武漢大學
WUHAN UNIVERSITY

测绘遥感信息工程国家重点实验室
State Key Laboratory of information Engineering in Surveying , Mapping and Remote Sensing

Cross-probabilistic weak-supervision learning for high-resolution land cover enhancement

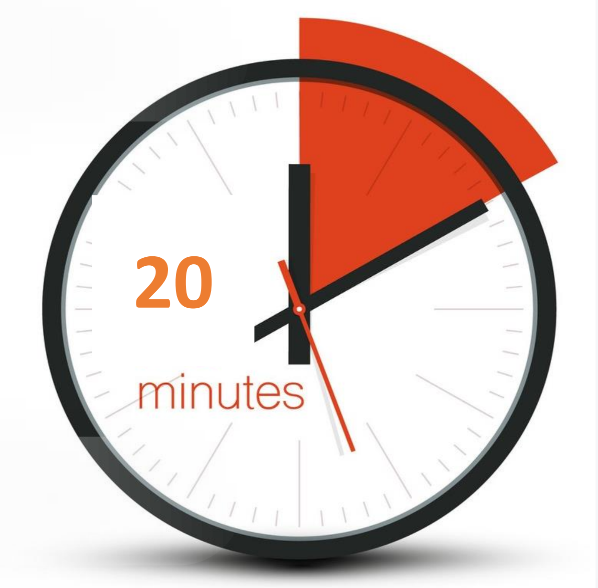
Boaz MWUBAHIMANA, MSc student, C22
~ Planetary Science Group ~

05/24/2024

i

Contents

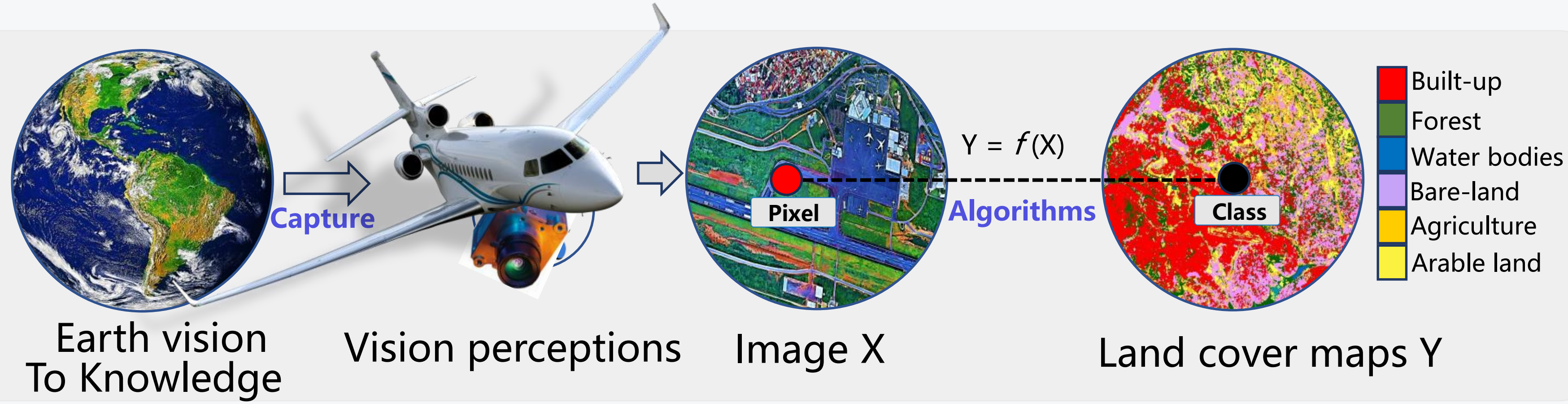
L i s t



01 Background

What is land-cover mapping?

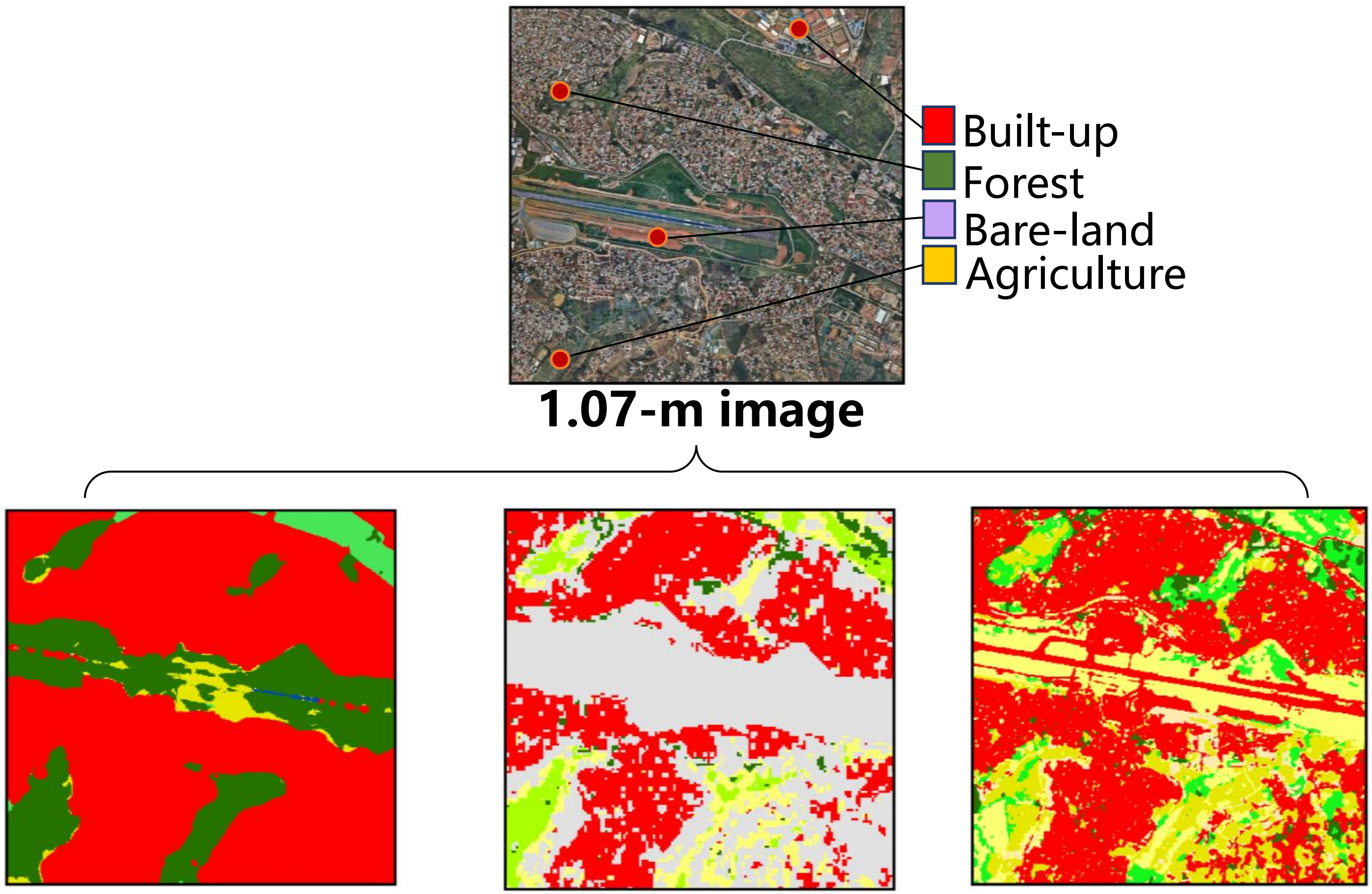
Land cover (maps) represent spatial information on different types (classes) of physical coverage of the Earth's surface,



e.g it may include transitions of land cover classes over time and hence captures land cover changes.

Why VHR LCM is needed?

HR resolution map is a precondition for fine-scale analysis



30-m map
Non detailed information

10-m map
Rough information

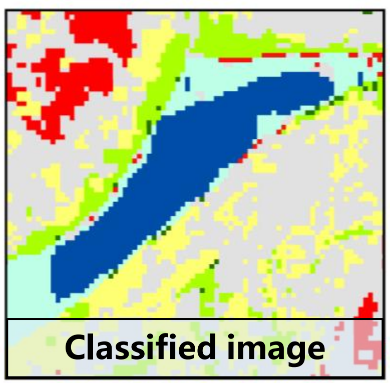
1-m map
Detailed information

Problem to address

- To conduct High resolution land cover map for fine-scale analysis

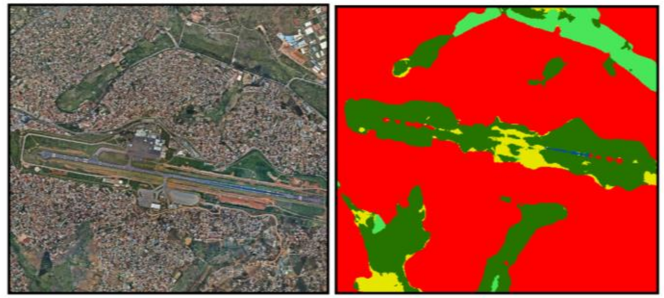
Overhead images

$X = \{ X_{ijc} \} \in R^{HxWxC}$

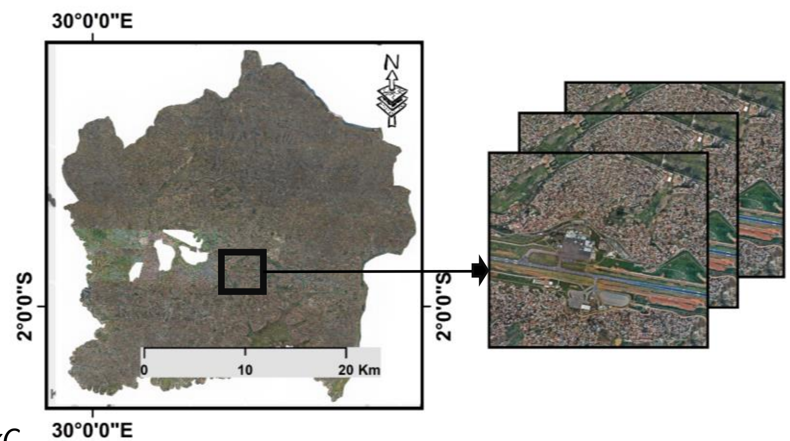


Classified image

$Y = \{ Y_{ijc} \} \in \{1,\ldots L\} R^{HxWxC}$



Multi-scale mismatch trainings



Huge volume dataset

① Lack of precise labels — Label demand

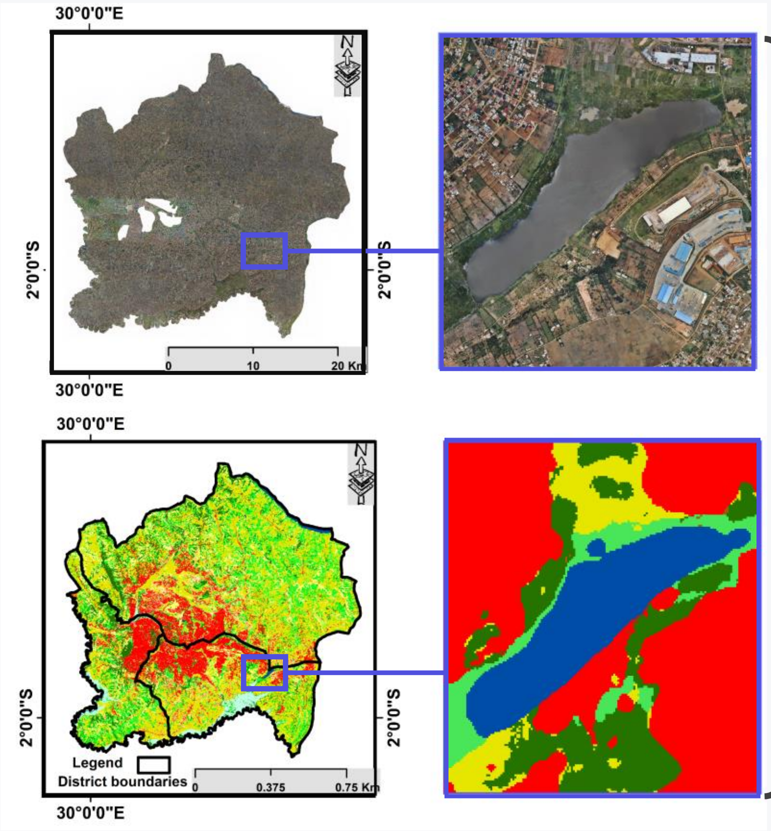
② Cross algorithms — Algorithm demand

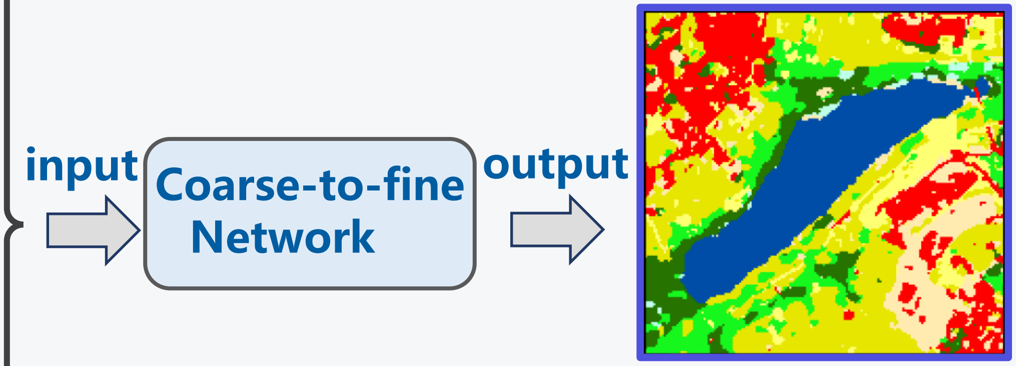
③ High-cost Clusters computation — computational demand

