WORKSHOP RAINSMORE/SWOT

SEMANTIC SEGMENTATION OF SATELLITE IMAGES VIA DEEP NEURAL NETWORKS

Valdivino Alexandre de Santiago Júnior

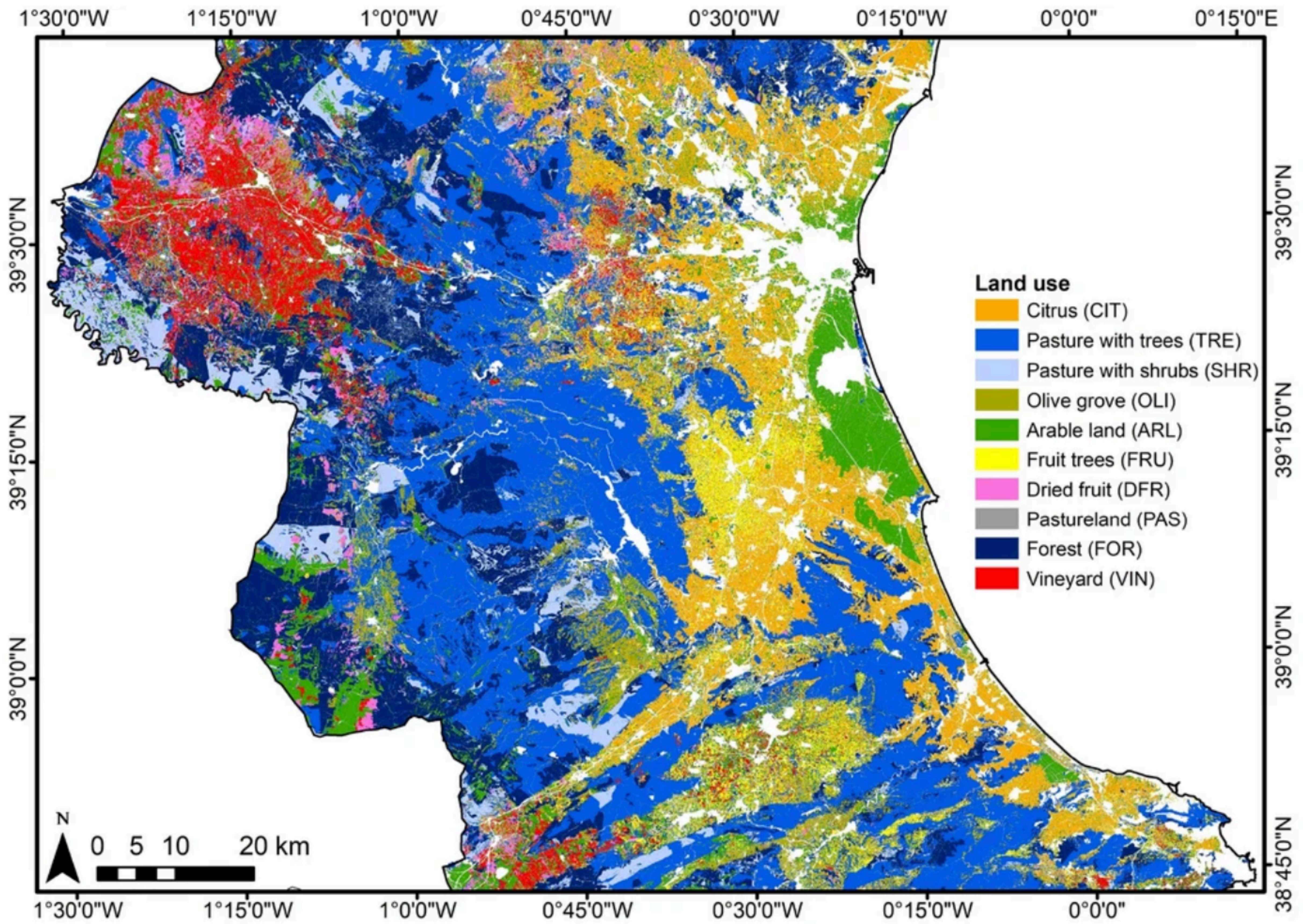
Coordenação de Pesquisa Aplicada e Desenvolvimento Tecnológico (COPDT)

Instituto Nacional de Pesquisas Espaciais (INPE)

São José dos Campos, SP, Brazil

SEMANTIC SEGMENTATION

- It is a **task**! Precisely, a dense prediction task!
- Pixel-based classification: every pixel matters!

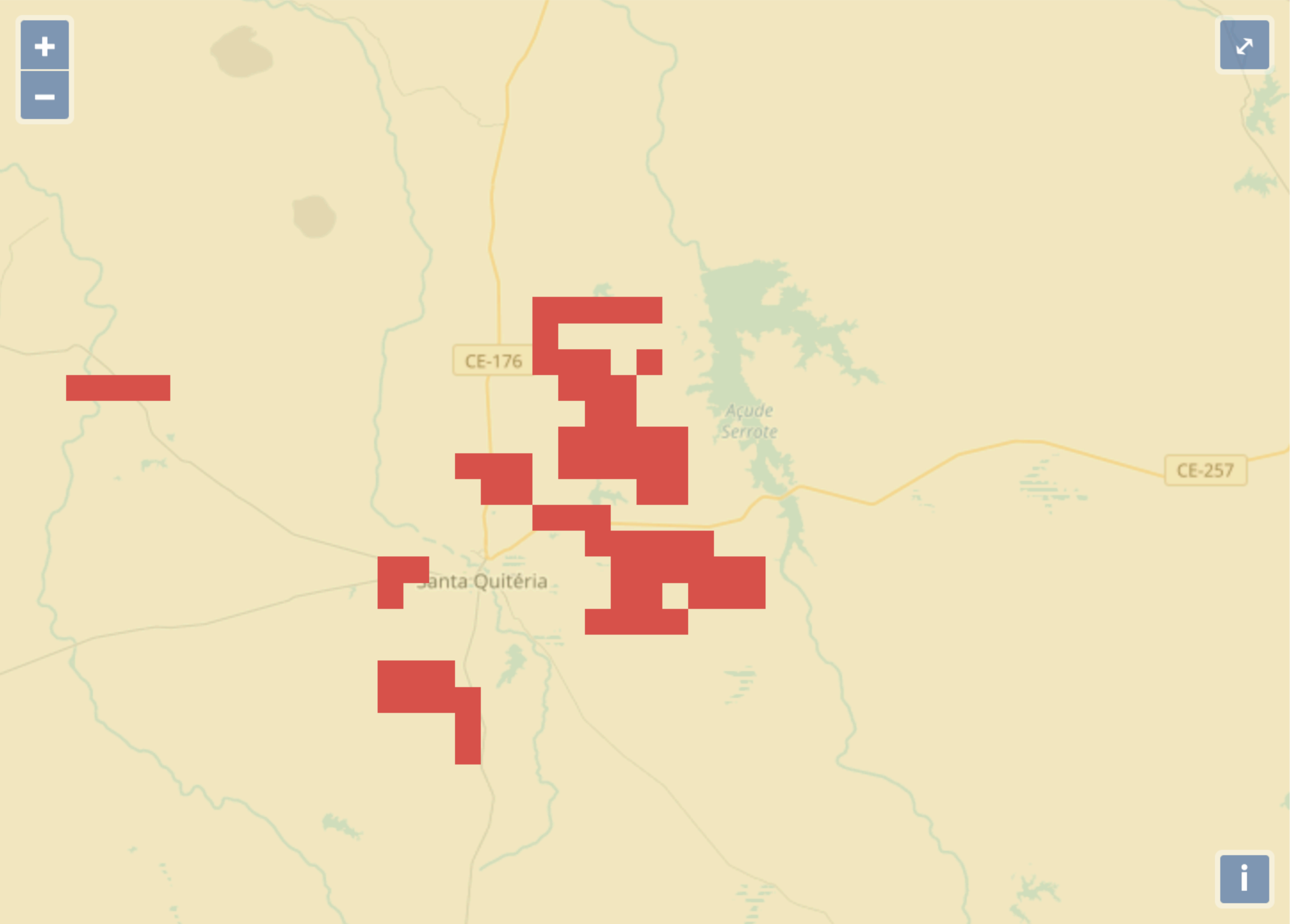


SEMANTIC SEGMENTATION

- Why is this important for remote sensing?

Mapa de área queimada

para o bioma Caatinga em Agosto de 2022



Área queimada (km²)

por bioma em Agosto de 2022

Ano	Mês	Amazônia	Caatinga	Cerrado	Mata Atlântica	Pampa	Pantanal	Total mensal
2022	1	2.273	47	277	452	453	81	3.583
2022	2	504	67	494	181	219	37	1.502
2022	3	64	28	508	156	36	56	848
2022	4	409	67	1.583	114	44	32	2.249
2022	5	1.696	322	6.783	371	85	408	9.665
2022	6	2.679	800	11.701	1.006	30	272	16.488
2022	7	4.154	1.302	17.045	2.856	46	545	25.948
2022	8	24.066	2.610	28.234	2.540	25	300	57.775
Total anual		35.845	5.243	66.625	7.676	938	1.731	118.058

* Veja nota explicativa no final da página.

INPE's QUEIMADAS Program: 1 km of spatial resolution.

Caatinga biome, august 2022.

