INTERNATIONAL CONFERENCE OF THE PCC - EXPERIENCE IN 3D, 4D AND AI IN CADASTRE

Using AI to improve data quality and citizen interaction

(May 2025)

Miguel Ancochea Nodal
miguel.ancochea@catastro.hacienda.gob.es

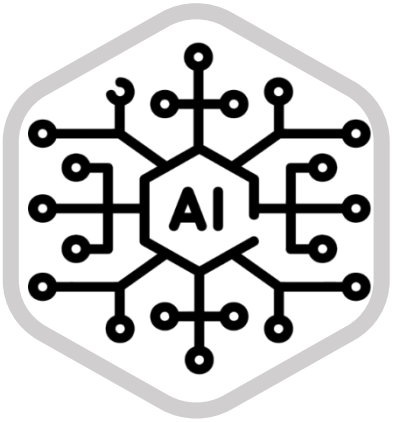
DG del Catastro
www.catastro.hacienda.gob.es

INDEX

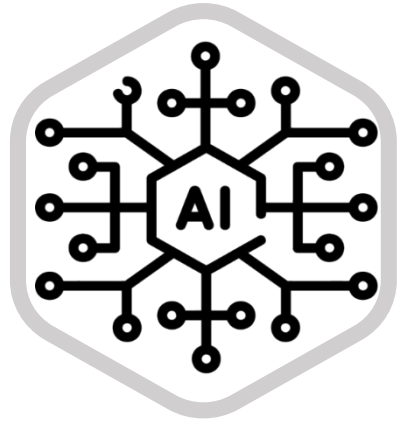
GENERAL CONSIDERATIONS

TYPOLOGIES

- A distinction must be made between at least two generations of AI.
- **«Traditional» AI vs "Generative" AI LLM**
- Traditional IAs
  - Neural networks, expert systems, chatbot, etc.
  - Well known scenario. Already regulated usage.
  - More 'visual' or 'geometric' uses of AI, which have a more intuitive application in the cadastral business compared with "more linguistic" ones.
  - We already have use cases implemented.



GENERAL CONSIDERATIONS



- "Generative" AI LLM
  - Regulatory standards in process - national, tax ministry and internal (Cadastre)
  - Principles: supervised, multidisciplinary governance, respect of protected data, internal use for decision making e.g. not for administrative resolutions, integrated in software applications.
  - Specific use cases (only prospective projects):

    - **Extraction of documentary information**, from documents in a dossier. E.g. summaries, analysis etc.

    - **Support for valuation**: Extraction of relevant information from façade photos, floor plans, cartography etc.

CITIZEN INTERACTION: CATIA

PURPOSE AND SCOPE OF THE TOOL

- Purpose: Improvement of citizen service.
- Interactive help (chatbot) in the Cadastral Website (SEC) through the use of specialized AI technologies (limited).
- Low-cost solution using open source SW (RASA) and in-house development.

CATIA
Asistente Virtual del Catastro

CITIZEN INTERACTION: CATIA

PURPOSE AND SCOPE OF THE TOOL

- CATIA guides the citizen by redirecting him to the already existing features within the Cadastral website (SEC).
- It is NOT a generative AI and it is NOT autonomous in learning, it is supervised.

CATIA
Asistente Virtual del Catastro

CITIZEN INTERACTION: CATIA

PURPOSE AND SCOPE OF THE TOOL

- Interactive auxiliary tool within an “omnichannel” service approach.
- Complementary to other attention and communication channels with citizens, such as the Cadastre Call Center and in-person.
- Project included in the framework of “*Measure 2 of Public Administrations' Digitalisation Plan - Improving citizens´ user experience*”.
- *“The aim of this measure is to offer a personalized, proactive and omnichannel citizen service model, through a single point of access both for obtaining information and to carry out procedures with the Public Administration. One of its main lines of action is the implementation of chatbots or intelligent virtual assistants for the highest impact use cases..”*

CATIA
Asistente Virtual del Catastro

CITIZEN INTERACTION: CATIA

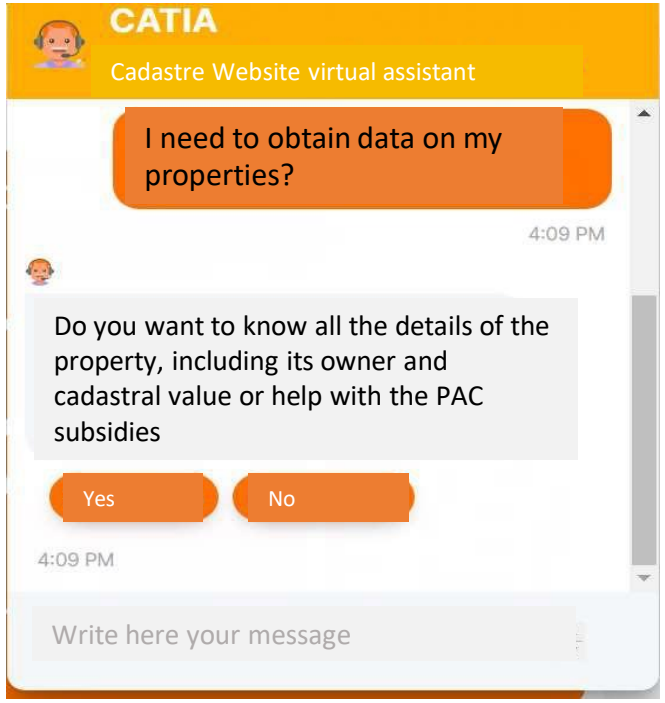
STRUCTURE

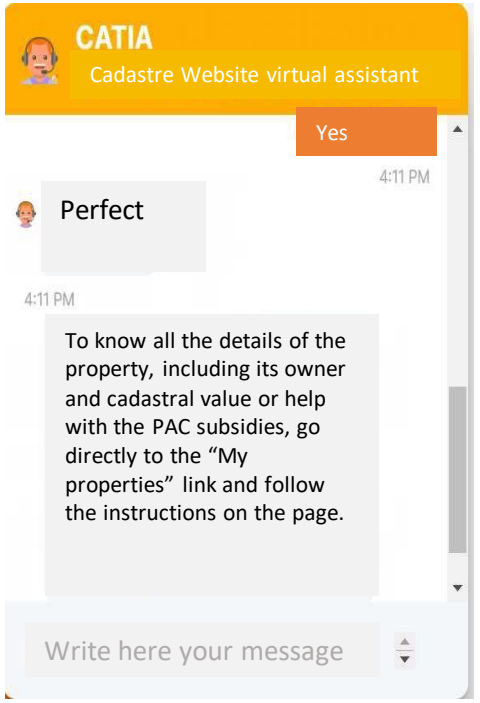
- Structured by topics:
  - *My Properties, Search, My Files, Notifications, My Procedures, Reference value, CSV Matching, Heirs, Suggestions*.
  - These topics are fed with questions, synonyms, keywords, etc.