Difficult to break barriers of spatial/spectral coarse resolution

Novel model to address the issue

VHRI + **Pseudo**-labels = HR LCP



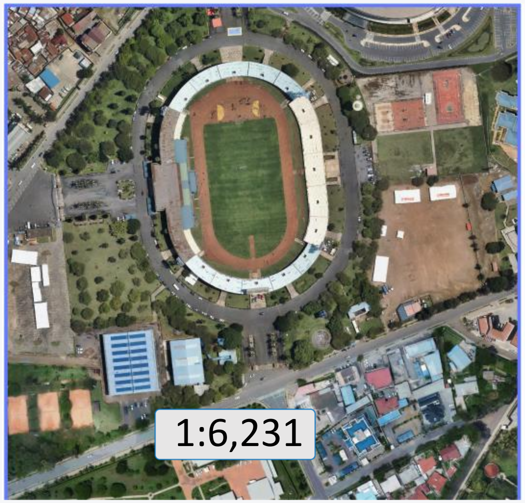


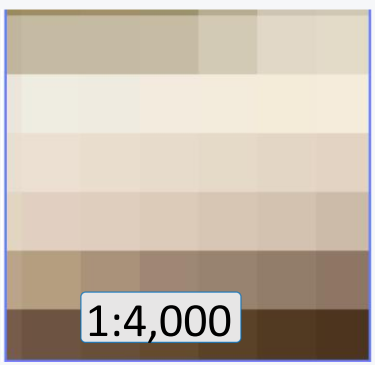
HR Land cover product

Weak-supervision, is all we needed! However,

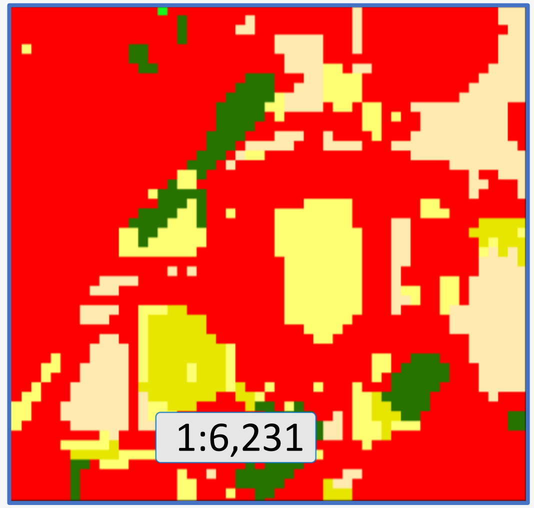
More challenges still hampering the race

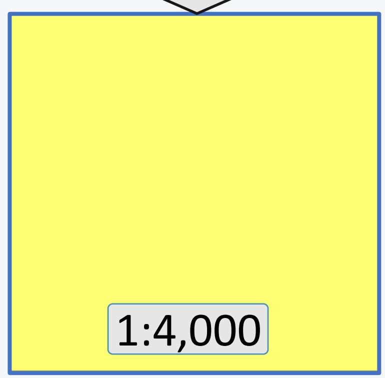
Challenge: Spatial resolution mismatch





36m X 36m pixels





1pixels

?

1:4,000

36m X 36m pixels

Almost 36 pixels represents one

02 Previous studies

Previous work

	Author*s	Problem	Dataset	Spectral	Spatial resolution	Methods	OA%
1	USGS-1990	GLC_P	MODIS	8	1km	CML	79
2	USGS-2000	GLC_P	MODIS-	8	500km	CML	
	Stanford university	GLC_P	Landsat		300	CML	67
3	Chen et., all 2015	GLC_P	GlobaLand30- Landsat	8	30m	CML	86
4	Zhang et all 2021	GLC_P	FCS30- Landsat	8	30m	CML	83
5	Van De Kerchove et., all 2021	GLC_P	ESA- Sentinel	11	10m	CML_R	73
6	Kaarra et., all 2021	GLC_P	ESRI-Sentinel	11	10m	CML_R	85
8	Gong et., all 2019	GLC-FC	Tsinghua university- Sentinel	11	10m	CML_R	83
9	Nduwayezu et.al2019	Ecology CD	Kigali- Worldview, Sentinel	4-11	10	CML	72
10	Li et., all 2023	NSLC_P, SinLoc1	GEI	3	1m	DL	73.6
11	- Ours -	CSLC_P	C2F (Ours), Drone image	4	Sub-meter	DL	-

GLC_P: Global land cover products
LC_P: land cover products
CD: Change detection
Very Low resolution
Low resolution
Low resolution
High resolution

GE: Google Earth Image
CML_R: Classical machine learning, Random forest
NSLC_P: National-Scale land cover products
Fine and detailed resolution

DL: Deep learning
C2F: Frome-Coarse-to- fine
Km, m: kilometer, meter

Research questions and Objectives

Question:

1. Can end-to-end deep learning model be supporting ongoing effort for spaceborne and satellite mission in development of high-resolution land cover products to local-scale?
2. Can end-to-end deep networks learning model be used to develop very high-resolution land cover products in complex urban feature patterns from mismatch of very high-resolution features and pseudo-labels?

Objectives

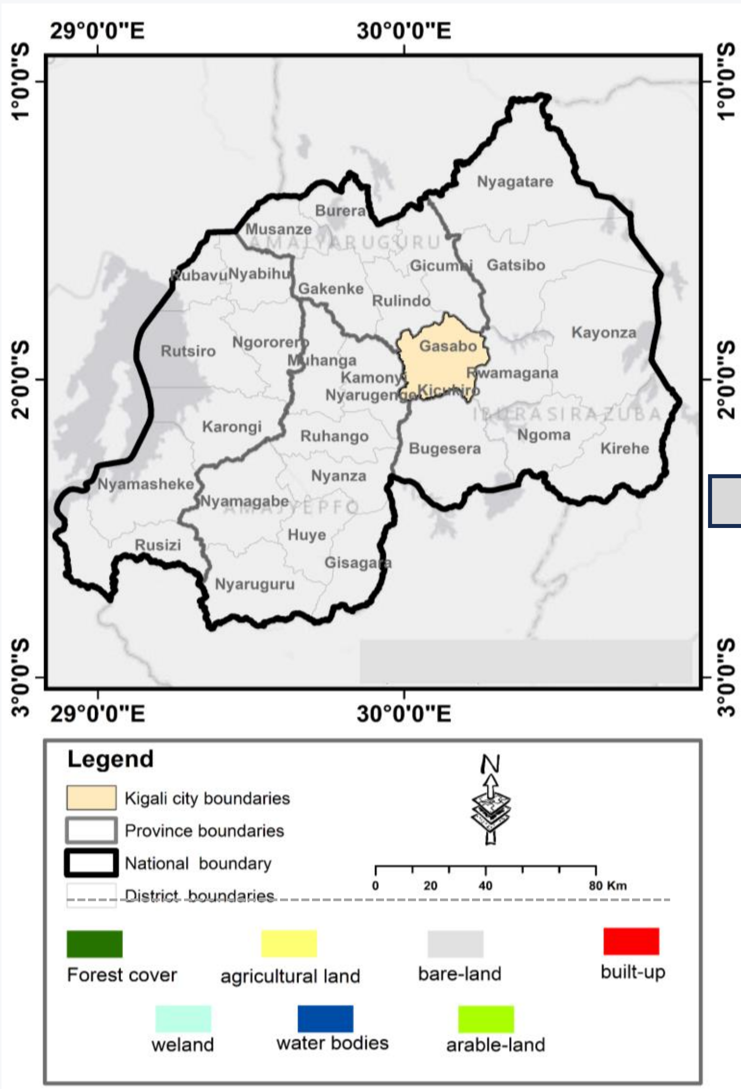
The primary objectives of this research:

1. To break barriers of spatial coarse resolution in development of high spatial-resolution benchmark dataset and high-resolution land cover map for deep learning model training.
2. 2. To develop comprehensive cross modalities to generate very high-resolution land cover map from mismatch between high resolution images features and pseudo labels. to provide detailed land cover that captures fine-grained features from complex urban features.

