

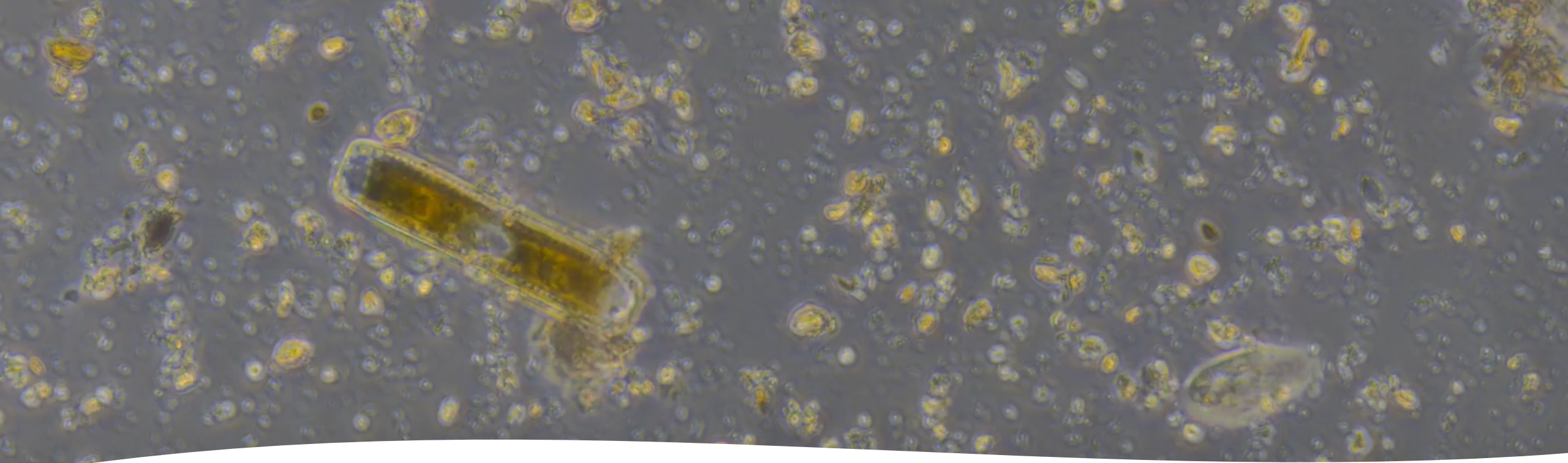
WATER QUALITY MONITORING USING HIGH RESOLUTION MICROSCOPIC IMAGES AND MACHINE LEARNING



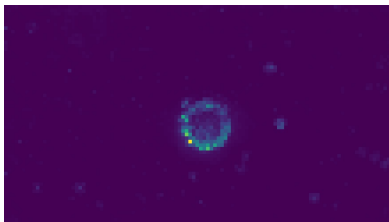
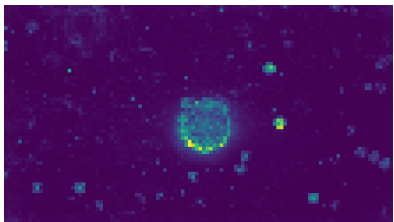
INTERNATIONAL CONFERENCE OF HIGH ENERGY PHYSICS

Renzo Valencia Oyarce – Machine learning and Metagenomics student.