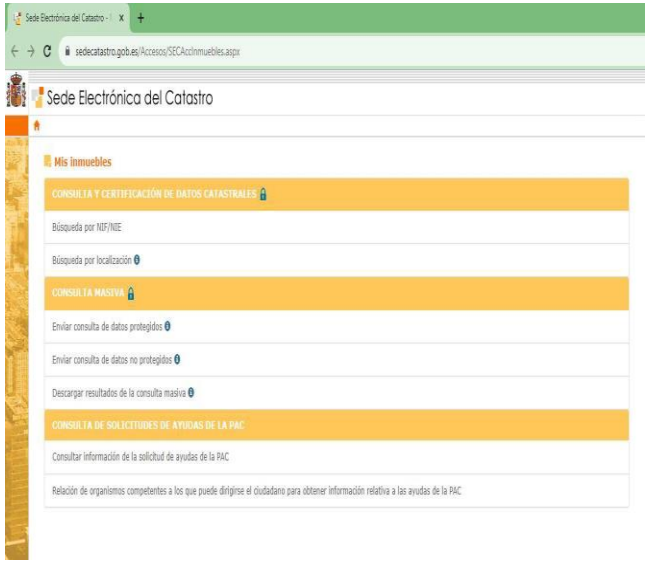
CITIZEN INTERACTION: CATIA

OPERATION

MY PROPERTIES







CATIA
Asistente Virtual del Catastro

DATA QUALITY: FACADE PHOTOS

The publication of facade photographs on the internet means special care must be taken to censor recognizable people's faces or car license plates.

This control can no longer be entrusted to individuals alone. It must be included in the automatic validations carried out by the management software (SIGCA).

- Select the basic software to be used (pre-trained models), test as prototypes.
- Creation of the basic features of face and plate number detection.
- Integration of these features within the software flow.

DATA QUALITY: FACADE PHOTOS

FACE DETECTION

DNN (Deep neural network) Prototype:

It is the most reliable and already recognized complex shapes. It uses neural network technology to view faces recognized by patterns (face, ears, eyelashes, moustaches, mouths...) and some parameters were added to improve its localization, such as switching images to black and white.

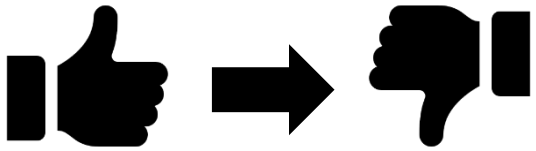
Positive correct with DNN prototype.

DATA QUALITY: FACADE PHOTOS

FACE DETECTION

The prototype chosen is not infallible, it requires parameterization.

A filter is made in case the faces which are detected are small or doubtful, so a minimum size, and a confidence limit of 97% is applied to confirm whether it is a face or not. This achieves 100% effectiveness. Any face detected is a face.



There are 2 faces and a lantern. This is a false positive. We remove the lantern, and we remove the faces due to their size.

