test set with accuracy - 0.95, F1 - 0.31 and precision - 0.53.
5. Tested proposed model against separate MSBC dataset with accuracy - 0.94, F1 - 0.37 and precision - 0.43.
6. Tested proposed model against randomly gathered Lithuanian regions. Results have shown a more precise change regions when comparing with base model.

Results allowed to reach certain conclusions:

1. Analysed data could be transformed into additional unsupervised or semi-supervised dataset. This also applies to images with lower visual bands quality.
2. Even though supervised learning would not benefit from the new data without manual labeling, knowledge transfer was proven to be another viable alternative.
3. Improved model architecture allowed the proposed model to consume the most notable concentrated input information and potentially help to identify changes.
4. The proposed methods demonstrated better performance on OSCD dataset and could be useful in remote sensing change detection tasks instead of the base model.
5. Model's testing against MSBC dataset has shown improvements even on totaly unreleated dataset. This shows that model did not overfit and is capable to predict unseen data equally well.
6. Testing against manually downloaded satellite imagery has demonstrated increased performance of proposed model. These results by no means allow making generalised conclusions or proving that the custom model is superior across the entire Lithuanian region. However, the proposed methodology could be reasoned to be related with the increase in selected imagery land-cover, land-use domains.

References

[22] Sentinel-2 overview, 2022. URL: https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-2/overview.

[AA19] Anju Asokan and J. Anitha. Change detection techniques for remote sensing applications: a survey. en. *Earth Science Informatics*, 12(2):143–160, 2019-06. ISSN: 1865-0473, 1865-0481. DOI: 10.1007/s12145-019-00380-5. URL: http://link.springer.com/10.1007/s12145-019-00380-5 (visited on 2023-04-29).

[BIS+23] Joshua Billson, Md Samiul Islam, Xinyao Sun, and Irene Cheng. Water Body Extraction from Sentinel-2 Imagery with Deep Convolutional Networks and Pixelwise Category Transplantation. en. *Remote Sensing*, 15(5):1253, 2023-02. ISSN: 2072-4292. DOI: 10.3390/rs15051253. URL: https://www.mdpi.com/2072-4292/15/5/1253 (visited on 2023-04-29).

[BP22] Wele Gedara Chaminda Bandara and Vishal M. Patel. A transformer-based siamese network for change detection. In *2022 IEEE INTERNATIONAL GEOSCIENCE AND REMOTE SENSING SYMPOSIUM (IGARSS 2022)*, IEEE International Symposium on Geoscience and Remote Sensing IGARSS, pp. 207–210. IEEE, 2022. ISBN: 978-1-6654-2792-0. DOI: 10.1109/IGARSS46834.2022.9883686. IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Kuala Lumpur, MALAYSIA, JUL 17-22, 2022.

[BSB+22] Abdul Basit, Muhammad Adnan Siddique, Muhammad Khurram Bhatti, and Muhammad Saquib Sarfraz. Comparison of CNNs and Vision Transformers-Based Hybrid Models Using Gradient Profile Loss for Classification of Oil Spills in SAR Images. en. *Remote Sensing*, 14(9):2085, 2022-04. ISSN: 2072-4292. DOI: 10.3390/rs14092085. URL: https://www.mdpi.com/2072-4292/14/9/2085 (visited on 2023-04-29).

[CLB+19] Rodrigo Caye Daudt, Bertrand Le Saux, Alexandre Boulch, and Yann Gousseau. Multitask learning for large-scale semantic change detection. en. *Computer Vision and Image Understanding*, 187:102783, 2019-10. ISSN: 10773142. DOI: 10.1016/j.cviu.2019.07.003. URL: https://linkinghub.elsevier.com/retrieve/pii/S1077314219300992 (visited on 2024-01-07).

[CPY+22] Hongjia Chen, Fangling Pu, Rui Yang, Rui Tang, and Xin Xu. Rdp-net: region detail preserving network for change detection. *IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING*, 60, 2022. ISSN: 0196-2892. DOI: 10.1109/TGRS.2022.3227098.

[DLB+18] Rodrigo Caye Daudt, Bertrand Le Saux, Alexandre Boulch, and Yann Gousseau. Urban change detection for multispectral earth observation using convolutional neural networks. In *IGARSS 2018 - 2018 IEEE INTERNATIONAL GEOSCIENCE AND REMOTE SENSING SYMPOSIUM*, IEEE International Symposium on Geoscience and

Remote Sensing IGARSS, pp. 2115–2118. Inst Elect & Elect Engineers; Inst Elect & Elect Engineers Geoscience & Remote Sensing Soc; European Space Agcy, 2018. ISBN: 978-1-5386-7150-4. 38th IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Valencia, SPAIN, JUL 22-27, 2018.

[DLB18a] Rodrigo Caye Daudt, Bertrand Le Saux, and Alexandre Boulch. Fully convolutional siamese networks for change detection, Inst Elect & Electron Engneers; IEEE Signal Processing Soc, 2018. 25th IEEE International Conference on Image Processing (ICIP), Athens, GREECE, OCT 07-10, 2018.

[DLB18b] Rodrigo Caye Daudt, Bertrand Le Saux, and Alexandre Boulch. Fully convolutional siamese networks for change detection. In *2018 25TH IEEE INTERNATIONAL CONFERENCE ON IMAGE PROCESSING (ICIP)*, IEEE International Conference on Image Processing ICIP, pp. 4063–4067. Inst Elect & Electron Engneers; IEEE Signal Processing Soc, 2018. ISBN: 978-1-4799-7061-2. 25th IEEE International Conference on Image Processing (ICIP), Athens, GREECE, OCT 07-10, 2018.

[DVD+21] David Duque-Arias, Santiago Velasco-Forero, Jean-Emmanuel Deschaud, François Goulette, Andres Serna, Etienne Decencière, and Beatriz Marcotegui. On Power Jaccard Losses for Semantic Segmentation: en. In *Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, pp. 561–568, Online Streaming, — Select a Country — . SCITEPRESS - Science and Technology Publications, 2021. ISBN: 978-989-758-488-6. DOI: 10.5220/0010304005610568. URL: https://www.scitepress.org/DigitalLibrary/Link.aspx?doi=10.5220/0010304005610568 (visited on 2023-05-07).

[EH11] H. M. El-Asmar and M. E. Hereher. Change detection of the coastal zone east of the nile delta using remote sensing. *Environmental Earth Sciences*, 62(4):769–777, 2011-02. ISSN: 1866-6299. DOI: 10.1007/s12665-010-0564-9. URL: https://doi.org/10.1007/s12665-010-0564-9.

