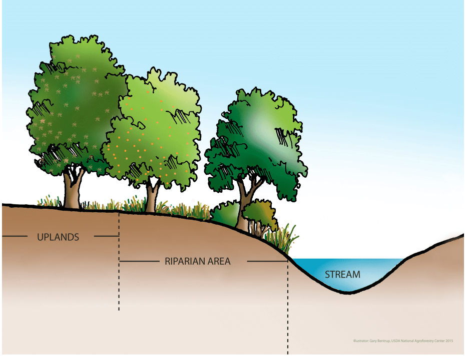
Multimodal Sub-Pixel Classification of Forest Vegetation in Riparian Zones

Luan Casagrande[1]
Roberto Hirata Jr.[1]

[1]Institute of Mathematics and Statistics
University of São Paulo
São Paulo, SP, Brazil;
Email: luanccasagrande@gmail.com, hirata@ime.usp.br

Riparian zones

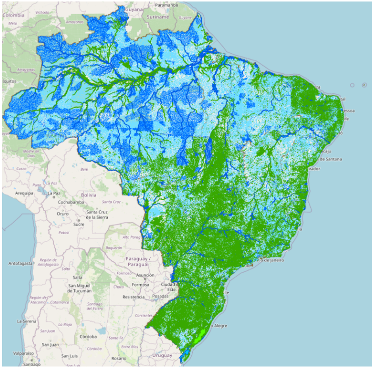
- Riparian zones are three-dimensional zones encompassing the following attributes [1]:
  - ◇ Hydrogeomorphic;
  - ◇ Vegetational;
  - ◇ Food-web.
- Riparian vegetation corresponds to all vegetation units along river networks;
- They perform several functions, such as [2,3]:
  - ◇ Natural corridors for terrestrial wildlife;
  - ◇ Water purification;
  - ◇ Reduce flood vulnerability;
  - ◇ Areas of recreation.



Riparian zones illustration [4].

Monitoring riparian zones

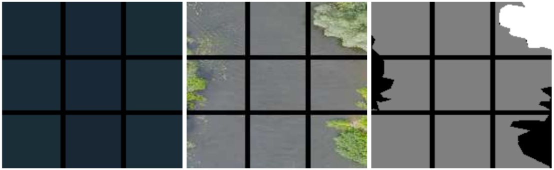
- Regulations were created aiming to protect these areas:
  - ◇ Forest code - Brazil;
  - ◇ Law nº 33/96 from Portugal.
- In 2017, R$ 3 billion in fines in Brazil [5];
- Complexity: 16.6Mha from Brazil is covered by water [6] and Brazil has around 6.9M agricultural units with an average area of 80.8ha [7].
- Quickly and accurately mapping riparian zones is necessary to guarantee that these regulations are being respected;



Water bodies (green) and watercourses (blue) [8].

Strategies to map riparian zones

- Manually map on site;
- Remote sensing:
  1. Unmanned Aerial system:
     - ◇ Very-high spatial resolution;
     - ◇ Most expensive option, low swath, and low spectral resolution.
  2. Satellite
     - ◇ Most affordable option, high swath, and medium spectral resolution;
     - ◇ Generally it has a poor spatial resolution.

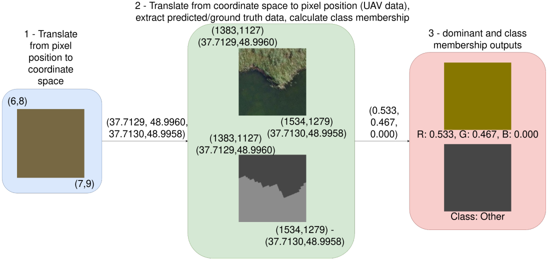


(a) *Satellite data.* **(b)** *UAV data.* **(c)** *Labels.*

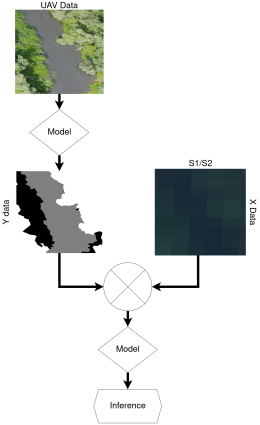
Comparison between satellite and UAV data against labels.

Strategy based on the synergy between UAV and Satellite

- UAVs' advantages appear to compensate for the satellites' disadvantages, and vice versa [9].
- Why not take advantage of both?
- Satellite has great potential to cover large areas and can provide a better spectral resolution;
- UAVs can be used to acquire very high resolution data and help us to better define the meaning of each pixel;



Example of dominant and class membership label calculation.



Proposed strategy diagram.

Dataset

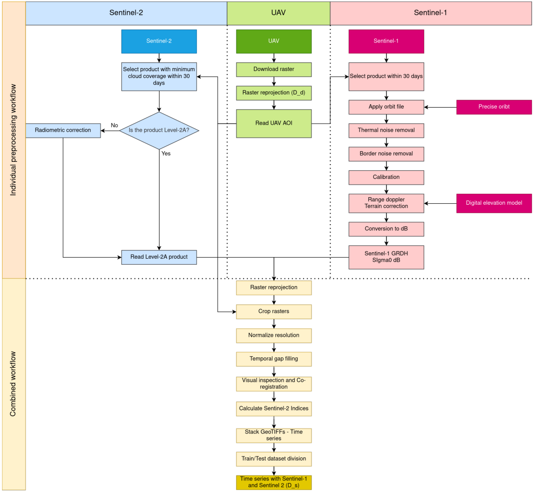
- **A new riparian-zones dataset that correlates UAV, Sentinel-1, and Sentinel-2 data is proposed in this work;**
- The dataset is composed of data acquired in 10 different places in Europe and South America with a main focus on Brazil (4 site locations) covering $3.605 km^2$;
- The dataset spans seven months, with data collected once a month;
- Classes: Water, forest and woodland, and other;
- Data sources:
  - ◇ UAV: OpenAerialMap (OAM);
  - ◇ Satellite: Open-Access Copernicus hub.