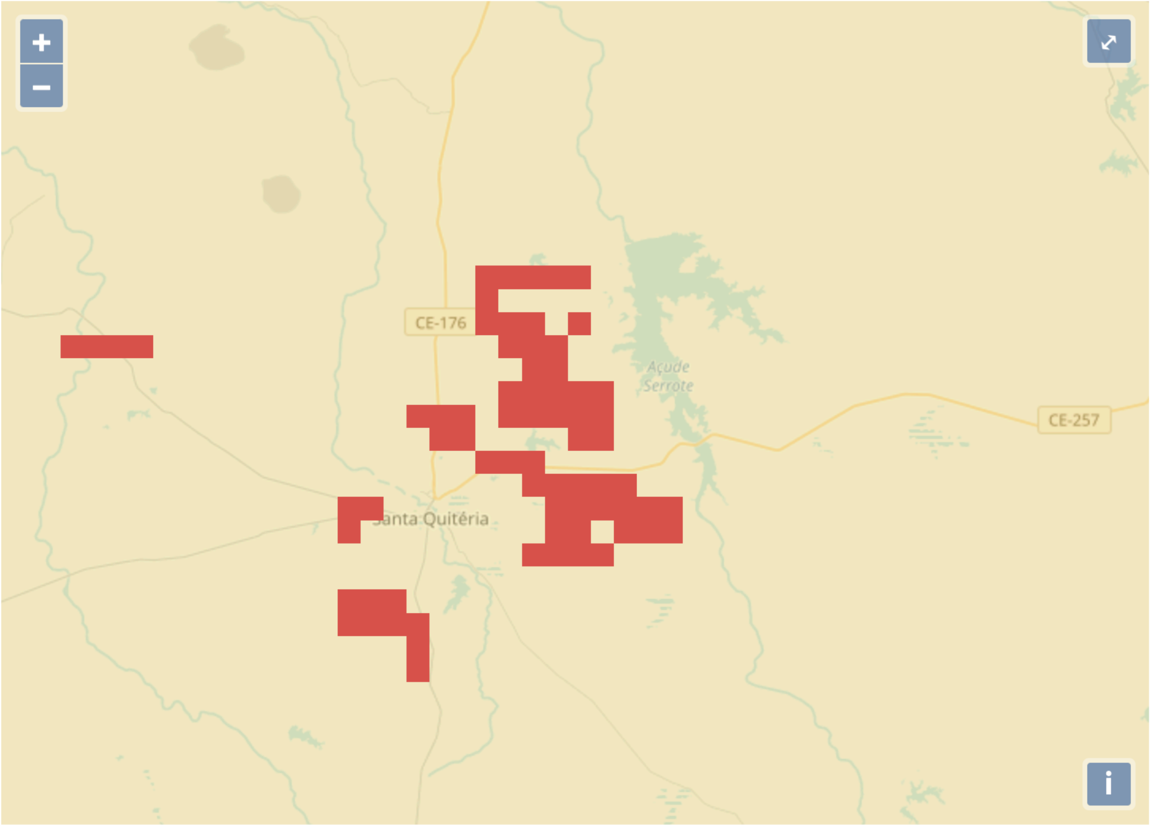
Source: https://queimadas.dgi.inpe.br/queimadas/aq1km/

SEMANTIC SEGMENTATION

- Why is this important for remote sensing?

Mapa de área queimada

para o bioma Caatinga em Agosto de 2022



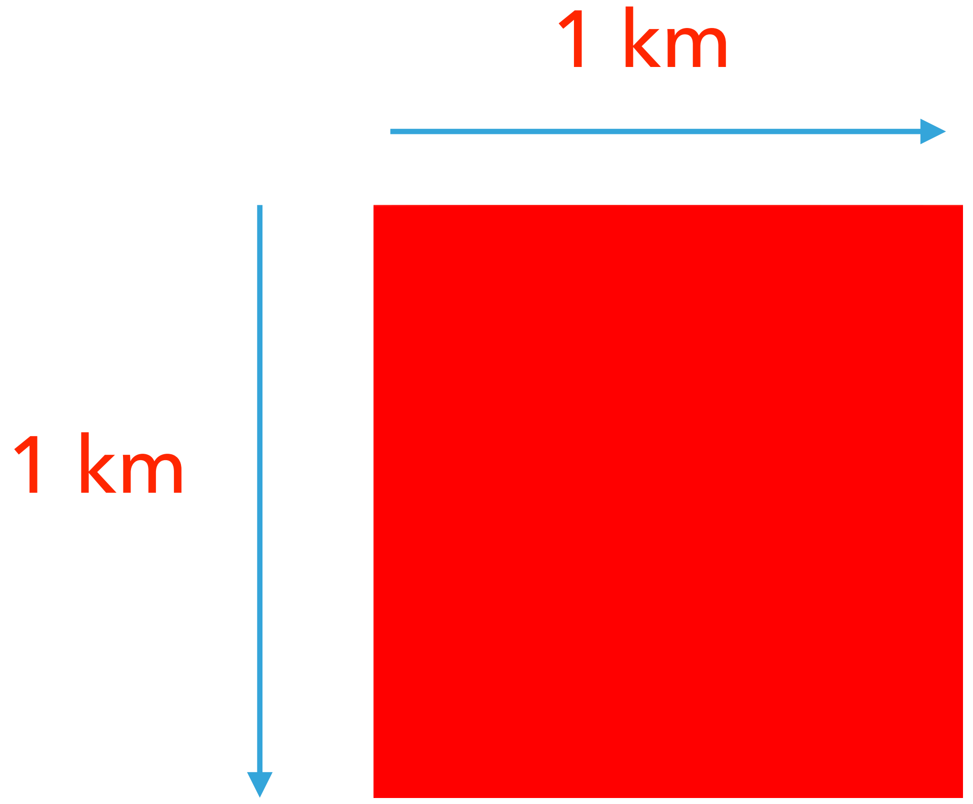
Área queimada (km²)

por bioma em Agosto de 2022

Ano	Mês	Amazônia	Caatinga	Cerrado	Mata Atlântica	Pampa	Pantanal	Total mensal
2022	1	2.273	47	277	452	453	81	3.583
2022	2	504	67	494	181	219	37	1.502
2022	3	64	28	508	156	36	56	848
2022	4	409	67	1.583	114	44	32	2.249
2022	5	1.696	322	6.783	371	85	408	9.665
2022	6	2.679	800	11.701	1.006	30	272	16.488
2022	7	4.154	1.302	17.045	2.856	46	545	25.948
2022	8	24.066	2.610	28.234	2.540	25	300	57.775
Total anual		35.845	5.243	66.625	7.676	938	1.731	118.058

* Veja nota explicativa no final da página.

Each pixel = 1 km^2.

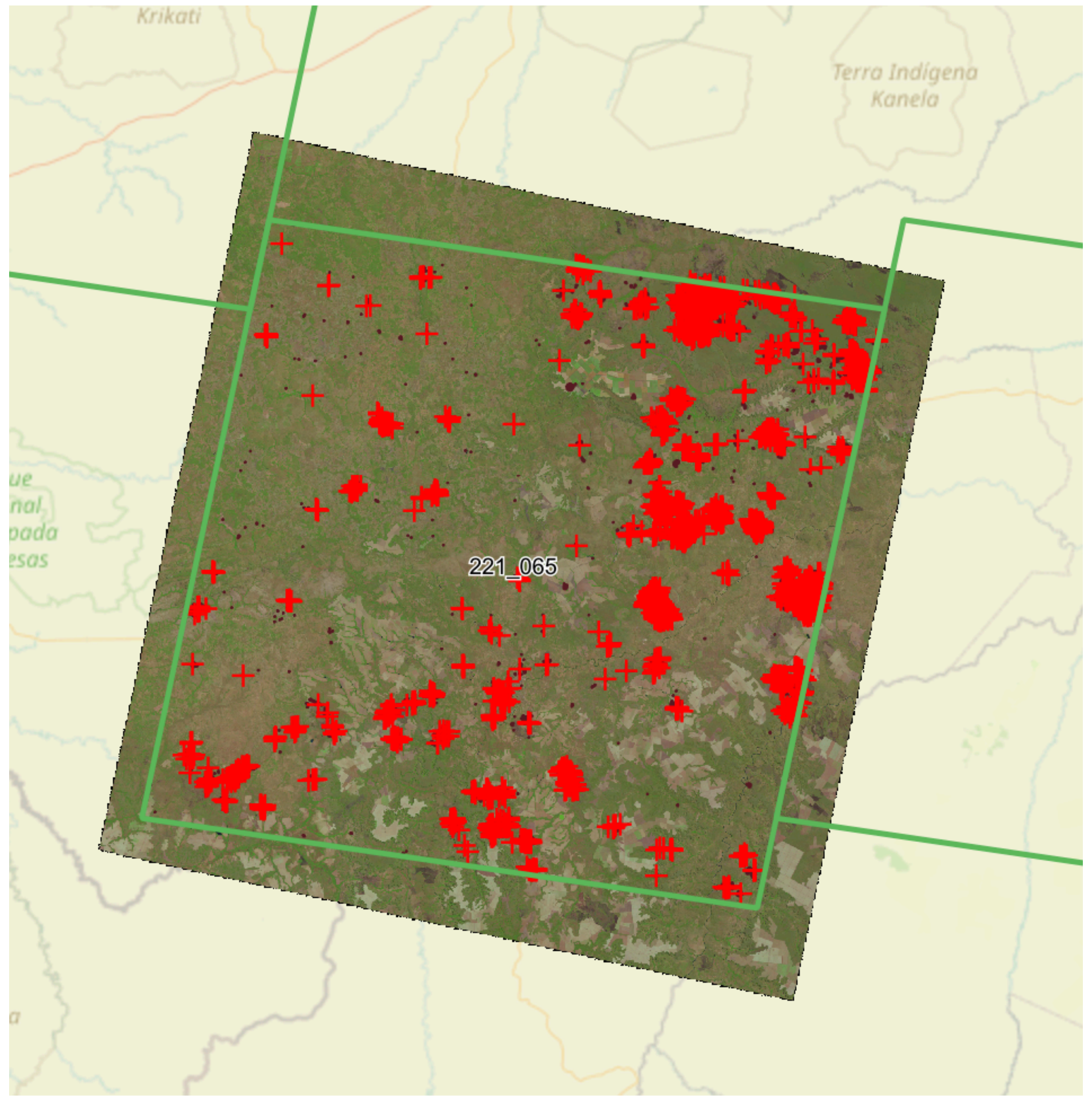


INPE's QUEIMADAS Program: 1 km of spatial resolution.

Caatinga biome, august 2022.

SEMANTIC SEGMENTATION

- QUEIMADAS also rely on 30 m spatial resolution images. Still low resolution.



