

CatLC: datos abiertos en multiresolución para la generación de productos y cubiertas del suelo con IA

Inteligencia Artificial en Observación de la Tierra

catlc@icgc.cat



del DATO al CONOCIMIENTO



Ciencia



Tecnología



Retos a
resolver



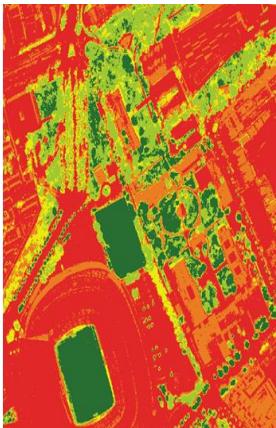
Valor
añadido



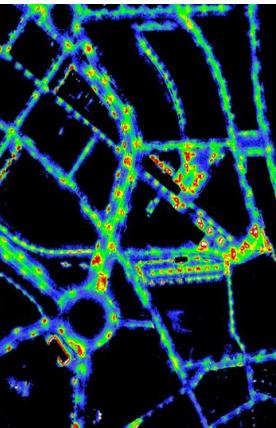
movimiento suelo



emisión térmica

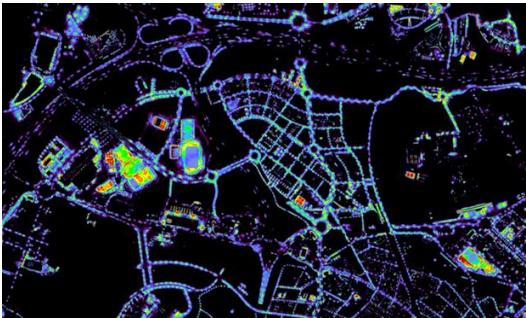


verde urbano

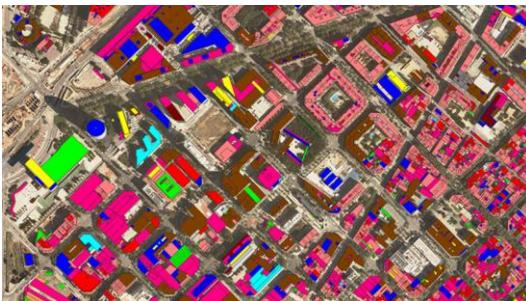


contaminación lumínica

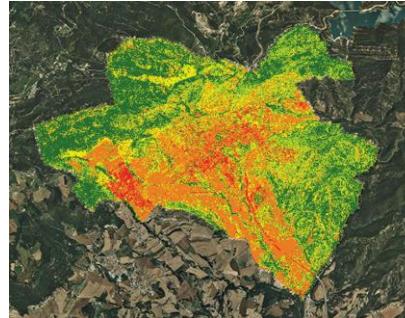
del DATO al CONOCIMIENTO



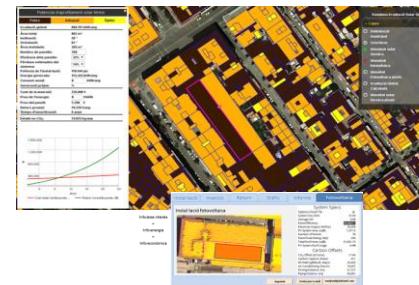
Contaminación lumínica.
Sensores hiperespectrales VNIR



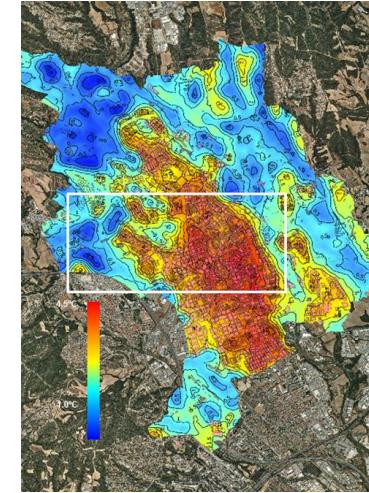
Clasificación de materiales.
Sensores hiperespectrales VNIR + Térmicos



Verde urbano. Cámaras fotogramétricas

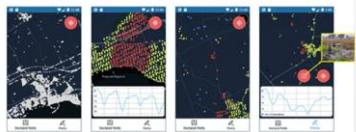


Potencial Solar. LiDAR

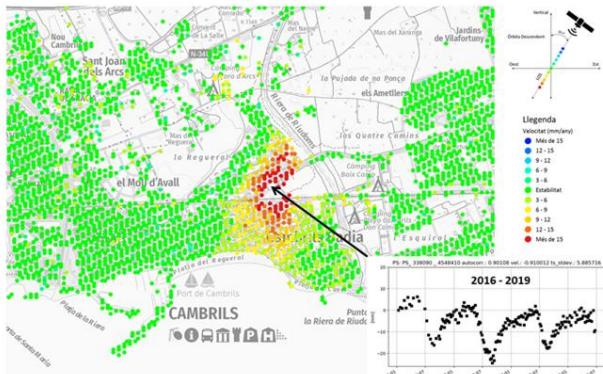


Isla de calor urbana.
Sensores hiperespectrales térmicos

del DATO al CONOCIMIENTO

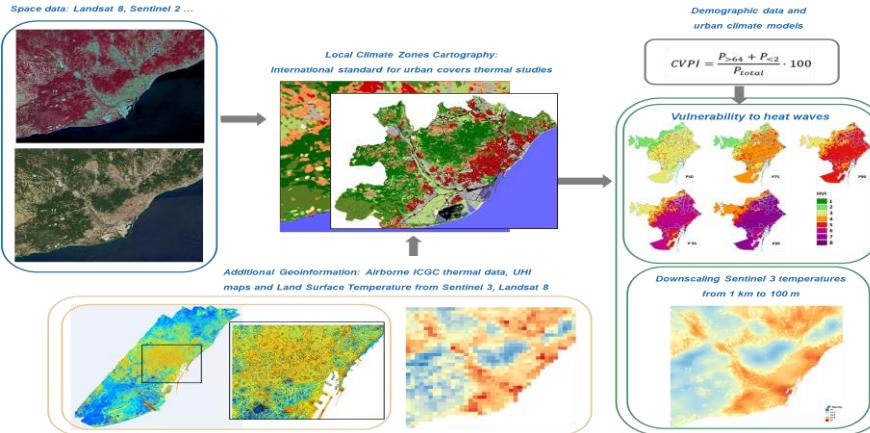


Quifer app wins top prize at 2020 Space App Camp



Monitoreo de acuíferos. Sensores Radar

Jornadas SIG Libre 2023. Girona, 14.06.2023



Vulnerabilidad a las olas de calor. Sensores hiperespectrales VNIR + Térmicos



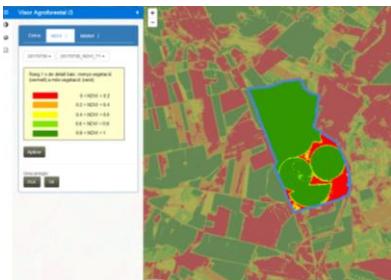
Exposición que analiza el papel de la OT en los ODS.

<https://www.space4ourplanet.org/>

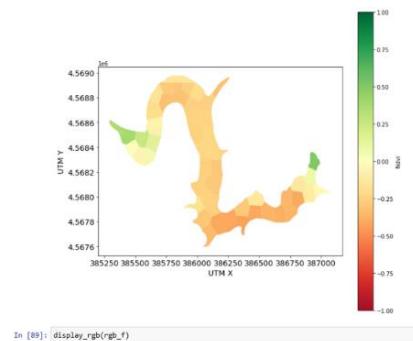
del DATO al CONOCIMIENTO



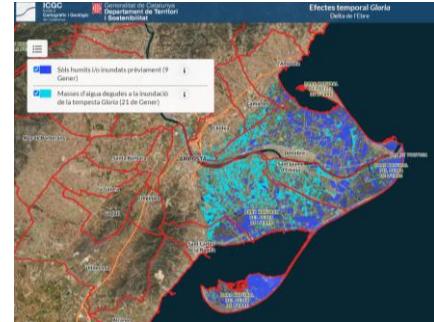
Visor mosaicos mensuales de imágenes **Sentinel-2**



