Машинное обучение и искусственный интеллект для решения научных и прикладных задач

Александр Ужинский
auzhinskiy@jinr.ru

Industry 4.0

The term “Industry 4.0” is used to signify the beginning of the fourth industrial revolution – the previous three being mechanical production, mass production, and then the digital revolution. It could be argued that Industry 4.0 is simply an amalgamation of the three previous eras in manufacturing, but Industry 4.0 is poised to be much more impactful than that.

Industrial Internet of Things (IIoT), Automation, **Artificial Intelligence, Big Data & Analytics, The Cloud, Cybersecurity, Simulations, Robotics, Smart manufacture, Mobile devices, Smart manufacture, etc.**

Basis - machine learning

X — множество *объектов* (точнее, их информационных описаний)
Y — множество *ответов* (оценок, предсказаний или прогнозов)
$y: X \to Y$ — неизвестная зависимость (target function)

Дано:
$\{x_1, \dots, x_\ell\} \subset X$ — *обучающая выборка* (training sample)
$y_i = y(x_i), \;\; i = 1, \dots, \ell$ — известные ответы

Найти:
$a: X \to Y$ — алгоритм, решающую функцию (decision function), приближающую y на всём множестве X

Весь курс машинного обучения — это конкретизация:

- как задаются объекты и какими могут быть ответы
- в каком смысле «a приближает y»
- как строить функцию a

Basis - machine learning

$f_j\colon X \to D_j,\;\; j = 1, \dots, n$ — признаки объектов (features)

Типы признаков:

- $D_j = \{0, 1\}$ — *бинарный* признак f_j
- $|D_j| < \infty$ — *номинальный* признак f_j
- $|D_j| < \infty$, D_j упорядочено — *порядковый* признак f_j
- $D_j = \mathbb{R}$ — *количественный* признак f_j

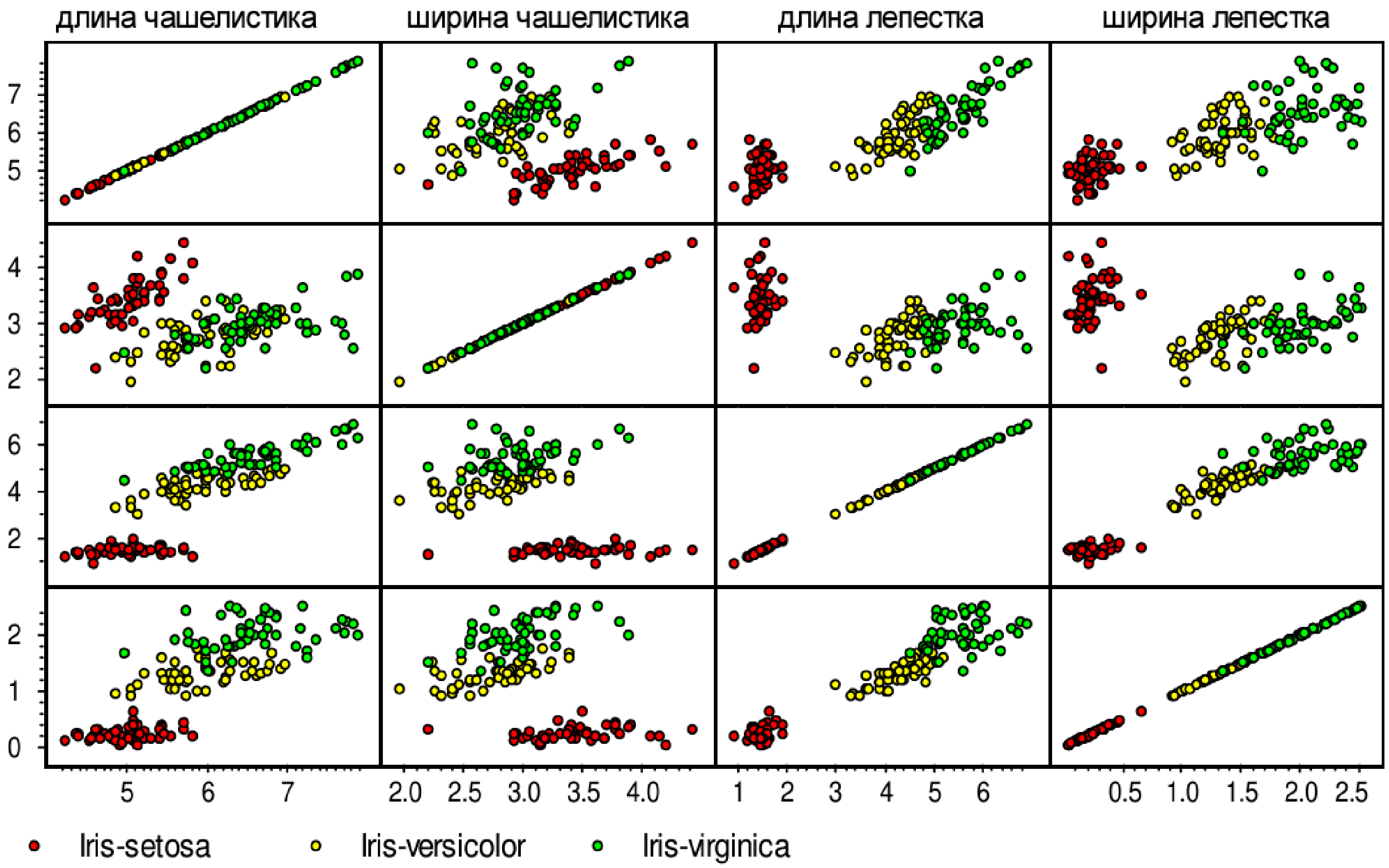
Вектор $\bigl(f_1(x), \dots, f_n(x)\bigr)$ — *признаковое описание* объекта x

Матрица «*объекты–признаки*» (feature data)

$$F = \bigl\| f_j(x_i) \bigr\|_{\ell \times n} = \begin{pmatrix} f_1(x_1) & \dots & f_n(x_1) \\ \dots & \dots & \dots \\ f_1(x_\ell) & \dots & f_n(x_\ell) \end{pmatrix}$$

Basis - machine learning

$n = 4$ признака, $|Y| = 3$ класса, длина выборки $\ell = 150$.



К. В. Воронцов (k.v.vorontsov@phystech.edu)

Basis - machine learning

Модель (predictive model) — параметрическое семейство функций

$$A = \{g(x, \theta) \mid \theta \in \Theta\},$$

где $g\colon X \times \Theta \to Y$ — фиксированная функция,
Θ — множество допустимых значений параметра θ

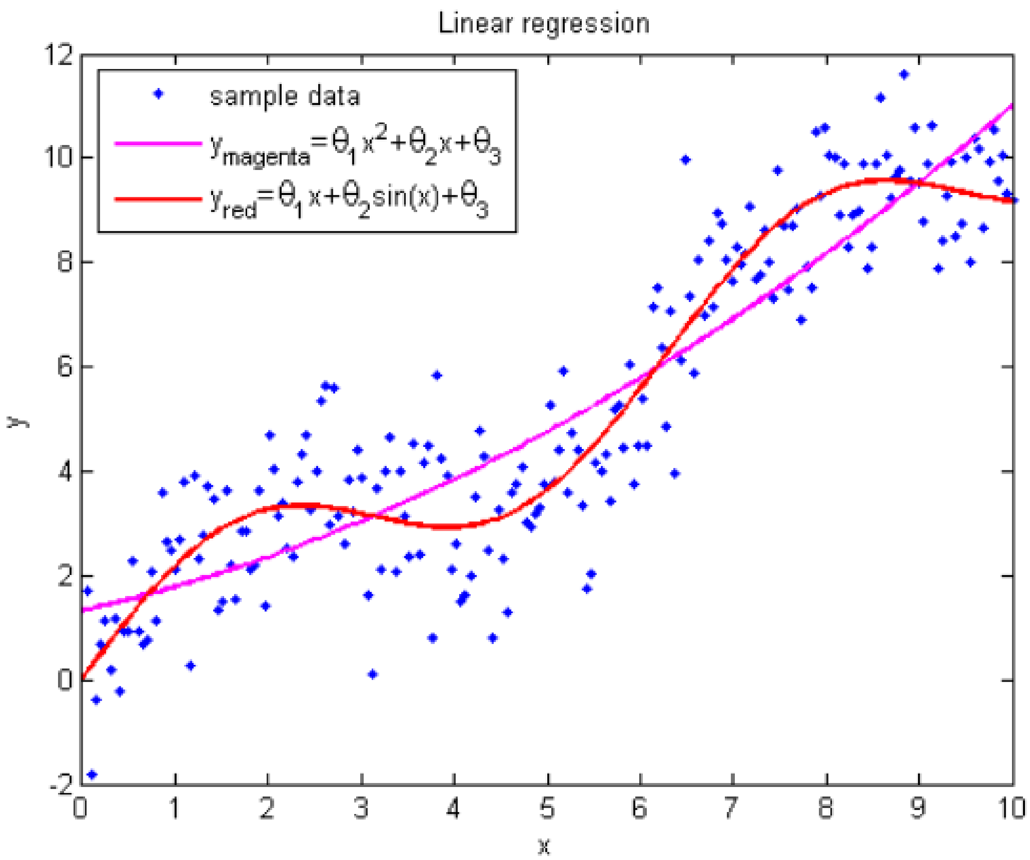
Пример.
Линейная модель с вектором параметров $\theta = (\theta_1, \dots, \theta_n) \in \mathbb{R}^n$:

$$g(x, \theta) = \sum_{j=1}^{n} \theta_j f_j(x) \text{ — для регрессии и ранжирования, } Y = \mathbb{R}$$

$$g(x, \theta) = \operatorname{sign} \sum_{j=1}^{n} \theta_j f_j(x) \text{ — для классификации, } Y = \{-1, +1\}$$

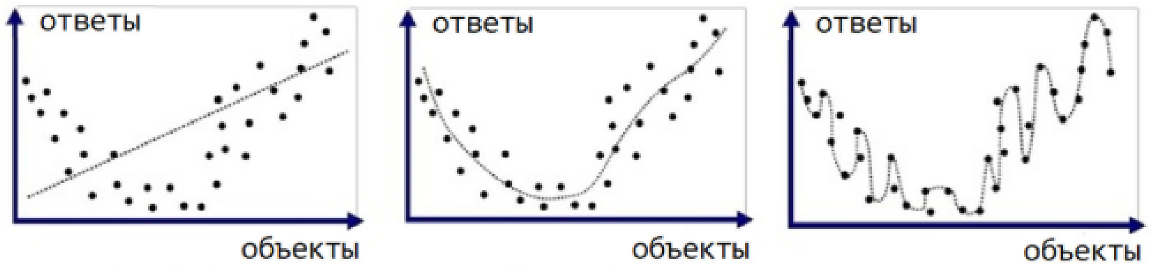
Basis - machine learning

$X = Y = \mathbb{R}$, $\ell = 200$, $n = 3$ признака: $\{x, x^2, 1\}$ или $\{x, \sin x, 1\}$



- генерация признаков (feature generation) обогащает модель
- на практике очень важно «правильно угадать модель»

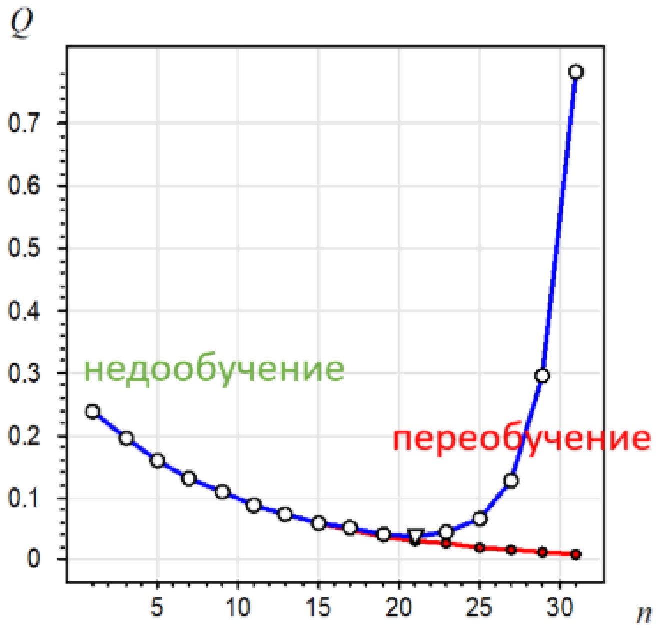
Basis - machine learning



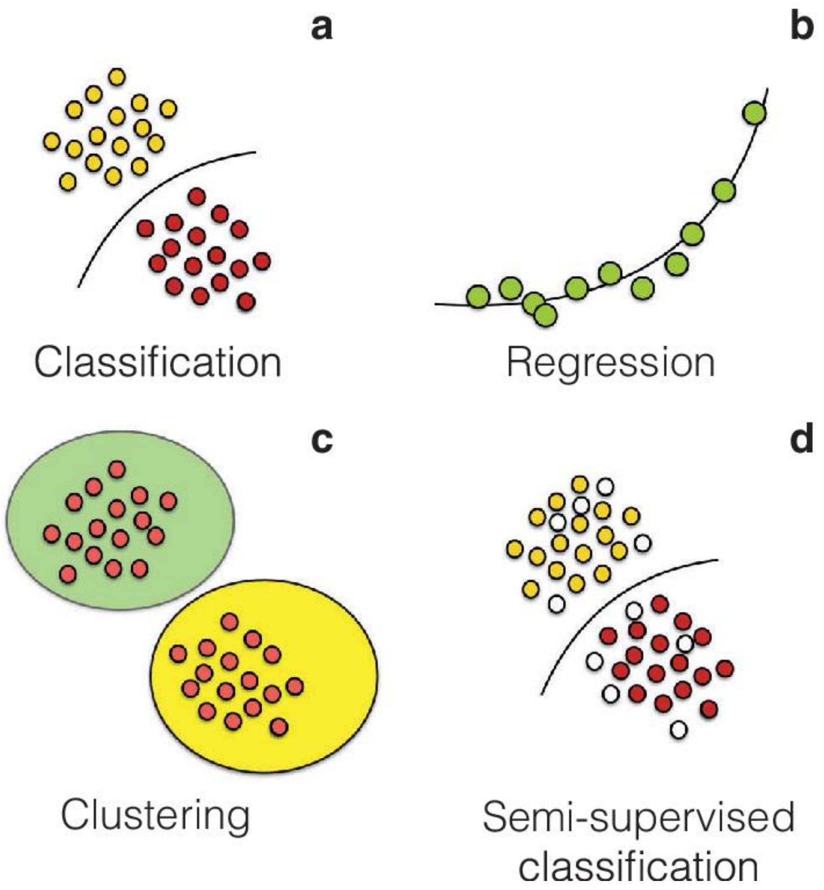
недообучение

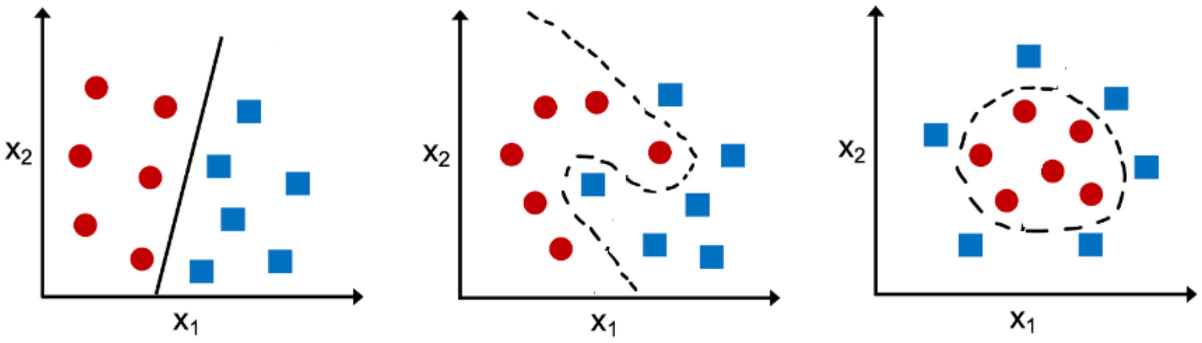
переобучение

- Недообучение (underfitting): модель слишком проста, недостаточное число параметров n
- Переобучение (overfitting): модель слишком сложна, избыточное число параметров n



Back to the machine learning





Classification

Basis - machine learning

Воронцов Константин Вячеславович

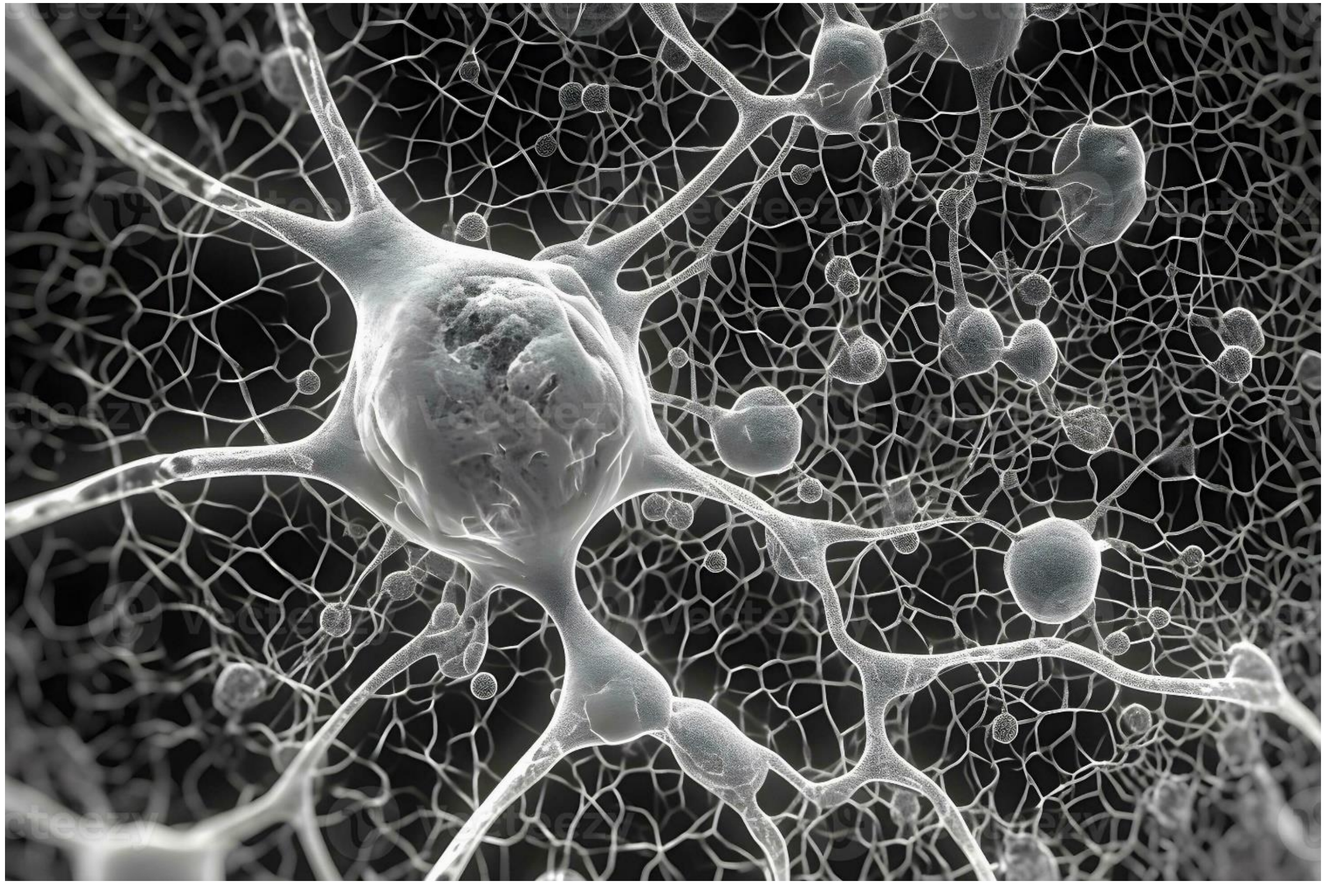
Зав. кафедрой ММП, профессор РАН, д.ф.-м.н.,
доцент ММП

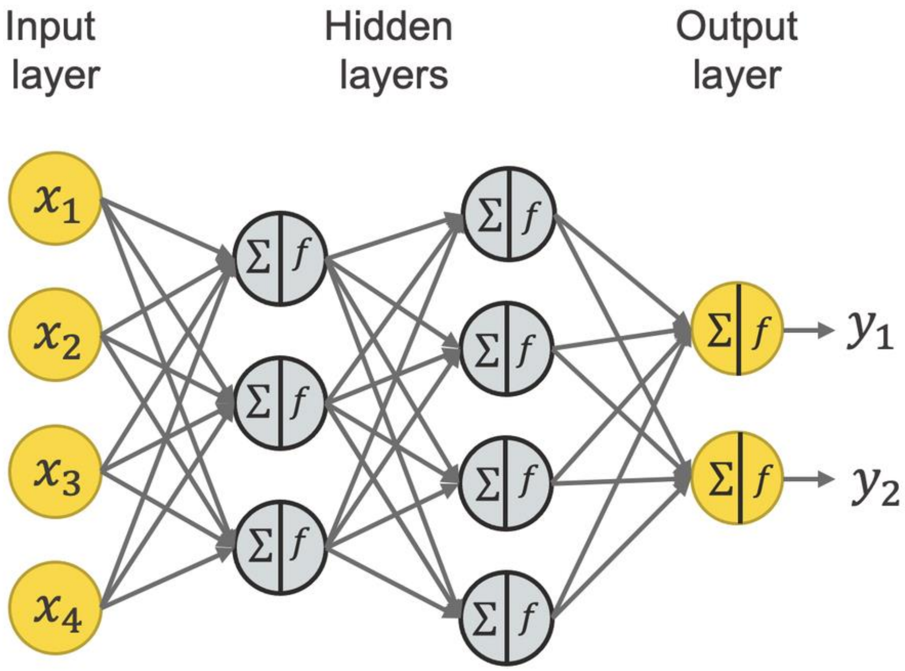
Математические основы машинного обучения

https://www.youtube.com/@MachineLearningIS

Neural networks

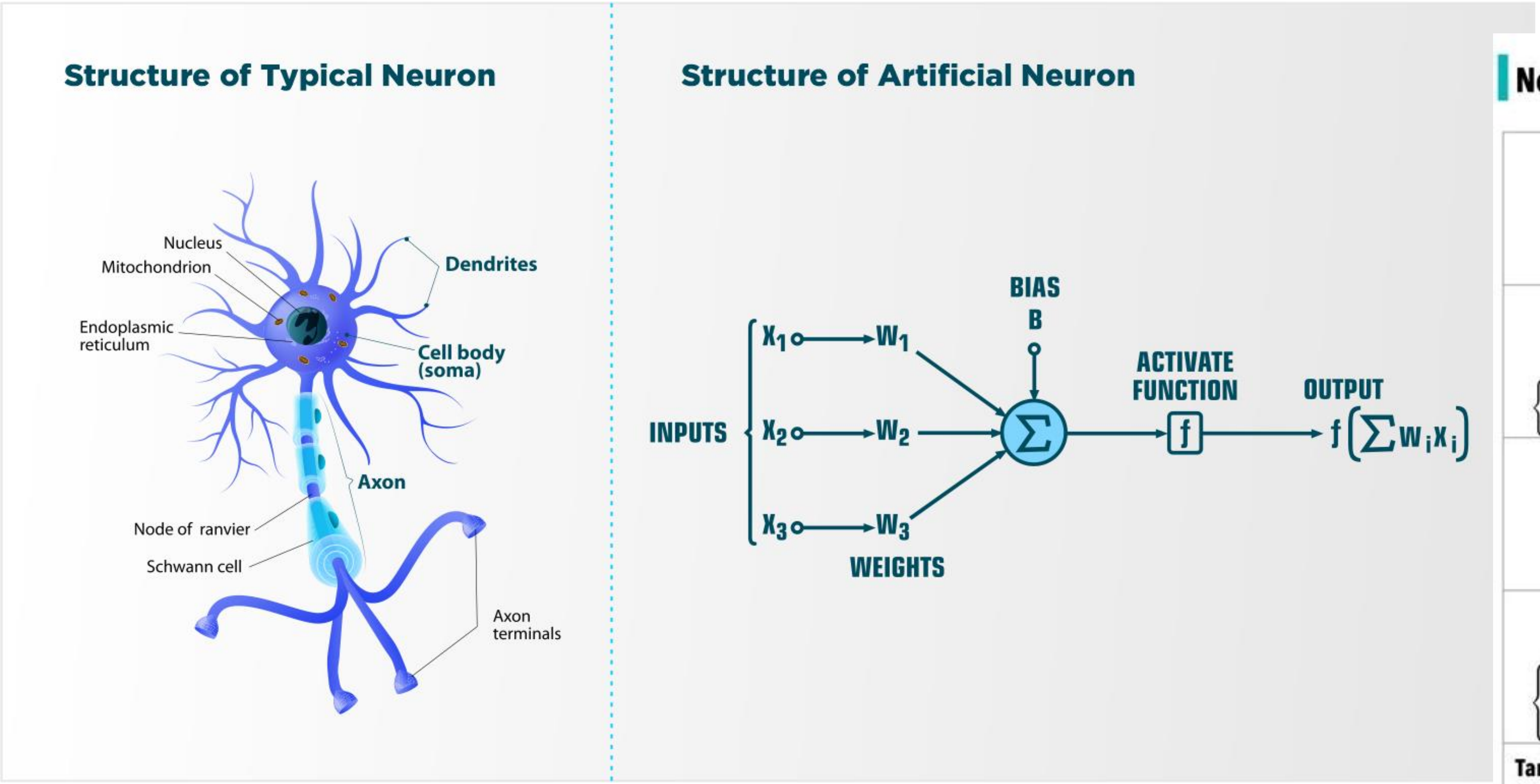
Neural networks are an example of a supervised learning algorithm and seek to approximate the function represented by your data

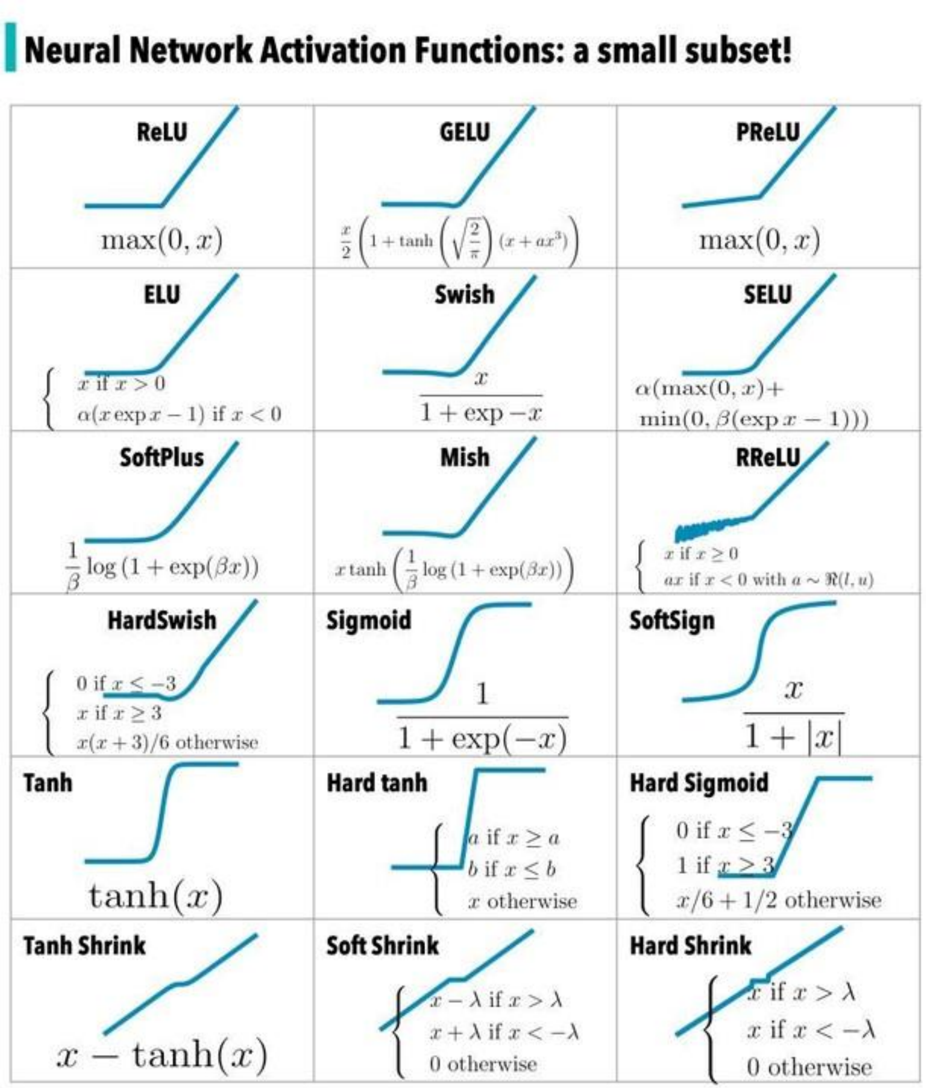




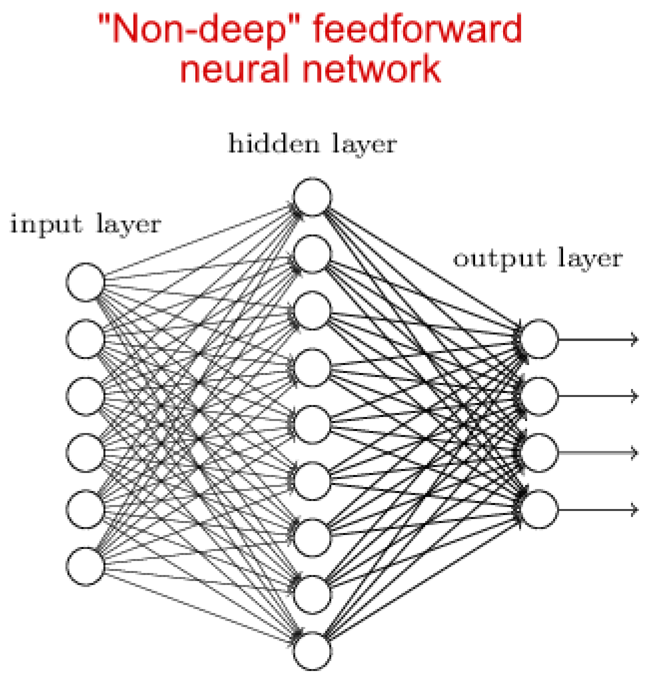
Farley and Wesley A. Clark were the first to simulate a Hebbian network in 1954 at MIT. They used computational machines, then called "calculators".