DATA QUALITY: FACADE PHOTOS

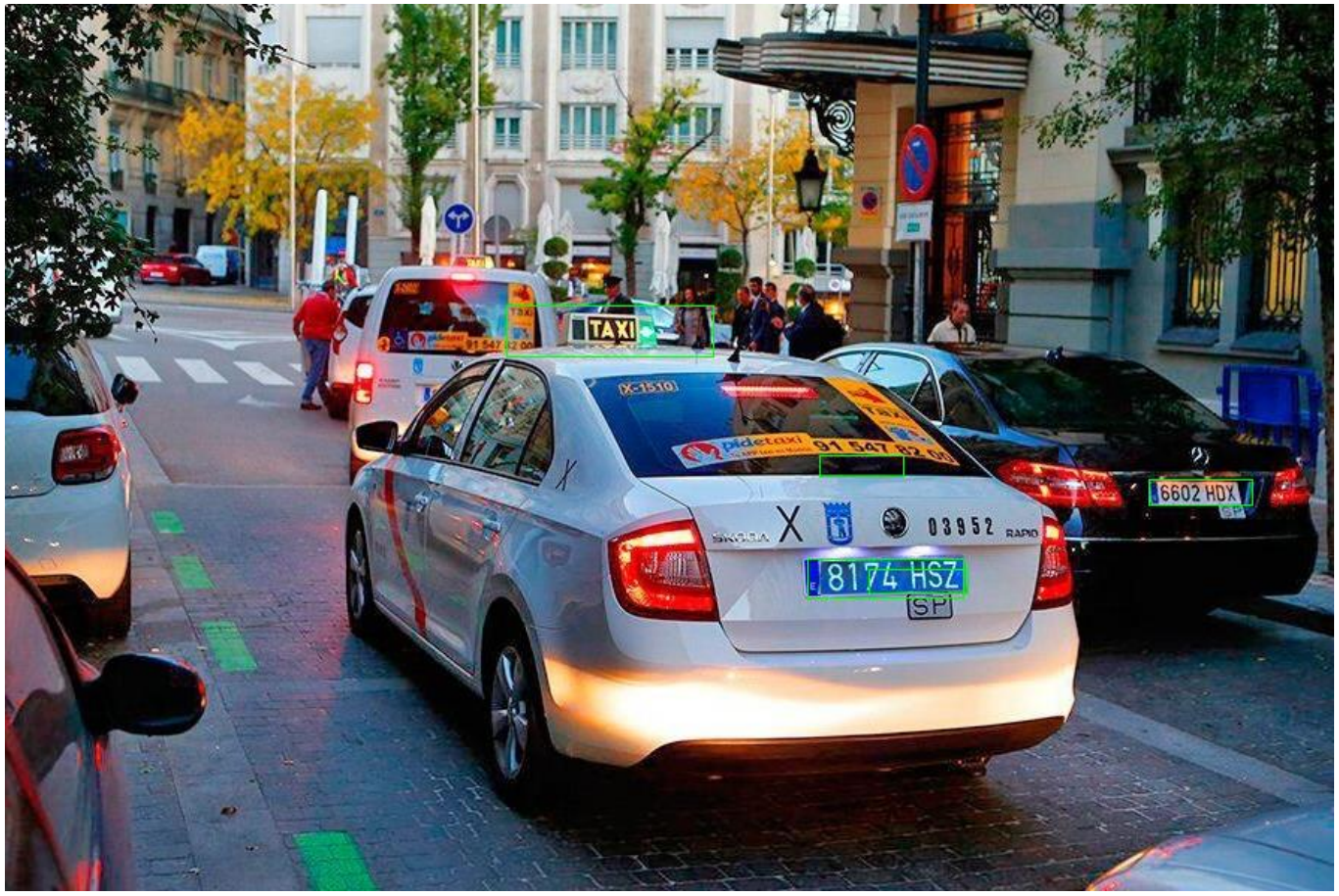
LICENSE PLATE DETECTION

Cascade Prototype:

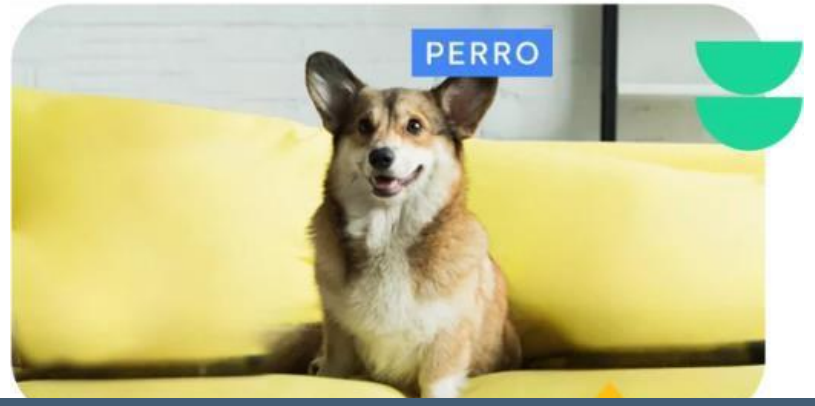
The one selected. It uses the *Cascade Classifier* recognition system. This system performed better at recognizing the plates and not random text in the image.

Refining the method of reading the letters in the image was also required. For this purpose we used *Tesseract*.

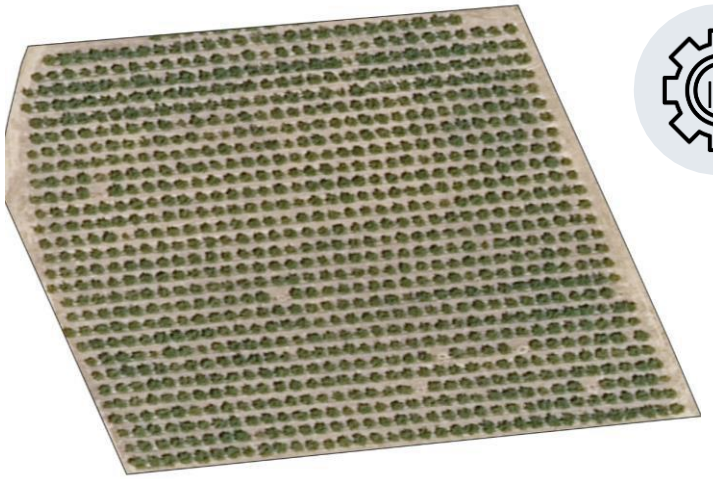
Also applied changes to the image, such as black and white, contrast and noise removal. We also rotated and inverted it in specific cases, as when it is a taxi or the number plate is rotated.