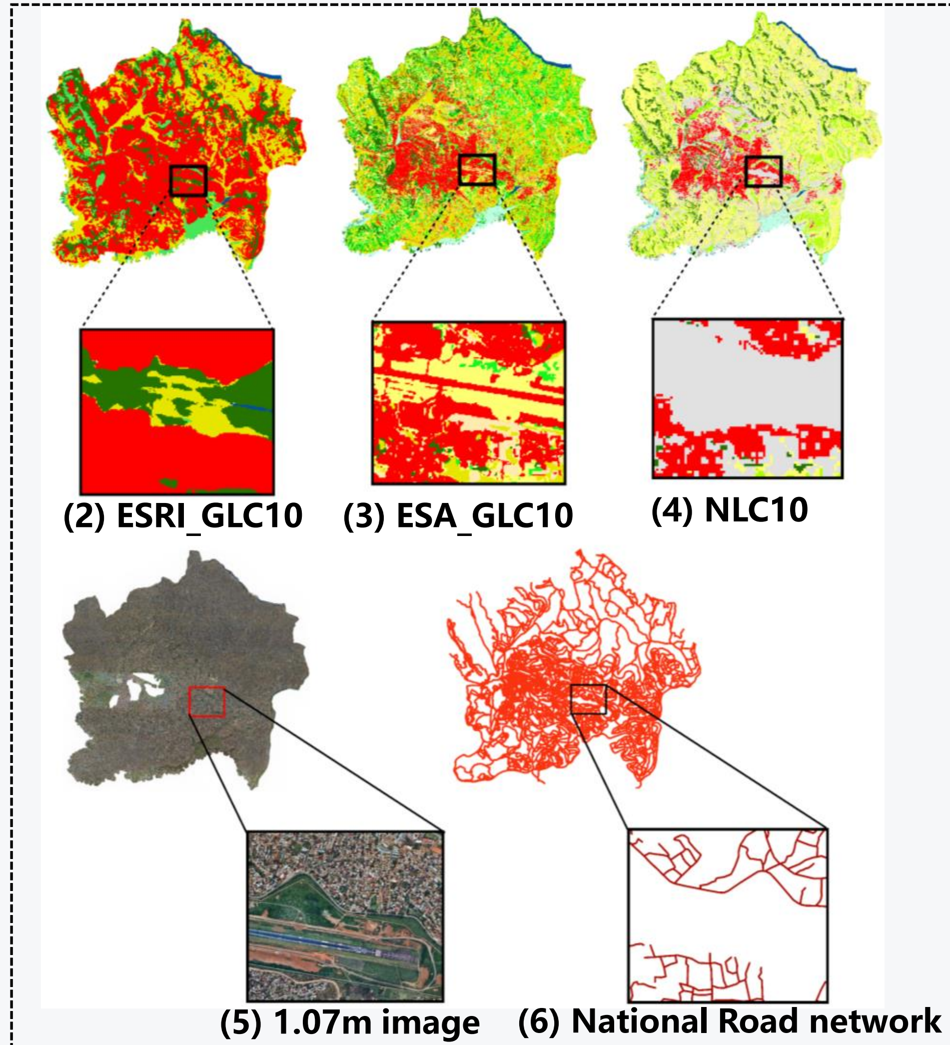
03

Datasets & Framework

Training Data acquisition



(1) Training distributions



Multisource dataset == 1.3TB

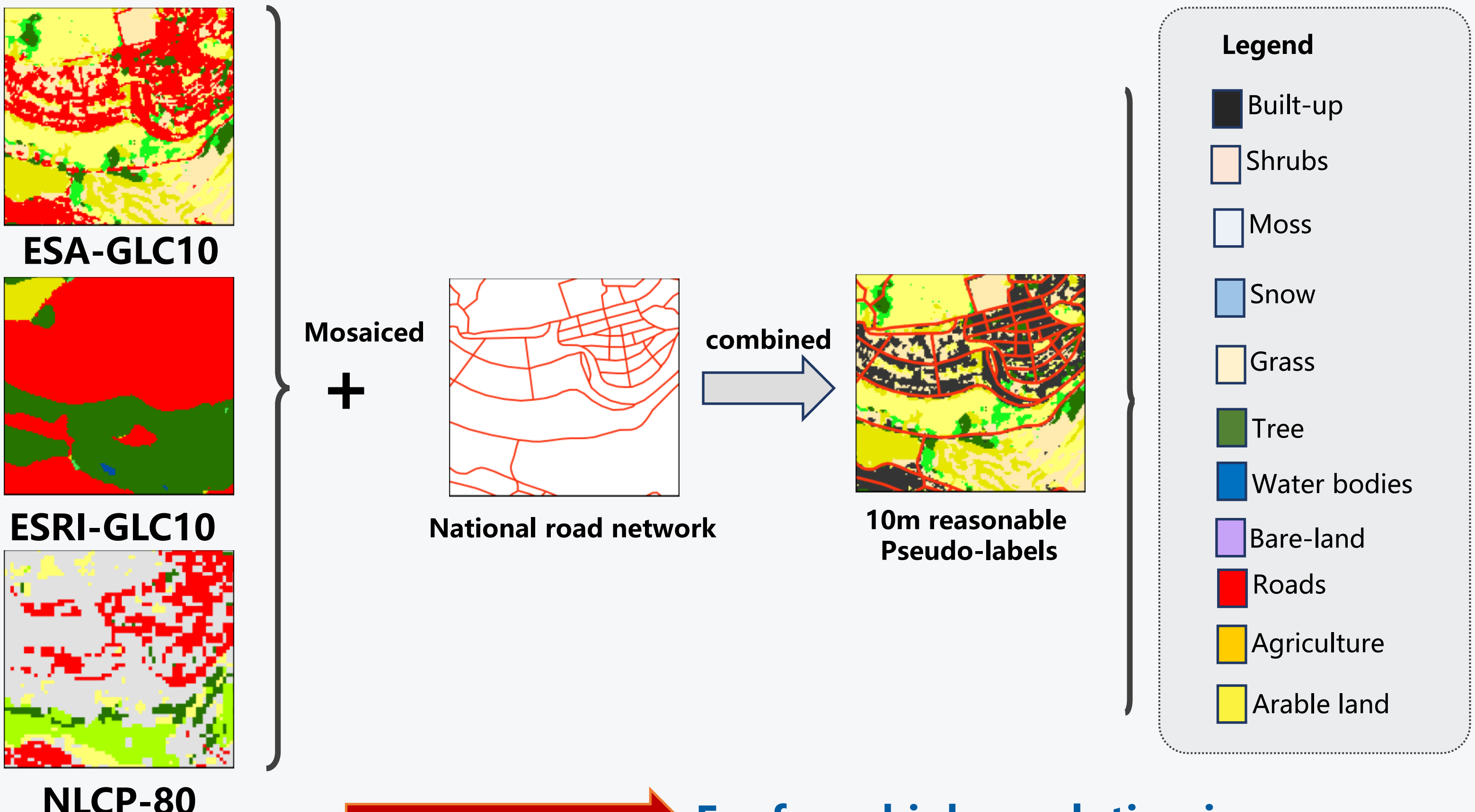
Training Data acquisition

Global and National land cover products classification scheme

	ESA GLC_P	ESRI GLC_P	NLC_P
Spatial coverage	Global	Global	National
Spatial resolution	10 m	10 m	8 m
Affiliation	European Space Agency	ESRI	National Land Agency
Land cover type Color lamp	Trees	Trees	Trees
	Scrub	Shrub	-
	Grass	Grass	Grass
	Crops	Crops	-
	Built area	Built-up	Built area
	Bare	Barren	Bare
	Snow and nice	Snow and nice	-
	Water	Water	Water
	Flooded vegetation		-
	Herbaceous Wetland	Flooded vegetation	Wetland
	Mangrove		-
	Moss and lichen		-
	ESA: Environment Science Institute, GLC_P: Global Land Cover Products, NLC_P: National Land Cover Product, Rwanda		

Training Data acquisition

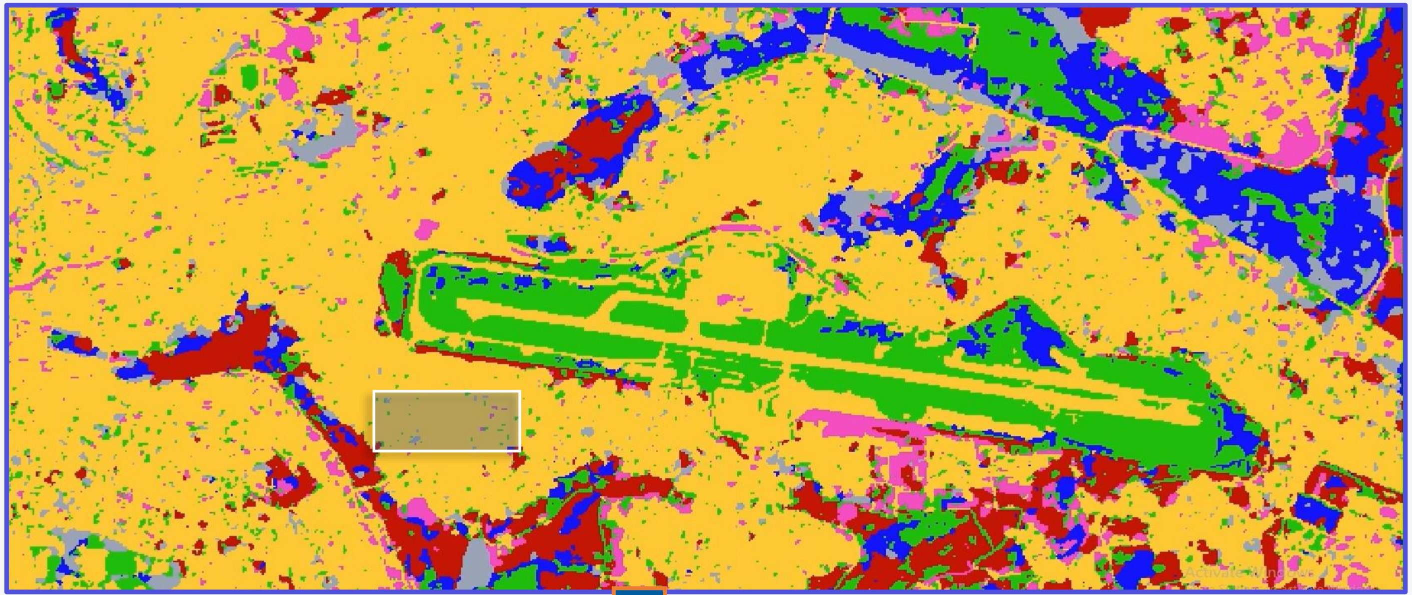
Overlays of National Roads network with GLCP



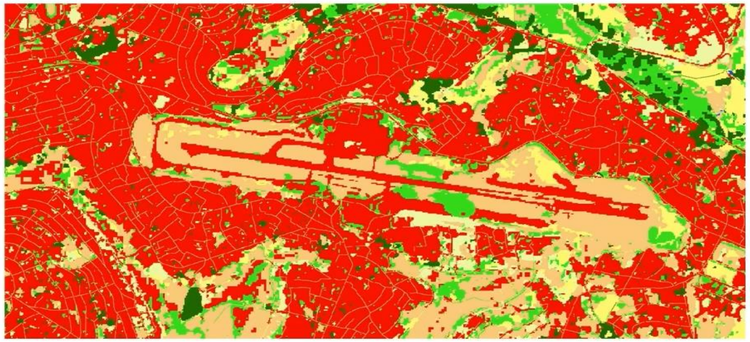
Far from high resolution image representations in high dimensions domains

Challenge: Spatial resolution mismatch

□ Before Superpixel

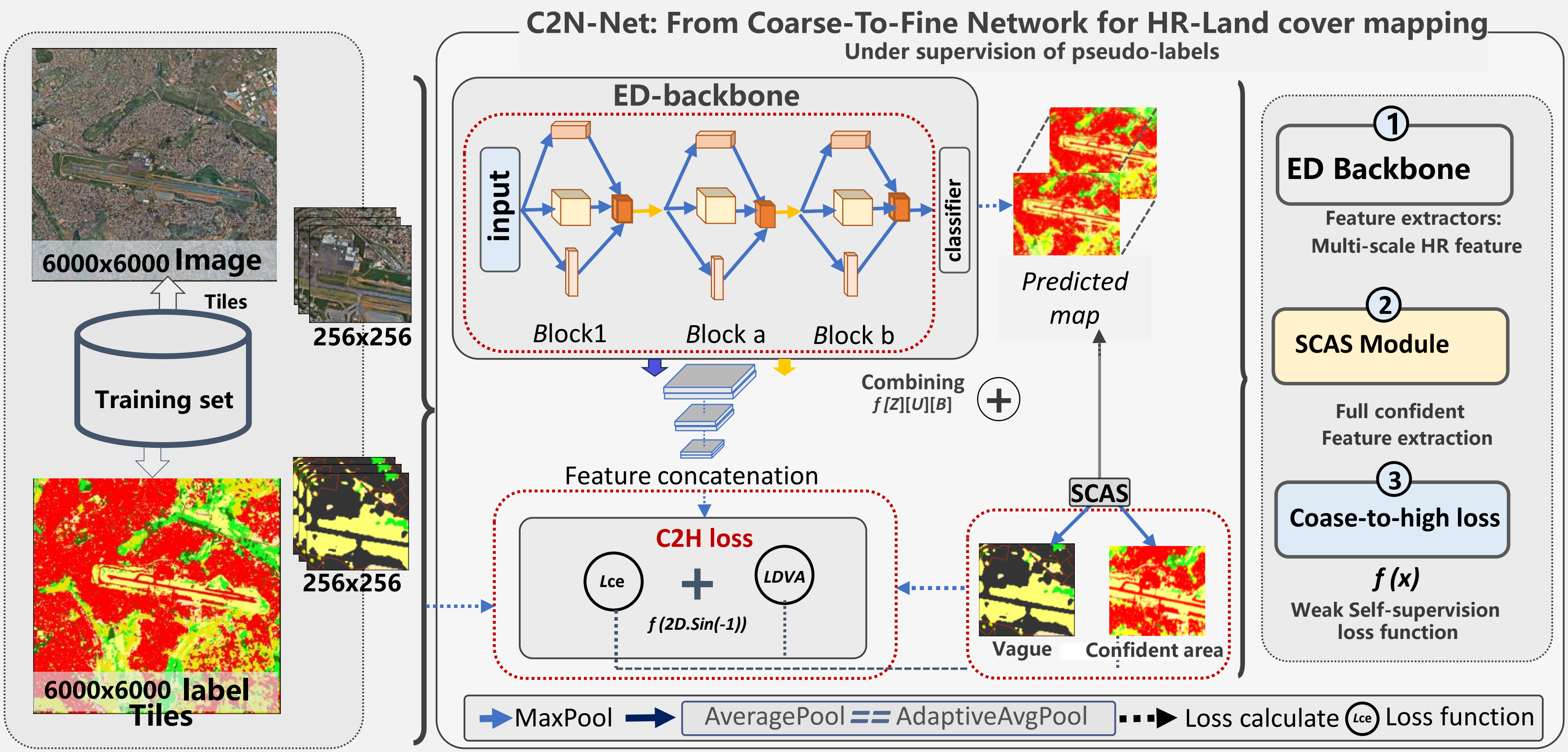


□ After Superpixel



From Coarse to Fine Network: C2Net

C2F-Net: composed of 3 major components:



Confident Hight Resolution features extracted from pseudo(noise)-labels under weak supervision

Experiment set up

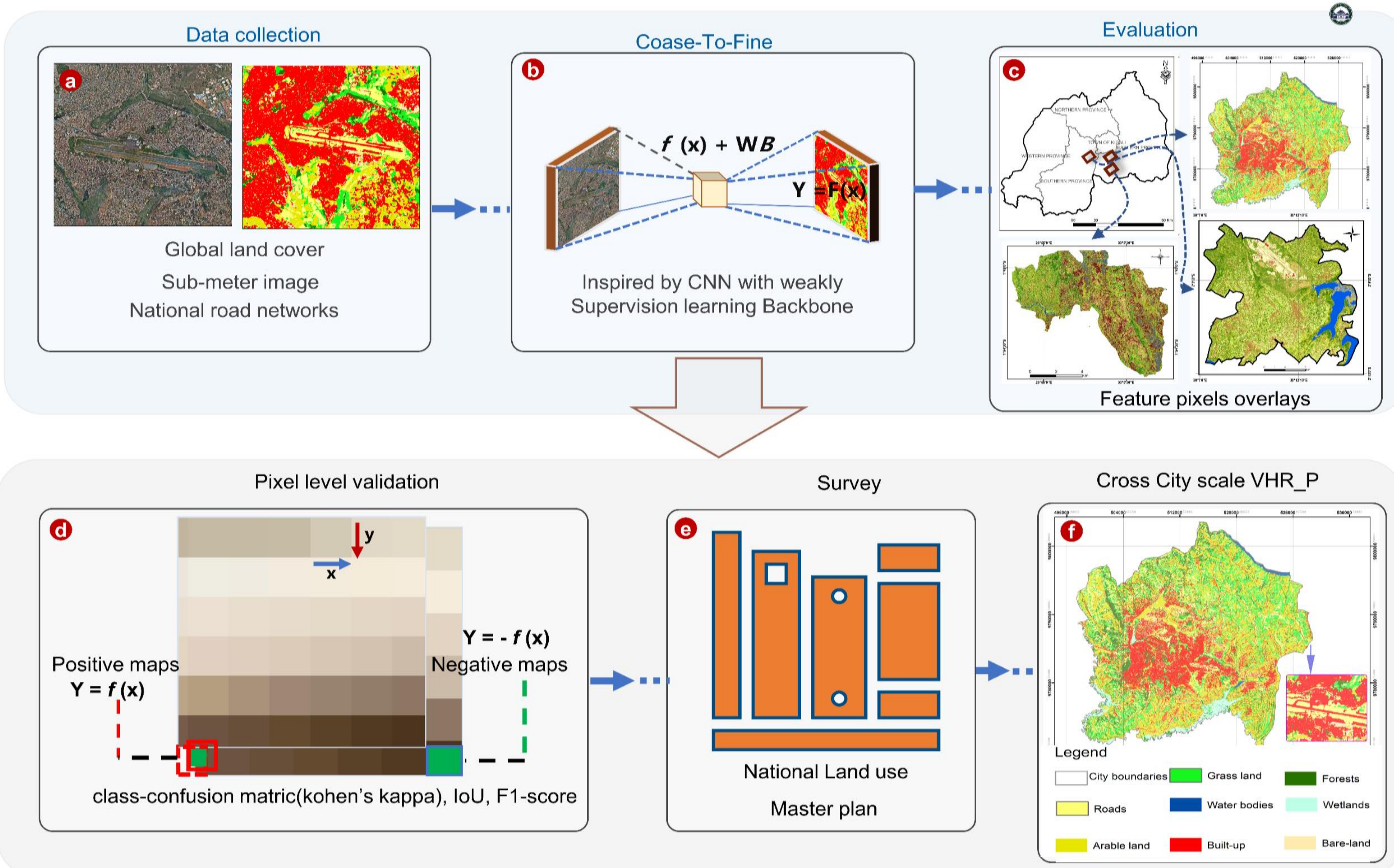
NVIDIA GeoForce RTX 3070 TI 16GB

PYTORCH

Deep Learning with PyTorch

- **Tile :** 6000
- **chips :** 256x256
- **Total training set :** 36,300
- **Batch size:** 64
- **Iterations :** 567.1875
- **Epochs :** 50
- **Optimizer:** Adam
- **Learning rate:** 0.001
- **Drop-out:** drop-out
- **Verbose:** 1
- **Activation *f*:** SoftMax