[FKG+22] Tautvydas Fyleris, Andrius Kriščiūnas, Valentas Gružauskas, Dalia Čalnerytė, and Rimantas Barauskas. Urban change detection from aerial images using convolutional neural networks and transfer learning. *ISPRS International Journal of Geo-Information*, 11(4):246, 2022-04. ISSN: 2220-9964. DOI: 10.3390/ijgi11040246. URL: http://dx.doi.org/10.3390/ijgi11040246.

[GHP+19] Jinqi Gong, Xiangyun Hu, Shiyan Pang, and Kun Li. Patch matching and dense crf-based co-refinement for building change detection from bi-temporal aerial images. *SENSORS*, 19(7), 2019-04. DOI: 10.3390/s19071557.

[GP] Dalia Grendaite and Linas Petkevičius. Identification of algal blooms in lakes in the Baltic states using Sentinel-2 data and artificial neural networks. en.

[IKW+18] Fabian Isensee, Philipp Kickingereder, Wolfgang Wick, Martin Bendszus, and Klaus H. Maier-Hein. Brain tumor segmentation and radiomics survival prediction: contribution to the brats 2017 challenge. In A Crimi, S Bakas, H Kuijf, B Menze, and M Reyes, editors, *BRAINLESION: GLIOMA, MULTIPLE SCLEROSIS, STROKE AND TRAUMATIC BRAIN INJURIES, BRAINLES 2017*, vol. 10670 of *Lecture Notes in Computer Science*, pp. 287–297. Univ Penn, Ctr Biomed Image Comp & Analyt; Univ Med Ctr Utrecht, Image Sci Inst, 2018. ISBN: 978-3-319-75238-9; 978-3-319-75237-2. DOI: 10.1007/978-3-319-75238-9_25. 3rd International Workshop on Brain-Lesion (BrainLes) held jointly at the Conference on Medical Image Computing for Computer Assisted Intervention (MICCAI), Quebec City, CANADA, SEP 14, 2017.

[JCF+23] Johannes Jakubik, Linsong Chu, Paolo Fraccaro, Carlos Gomes, et al. Prithvi-100M, 2023-08. DOI: 10.57967/hf/0952.

[KCR+17] S. Kaliraj, N. Chandrasekar, K. K. Ramachandran, Y. Srinivas, and S. Saravanan. Coastal landuse and land cover change and transformations of kanyakumari coast, india using remote sensing and gis. *EGYPTIAN JOURNAL OF REMOTE SENSING AND SPACE SCIENCES*, 20(2):169–185, 2017-12. ISSN: 1110-9823. DOI: 10.1016/j.ejrs.2017.04.003.

[KKS+21] Krishna Karra, Caitlin Kontgis, Zoe Statman-Weil, Joseph C. Mazzariello, Mark Mathis, and Steven P. Brumby. Global land use / land cover with Sentinel 2 and deep learning. In *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, pp. 4704–4707, 2021-07. DOI: 10.1109/IGARSS47720.2021.9553499. ISSN: 2153-7003.

[KLZ+18] Ling Ke, Yukun Lin, Zhe Zeng, Lifu Zhang, and Lingkui Meng. Adaptive change detection with significance test. *IEEE Access*, 6:27442–27450, 2018. DOI: 10.1109/ACCESS.2018.2807380.

[LGG+20] Tsung-Yi Lin, Priya Goyal, Ross Girshick, Kaiming He, and Piotr Dollar. Focal Loss for Dense Object Detection. en. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(2):318–327, 2020-02. ISSN: 0162-8828, 2160-9292, 1939-3539. DOI: 10.1109/TPAMI.2018.2858826. URL: https://ieeexplore.ieee.org/document/8417976/ (visited on 2024-01-07).

[LJ20] Jin Liu and Shunping Ji. A Novel Recurrent Encoder-Decoder Structure for Large-Scale Multi-View Stereo Reconstruction From an Open Aerial Dataset. en. In *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6049–6058, Seattle, WA, USA. IEEE, 2020-06. ISBN: 978-1-72817-168-5. DOI: 10.1109/CVPR42600.2020.00609. URL: https://ieeexplore.ieee.org/document/9156578/ (visited on 2024-01-07).

[LVV+18] M. A. Lebedev, Y. V. Vizilter, O. V. Vygolov, V. A. Knyaz, and A. Y. Rubis. Change Detection in Remote Sensing Images Using Conditional Adversarial Networks. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 422:565–571, 2018-05. DOI: 10.5194/isprs-archives-XLII-2-565-2018.

[LXG+22] Jia Liu, Wenjie Xuan, Yuhang Gan, Yibing Zhan, Juhua Liu, and Bo Du. An end-to-end supervised domain adaptation framework for cross-domain change detection, 2022-12. DOI: 10.1016/j.patcog.2022.108960.

[LZZ+22] Haoyang Li, Fangjie Zhu, Xiaoyu Zheng, Mengxi Liu, and Guangzhao Chen. MSCDUNet: A Deep Learning Framework for Built-Up Area Change Detection Integrating Multispectral, SAR, and VHR Data. en. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15:5163–5176, 2022. ISSN: 1939-1404, 2151-1535. DOI: 10.1109/JSTARS.2022.3181155. URL: https://ieeexplore.ieee.org/document/9791854/ (visited on 2024-05-19).

[MDA18] Nouha Mezned, Belgacem Dkhala, and Saadi Abdeljaouad. Multitemporal and multisensory landsat etm+ and oli 8 data for mine waste change detection in northern tunisia. *JOURNAL OF SPATIAL SCIENCE*, 63(1):135–153, 2018. ISSN: 1449-8596. DOI: 10.1080/14498596.2017.1345666.

[MS21] Mina Mohammadi and Alireza Sharifi. Evaluation of convolutional neural networks for urban mapping using satellite images. *JOURNAL OF THE INDIAN SOCIETY OF REMOTE SENSING*, 49(9):2125–2131, 2021-09. ISSN: 0255-660X. DOI: 10.1007/s12524-021-01382-x.

[NH18] Virginia Ng and Daniel Hofmann. Scalable feature extraction with aerial and satellite imagery. In *Proceedings of the 17th Python in Science Conference (SciPy 2018)*, pp. 145–151, 2018-01. DOI: 10.25080/Majora-4af1f417-015.

[NH20] Truong Linh Nguyen and DongYeob Han. Detection of Road Surface Changes from Multi-Temporal Unmanned Aerial Vehicle Images Using a Convolutional Siamese Network. en. *Sustainability*, 12(6):2482, 2020-03. ISSN: 2071-1050. DOI: 10.3390/su12062482. URL: https://www.mdpi.com/2071-1050/12/6/2482 (visited on 2023-05-08).