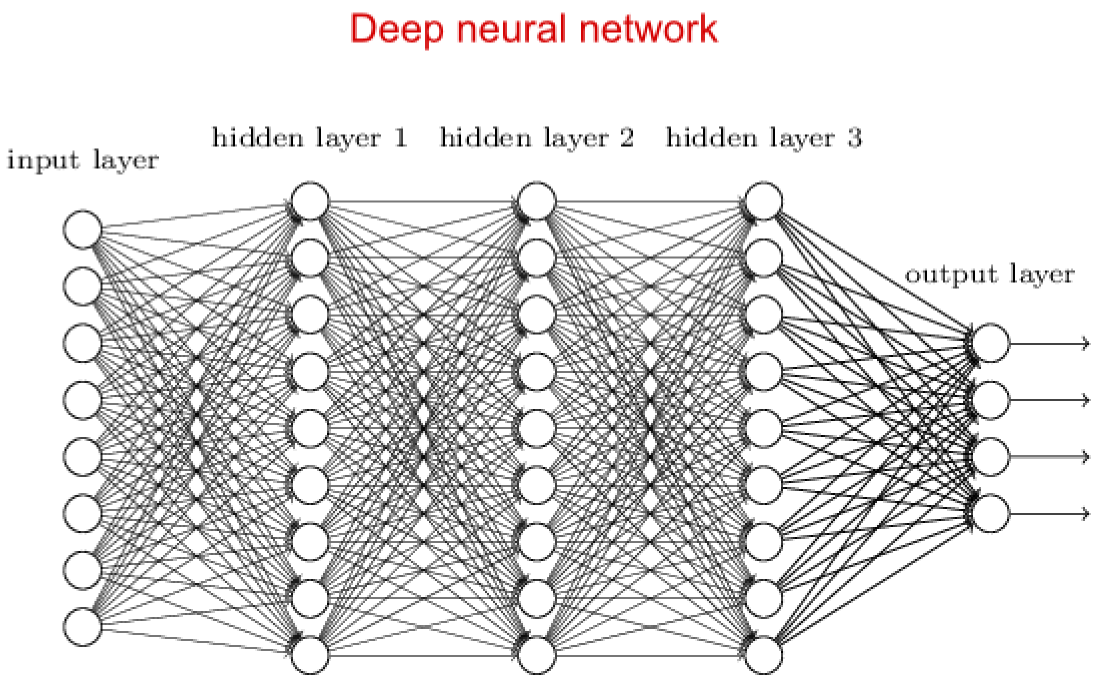
Neural networks





Deep neural networks

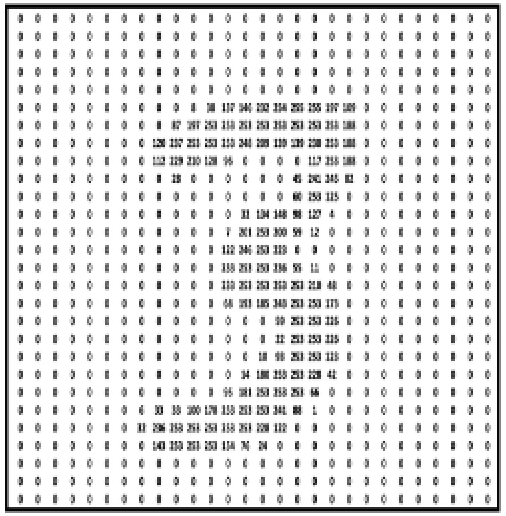




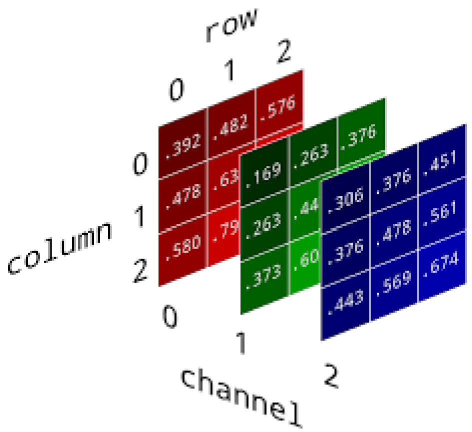
A neural network with multiple hidden layers and multiple nodes in each hidden layer is known as a **deep learning system** or a **deep neural network**. Deep learning is the development of deep learning algorithms that can be used to train and predict output from complex data.

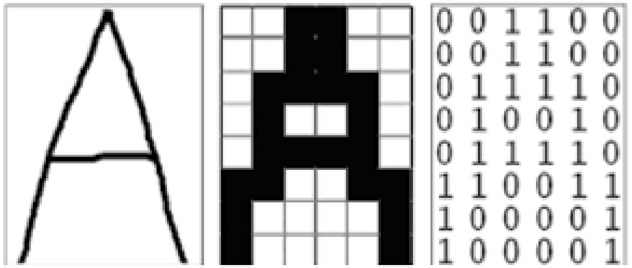
What computer sees?

What we see



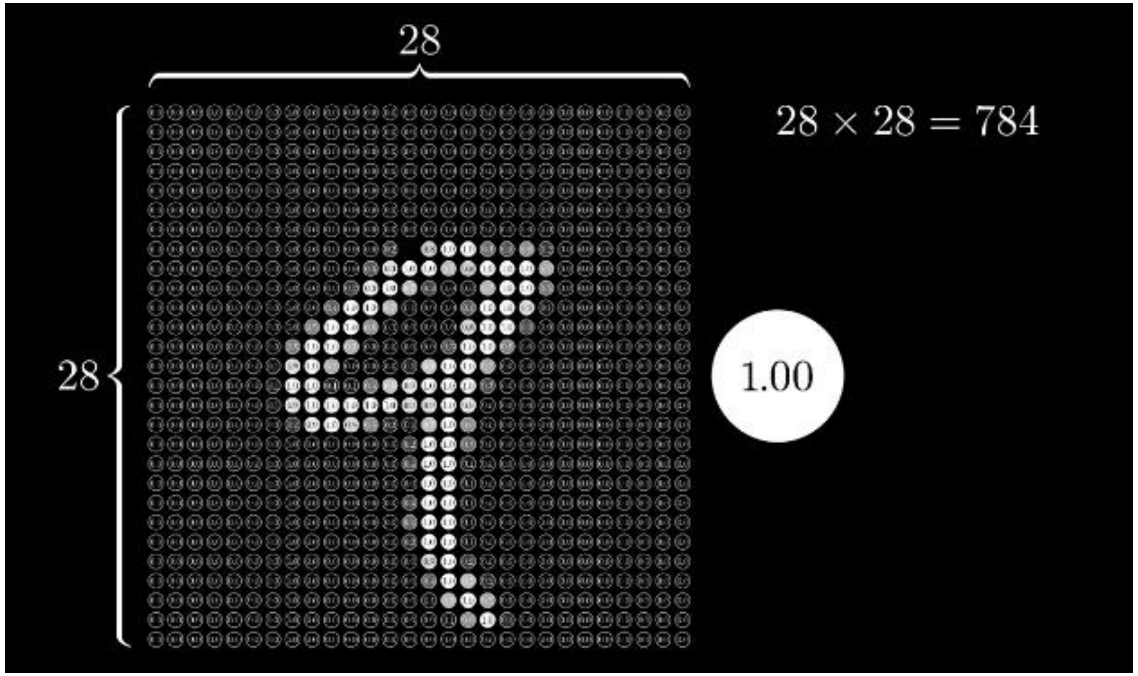
What a computer sees

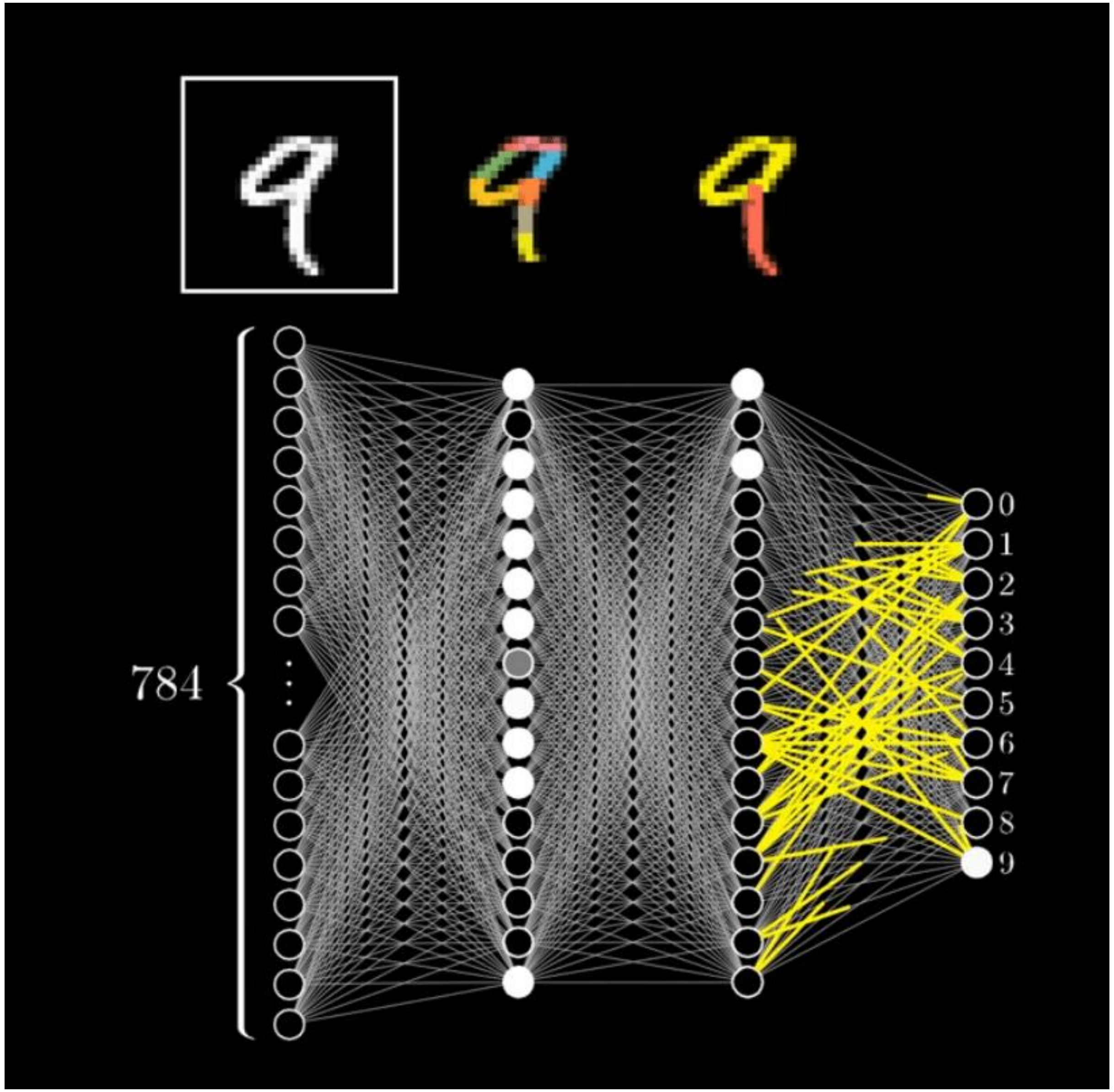




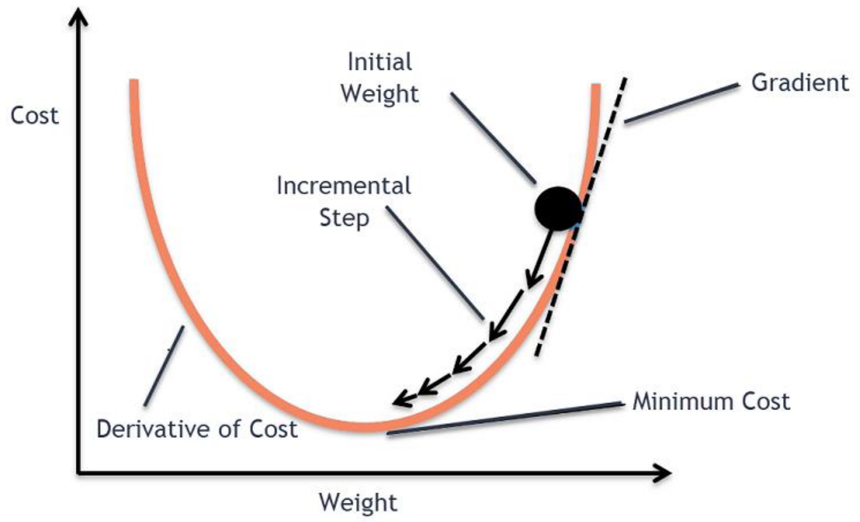
The image itself is a source of information.

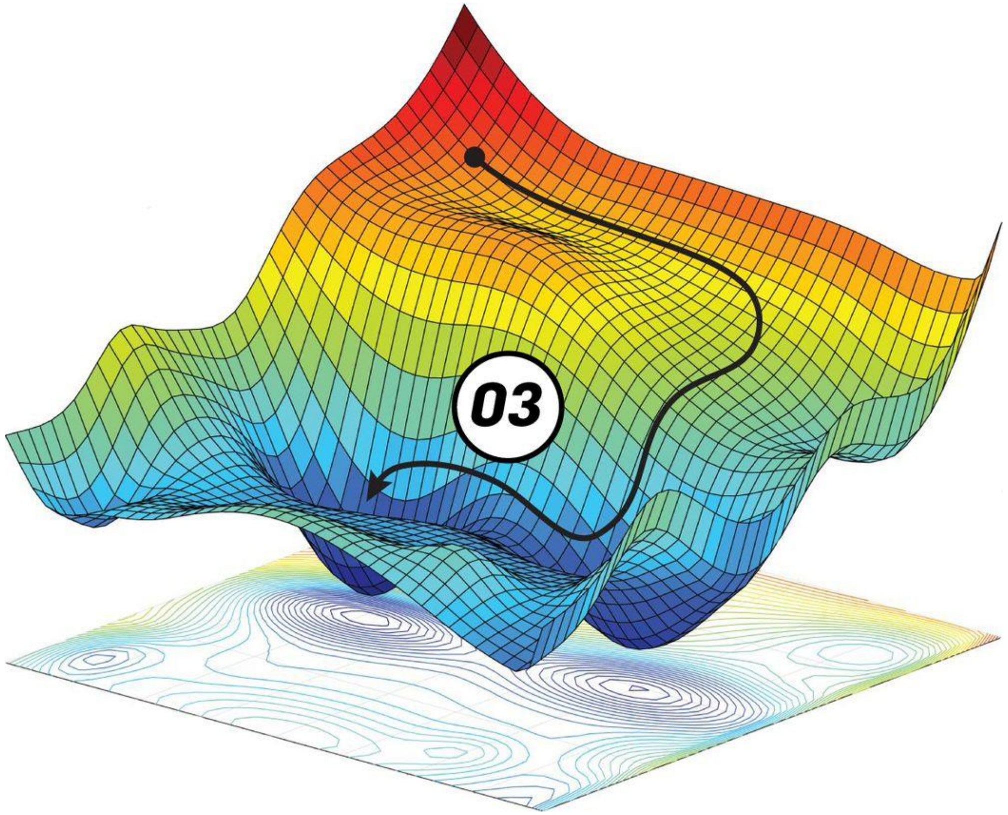
Classification of handwritten digits





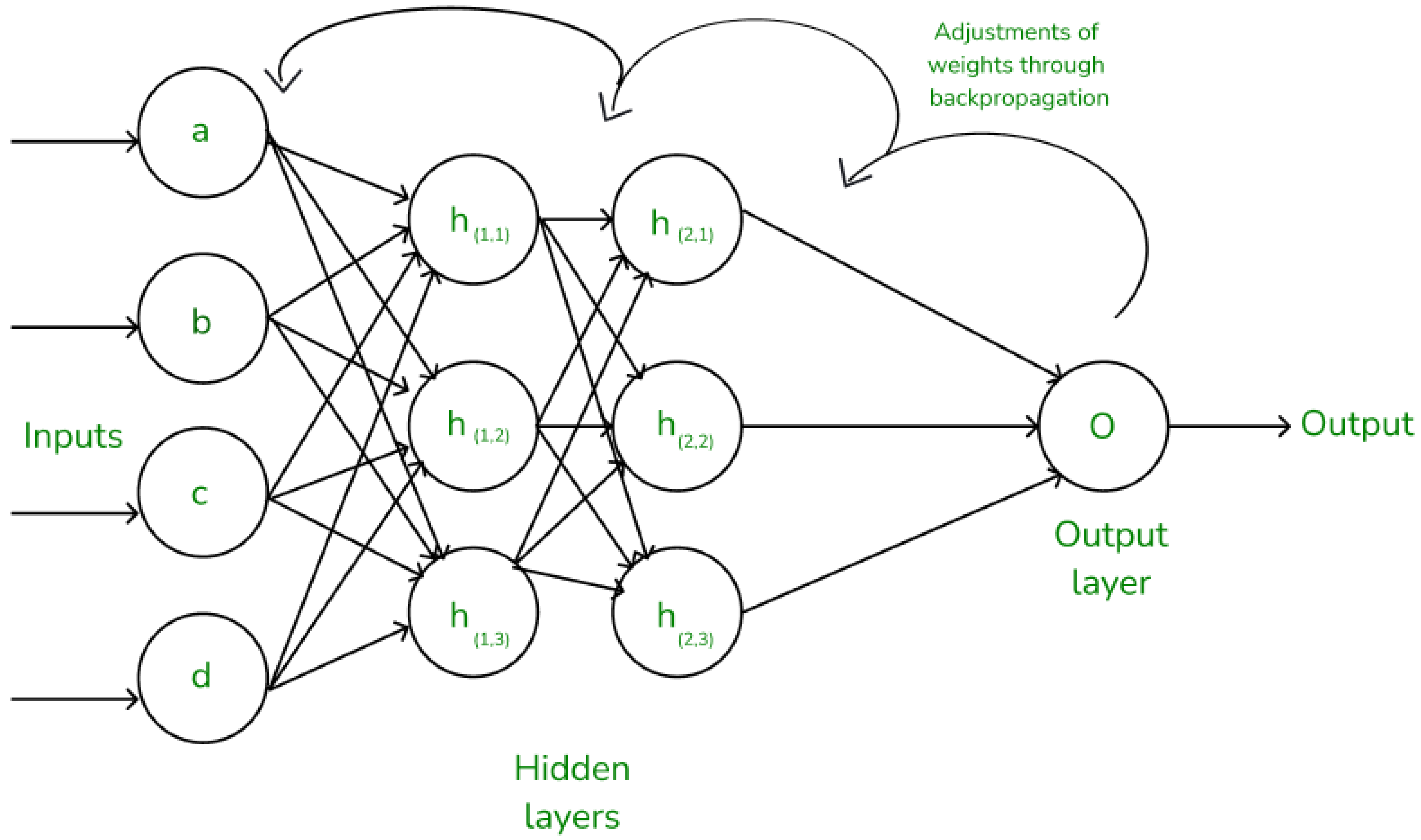
Basis - gradient descent





Gradient descent is an optimization algorithm which is commonly-used to train machine learning models and neural networks. It trains machine learning models by minimizing errors between predicted and actual results.

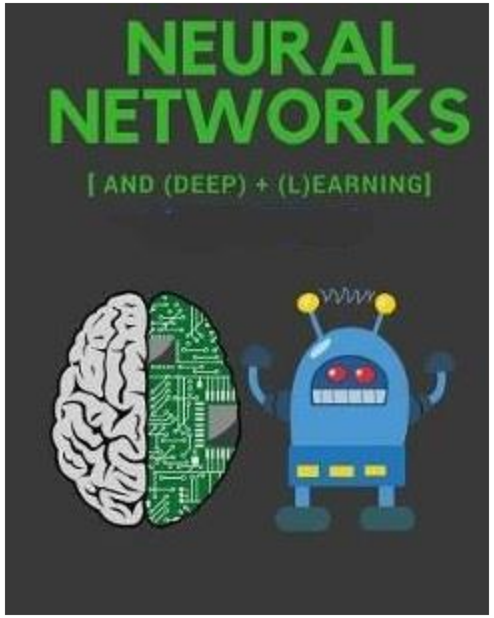
Basis - backpropagation



In machine learning, backpropagation is a gradient estimation method used to train neural network models. The gradient estimate is used by the optimization algorithm to compute the network parameter updates.

Neural Networks and Deep Learning 2019

Michael Nielsen

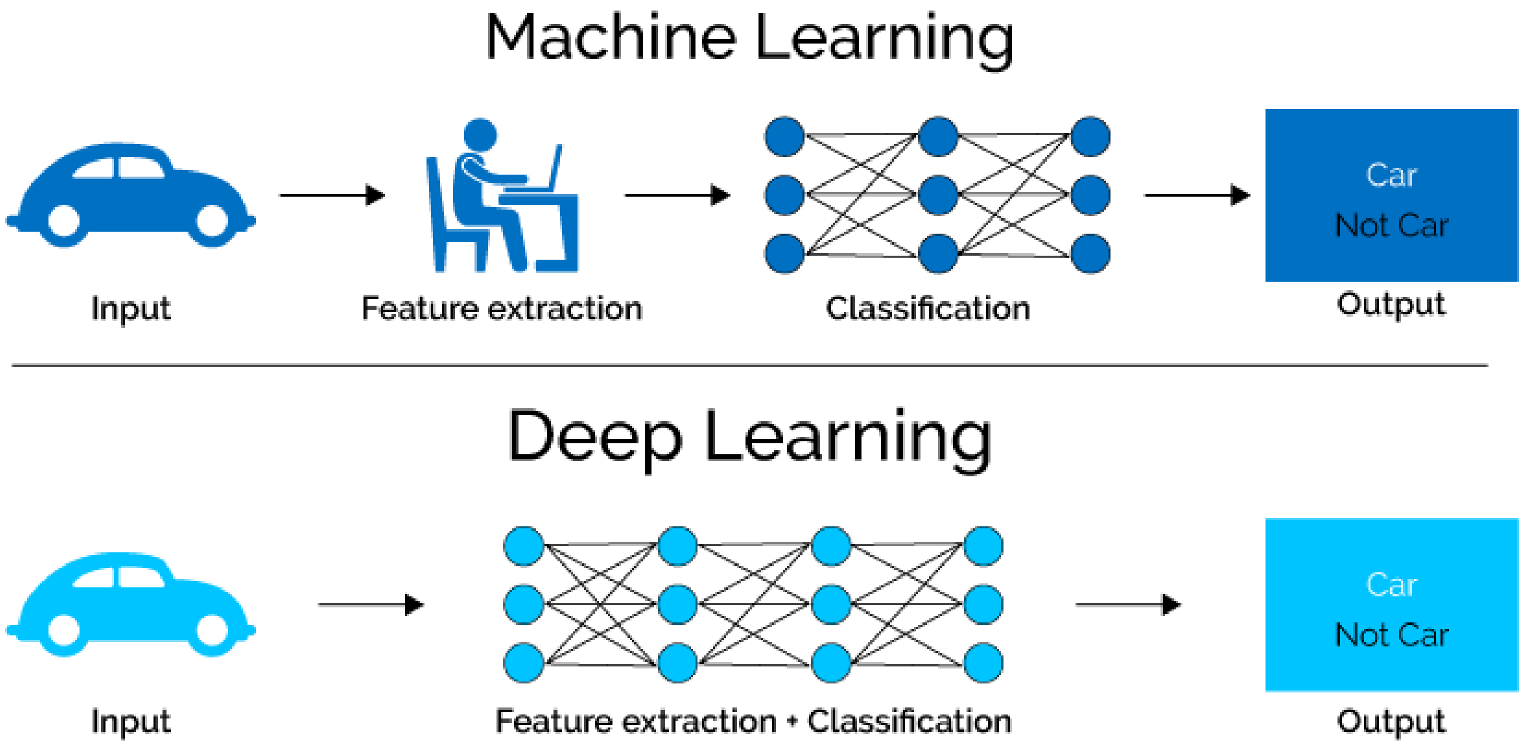


http://neuralnetworksanddeeplearning.com/index.html

ML vs. Deep Learning

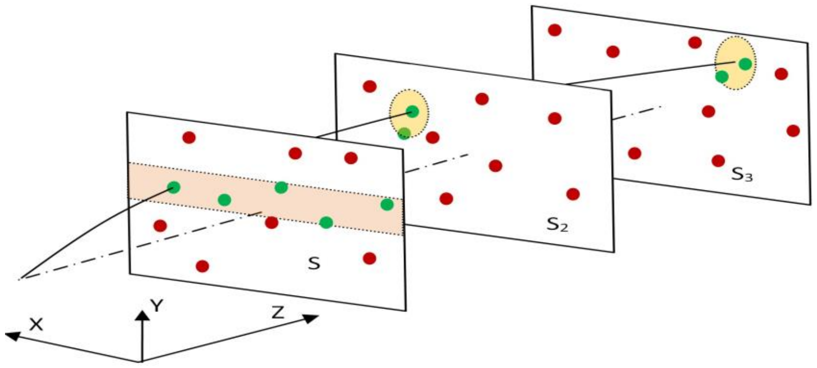
Deep learning (DL) is a machine learning subfield that uses multiple layers for learning data representations

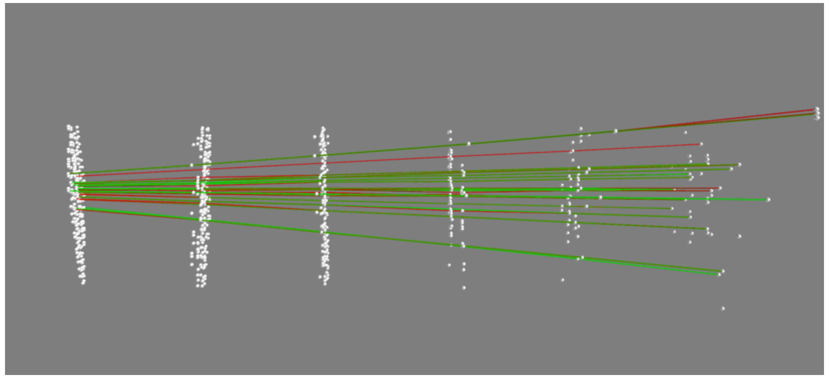
DL is exceptionally effective at learning patterns



HybriLIT & SUPERCOMPUTER “Govorun”