C2F training evaluation and validation



04

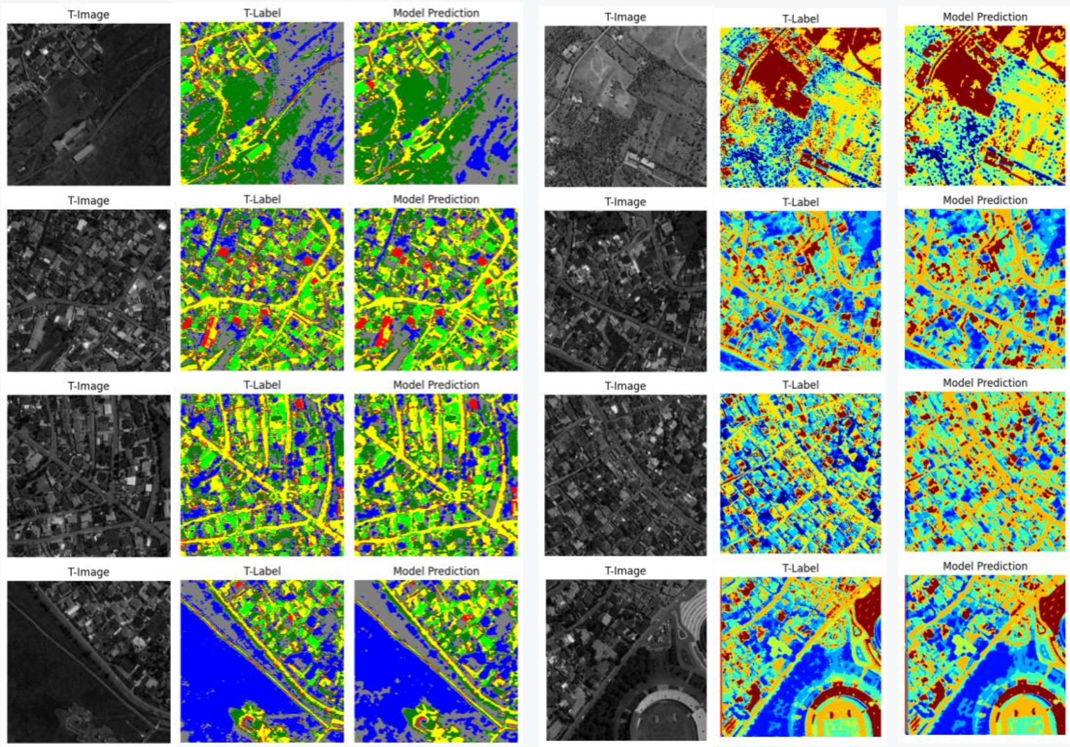
Results

01

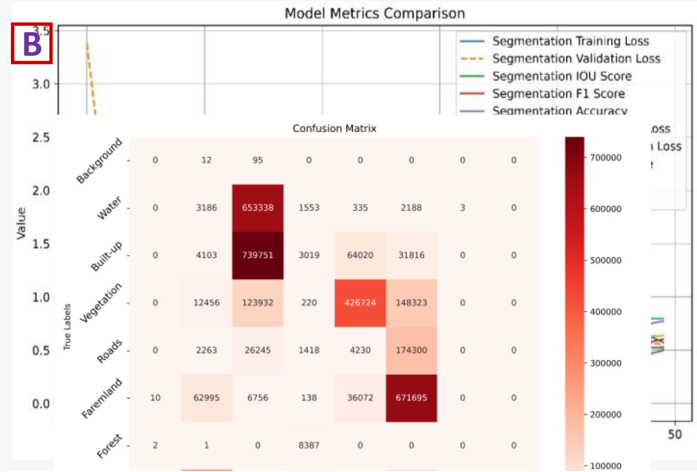
Proposed architecture behavior

RESULTS- C2FNet behavior-Kigali

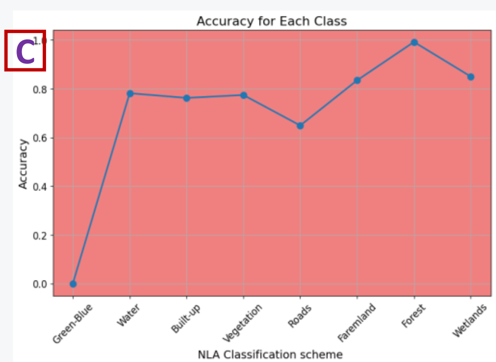
A



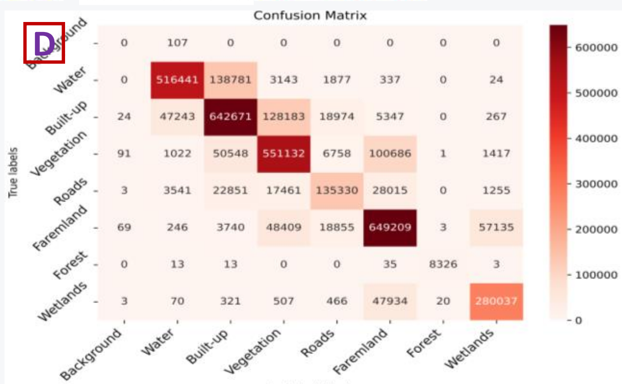
B



C



D



E

No	Classes	Per-class accuracy (%)	Recall	Precision	F1-score	support
0	Background	0	0.00	0.00	0.00	107
1	Water	78	0.91	0.78	0.84	660603
2	Built-up	76	0.75	0.76	0.76	842709
3	Vegetation	77	0.74	0.77	0.75	711655
4	Roads	64	0.74	0.65	0.69	208456
5	Fam-land	83	0.78	0.83	0.81	777666
6	Forest	99	1.00	0.99	0.99	8390
7	wetlands	85	0.82	0.85	0.84	329358
8	Macro avg average		71	72	71	3538944
9	Weighted average			79	79	3538944
10	MIoU					
11	OA (%)		80			
12	KAPPA		78			
13	AA (%)		81.2			

A Ablation results | B Training metrics | C Per class accuracy | D Confusion matrix | E C2F_Prediction

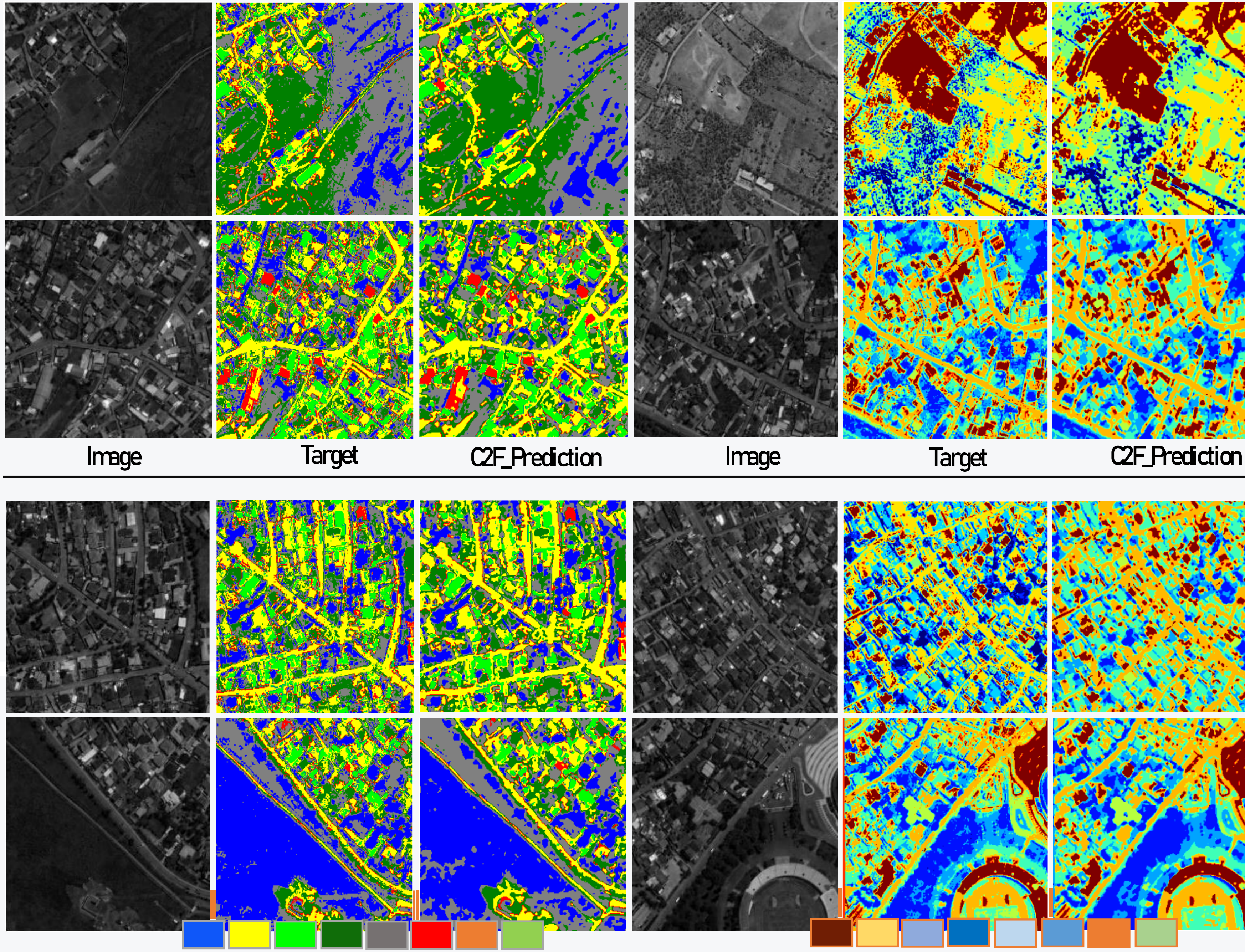
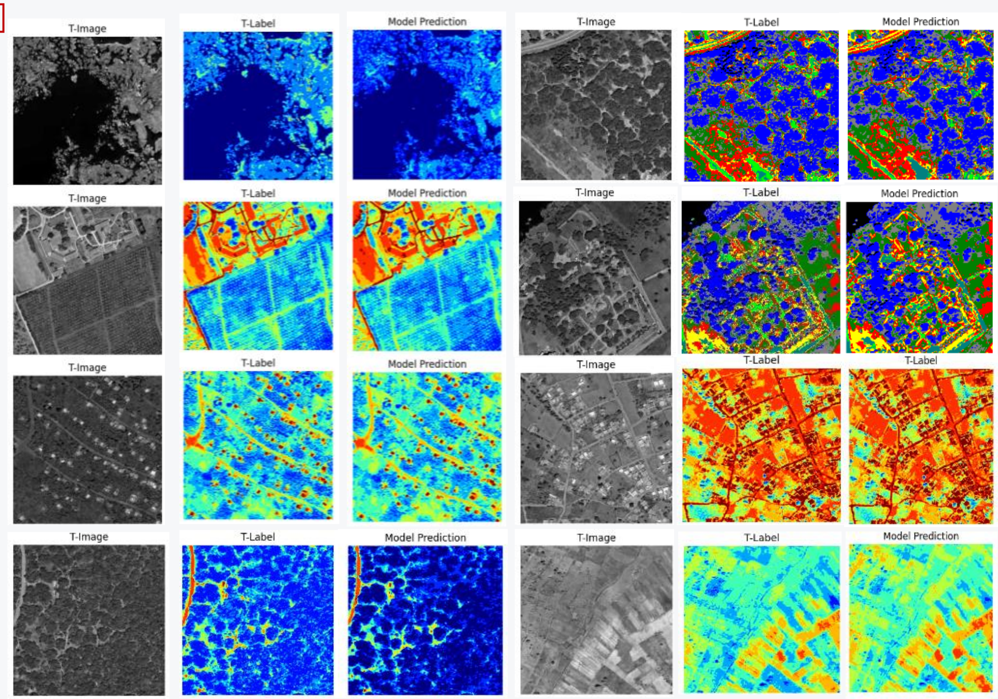


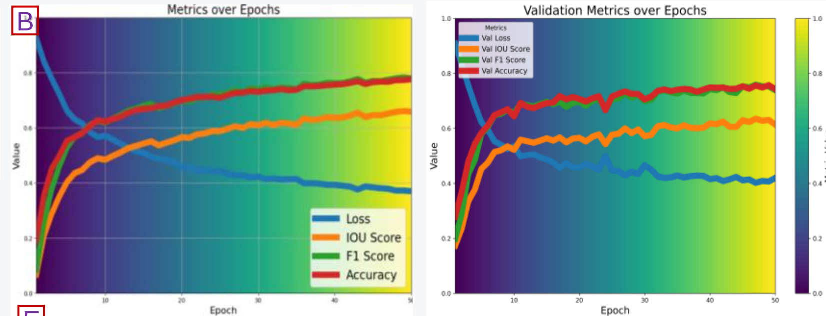
Image
Target
C2F_Prediction
Image
Target
C2F_Prediction