[PZG19a] Daifeng Peng, Yongjun Zhang, and Haiyan Guan. End-to-end change detection for high resolution satellite images using improved unet++. *Remote Sensing*, 11(11):1382, 2019.

[PZG19b] Daifeng Peng, Yongjun Zhang, and Haiyan Guan. End-to-End Change Detection for High Resolution Satellite Images Using Improved UNet++. en. *Remote Sensing*, 11(11):1382, 2019-06. ISSN: 2072-4292. DOI: 10.3390/rs11111382. URL: https://www.mdpi.com/2072-4292/11/11/1382 (visited on 2023-05-01).

[RAD+20] Federico Ronci, Corrado Avolio, Mauro Di Donna, Massimo Zavagli, Veronica Piccialli, and Mario Costantini. Oil Spill Detection from SAR Images by Deep Learning. en. In *IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium*, pp. 2225–2228, Waikoloa, HI, USA. IEEE, 2020-09. ISBN: 978-1-72816-374-1. DOI: 10.1109/IGARSS39084.2020.9323590. URL: https://ieeexplore.ieee.org/document/9323590/ (visited on 2023-04-29).

[RCS+21] Maryam Rahnemoonfar, Tashnim Chowdhury, Argho Sarkar, Debvrat Varshney, Masoud Yari, and Robin Roberson Murphy. FloodNet: A High Resolution Aerial Imagery Dataset for Post Flood Scene Understanding. en. *IEEE Access*, 9:89644–89654, 2021. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2021.3090981. URL: https://ieeexplore.ieee.org/document/9460988/ (visited on 2024-01-07).

[RFB15] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: convolutional networks for biomedical image segmentation. N Navab, J Hornegger, WM Wells, and AF Frangi, editors, Tech Univ Munich; Friedrich Alexander Univ Erlangen Nuremberg, 2015. DOI: 10.1007/978-3-319-24574-4_28. 18th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Munich, GERMANY, OCT 05-09, 2015.

[RL98] Merrill K Ridd and Jiajun Liu. A comparison of four algorithms for change detection in an urban environment. *Remote Sensing of Environment*, 63(2):95–100, 1998. ISSN: 0034-4257. DOI: https://doi.org/10.1016/S0034-4257(97)00112-0. URL: https://www.sciencedirect.com/science/article/pii/S0034425797001120.

[SHQ+19a] M. Schmitt, L. H. Hughes, C. Qiu, and X. X. Zhu. AGGREGATING CLOUD-FREE SENTINEL-2 IMAGES WITH GOOGLE EARTH ENGINE. en. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV-2/W7:145–152, 2019-09. ISSN: 2194-9050. DOI: 10.5194/isprs-annals-IV-2-W7-145-2019. URL: https://www.isprs-ann-photogramm-remote-sens-spatial-inf-sci.net/IV-2-W7/145/2019/ (visited on 2023-03-09).

[SHQ+19b] M. Schmitt, L. H. Hughes, C. Qiu, and X. X. Zhu. SEN12MS – A CURATED DATASET OF GEOREFERENCED MULTI-SPECTRAL SENTINEL-1/2 IMAGERY FOR DEEP LEARNING AND DATA FUSION. en. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV-2/W7:153–160, 2019-09. ISSN: 2194-9050. DOI: 10.5194/isprs-annals-IV-2-W7-153-2019. URL: https://isprs-annals.copernicus.org/articles/IV-2-W7/153/2019/ (visited on 2024-01-07).

[SHZ18] M. Schmitt, L. H. Hughes, and X. X. Zhu. THE SEN1-2 DATASET FOR DEEP LEARNING IN SAR-OPTICAL DATA FUSION. en. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, IV-1:141–146, 2018-09. ISSN: 2194-9050. DOI: 10.5194/isprs-annals-IV-1-141-2018.

URL: https://isprs-annals.copernicus.org/articles/IV-1/141/2018/ (visited on 2024-01-07).

[SLV+17] Carole H. Sudre, Wenqi Li, Tom Vercauteren, Sebastien Ourselin, and M. Jorge Cardoso. Generalised Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations. en. In M. Jorge Cardoso, Tal Arbel, Gustavo Carneiro, Tanveer Syeda-Mahmood, et al., editors, *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Vol. 10553, pp. 240–248. Springer International Publishing, Cham, 2017. ISBN: 978-3-319-67557-2 978-3-319-67558-9. DOI: 10.1007/978-3-319-67558-9_28. URL: http://link.springer.com/10.1007/978-3-319-67558-9_28 (visited on 2024-01-07). Series Title: Lecture Notes in Computer Science.

[SPS+19] Vaughn Smith, Carlos Portillo-Quintero, Arturo Sanchez-Azofeifa, and Jose L. Hernandez-Stefanoni. Assessing the accuracy of detected breaks in landsat time series as predictors of small scale deforestation in tropical dry forests of mexico and costa rica. *REMOTE SENSING OF ENVIRONMENT*, 221:707–721, 2019-02. ISSN: 0034-4257. DOI: 10.1016/j.rse.2018.12.020.

[STL+20] Jinyuan Shao, Lina Tang, Ming Liu, Guofan Shao, Lang Sun, and Quanyi Qiu. BDD-Net: A General Protocol for Mapping Buildings Damaged by a Wide Range of Disasters Based on Satellite Imagery. en. *Remote Sensing*, 12(10):1670, 2020-05. ISSN: 2072-4292. DOI: 10.3390/rs12101670. URL: https://www.mdpi.com/2072-4292/12/10/1670 (visited on 2023-04-29).

[SUG22] Edvinas Stonevicius, Giedrius Uselis, and Dalia Grendaite. Ice detection with sentinel-1 sar backscatter threshold in long sections of temperate climate rivers. *REMOTE SENSING*, 14(7), 2022-04. DOI: 10.3390/rs14071627.

[SWZ+21] Hua Su, An Wang, Tianyi Zhang, Tian Qin, Xiaoping Du, and Xiao-Hai Yan. Super-resolution of subsurface temperature field from remote sensing observations based on machine learning. *INTERNATIONAL JOURNAL OF APPLIED EARTH OBSERVATION AND GEOINFORMATION*, 102, 2021-10. ISSN: 1569-8432. DOI: 10.1016/j.jag.2021.102440. Number: 102440 tex.eissn: 1872-826X tex.orcid-numbers: Su, Hua/0000-0003-0280-3926 Wang, An/0000-0003-0361-7543 tex.researcherid-numbers: Su, Hua/GSI-7526-2022 Wang, Anzhong/AAS-9614-2020 tex.unique-id: WOS:000685083000002.