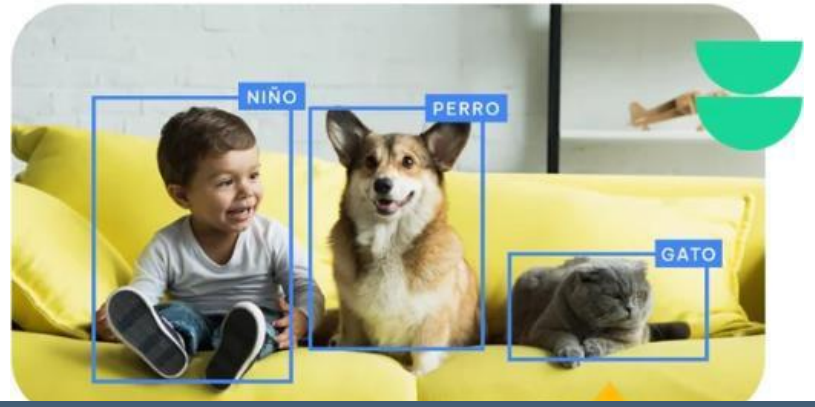


DATA QUALITY: BUILDING DETECTION AND CROP CLASSIFICATION



CLASSIFICATION

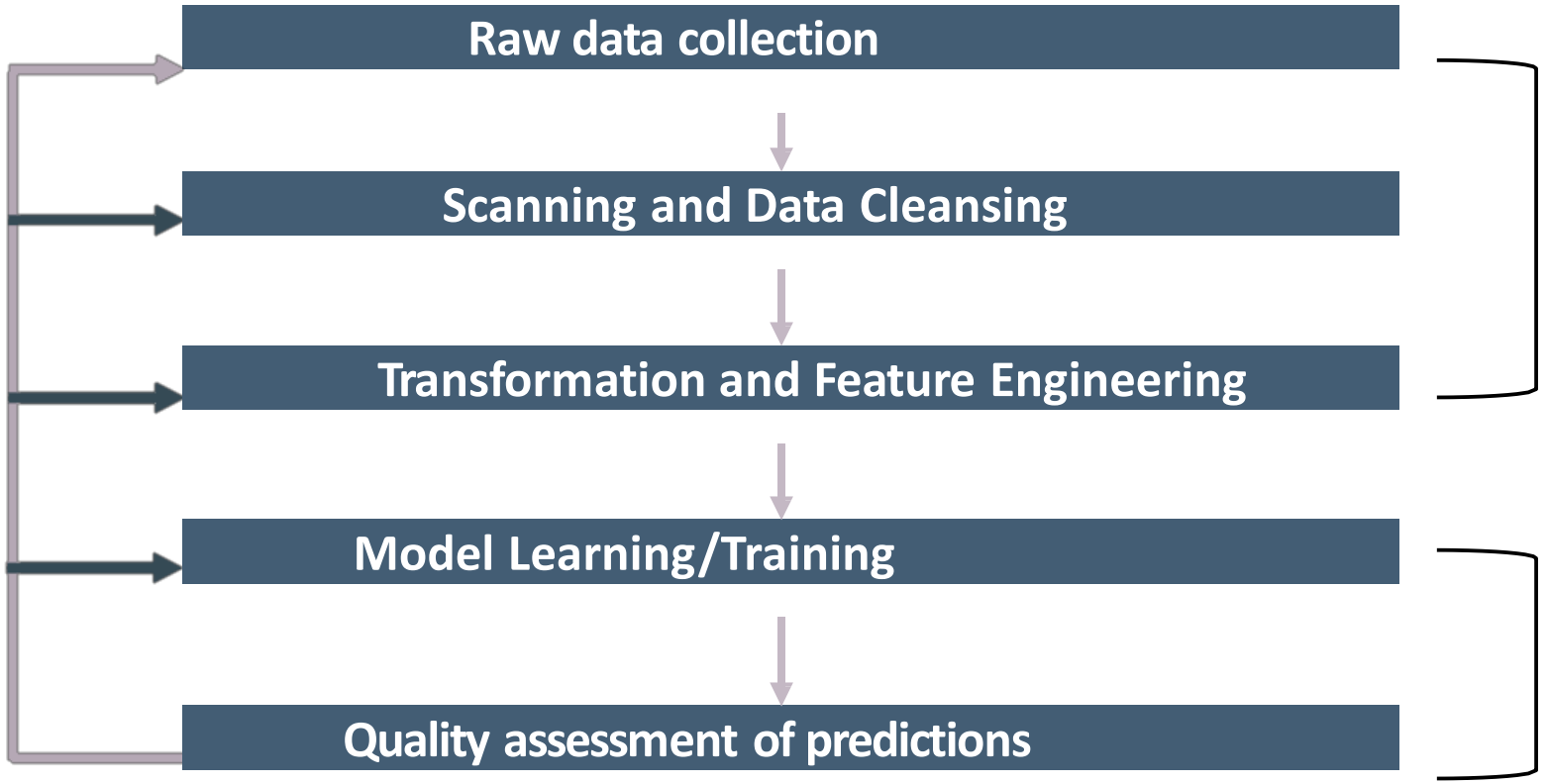




OBJECT DETECTION

DATA QUALITY: BUILDING DETECTION AND CROP CLASSIFICATION

PHASES IN PROJECT



90%

BUSINESS KNOWLEDGE

10%

DATA QUALITY: BUILDING DETECTION AND CROP CLASSIFICATION

2. BUILDING DETECTION

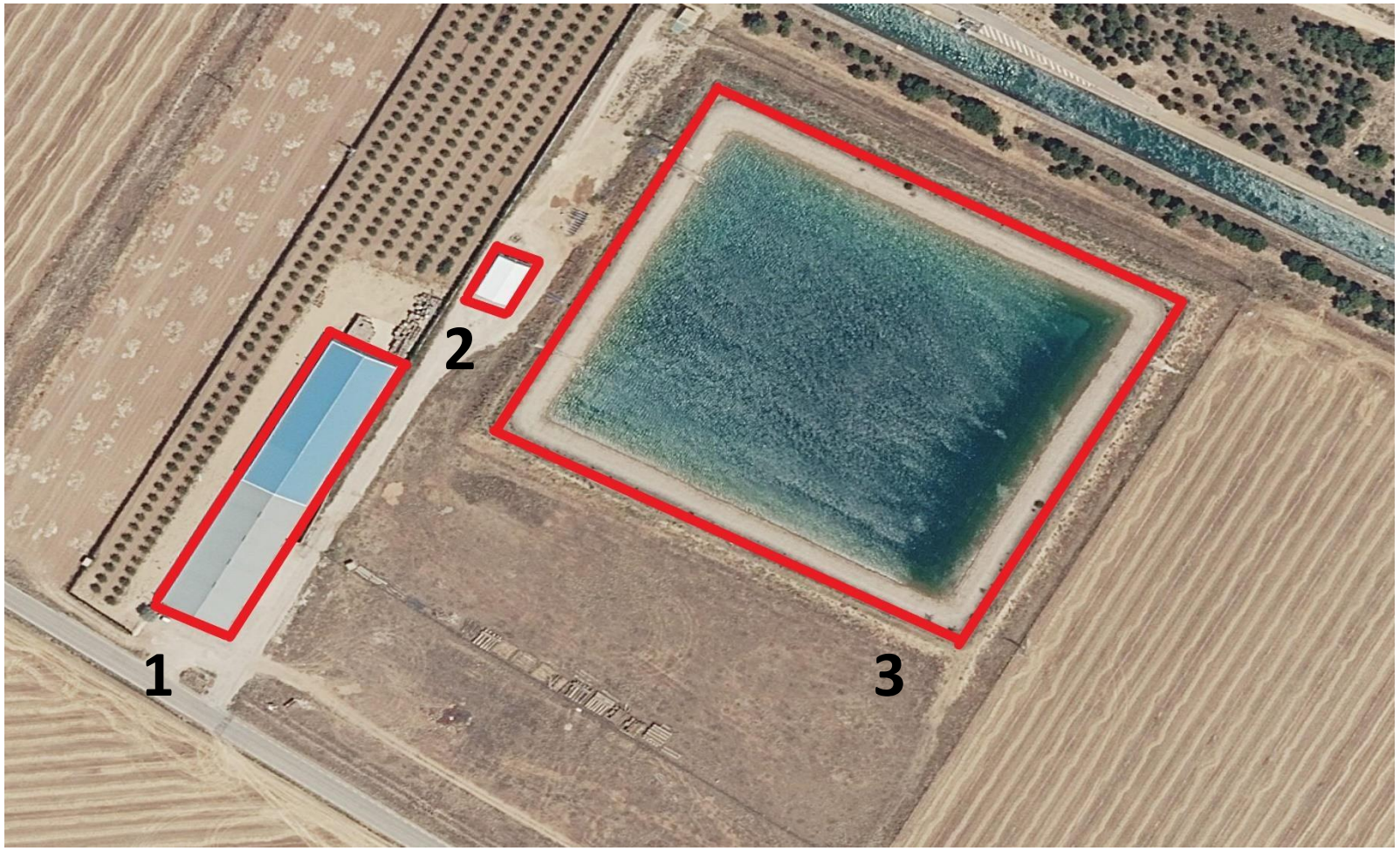
OBJECTIVE **DETECT CONSTRUCTION OMISSIONS**