RESULTS-C2FNet behavior-Bugesera

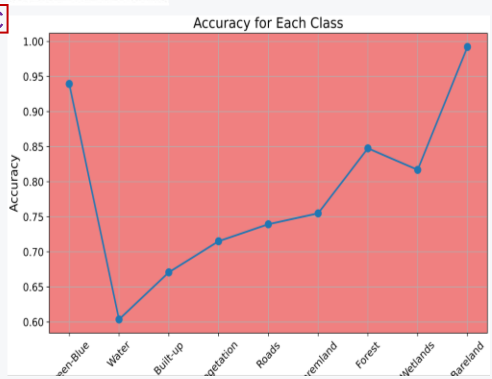
A



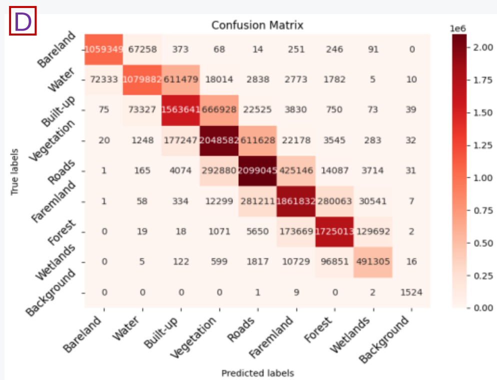
B



C



D



E

No	Classes	Per-class accuracy (%)	Recall	Precision	F1-score	Cohen's kappa	support
0	Background	0.00	0.00	0.00	0.00	0.00	107
1	Water	0.60	0.60	0.88	0.72	0.62	1789116
2	Built-up	0.67	0.67	0.66	0.67	0.60	2331188
3	Vegetation	0.71	0.72	0.67	0.69	0.59	2864763
4	Roads	0.73	0.74	0.69	0.72	0.67	2839143
5	Fam-land	0.75	0.75	0.74	0.75	0.72	2466346
6	Forest	0.84	0.85	0.81	0.83	0.63	2035134
7	Wetlands	0.81	0.82	75	0.78	0.75	601444
8	Bare lands	0.99	0.99	0.92	0.95	0.74	1536
1	Macro avg	-	0.79	0.79	0.78	-	16056320
2	Weighted average	-	0.75	0.74	0.74 -	-	16056320 -
3	MIoU	0.67	-	-	-	-	-
4	OA (%)	0.78	0.80	-	-	-	-
5	KAPPA	0.76	0.78	-	-	-	-
6	AA (%)	0.82	0.812	-	-	-	-
7	Training time	13hrs	-	-	-	-	-
8	Testing time	14seconds		-	-		-

A Ablation results

B Training metrics

C Per class accuracy

D Confusion matrix

E C2F_Prediction

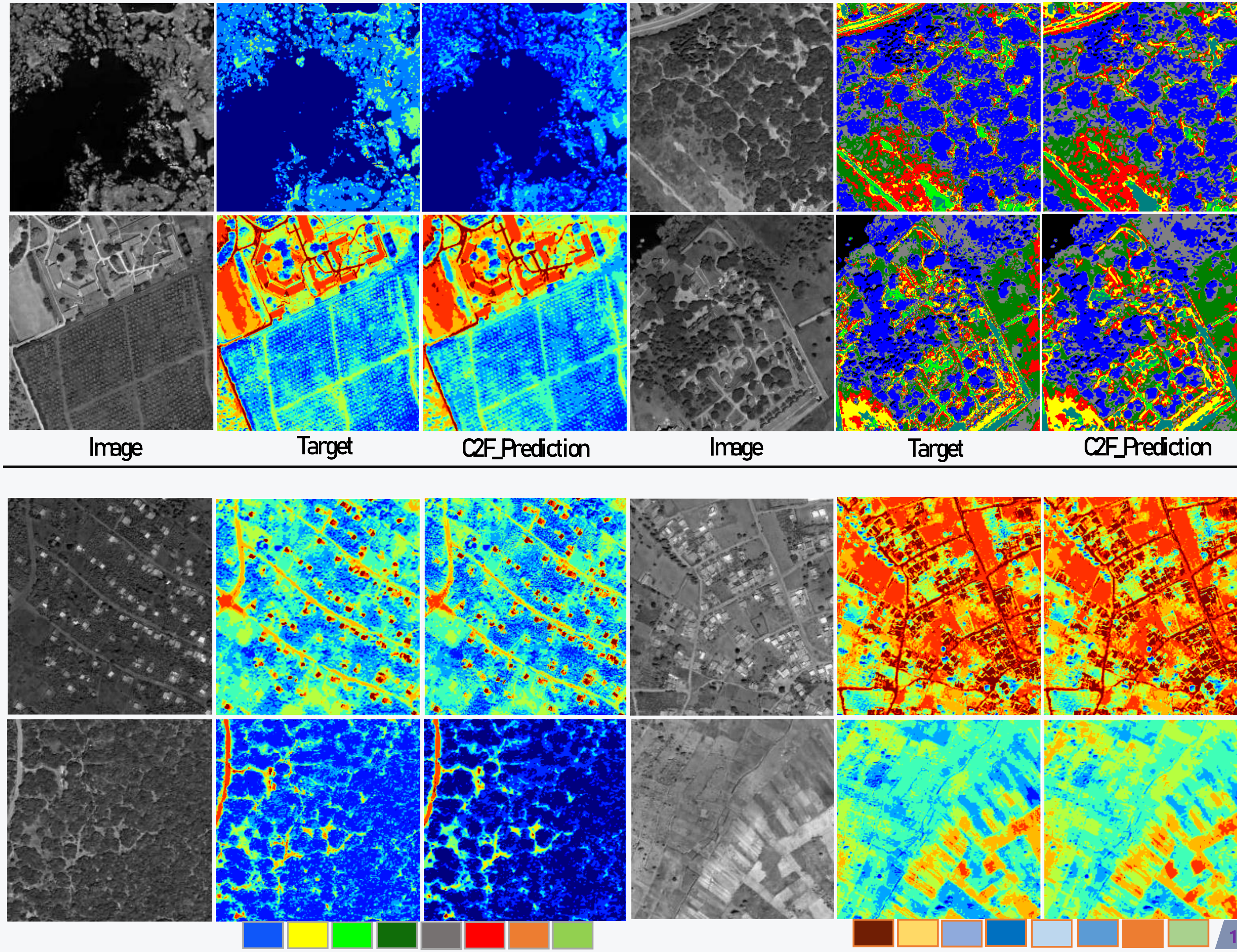
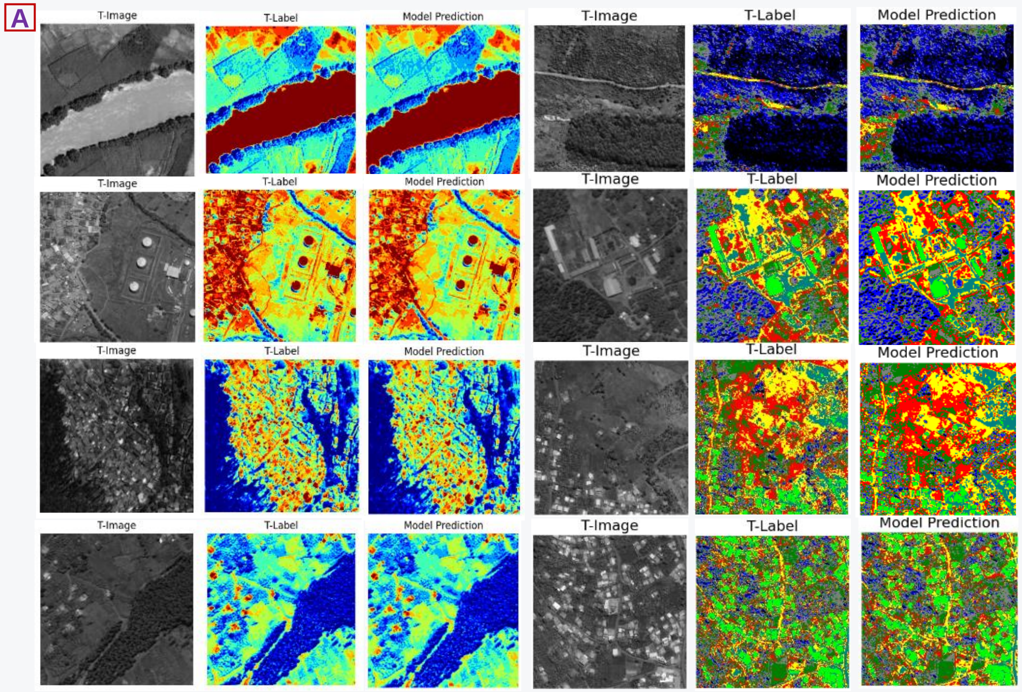
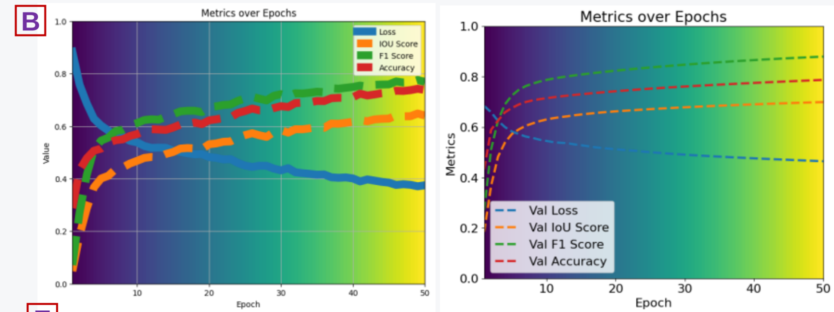
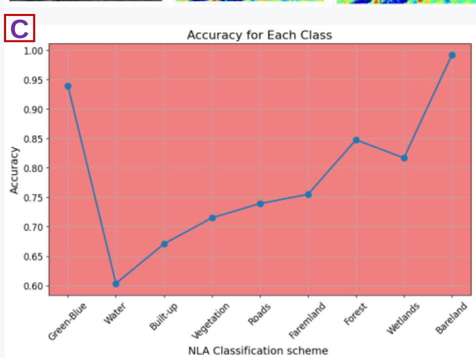


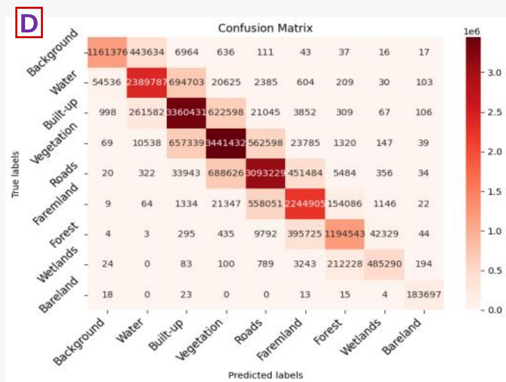
Image
Target
C2F_Prediction
Image
Target
C2F_Prediction