[SYL+23] Bella Specktor-Fadida, Bossmat Yehuda, Daphna Link-Sourani, Liat Ben-Sira, Dafna Ben-Bashat, and Leo Joskowicz. Contour Dice Loss for Structures with Fuzzy and Complex Boundaries in Fetal MRI. en. In Leonid Karlinsky, Tomer Michaeli, and Ko Nishino, editors, *Computer Vision – ECCV 2022 Workshops*. Vol. 13803, pp. 355–368. Springer Nature Switzerland, Cham, 2023. ISBN: 978-3-031-25065-1 978-3-031-25066-8. DOI: 10.1007/978-3-031-25066-8_19. URL: https://link.

springer.com/10.1007/978-3-031-25066-8_19 (visited on 2024-01-07). Series Title: Lecture Notes in Computer Science.

[TKW+22] Aysim Toker, Lukas Kondmann, Mark Weber, Marvin Eisenberger, et al. DynamicEarthNet: Daily Multi-Spectral Satellite Dataset for Semantic Change Segmentation. en. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 21126–21135, New Orleans, LA, USA. IEEE, 2022-06. ISBN: 978-1-66546-946-3. DOI: 10.1109/CVPR52688.2022.02048. URL: https://ieeexplore.ieee.org/document/9879359/ (visited on 2024-01-07).

[TZZ+22] Shiqi Tian, Yanfei Zhong, Zhuo Zheng, Ailong Ma, Xicheng Tan, and Liangpei Zhang. Large-scale deep learning based binary and semantic change detection in ultra high resolution remote sensing imagery: From benchmark datasets to urban application. en. *ISPRS Journal of Photogrammetry and Remote Sensing*, 193:164–186, 2022-11. ISSN: 09242716. DOI: 10.1016/j.isprsjprs.2022.08.012. URL: https://linkinghub.elsevier.com/retrieve/pii/S0924271622002210 (visited on 2024-01-07).

[WSH+22] Han Wu, Huina Song, Jianhua Huang, Hua Zhong, Ronghui Zhan, Xuyang Teng, Zhaoyang Qiu, Meilin He, and Jiayi Cao. Flood Detection in Dual-Polarization SAR Images Based on Multi-Scale Deeplab Model. en. *Remote Sensing*, 14(20):5181, 2022-10. ISSN: 2072-4292. DOI: 10.3390/rs14205181. URL: https://www.mdpi.com/2072-4292/14/20/5181 (visited on 2023-04-29).

[WZM+21] Junjue Wang, Zhuo Zheng, Ailong Ma, Xiaoyan Lu, and Yanfei Zhong. LoveDA: a remote sensing land-cover dataset for domain adaptive semantic segmentation. Zenodo, 2021-10. DOI: 10.5281/zenodo.5706578. URL: https://doi.org/10.5281/zenodo.5706578.

[XWJ+21] Shao Xiang, Mi Wang, Xiaofan Jiang, Guangqi Xie, Zhiqi Zhang, and Peng Tang. Dual-Task Semantic Change Detection for Remote Sensing Images Using the Generative Change Field Module. en. *Remote Sensing*, 13(16):3336, 2021-08. ISSN: 2072-4292. DOI: 10.3390/rs13163336. URL: https://www.mdpi.com/2072-4292/13/16/3336 (visited on 2023-05-08).

[YLX+15] Daiqin Yang, Zimeng Li, Yatong Xia, and Zhenzhong Chen. Remote sensing image super-resolution: challenges and approaches. In *2015 IEEE International Conference on Digital Signal Processing (DSP)*, pp. 196–200, 2015. DOI: 10.1109/ICDSP.2015.7251858.

[ZMH+22] Jiding Zhai, Chunxiao Mu, Yongchao Hou, Jianping Wang, Yingjie Wang, and Haokun Chi. A Dual Attention Encoding Network Using Gradient Profile Loss for Oil Spill Detection Based on SAR Images. en. *Entropy*, 24(10):1453, 2022-10. ISSN: 1099-4300. DOI: 10.3390/e24101453. URL: https://www.mdpi.com/1099-4300/24/10/1453 (visited on 2023-05-08).

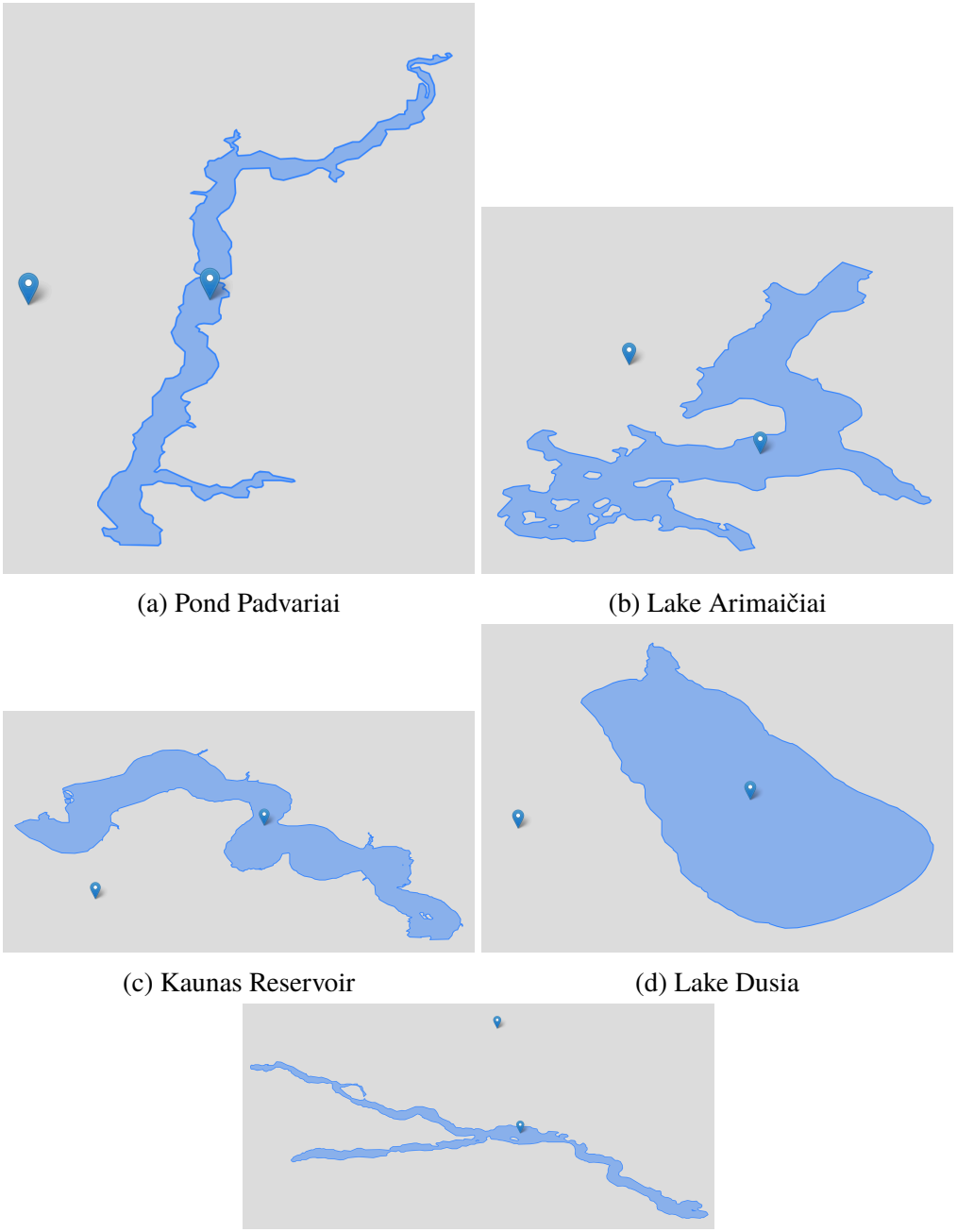
[ZMZ+21] Zhuo Zheng, Ailong Ma, Liangpei Zhang, and Yanfei Zhong. Change is everywhere: single-temporal supervised object change detection in remote sensing imagery. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 15193–15202, 2021.