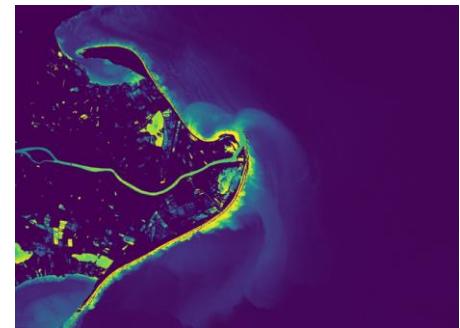
Visor agroforestal



Monitoreo inundaciones. **Imágenes SAR**



Calidad del agua en masas continentales

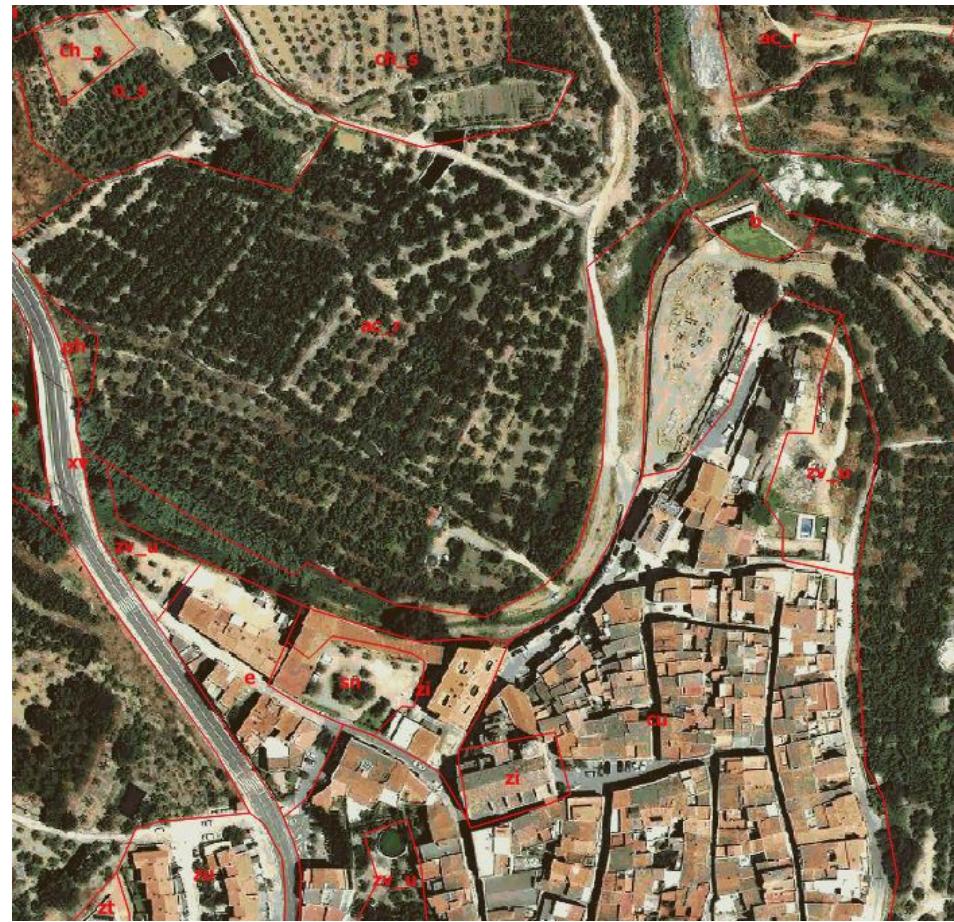


Contenido de clorofila del agua

MCSC - metodología

fotointerpretación de ortofotos de escalas grandes con información auxiliar y aprovechamiento de otras fuentes existentes

área mínima: 500 m²



MCSC - metodología

fotointerpretación de ortofotos de escalas grandes con información auxiliar y aprovechamiento de otras fuentes existentes

área mínima: 500 m²

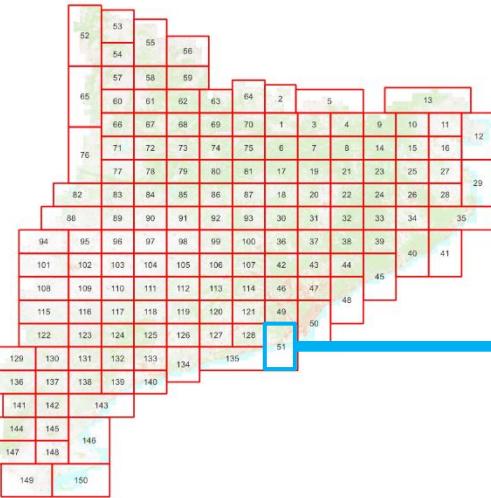


41 categorías (obligatorias)

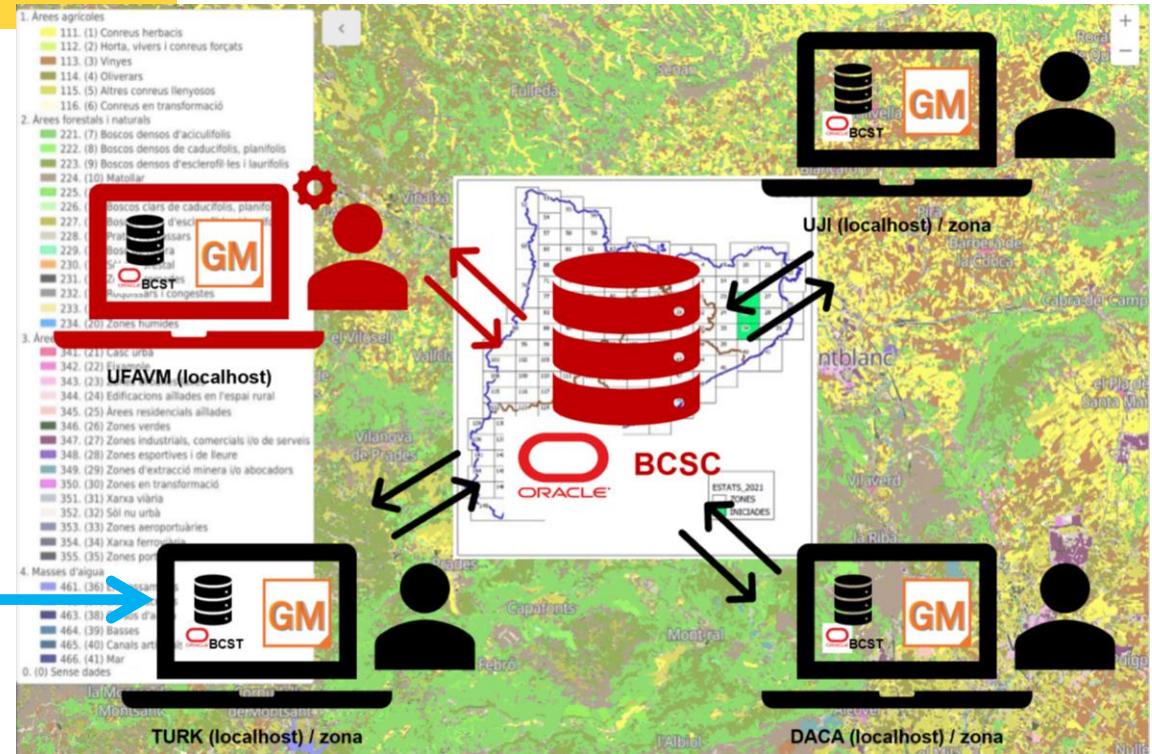
31 atributos (opcionales)



Arquitectura



1. Àrees agrícoles
 111. (1) Coneus herbàcis
 112. (2) Horta, vivers i coneus forçats
 113. (3) Vinyes
 114. (4) Oliveras
 115. (5) Altres coneus llenyosos
 116. (6) Coneus en transformació
2. Àrees forestals i naturals
 221. (7) Boscos densos d'acículifolis
 222. (8) Boscos densos de caducifolis, planifolis
 223. (9) Boscos densos d'escièrrofles i laurifolis
 224. (10) Matorral
 225.
 226.
 227.
 228.
 229. Prats d'estiu
 230. S'ldr. S'ldr. S'ldr.
 231. Rouredes i conges
 232.
 233.
 234. (20) Zones humides
3. Àrees urbanes
 341. (21) Carrer urbà
 342. (22) Exemple
 343. (23) Carrer
 344. (24) Edificacions afiliades en l'espai rural
 345. (25) Àrees residencials afiliades
 346. (26) Zones verdes
 347. (27) Zones industrials, comercials i/o de serveis
 348. (28) Zones esportives i de lleure
 349. (29) Zones d'extracció minera i/o abocadors
 350. (30) Zones en transformació
 351. (31) Xarxa viària
 352. (32) Sòl nu urbà
 353. (33) Zones aeroportuàries
 354. (34) Xarxa ferroviària
 355. (35) Zones portuàries
4. Masses d'aigua
 461. (36)
0. (0) Sense dades