INCLUDES

- LOCATION ON CARTOGRAPHY
- CONSTRUCTION TYPE

  (PHASE 1)
  - Single-family homes
  - Swimming pools
  - Agricultural buildings and warehouses
  - Greenhouses
  - Irrigation ponds

DATA QUALITY: BUILDING DETECTION AND CROP CLASSIFICATION

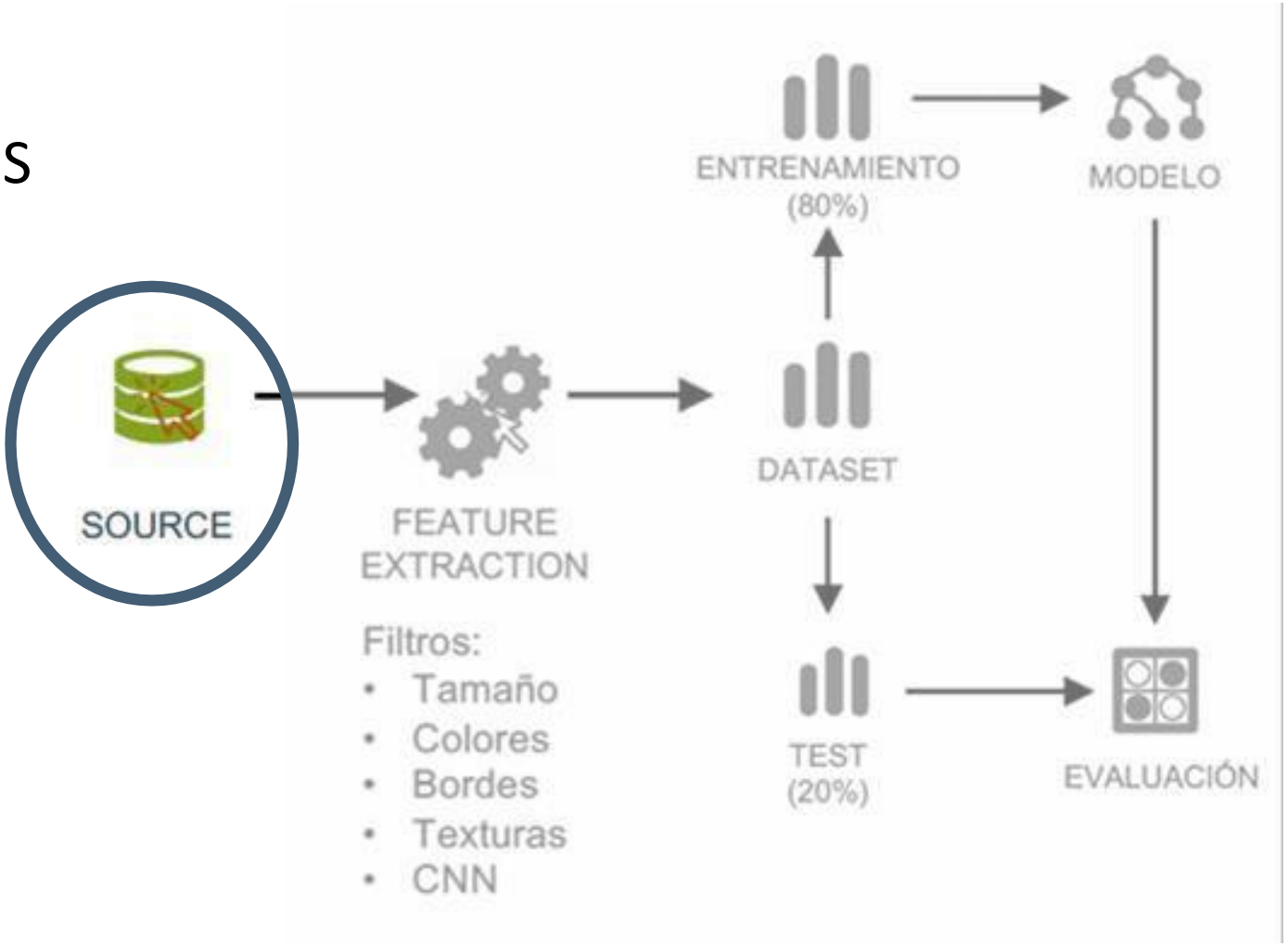


1. BUILDINGS
2. AGRICULTURAL
3. IRRIGATION POND

DATA QUALITY: BUILDING DETECTION AND CROP CLASSIFICATION

15.000 IMAGES

- LABELED CONSTRUCTIONS (SUPERVISED)
- PNOA ORTHOPHOTOS
- SCALE 1:1000
- A4 SIZE



DATA QUALITY: BUILDING DETECTION AND CROP CLASSIFICATION

IMAGE PROCESSING

- EDGE FILTER
- TEXTURE FILTER
- SIZE & COLOR

DATA QUALITY: BUILDING DETECTION AND CROP CLASSIFICATION

PRETRAINED CONVOLUTIONAL NEURAL NETWORKS (CNN)

