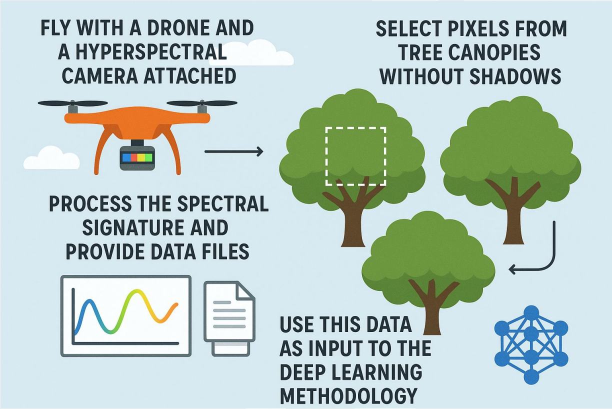
# Optimizing Picual Olive Variety Recognition through Deep Learning and Hyperspectral Imaging in Precision Agriculture





## Authors

**Alba Gómez** (Researcher, University of Jaén)

**Ruth M. Córdoba** (Researcher, University of Jaén)

**Juan J. Cubillas** (Inform. and Commun. Technologies applied to Education, International University of La Rioja)

**Lidia M. Ortega** (Computer Science. University of Jaén)

# Index


1. **The Challenge of Olive Variety Classification in Mixed Groves**
2. **Agricultural Context and Motivation of the Study**
3. **Research Objectives**
4. **Background and Related Work**
5. **General Methodology**
   - **5.1.** Hyperspectral Data Acquisition and Preparation
   - **5.2.** Canopy Segmentation and Spectral Filtering
   - **5.3.** Creation of Training, Validation, and Test Sets
6. **Deep Learning Model Architecture**
7. **Model Results**
8. **Discussion and Analysis of the Results**

# Problems in Olive Variety Classification in Mixed Groves

- ➔ Manual classification is labor-intensive and impractical for large groves.
- ➔ There are subtle differences between varieties, requiring advanced techniques.
- ➔ Only agricultural technicians and experienced farmers are able to distinguish varieties


**IMPORTANCE: Accurate variety identification enhances olive oil quality and resource efficiency**

# The Challenge of Olive Variety Classification in Mixed Groves

- **Impact**: affects irrigation, pruning, fertilization, and pest control strategies.
- **Need**: automated, accurate variety identification.
- **The objectives related with Precision Agriculture:**
  - Optimizing resources for sustainable farming.
  - Enhancing olive oil quality by controlling variety mixtures.
  - Improve crop yields and methodologies in large-scale groves.

# Agricultural Context and Motivation of the Study

## LOCATION

Jaén, Spain

World's leading olive oil producer.

## KEY VARIETIES

Picual and Arbequina dominate.


## STUDY SITE

IFAPA Venta del Llano Center

Mengíbar, Jaén.

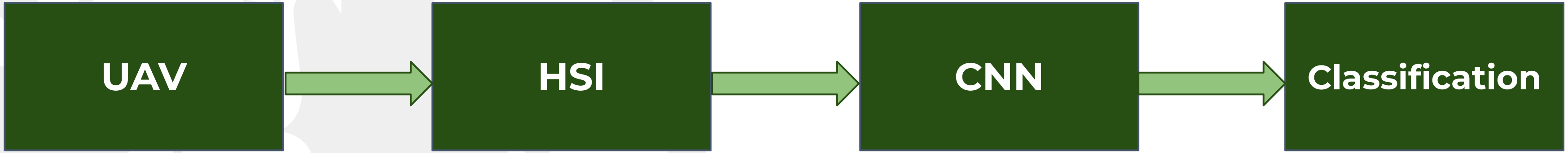
# Research objectives

➔ Utilize UAV-based hyperspectral imaging for high-resolution spectral data.

➔ Develop a deep learning-based approach for olive variety classification.