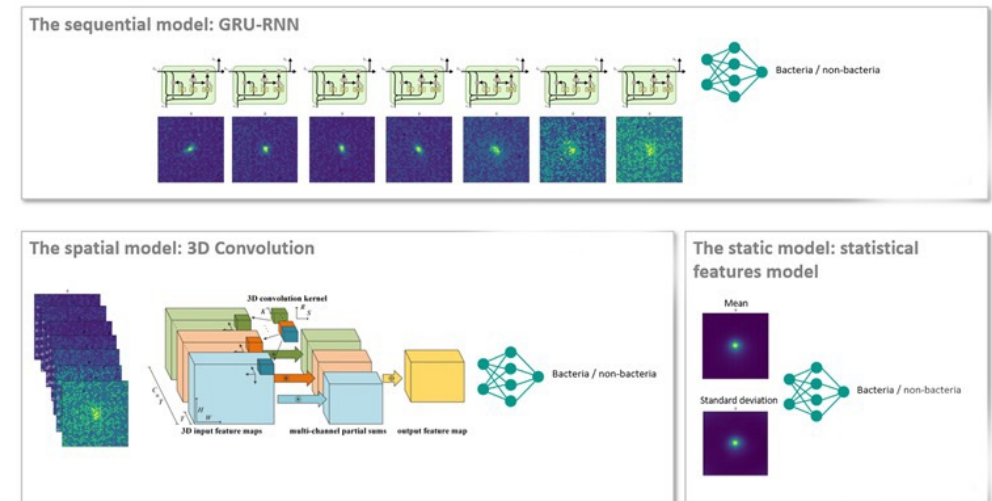
Research funded with the cooperation of **ST>RT-UPCHILE** and  **nVIDIA** INCEPTION PROGRAM



- The new methodologies for microbiological analysis are based on the ravages that climate change is causing in the world.
- The traditional methodology considers extended times from sampling to identification of a hydrobiological pest, which affects risk management against these contaminants in large bodies of raw water.
- A cyanobacteria is approx. 200 μm , which is easily recordable, so using federated learning algorithms and transformed vision it is possible to detect and quantify these systems in real time.



- The Transformation Visualization Module by federated learning has the ability to focus learning based on machine knowledge and not on data.
- The Transformed visualization architecture was designed with the ability to visualize the micro metabolism of cyanobacteria giving an additional feature to the identification of particle and microbiome.
- The identification architecture uses a GRU-RNN sequential model to observe the movement of the object of interest, subsequently it uses a 3D-CNN spatial model for the focus and finally a statistical model to calculate the detected probability of the object.



Procedure, development and results

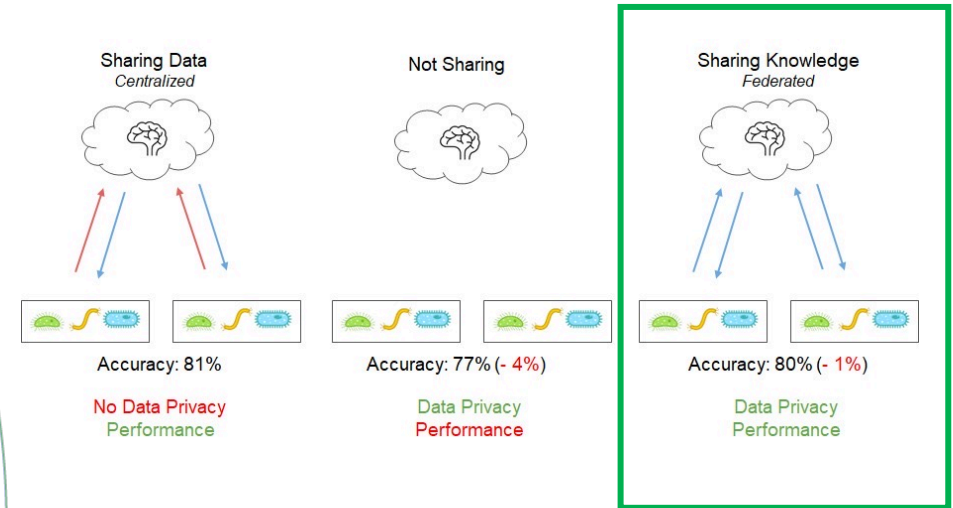
LABORATORY ON A CHIP, the system is implemented on a 1 to 4 channel microfluidic chip of 200 μm area so that the observation has a component of counting and identification.

THE MICROSCOPE used for the analysis was an inverted ZEISS AXIO VERT A.1 with a 40x objective and axiocam 208 color 4K and a peristaltic pressure pump with a flow rate of 0.5 milliliters per second.

MACHINE LEARNING PROCESSING was performed on an Nvidia RTX 3070 Ti GPU with 16GB of Ram. Depending on the microbiological analysis, the response times ranged from 8 microseconds to 3 hours per iteration.