[ZS20] Min Zhang and Wenzhong Shi. A Feature Difference Convolutional Neural Network-Based Change Detection Method. en. *IEEE Transactions on Geoscience and Remote Sensing*, 58(10):7232–7246, 2020-10. ISSN: 0196-2892, 1558-0644. DOI: 10.1109/TGRS.2020.2981051. URL: https://ieeexplore.ieee.org/document/9052762/ (visited on 2023-05-01).

[ZYD+22] Xue Zhao, Yang Yang, Fuzhou Duan, Miao Zhang, Guofu Jiang, Xing Yan, Shisong Cao, and Wenji Zhao. Identification of construction and demolition waste based on change detection and deep learning. *INTERNATIONAL JOURNAL OF REMOTE SENSING*, 43(6):2012–2028, 2022-03. ISSN: 0143-1161. DOI: 10.1080/01431161.2022.2054296.

Appendix Nr. 1

5 selected lakes with points



(a) Pond Padvariai

(b) Lake Arimaičiai

(c) Kaunas Reservoir

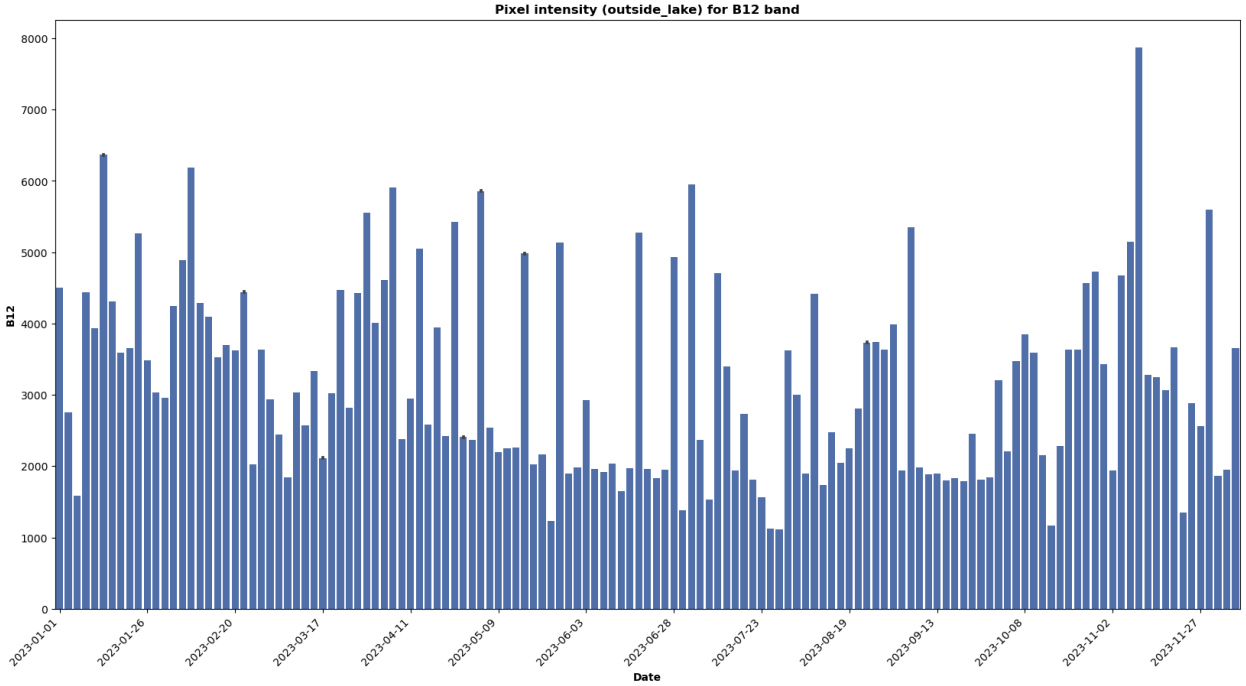
(d) Lake Dusia

(e) Lake Asveja

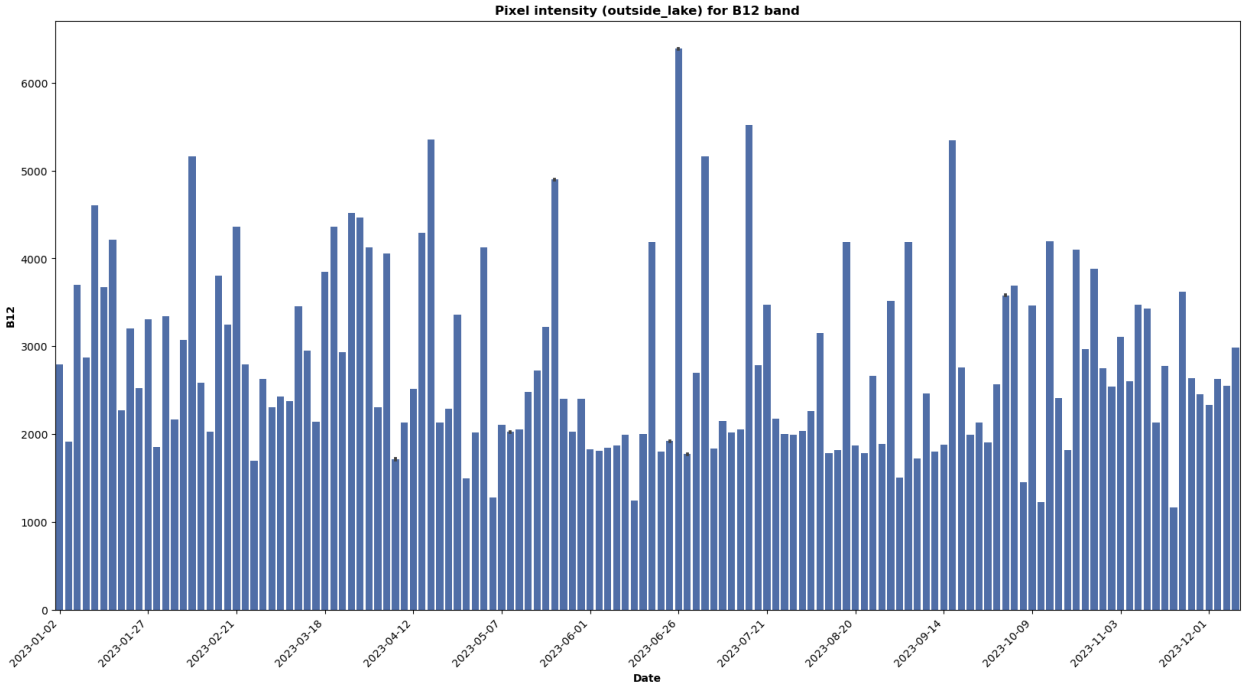
Figure 26. 5 selected water bodies with points that are selected for analysis

Appendix Nr. 2

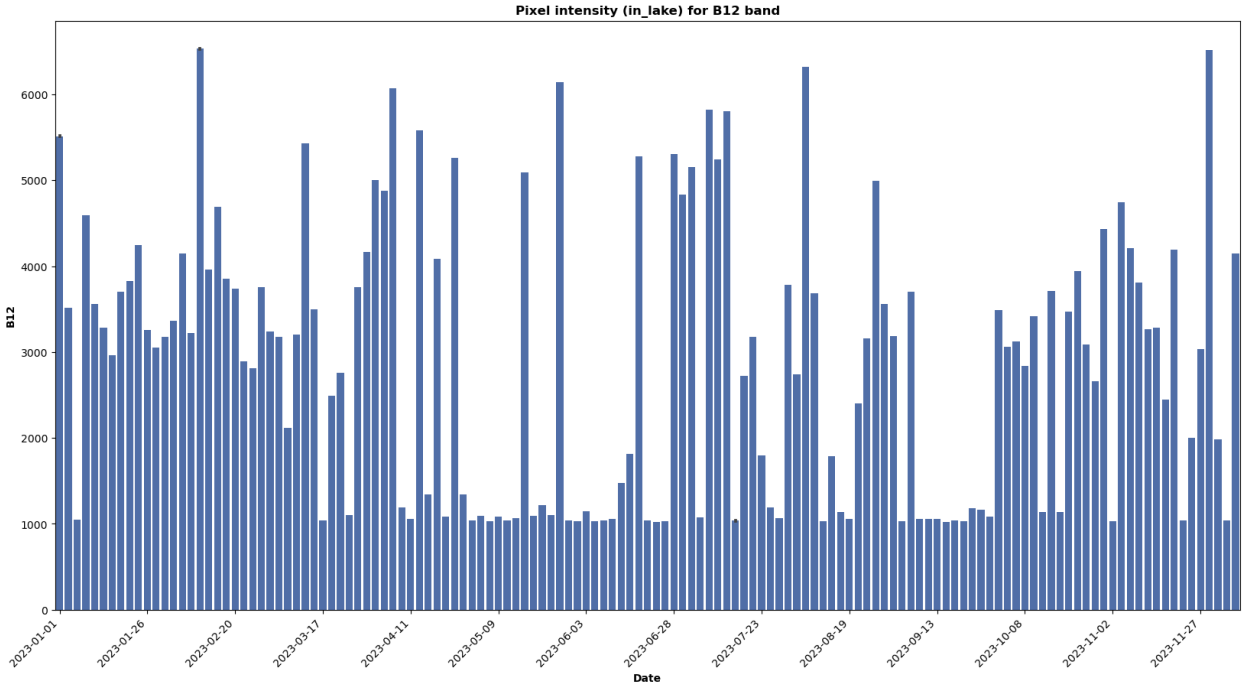
The closes and the furthest water body form Baltic sea



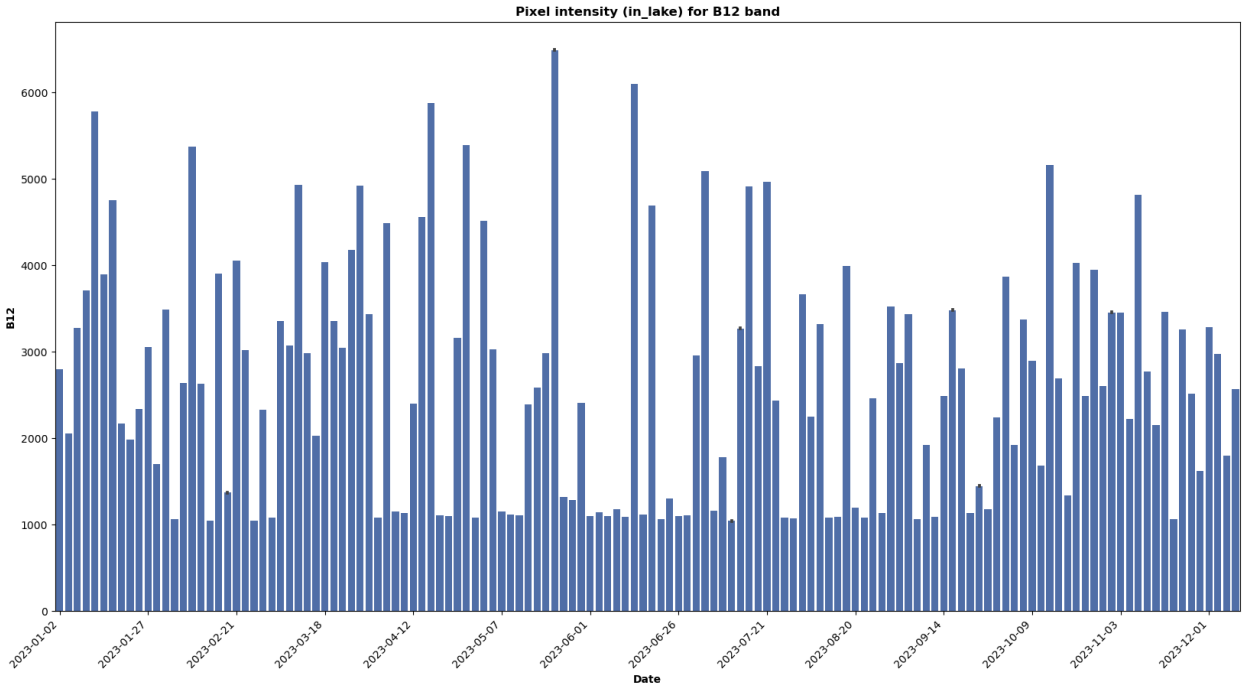
(a) 2023 outside lake Asveja



(b) 2023 outside pond Padvariai



(a) 2023 inside lake Asveja



(b) 2023 inside pond Padvariai

Appendix Nr. 3
Running median distribution with data (2019-2013)