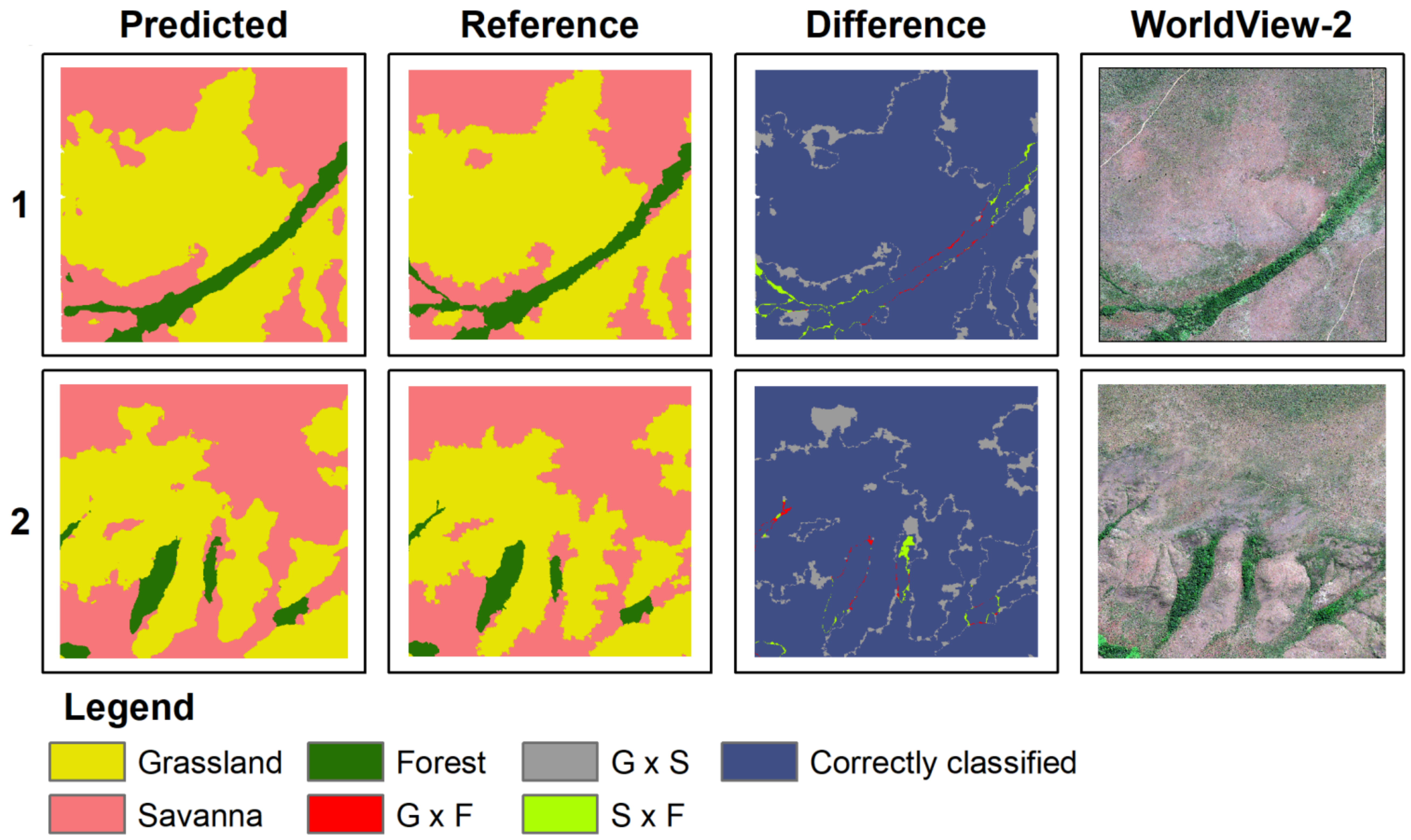
- But, today we have very high-resolution sensors in satellites.
  - Ex: KOMPSAT 3/3A satellite, Korea Aerospace Research Institute (KARI): 50 - 70 cm.

WORKFLOW FOR IMAGE SEMANTIC SEGMENTATION AT INPE (SIMPLIFIED)

- 1.) Download the images (separated bands) from:
    - http://www2.dgi.inpe.br/catalogo/explore.
- 2.) Compose the bands (QGIS).
- 3.) Generate images patches by cropping the entire scene. It results in patches of dimensions of 128 x 128, 224 x 224, 256 x 256 pixels.
- 4.) Generate the masks for each image patch.
- 5.) Train the deep neural network (DNN) model for semantic segmentation. The model receives **image patches and the corresponding masks of the training and validation datasets**.
- 6.) Evaluate the performance of the model (inference phase). The model receives **only the image patches of the test dataset and predicts the masks**. Metrics: pixel-based accuracy, pixel-based F1-score, mean Intersection over Union (IoU).

PATCHES AND MASKS

- During the training phase, the DNN receives …



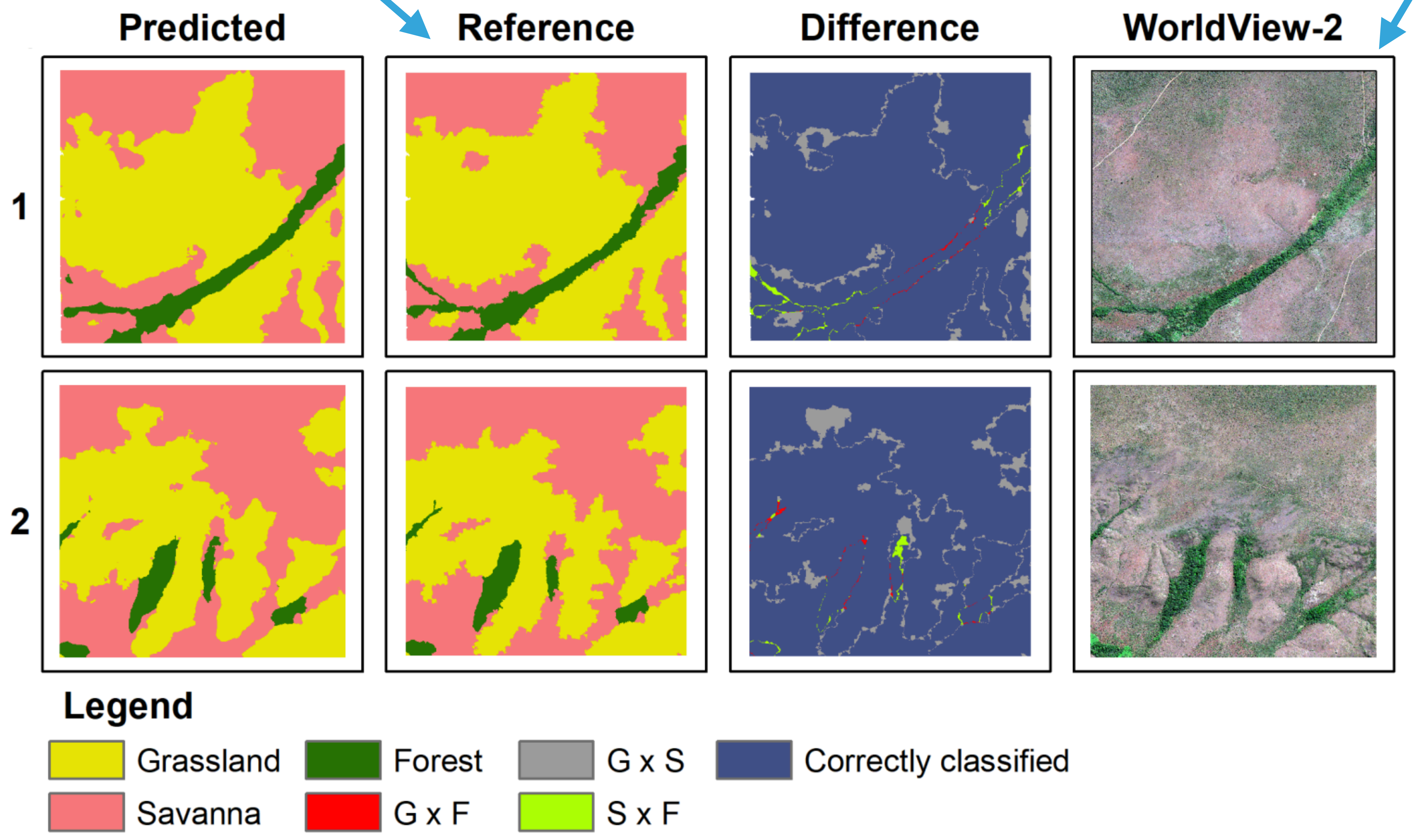
Source: Neves (2021) [3]

PATCHES AND MASKS

- During the training phase, the DNN receives the image patches and masks.

Masks (Ground Truth)

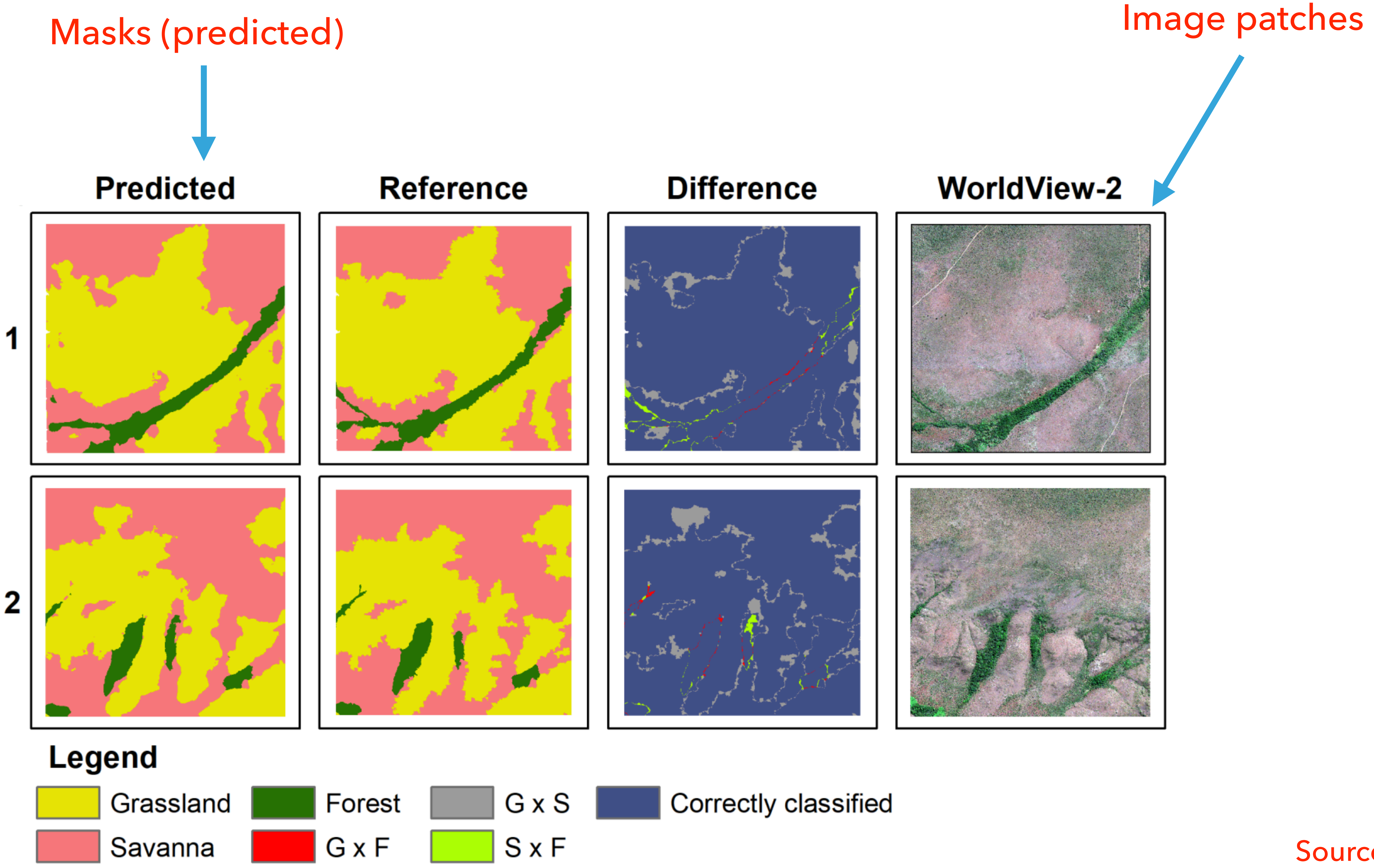
Image patches



Source: Neves (2021) [3]