Location	Latitude	Longitude	UAV Date	Res.	Area	Test	Train
Croatia	43.4048	16.7895	2020-06-20	3	5.2	33.9	66.1
Russia (1)	52.7210	44.3982	2017-08-17	2	11.8	32.1	67.9
Russia (2)	54.6802	35.0805	2017-07-15	6	100.5	30.9	69.1
Ukraine	48.9939	37.71018	2020-07-30	5	41.2	25.0	75.0
Belarus	53.9612	27.5941	2020-06-20	3	22.0	13.6	86.4
Denmark	54.9677	11.5607	2020-04-11	3	13.1	28.0	72.0
Brazil (1)	-21.6771	-43.3120	2019-09.27	7	44.6	30.2	69.8
Brazil (2)	-10.9733	-58.3108	2020-01-28	4	33.8	32.0	68.0
Brazil (3)	-9.5881	-60.2143	2020-01-27	4	41.7	33.2	66.8
Brazil (4)	-10.6224	-58.0940	2020-01-28	2	46.6	34.6	65.4

Drone and Sentinel-2 details. The total UAV area is $3.605 km^2$.

Dataset - preprocessing workflow

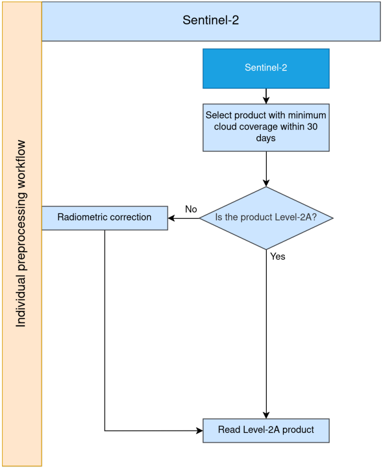
- Our proposal for dataset pre-processing workflow has two main phases:
  - Individual preprocessing workflow;
  - Combined workflow;
- UAV preprocessing workflow is simplified:
  - Reprojection (EPSG:4326);
  - Area of interest (AOI) extraction;



Dataset composition workflow.

Dataset - Sentinel 2 preprocessing workflow

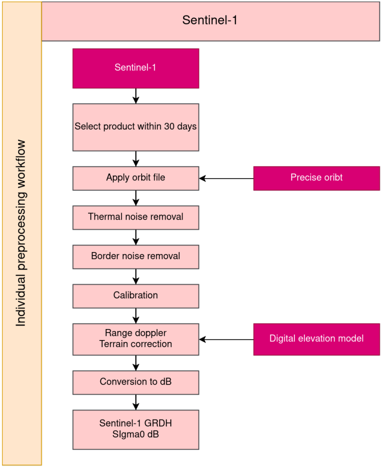
- Only images with minimal/no visible cloud cover over the area of interest were used;
- Only Level-2A products were used in the workflow since they provide the bottom of atmosphere reflectance;
- Level-1C (L1C) top of atmosphere are corrected using Sen2Cor toolbox from European Space Agency, which performs:
  - ◇ Atmospheric correction;
  - ◇ Terrain correction;
  - ◇ Cirrus correction.



Sentinel-2 workflow.

Dataset - Sentinel 1 preprocessing workflow

- Level-1 Ground Range Detected (GRD) characteristics:
  - ◇ Interferometric wide mode;
  - ◇ Descending orbit direction;
  - ◇ Dual VV+VH polarisation.
- SNAP from the European Space Agency was used to pre-process.



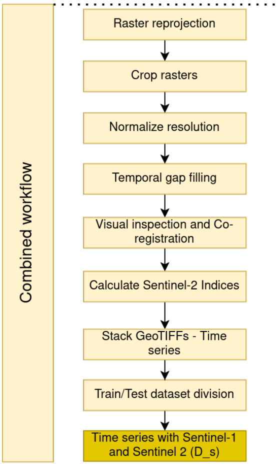
Sentinel-1 workflow.

Dataset - Combined workflow

- All rasters are reprojected to EPSG:4326;
- Rasters that have a resolution different than 10m are re-projected;
- Missing rasters are filled using temporal gap filling, where the interpolated reflectance value ρ_j (between ρ_i and ρ_k) is computed by:

$$\rho_j = \frac{(t_j - t_i) * (\rho_k - \rho_i)}{(t_k - t_i)} + \rho_i \tag{1}$$

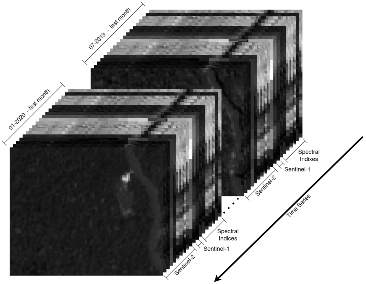
- Assess co-registration needs based on visual inspection: co-registration was not necessary;



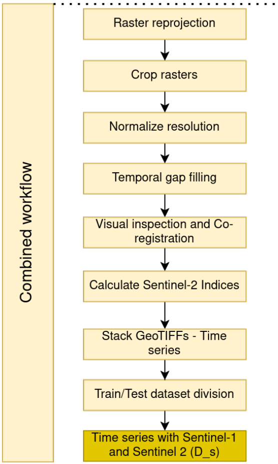
Combined workflow.

Dataset - Combined workflow (Cntd.)

- Spectral indices calculation: NDVI, IB, SAVI, NDWI, EVI, EVI2, GNDVI, and NDMI;
- Finally, data is stacked and split into training and testing sets.



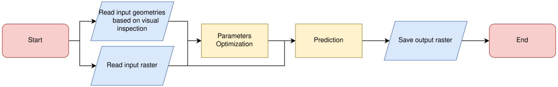
Brazil (3) data cube. Per month there are 20 bands: 10 Sentinel-2 (B, G, R, RedEdge at 704nm, 740nm, 783nm, and 865nm, NIR, SWIR bands at 1610nm and 2190nm), Sentinel-1, and 8 Spectral indices, totaling 140 bands.



Combined workflow.

Ground truth composition

- Ground truth data was composed based on UAV data;
- k-means clustering with k-means++ initialization followed by visual inspection and manual fixing were used to compose labels;

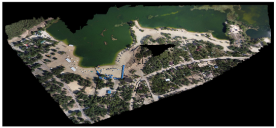


Flowchart ground truth composition algorithm.