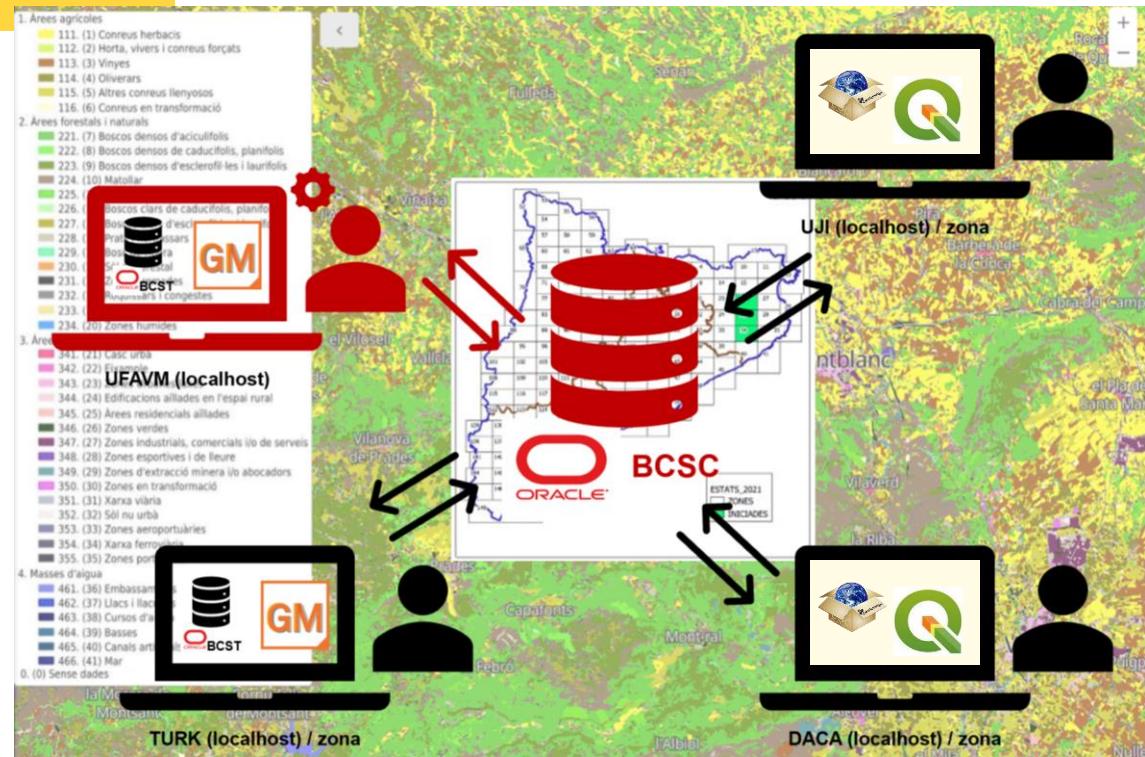


Arquitectura actual

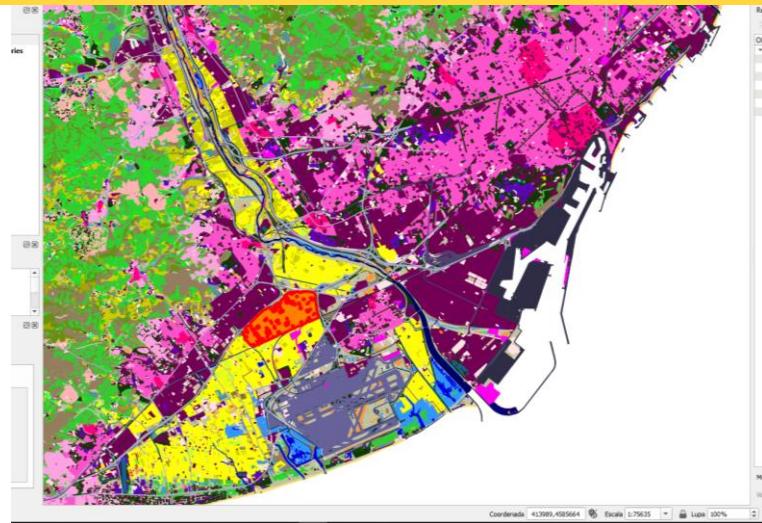
Migración a QGIS



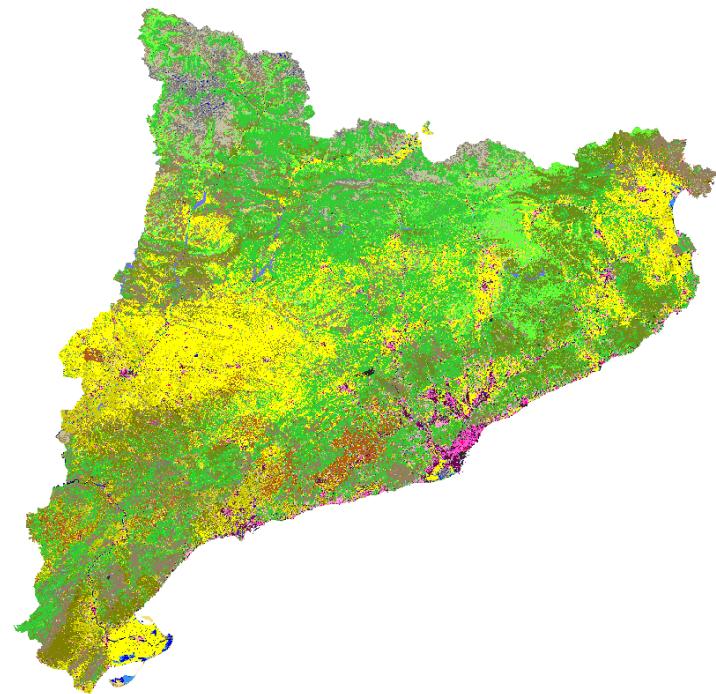
del entorno
fotointerpretación



MCSC (2009/2018)



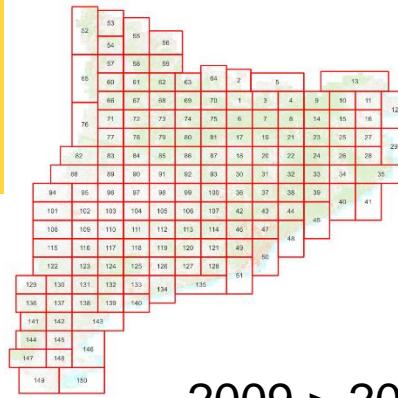
- GeoPackage
- Leyenda incorporada
- Tabla de colores
- 1.6 GB



- GSD 1 m
- Big TIFF
- 1 banda + tabla colores
- 1.8 Gb

MCSC - metodología

Pub(Y): $\frac{3}{4}$ GC + $\frac{1}{4}$ PC
ortofoto Y-1



2009 > 2018

GC: actualización **cambios > 2 ha**

2018 > 2022

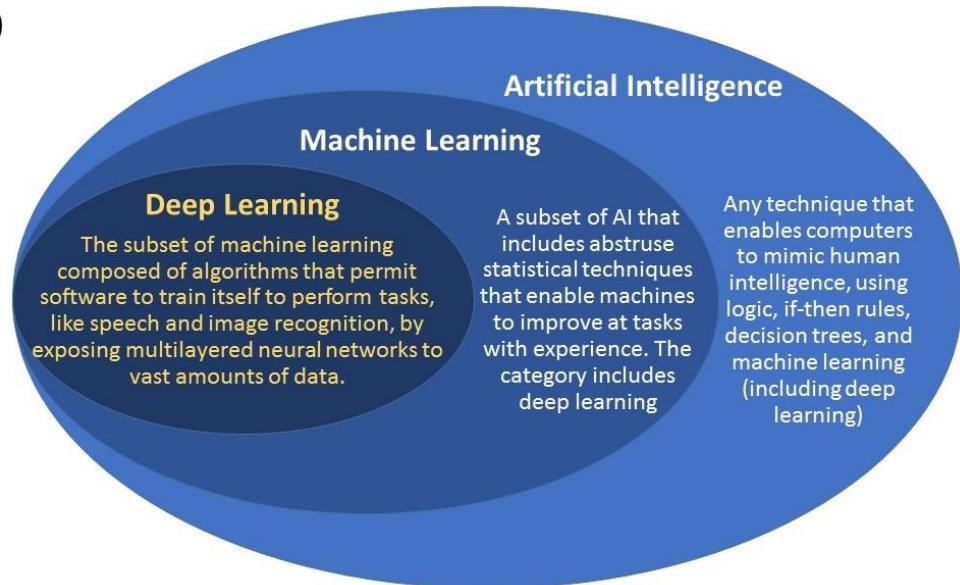
PC: actualización al **detailed (PC)**.

Pub(2023) (?)

Inteligencia Artificial

¿ Por qué?

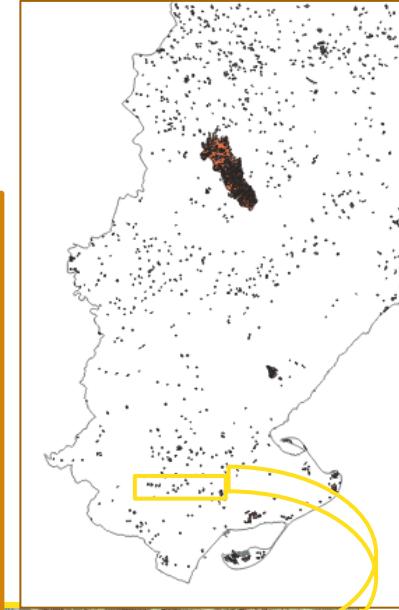
- ✓ **Necesidad de automatizar procesos**
- ✓ Entornos de computación (propios y nube)
- ✓ Volumen de datos
- ✓ Requerimientos latencia/reactividad



Detección Cambios 2018-19

Marcas > 2 ha :

1. Sentinel-2 (todo el territorio) identificadas automáticamente
2. *Deep Learning* (ortos 2018-2019) para zonas urbanas
3. Diferencias SIGPAC entre el 2018 y 2019, mayores de 2 ha.
4. Actualización de marcas BT5M
5. Actualización de marcas de Incendios
6. Zonas en construcción (ZC) i zonas en transformación (ZT), mayores de 2 ha, actualizadas en la versión anterior.