PATCHES AND MASKS

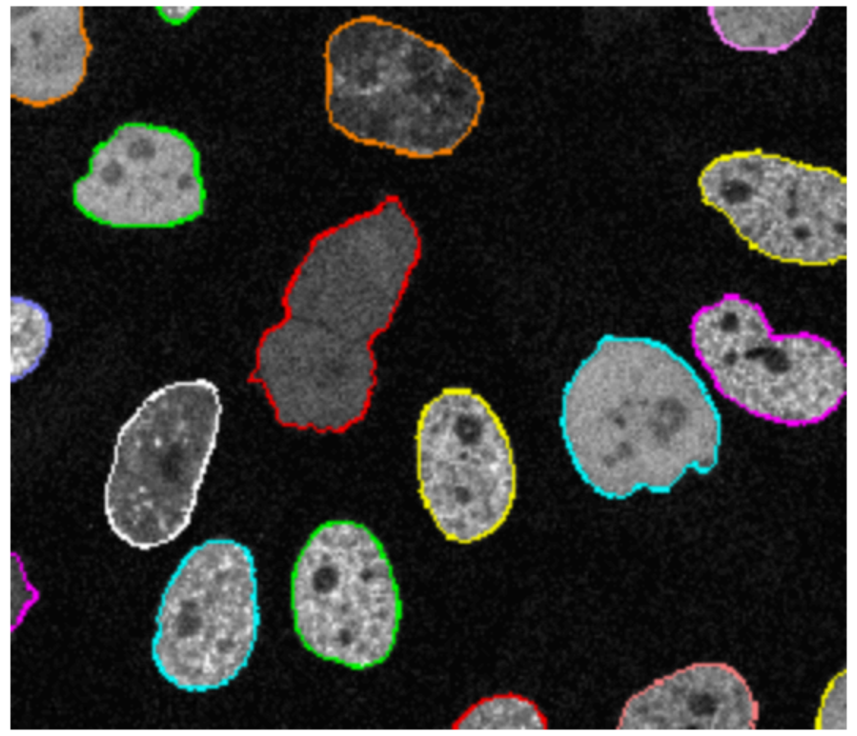
- During the inference phase, the DNN receives the image patches and predicts the masks.



Source: Neves (2021) [3]

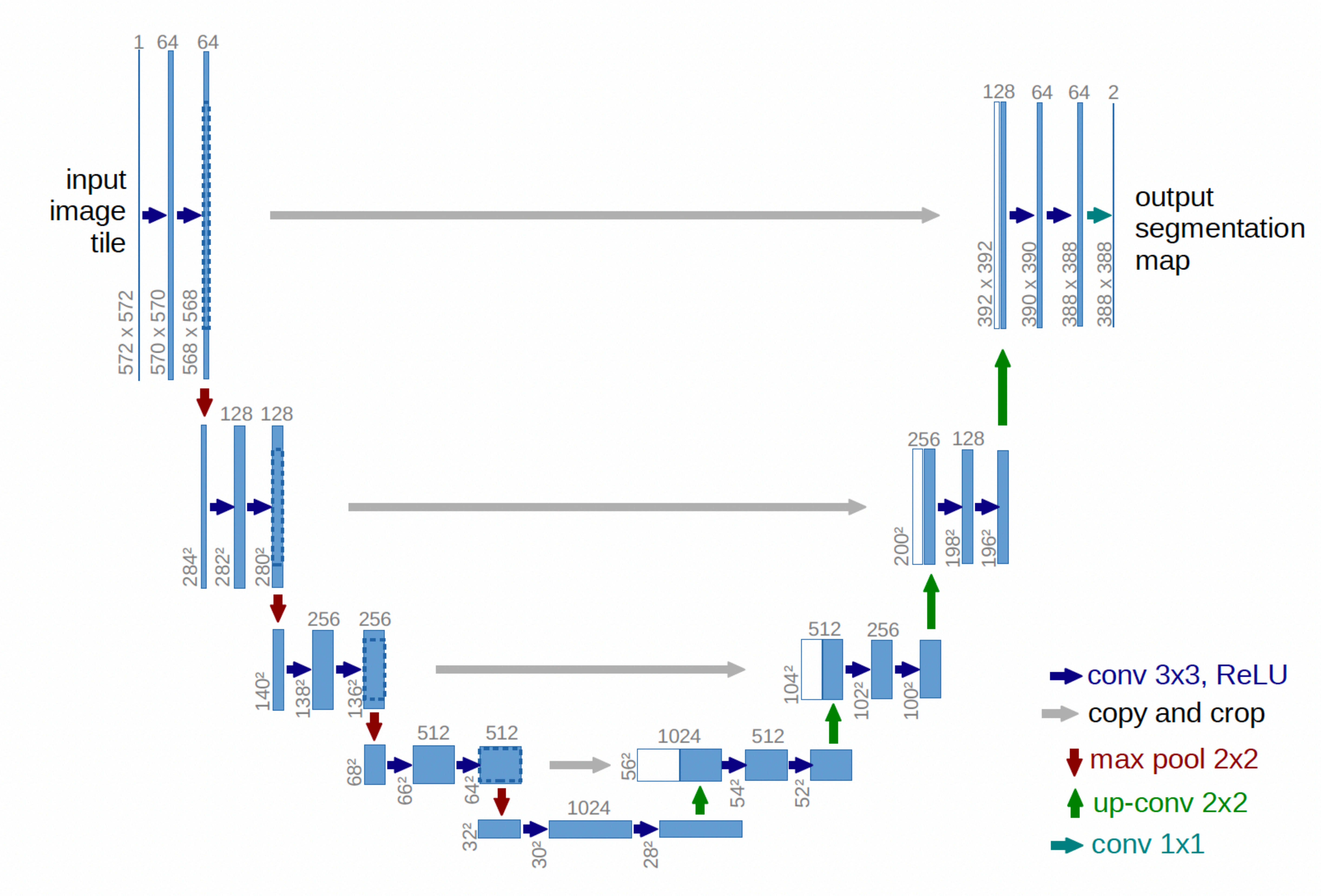
U-NET

- It is a convolutional neural network (CNN).
- U-Net's main characteristics [1]:
  - Suitable for semantic segmentation. But it can also be used for classification;
  - Contracting path to capture context and symmetric expanding path for precise localisation (U-shape);
  - Strong use of data augmentation (few images);
  - 23-layer deep fully convolution network (FCN).



Cell Tracking - IEEE International Symposium on Biomedical Imaging (ISBI 2015): 1st place.

U-NET



1
64
64
input image tile
572 x 572
570 x 570
568 x 568
128
128
284²
282²
280²
256
256
140²
138²
136²
512
512
68²
66²
64²
1024
32²
30²
28²
1024
512
56²
54²
52²
512
256
104²
102²
100²
256
128
200²
198²
196²
128
64
64
2
392 x 392
390 x 390
388 x 388
388 x 388
output segmentation map
conv 3x3, ReLU
copy and crop
max pool 2x2
up-conv 2x2
conv 1x1

U-NET

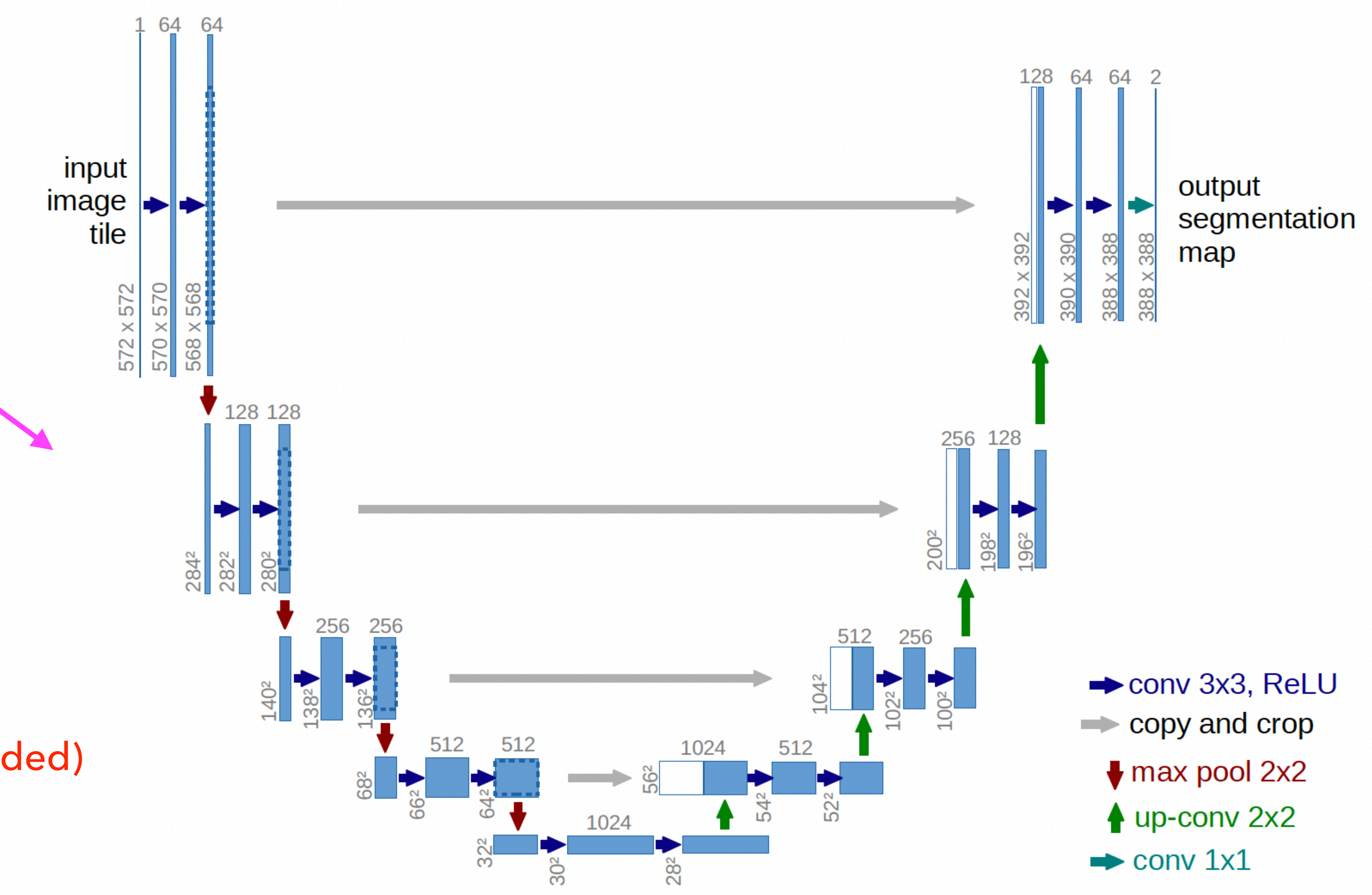
- Semantic segmentation classifies at pixel-level. Good: spatial dimensions of the input and output are the same.
  - Channel dimension at one output pixel can hold the classification results for the input pixel at the same spatial position.