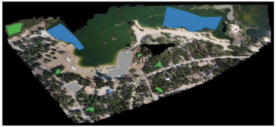
- A small percentage of data was labeled to be used as a reference by the clustering approach;
- Number of clusters and best vegetation index or color space band are selected using a cartesian grid search.

Parameter	Range	Step
Vegetation Indexes or Color Spaces	Excess of Green (Woebbecke *et al.*, 1995) RGBVI (Bendig *et al.*, 2015) Hue from HSV A from CIELAB B from CIELAB	NA
Number of clusters	2-10	1

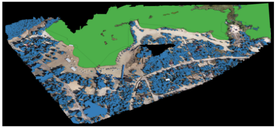
Flowchart ground truth composition algorithm.



(a) *Input UAV data - Donetsk Oblast, Ukraine.*



(b) *Input annotations over the UAV Data.*

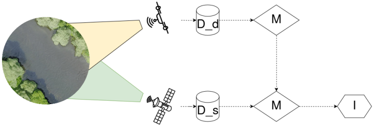


(c) *Ouptut from semi-supervised approach.*

Inputs and output - ground truth.

UAV and Satellite Synergy for Riparian Zone Classification

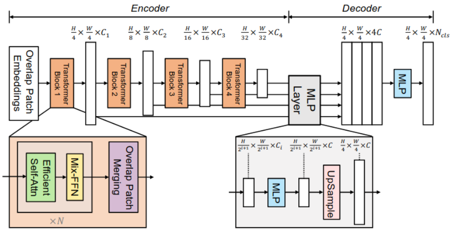
- A new approach using the combination of data acquired by UAVs and Satellite, and SegFormer and a Deep learning based class membership (CM) classifier. Main differences:
  - ◇ Replace labor-intensive OBIA based approach;
  - ◇ Optimize NN architecture using Neural Architecture Search (NAS);
  - ◇ SAR and Spectral indices;
  - ◇ Temporal data.
- Inputs:
  - ◇ 3 Bands GeoTIFF VHR UAV data and
  - ◇ 20(140) Bands GeoTIFF S2/S1 data;
- Output: GeoTIFF raster describing class membership;



Proposed approach pipeline. D is data, M is model, I is interpretation, d is drone, and s is satellite. Adapted from Carbonneau et al. (2020) [11].

Semantic Segmentation Model - SegFormer - UAV

- SegFormer (B4) consists of two main modules:
  - ◇ A hierarchical transformer encoder - high and low resolution features;
  - ◇ A lightweight All Multilayer Perceptron (All-MLP) decoder to fuse these multi-level features.
- Main parameters:
  - ◇ Input patches (512x512 pixels);
  - ◇ Data augmentation was used in the training phase;
  - ◇ Trained using: Adam optimizer and Sparse Categorical Crossentropy loss;
  - ◇ LR (0.0001) with learning decay by 0.1 - patience 10 epochs;
  - ◇ Early stopping - 30 epochs of patience based on validation loss.



SegFormer structure. Extracted from Xie et al. (2021).

Class membership classifier - NAS - Satellite

- DNNs with and without convolutional layers are evaluated;
- NAS is used to design and optimize the architectures. The search spaces are parameterized by:
  - ◇ The number of layers;
  - ◇ The type of operation;
  - ◇ The number of filters, kernel size, number of units, and/or other specific parameters.
- Regularized evolution (population 100, mutation probability 0.05, 2000 trials) is used as a search strategy.

	CNN	DNN
Num. of Dense Layers	1-5	
Num. of Neurons	8,16,32,64,128,256,512,1024,2048	
Num. of Conv. Layers	1-5	NA
Depthwise Separable flag	0,1	NA
Kernel Size	1,3	NA
Num. of out channels	2,4,6,8,16,32,64	NA
Activation function	ReLU, LeakyReLU, Hardswish, None	
Droupout value	0,0.25,0.5,0.75	
Window size	1,3,5,7,9,11	NA

Search space based on chain-structured neural networks, where NA means not applicable.

Summary of experiments

1. Semantic segmentation - SegFormer;
2. Calibrated classifier:
   - Model performance without temporal data;
   - Model performance with temporal data;
   - Ground truth as the target variable.

Calibration input Temporal resolution	SegFormer input Without temporal data		SegFormer input With temporal data		GT input Without temporal data		GT input With temporal data	
input data//model type	CNN	DNN	CNN	DNN	CNN	DNN	CNN	DNN
Sentinel-1	2	2	14	14	-	-	-	-
Sentinel-2	10	10	70	70	-	-	-	-
Sentinel-1 Spectral indices	10	10	70	70	-	-	-	-
Sentinel-2 Spectral indices	18	18	126	126	-	-	-	-
Sentinel-1 Sentinel-2	12	12	84	84	-	-	-	-
Sentinel-1 Sentinel-2 Spectral indices	20	20	140	140	20	-	140	-

Number of features per combination evaluated, where w/o means without and w/ means with.

Semantic segmentation - SegFormer

- Model has learned meaningful patterns - almost perfect level of agreement;
- Class water has the lowest IoU;
- Such differences between IoUs per class impact the calibrated classifier performance.
- Best results: Brazil (2) and (4). Worst results: Denmark and Brazil (3).

Class	Other	Water	Woodland and forest
IoU	0.868	0.784	0.950
OA	95.2%		
κ	0.910		

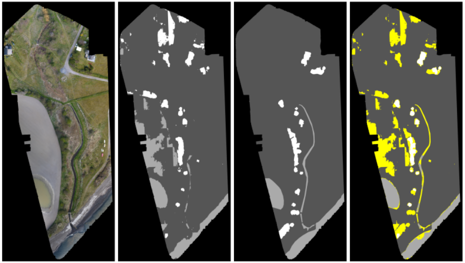
Class-specific IoU, OA, and kappa statistic - SegFormer model.

location	IoU - other	IoU - Water	IoU - Woodland and forest	OA (%)	kappa
Croatia	0.848	0.572	0.735	88.1	0.739
Russia (1)	0.926	0.729	0.859	94.3	0.873
Russia (2)	0.851	0.900	0.831	92.5	0.885
Ukraine	0.838	0.931	0.696	90.7	0.851
Belarus	0.937	0.904	0.890	95.8	0.917
Denmark	0.883	0.449	0.270	89.1	0.508
Brazil (1)	0.947	0.797	0.001	95.3	0.812
Brazil (2)	0.081	0.981	0.998	99.8	0.984
Brazil (3)	0.294	0.905	0.904	91.0	0.635
Brazil (4)	0.023	0.969	0.998	99.8	0.973

Class-specific IoU, OA, and kappa statistic for the different locations of the dataset.