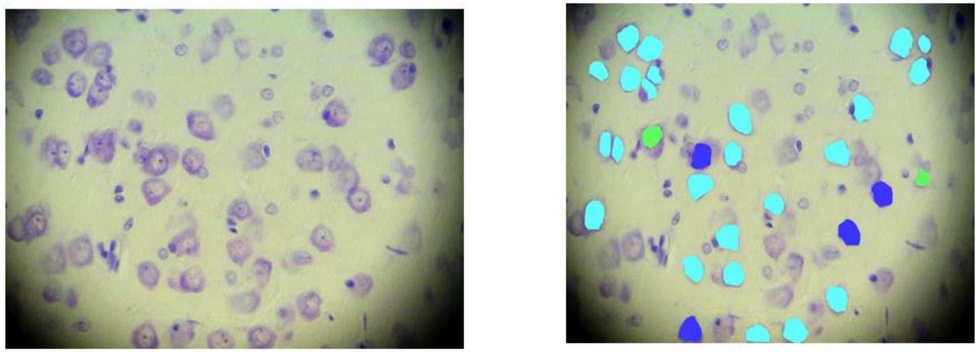
Использование нейросетей для трекинга в ФВЭ



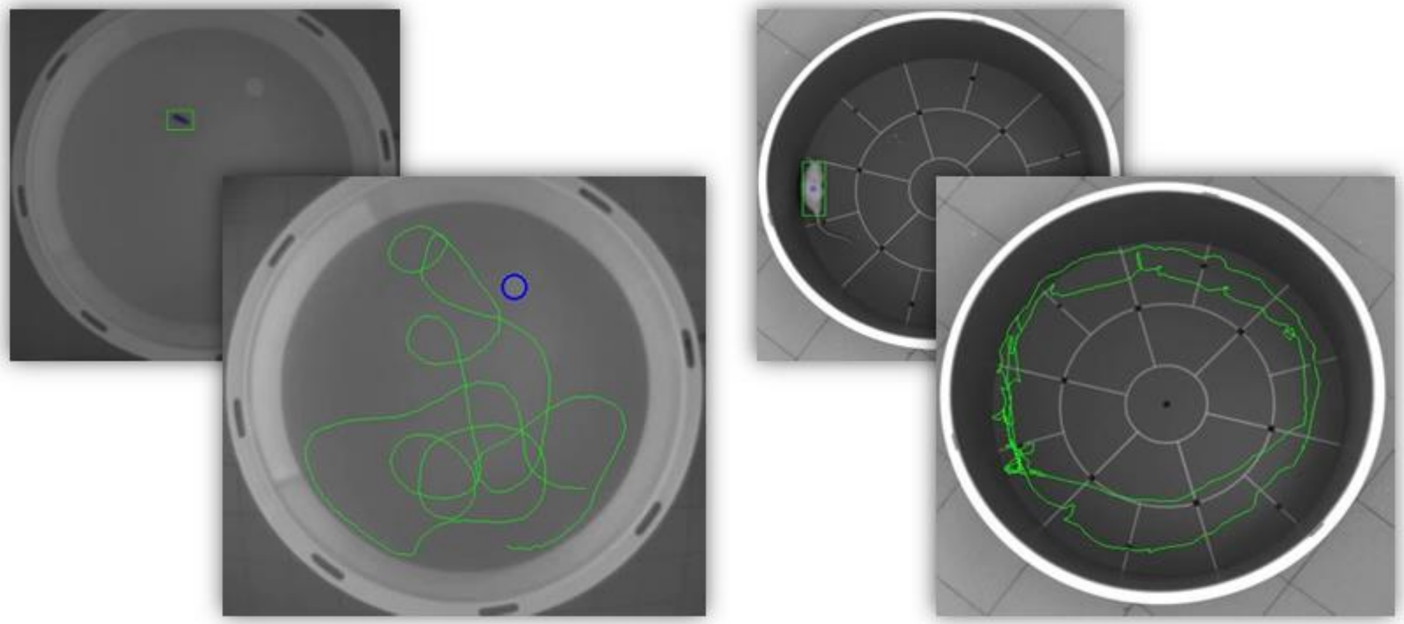


проф. Г. А. Ососков – следующая лекция

Radiation biology tasks



МОРФОЛОГИЧЕСКИЙ АНАЛИЗ



ПОВЕДЕНЧЕСКИЙ АНАЛИЗ

Контроль загрязнения тяжелыми металлами

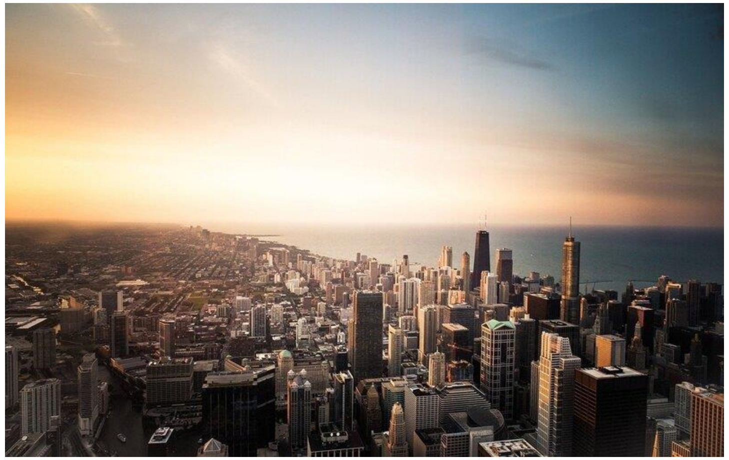
Environmental Concerns

The environmental problems like global warming, acid rain, air pollution, urban sprawl, waste disposal, ozone layer depletion, water pollution, climate change and many more affect every human, animal, and nation on this planet.

Over the last few decades, the exploitation of our planet and the degradation of our environment has gone up at an alarming rate. As our actions have been not in favor of protecting this planet, we have seen natural disasters striking us more often in the form of flash floods, earthquakes, blizzards, tsunamis, and cyclones.





Air pollution

Air pollution has a significant **negative impact** on the various components of ecosystems, **human health**, and ultimately, causes significant **economic damage**.

More than nine out of 10 of the world's population – 92% – lives in places where **air pollution exceeds safe limits**, according to research from the World Health Organization (WHO).

There are regional and international **environment control programs**. They use different techniques and tools but as a result, they all want to understand **what is the current situation** and how it will evolve.

Approaches

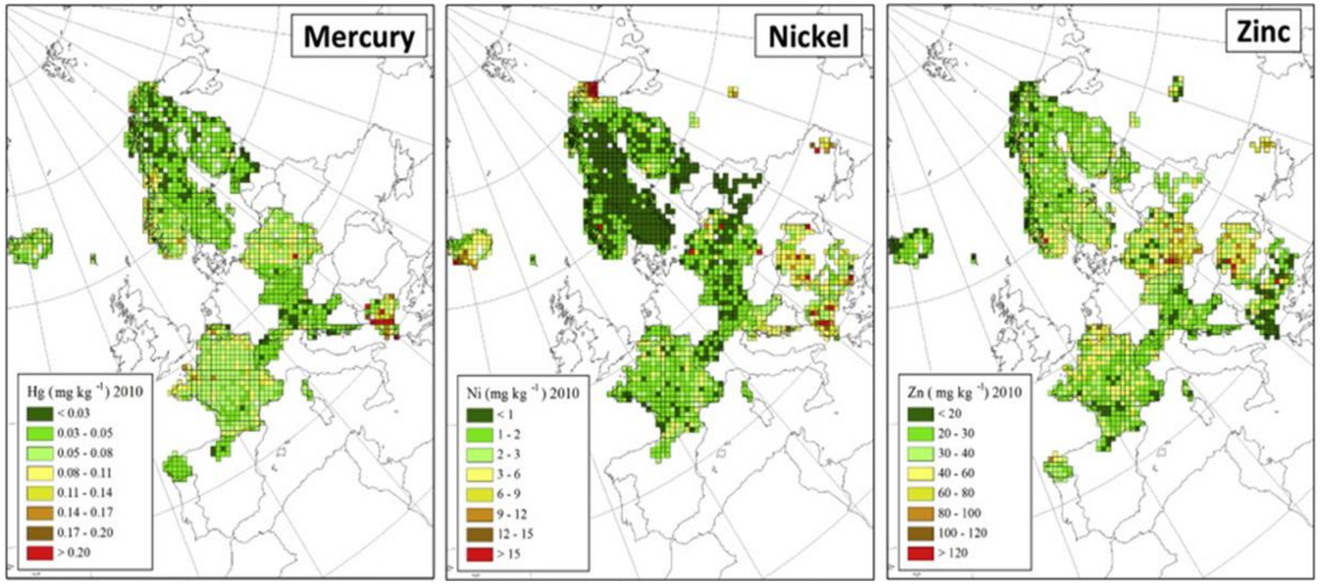
Generally, studies are based on the data obtained at the sampling sites in manual or automatic mode. The collected material is analyzed using various techniques in the field or in special laboratories. Air quality (AQ) monitoring stations provide information about regulatory air pollutants such as gaseous pollutants, PM, but rarely about heavy metals. To get detailed information samples should be processed in laboratories.

After collection the data are aggrigated and interpreted, and quite often the results are ambiguous and require the involvement of experts.

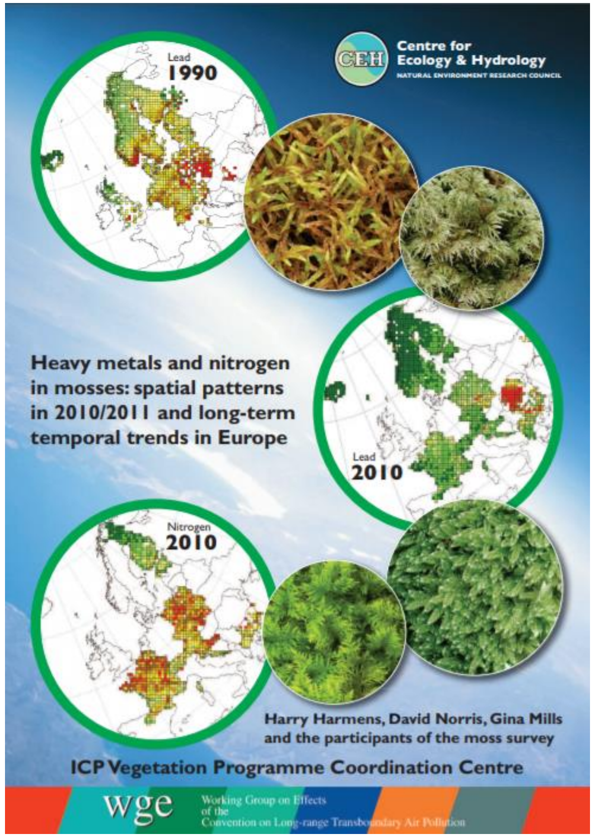
The level of automation and adoption of information technology in environmental monitoring programs is constantly increasing, although it lags far behind areas where the use of modern technology can lead to rapid economic impact.

ICP Vegetation

The aim of the **UNECE International Cooperative Program (ICP)** Vegetation in the framework of the United Nations Convention on Long-Range Transboundary Air Pollution is to **identify the main polluted areas of Europe**, produce regional maps and further develop the understanding of the long-range transboundary pollution. Atmospheric deposition study of heavy metals, nitrogen, persistent organic compounds (POPs) and radionuclides is based on the analysis of naturally growing mosses through moss surveys carried out **every 5 years**. The program is realized in **43 countries of Europe and Asia**. Mosses are collected at thousands of sites

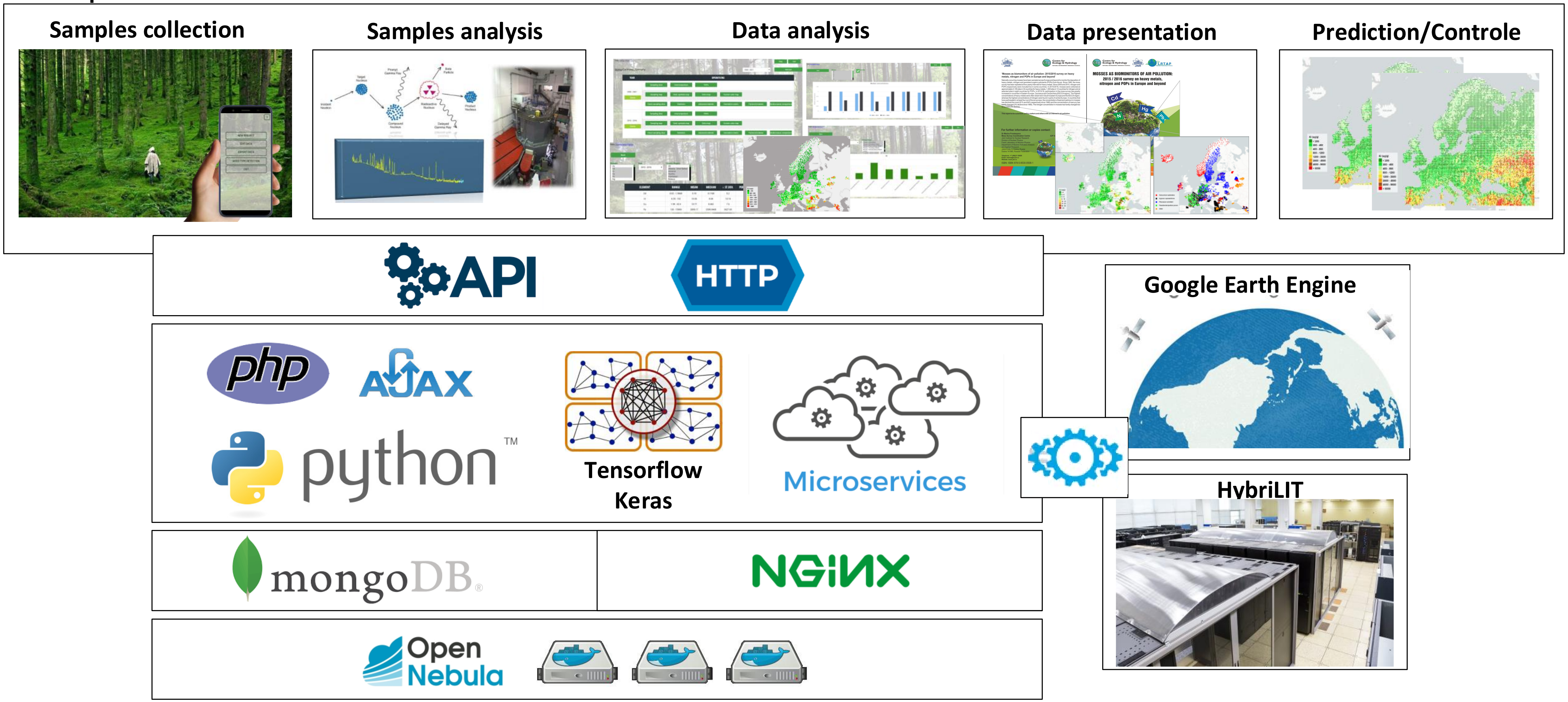


Examples of distribution maps in Atlas 2010



Since 2014 the JINR Frank Laboratory of Neutron Physics sector of neutron activation analysis is the **coordinator of the ICP** Vegetation program

The platform

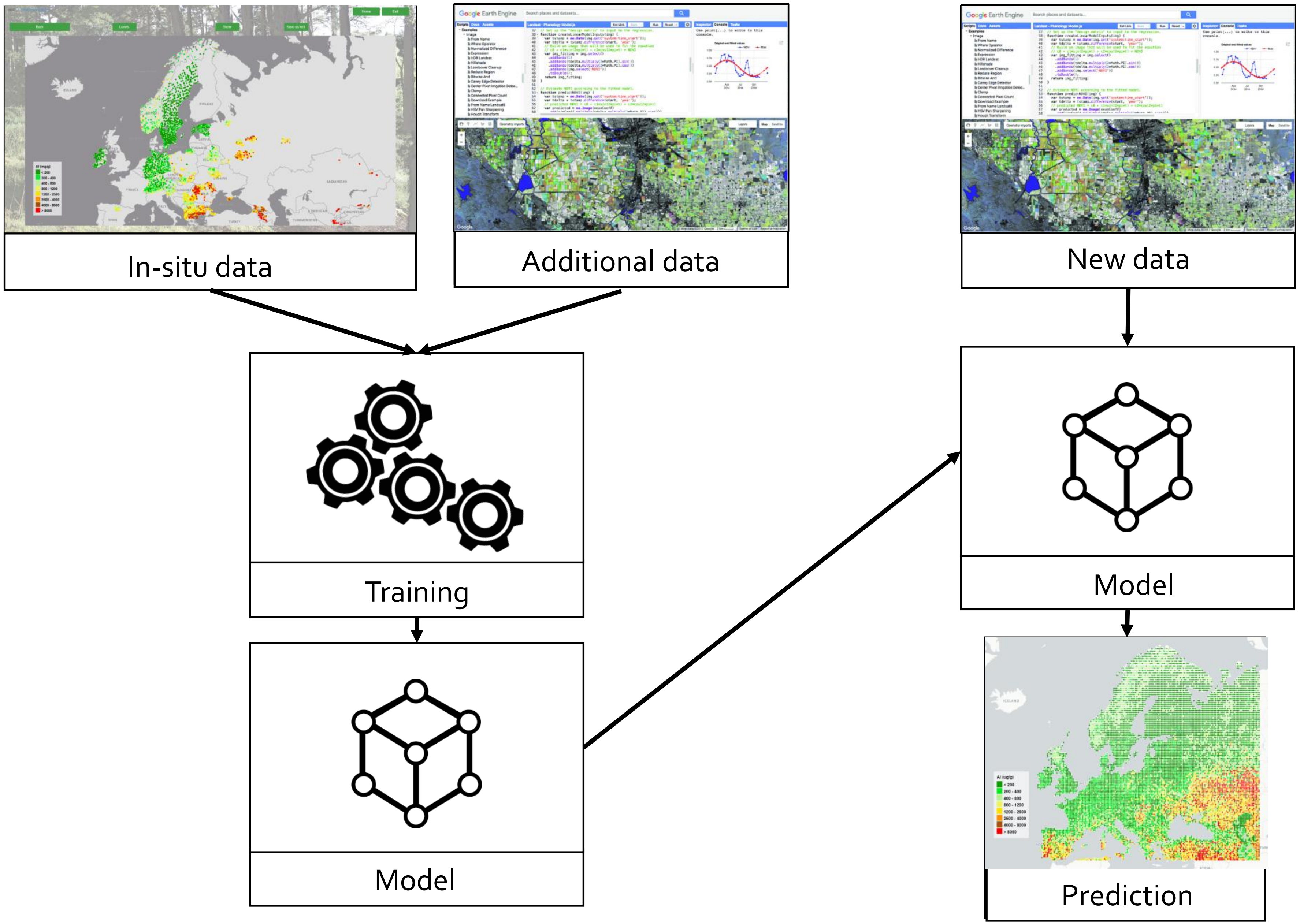


Since the launch of the first version of the platform, a mobile application has been developed to simplify the process of collecting and verifying data, deep learning models for image classification and pollution prediction based on remote sensing data, various functional blocks implemented in a microservice architecture to automate a number of operational tasks, and the analytical capabilities of the system are also expanded.

DMS. Atlas 2015-2016

The project, designed to automate typical operations with data and the creation of regional maps, absorbed more and more new approaches and technologies and now it may be classified as intelligent platforms.

Machine learning (Supervised learning)



Satellite programs

Source: NASA

The Earth Engine Data Catalog

Landsat & Sentinel 1, 2
10-30m, weekly

MODIS
250m daily

Vector Data
WDPA, Tiger

Terrain & Land Cover

Weather & Climate
NOAA NCEP, OMI, ...

... and upload your own vectors and rasters

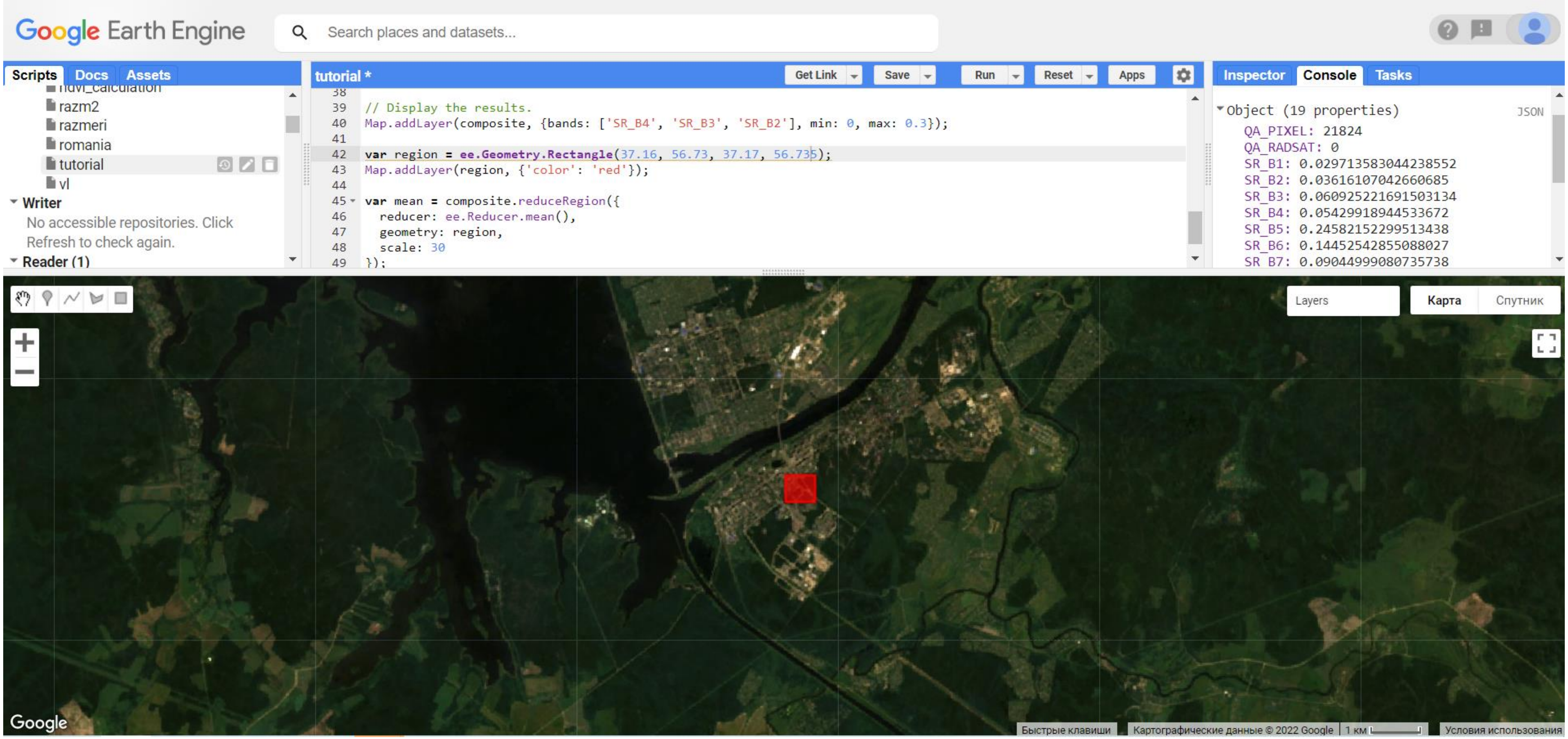
> 200 public datasets

> 5 million images

> 4000 new images every day

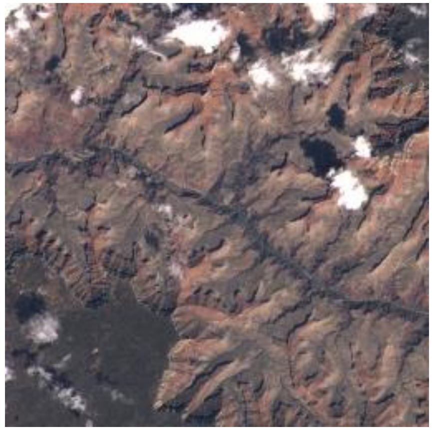
> 80 petabytes of data

Reducer in action



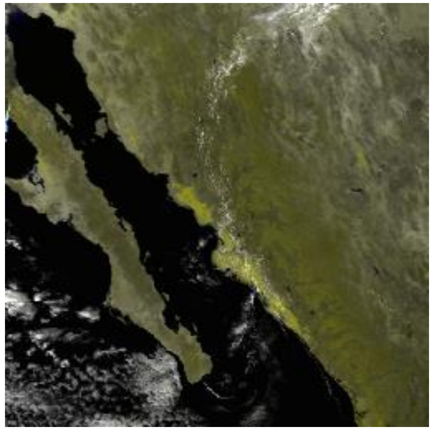
Google Earth Engine

There are more than **100 satellite programs** and modeled datasets. Google Earth Engine has **JavaScript online editor** to create and verify code and **python API** to communicate with user's applications.

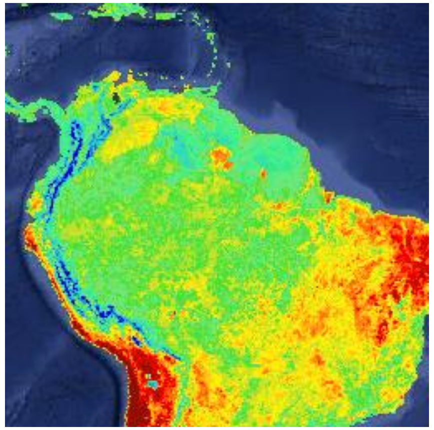


Landsat (15-30m Resolution)

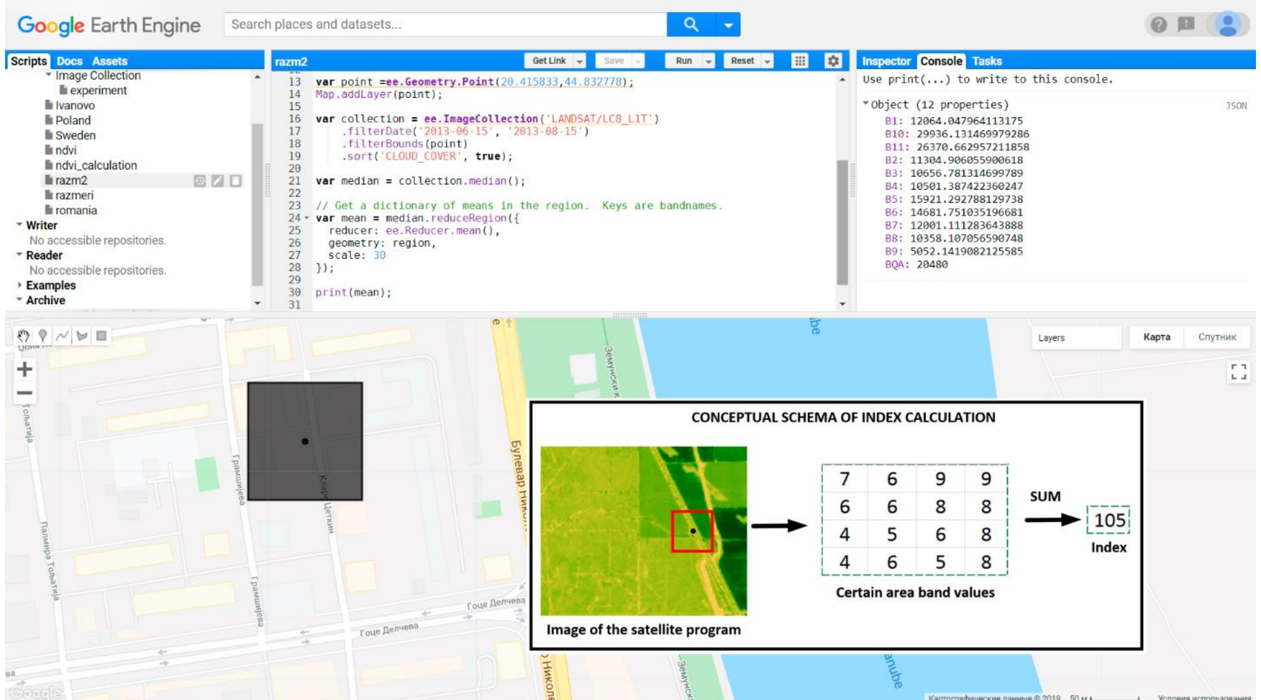
Modis (250-500m Resolution)



Sentinel (250-500m Resolution)

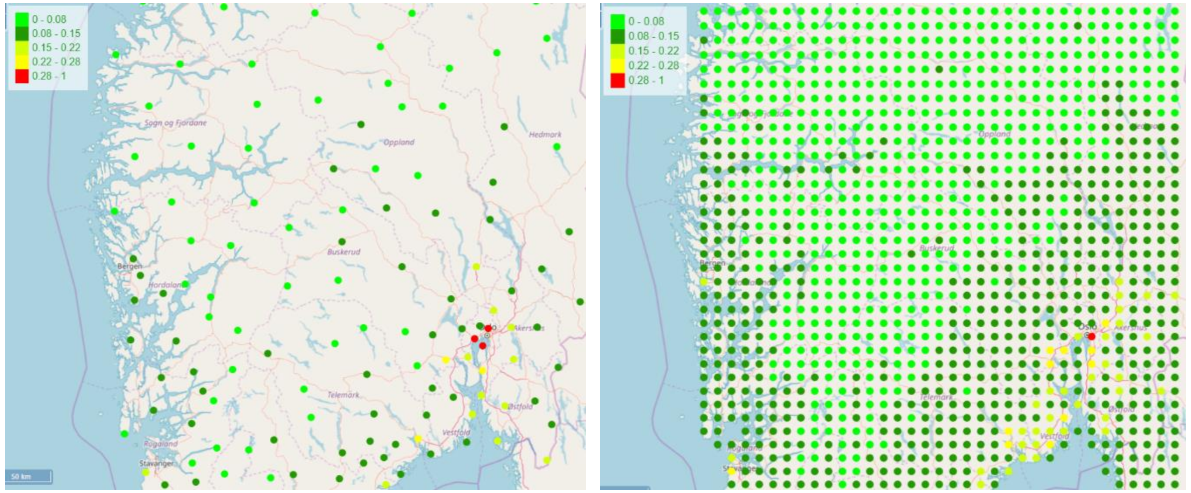


The MOD11A2 V6 average 8-day land surface temperature (LST) in a 1200 x 1200 kilometer grid.