Algunas marcas:

- ◀ Sentinel 2
- ▶ SIGPAC
- ↙ Deep Learning



Marcas DL 2009-19

Zonas urbanas

Zonas agrícolas y forestales, a estudiar

Máscara MCSC 2009
(orto 2009)



Marcas de cambio DL
(orto 2009)



Marcas de cambio DL
(orto 2019)



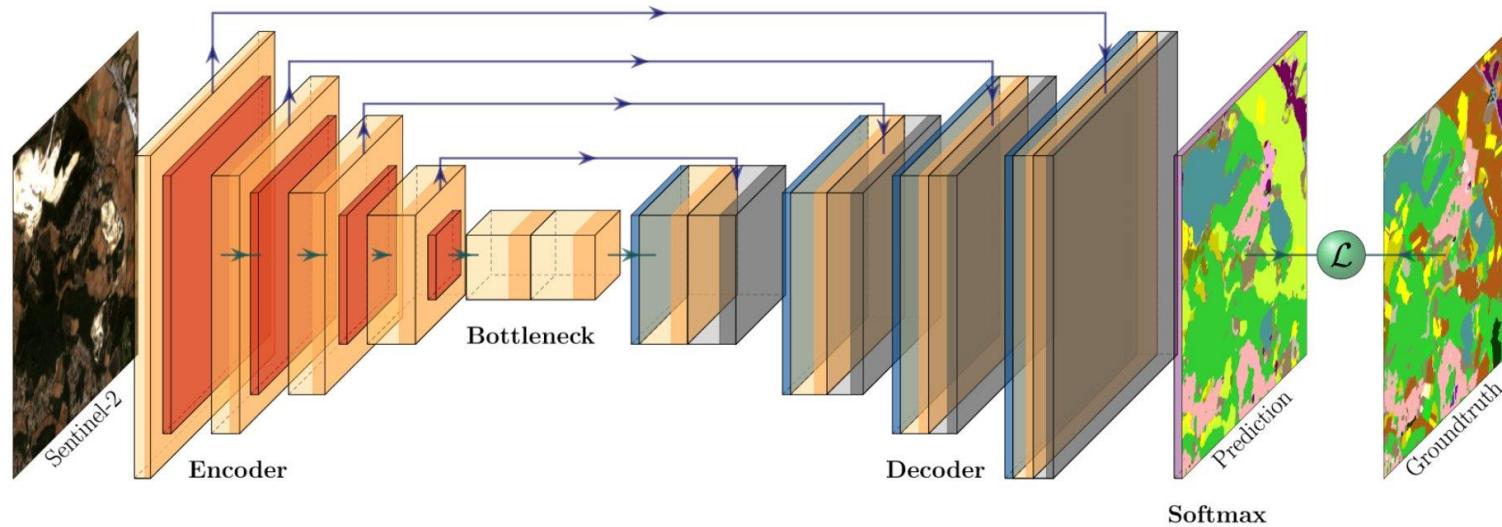
Azul: Detectado en el DL pero no consta en el MCSC 2009

Verde: Consta en el MCSC 2019 pero no detectado en el DL

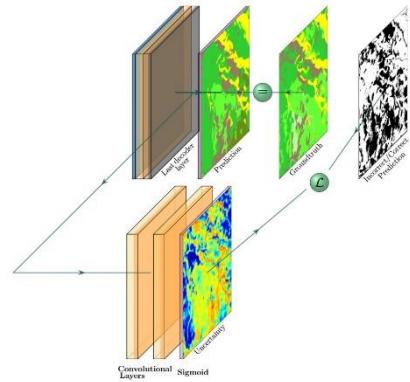
La suma de las dos marcas son los cambios detectados a partir de técnicas DL

Deep Learning + human-in-the-loop

Arquitectura U-Net

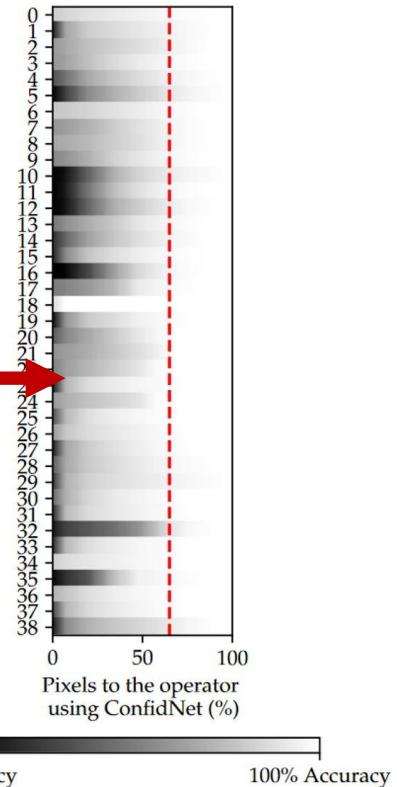


Deep Learning + human-in-the-loop



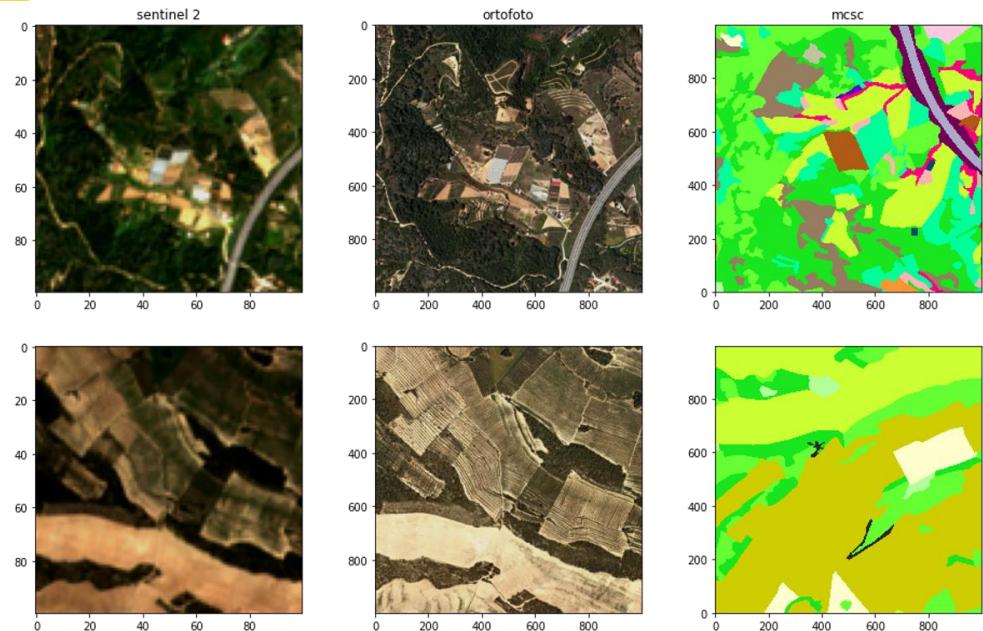
**Estimación de 90%
de exactitud global
enviando el 50% de
etiquetas a
fotointérpretes**

- Además de clasificar, queremos saber cómo de segura está la red de su predicción.
- Entrenamiento, en paralelo, de una red para aprender cómo de segura está la red original.
- Dos salidas: una predicción para cada píxel y otra con qué certeza se ha realizado.



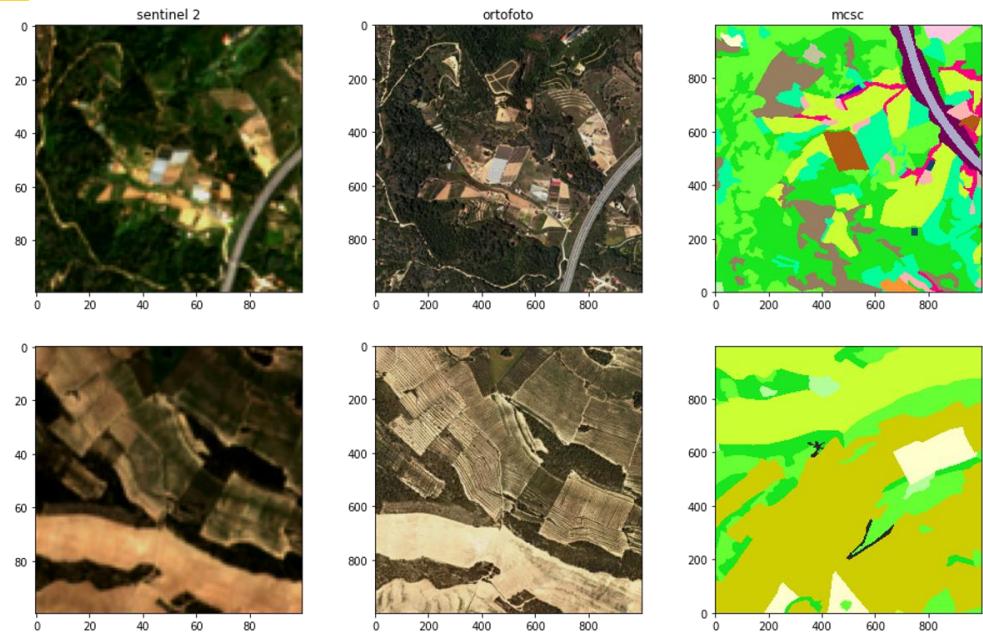
DL multiresolución

- Creación red neuronal con 2 entradas: Sentinel-2 y ortofoto, ambas con una buena resolución espacial.
- *Next Step 1:* combinación con CHM, DEM y Sentinel-1.
- *Next Step 2:* combinación definitiva con LiDAR.
- ???



DL multiresolución

- Creación red neuronal con 2 entradas: Sentinel-2 y ortofoto, ambas con una buena resolución espacial.
- *Next Step 1:* combinación con CHM, DEM y Sentinel-1.
- *Next Step 2:* combinación definitiva con LiDAR.
- ???