- Thus, some types of CNN layers can be used to increase (upsample) the spatial dimensions of intermediate feature maps, after the spatial dimensions are reduced (downsampled) by previous CNN layers.

U-NET

Contracting path

(decrease dimension size).

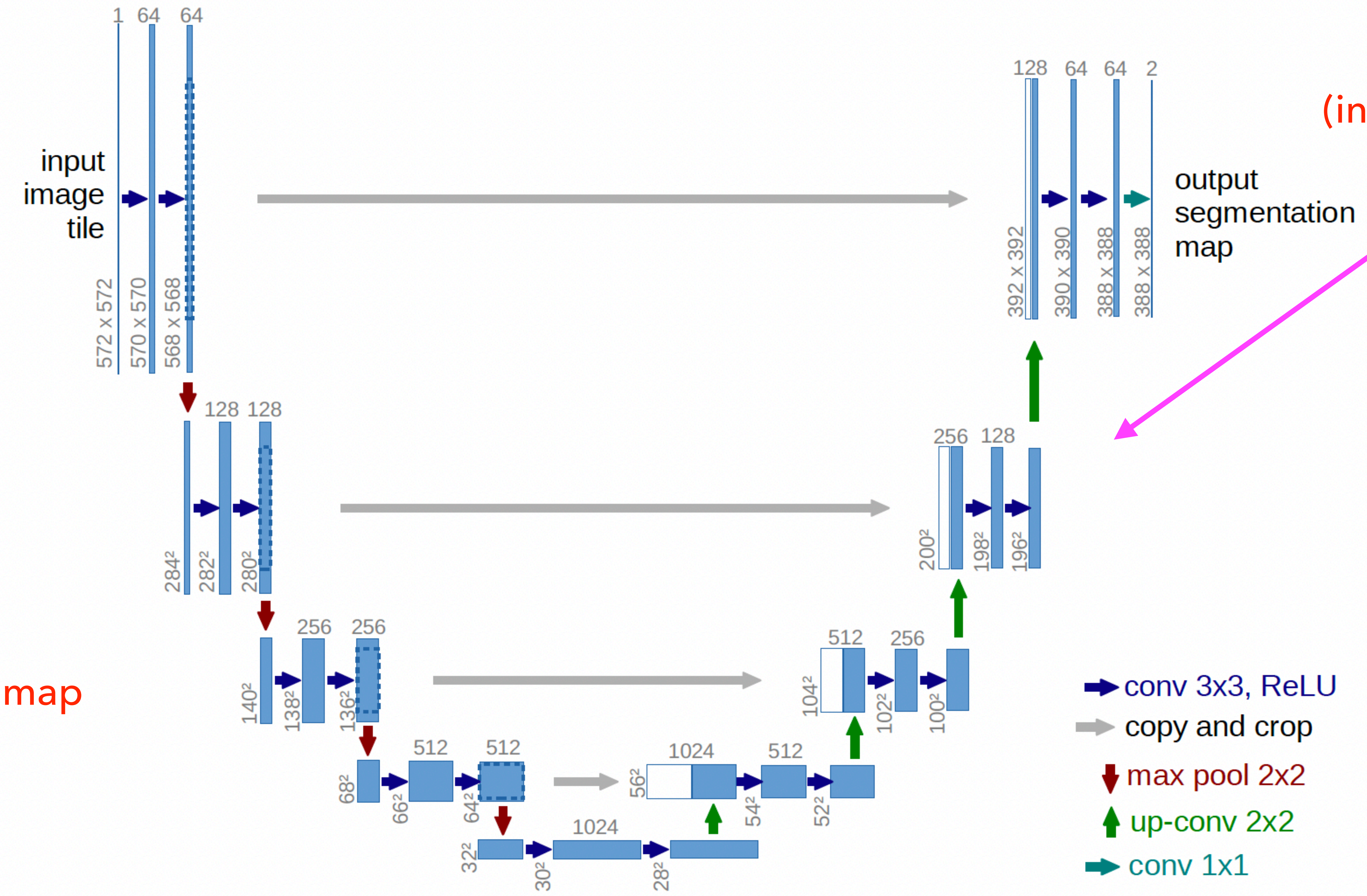


Two 3x3 convolutions (unpadded)

ReLU

2x2 max pooling (stride 2).

U-NET



Expanding path

(increase dimension size).

2x2 conv. (up-conv)