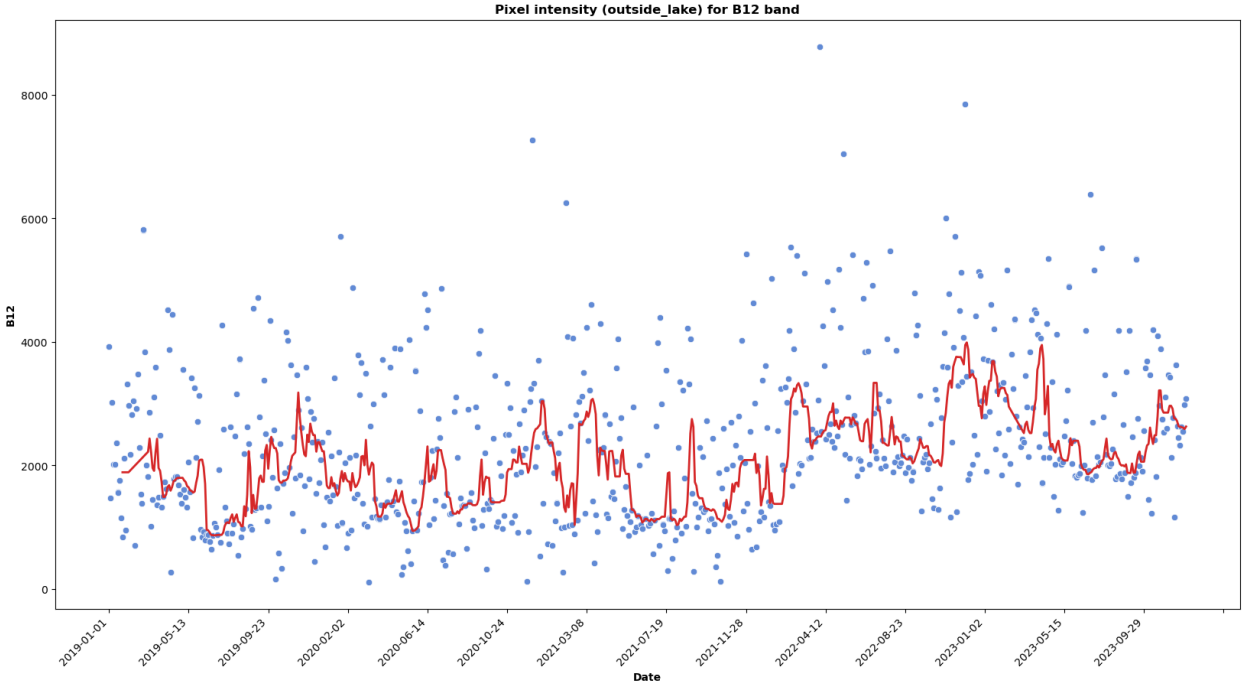
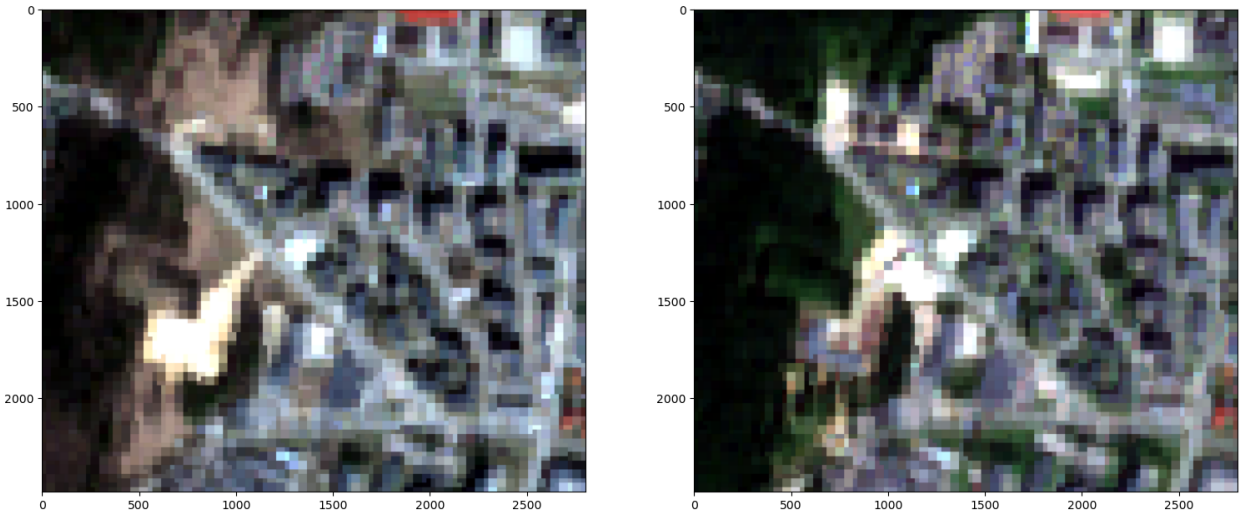
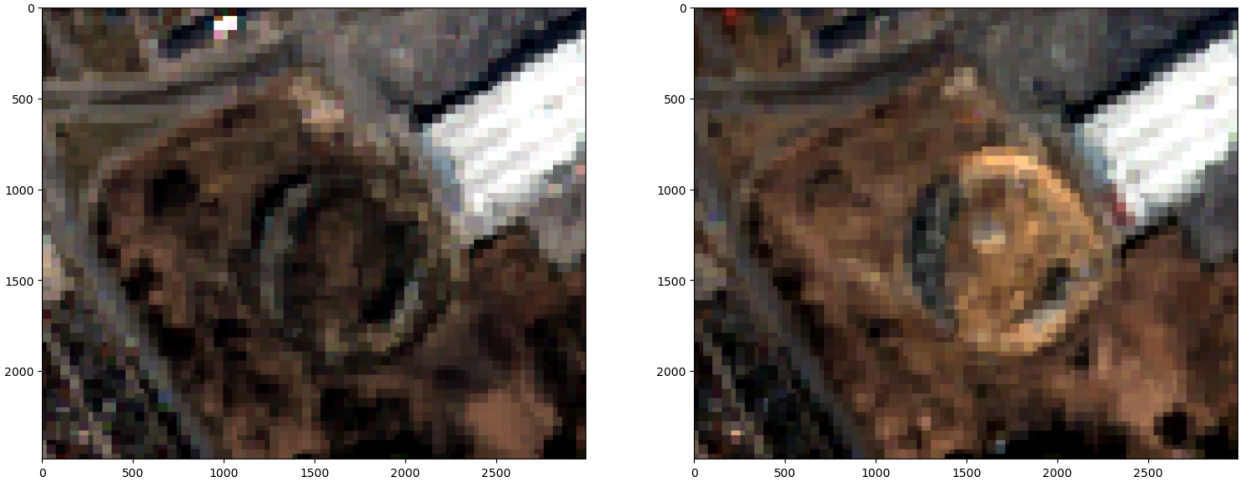