Doctorado Industrial 2018-DI-0045

CatLC (2018) DL multiresolución

Ortofoto RGB

GSD 1 m

252.3 GB

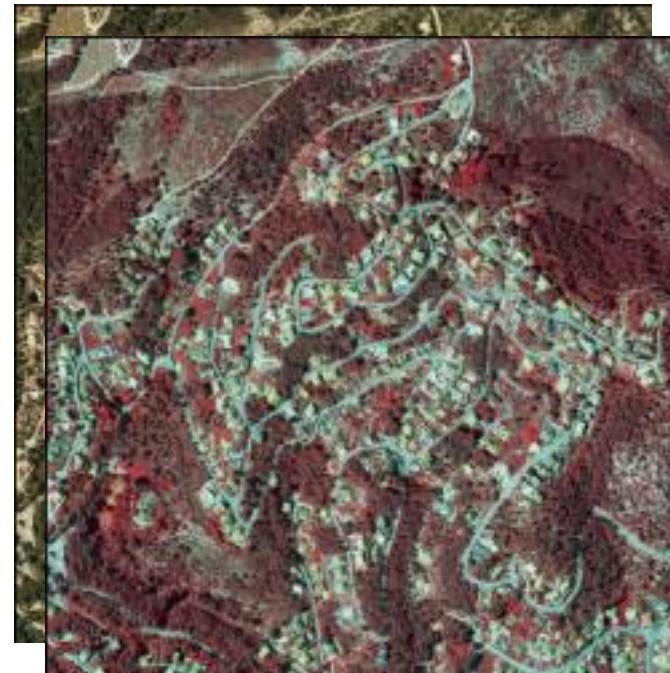


CatLC (2018) DL multiresolución

Ortofoto IRC

GSD 1 m

84.1 GB



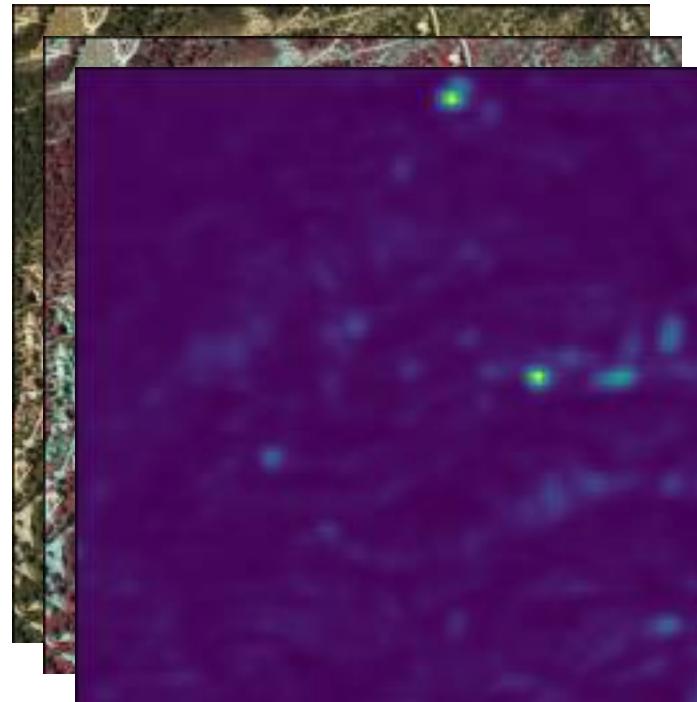
CatLC (2018) DL multiresolución

Sentinel-1

GSD 10 m

3.4 GB

Enero-Diciembre + anual



CatLC (2018) DL multiresolución

Sentinel-2

GSD 10 m

16.8 GB

Abril + Agosto

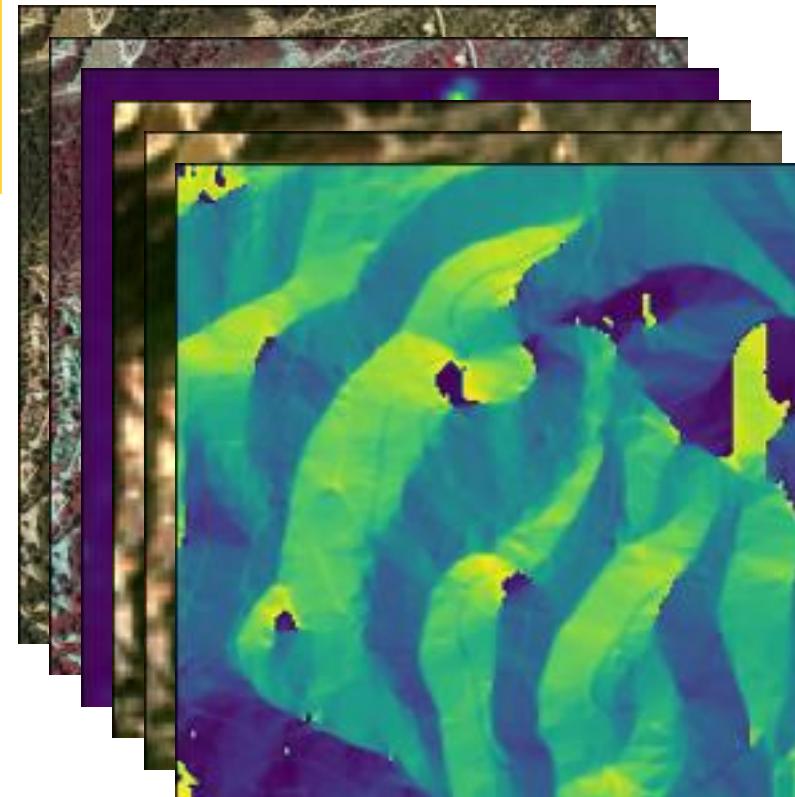


CatLC (2018) DL multiresolución

Aspect

GSD 5 m

13.4 GB



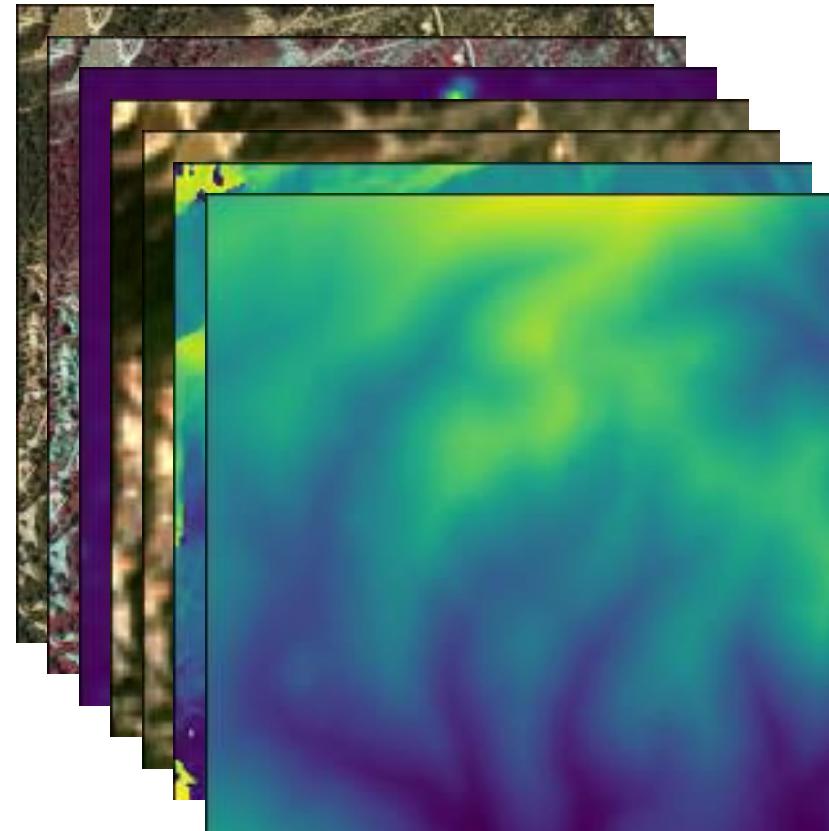