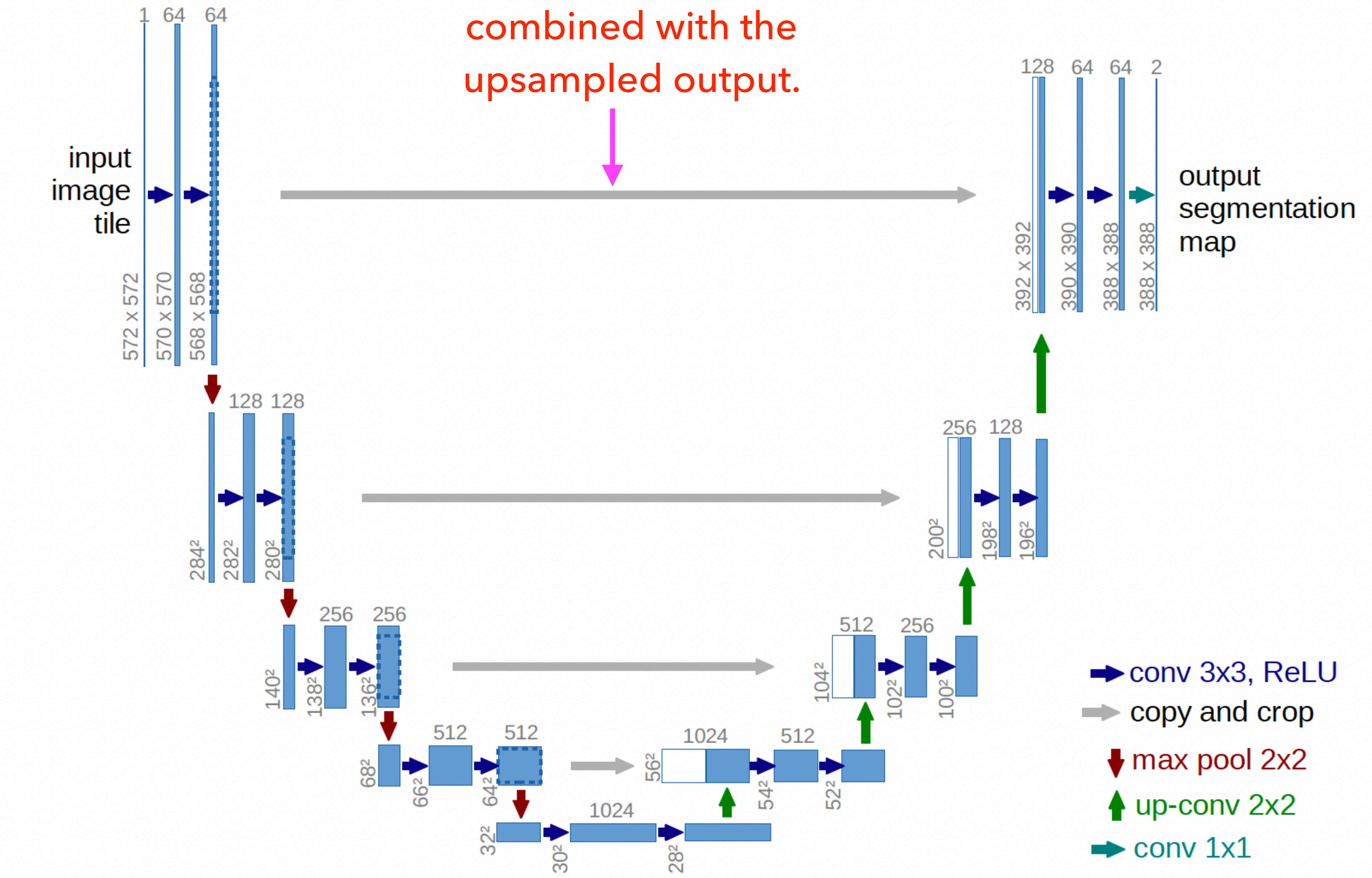
Concatenation w/cropped feat. map

Two 3 x 3 convolutions

ReLU

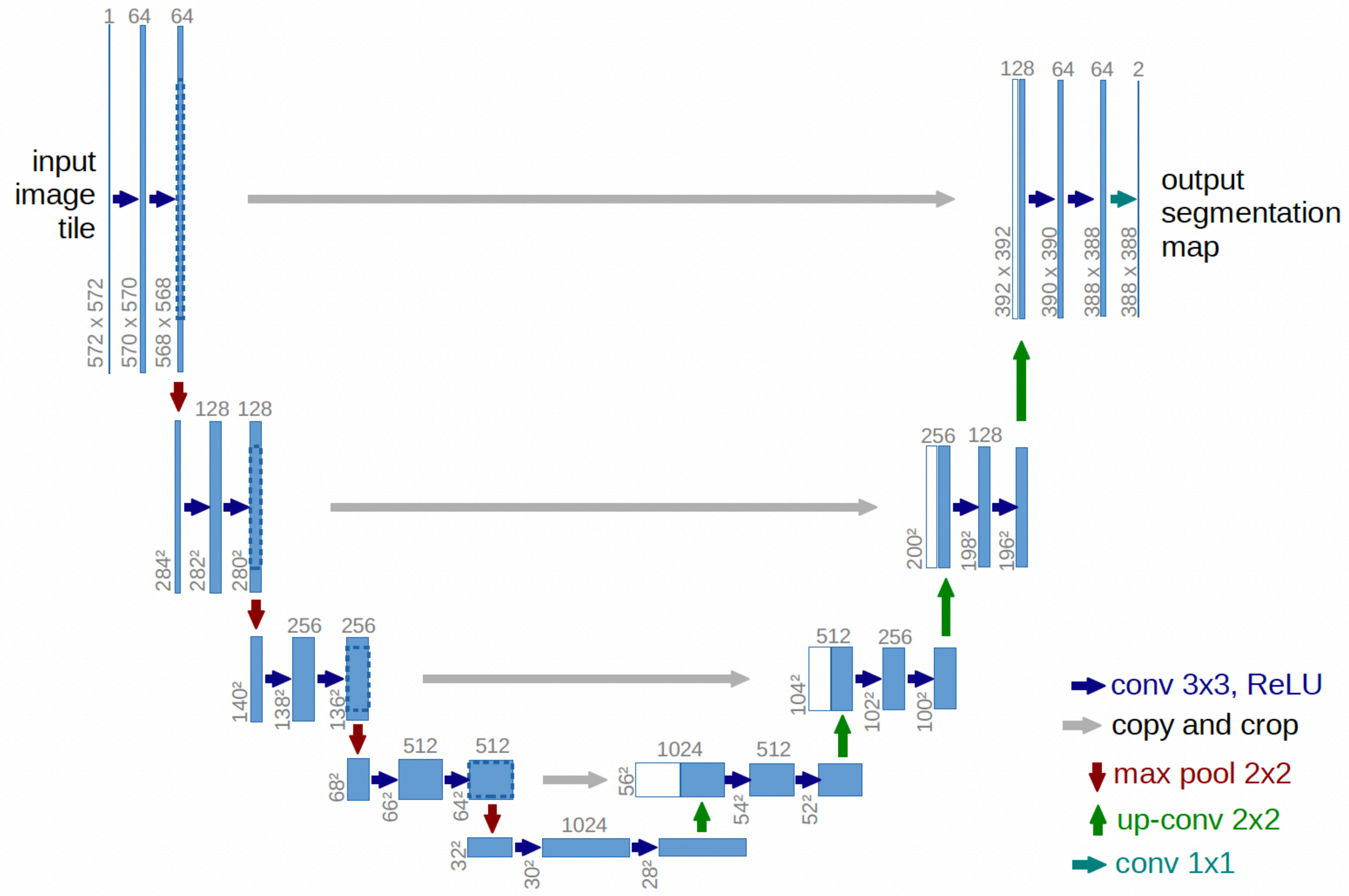
U-NET

Localisation: high resolution features from the contracting path are combined with the upsampled output.



No FC layer at all!

U-NET

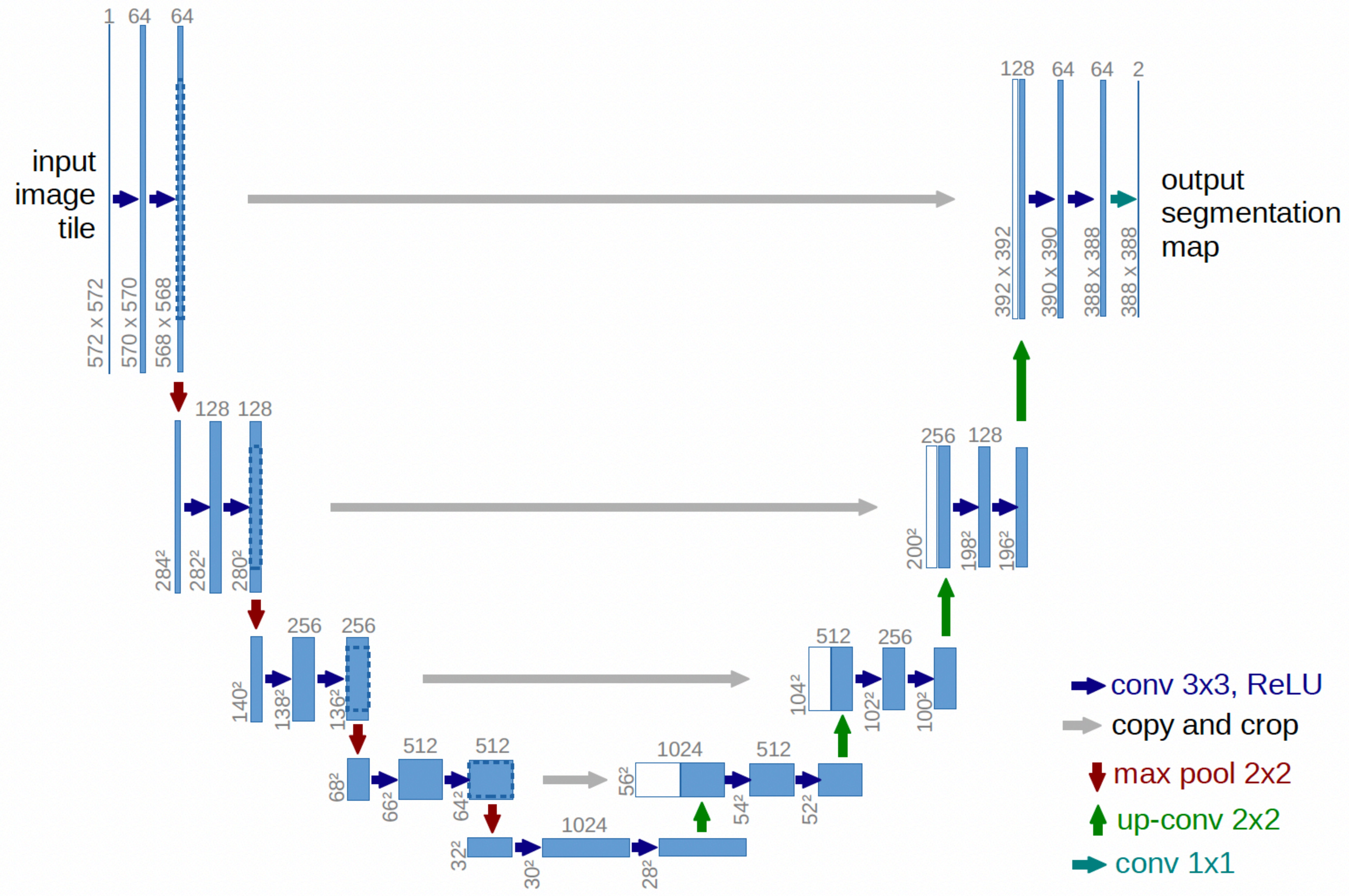


Biomedical domain (others too): not huge number of images available.

Localisation: label for each pixel and not for the entire image (**each pixel matters**)!

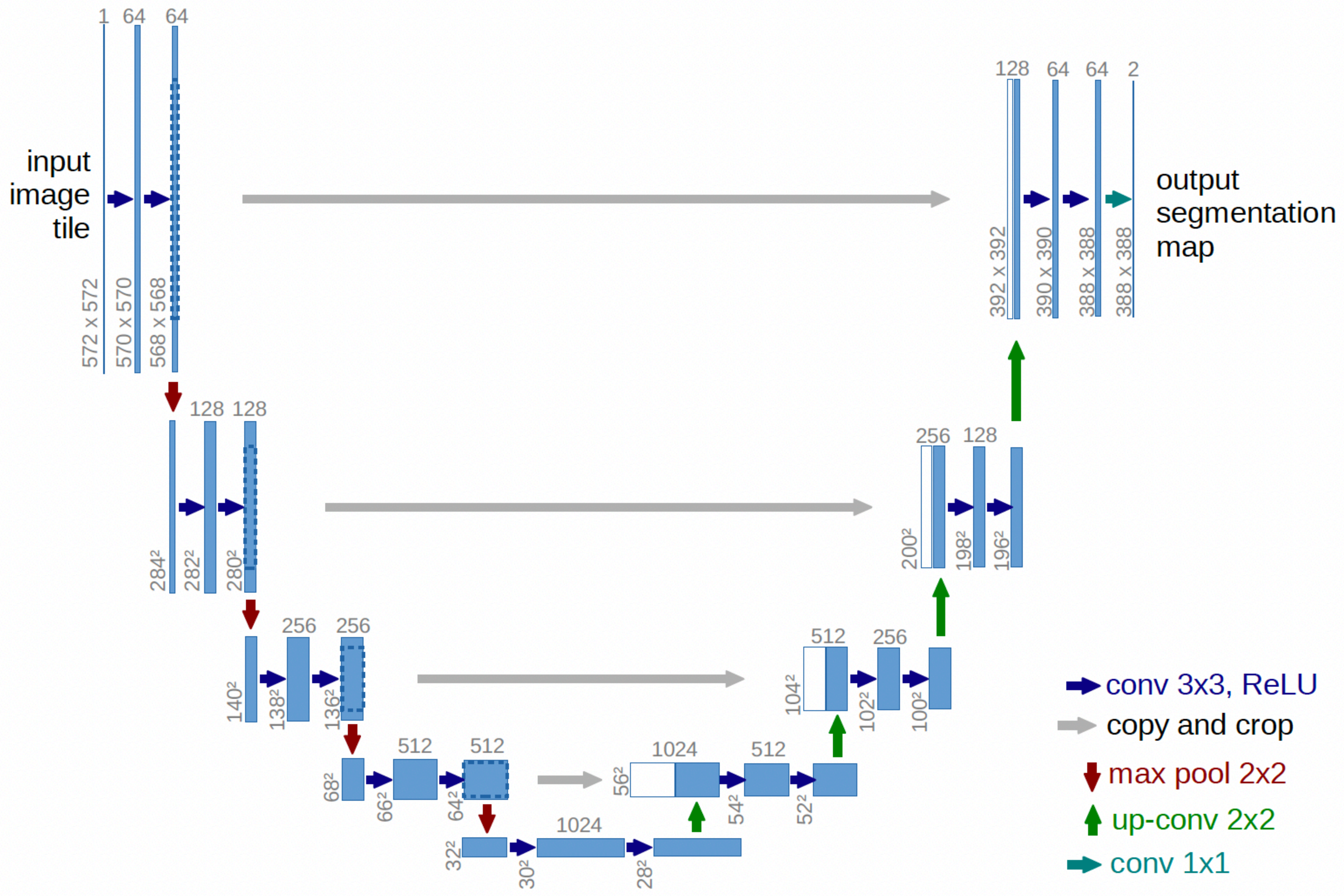
Label for each pixel: using a local region (patch) around the pixel as input.

U-NET



Training: in order to minimise the overhead and make maximum use of the GPU memory, large input tiles were selected over a large batch size.
batch = 1 image.