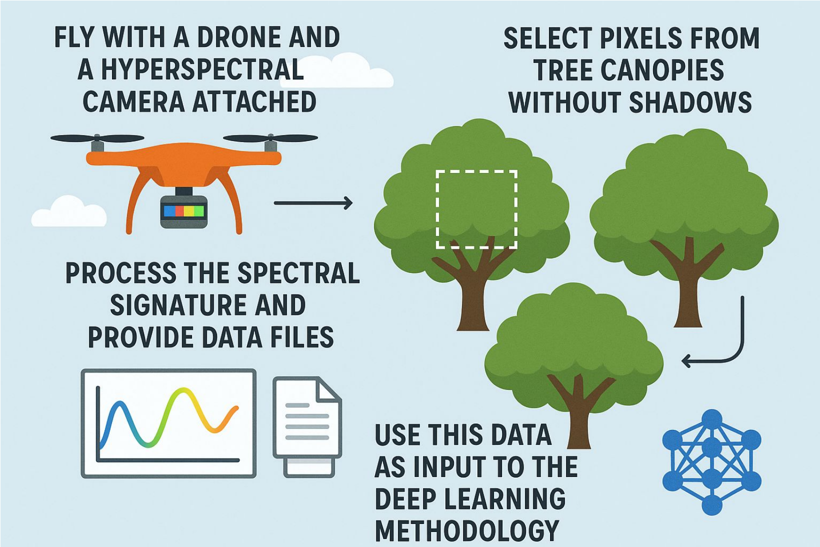
➔ Achieve high accuracy in automatic classification Picual vs. non-Picual varieties.

➔ Enable scalable, sustainable precision agriculture solutions.



# General methodology

- Fly with a drone and a hyperspectral camera attached
- Select pixels from tree canopies without shadows
- Process the spectral signature and provide data files
- Use this data as input to the Deep Learning methodology



# Materials and Methods

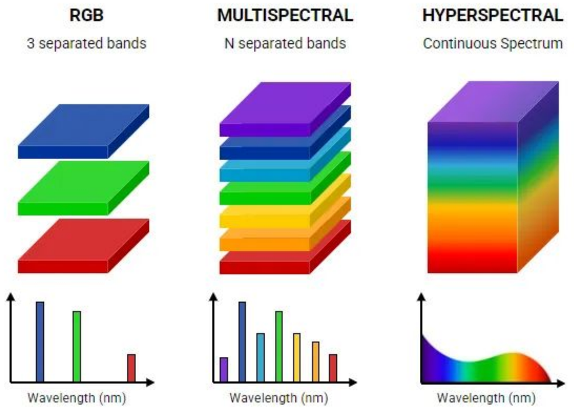
## Hyperspectral Imaging

➔ **HSI Overview:** captures 270 spectral bands (400–1000 nm).

➔ **Advantages:**

- Detects subtle reflectance differences vs. multispectral imaging.
- Used for vegetation monitoring, disease detection, and classification.

➔ **Relevance:** ideal for distinguishing olive varieties.