CatLC (2018) DL multiresolución

DEM

Digital Elevation Model

GSD 1 m

13.4 GB

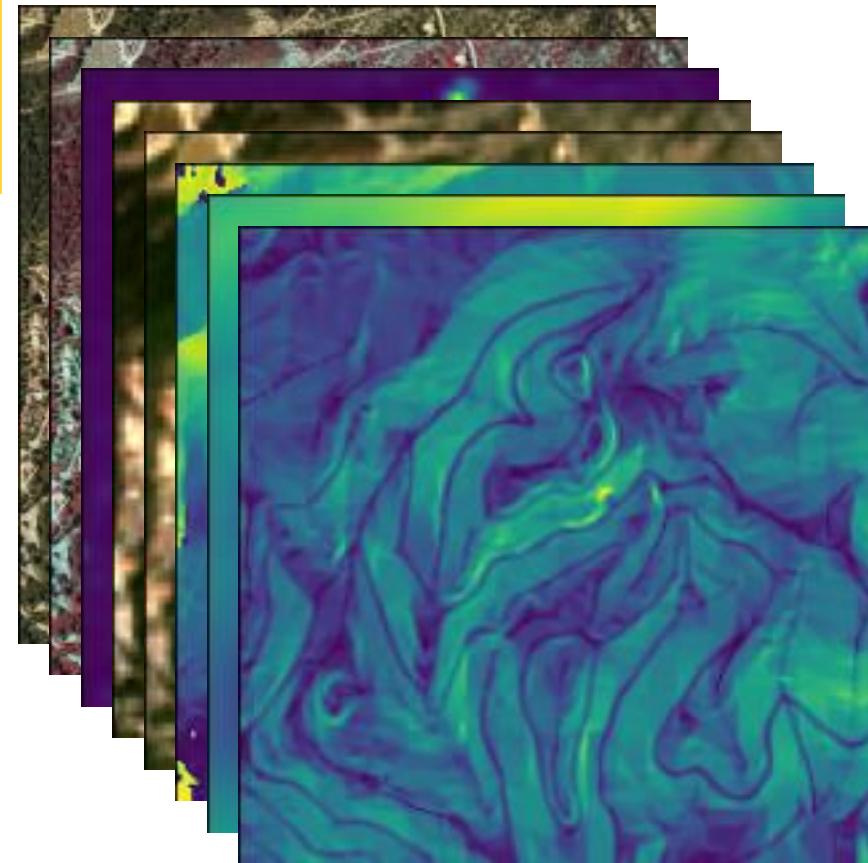


CatLC (2018) DL multiresolución

Slope

GSD 5 m

13.4 GB



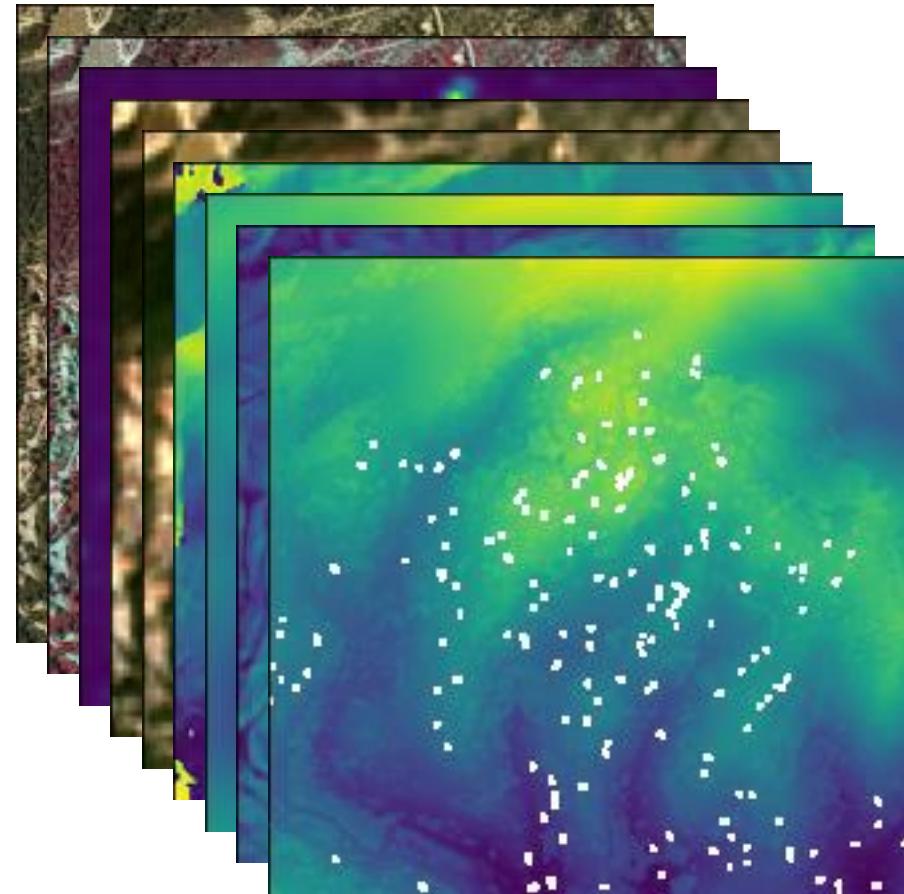
CatLC (2018) DL multiresolución

DSM

Digital Surface Model

GSD 1 m

336.4 GB



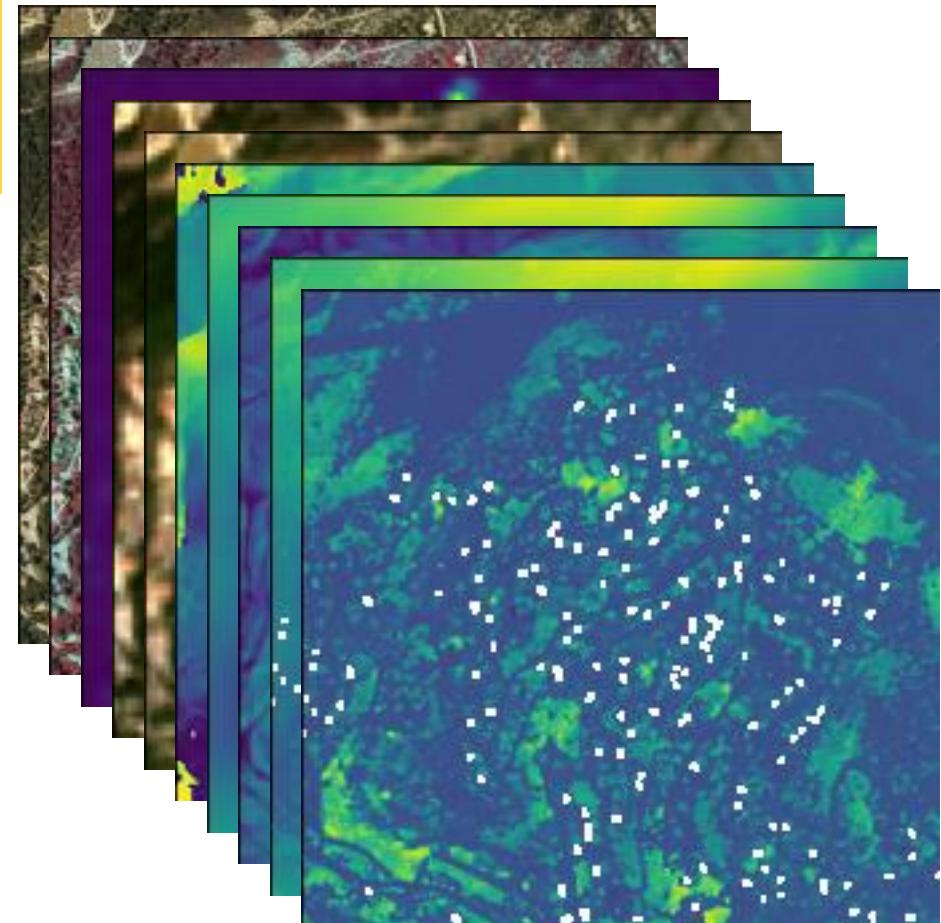
CatLC (2018) DL multiresolución

CHM

Canopy Height Model

GSD 1 m

336.4 GB



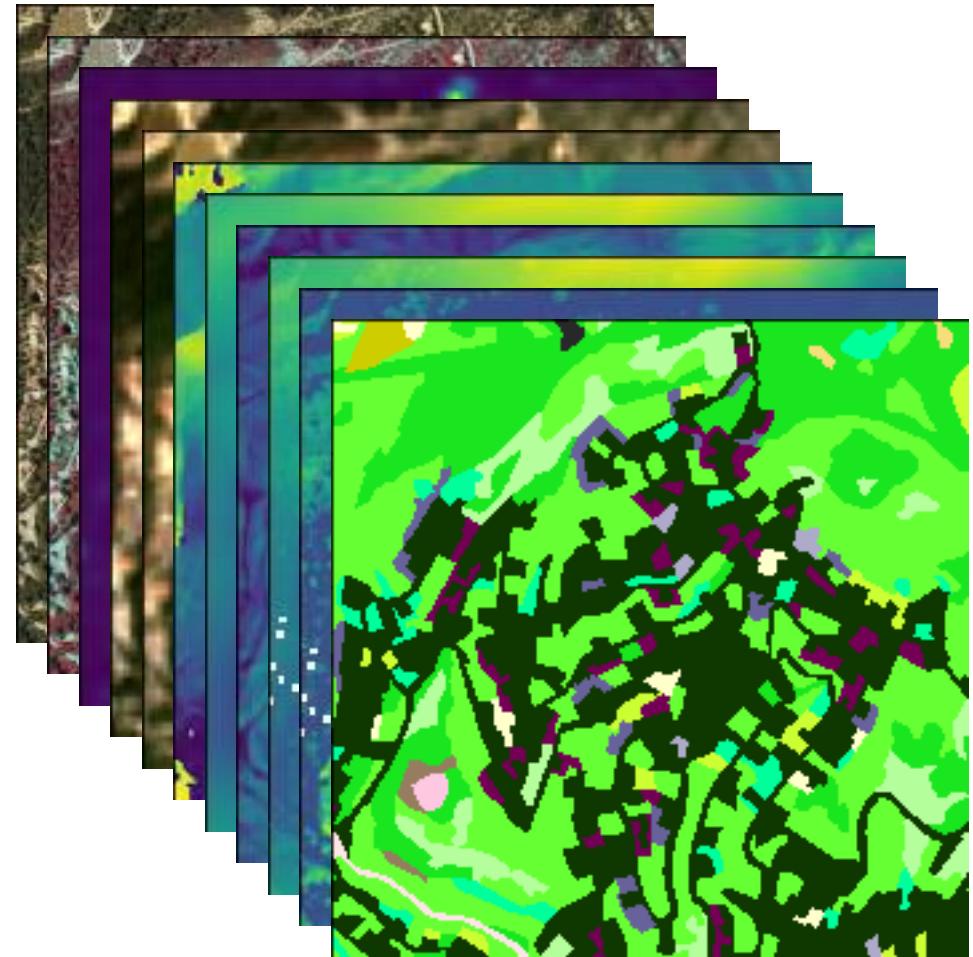
CatLC DL multiresolución

MCSC 2018

verdad (*ground-truth*)