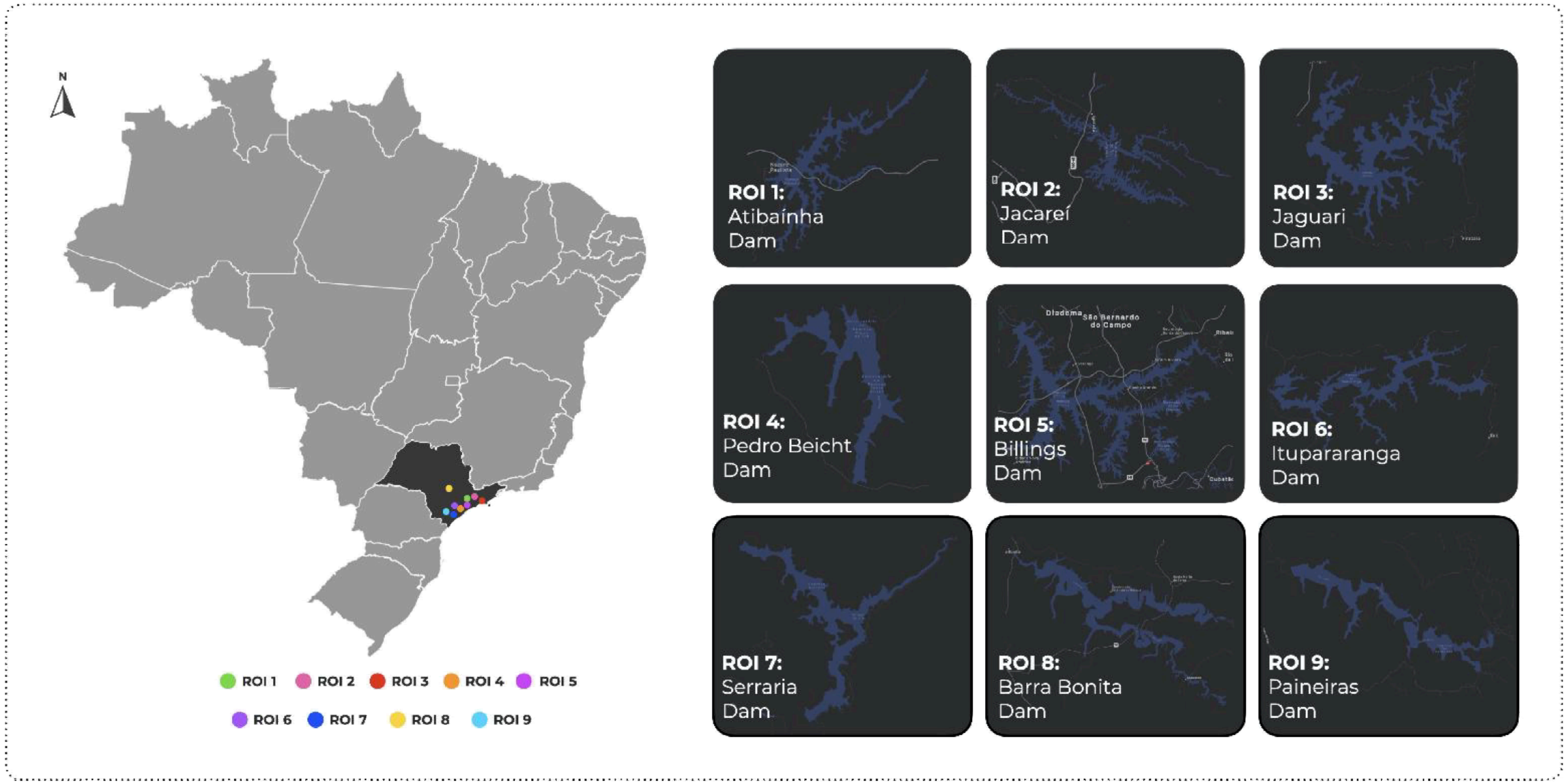
U-NET



Currently, there are several versions of the U-Net:
See Ref [2].

CNN FOR CLASSIFYING WATER VOLUME OF DAMNS

- Contextual classification (not semantic segmentation here).



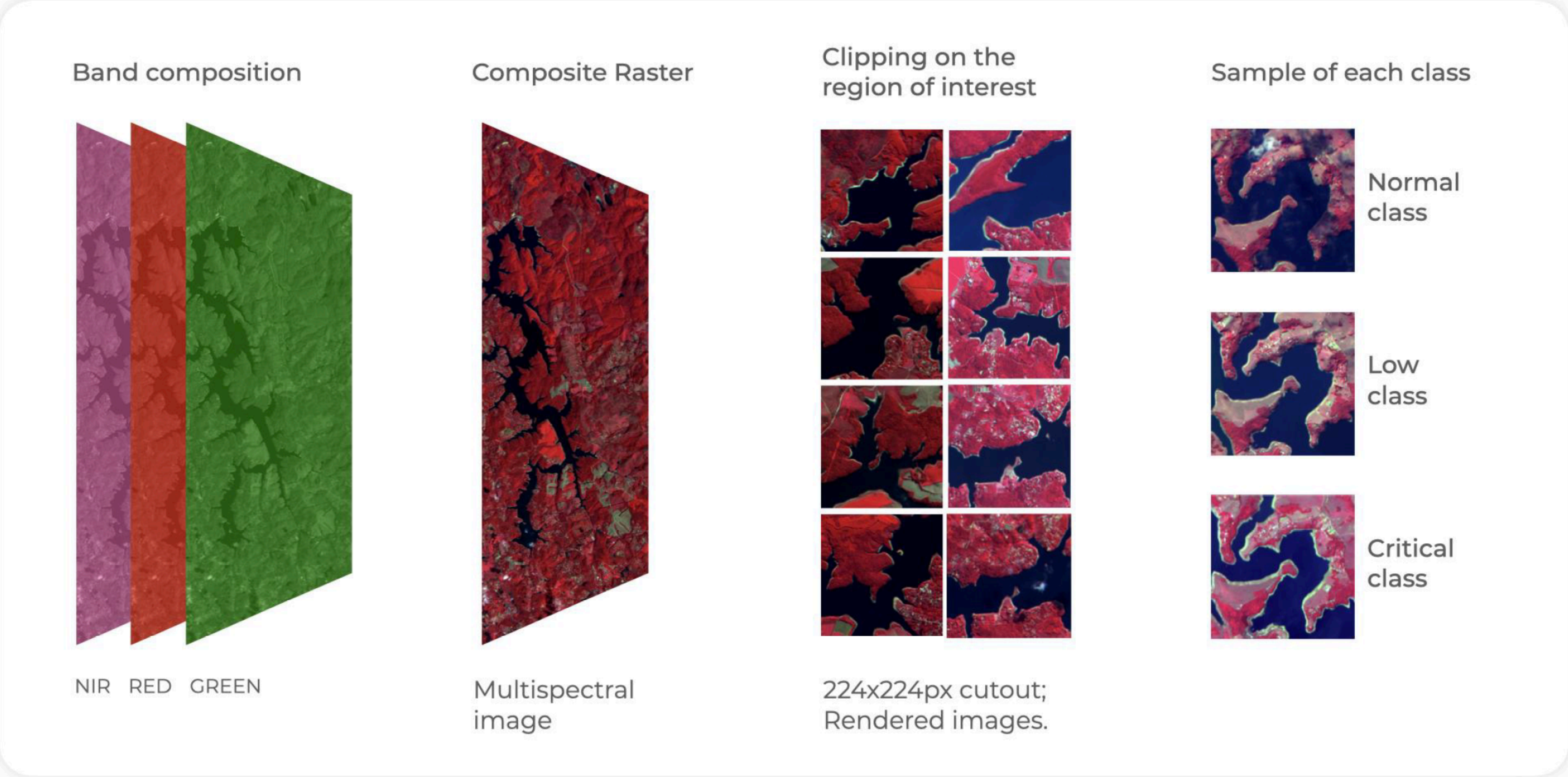
Study site:

Nine damns of state of São Paulo.

Source: Miranda et al. (2021) [4]

CNN FOR CLASSIFYING WATER VOLUME OF DAMNS

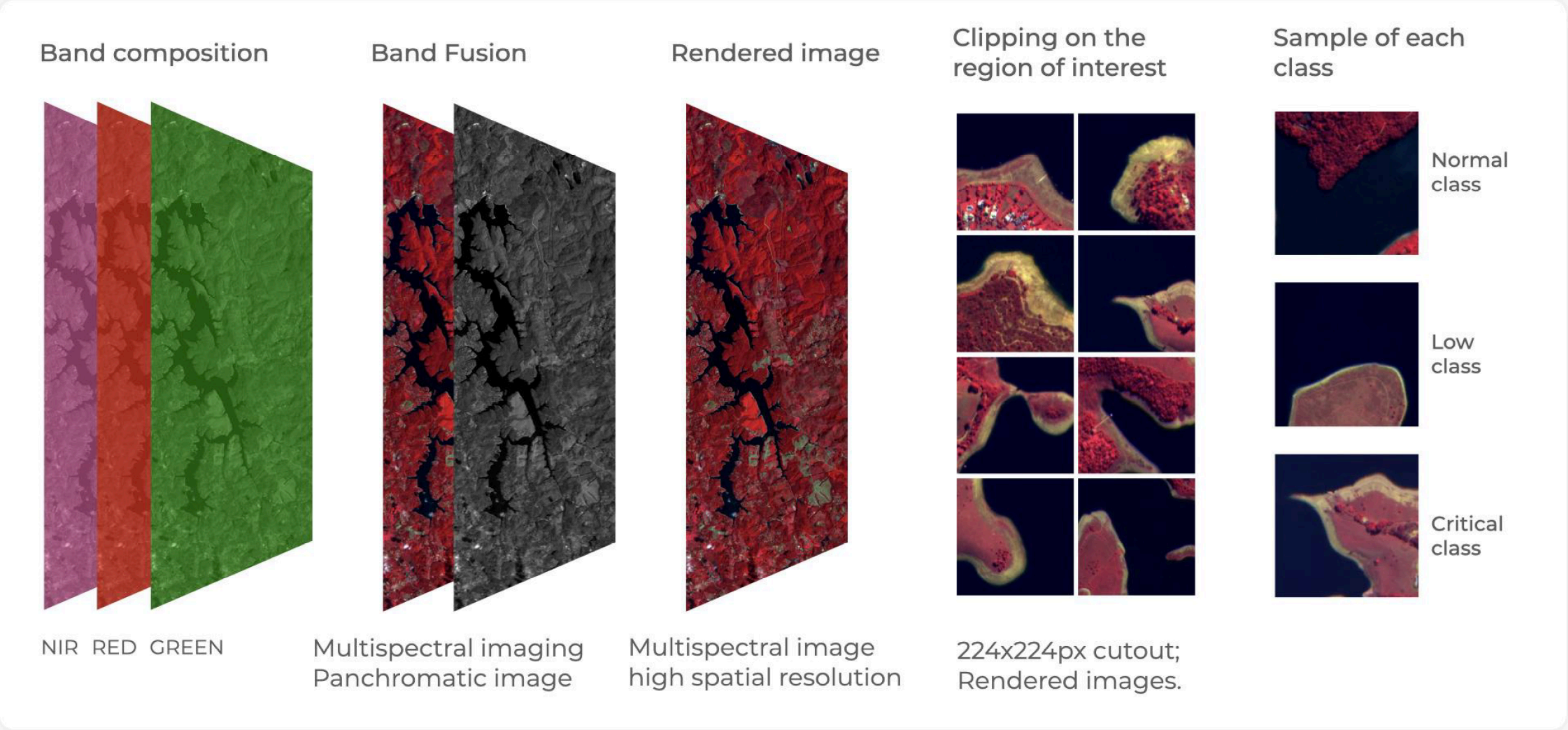
- Training dataset: CBERS-4 PAN10M sensor (10 m of spatial resolution).