RESULTS-C2FNet behavior-Kamonyi





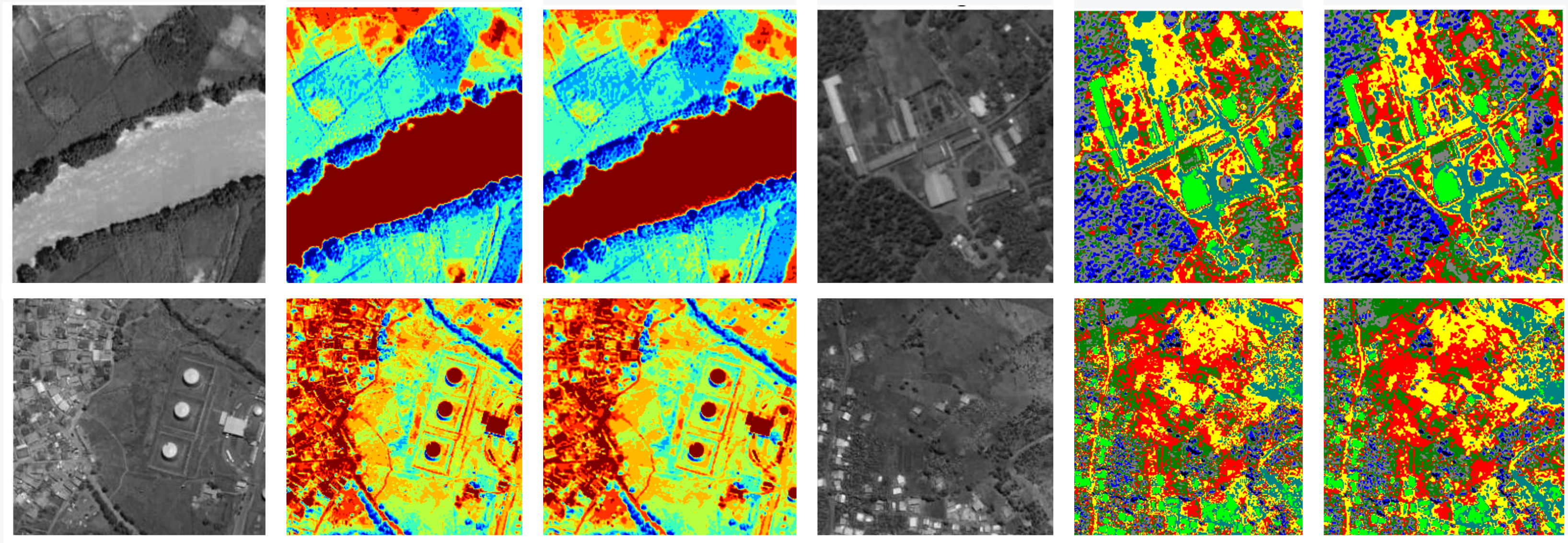




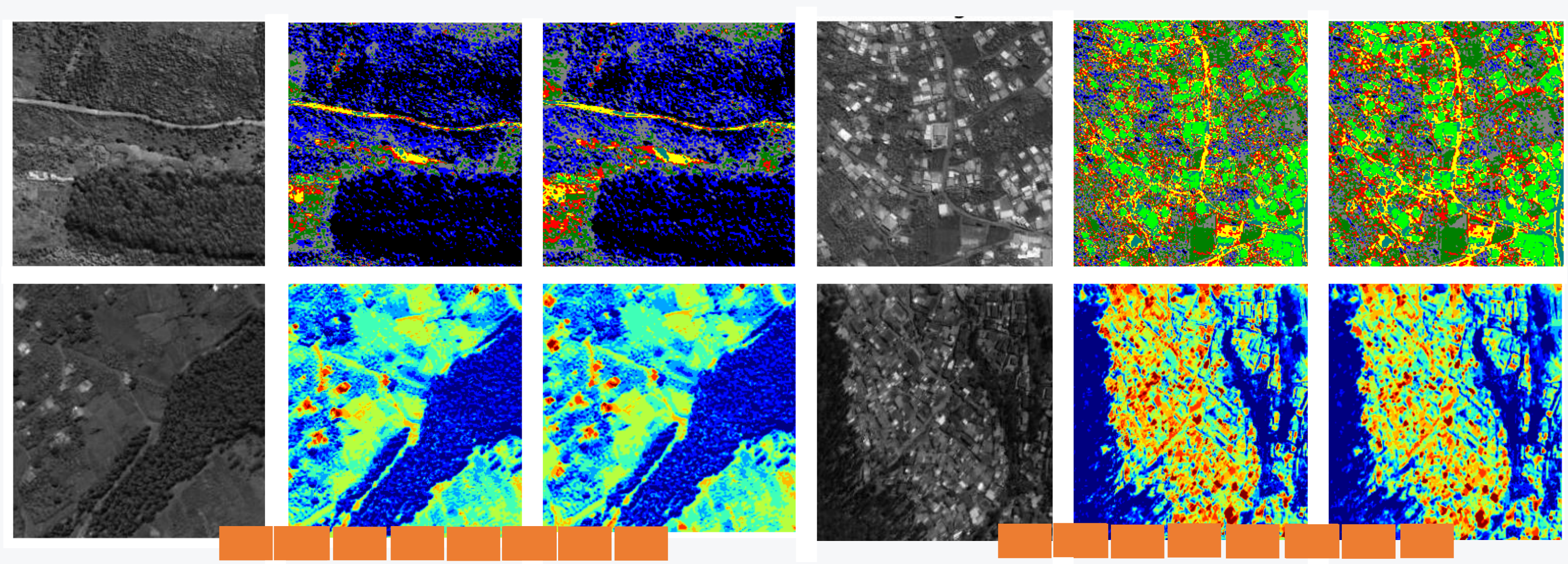
E

No	Classes	Per-class accuracy (%)	Recall	Precision	F1-score	support
0	Background	0	0.00	0.00	0.00	183770
1	Water	0.75	76	0.77	0.76	660603
2	Built-up	0.78	0.79	0.71	0.74	3162982
3	Vegetation	0.73	0.73	0.72	0.73	4270988
4	Roads	0.72	0.72	0.73	0.73	4697267
5	Fam-land	0.75	0.75	0.72	0.74	4273498
6	Forest	0.72	0.73	0.76	0.74	2980964
7	Wetlands	0.69	0.69	0.92	0.79	701951
8	Bare land	0.72	0.72	0.95	0.82	1612834
1	Macro avg	-	0.77	0.81	0.78	23527424
2	Weighted average	-	0.75	0.75	0.75 0.75	23527424 -
3	MIoU	0.65	-	-	-	-
4	OA (%)	0.75	0.774	-	-	-
5	KAPPA	-	0.78	-	-	-
6	AA (%)	0.74	0.742	-	0.75	-
7	Training time	16hrs	-	-	-	-
8	Testing time	732ms	-	-	-	-

A Ablation results
B Training metrics
C Per class accuracy
D Confusion matrix
E C2F_Prediction



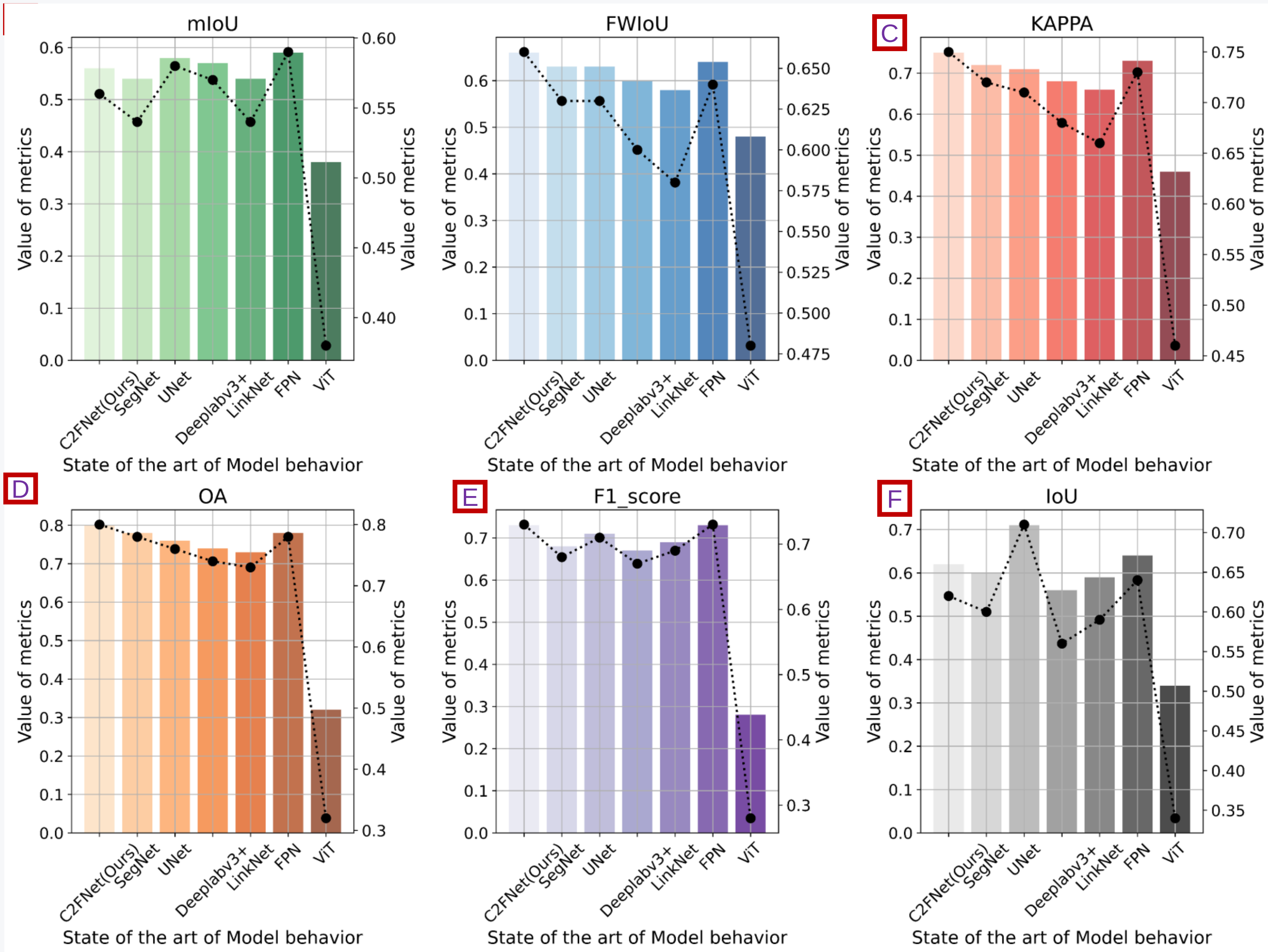
(a) Image ; (b) Target ; (c) C2F_Prediction ; (d) Image ; (e) Target ; C2F_Prediction



04 Results

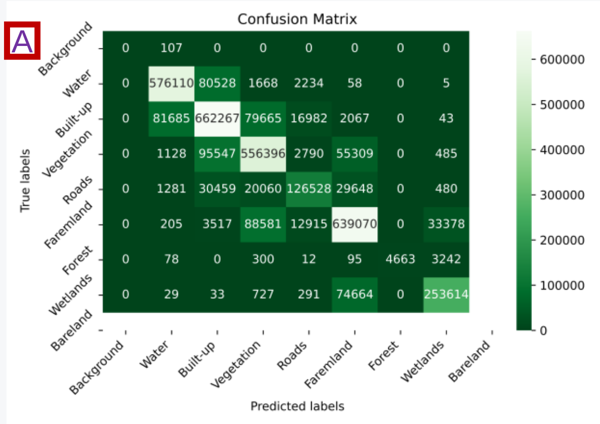
02 State of the Art-architecture comparison