GSD 1 m

84.1 GB

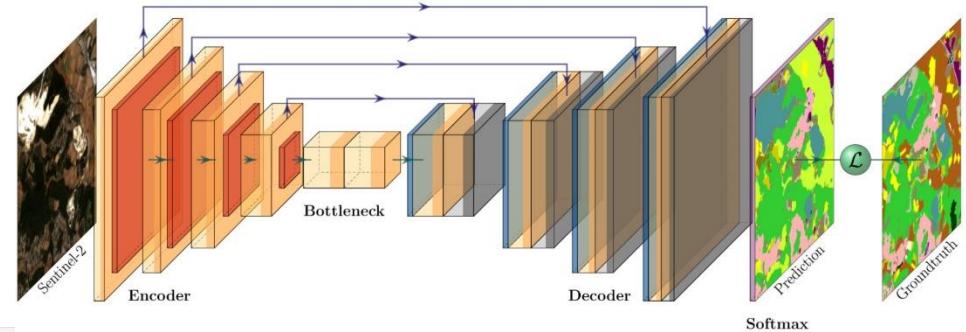


CatLC red neuronal U-Net

IMPORT

In [1]:

```
import rasterio
from rasterio.windows import Window
from torch.utils.data import Dataset, DataLoader
from torchvision import transforms, utils
import torch
from torch import optim
import torch.nn.functional as F
import torch.nn as nn
import numpy as np
from unet_parts import *
from datetime import datetime
from tqdm import tqdm
import segmentation_models_pytorch as smp
from torch.utils.tensorboard import SummaryWriter
import matplotlib.pyplot as plt
import pandas as pd
from utils import metrics
import matplotlib.patches as mpatches
from torchmetrics import MetricCollection, Accuracy, Precision, Recall, IoU, ConfusionMatrix
from skimage.exposure import rescale_intensity
import pickle
import tikzplotlib
```



unet_catlc.ipynb

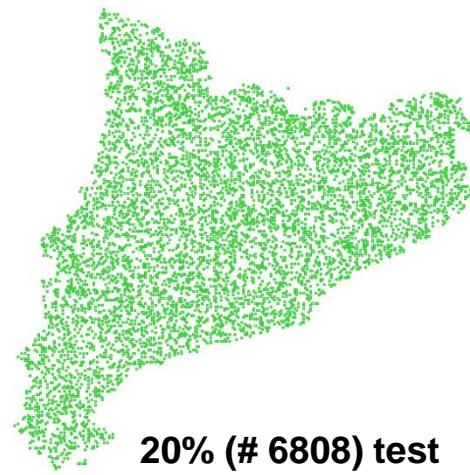
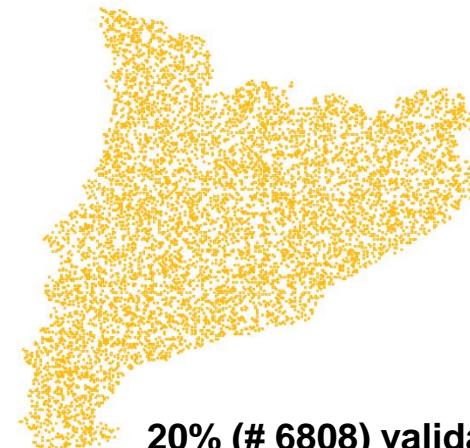
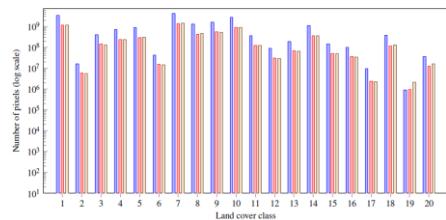
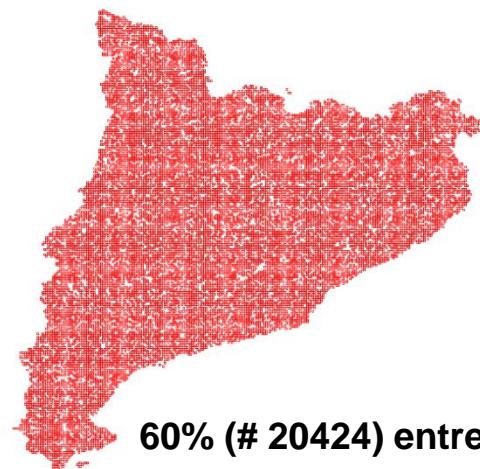
github.com/OpenICGC/CatLC

PyTorch

CatLC *training data*

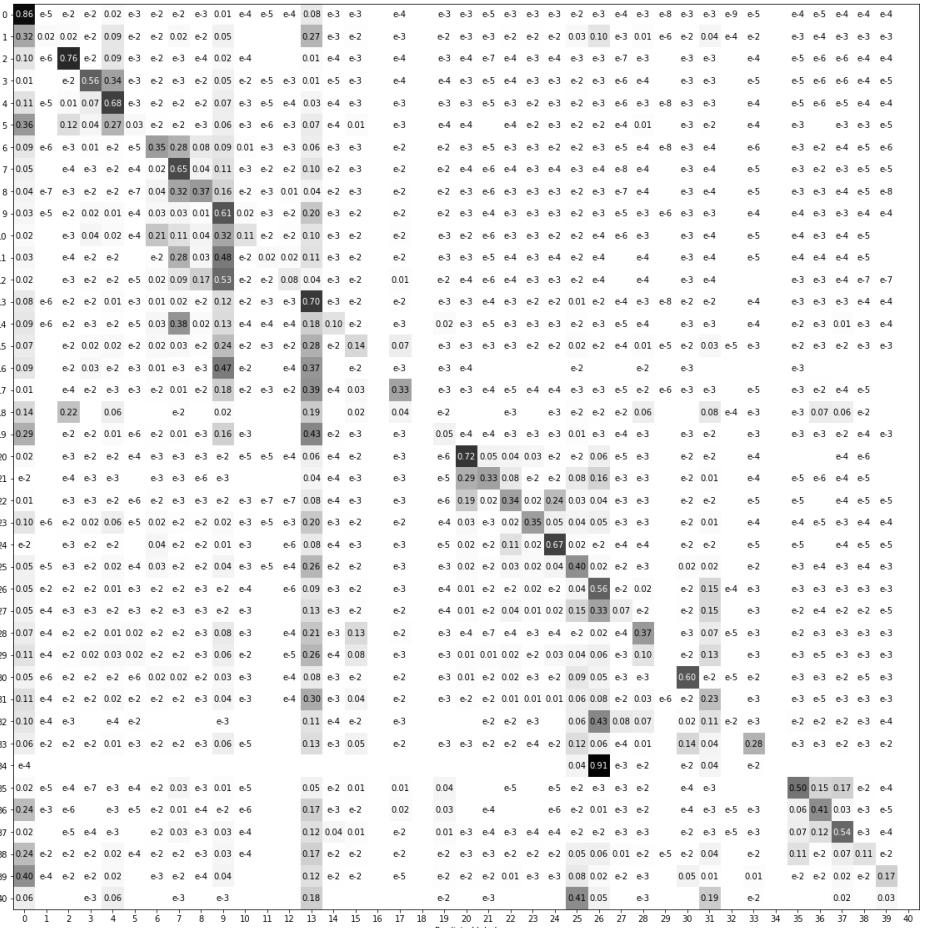
41 categorías

Agricultural area	Urban area	Water bodies
01 Herbaceous crops	14 Meadows and grasslands	29 Sports and leisure areas
02 Orchard, plant nurseries...	15 Shore forest	30 Mining or landfills
03 Vineyards	16 Bare forest soil	31 Areas in transformation
04 Olive groves	17 Burned areas	32 Road network
05 Other woody crops	18 Rocky	33 Urban bare ground
06 Crops in transformation	19 Beaches	34 Beaches
Forest area	20 Wetlands	35 Airport areas
07 Dense coniferous forests	21 Urban area	36 Railway network
08 Dense deciduous forests	22 "Eixample"	37 Port areas
09 Dense forests of sclerophylls	23 Lax Urban Areas	38 Reservoir
10 Scrub	24 Isolated buildings	39 Lakes and lagoons
11 Clear coniferous forests	25 Isolated residential areas	40 Watercourses
12 Clear deciduous forests	26 Green areas	41 Rafts
13 Clear forests of sclerophylls	27 Industrial or commercial	42 Artificial channels



CatLC - escenarios

1. ortoRGB + ortoIRC



CatLC - escenarios

1. ortoRGB + ortoIRC
2. **Sentinel-2 (Ab. + Ag)**

CatLC - escenarios

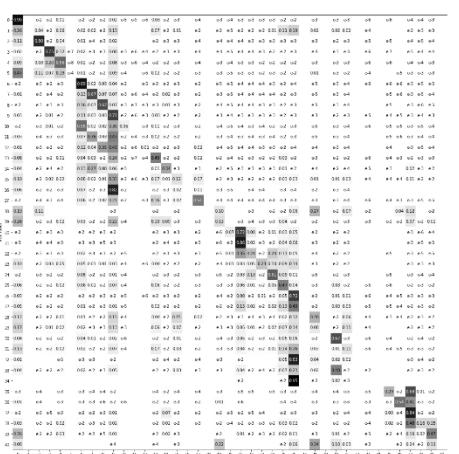
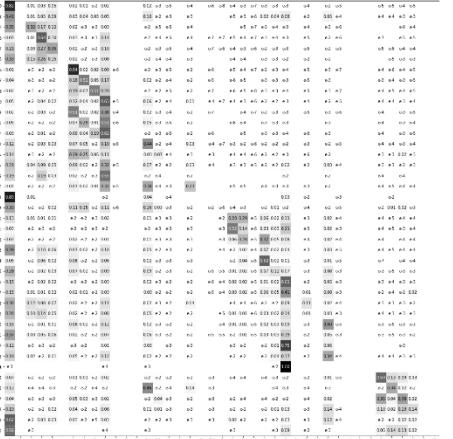
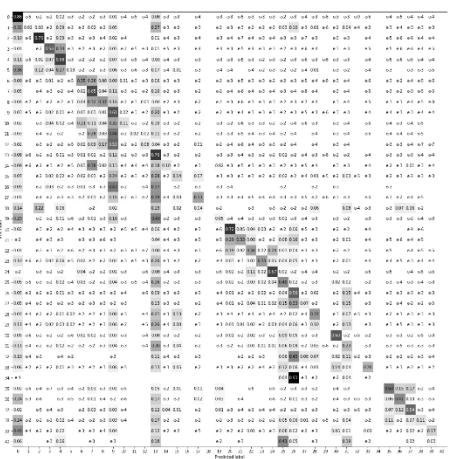
1. ortoRGB + ortoIRC
 2. Sentinel-2 (Ab. + Ag.)
 3. CatLC dataset