Source: Miranda et al. (2021) [4]

CNN FOR CLASSIFYING WATER VOLUME OF DAMNS

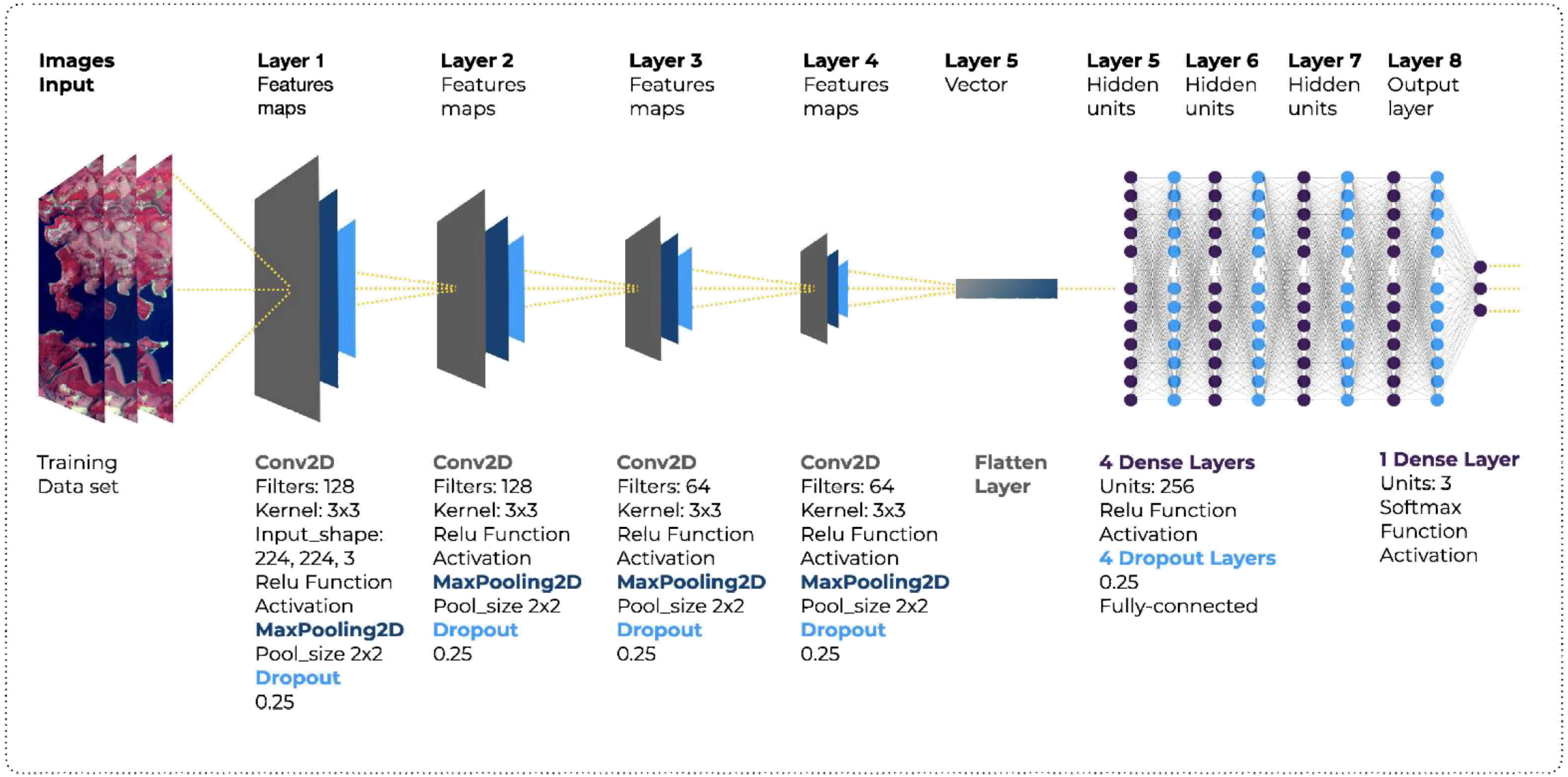
- Test dataset: CBERS-4A WPM sensor (2 m of spatial resolution after pan sharpening/fusion).



Source: Miranda et al. (2021) [4]

CNN FOR CLASSIFYING WATER VOLUME OF DAMNS

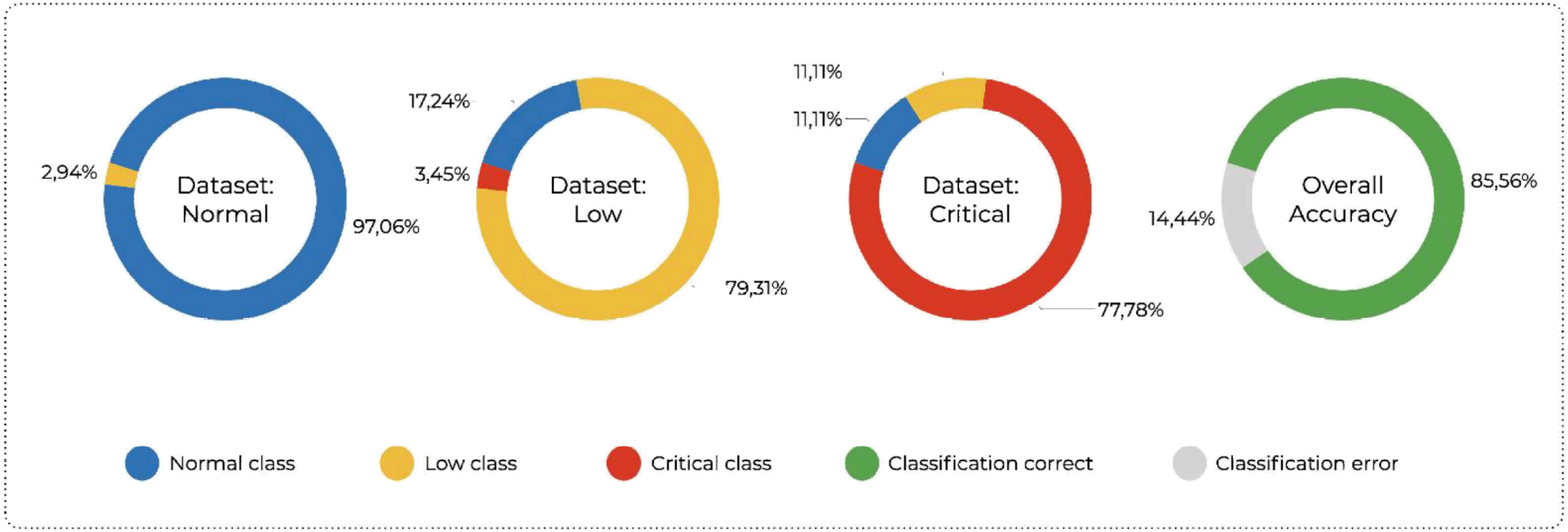
- Our CNN: CerraNet.



Source: Miranda et al. (2021) [4]

CNN FOR CLASSIFYING WATER VOLUME OF DAMNS

- Results.



Source: Miranda et al. (2021) [4]

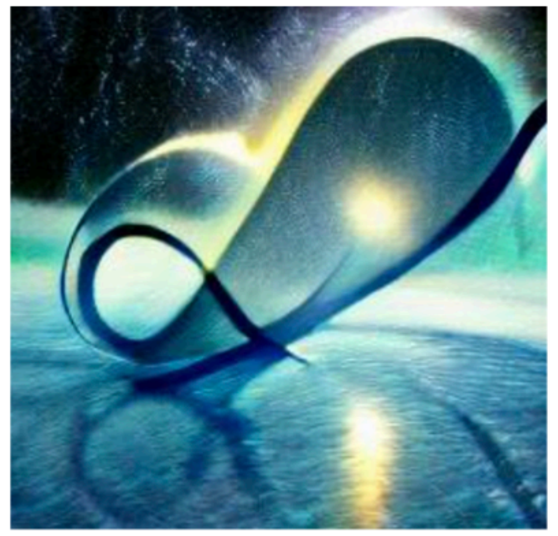
SOME AVAILABLE SOURCE CODE AND DATASETS

- General deep learning code (semantic segmentation, contextual classification, ...): https://github.com/vsantjr (Valdivino Santiago Júnior).
- CerraNet CNN: https://github.com/MirandaMat/cerraNet-v2 (Mateus Miranda).
- CerraData dataset (Cerrado biome) and code: https://github.com/ai4luc/CerraData-code-data (Mateus Miranda, Lucas Silva, Samuel Santos, Valdivino Santiago Júnior, Thales Körting, Jurandy Almeida).
- Time series prediction via deep neural networks: https://github.com/RenatoMaximiano/ETOUNN (Renato Maximiano).
- WorCAP's Hackathon 2022: https://www.kaggle.com/competitions/hackathon-worcap-2022.

PROJECT IDEEPS

- Classificação de imagens via redes neurais profundas e grandes bases de dados para aplicações aeroespaciais.

Project IDeepS



Source: https://github.com/vsantjr/IDeepS

REFERENCES

- [1] O. Ronneberger, P. Fischer, and T. Brox. 2015. U-Net: Convolutional Networks for Biomedical Image Segmentation. arXiv:1505.04597 [cs.CV].
- [2] N. Siddique, S. Paheding, C. P. Elkin, and V. Devabhaktuni. 2021. U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications. IEEE Access 9 (2021), 82031–82057. https://doi.org/10.1109/ACCESS.2021.3086020.
- [3] A. K. Neves. Hierarchical mapping of Brazilian Savanna (Cerrado) physiognomies based on Deep Learning. 2021. 96 p. IBI: <8JMKD3MGP3W34R/44DTSUS>. (sid.inpe.br/mtc-m21c/2021/03.30.18.49-TDI). Tese (Doutorado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2021. Disponível em: <ibi:8JMKD3MGP3W34R/44DTSUS>.

REFERENCES

- [4] M. S. Miranda, R. S. Maximiano, V. A. Santiago Júnior, T. S. Körting, L. M. G. Fonseca. Classification of the water volume of dams using heterogeneous remote sensing images through a deep convolutional neural network. In: XXII Brazilian Symposium on Geoinformatics (GEOINFO), 2021, Sao Jose dos Campos. Proceedings of the XXII GEOINFO, 2021. v. 1. p. 179-188.

THANK YOU!

- E-mail: valdivino.santiago@inpe.br.
- Personal web page: http://www.lac.inpe.br/valdivino/.