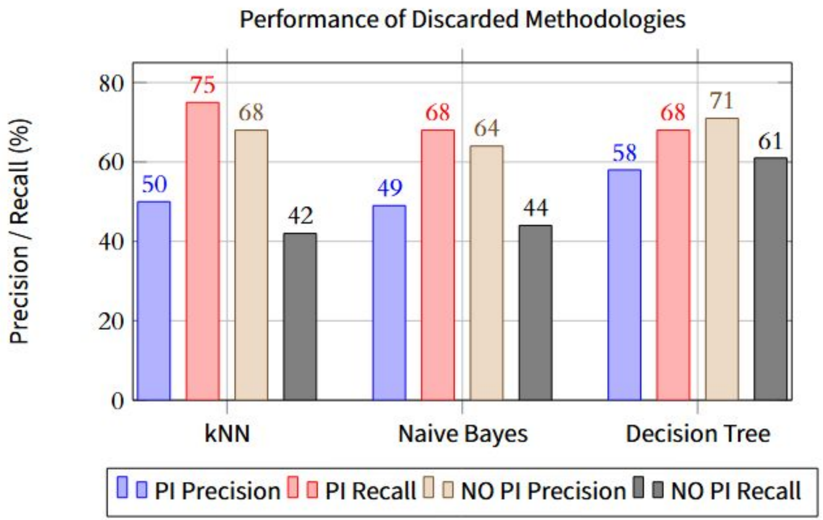

# Materials and Methods

## Deep Learning

- ➔ Convolutional Neural Networks (CNNs) handle high-dimensional HSI data.
- ➔ Automatic feature extraction reduces preprocessing.

**Comparison**: outperforms traditional methods (k-NN, Naïve-Bayes, Decision Tree).

**Gap**: limited DL research for olive variety classification.



# Specific Methodology

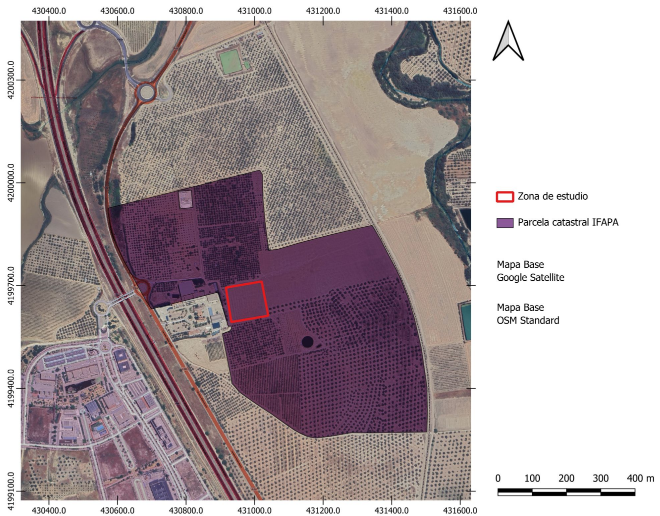
## Dataset

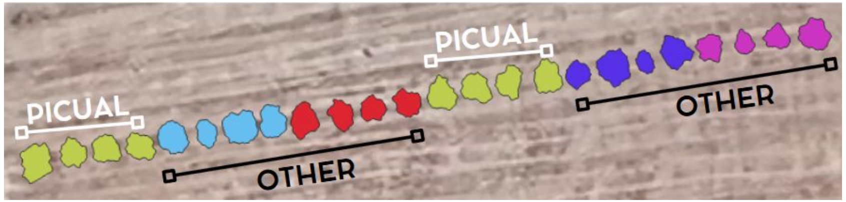
**Study Site:** IFAPA farm, Mengíbar, Jaén
14 rows, 24 trees each.

**Steps:**

1. Hyperspectral data acquisition via UAV + LiDAR.
2. Canopy segmentation and spectral filtering.
3. Dataset preparation including augmentation (training/validation/test).
4. CNN modeling for Picual classification.

**Dataset:** 264 Picual and 264 non-Picual trees.



# Specific Methodology
## Hyperspectral Data Acquisition and Preparation

**Acquisition:**

UAV with NanoHyperspec camera and LiDAR sensor.

Flight: 30m altitude, 5 m/s speed, optimized overlap.

**Data:** 270 spectral bands (400–1000 nm).

**Preparation:**

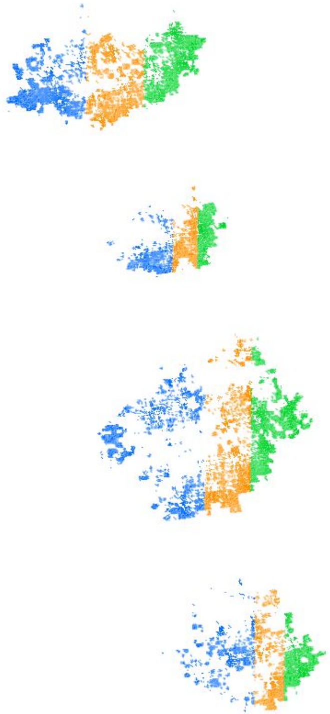
Reflectance calibration using Headwall SpectralView™.