RESULTS-C2FNet behavior over metrics

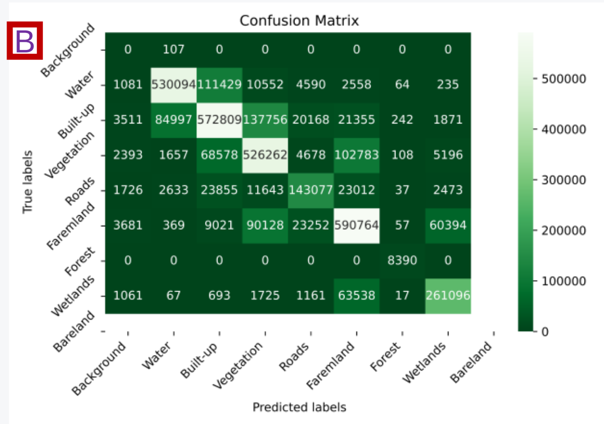


RESULTS-C2FNet behavior

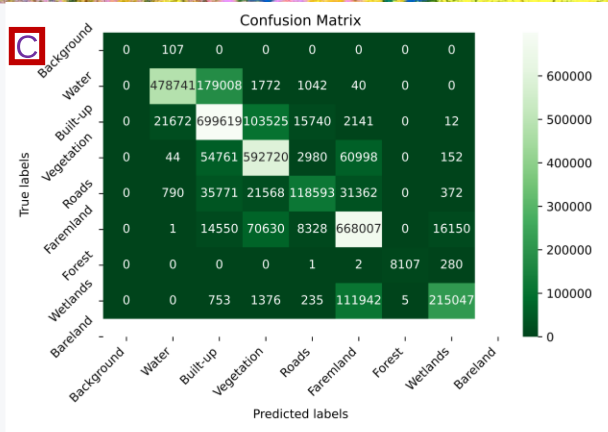
A



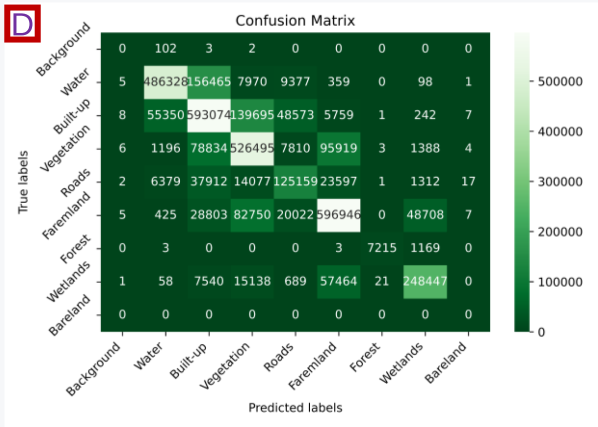
B



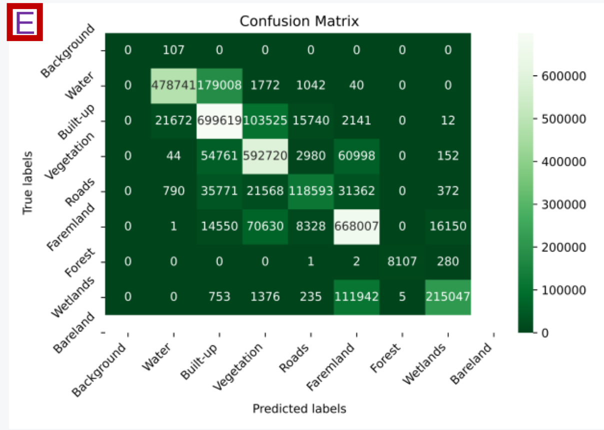
C



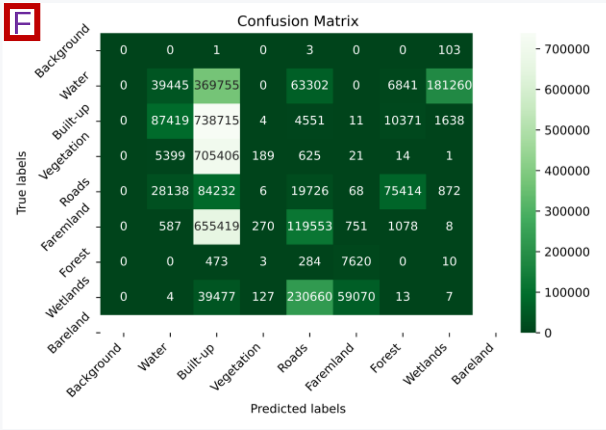
D



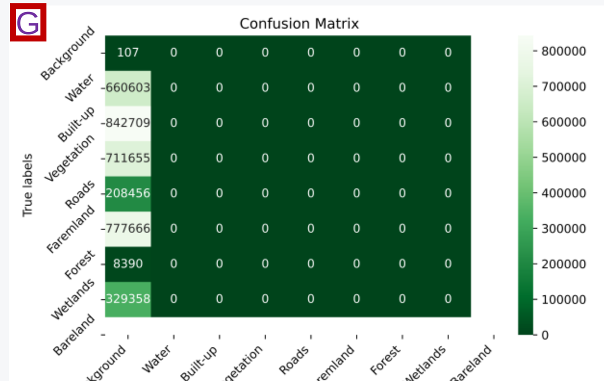
E



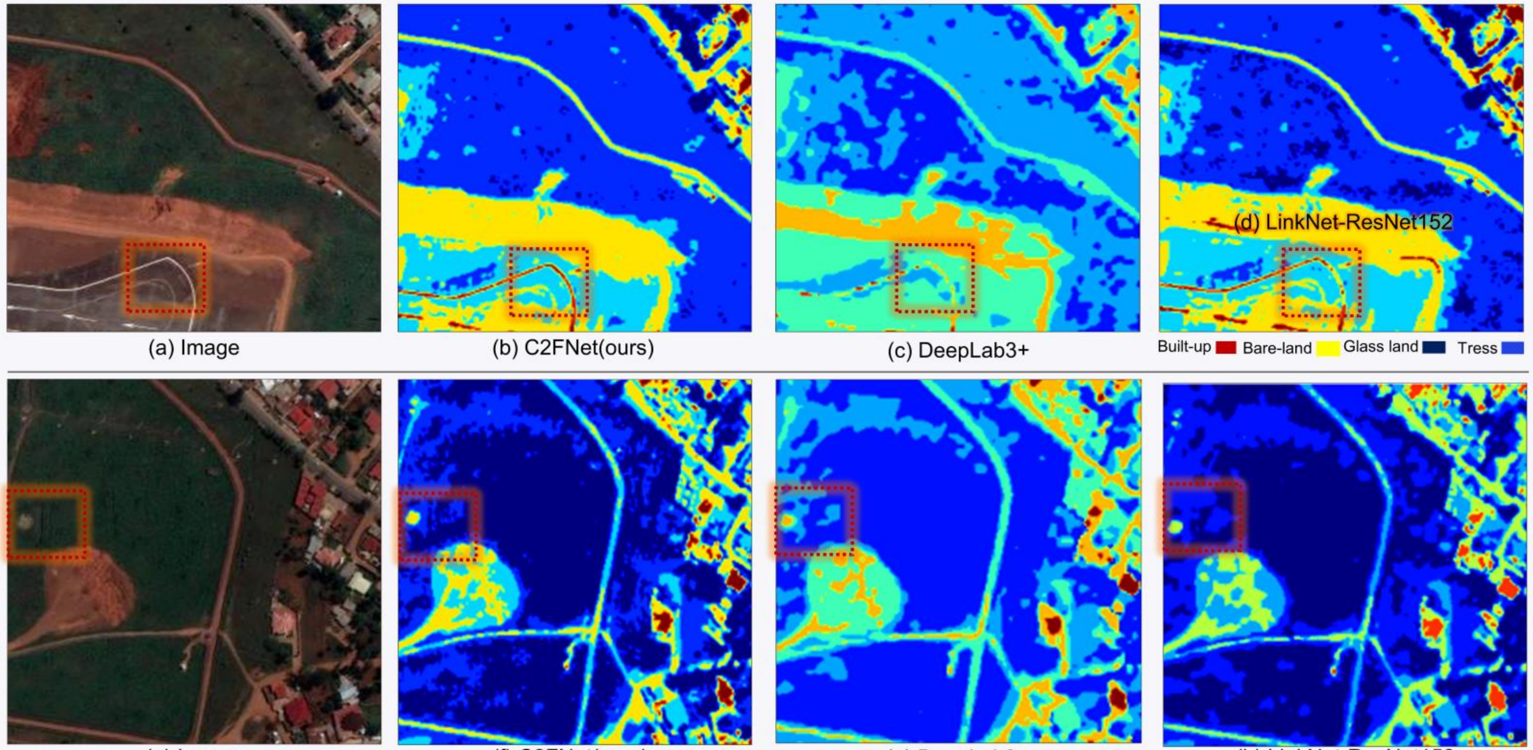
F



G



RESULTS- Visual comparative-1



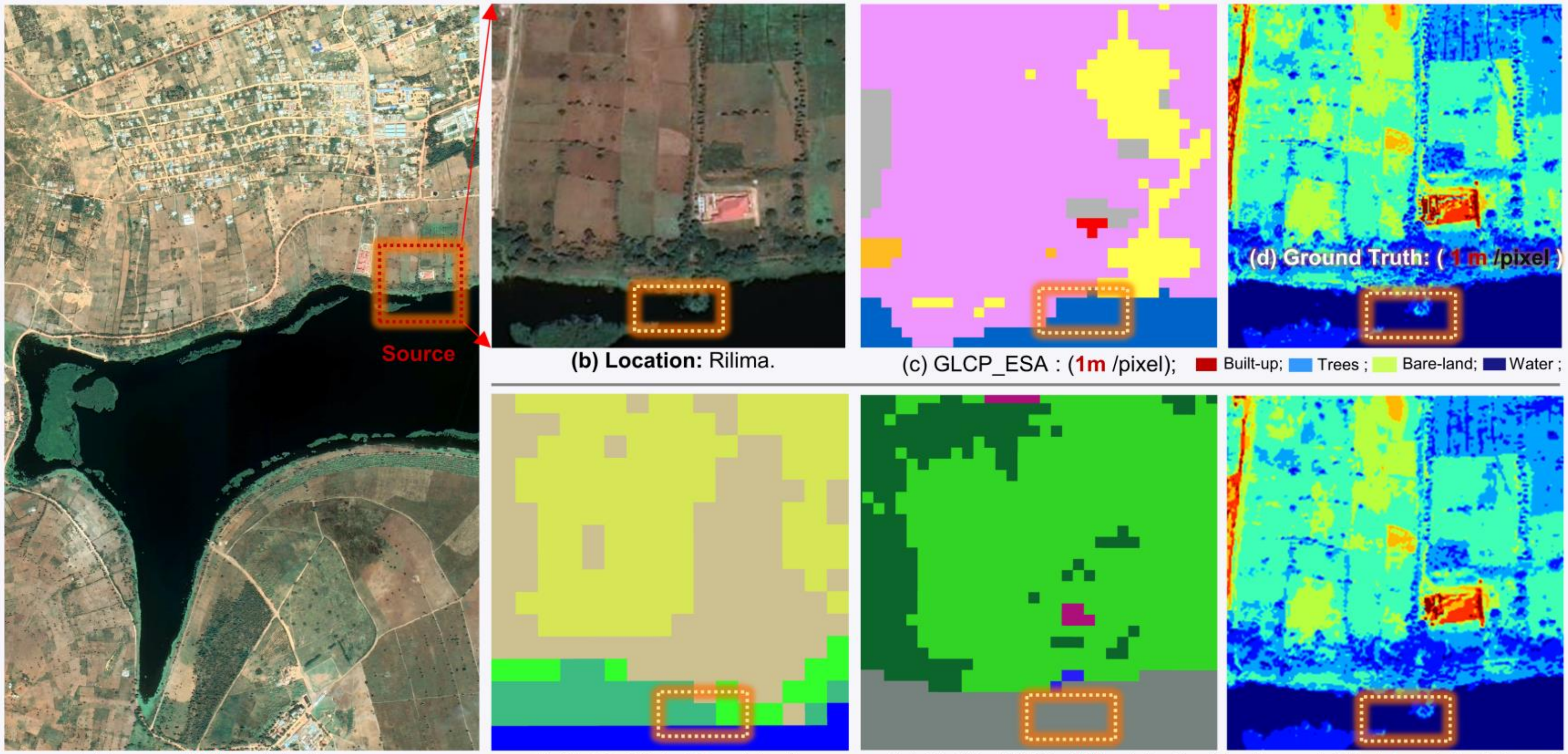
(a) Image (b) C2FNet(ours) (c) DeepLab3+

(e) Image (f) C2FNet(ours) (g) DeepLab3+ (h) LinkNet-ResNet152

No	Models	Backbones	Ep	TP(M)	TM	AA	L	IoU	F1-s	OA
1	**C2FNet (Ours)**	**Vanilla**	**50**	**24.45**	**7**	**0.821**	**0.36**	**0.62**	**0.72**	**0.79**
2	SegNet	ResNet50	-	28.64	18’56’’	0.78	0.41	0.60	0.69	0.76
3	ViT	PE, PE, TE	-	24.45-3	2’17’’	0.87	0.88	0.03	0.033	0.87
4	DeepLabV3+	ResNet50	-	42.004	17’58’’	0.75	0.42	0.56	0.67	0.72

Ep= Epochs, TM= model’s “Training time”, AA=User’s accuracy, L=Model training “loss”, PAE= Transformer’s “Patch embedding”, POE= Transformer’s “Positioning encodings”, TE= Transformer’s encoders and (-) the same as above.

RESULTS- Visual comparative- 2

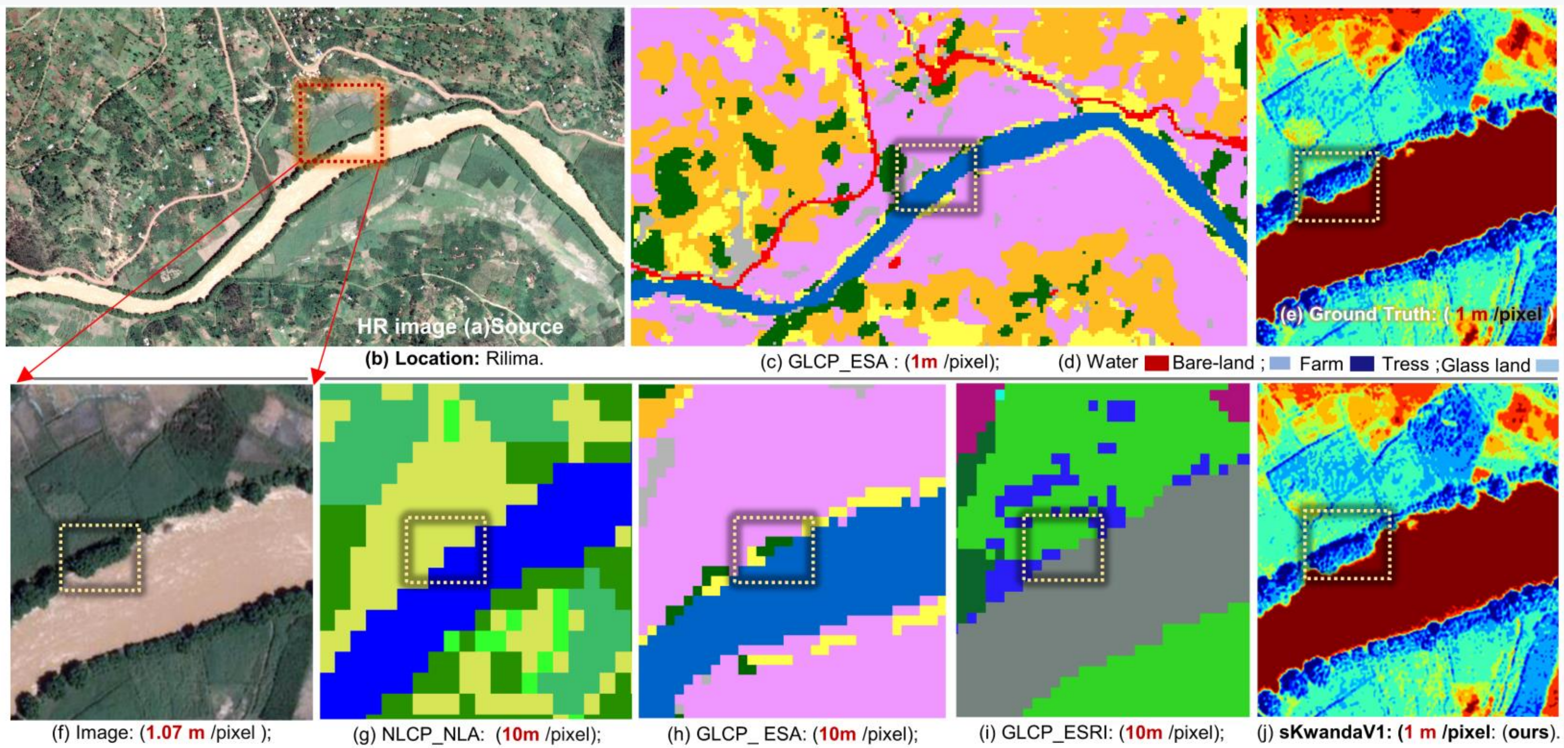


(e) Image: (1.07 m /pixel); (f) NLCP_NLA: (10m /pixel); (g) GLCP_ ESRI : (10m /pixel); (b) **sKwandaV1: (1 m /pixel**: (**ours**)

No	Models	Region	TP(M)	TM	L	VL	VIoU	VLA-F1-s	VA
1	C2FNet	Kigali	9.15	7'2''	0.387	0.453	0.582	0.707	0.78
		Bugesera	24.45	8'34''	0.369	0.420	0.614	0.739	0.743
		Kamonyi	29.05	4'14''	0.500	0.370	0.287	0.404	0.646

RESULTS- Visual comparative-3

Nº	Model	C2FNet	SegNet	UNet	Deeplabv3+	LinkNet	FPN	ViT
1	mIoU	0.56	0.54	0.58	0.57	0.54	0.59	0.38
2	FWIoU	0.66	0.63	0.63	0.6	0.58	0.64	0.48
3	KAPPA	0.75	0.72	0.71	0.68	0.66	0.73	0.46
4	OA	0.8	0.78	0.76	0.74	0.73	0.78	0.32
5	F1_score	0.73	0.68	0.71	0.67	0.69	0.73	0.28
6	IoU	0.62	0.6	0.71	0.56	0.59	0.64	0.34