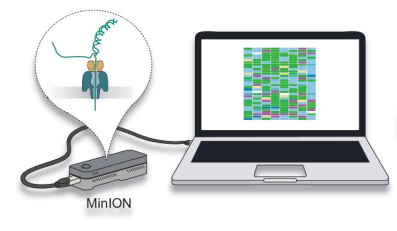
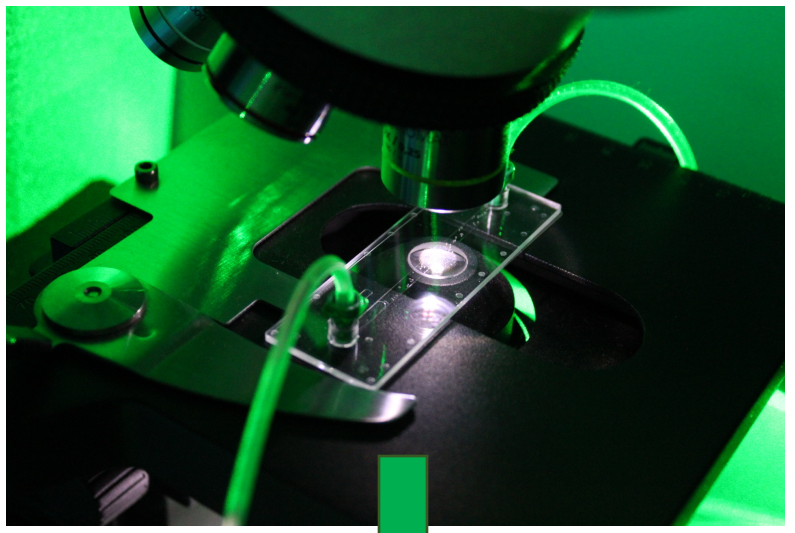
THE MICROBIOLOGICAL ANALYSIS replicated were those described by the WHO within the water resources monitoring plan and are the percentage of Biovolume of cyanobacteria and cells per milliliter.

The architectures compared to supervised and transfer learning can achieve high percentages of precision and accuracy but they are a risk for computer security since they centralize all the data for its operation.

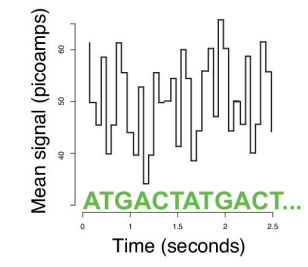


While the federated learning system obtains similar results but manages to ensure the privacy, precision and accuracy of the ML architecture only by **sharing the knowledge of the machine.**

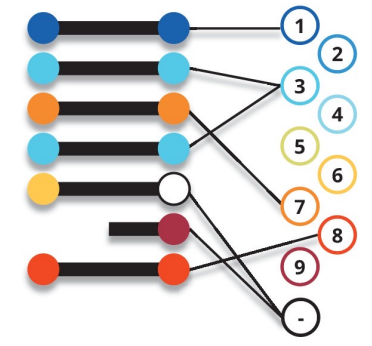
Detection with Machine Learning and Identification with Metagenomics



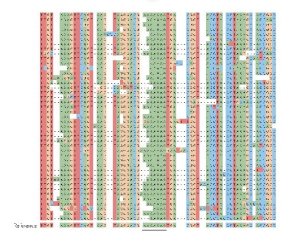
DNA Sequencing
R9.4.1 flow cell with LSK109



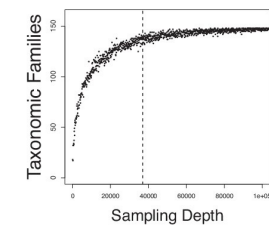
Basecalling
Guppy 3.15 (FlipFlop)



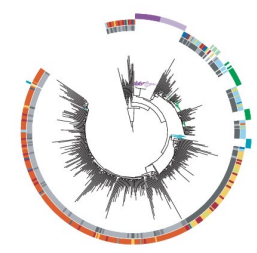
Barcode assignment
Porechop 0.2.4



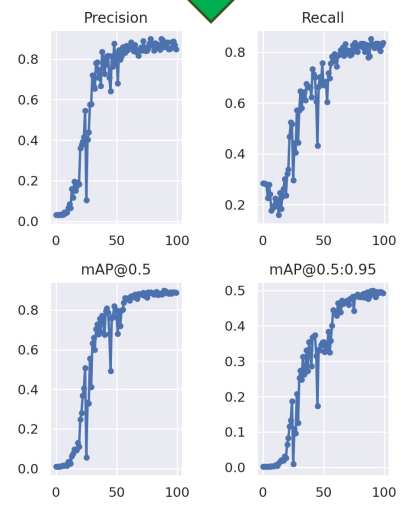
Taxonomic classification
Minimap2 (k = 15) vs. SILVA 132