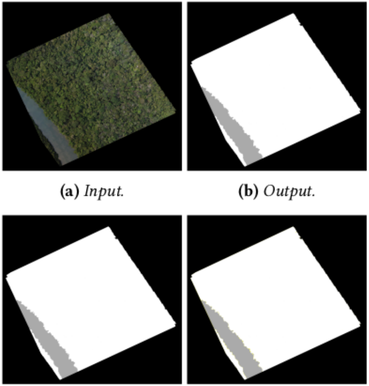
Semantic segmentation - SegFormer (Cntd.)

- Brazil (2): miss-predictions happened in the boundary between classes;
- Denmark: major miss-predictions for classes water and woodland and forest;
- The errors from the second case are significant enough to impact dominant/sub-dominant classes.



(a) *Input.* (b) *Output.* (c) *Ground truth.* (d) *Difference.*

Denmark - evaluation area.



(a) *Input.* (b) *Output.*

(c) *Ground truth.* (d) *Difference.*

Brazil (2) - evaluation area.

Without Temporal Data - Dominant class

- Our bests: CNN-based S1-S2-SIs and DNN-based S1-SIs;
- Our DNN and CNN architectures using only S2 input outperform related works with statistically significant differences - NAS impact;
- Our bests significantly outperform DNN-S2; S1 or S1, SIs, and spatial resolution can improve dominant class results.

Convolutional Neural Network						
	S1	S2	S1-SIs	S2-SIs	S1-S2	**S1-S2-SIs**
OA	81.39	95.44	95.69	95.53	95.60	**95.76**
κ	0.616	0.908	0.914	0.910	0.913	**0.915**
Deep Neural Network						
OA	72.05	95.22	**95.60**	95.37	95.46	95.53
κ	0.384	0.904	**0.912**	0.907	0.909	0.911

OA and κ statistic for evaluated combinations without temporal resolution.

	Furuya *et al.* (2020)				Carbonneau *et al.* (2020)				Our S2		Our Best	
M	SVM	DT	RF	NB	C-1	C-2	CM-1	CM-2	CM1	CM-2	CM1	**CM-2**
OA	90.71	92.69	93.01	85.86	94.30	93.88	93.97	94.10	95.22	95.44	95.60	**95.76**
κ	0.808	0.854	0.859	0.724	0.884	0.877	0.877	0.880	0.904	0.908	0.912	**0.915**

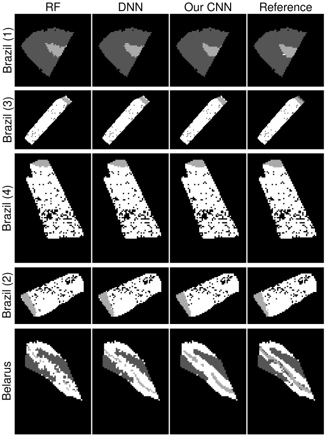
Related works and our best OA and κ statistic, where C is crispy, CM is class membership, 1 is DNN, 2 is CNN classifier, and S2 is Sentinel 2.

Without Temporal Data - Dominant class (Cntd.)

- Major difference in the Belarus location: our results are less noisy;
- Our best was the only one capable to partially predict water body in Brazil 3;
- Denmark: Worst SS performance. Only case that our solution was not superior.

	RF - Furuya *et al.* (2020)					DNN - Carbonneau *et al.* (2020)					CNN-S1+S2+SIs				
	OA	k	C1	C2	C3	OA	k	C1	C2	C3	OA	k	C1	C2	C3
Belarus	73.3	0.528	0.660	0.153	0.603	80.4	0.657	0.715	0.425	0.691	**85.4**	**0.750**	**0.748**	**0.642**	**0.766**
Brazil (1)	97.1	0.848	0.969	0.780	0.000	97.2	0.853	0.970	0.781	0.000	**97.9**	**0.879**	**0.976**	**0.819**	**0.000**
Brazil (3)	96.5	0.735	0.000	0.763	0.974	97.2	0.785	0.000	0.802	0.980	**97.6**	**0.816**	**0.111**	**0.810**	**0.982**
Brazil (4)	98.8	0.878	0.000	0.884	0.987	99.5	0.944	0.000	0.899	0.994	**99.8**	**0.976**	**0.000**	**0.955**	**0.997**
Brazil (2)	98.1	0.906	-	0.869	0.980	98.5	0.929	-	0.890	0.983	**99.2**	**0.961**	-	**0.933**	**0.991**
Denmark	**100.0**	**1.000**	**1.000**	**1.000**	-	**100.0**	**1.000**	**1.000**	**1.000**	-	88.9	0.727	0.857	0.667	-
Croatia	80.0	0.615	0.667	0.667	-	80.0	0.615	0.667	0.667	-	**100.00**	**1.000**	**1.000**	**1.000**	-
Ukraine	77.1	0.604	0.688	0.716	0.226	75.7	0.599	0.640	0.775	0.254	**82.0**	**0.671**	**0.729**	**0.818**	**0.162**
Russia (2)	**100.0**	**1.000**	**1.000**	-	-	**100.0**	**1.000**	**1.000**	-	-	**100.00**	**1.000**	**1.000**	-	-
Russia (1)	67.3	0.449	0.350	0.429	0.615	67.3	0.381	0.500	0.222	0.628	**74.5**	**0.504**	**0.500**	**0.421**	**0.690**

OA, κ statistic, and IoU, where C1 represents class other, C2 class water, and C3 woodland and forest. Cells filled with dash (-) represent that the value is not applicable.