Geometric correction with LiDAR-derived DEM.

**Outcome:** High-quality dataset of 24 olive trees.

# Specific Methodology

## Canopy Segmentation and Spectral Filtering



**Canopy Segmentation:**

**Method:** Enhanced Vegetation Index (EVI) and DBSCAN clustering.

**Outcome:** individual tree canopies delineated with unique identifiers.

**Spectral Filtering:**

**Method:** remove noisy pixels (e.g., shadows) using NIR reflectance.

**Outcome:** spectrally stable dataset for CNN modeling.

**Data Augmentation:**

**Method:** obtain several individuals from an unique olive tree with Kmeans

**Outcome:** An augmented set of olive trees

# Specific Methodology

## Creation of Training, Validation, and Test Sets

**Dataset:** 264 Picual and 264 non-Picual trees (balanced).

**Training:** 80% (180 Picual, 186 non-Picual).

**Validation:** 20% (49 Picual, 43 non-Picual).

**Test:** 35 Picual, 35 non-Picual.

**Purpose:** ensure diverse, representative samples to prevent overfitting.

| Dataset | Train Data | Validation Data | Test Data | Total |
|---|---|---|---|---|
| PI | 180 | 49 | 35 | 264 |
| NO PI | 186 | 43 | 35 | 264 |

# Deep Learning Model Architecture

## General View

**Model:** 1D CNN for Picual vs. non-Picual classification.

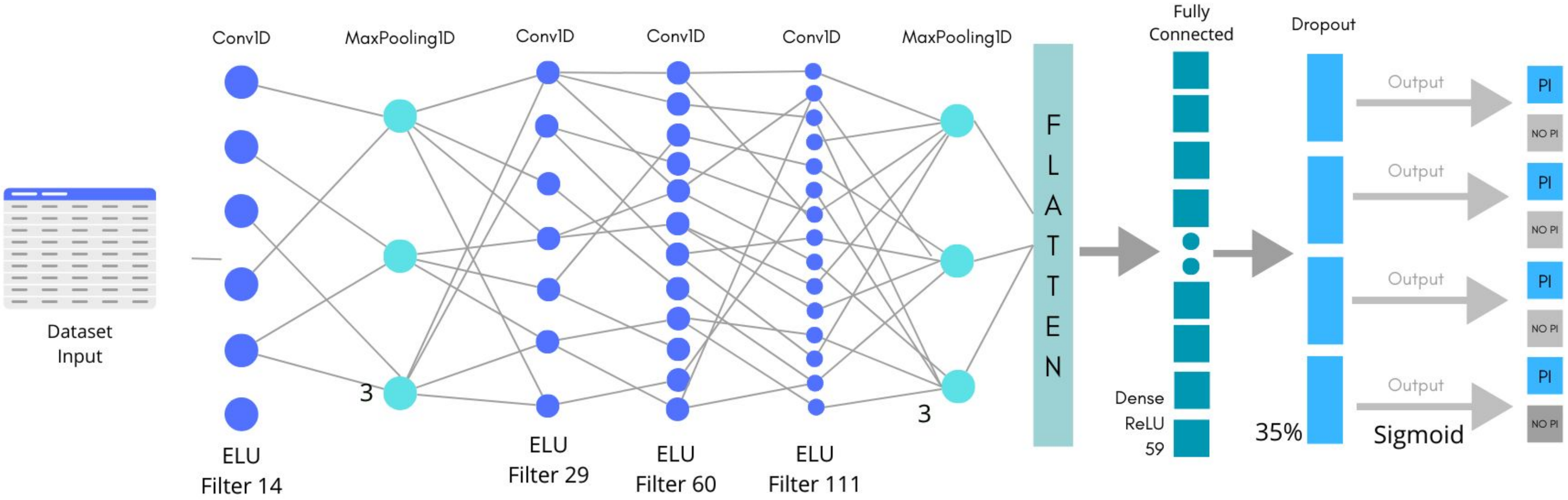
**Architecture:**

4 convolutional layers (ELU activation).
2 max-pooling layers (pool size = 3).
Flatten layer
Fully connected layer (ReLU) with dropout.
Sigmoid output for binary classification.