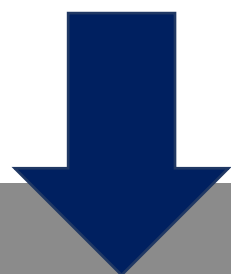
Rarefaction
to 37,000 reads



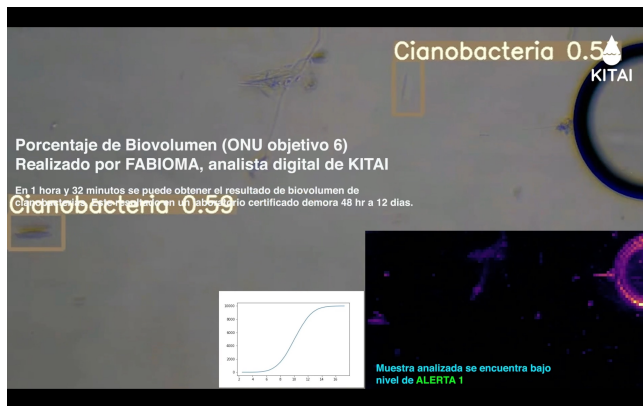
Metagenomic analyses
phylogenetics, spatiotemporal clustering, hydrochemical associations



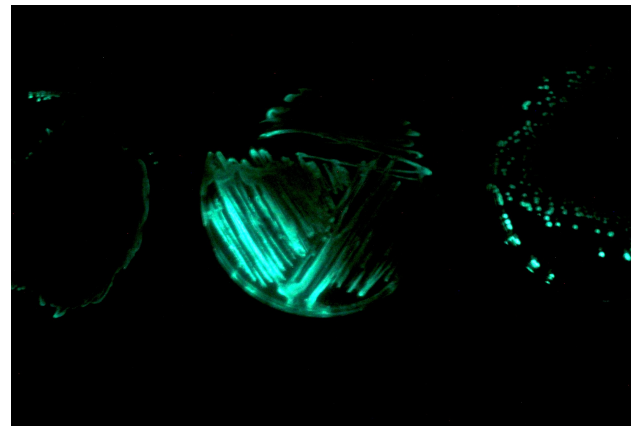
to finally be able...



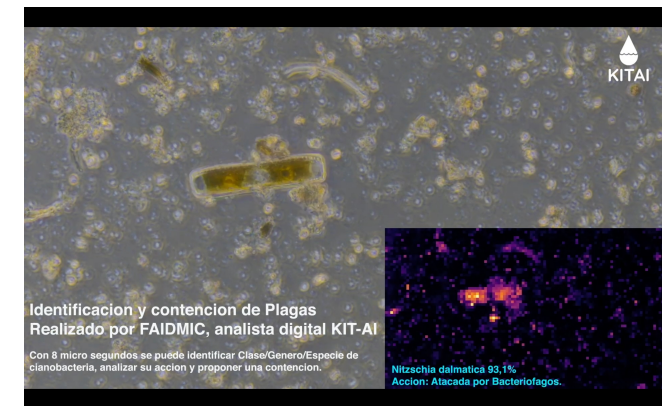
Final Visualization of the first Microbiology Laboratory on a Chip



1



3



2

- 1- Microbiological analysis of cells per milliliter of sample at a flow rate of 1 ml * sec. Accuracy: 98.75% | Sensibility: 98.135%.
- 2- Microbiological analysis of identification and containment of hydrobiological plates with VC and ML system. Accuracy: 95.32% | Sensibility: 93.21%.
- 3- Containment of hydrobiological pest using metagenomics and genetics directed with CRISPR - CAS13. 99.32% effective within the first 24 hours of detection by the AI system.