Figure 29. Outside pond Padvariai

Appendix Nr. 4
Chosen images of 4 regions

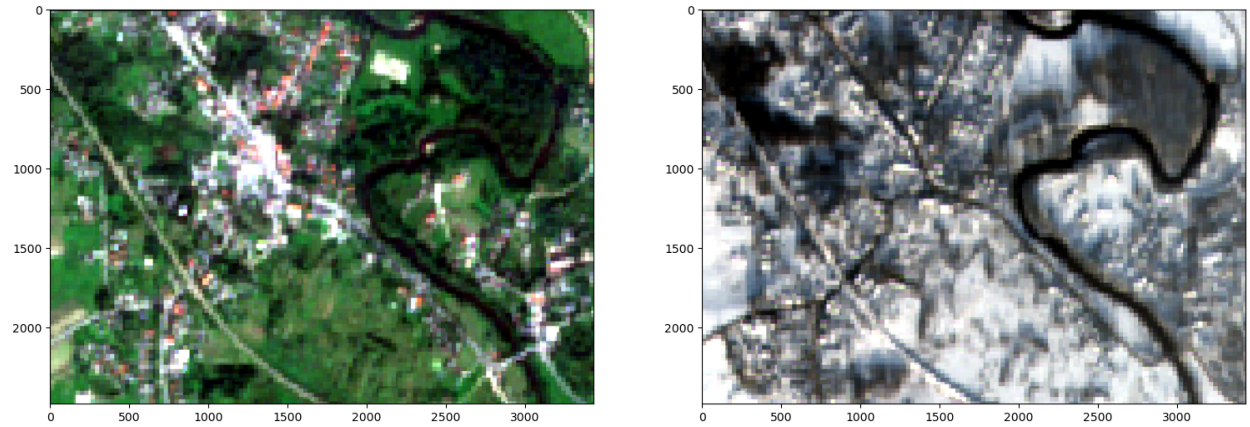


(a) Bajorai district. Left image - 2019-04-02, right image - 2022-08-12

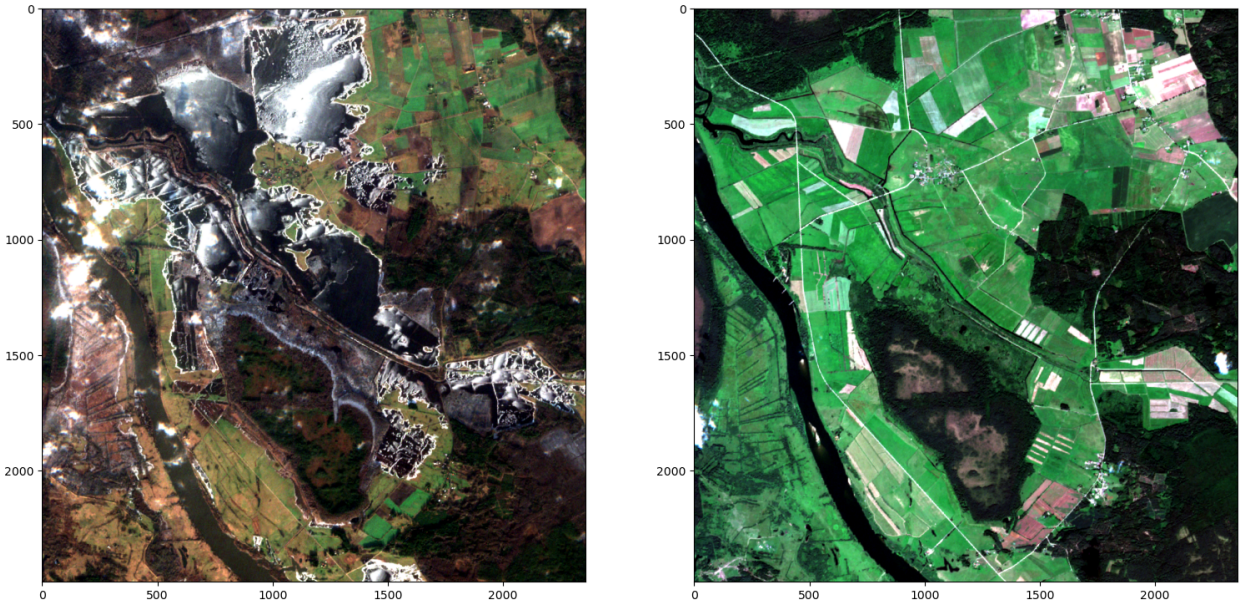


(b) National stadium constructions. Left image - 2019-03-23, right image - 2022-03-20

Figure 30. Images A (left column) and images B (right column) that are to be used in analysis.



(a) Minija river floods. Left image - 2023-09-09, right image - 2023-12-03



(b) Šyša river floods. Left image - 2023-01-12, right image - 2023-08-12

Figure 31. Images A (left column) and images B (right column) that were used in analysis.

Appendix Nr. 5
Manual evaluation points A and E

(a) FC-Siam-Diff

(b) Custom

Figure 32. Point A detected changes by each model

(a) FC-Siam-Diff

(b) Custom

Figure 33. Point B detected changes by each model

(a) FC-Siam-Diff

(b) Custom

Figure 34. Point C detected changes by each model

(a) FC-Siam-Diff

(b) Custom

Figure 35. Point D detected changes by each model