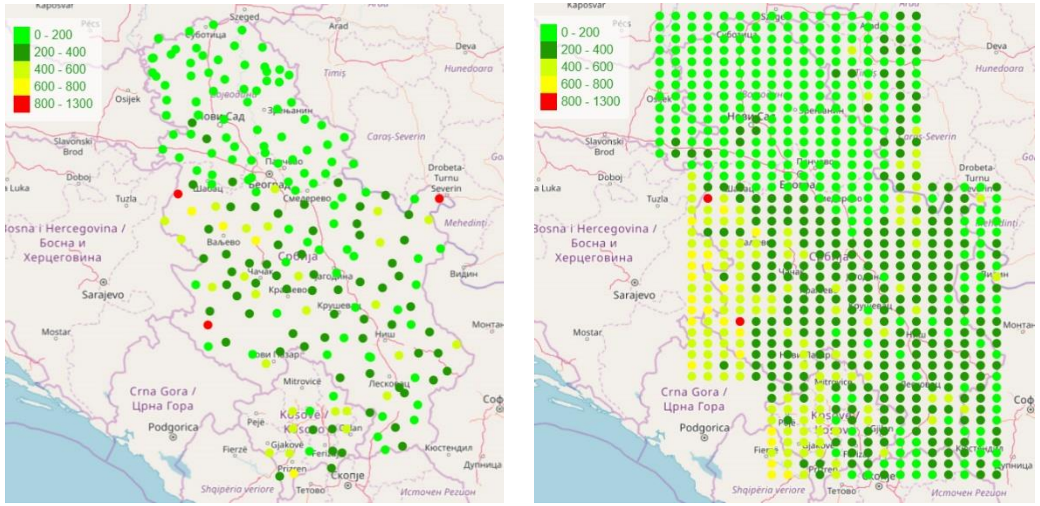
Specify program and time-period to get a collection of images, for example, program – "MODIS/006/MOD09A1" from 2013-06-15 to 2013-08-15 (the period relevant for in situ biomonitoring). Then, define the analyzed area, for example, a square kilometer, with center at the coordinates where sampling was performed. During the satellite data collection, under the bands (channels) of the median image, we execute some mathematical functions (max, min, median, etc.) and get the numerical values.

Results on the regional level

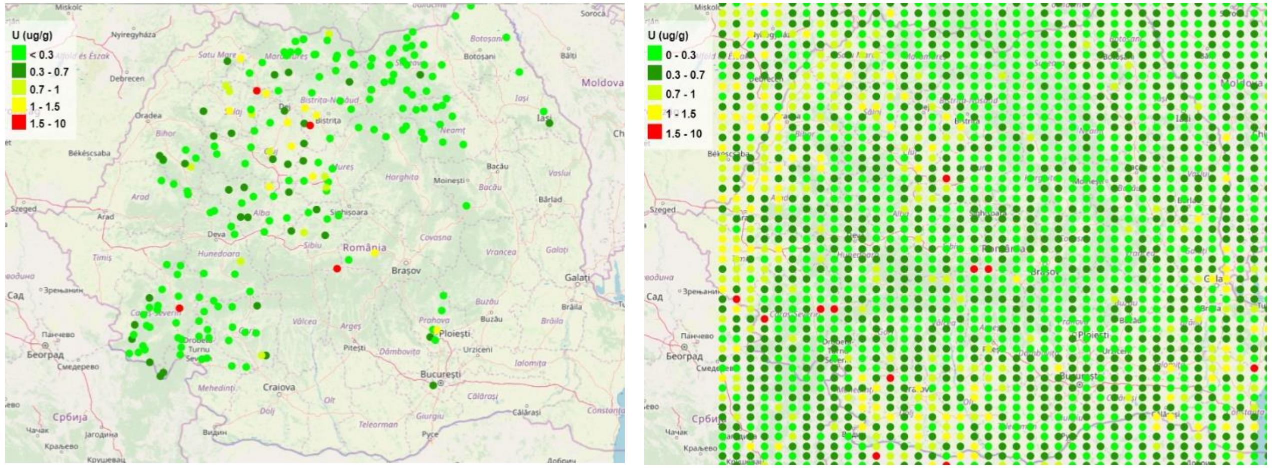


Sb at Norway. Left – real life, right - prediction

Candidates for modeling:
Al, As, Cr, Cu Fe, Mn, Ni, Pb, V, Sb, U ...



Mn at Serbia. Left – real life, right - prediction



U at Romania. Left – real life, right - prediction

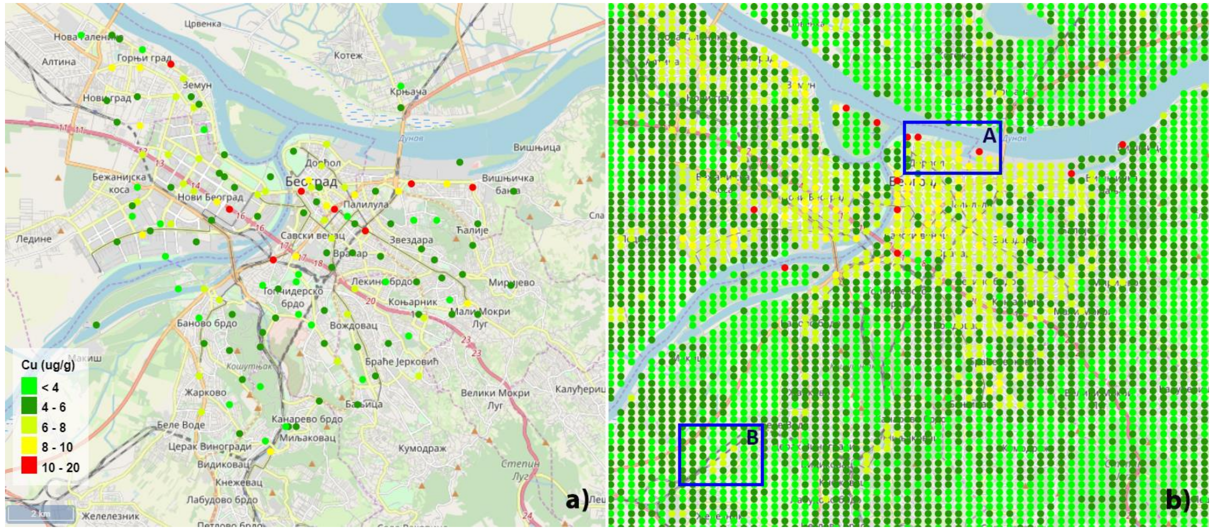


Figure 3. Concentration of Cu in the summer of 2013 (Belgrade): a) real measurements, and b) prediction values; area A represents central part of Old Belgrade with permanently high traffic flow; area B represents a large railway terminal

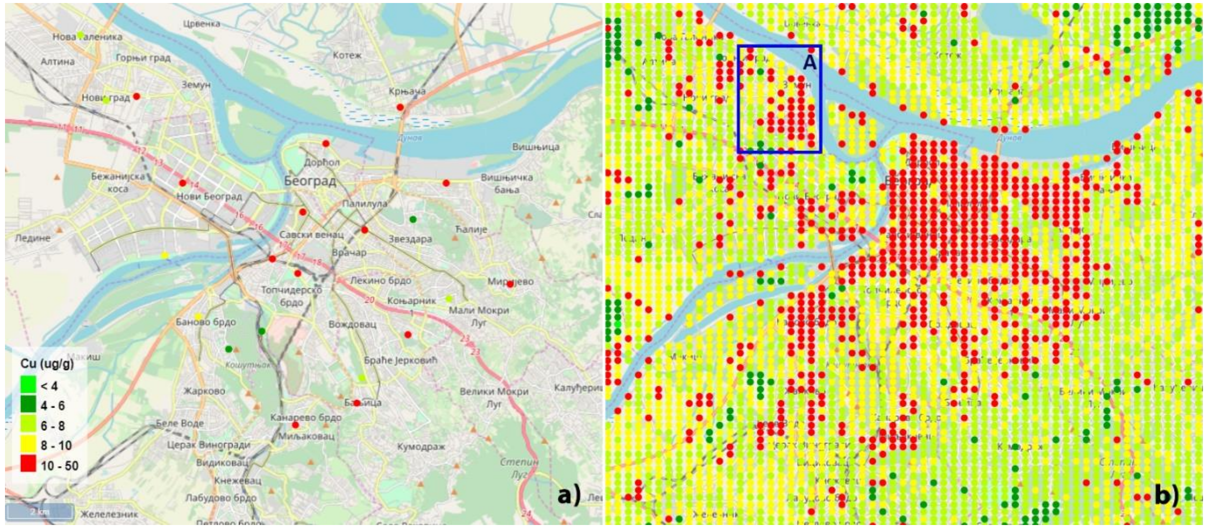


Figure 4. Concentration of Cu in the winter season 2013/2014 (Belgrade): a) real measurements, and b) prediction values; area A represents an old city core highly polluted in winter season

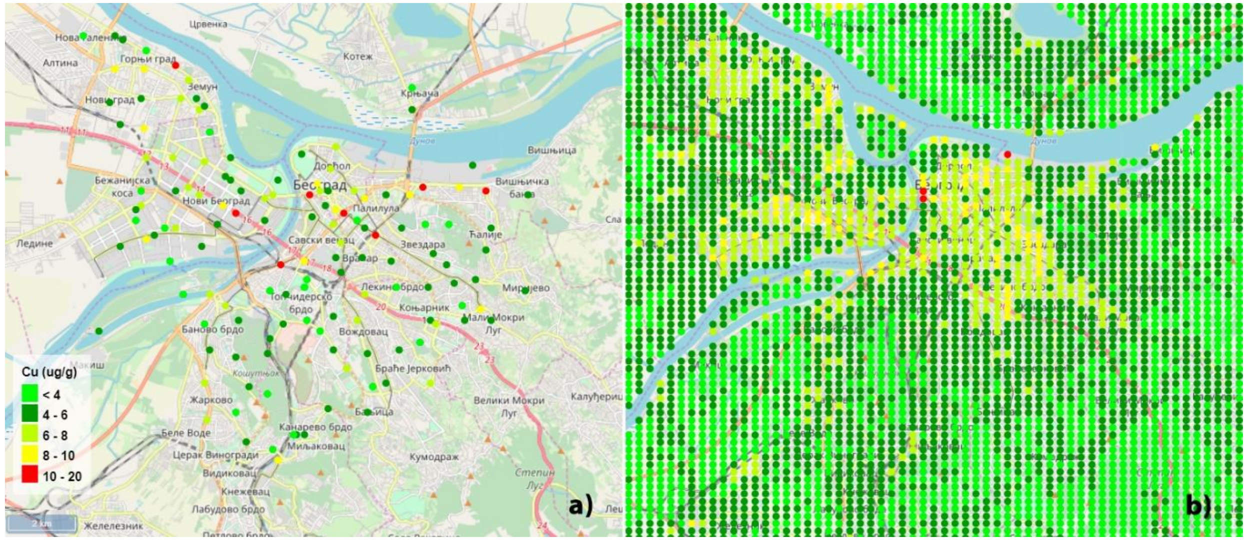
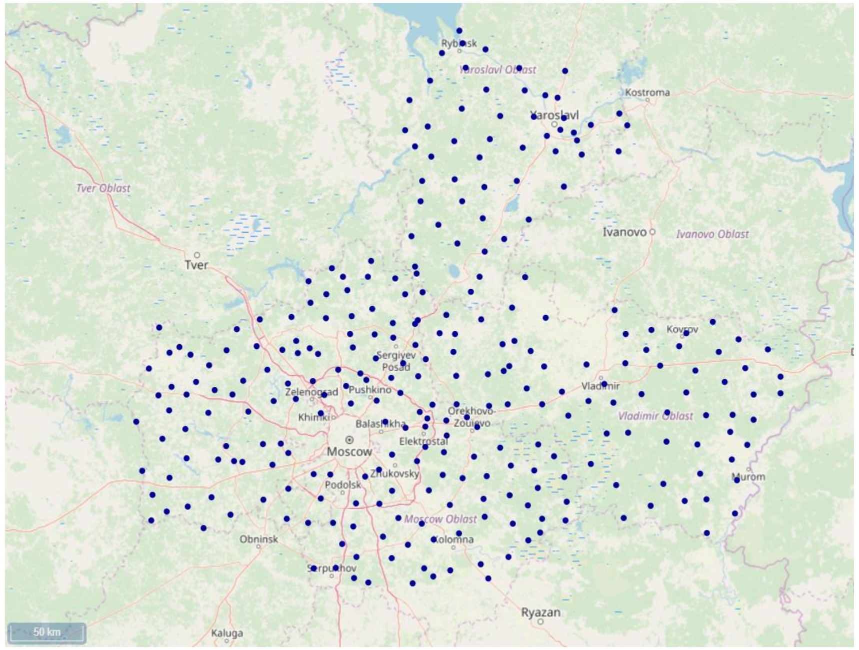


Figure 5. Concentration of Cu in Belgrade: a) biomonitoring measurements in the summer of 2013, and b) prediction for 2018

Prediction of air pollution by potentially toxic elements over urban area by combining satellite imagery, Moss Biomonitoring Data and Machine Learning
A. Uzhinskiy,
M. Anitich Urorevice, M. Frontasyeva, Ciencia e Tecnica Vitivinicola Journal, ISSN:2416-3953, 12, 35, 2020

Machine learning and neural networks



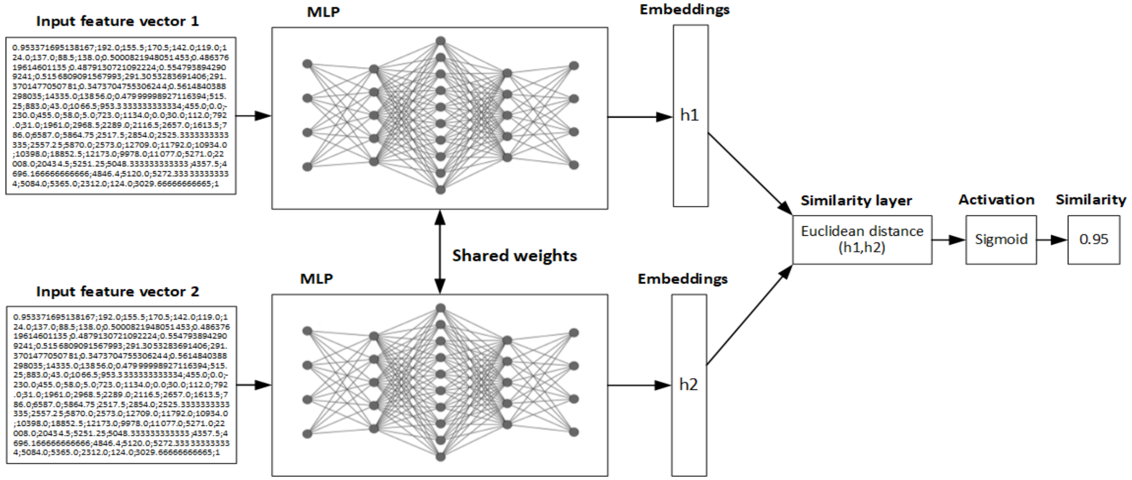
We use the information on 73, 53, and 156 samples from the Vladimir, Yaroslavl, and Moscow regions gathered in 2018 - 2019.

The indices are gathered based on data from 13 programs for 281 sampling sites, and their linkage with the concentration of 18 heavy metals is verified. Altogether 9 HMs, i.e., Al, Fe, Sb, Na, Sc, Sm, Tb, Th, and U, look very prospective for modeling.

We examine three approaches: Gradient Boosting, Multilayer perceptron, and Siamese network.

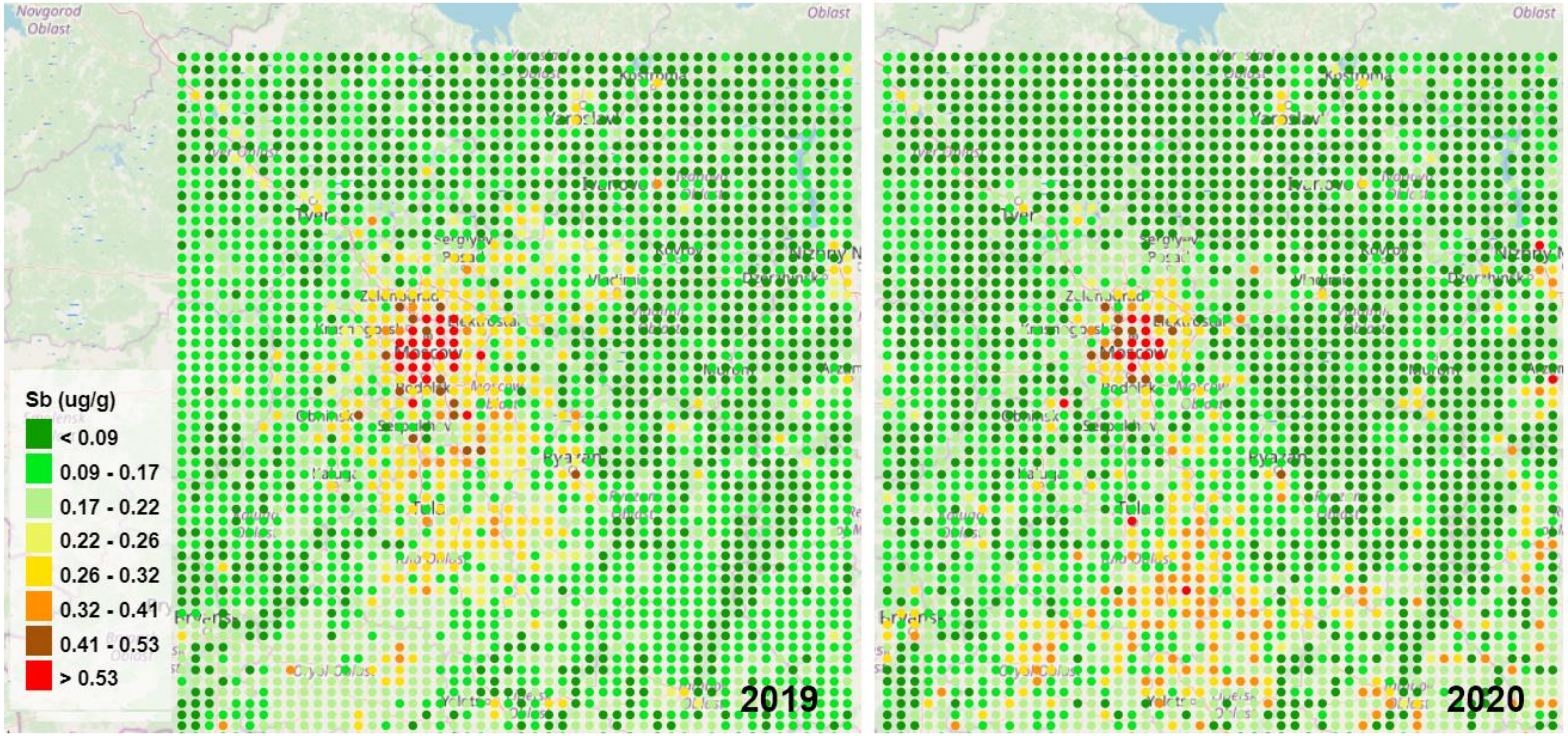
	Al		Fe		Sb	
	Acc si	Acc ai	Acc si	Acc ai	Acc si	Acc ai
GB	0.91	0.92	0.92	0.93	0.94	0.94
MLP	0.89	0.91	0.92	0.92	0.89	0.92
SNN	0.92	0.93	0.93	0.93	0.93	0.94

Table 2. Mean accuracy of the models. GB is gradient boosting. MLP is the multilayer perceptron. SNN is the Siamese neural network. Acc Si is the accuracy on the selected indices. Acc Al is the accuracy on all indices.



Siamese network architecture

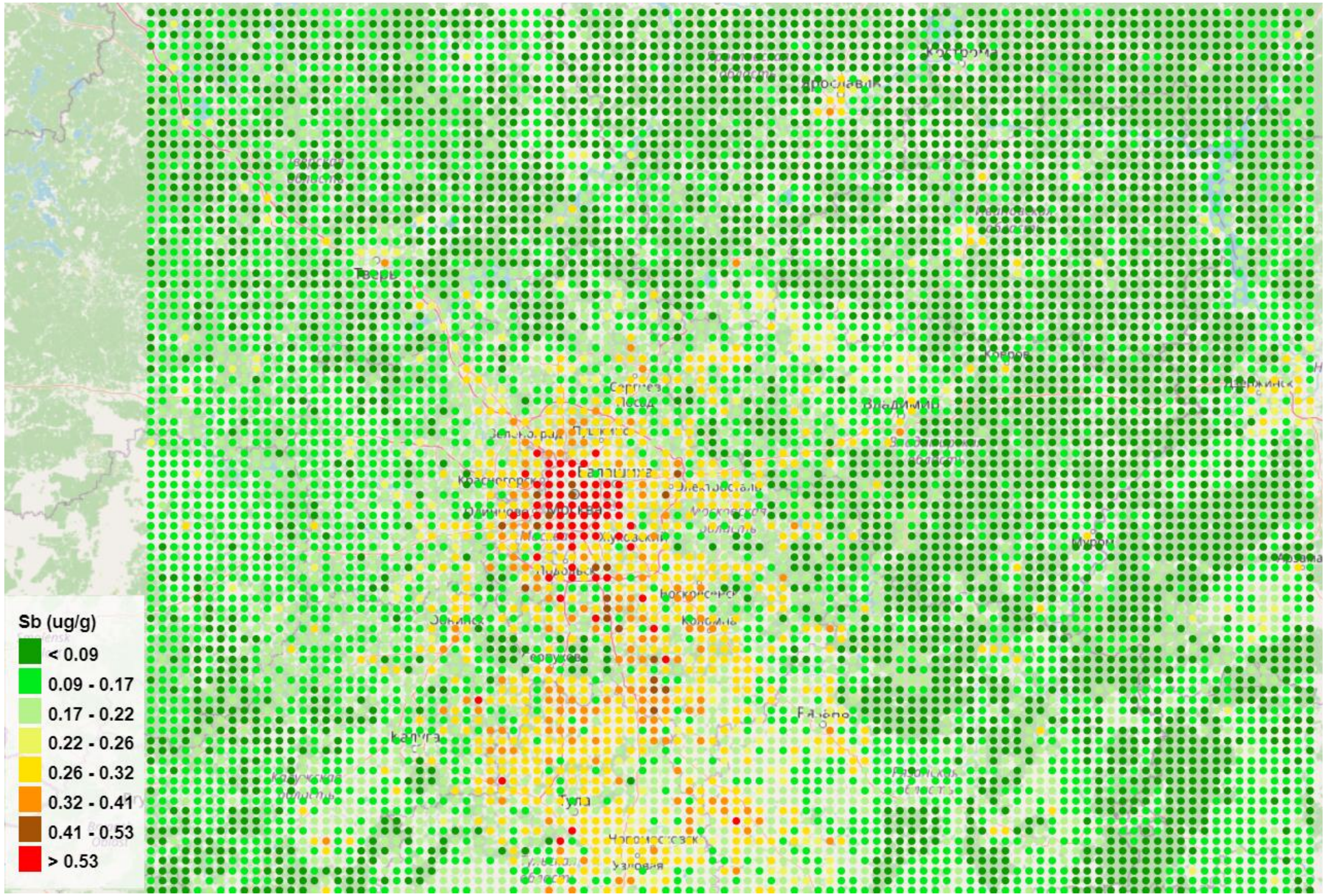
Results (2019 – 2020)



Sb contamination prediction for 2019 (left) and 2020 (right)

The lockdown in Russia that lasted for approximately 1.5 months imposed different limitations. Most of the limitations restricted the movement activities of the population. According to the official statistics, industrial production in Russia decreased by 2.9% from the past, by the end of 2020.

Results (High spatial resolution)



High spatial resolution of the SNN model prediction of Sb contamination

Moscow is a thickly populated city, and the population is increasing at a fast pace. Published information reveals, there are about 12.5 million habitants in Moscow. Therefore the Sb contamination level there is bound to be very high.

The map also reveals clusters of hot spots in large cities, such as Tula, Kaluga, Vladimir, Tver, Nizhny Novgorod, Yaroslavl, etc. It is also seen that from Sergiyev Posad to the north direction, the contamination level is rather low, except Yaroslavl, where the working oil refinery is located.

The Tula region stands out on the map. There is a multitude of industrial enterprises located in the region, i.e., chemical, metallurgical, and machine-building, besides several large thermal power plants. Huge transport nodes and federal freeways are seen, rather clearly, on the map.

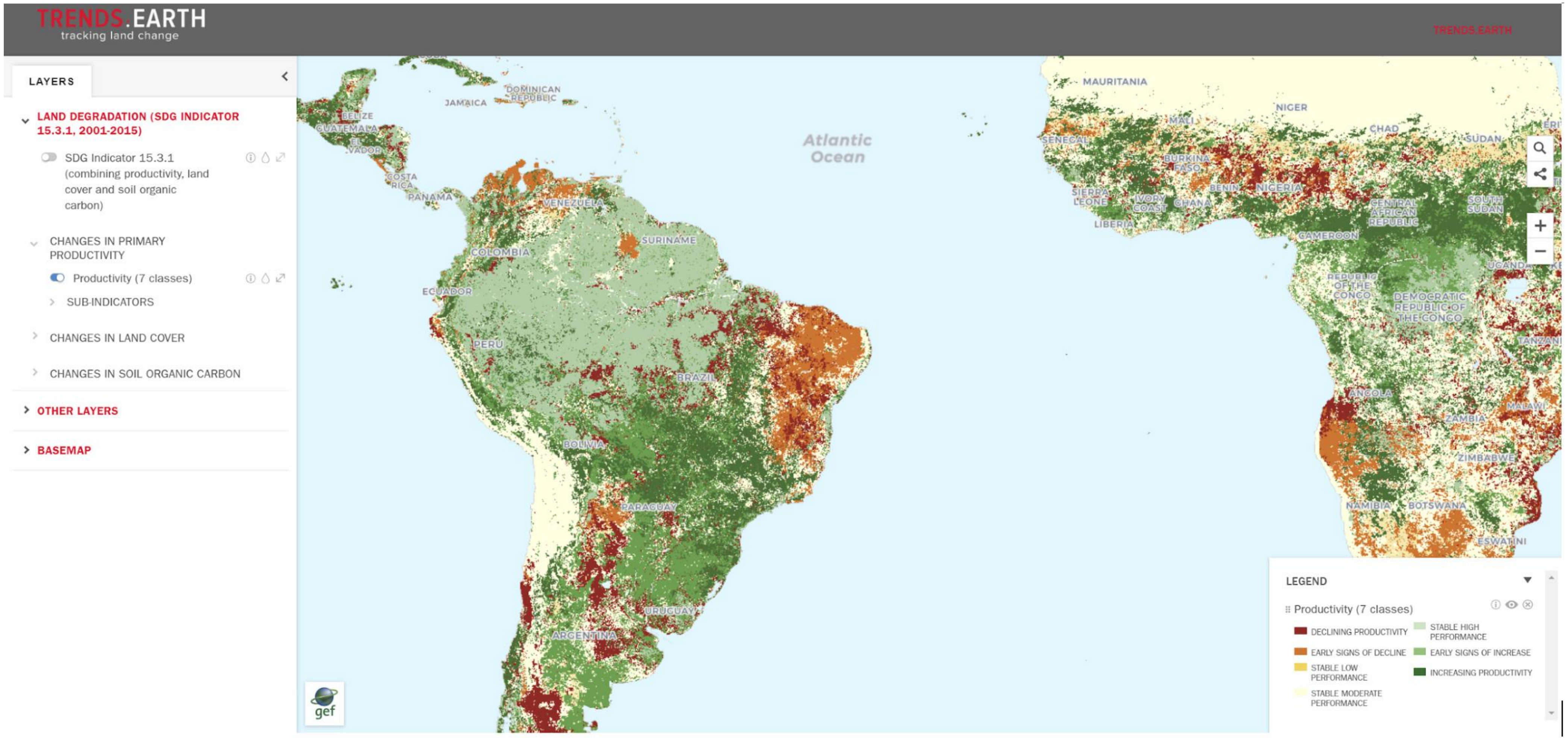
Uzhinskiy A.; Vergel K.; Central Russia heavy metal contamination model based on satellite imagery and machine learning, Computer Optics 2023; 47(1): 137-151. DOI: 10.18287/2412-6179-CO-1149.