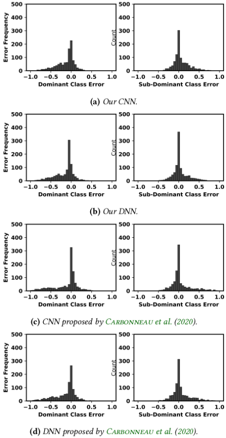
Maps representing the dominant class: RF proposed by Furuya et al. (2020) [10], DNN proposed by Carbonneau et al. (2020) [11], and our best.

Without Temporal Data - Class membership

- CNN-based model proposed by Carbonneau et al. (2020) [11] produced the lowest MAE for the D class;
- Our DNN-based architecture achieved the lowest MAE for the SD class when considering all pixels;
- Our models achieved the lowest MAE for D and SD classes, showing higher potential to predict subtle changes.

		Paper Carbonneau *et al.* (2020)		Our	
	Heterogeneous + homogeneous pixels				
	Metric	DNN	CNN	DNN	CNN
D	ME	-8.87	**-5.42**	-9.52	-8.05
	MAE	9.84	**6.95**	10.15	8.85
SD	ME	3.56	2.40	**2.11**	2.71
	MAE	5.19	4.24	**3.65**	4.11
	Heterogeneous pixels				
D	ME	-13.46	-11.20	-15.32	**-13.3**
	MAE	18.61	19.34	18.69	**17.57**
SD	ME	5.24	4.91	**3.98**	5.69
	MAE	13.94	14.77	**12.20**	13.16

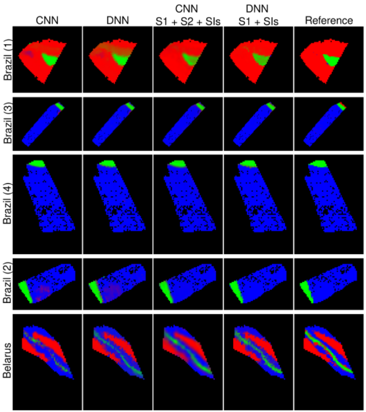
ME and MAE for dominant (D)/sub-dominant (SD) classes.



Error frequency per model.

Without Temporal Data - Class membership (Cntd.)

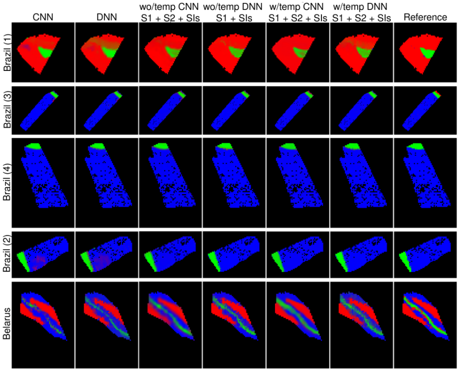
- Our model's certainty is smaller than the ones proposed in related works;
- Belarus: our models have managed to predict the dominant class water in a more accurate way;
- Our models had a better performance in sub-dominant class prediction for Brazil (1) and (2);
- Misprediction clusters in related works (SD) may misguide conservation efforts and undermine confidence in the maps.



Maps show class percentages without temporal data. CNN and DNN represent CM models proposed by Carbonneau et al., 2020 [11]. Green pixels indicate water, red for other classes, and blue for woodland/forest. Each pixel contains a mix of all classes.

With Temporal Data - Class membership (Cntd.)

- CNN-based results are smoother than DNN-based results;
- We can notice major differences in Belarus:
  - Our models without temporal data predicted the river more accurately than the reference;
  - Same relation between with and without temporal data;



Maps representing a class percentage. CNN and DNN represent class-membership models proposed by Carbonneau et al., 2020 [11]. Reference is reference labels for test areas. Green pixels represent water red other, and blue woodland and forest. Each pixel is represented by a mixture of all classes.

Conclusions

- This work presents a novel two-stage approach that takes advantage of the synergy between UAV and Satellite using a CNN-based class membership classifier calibrated by SegFormer predictions made in UAV data;
- A new dataset combining UAV, Sentinel-1, and Sentinel-2 data collected from ten sites across Europe and South America is introduced in this work;
- Compared to the reproduced works, our best combinations produced distinctly different and superior results for the dominant class;
- We can predict class membership before it is dominant, aiding resource management in proactive conservation;
- Besides being superior, our approach reduces human impact in the pipeline through a semantic segmentation model;

Conclusions

- The comparisons performed show that:
  - Neural architecture search effectively addressed variations in the target variable, yielding a superior result with statistically significant differences compared to Carbonneau et al. (2020) [11];
  - Incorporating S1 data, or S1, SIs, and spatial resolution can improve dominant class results;
  - The inclusion of temporal data in the proposed model had a significant impact on the performance particularly when using 3D CNNs (Spatial-temporal CNN);
- Limitations:
  - Limited search space: search space for parameter optimization may have constrained the model's performance potential;
  - Cloud coverage: Products impacted by cloud coverage were not included in this study;
  - Seasonal variability: We have not evaluated the solution's potential to handle seasonal variability.
- Future works:
  - Increase search space;
  - Evaluation of spatial-contextual models (Semantic segmentation step);
  - Extend temporal analysis and investigate RNN alternatives for temporal patterns;
  - Increase dataset aiming to evaluate solution's potential in seasonal variability;

References

1 G. Stanley et al. An Ecosystem Perspective of Riparian Zones. BioScience, v. 41, n. 8, p. 540–551, Set 1991. Acesso em: 9 maio 2020.

2 B. Grizzetti et al. "Relationship between ecological condition and ecosystem services in European rivers, lakes and coastal waters". Science of The Total Environment 671 (June 2019), pp. 452–465.

3 Lorna J. Cole, Jenni Stockan, and Rachel Helliwell. "Managing riparian buffer strips to optimize ecosystem services: a review". Agriculture, Ecosystems, and Environment 296 (July 2020), p. 106891.

4 Nature works park. Riparian Buffer. In: Riparian Buffer. Available in: https://natureworkspark.org/project/riparian-buffer/. Last access: 29 mar. 2023.

5 Agência Câmara de Notícias. Governo vai reduzir em até 60% valor das multas por crimes ambientais, diz ministro do Meio Ambiente. 2017. url: https: //www.camara.leg.br/noticias/520617-governo-vai-reduzir-em-ate-60-valor-das-multas-por-crimes-ambientais-diz-ministro-do-meio-ambiente/.