- MOBILENETV2
- RESNET-18

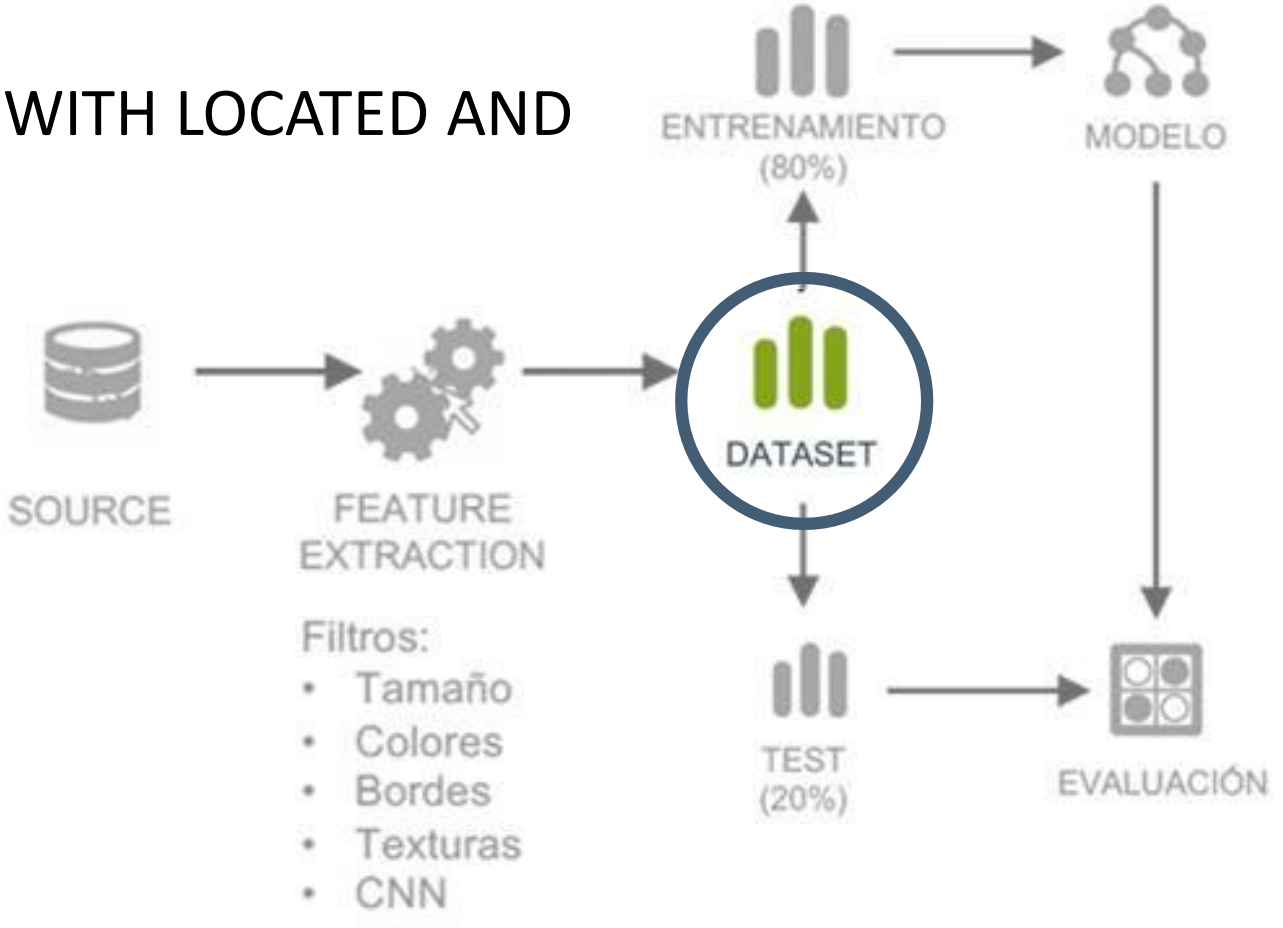
DATA QUALITY: BUILDING DETECTION AND CROP CLASSIFICATION

DATASET

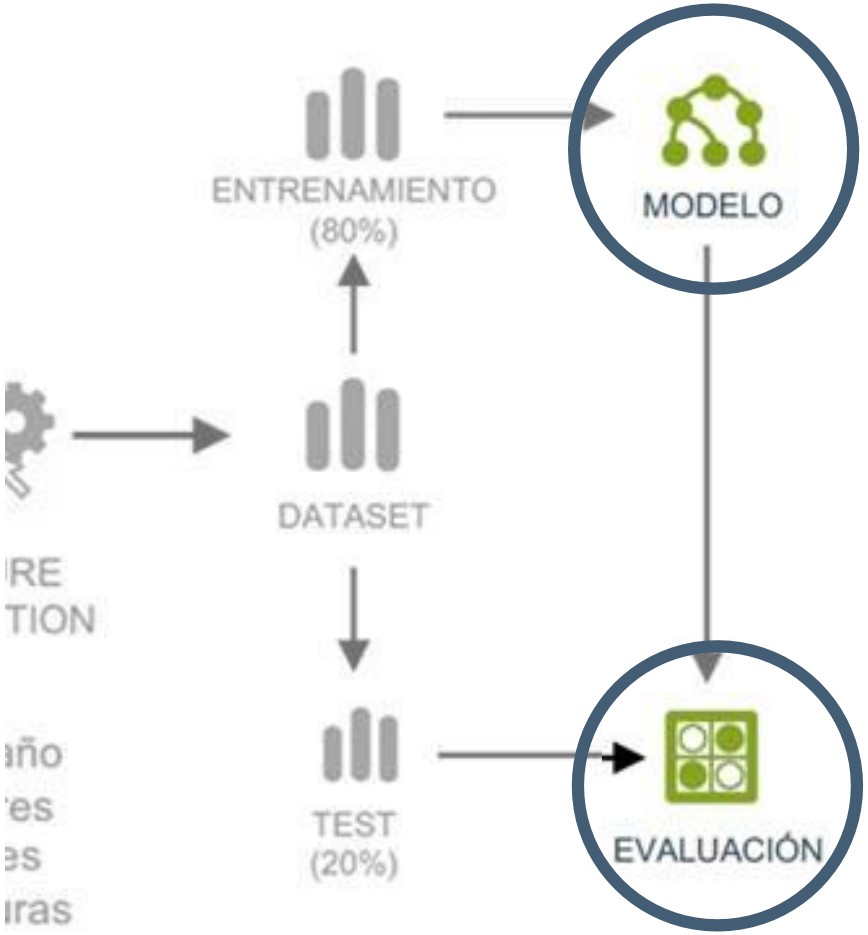
- 15,000 ORTHOPHOTOS WITH LOCATED AND IDENTIFIED BUILDINGS