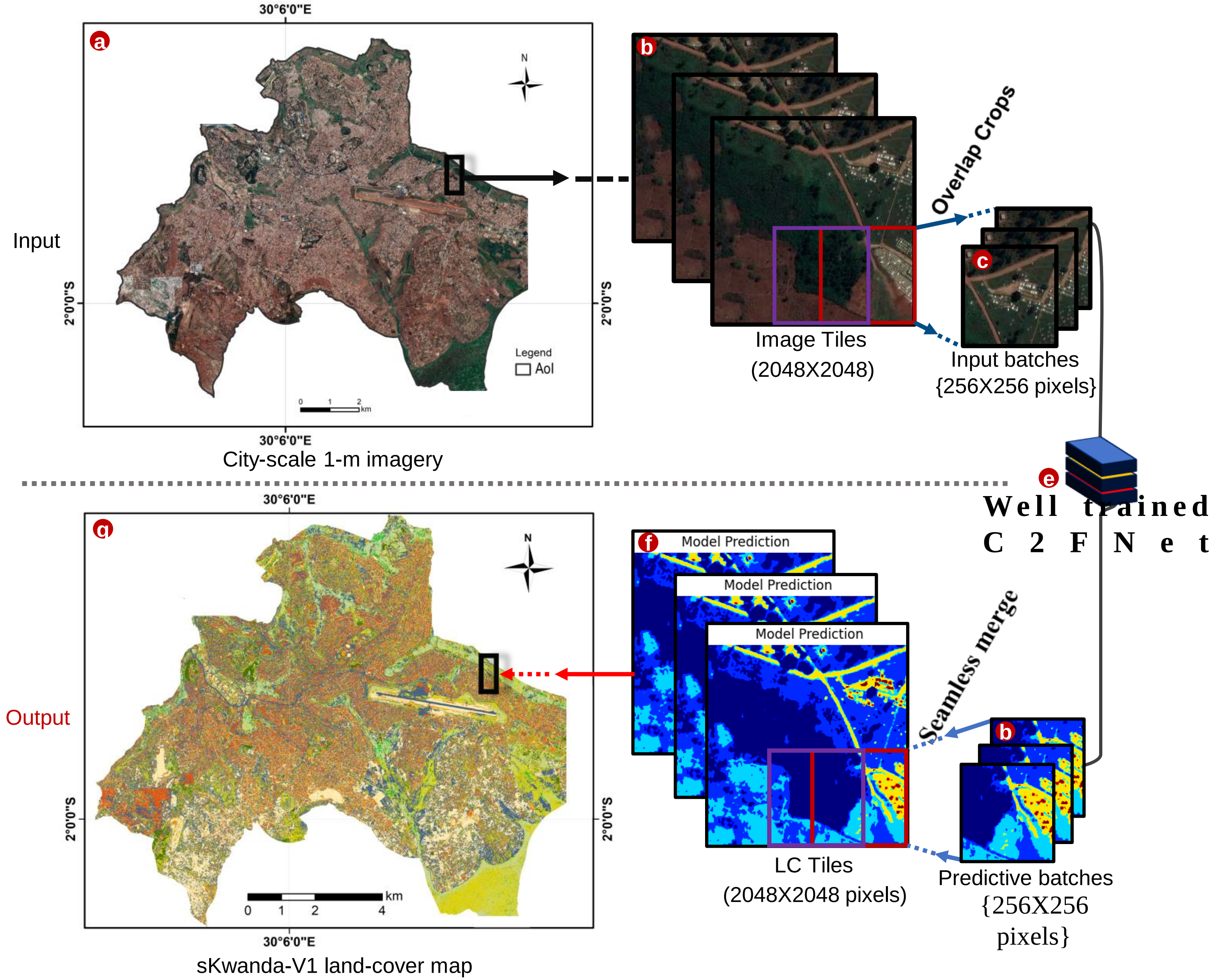
04

02

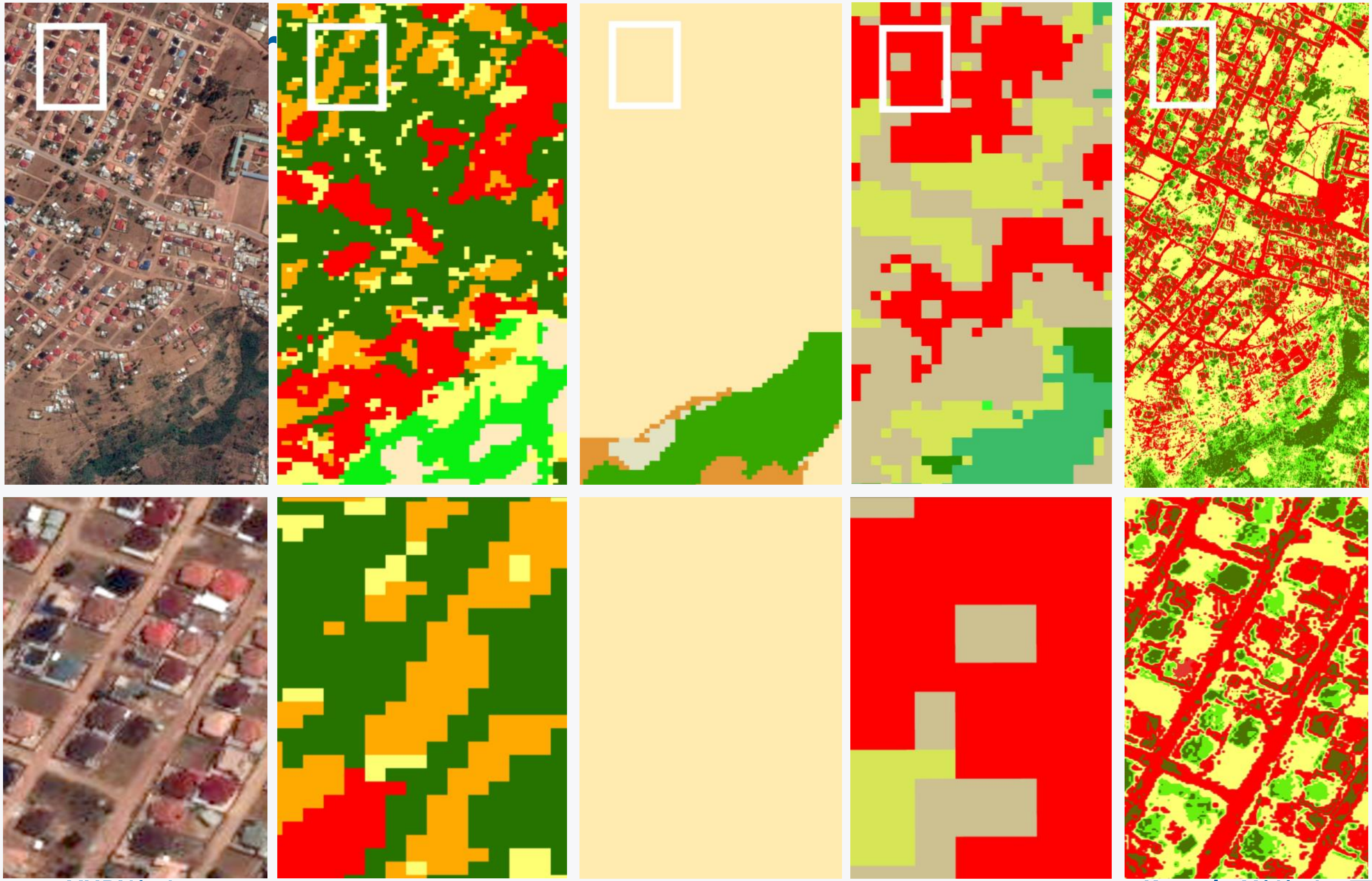
Results

Practical mapping

sKwanda-V1 Product (by C2F)

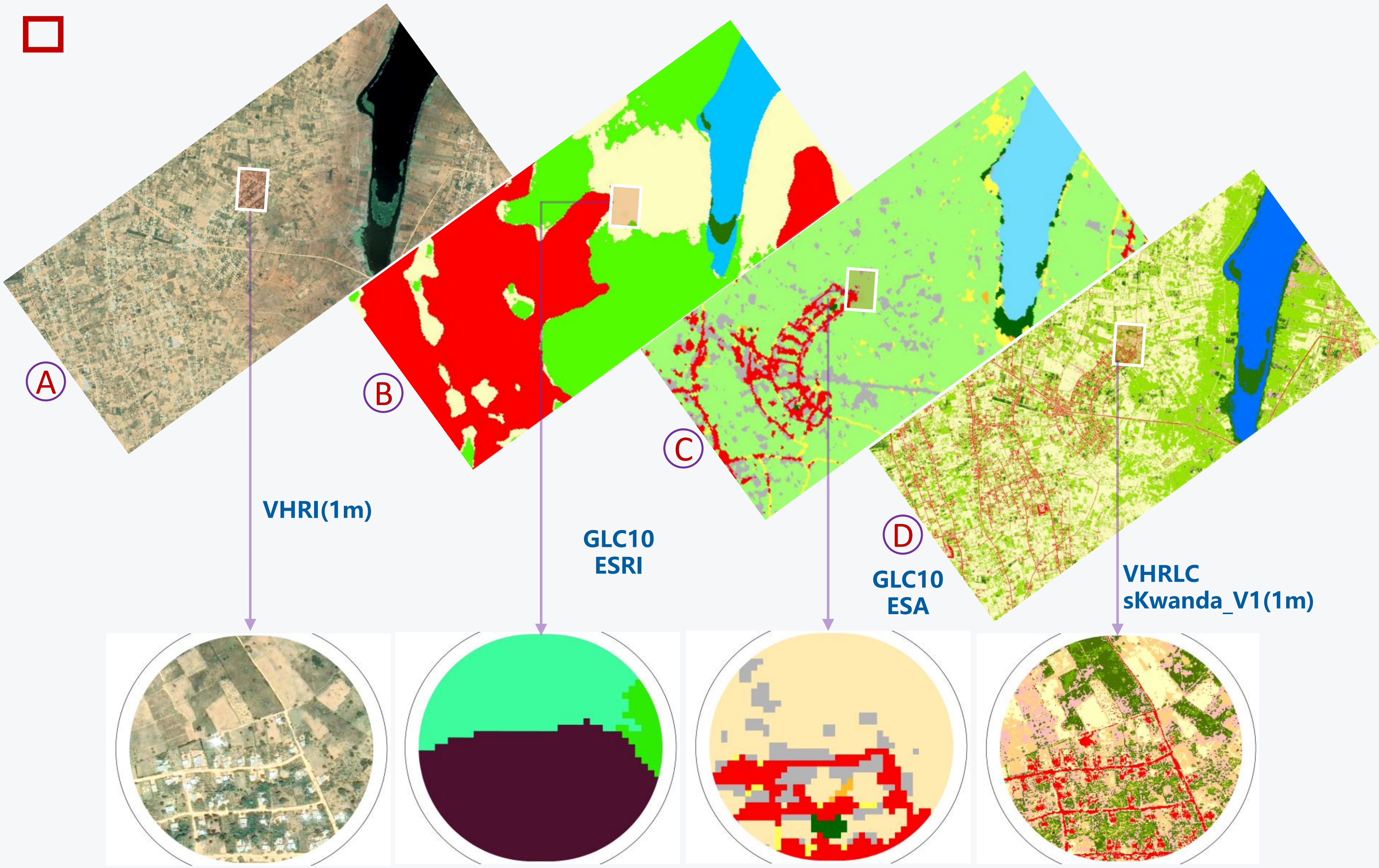


sKwanda-V1 Product Kamonyi dataset



VHRI(1m) | GLC10 ESRI | GLC10 ESA | NLC8 | sKwanda_V1(1m)

RESULTS-C2FNet behavior-Bugesera



Conclusion

- Leveraging a combination of freely accessible data, including 2 GLC10 products, National Land Authority records, Rwanda Transport Agency data, and Google Earth imagery, sKwanda_v1 stands out as a high-resolution dataset, covering both urban and rural areas.
- sKwanda_v1 sets a new standard in land-cover datasets, offering superior accuracy in delineating land-cover edges and depicting dense urban patterns compared to existing products. Quantitative assessments reveal an impressive overall accuracy of 80.01% with a kappa coefficient of 0.75, consistently demonstrated across diverse geographical regions and aligning well with official government survey data. It serves as a valuable resource for a multitude of applications requiring precise land-cover information.

Therefore, the contribution of this study can be summarized as: a) To improve the accuracy and robustness of land cover mapping, especially in areas with complex features.; b) To maximize the utilization of available data sources for advanced high-resolution land cover mapping. The results also contribute to predictions accurately reflect adding new classes such as low vegetation, trees and traffic patterns and cityscape layouts.

04 References

References

- [1] Li, Z., Zhang, H., Lu, F., Xue, R., Yang, G., & Zhang, L. (2022). Breaking the resolution barrier: A low-to-high network for large-scale high-resolution landcover mapping using low-resolution labels. ISPRS Journal of Photogrammetry and Remote Sensing, 192, 244–267. https://doi.org/10.1016/j.isprsjprs.2022.08.008

- [2] Aryal, K., Apan, A., & Maraseni, T. (2023). Comparing global and local land cover maps for ecosystem management in the Himalayas. Remote Sensing Applications: Society and Environment, 30. https://doi.org/10.1016/j.rsase.2023.100952
- [3] Bartholomé, E., & Belward, A. S. (2005). GLC2000: A new approach to global land cover mapping from earth observation data. International Journal of Remote Sensing, 26(9), 1959–1977. https://doi.org/10.1080/01431160412331291297
- [4] Li, Z., Zhang, H., Lu, F., Xue, R., Yang, G., & Zhang, L. (2022). Breaking the resolution barrier: A low-to-high network for large-scale high-resolution land-cover mapping using low-resolution labels. ISPRS Journal of Photogrammetry and Remote Sensing, 192, 244–267. https://doi.org/10.1016/j.isprsjprs.2022.08.008
- [5] Ian Goodfellow, Yoshua Bengio, and Aaron Courville: Deep learning The MIT Press, 2016, 800 pp, ISBN: 0262035618
- Winkler, K., Fuchs, R., Rounsevell, M., & Herold, M. (2021). Global land use changes are four times greater than previously estimated. Nature Communications, 12(1). https://doi.org/10.1038/s41467-021-22702-2