6 A dinâmica da superfície de água do território brasileiro. Projeto mapbiomas, 2021.

7 Husson, E.; Ecke, F.; Reese, H. Comparison of manual mapping and automated object-based image analysis of non-submerged aquatic vegetation from very-high-resolution UAS images. Remote Sensing, MDPI AG, v. 8, n. 9, p. 724, sep 2016.

8 Agência Nacional de Águas e Saneamento Básico. Sistema Nacional de Informações sobre Recursos Hídricos (SNIRH). 2023. url: https://www.snirh.gov.br/.

9 Emilien Alvarez-Vanhard, Thomas Corpetti, and Thomas Houet. "UAV and satellite synergies for optical remote sensing applications: a literature review". Science of Remote Sensing 3 (June 2021), p. 100019.

10 Danielle Elis Garcia Furuya et al. "A machine learning approach for mapping forest vegetation in riparian zones in an atlantic biome environment using sentinel-2 imagery". Remote Sensing 12.24 (Dec. 2020), p. 4086.

11 P. E. Carbonneau et al. "UAV-based training for fully fuzzy classification of sentinel-2 fluvial scenes". Earth Surface Processes and Landforms 45.13 (Aug. 2020), pp. 3120–3140.

Combining UAV and Multi-Source Satellite Data for Sub-Pixel Classification of Forest Vegetation in Riparian Zones

Thank you!

Luan Casagrande
Advisor: Prof. Dr. Roberto Hirata Jr.

Related works

- Two works were reproduced for comparison purposes;
- Furuya et al. proposed a comparison between multiple classifiers to map Riparian zones using Sentinel-2 data:
  - ◇ Evaluated Random Forest (RF), Decision Tree (DT), Support Vector Machine (SVM), and Normal-Gaussian Bayes (NB) to classify;
  - ◇ Best results achieved with DT;
  - ◇ Problems reported: Sparse vegetation, soil brightness, and types of vegetation covers that were not included in the training dataset.
  - ◇ Main differences: Synergy, sub-pixel classification, additional data types, and time-series.
- Carbonneau et al. proposed a comparison between fuzzy and crispy classifiers to classify fluvial scenes using Sentinel-2 data calibrated by UAV:
  - ◇ Evaluated CNNs and DNNs - fuzzy and crispy classifiers;
  - ◇ Best results achieved with CNN - Fuzzy classifier;
  - ◇ Problems reported: a high percentage of vegetation in wetted areas, limited dataset (regional scale), and lack of significant seasonal variability in the data.
  - ◇ Main differences: target variable - satellite model, additional data types, time-series, neural architecture search.

With Temporal Data - Dominant class

- Our bests: CNN and DNN based S1-S2-SIs;
- In 3 out of 6 cases, the improvement between without and with temporal data is statistically significant (CNNs) - Combinations with S1;
- Temporal data does not significantly enhance dominant class accuracy for most combinations using DNN-based architecture;
- Our best models with temporal data are better and statistically different than related works.

	Convolutional Neural Network				Deep Neural Network			
	w/temp		wo/temp		w/temp		wo/temp	
	OA	κ	OA	κ	OA	κ	OA	κ
S1	85.48	0.694	81.39	0.616	72.52	0.384	72.05	0.384
S2	95.76	0.915	95.44	0.908	95.15	0.903	95.22	0.904
S1-SIs	96.13	0.923	95.69	0.914	95.55	0.911	**95.60**	**0.912**
S2-SIs	95.82	0.916	95.53	0.910	95.37	0.907	95.37	0.907
S1-S2	96.14	0.923	95.60	0.913	95.38	0.908	95.46	0.909
S1-S2-SIs	**96.18**	**0.924**	**95.76**	**0.915**	**95.62**	**0.912**	95.53	0.911

OA and κ statistic for evaluated combinations with temporal resolution.

	FURUYA *et al.* (2020)	CARBONNEAU *et al.* (2020)	Our best wo/temp	Our best w/temp
M	RF	C-1	CNN - S1-S2-SIs	CNN - S1-S2-SIs
OA	93.01	94.30	95.76	96.18
κ	0.859	0.884	0.915	0.924

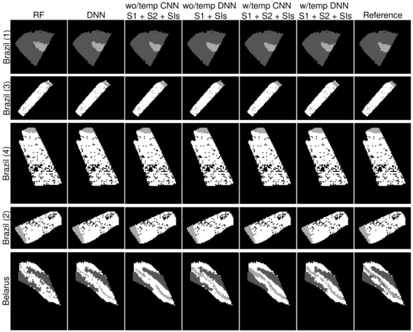
OA and κ statistic for best results achieved for reference works, without temporal data, and with temporal data.

With Temporal Data - Dominant class (Cntd.)

- Our best CNN: best performance in 6/10 locations;
- Without temporal data was superior in 3/10 locations. Two are among the smallest areas;
- Our CNN with temporal data underperforms related works in only one case (Denmark);
- Belarus: Significant differences appear over or near the river, with the best temporal-data combination closest to the reference.

	wo/temp				w/temp			
	DNN - S1-SIs		CNN - S1-S2-SIs		DNN - S1-S2-SIs		CNN - S1-S2-SIs	
	OA	κ	OA	κ	OA	κ	OA	κ
Belarus	80.4	0.657	85.4	0.750	85.0	0.749	**85.8**	**0.766**
Brazil (1)	97.2	0.853	97.9	0.879	98.0	0.887	**98.4**	**0.906**
Brazil (3)	97.2	0.785	97.6	0.816	97.6	0.815	**98.7**	**0.902**
Brazil (4)	99.5	0.944	**99.8**	**0.976**	99.5	0.944	99.7	0.967
Brazil (2)	98.5	0.929	99.2	0.961	99.3	0.966	**99.6**	**0.977**
Denmark	**100.0**	**1.000**	88.9	0.727	77.8	0.400	88.9	0.727
Croatia	80.0	0.615	**100.0**	**1.000**	80.0	0.615	80.0	0.545
Ukraine	75.7	0.599	82.0	0.671	**82.0**	**0.694**	81.3	0.670
Russia (2)	**100.0**	**1.000**	**100.0**	**1.000**	**100.0**	**1.000**	**100.0**	**1.000**
Russia (1)	67.3	0.381	74.5	0.504	74.5	0.557	**80.0**	**0.619**

OA and κ statistic for best combinations with and without temporal data. Cells filled with dash (-) represent that the value is not applicable.