В настоящее время ведутся исследования в области прогнозирования загрязнения почвы в долинах Нила (Египет)

Alternative projects

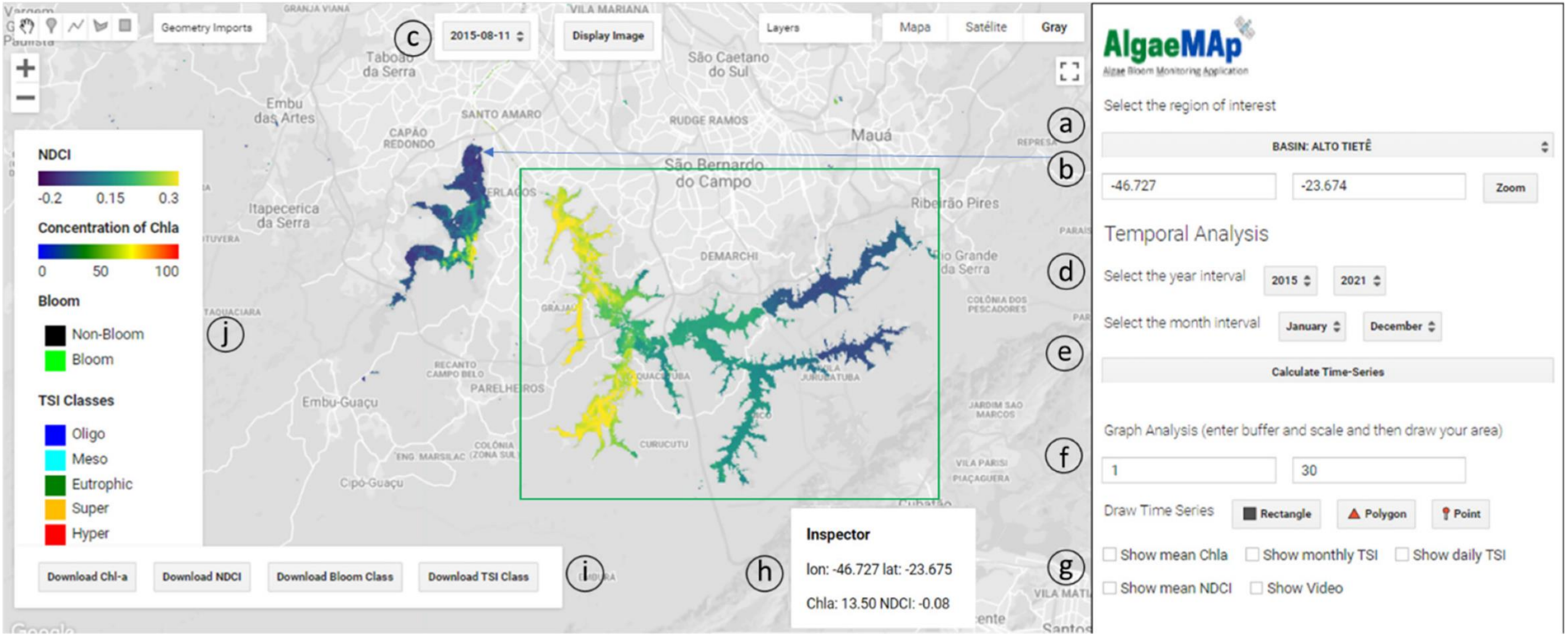
Trends.Earth

Through this project, we will develop a cloud-based platform dedicated to mapping land degradation which identifies potential land restoration opportunities at national to regional scale, allowing communities to prioritize areas to protect, manage, and restore in order to achieve land degradation neutrality.



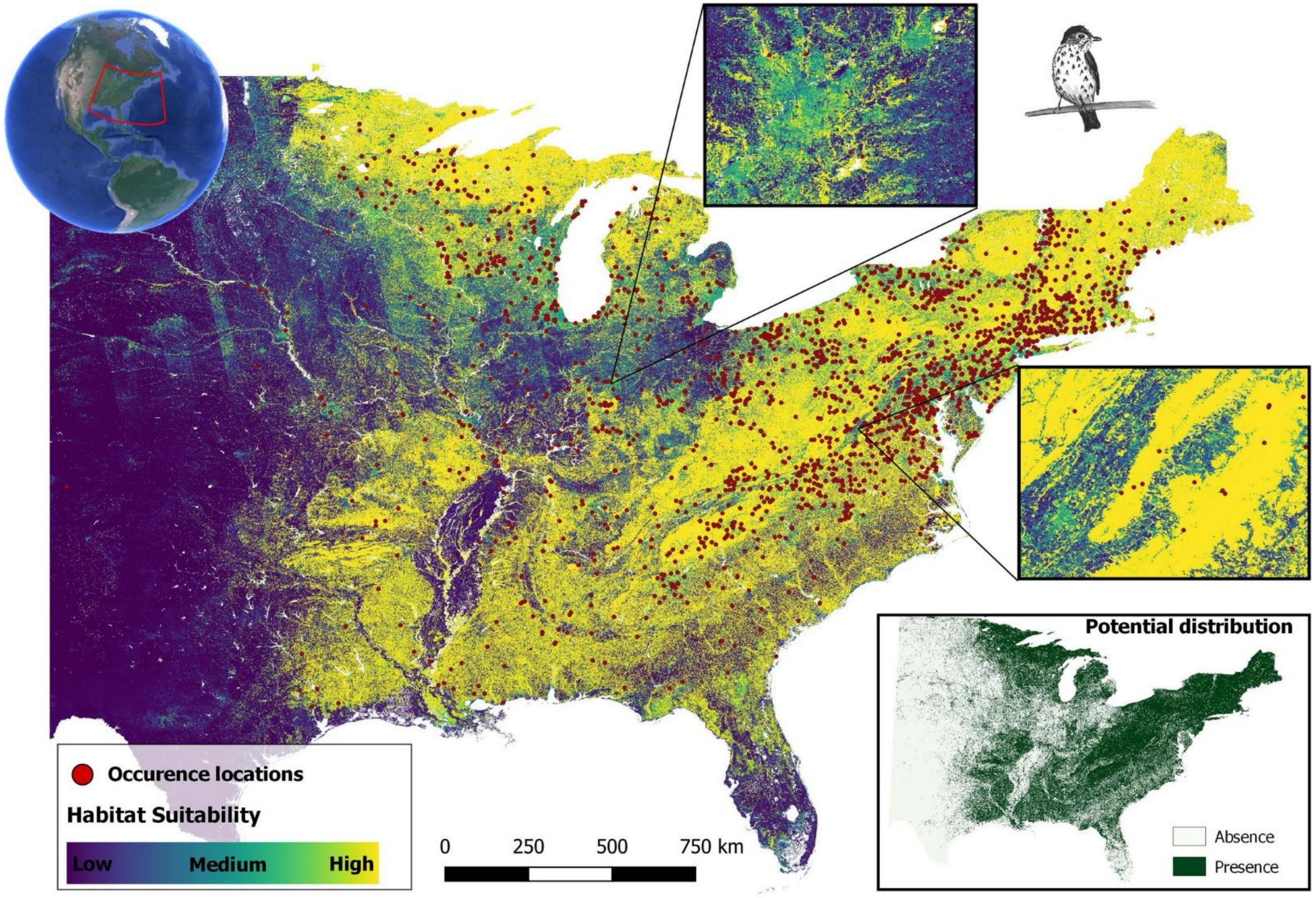
Alert System for Algal Bloom

Algae blooms occur when certain kinds of algae grow very quickly, forming patches, or "blooms," in the water. These blooms can be indicators of water degradation and emit powerful toxins that can endanger human and animal health.



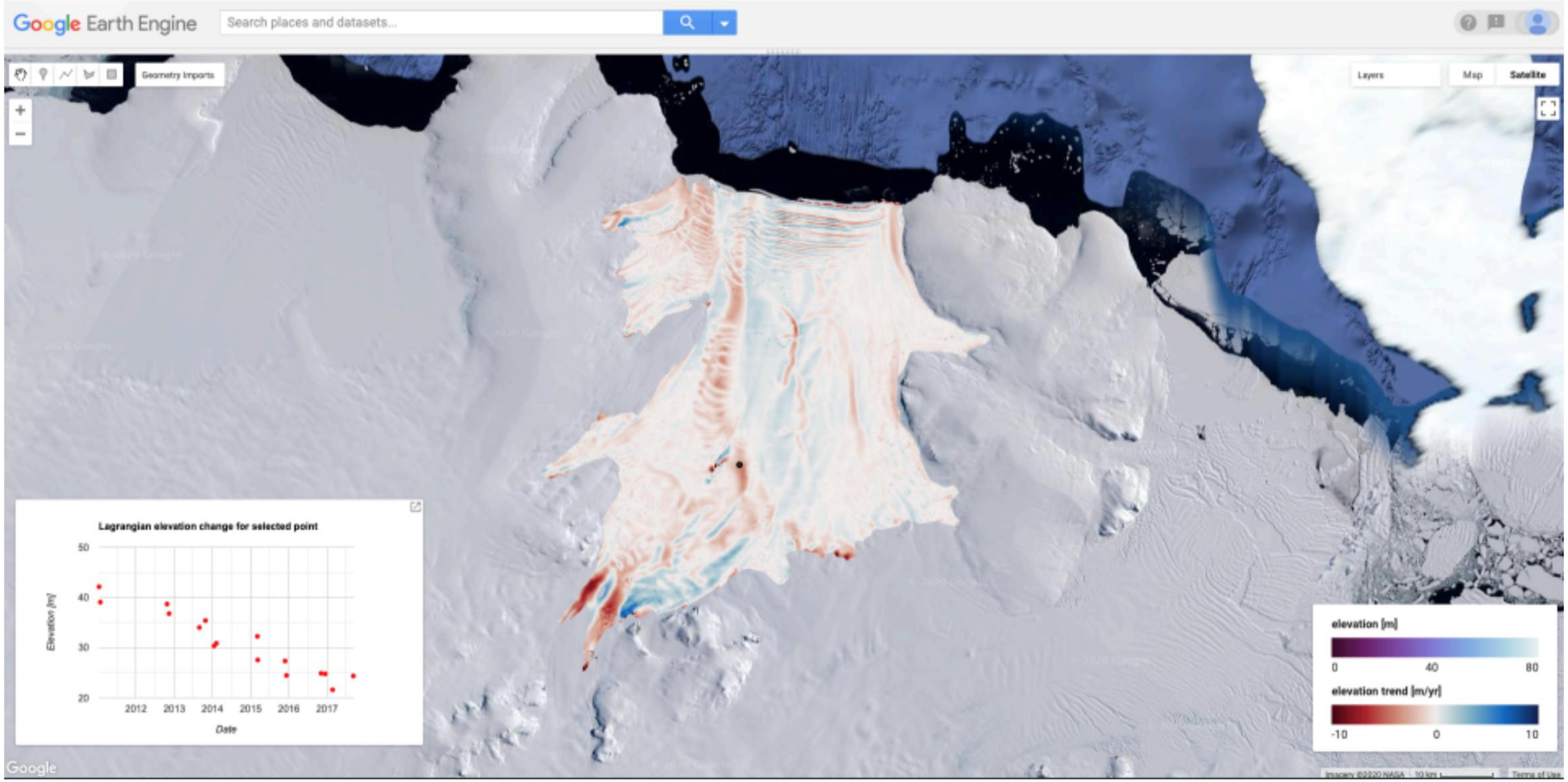
Essential Biodiversity Variables - ScaleUp

We implemented a workflow for species distribution modelling in GEE that includes importing species occurrence data into the GEE platform, selecting and preparing predictor variables, and performing model fitting with spatial or temporal split-block cross-validation techniques.



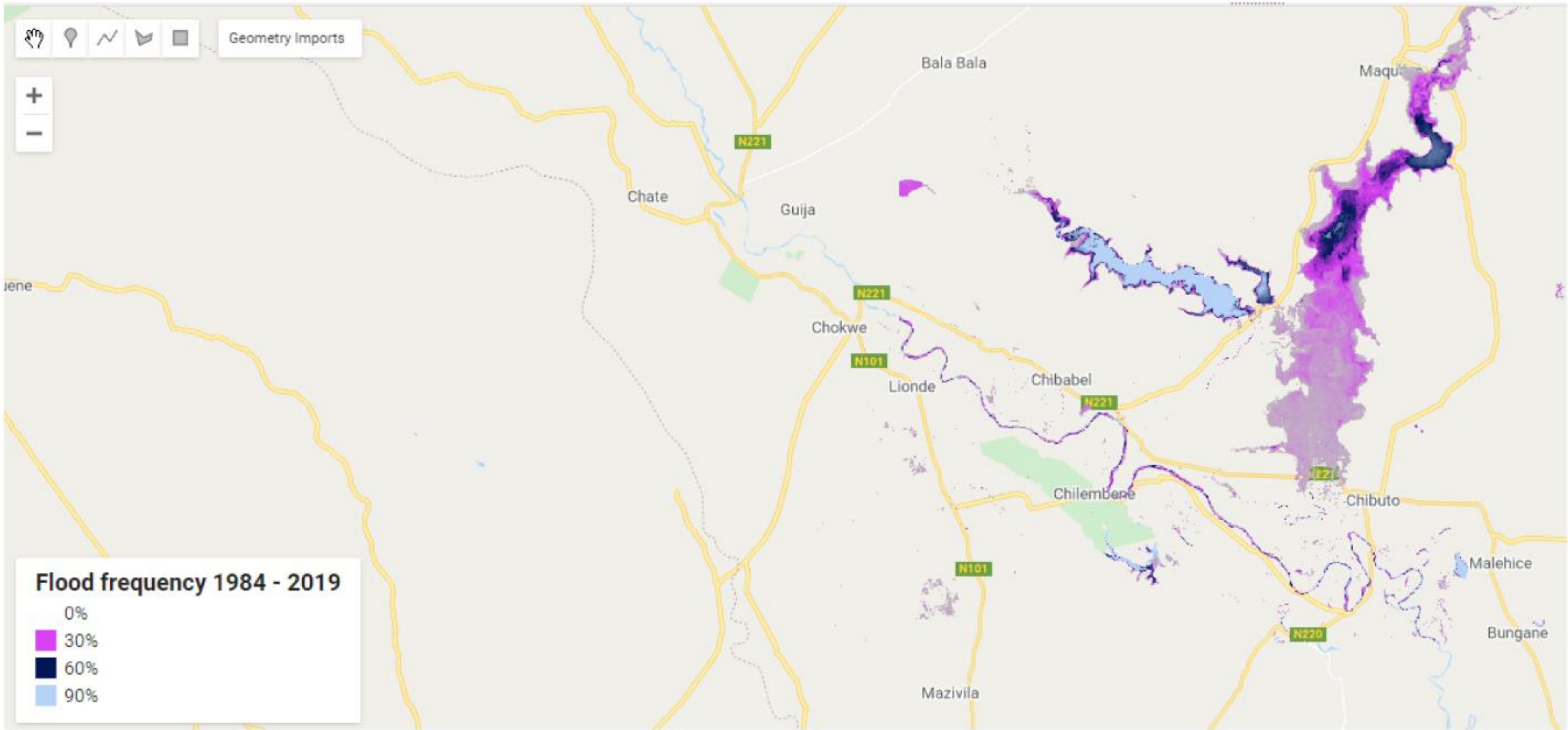
GEE-interface for analyzing basal melt over Dotson ice shelf

A platform to analyze and demonstrate the use of satellite data over Antarctic ice shelves. This will allow insight in the surface, subsurface and basal conditions of these ice shelves, which are major sources of uncertainty in future sea level projections from Antarctica.



Supporting water and disaster risk management in Mozambique

The project HydroPC focused on co-development and application of innovative data technologies and comprehensive training of beneficiaries to support water and disaster risk management in Mozambique.

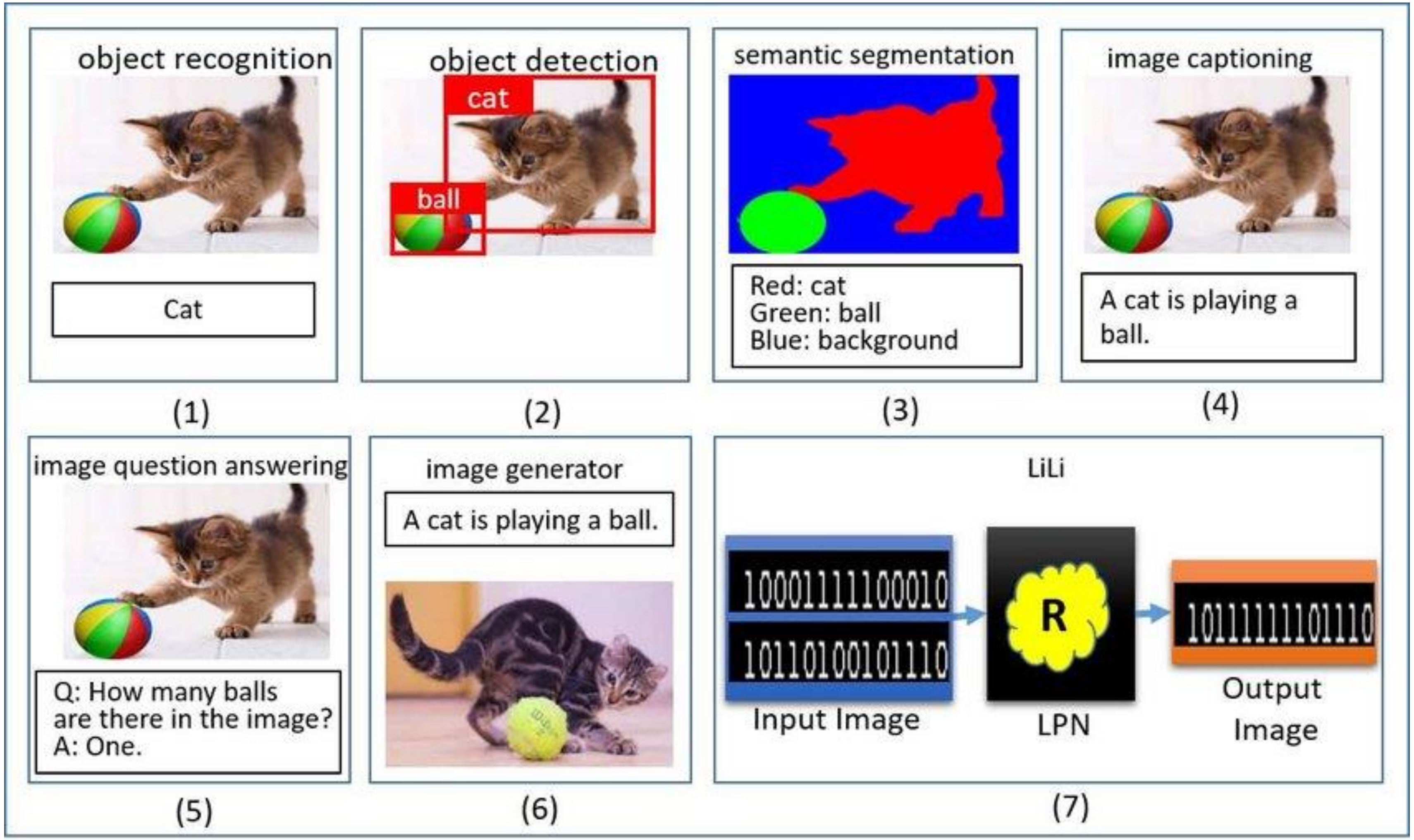


Что можете сделать вы?

- Проект по контролю за качеством воздуха с использованием открытых данных с метео-станций
- Проект по контролю состояния водных ресурсов
- Проект по отслеживанию изменений (леса, реки, использование земельных ресурсов и т.д.)
- Ваш вариант ...

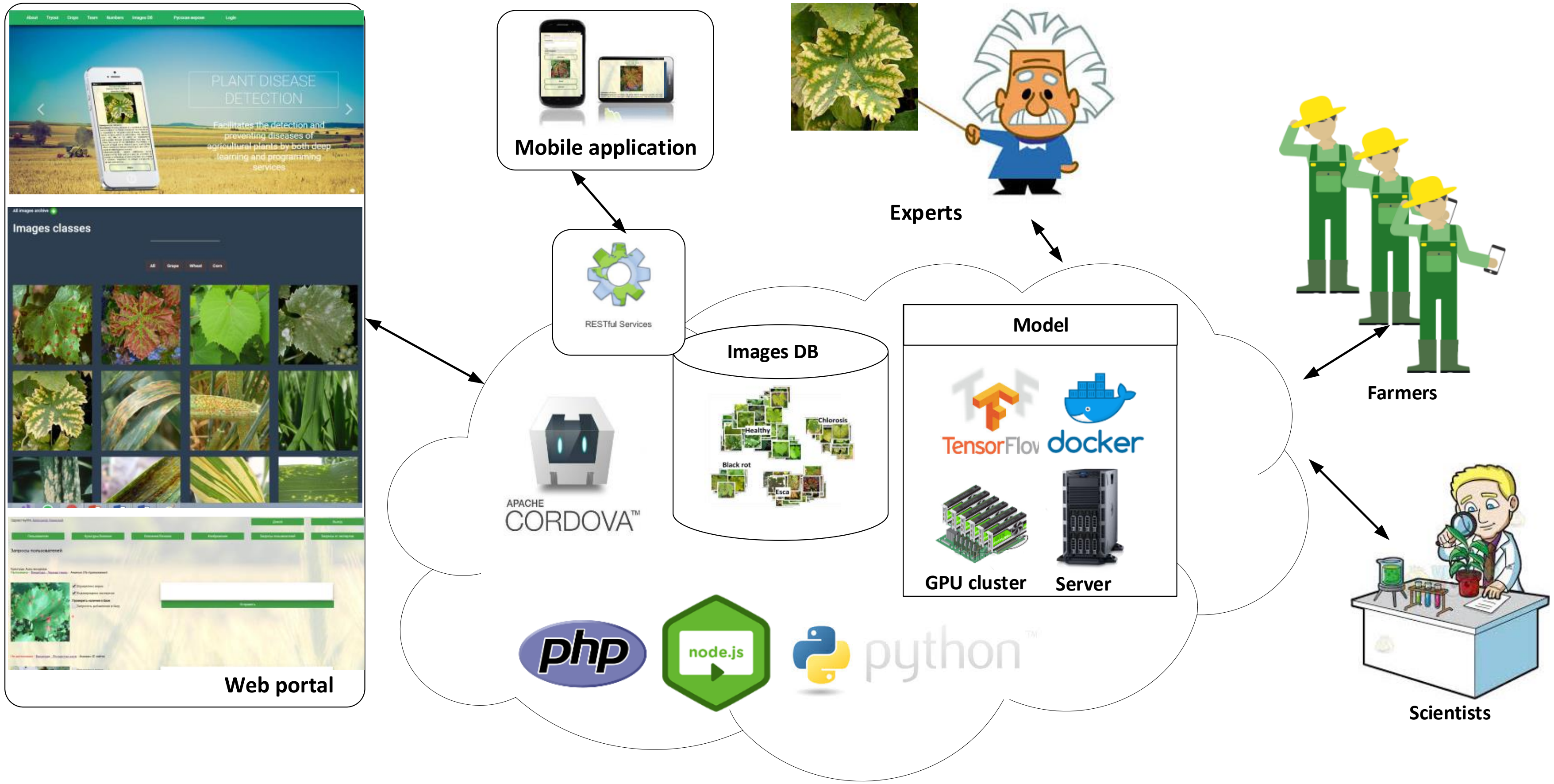
Использование нейросетей в задачах сельского хозяйства

Tasks



object recognition
Cat
(1)
object detection
cat
ball
(2)
semantic segmentation
Red: cat
Green: ball
Blue: background
(3)
image captioning
A cat is playing a ball.
(4)
image question answering
Q: How many balls are there in the image?
A: One.
(5)
image generator
A cat is playing a ball.
(6)
LiLi
1000111110001O
1011010010111O
R
1011111110111O
Input Image
LPN
Output Image
(7)

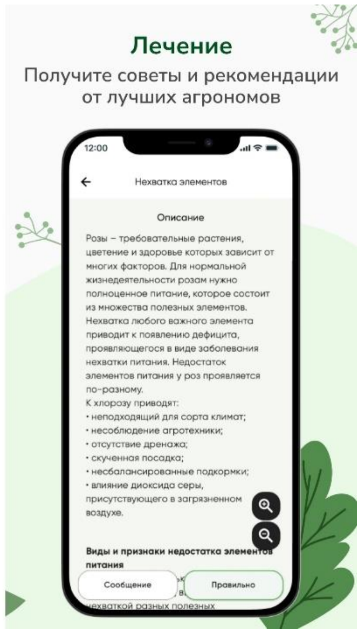
Palnts disease detection platform



PDDP consists of a set of interconnected services and tools developed, deployed and hosted with the help of the JINR cloud infrastructure. Our web-portal (pdd.jinr.ru – old. Doctorp.org - new), was developed with the Node.js and PHP. It provides not only a web-interface but also the API for third-party services. We have the Pytorch model in the Docker realized as a Tensorflow serving. The model can work at the virtual server, or at a GPU cluster.

We have a mobile App for Android that was developed using the Flutter, so we could build it for iOS, and Windows.