  +

- INFORMATION EXTRACTED THROUGH FILTERS AND CNN



DATA QUALITY: BUILDING DETECTION AND CROP CLASSIFICATION



MODEL DEVELOPMENT

80% DATA SET

+

MODEL VALIDATION

20% DATA SET

EXTERNAL VALIDATION

DATA QUALITY: BUILDING DETECTION AND CROP CLASSIFICATION

3. CROP CLASSIFICATION

OBJECTIVE: **CADASTRAL QUALIFICATION ASSIGNMENT(CC)**

- 30 CADASTRAL QUALIFICATIONS
- AT SUBPLOT LEVEL
- ANALYSIS OF THE BEST PERFORMING MODEL:

  PARTICULAR MODEL **vs** GENERAL MODEL.

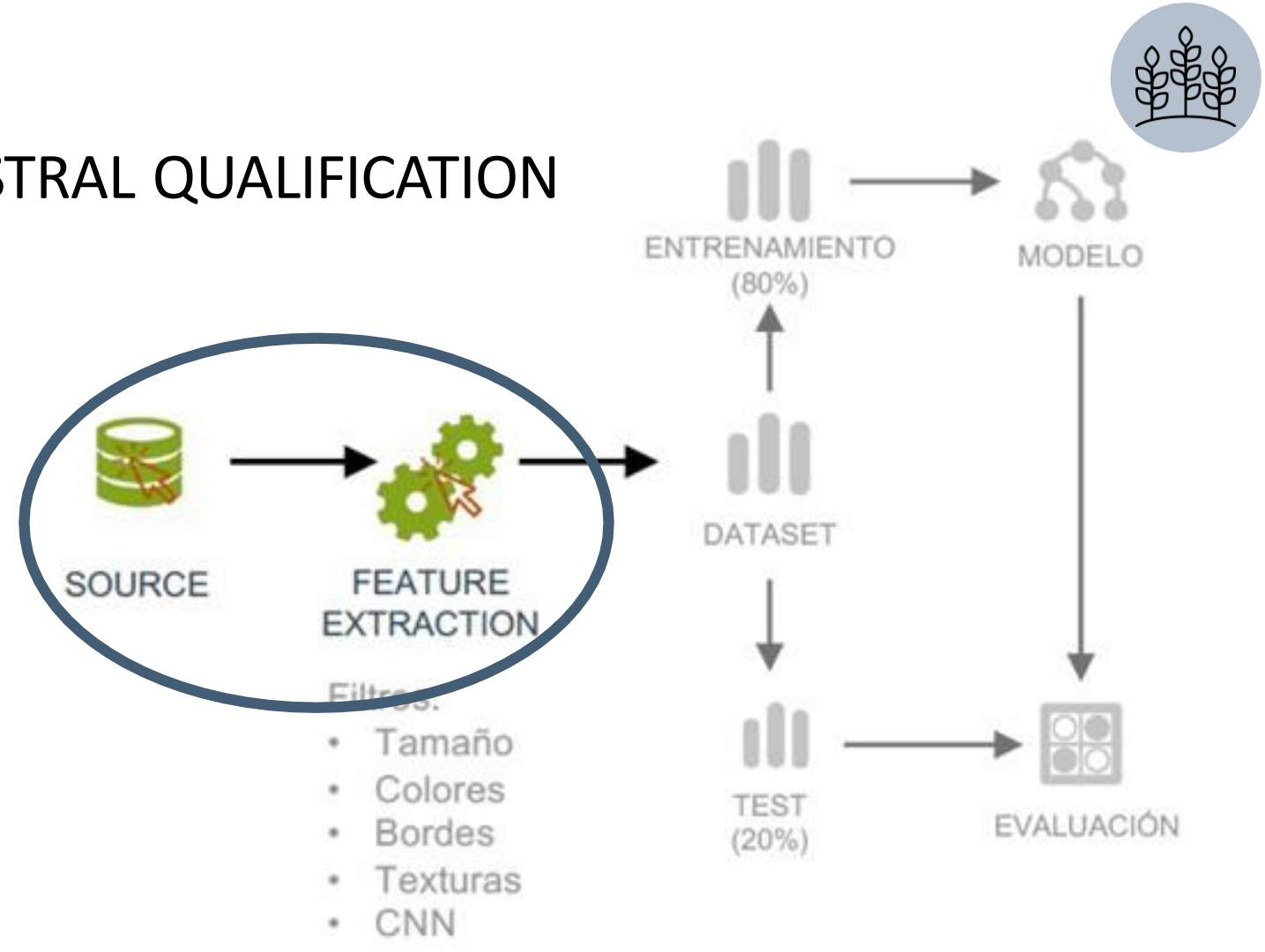
DATA QUALITY: BUILDING DETECTION AND CROP CLASSIFICATION