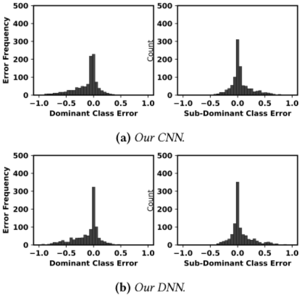
Maps representing the dominant class: RF proposed by Furuya et al. (2020) [10], DNN proposed by Carbonneau et al. (2020) [11], and our best (with and without temporal data).

With Temporal Data - Class membership

- Considerable progress in terms of M for the dominant class with and without homogeneous pixels;
- Small progress for the SD class in heterogeneous pixels classified by the CNN, but a worse performance by the DNN model;

		Paper Carbonneau *et al.* (2020)		Our wo/temp		our w/temp	
	Heterogeneous + homogeneous pixels						
	Metric	DNN	CNN	DNN	CNN	DNN	CNN
D	ME	-8.87	**-5.42**	-9.52	-8.05	-6.43	-7.17
	MAE	9.84	**6.95**	10.15	8.85	7.40	7.97
SD	ME	3.56	2.40	**2.11**	2.71	2.72	3.22
	MAE	5.19	4.24	**3.65**	4.11	4.08	4.52
	Heterogeneous pixels						
D	ME	-13.46	-11.20	-15.32	-13.30	-12.18	**-12.49**
	MAE	18.61	19.34	18.69	17.57	17.30	**16.78**
SD	ME	5.24	4.91	**3.98**	5.69	5.98	5.83
	MAE	13.94	14.77	**12.20**	13.16	13.21	12.81

ME and MAE for dominant (D)/sub-dominant (SD) classes.



Error frequency per model.

Ground Truth - Target Variable

- Performance for models with GT as input is worse than SS, but differences are not statistically significant;
- Optimization with only 2000 combinations, which is considerably smaller than the search space;
- Models with GT achieved the lowest MAE for D and SD classes;
- Same pattern mentioned in the previous two slides: temporal data has helped to improve the results.

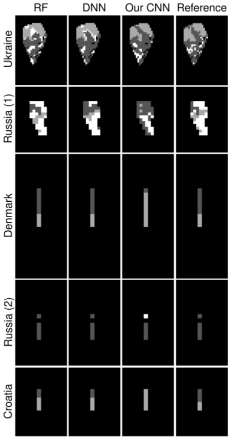
		Semantic Segmentation		GT	
	Heterogeneous + homogeneous pixels				
	Metric	wo/temp	w/temp	wo/temp	w/temp
D	ME	-8.05	-7.17	**-6.57**	**-4.36**
	MAE	8.85	7.97	**7.61**	**5.48**
SD	ME	2.71	3.22	**2.62**	**1.98**
	MAE	4.11	4.52	**4.08**	**3.53**
	Heterogeneous pixels				
D	ME	-13.30	-12.49	**-11.43**	**-10.21**
	MAE	17.57	16.78	**16.97**	**16.18**
SD	ME	5.69	**5.83**	**5.04**	5.14
	MAE	13.16	**12.81**	**12.85**	13.40

OA and κ statistic for the best combinations with and without temporal data, using both the output from SS and GT as the target variable,

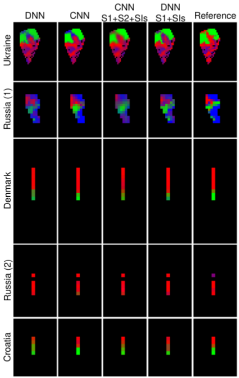
	Semantic Segmentation				Ground Truth			
	w/temp		wo/temp		w/temp		wo/temp	
	OA	κ	OA	κ	OA	κ	OA	κ
S1-S2-SIs	**96.18**	**0.924**	**95.76**	**0.915**	96.07	0.922	95.60	0.913

ME and MAE for dominant/sub-dominant classes.

Appendix - Calibrated classifier - additional outputs

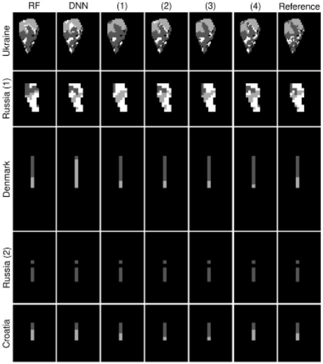


Additional maps representing the dominant class: RF proposed by Furuya et al. (2020), DNN proposed by Carbonneau et al. (2020) and our best.

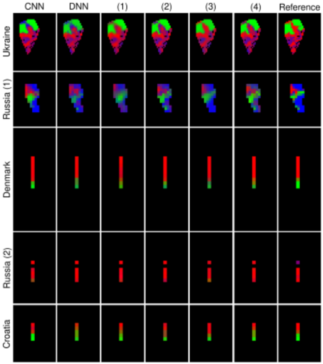


Additional maps representing a class percentage without temporal data. CNN and DNN represent class-membership models proposed by Carbonneau et al., 2020.

Appendix - Calibrated classifier - additional outputs (Cntd.)



Additional maps representing the dominant class: RF proposed by Furuya et al. (2020), DNN proposed by Carbonneau et al. (2020), and our best. (1) is without temporal CNN combination (S1 + S2 + SIs), (2) is without temporal DNN combination (S1 + SIs), (3) is with temporal CNN combination (S1 + S2 + SIs), and (4) is with temporal CNN combination (S1 + S2 + SIs).



Maps representing a class percentage without temporal data. CNN and DNN represent class-membership models proposed by Carbonneau et al., 2020. (1) is without temporal CNN combination (S1 + S2 + SIs), (2) is without temporal DNN combination (S1 + SIs), (3) is with temporal CNN combination (S1 + S2 + SIs), and (4) is with temporal CNN combination (S1 + S2 + SIs).