**Optimization:** Bayesian optimization for hyperparameters.

# Deep Learning Model Architecture

## General View



# Deep Learning Model Architecture

## Hyperparameter Optimization

**Approach:** manual tuning followed by *Bayesian* optimization.

**Parameters:**

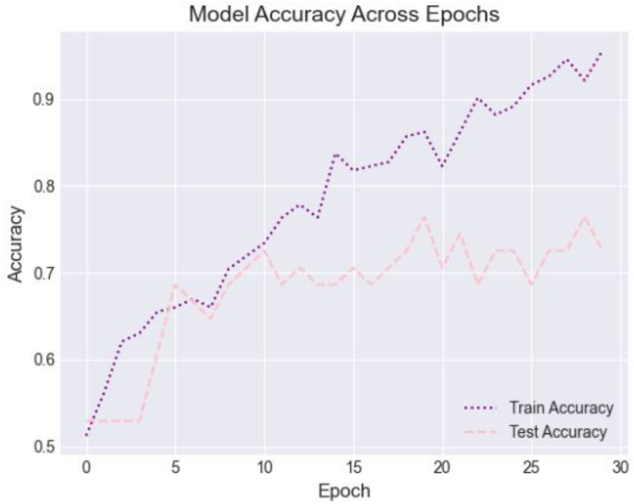
- ➔ **Filters:** 14, 29, 60, 111
- ➔ **Kernel Size:** 4
- ➔ **Dense Neurons:** 59
- ➔ **Pool Size:** 3
- ➔ **Batch Size:** 32
- ➔ **Epochs:** 50
- ➔ **Patience:** 10

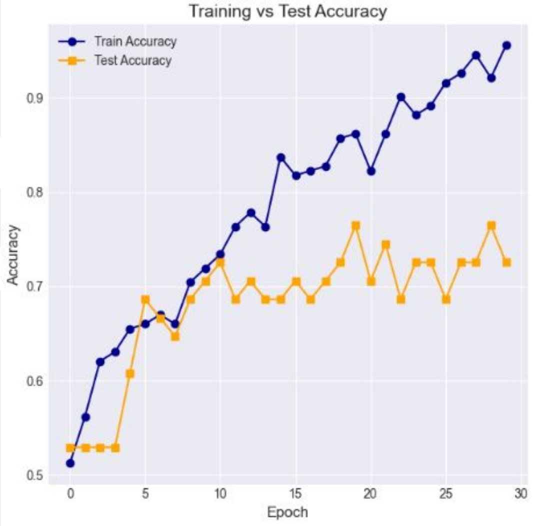
**Benefit:** enhanced model accuracy and efficiency.

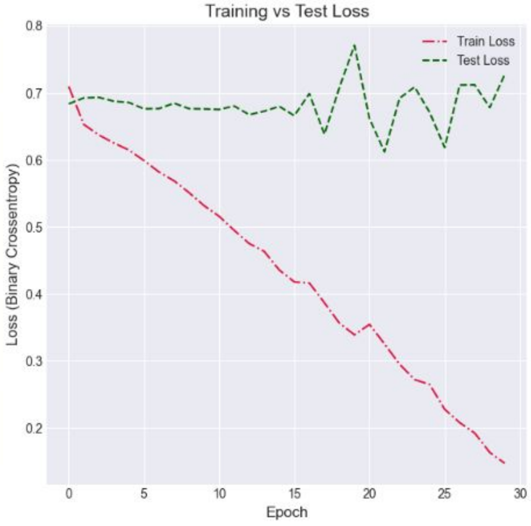
| Picual | |
|---|---|
| **Hyperparameters** | **Value** |
| Filter 1 | 14 |
| Filter 2 | 29 |
| Filter 3 | 60 |
| Filter 4 | 111 |
| Kernel size | 4 |
| Dense Neurons | 59 |
| Pool Size | 3 |
| Batch Size | 32 |
| Epochs | 50 |
| Patience | 10 |