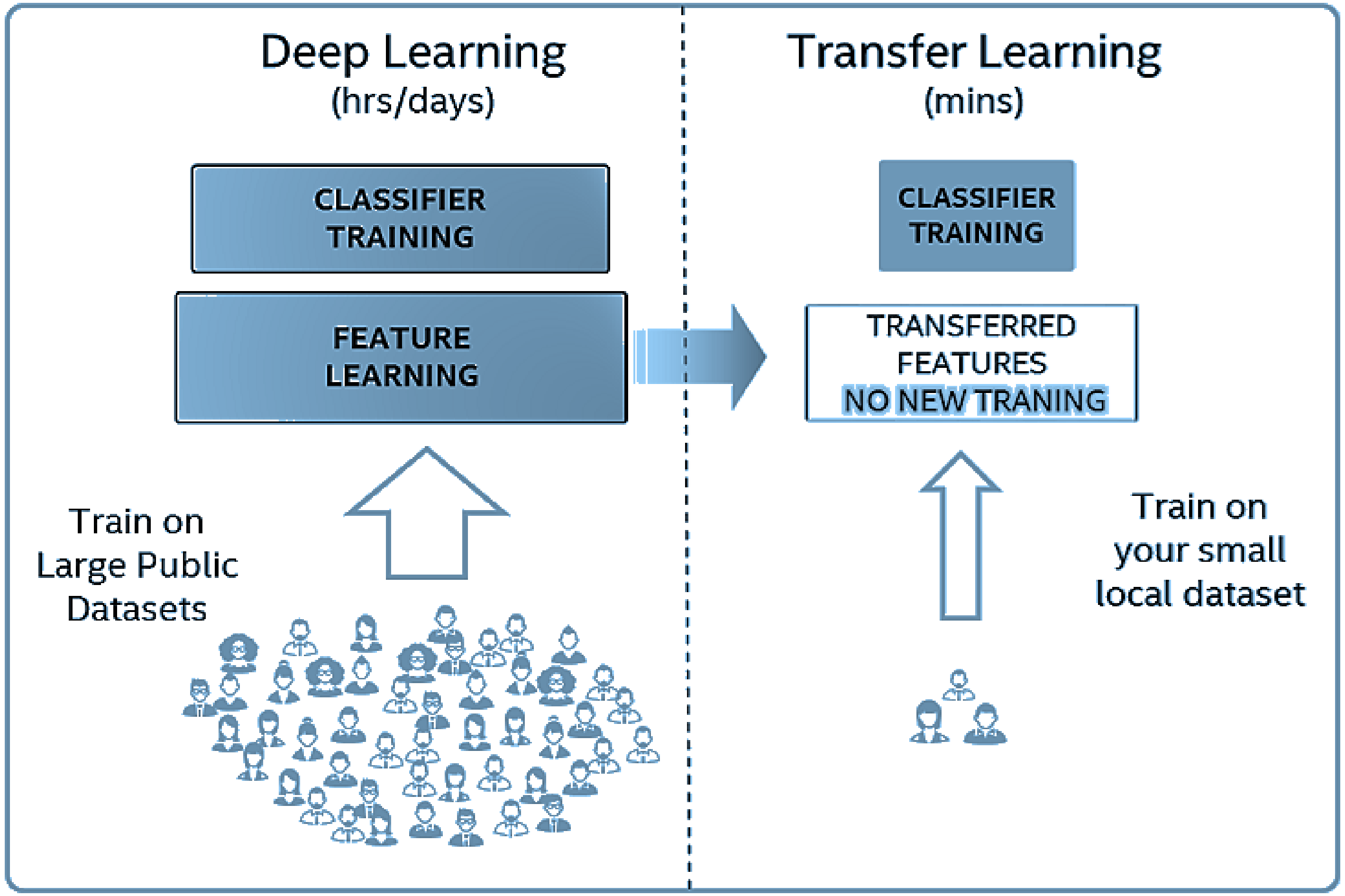
Plants doctor




```
{
  "error":0,
  "general_predictions":[
    {...},
    {...},
    {...}
  ],
  "custom_predictions":[
    {...},
    {
      "type":"розы",
      "prediction":[
        {
          "name":"Мозаика",
          "sample":"http://usrbb.ru/botainika/diseases/rmv1.jpg",
          "description":"<p>Это наиболее распространенное заболевание вирусного типа. Узнать его очень просто – листья покрываются хлоротичными пятнами и узорами, которые могут перейти в прожилковый хлороз. Листья деформируются (узколистность, курчавость, морщинистость) и постепенно опадают. Рост побегов замедляется, и они не вызревают. Побеги, которые больны, вырезают. При дальнейшем распространении болезни растение уничтожают
```

Basis

HOW TO LEARN IF LACK OF DATA
TRANSFER LEARNING



- **FIND A DEEP NEURAL NETWORK PRETRAINED ON A BIG DATASET**
- Replace the classification layer with a layer appropriate for your task
- Finetune the new classifier on specific data
- Voila! Use the new model for inference

ImageNet

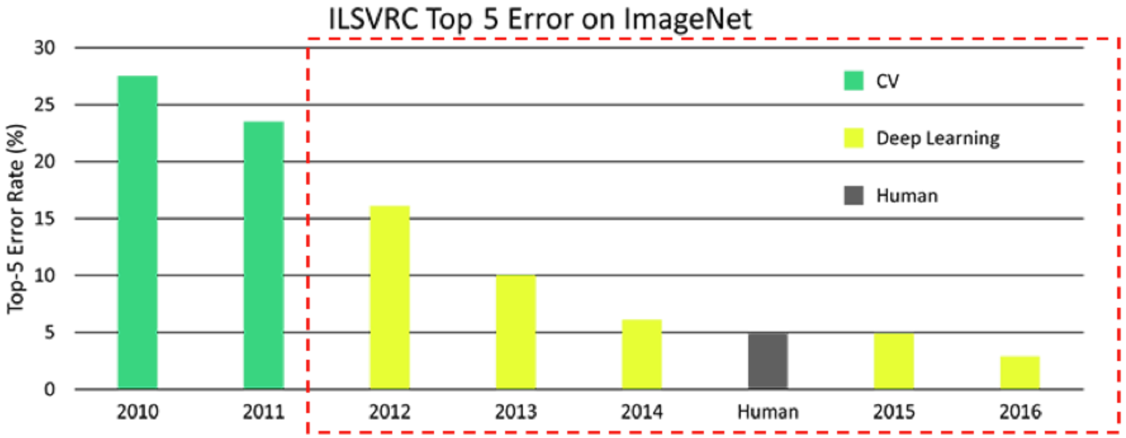
This dataset spans 1000 object classes and contains 1,281,167 training images, 50,000 validation images and 100,000 test images.

Since 2010 the dataset is used in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC), a benchmark in image classification and object detection.

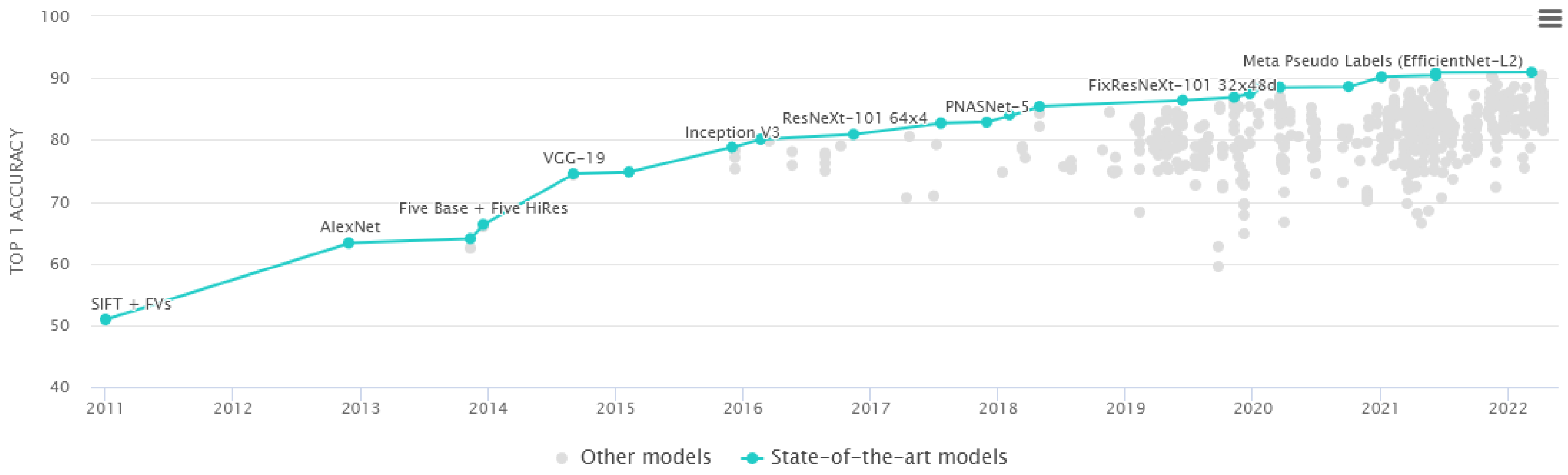
Deng J.; Dong W.; Socher R.; Li L.J., Li K.; Fei-Fei L. ImageNet: A large-scale hierarchical image database, 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 2009, pp. 248-255, doi: 10.1109/CVPR.2009.5206848.

Basis

IMAGENET Large Scale Visual Recognition Challenge (ILSVRC)



The introduction of Deep Learning techniques drove performance on image categorization from 30% error rates in 2010, down to <2% in 2017



One-shot / Few-shot Learning:

Few-shot learning (FSL) is a branch of supervised machine learning, focused on learning from limited number of examples[1]

Inspired by human/animal ability to rapidly generalize from few examples.

Typical scenarios[1]:

- Learning for rare cases – when obtaining labelled data is hard or impossible
- Reducing data gathering effort and computational cost

Variations:

- (<50, e.g. 5,10)-shot learning: General case, where only few examples are given
- 1-shot learning: Extreme case, where only one data example is given (e.g. 1 image per class)
- 0-shot learning: Instead of image, we have description of new class
- 'Less than 1'-shot learning[2]: N classes, M examples, M<N, use soft-labels

Why does it work?

In ML, experience is gained by fitting data.

In FSL we also have prior knowledge[1]. Two training phases:

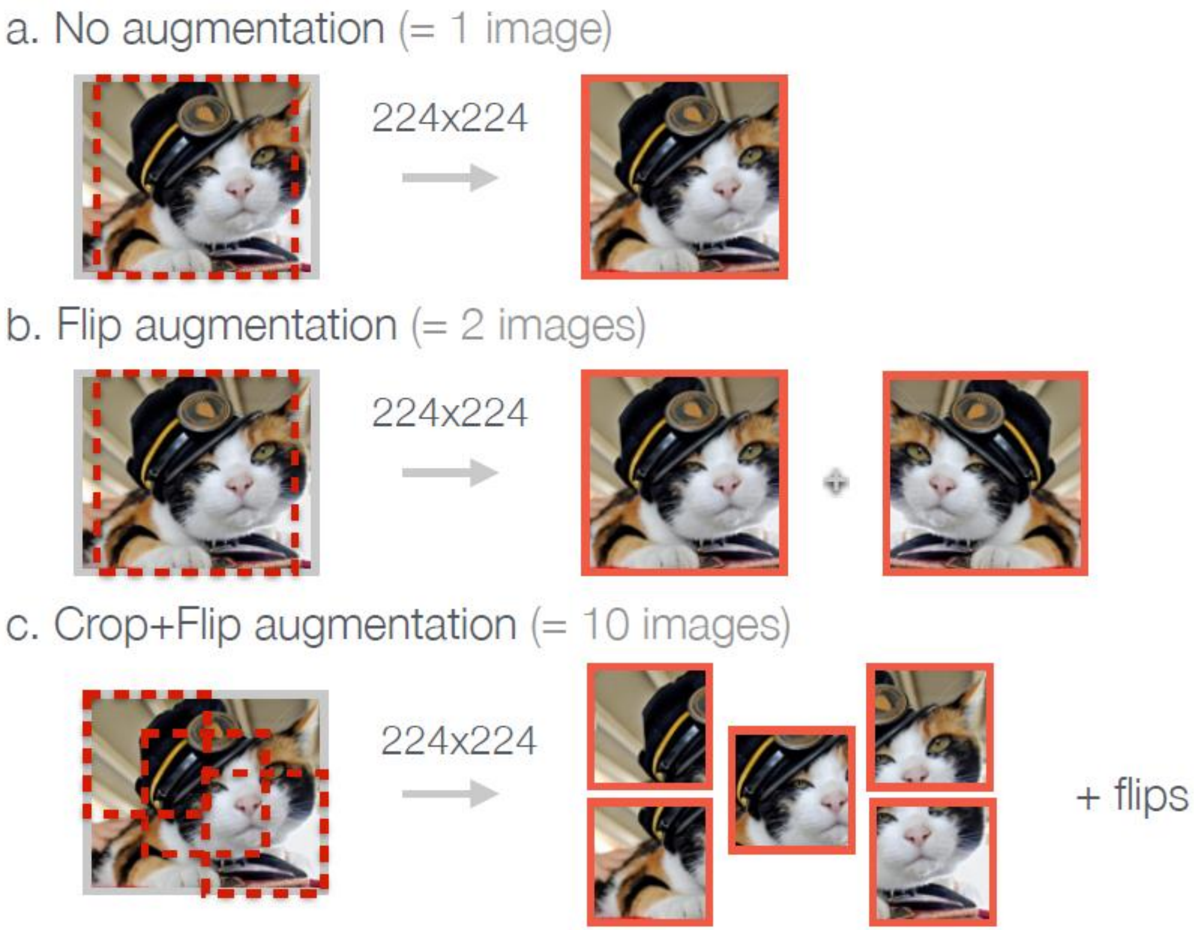
1. Gain prior knowledge by fitting large amount of examples that are similar to goal task
2. Gain experience on small dataset for goal task.

Autoaugmentation

Random Search learner This learner is a purely randomised searcher.

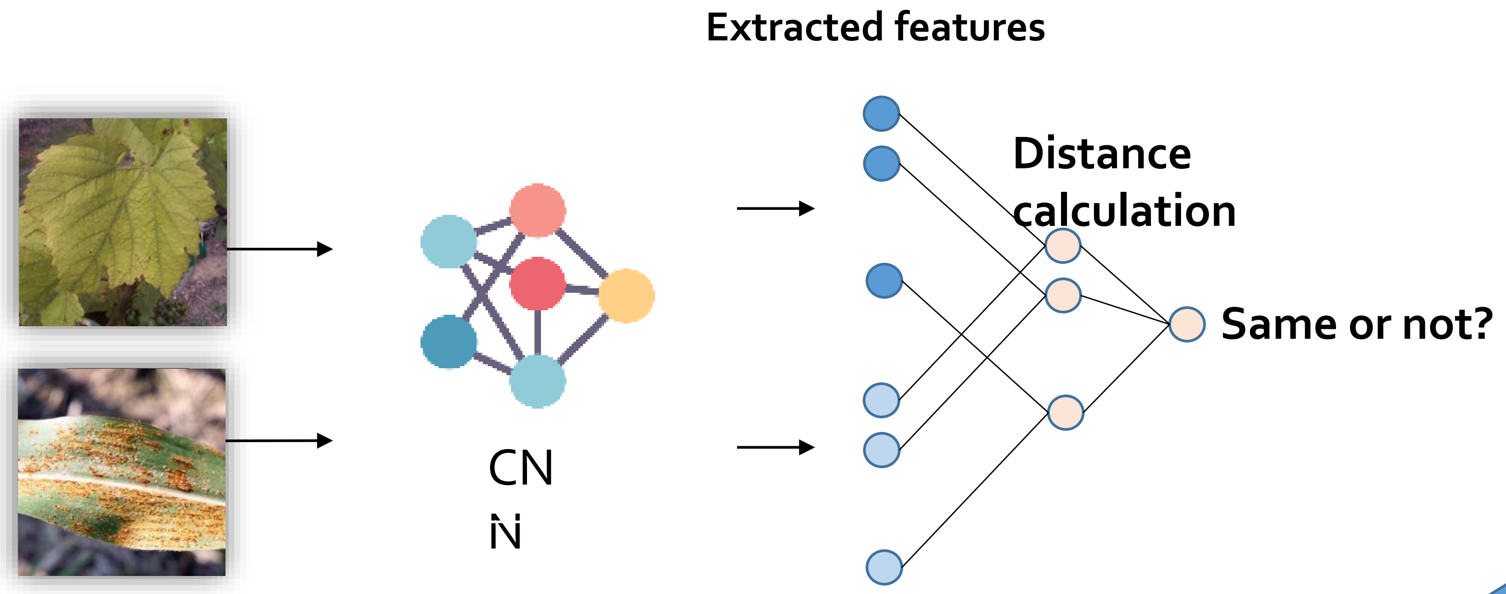
The genetic learner has similar elements to the Random Search learner, but uses information from previous sub-policies when generating new ones to more efficiently search for optimal augmentation parameters.

GRU with PPO updating Agent (gru learner.py) An GRU controller was used, which output a policy in the form of a length 10 sequence of vectors, each vector representing a operation. The GRU controller was updated using proxima policy optimization(PPO), using the accuracy of the child network as the reward value. In the context of the PPO update, which was developed in the reinforcement learning literature, the subpolicies are the ’actions’ of RL agent.



How to learn if lack of data - Siamese Networks

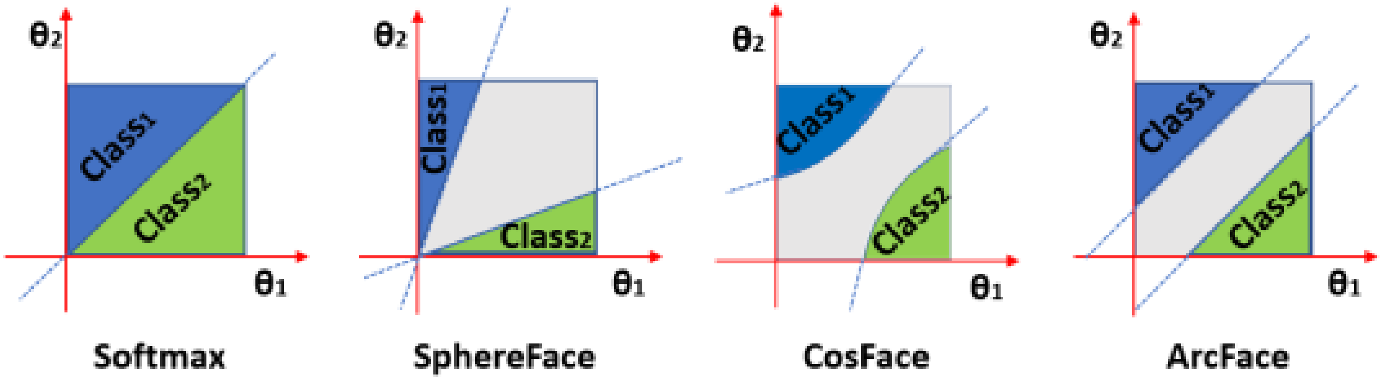
Siamese networks is a part of **one-shot learning** approach. One shot-learning aims to learn information about object categories from one, or only a few, training samples/images



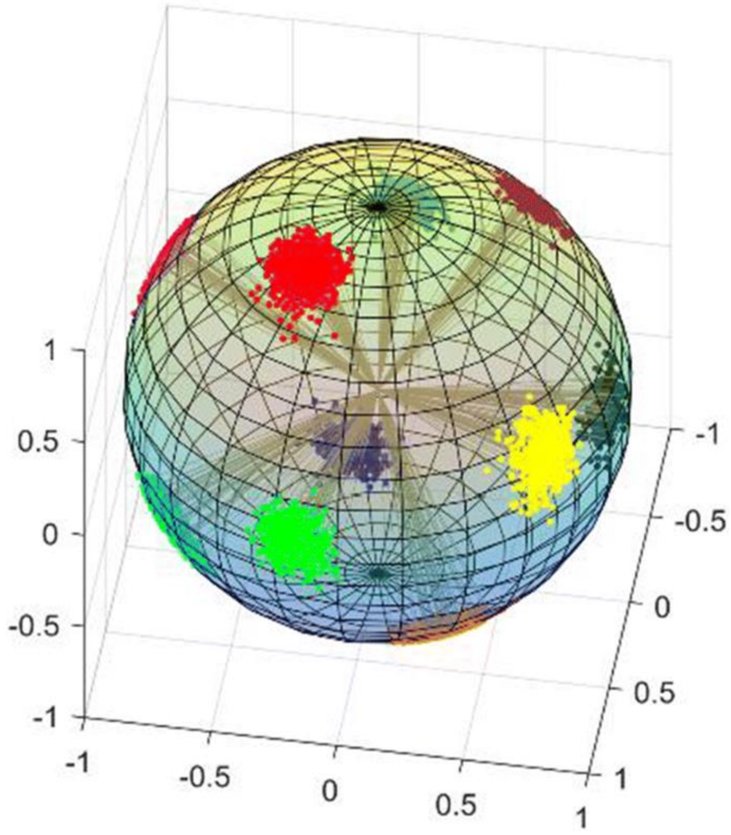
- Train siamese network on your own paired data
- Take the trained twin and append a classifier on top of it
- Finetune the new classifier on specific data appropriate for your task

Same as for transfer learning

Geometrical loss functions: arcface sphereface cosface



ArcFace, or **Additive Angular Margin Loss**, is a loss function used in face recognition tasks. The [softmax] is traditionally used in these tasks. However, the softmax loss function does not explicitly optimise the feature embedding to enforce higher similarity for intraclass samples and diversity for inter-class samples, which results in a performance gap for deep face recognition under large intra-class appearance variations.



Stage 2 – Validation of the generalization abilities (400 images)

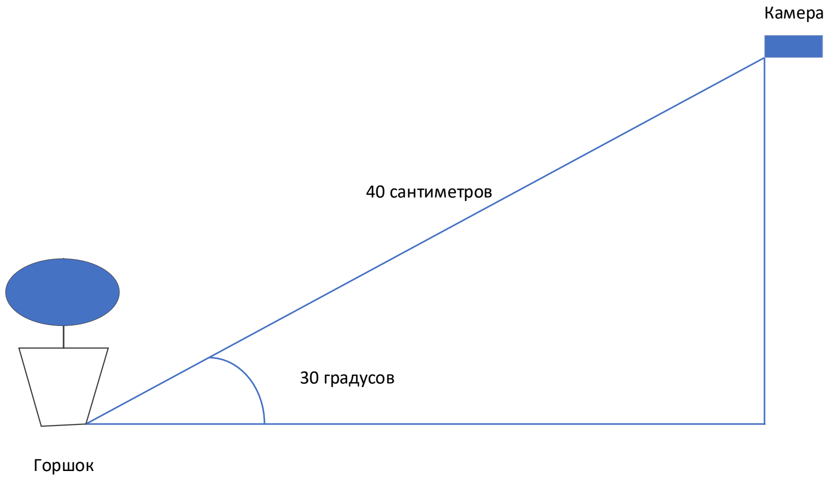
Custom datasets of real-life images are utilized, comprising over 4,000 samples across 68 classes of plant diseases, pests, and their effects.

Table presents the evaluation results of models on a dataset consisting of 400 hard-case images. Additionally, to illustrate the advantages of few-shot learning methods over Transfer Learning methods, the accuracy of networks using ImageNet weights with the same classification architecture, referred to as “Vanilla TL,” was also evaluated.

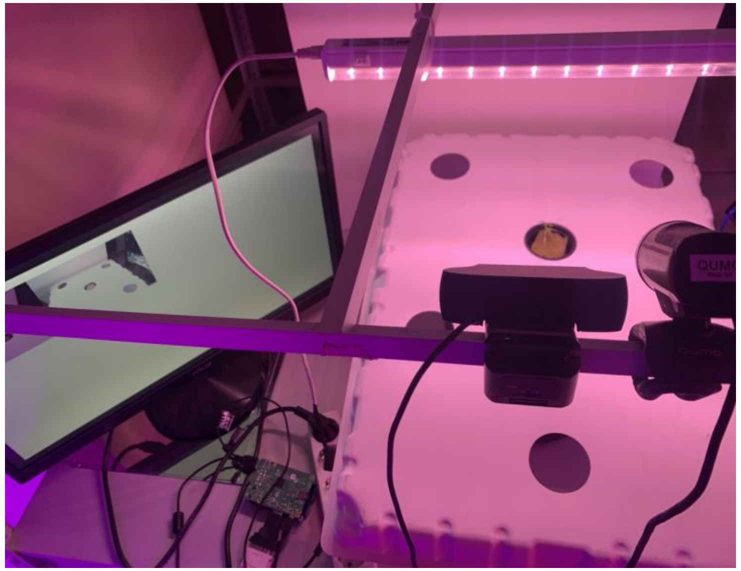
	ConvNeXt_small (191.7 Mb)		EfficientNet_B3 (47.2 Mb)		MobileNetV2 (13.6 Mb)		ResNeXt50_32X4D (95.8 Mb)	
	Vanilla	Norm	Vanilla	Norm	Vanilla	Norm	Vanilla	Norm
Contrastive	44%	46.4%	43.4%	46.2%	38.4%	40%	37.6%	43%
Triplet	58.2%	58.6%	54.4%	53.4%	39.4%	40.2%	47.4%	51.8%
Quadruple	58.6%	62.4%	55%	53.4%	40.6%	40.6%	50.4%	51.6%
SphereFace	72.2%	75.8%	57.6%	62%	45.8%	43.4%	53.4%	55.2%
CosFace	71.2%	73.6%	56.8%	59.8%	43.6%	43.8%	52.6%	49.6%
ArcFace	70%	67.8%	55%	57.2%	40.6%	41.2%	51.6%	49.6%
Vainilla TL	43.2%		37%		32%		38.6%	

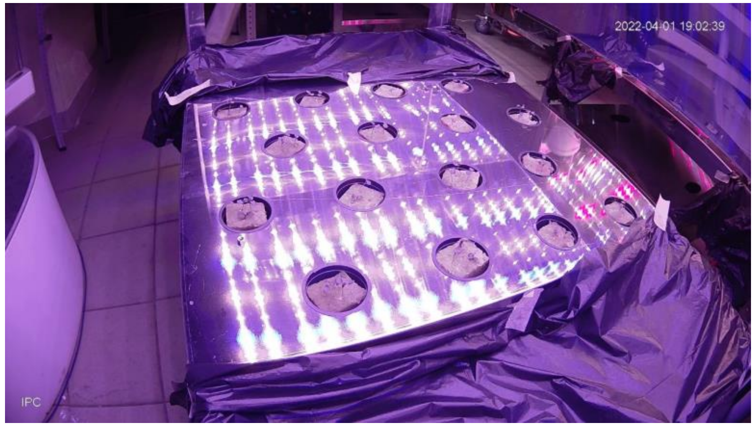
Plant state tracking

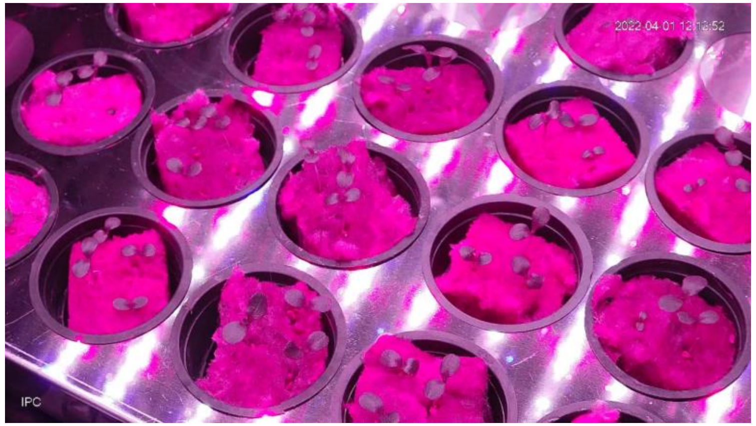
Joint project with the Temiryazev Academy within the framework of the project World-class Scientific Center “Agrotechnologies of the Future”



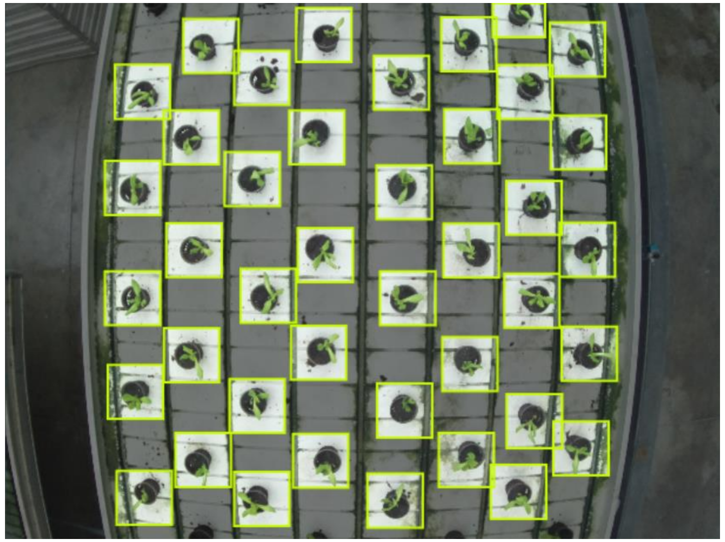
- **Classification of the degree of development of the plant.**
- **Determination of the weight group of the plant.**