2.500 IMAGES PER CADASTRAL QUALIFICATION

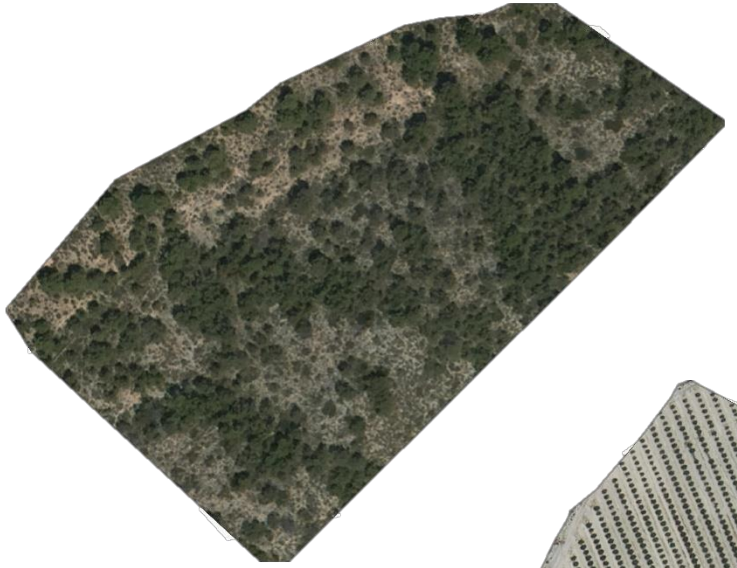
- SUPERVISED
- PNOA ORTHOPHOTOS
- SCALE 1:1000
- SUBPLOT

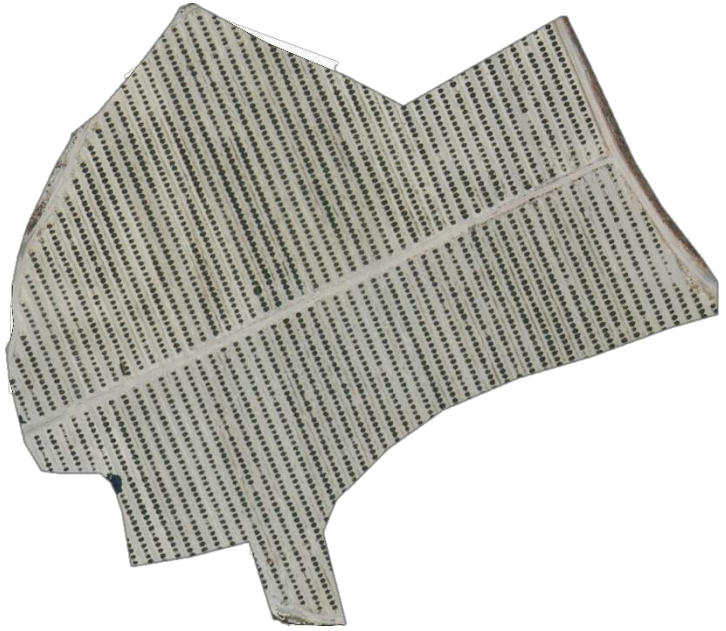
+ IMAGE PROCESSING

- FILTERS + CNN



DATA QUALITY: BUILDING DETECTION AND CROP CLASSIFICATION





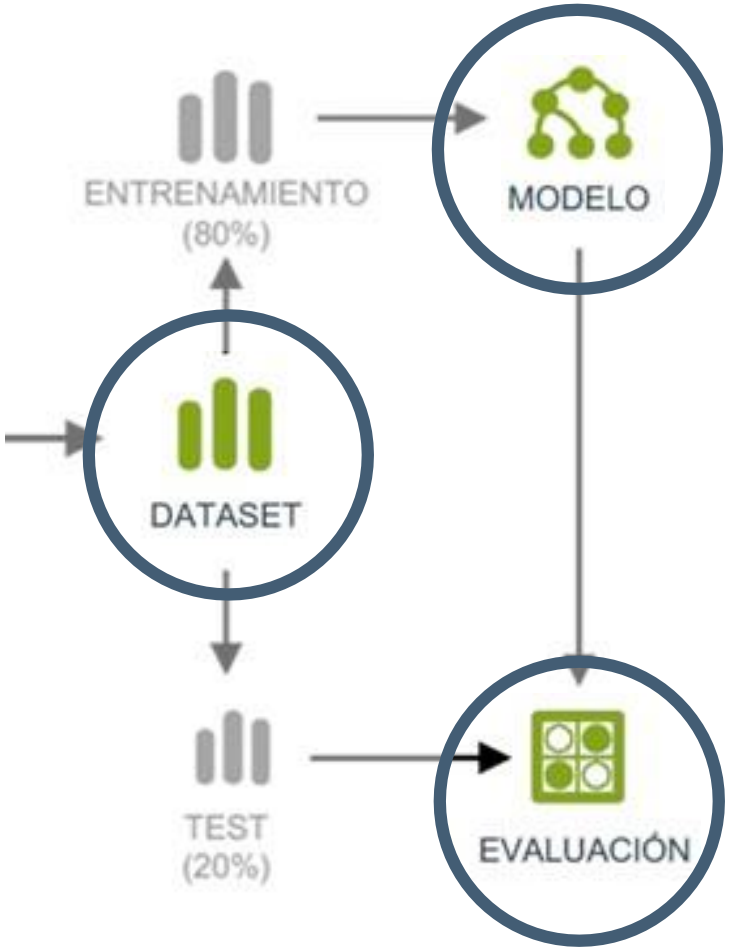
DATA QUALITY: BUILDING DETECTION AND CROP CLASSIFICATION

DATA SET

- 75,000 ORTHOPHOTOS WITH IDENTIFIED CC

+

- INFORMATION EXTRACTED THROUGH FILTERS AND CNN



MODEL DEVELOPMENT

80% DATA SET

+

MODEL VALIDATION

20% DATA SET

EXTERNAL VALIDATION

GRACIAS

Thank you for your attention

SUBDIRECCIÓN GENERAL DE TECNOLOGÍA Y DESPLIEGUE DIGITAL

SGTeDD

Miguel Ancochea Nodal
miguel.ancochea@catastro.hacienda.gob.es

DG del Catastro
www.catastro.hacienda.Gob.es