Appendix - McNemar's test - wo. temporal data

	Furuya et al. (2020)				Carbonneau et al. (2020)			
	SVM	DT	RF	NB	C-1	C-2	CM-1	CM-2
CNN-S2	176.50	72.61	64.09	383.47	19.71	30.61	32.49	32.10
DNN-S2	162.28	73.74	65.36	373.43	26.32	21.66	35.30	16.91
CNN-S1-S2-VIs	211.74	89.25	80.29	398.02	32.49	44.95	45.09	43.58
DNN-S1-SIs	187.46	101.19	87.29	402.75	36.53	33.60	49.51	29.11

McNemar's Test Results (χ) for related works against our best and our with Sentinel 2.

	Furuya et al. (2020)				Carbonneau et al. (2020)			
	SVM	DT	RF	NB	C-1	C-2	CM-1	CM-2
CNN-S2	2.81E-40	1.58E-17	1.19E-15	2.18E-85	9.00E-06	3.15E-08	1.20E-08	1.46E-08
DNN-S2	7.01E-41	8.89E-18	6.23E-16	3.35E-83	4.93E-09	9.10E-08	1.07E-11	7.22E-07
CNN-S1-S2-VIs	5.74E-48	3.47E-21	3.24E-19	1.49E-88	1.20E-08	2.02E-11	1.88E-11	4.06E-11
DNN-S1-SIs	2.46E-38	1.44E-20	8.02E-19	6.49E-89	7.87E-09	4.59E-07	1.15E-10	4.02E-06

McNemar's Test Results (p-values) for related works against our best and our with Sentinel 2.

Appendix - McNemar's test - w. temporal data

		CNN w/temp						DNN w/temp					
		S1	S1-SIs	S2	S1-S2	S1-S2-SIs	S2-SIs	S1	S1-SIs	S2	S1-S2	S1-S2-SIs	S2-SIs
1*	C-DNN	1.01E+02	2.23E+02	1.86E+02	2.16E+02	2.04E+02	1.84E+02	6.75E+02	1.84E+02	1.53E+02	1.75E+02	1.85E+02	1.72E+02
	C-CNN	2.99E+02	5.18E+01	3.22E+01	4.81E+01	4.24E+01	2.94E+01	9.92E+02	3.43E+01	1.94E+01	3.16E+01	4.02E+01	3.54E+01
	CM-DNN	2.71E+02	6.20E+01	4.61E+01	6.08E+01	5.90E+01	4.24E+01	9.56E+02	3.19E+01	1.86E+01	2.60E+01	3.44E+01	2.56E+01
	CM-CNN	2.83E+02	6.54E+01	4.51E+01	5.95E+01	5.50E+01	4.02E+01	9.57E+02	4.48E+01	3.03E+01	4.30E+01	5.16E+01	4.41E+01
2*	SVM	2.99E+02	5.98E+01	4.40E+01	5.83E+01	5.68E+01	3.73E+01	9.91E+02	2.71E+01	1.49E+01	2.20E+01	2.97E+01	2.13E+01
	RF	2.20E+02	1.01E+02	8.09E+01	9.83E+01	9.74E+01	7.34E+01	8.81E+02	8.33E+01	6.84E+01	8.09E+01	8.81E+01	7.50E+01
	NB	4.26E-01	4.47E+02	4.11E+02	4.30E+02	4.41E+02	4.13E+02	3.42E+02	4.02E+02	3.88E+02	3.92E+02	4.11E+02	3.95E+02
	DT	1.84E+02	1.12E+02	9.27E+01	1.13E+02	1.08E+02	8.99E+01	8.36E+02	9.34E+01	7.59E+01	8.52E+01	9.97E+01	8.44E+01

McNemar's Test Results (χ) for combinations with against without temporal data.

		CNN w/temp						DNN w/temp					
		S1	S1-SIs	S2	S1-S2	S1-S2-SIs	S2-SIs	S1	S1-SIs	S2	S1-S2	S1-S2-SIs	S2-SIs
1*	C-DNN	4.33E-67	6.10E-13	1.42E-08	4.01E-12	7.30E-11	5.77E-08	1.19E-217	4.84E-09	1.05E-05	1.85E-08	2.31E-10	2.67E-09
	C-CNN	7.55E-61	3.43E-15	1.11E-11	6.40E-15	1.56E-14	7.44E-11	6.55E-210	1.67E-08	1.61E-05	3.48E-07	4.38E-09	4.14E-07
	CM-DNN	1.33E-63	6.20E-16	1.88E-11	1.22E-14	1.18E-13	2.34E-10	4.20E-210	2.16E-11	3.62E-08	5.58E-11	6.83E-13	3.13E-11
	CM-CNN	4.66E-67	1.04E-14	3.20E-11	2.20E-14	4.95E-14	1.02E-09	2.02E-217	1.90E-07	1.13E-04	2.76E-06	5.06E-08	4.02E-06
2*	SVM	7.67E-24	2.49E-50	2.12E-42	6.93E-49	2.72E-46	7.81E-42	7.54E-149	7.24E-42	3.74E-35	6.85E-40	4.90E-42	2.69E-39
	RF	9.54E-50	1.00E-23	2.43E-19	3.62E-23	5.81E-23	1.05E-17	1.35E-193	7.12E-20	1.30E-16	2.42E-19	6.09E-21	4.80E-18
	NB	5.14E-01	3.58E-99	2.01E-91	1.80E-95	6.27E-98	8.66E-92	2.12E-76	1.83E-89	1.88E-86	3.88E-87	2.77E-91	6.51E-88
	DT	5.29E-42	3.02E-26	5.99E-22	1.98E-26	2.55E-25	2.48E-21	8.55E-184	4.34E-22	2.93E-18	2.66E-20	1.78E-23	4.02E-20

McNemar's Test Results (p-values) for combinations with against without temporal data.

Appendix - McNemar's test - wo. temporal data

SS/GT	w/temp	wo/temp
w/temp	0.22	6.00
wo/temp	1.84	0.47

McNemar's Test Results (χ) - GT against SS as target variable.

SS/GT	w/temp	wo/temp
w/temp	0.64	0.01
wo/temp	0.17	0.49

McNemar's Test Results (p-values) - GT against SS as target variable