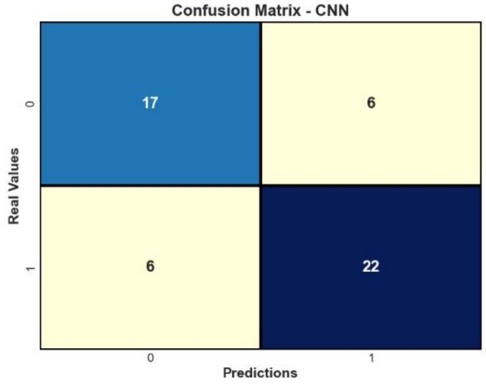
# Model Results

| Class | Precision | Recall | F1-Score |
|---|---|---|---|
| Picual | 0.84 | 0.96 | 0.90 |
| No Picual | 0.94 | 0.79 | 0.86 |
| **Accuracy** | | 0.88 | |
| **Macro Avg** | 0.89 | 0.87 | 0.88 |
| **Weighted Avg** | 0.89 | 0.88 | 0.88 |



# Discussion and Analysis of the Results

## Key Findings:

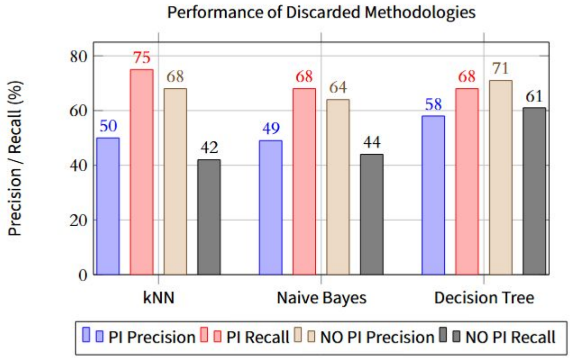
- ➔ CNN achieves 88% accuracy, outperforms traditional methods.
- ➔ High recall (96%) for Picual ensures reliable identification.

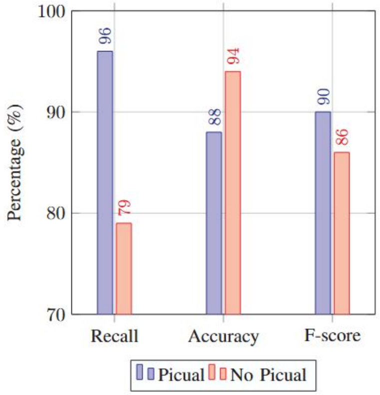
## Strengths:

- ➔ Automated feature extraction simplifies processing.
- ➔ Scalable UAV-based HSI.

## Limitations:

- ➔ Single farm testing limits generalizability.
- ➔ Environmental condition impacts.



# Discussion and Analysis of the Results

## Future Work

- ➔ Expand to **more varieties and farms**.
- ➔ Test under **diverse environmental conditions**.
- ➔ Explore **2D/3D CNNs** and **integrate LiDAR**.
- ➔ **Automate regional species cataloging.**

# Thanks for your attention

# Any questions??

Grupo de Gráficos y Geomática de Jaén

Universidad de Jaén

**Alba Gómez Liébana | aglieban@ujaen.es**

**Lidia Ortega Alvarado | lidia@ujaen.es**

**Ruth Córdoba Ortega | rcortega@ujaen.es**

**Juan J. Cubillas | juanjose.cubillas@unir.net**