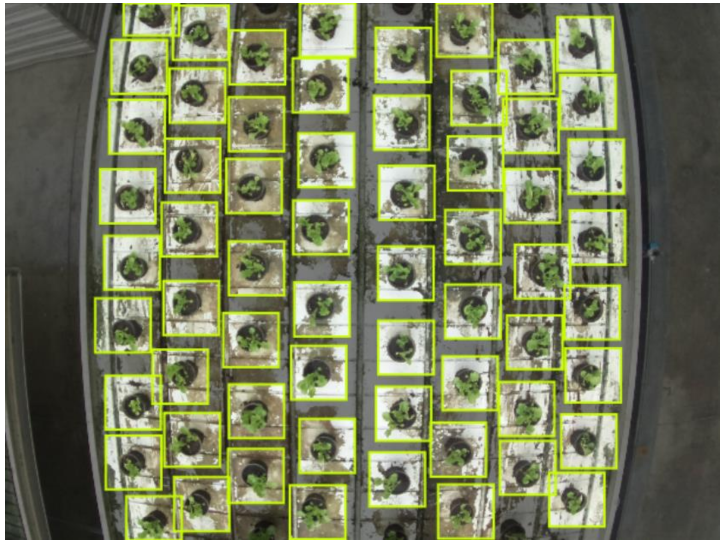


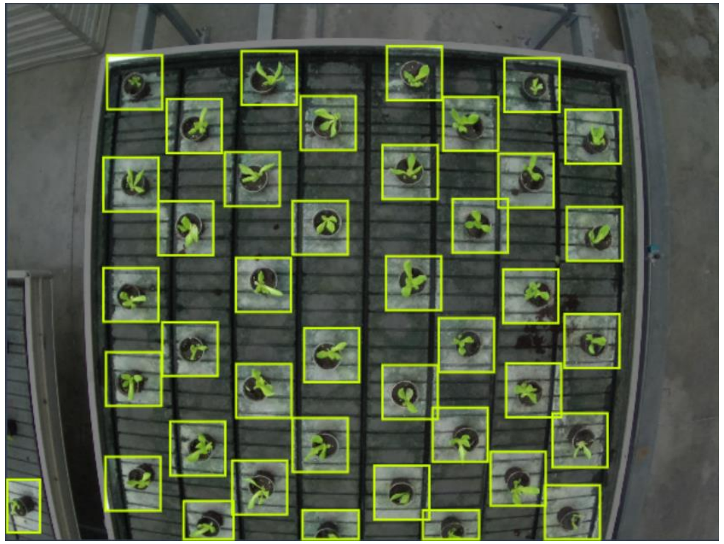




Salads classification







Object detection – 1 class
Classification – 6 classes

Accuracy > 99%

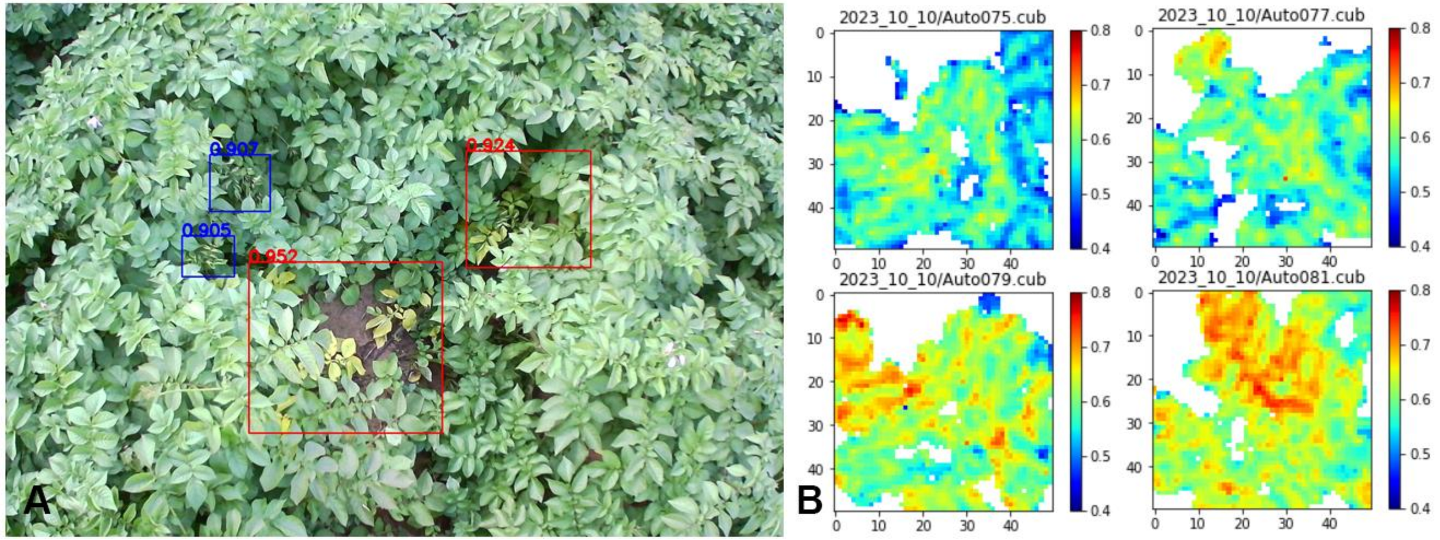
Тестировалась на различных моделях семейств YOLO V5, YOLO V7, YOLO V8, YOLO NAS

1я модель 92 изображения, 200 объектов
2я модель 92 изображения из нового сезона, 170 объектов
3я модель (объединенная 1я и 2я)

Модель	Precision	Recall	mAP50
V1	0.685	0.417	0.516
V2	0.635	0.531	0.586
V3	0.728	0.588	0.631

1я модель 124 изображения, 194 объектов

	Precision	Recall	mAP50
V1	0.827	0.567	0.703

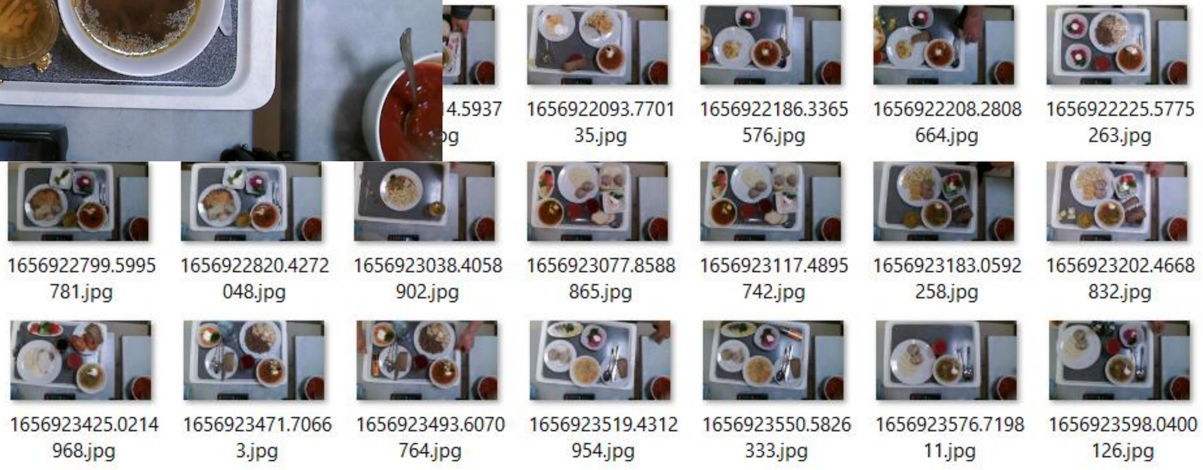


Set of images

Data collection was carried out in automatic mode using raspberry pi 4, a digital camera with manual focus, and sonar.

A set of images in 5 days:

Dining room LIT JINR



-  446 images
-  More than 150 different classes of objects

Examples

Input images

Processing

Output images

Что можете сделать вы?

- Исследование различных подходов к расширению обучающей выборки
- Использование Generative adversarial networks (GANs) для расширения базы изображений
- Оценка эффективности различных object detection подходов на различных задачах
- Использование методов классификации и детекции для решения прикладных задач в различных областях
- Ваш вариант …

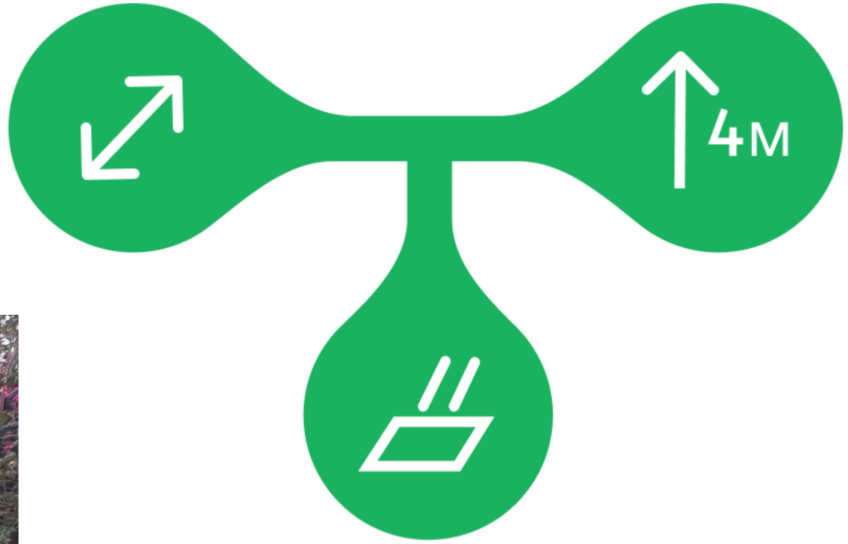
Russian greenhouse market
> 30 big complexes
> 1300 hectares
> 1300 thousand tons of vegetables
cucumbers — 48%
tomatoes - 46%
salads, greens, etc. - 6%

Greenhouse complexes

Greenhouse complexes play a pivotal role in year-round production of various crops A crucial challenge in greenhouse operations involves early pest detection and precise localization on leaves. Additionally, tasks like plant accounting, yield estimation, and environmental control at different locations are of paramount importance.

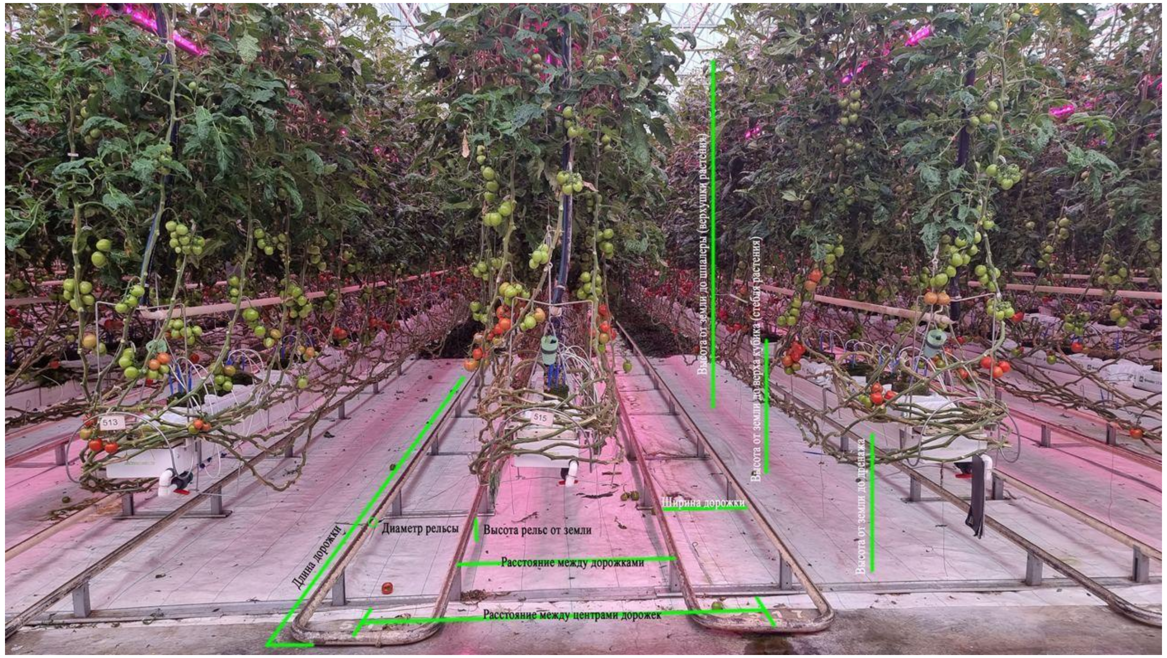
Issues

Greenhouse dimensions

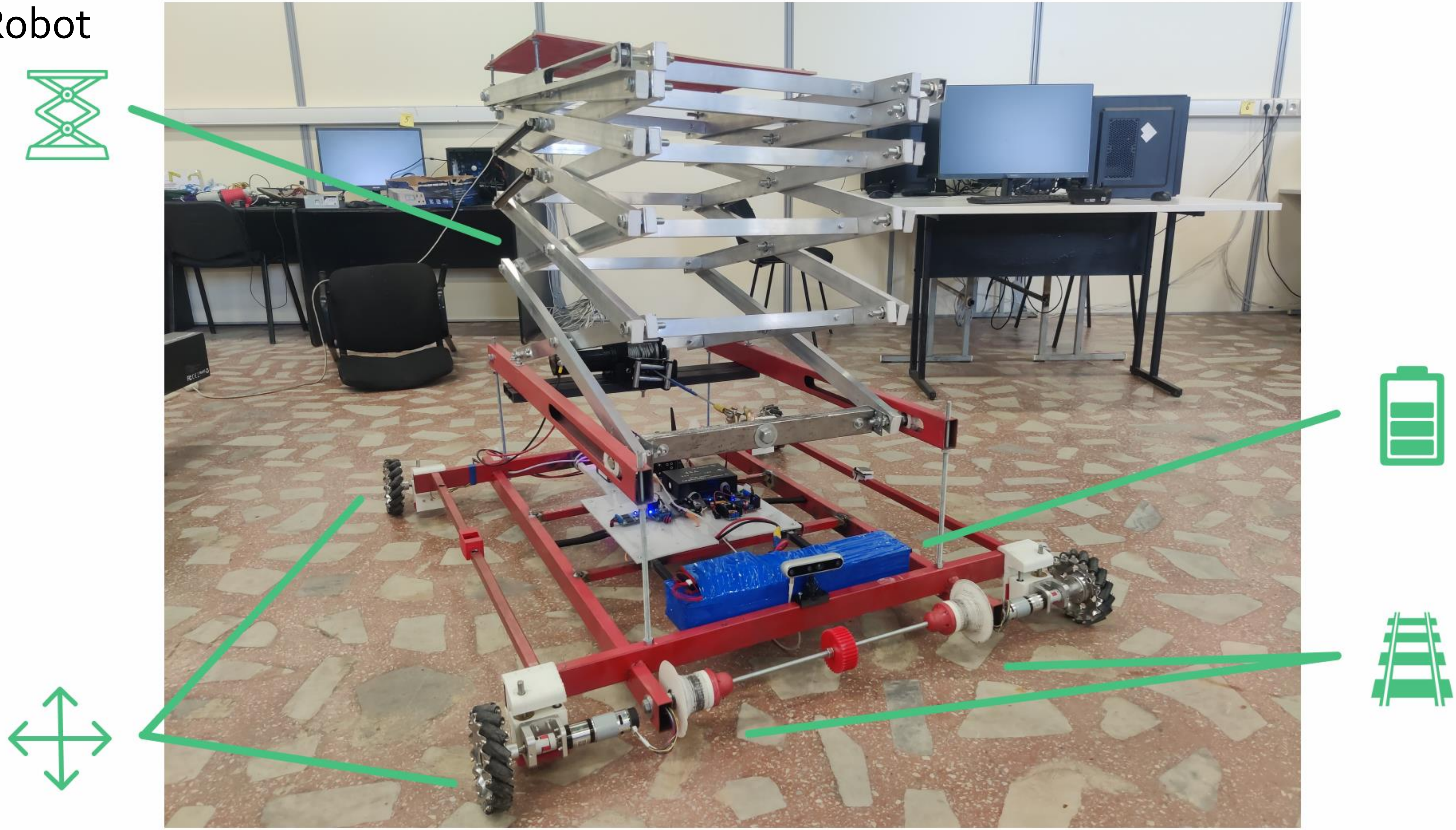


Diverse tasks at high altitude

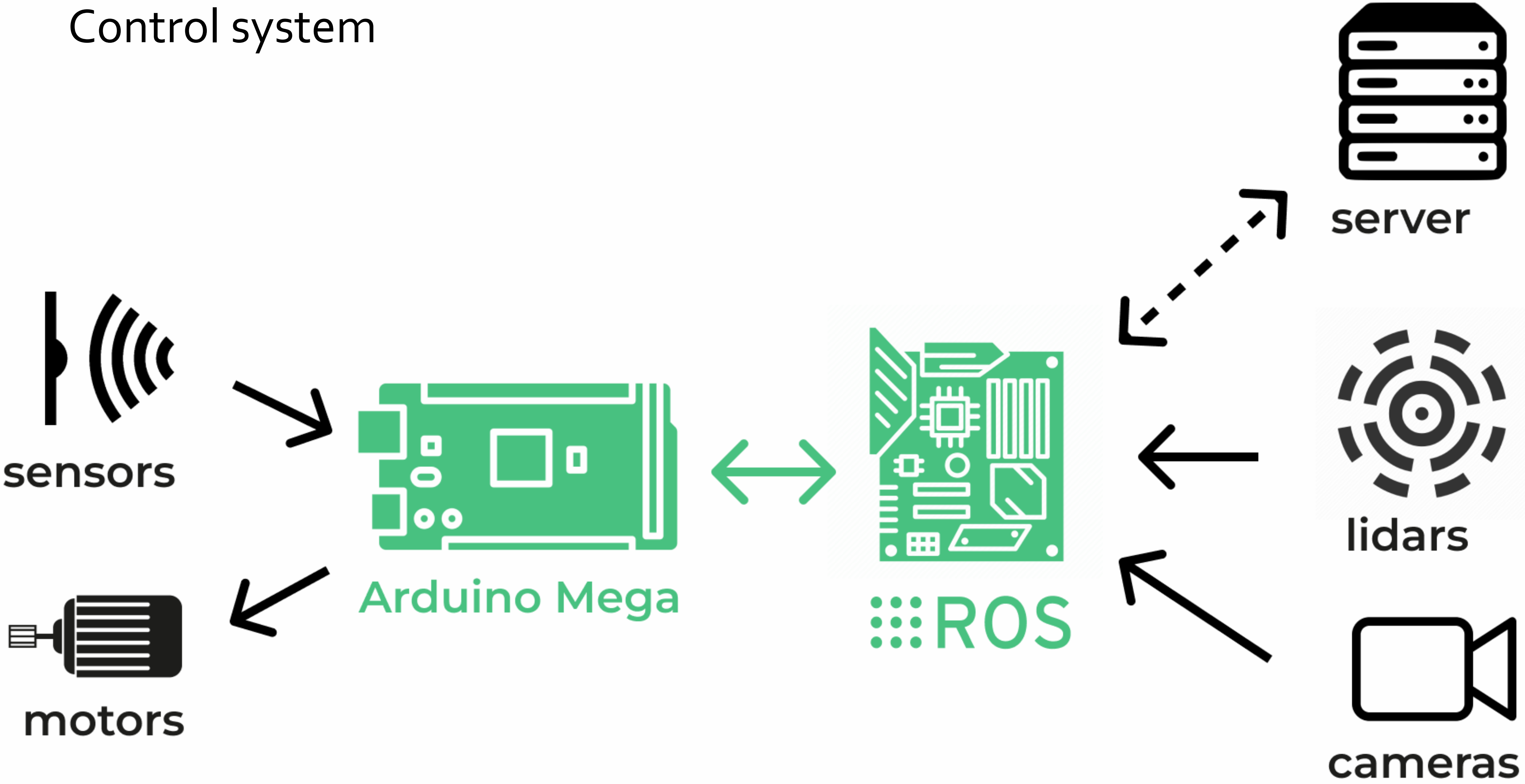
Two methods of movement - rails and concrete platform



Robot



Control system



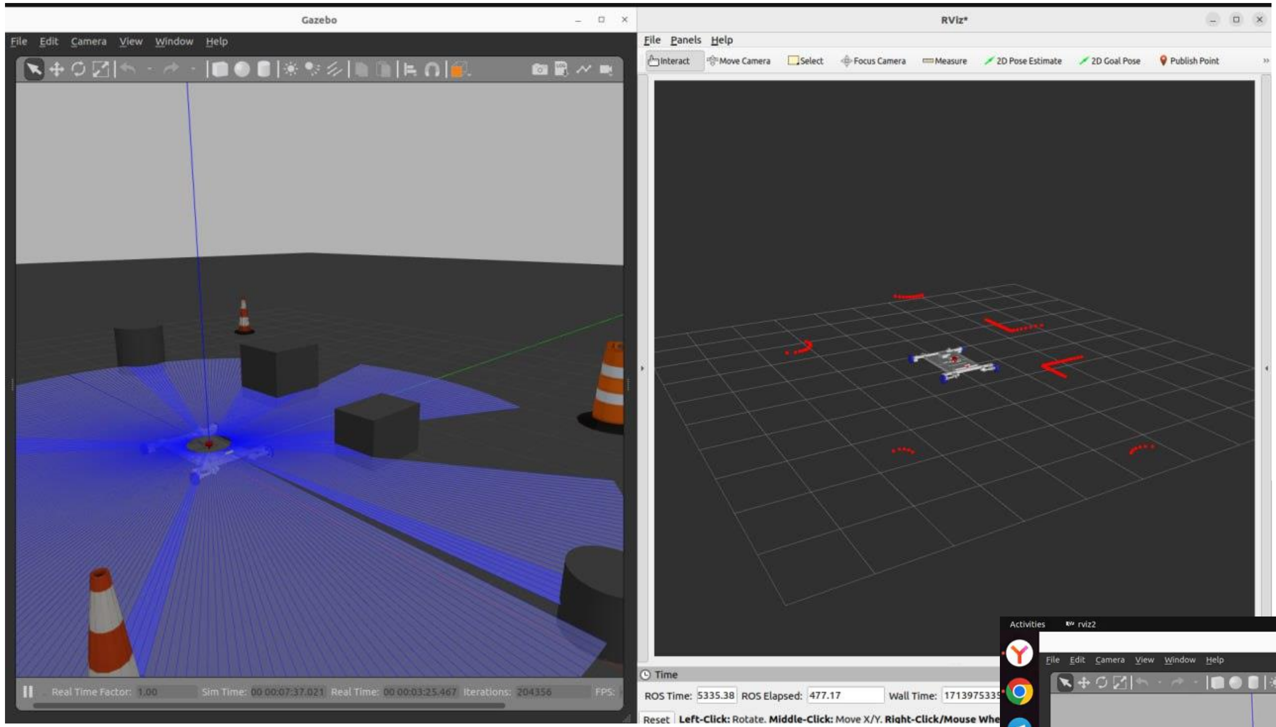
Navigation

QR codes

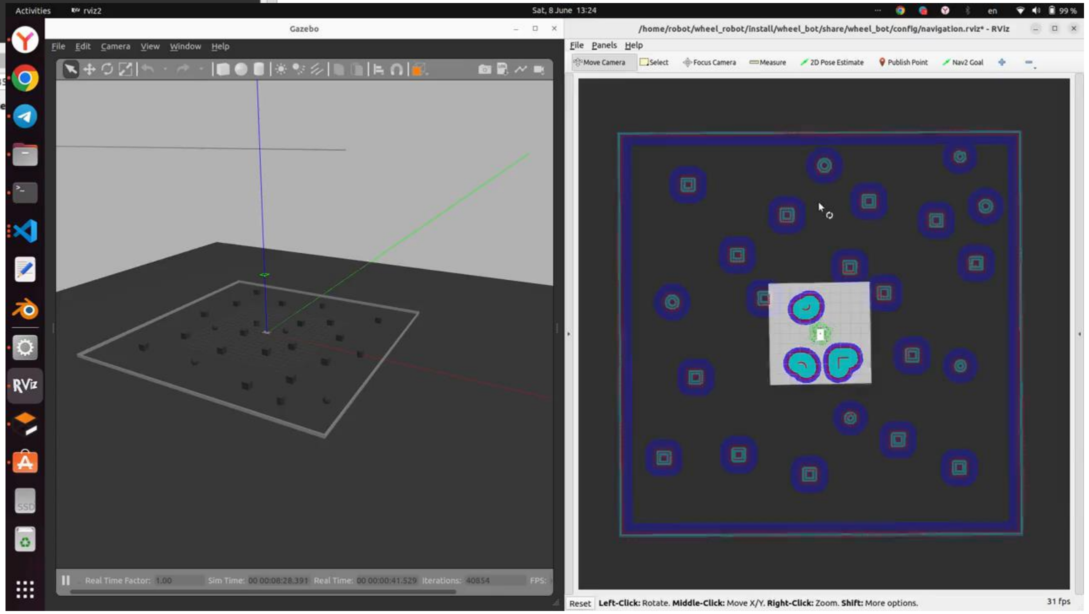
Point cloud

IMU
Odometry

Lidars

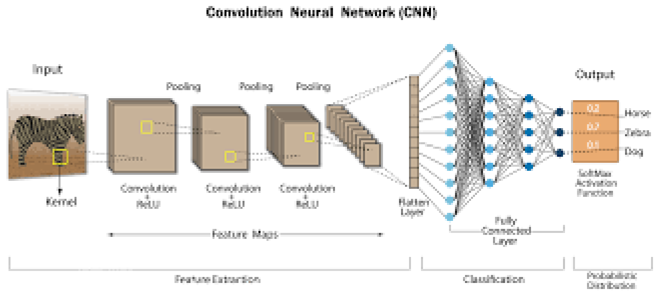

Gazebo
RViz*
File Edit Camera View Window Help
File Panels Help
Interact
Move Camera
Select
Focus Camera
Measure
2D Pose Estimate
2D Goal Pose
Publish Point
Time
ROS Time:
ROS Elapsed:
Wall Time:
Reset
Left-Click: Rotate. Middle-Click: Move X/Y. Right-Click/Mouse Whe

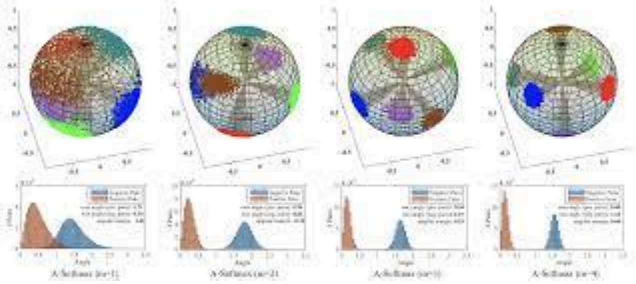
ROS2


Gazebo
/home/robot/wheel_robot/install/wheel_bot/share/wheel_bot/config/navigation.rviz* - RViz
File Edit Camera View Window Help
File Panels Help
Reset
Left-Click: Rotate. Middle-Click: Move X/Y. Right-Click: Zoom. Shift: More options.

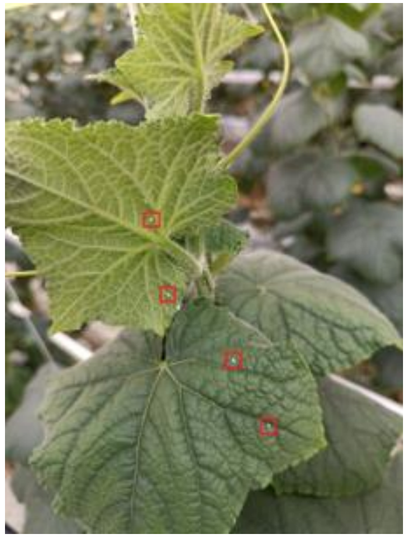
YOLO

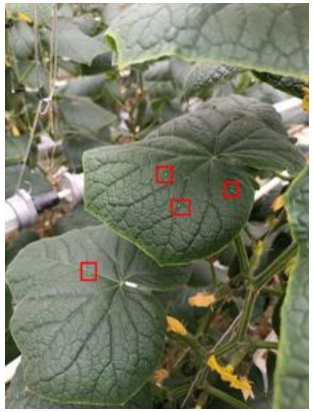
8, 9, 10



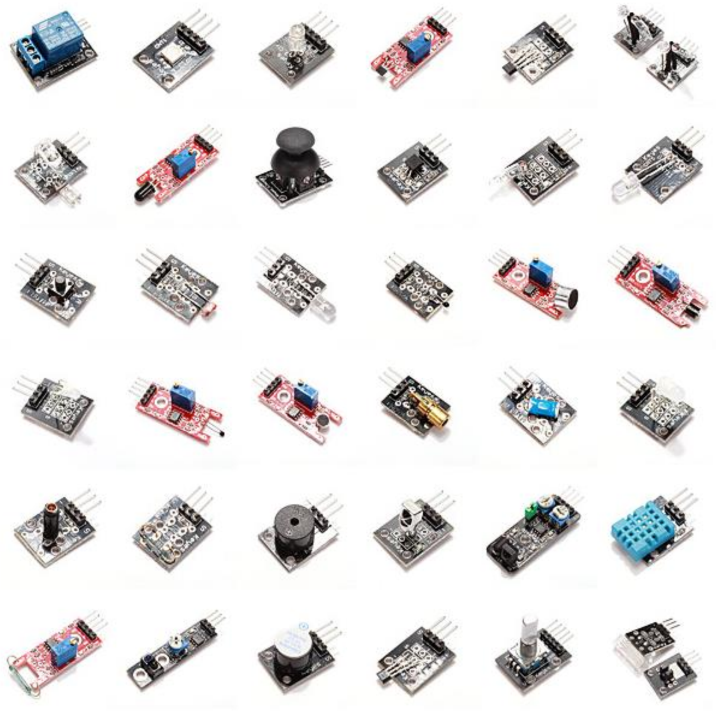


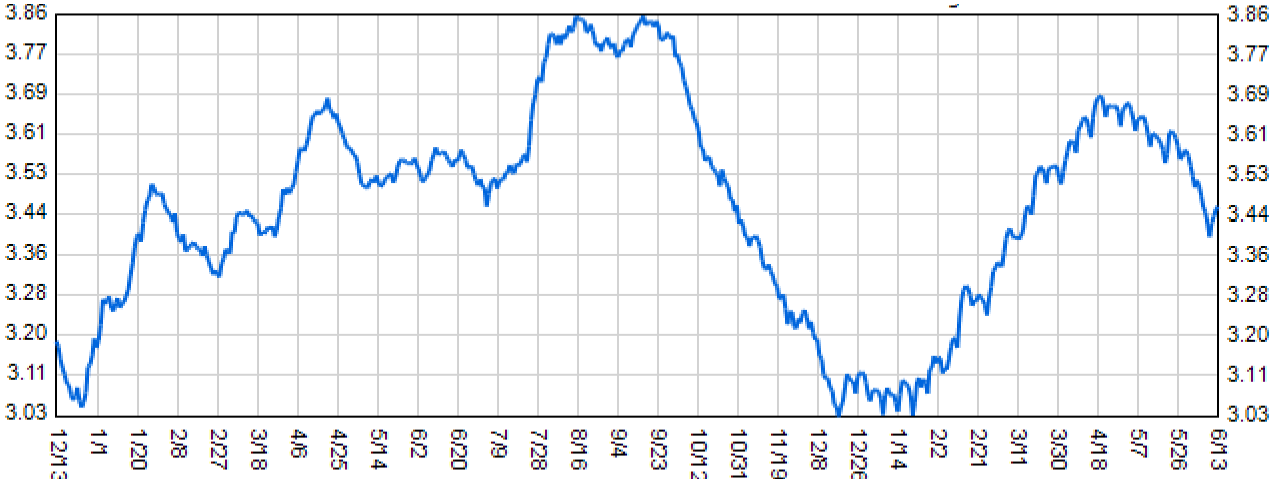
Uzhinskiy, A.; Ososkov, G.; Goncharov, P.; Nechaevskiy, A.; Smetanin, A. Oneshot learning with triplet loss for vegetation classification tasks. *Comput. Opt.* 2021, *45*, 608–614.