CatLC - escenarios

1. ortoRGB + ortoIRC

2. Sentinel-2 (Ab. + Ag.)

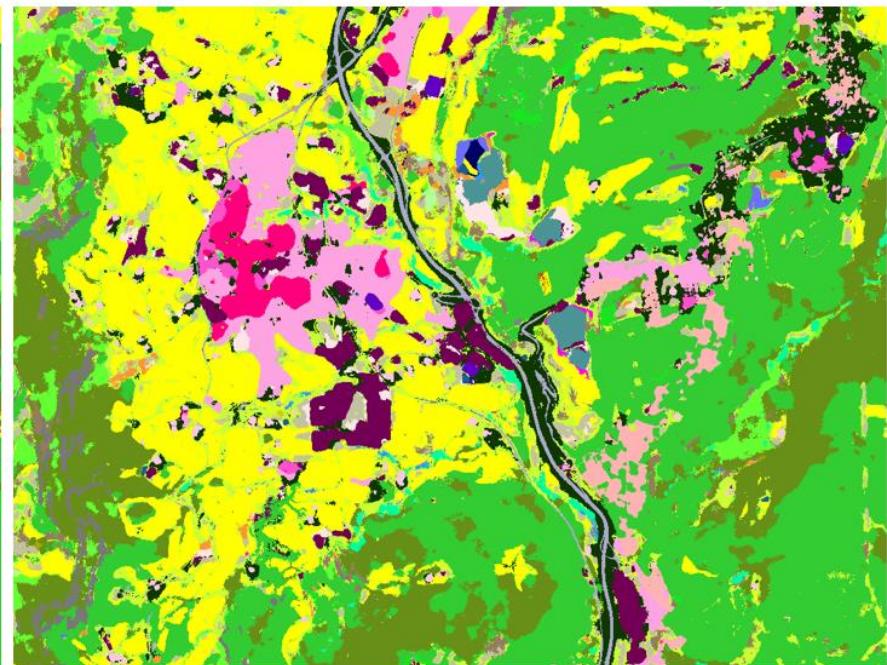
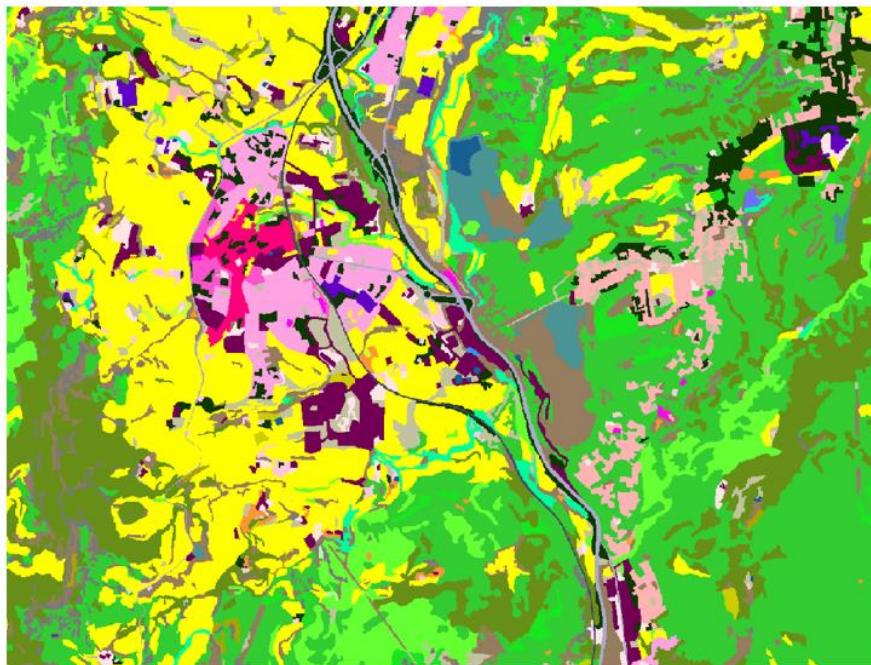
3. CatLC dataset



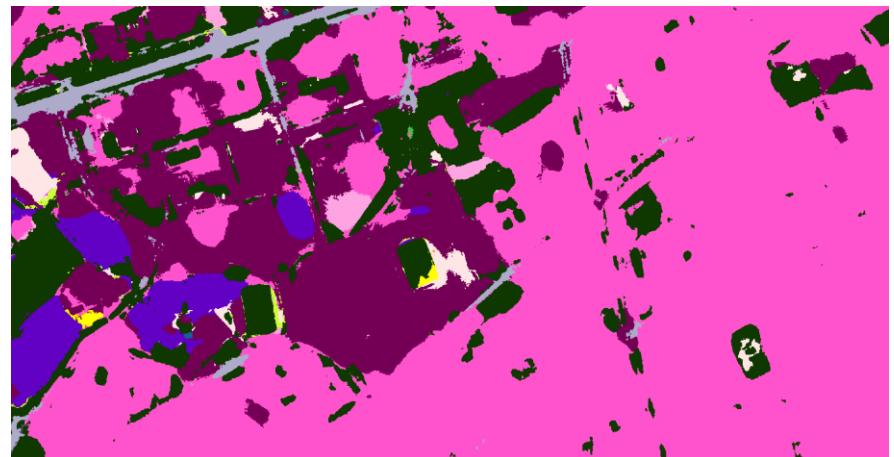
MCSC vs DL



MCSC vs DL



MCSC vs DL



MCSC vs DL



MCSC vs DL

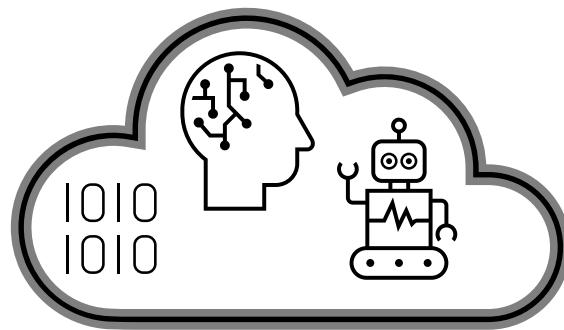


CatLC - website



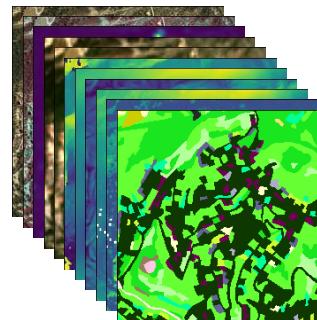
- ✓ <https://www.icgc.cat/en/Downloads/Maps-in-image-format/Catalonia-Multi-resolution-Landcover-Dataset-CatLC>
- ✓ Publicado en Nature Scientific Data.
- ✓ Datos Abiertos.

CatLC - usuarios



AI community

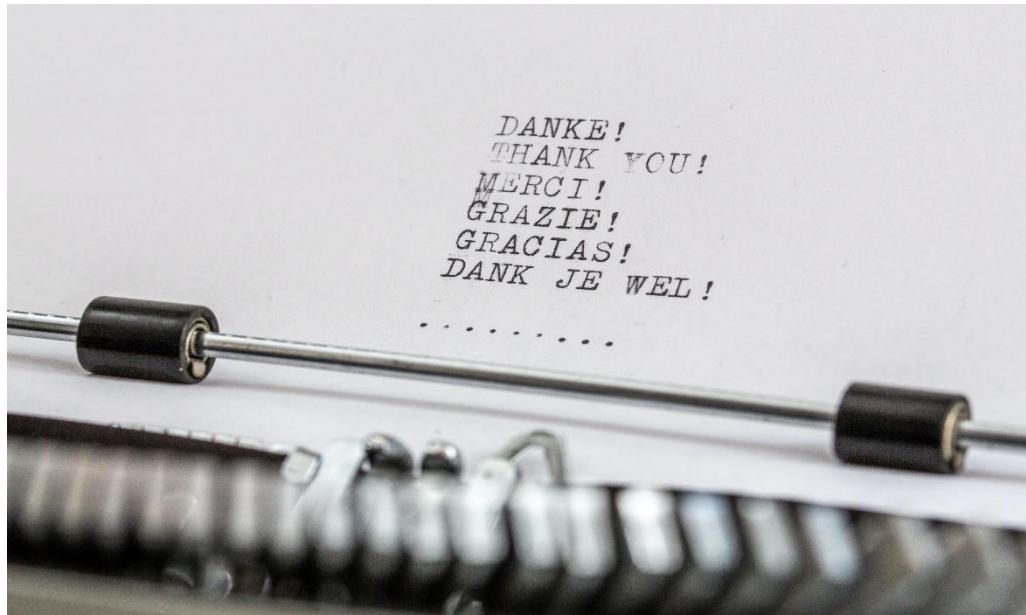
EO dataset – high quality



EO community

AI code sample





catlc@icgc.cat

Institut Cartogràfic i Geològic de Catalunya

Parc de Montjuïc,
E-08038 Barcelona

41°22'12" N, 2°09'20" E (ETRS89)

-  www.icgc.cat
-  icgc@icgc.cat
-  twitter.com/ICGCat
-  facebook.com/ICGCat

Tel. (+34) 93 567 15 00
Fax (+34) 93 567 15 67