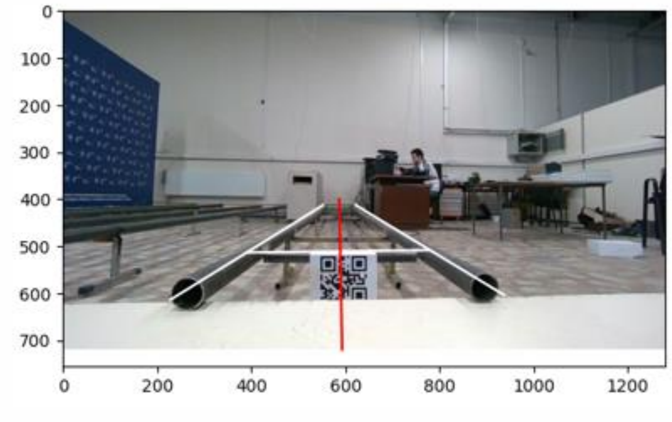
Sensors:
Temperature,
Humidity,
Light,
Carbon dioxide, Nitrous oxide, Hydrogen,
Hydrocarbon gases, etc

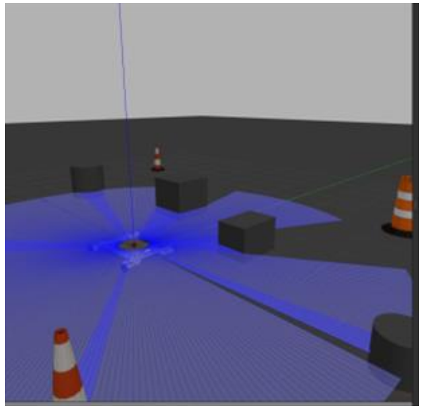





2 ряд
Дата: 21.10.2023
Время: 23:44 - 23:56
Растений: 36
Увядших растений: 3
Проблемы: Белокрылка (п. 10м)
1
2
3
4
5
6
7
8


2 ряд
Дата время: 21.10.2023 23:45
Позиция: 10м
Проблема: Белокрылка
1
2
3
4
5
6
7
8


0
100
200
300
400
500
600
700
0
200
400
600
800
1000
1200



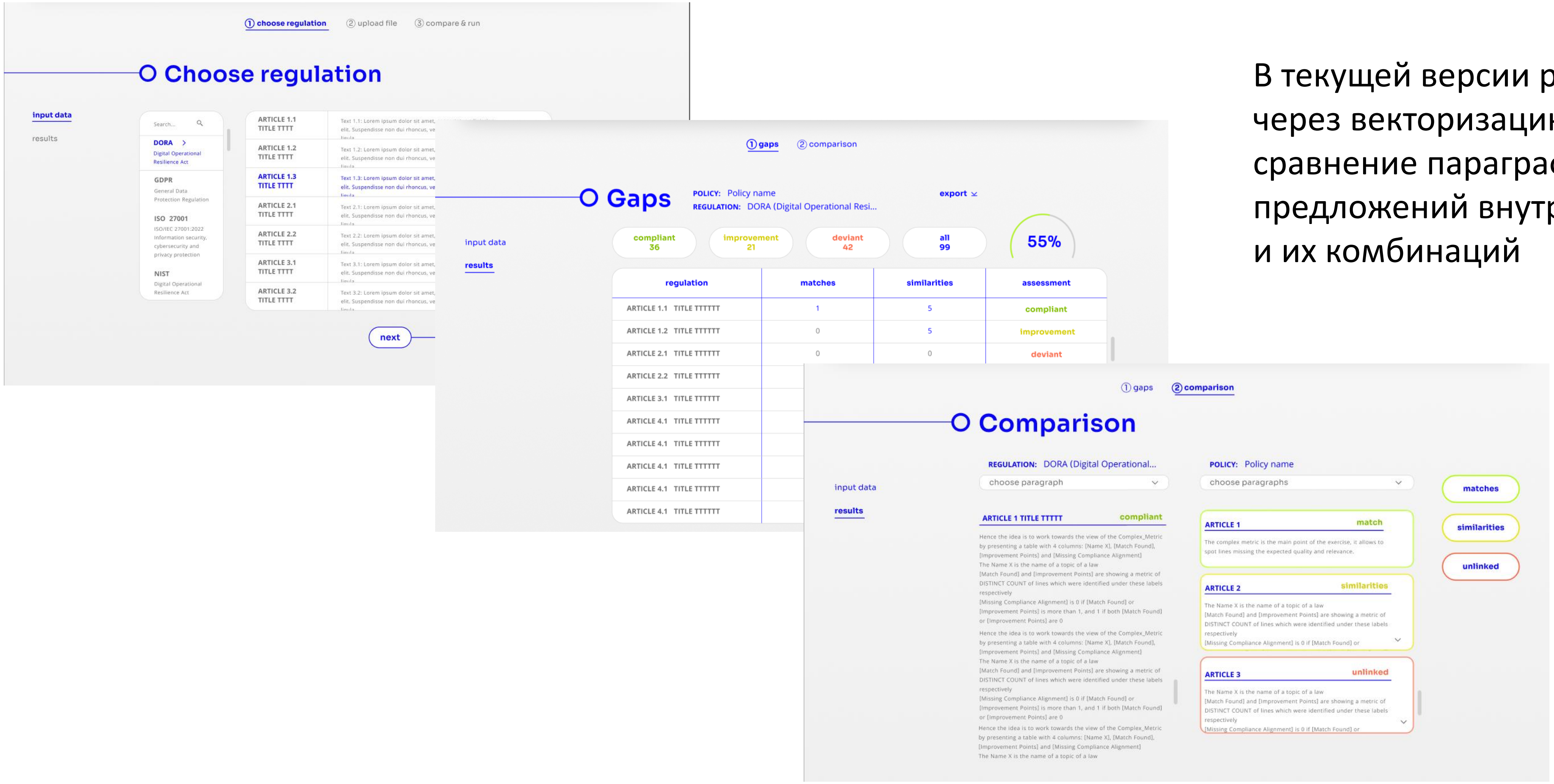

5 ряд
Дата время: 21.10.2023 23:54
Позиция: 6м
Проблема: Трипс
1
2
3
4
5
6
7
8

Что можете сделать вы?

- Широкий спектр различных задач в области робототехники - навигация, локализация, картографирование, движение
- Оценка эффективности различных object detection подходов для отслеживания болезней и вредителей в тепличных комплексах
- Интеграция датчиков для измерения различных параметров среды в тепличных комлпексах
- Ваш вариант

Анализ похожести текстов

Сравниваем по частям текст внутренней документации компании с законами или стандартами и находим какие части закона/стандарта схожи с регламентами компании. Те части закона, которые никак не отражены во внутренней документации, являются потенциальными уязвимостями при проверке со стороны аудитора и их необходимо описать в регламентах компании.



В текущей версии реализовано через векторизацию **TF-IDF.** Идет сравнение параграфов текста, предложений внутри параграфов и их комбинаций

Что можете сделать вы?

- Исследование альтернативных подходов к решению задачи:
  - сплит-модели
  - кросс-энкодеры типа MonoBERT
  - LLM + LLaMa3 и Mixtral
  - ваш вариант …

Спасибо за внимание!

email: auzhiskiy@jinr.ru
https://t.me/bigzmey

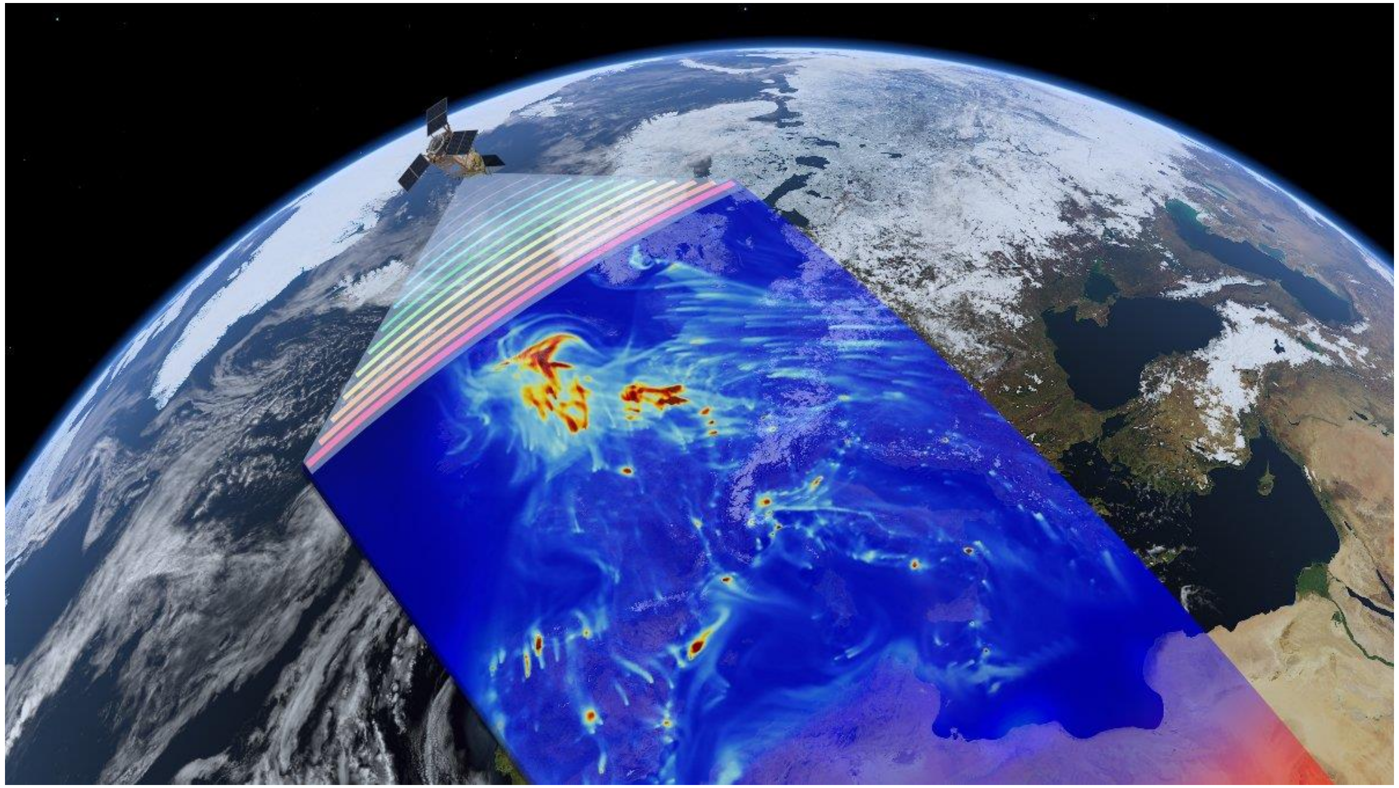
Additional information

Коммерческие спутниковые программы

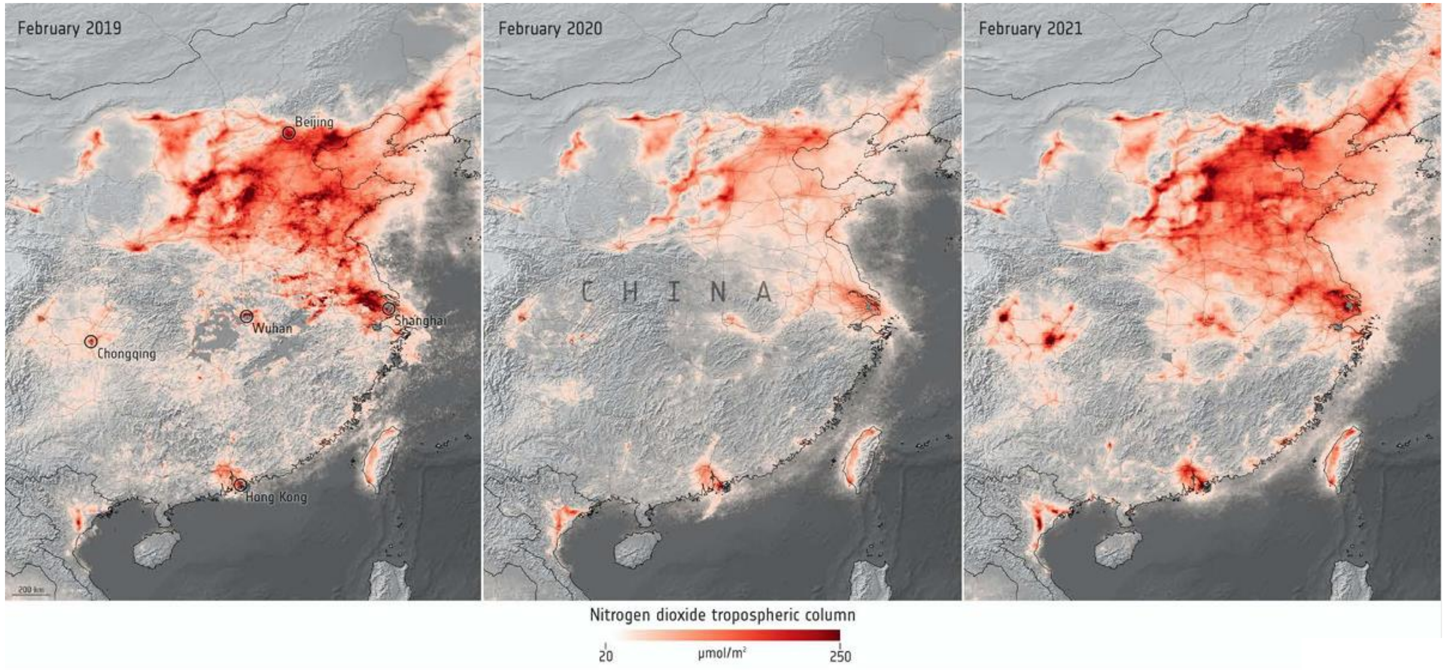
Sentinel-5

The Sentinel-5 mission is part of the European Earth Observation Programme "Copernicus" which is a coordinated and managed by the European Commission (EC). The space component of the Copernicus observation infrastructure is developed under the aegis of the European Space Agency (ESA).

Sentinel-5 is focused on air quality and composition-climate interaction with the main data products being O_3, NO_2, SO_2, HCHO, CHOCHO and aerosols. Additionally Sentinel-5 will also deliver quality parameters for CO, CH_4, and stratospheric O_3 with daily global coverage for climate, air quality, and ozone/surface UV applications.



Sentinel-5

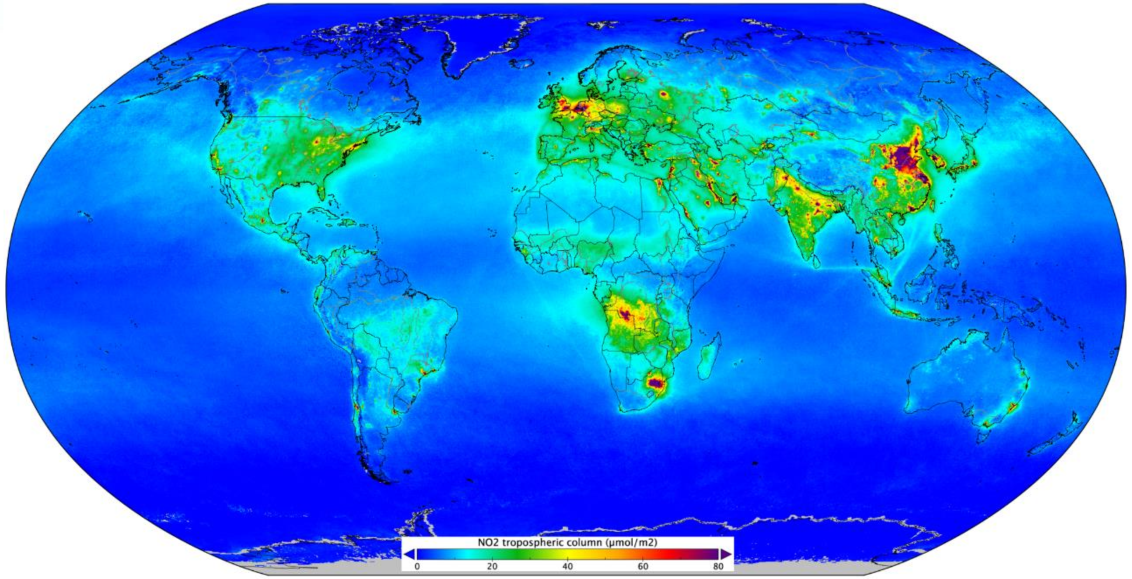


contains modified Copernicus Sentinel data (2019-21), processed by ESA

The Sentinel-5 mission consists of high resolution spectrometer system operating in the ultraviolet to shortwave infrared range with 7 different spectral bands: UV-1 (270-300nm), UV-2 (300-370nm), VIS (370-500nm), NIR-1 (685-710nm), NIR-2 (745-773nm), SWIR-1 (1590-1675nm) and SWIR-3 (2305-2385nm).

Resolution ~ 1114 m,
Orbital cycle 16 days.

Sentinel 7 (CO2M - Copernicus Anthropogenic Carbon Dioxide Monitoring) 2025 - 2026

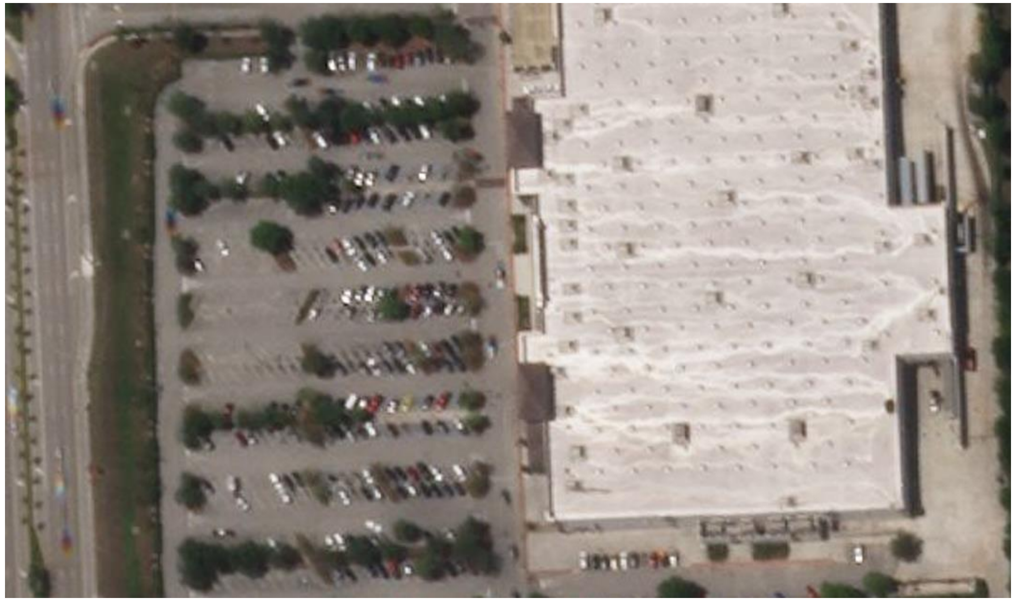


Nitrogen dioxide worldwide 12/03/2019

Commercial programs - www.planet.com

Formula 1 Miami Grand Prix

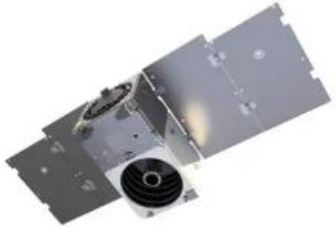
2019
Daily, less than 3m resolution, 4 spectral channels



Commercial programs - www.planet.com

Now



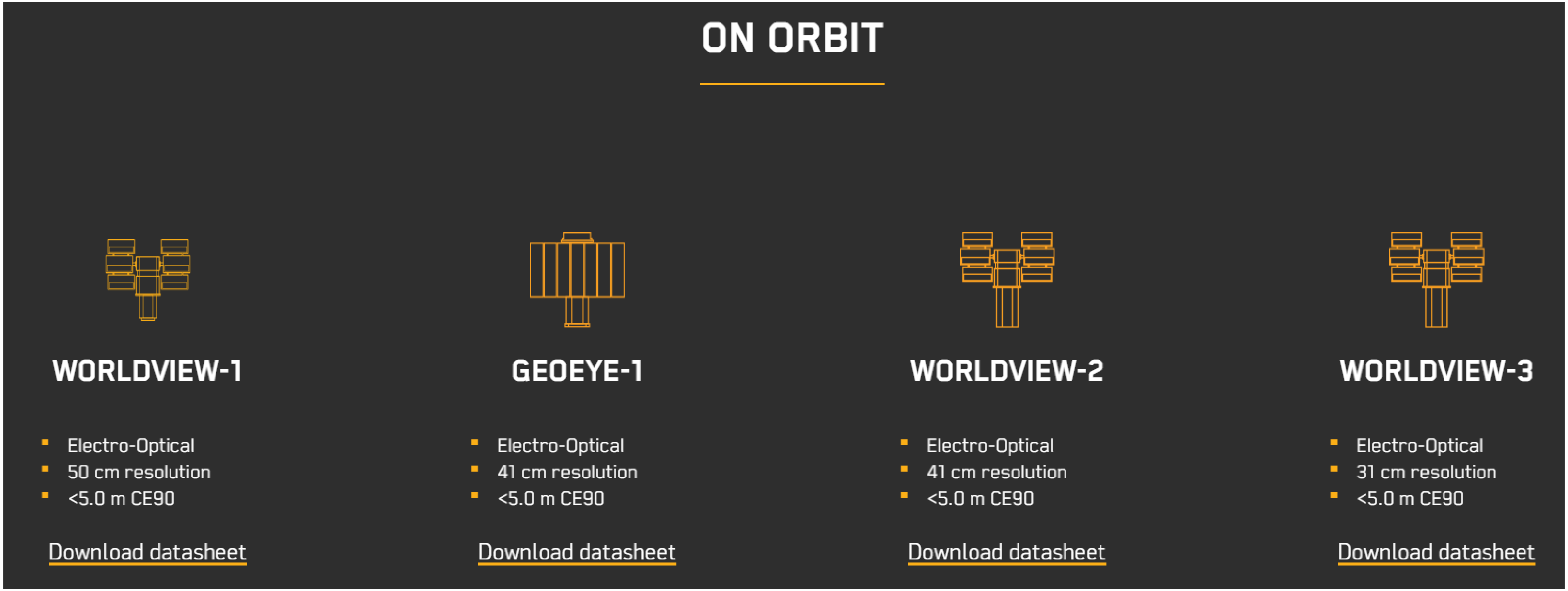



PlanetScope Always-on Monitoring	SkySat High-Resolution Tasking	Pelican Very High Resolution	Hyperspectral Broad Spectral Range
~200 Satellites	**21** Satellites	**~32** Satellites	**2** Satellites
3.7 m (3.0 NIIRS) GSD	**0.5 m (4.0-5.0 NIIRS)** GSD	**0.3 m (>5.5 NIIRs at-nadir)** GSD	**30 m** GSD
8 Spectral Bands	**RGB, Pan and NIR** Spectral Bands	**7** Spectral Bands	**400-2500nm (5nm spacing)** Spectral Bands
Not required Tasking	**Sub-daily** Tasking	**Up to 12 revisits/day** Tasking	**Tasking Required** Tasking

Commercial programs - maxar.com

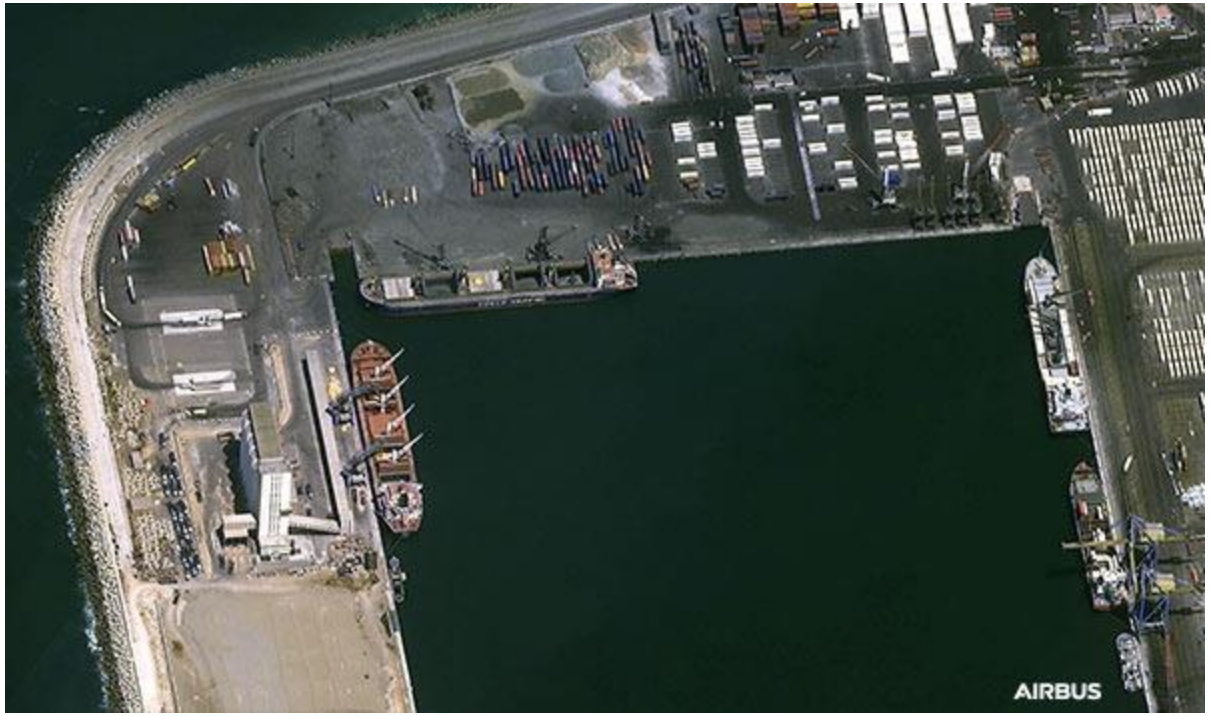
1999
~ 3 days, 3m resolution, 3 spectral channels

Commercial programs - maxar.com



From 2007	From 2008	From 2009	From 2014
~ 2 days	< 2 day	~ 1.2 day	< 1 day
1 spectral channel	5 spectral channels	9 spectral channels	29 spectral channels

Commercial programs - Airbus



Commercial programs - Airbus