**ANALISI SPETTRALE E CLASSIFICAZIONE DELLE SPECIE ARBOREE CON TECNICHE DI IMAGE PROCESSING E MACHINE LEARNING SU IMMAGINI SATELLITARI**

**INTRODUZIONE**

Negli ultimi anni, l'utilizzo delle immagini satellitari per l'analisi ambientale ha registrato un notevole incremento grazie ai progressi tecnologici nel campo del telerilevamento e dell'elaborazione delle immagini; il nostro progetto si propone di approfondirne i concetti chiave ed applicarli ad uno specifico caso: quello della classificazione degli ulivi sul territorio Pugliese.

Il nostro studio, partendo proprio da immagini satellitari ad alta risoluzione, tenta di integrare avanzate tecniche di elaborazione con modelli di machine learning, offrendo una soluzione automatizzata per la classificazione delle specie arboree; tutti i codici e le funzioni utili ai nostri scopi, verranno scritti nel linguaggio di programmazione Matlab.

Gli obiettivi principali sono:

* In primis, creare un database sufficientemente informativo e completo che, con delle features mirate, permetta in seguito all’algoritmo di discriminare al meglio tra le varie specie di ulivi.
* Successivamente creare diversi modelli di Machine Learning in grado di effettuare la vera e propria classificazione delle specie di alberi partendo dal dataset fornito.
* Infine, confrontare l’efficacia dei sopracitati modelli, valutando le performance in termini di accuratezza, precisione e robustezza.

I risultati ottenuti potrebbero avere interessanti implicazioni per il monitoraggio ecologico e la gestione sostenibile delle risorse forestali, supportandone le decisioni nella pianificazione; inoltre, migliorerebbe la nostra capacità di monitorare la biodiversità, rilevare cambiamenti ecologici e identificare specie a rischio.

**STATO DELL’ARTE**

Come già accennato nell’introduzione, negli ultimi anni abbiamo assistito a notevoli progressi nel mondo del telerilevamento, con l’introduzione di nuove tecniche e tecnologie che hanno migliorato l’accuratezza e l’affidabilità con la quale effettuare stime sulla vegetazione.

Nell’analisi della letteratura scientifica abbiamo individuato diversi strumenti e componenti hardware adatti allo scatto di immagini multispettrali; ad esempio, in [1] vengono usate immagini multispettrali satellitari QUICKBIRD e IKONOS (tramite satellite IKONOS-2).

Un’alternativa potrebbe essere quella esaminata in [2] e [3] dove il detecting delle chiome non viene supportato da immagini ottenute da satellite, bensì da un UAV Italdron 4HSE EVO (drone multi-rotore) da un’altezza di circa 70 metri su cui sono montate diverse fotocamere:

* Fotocamera Multispettrale a cinque bande MicaSense RedEdge-M.
* Fotocamera Termica FLIR Vue Pro 640 (per la cattura di immagini termiche ad alta risoluzione).
* Fotocamera Visibile ad Alta Risoluzione Sony α7r.

Questo tipo di strumentazione è particolarmente efficace nel bilanciare la qualità delle immagini (risoluzione) e l'efficienza del volo (resa dei rilievi aerei), considerando le caratteristiche specifiche degli oliveti.

Per quanto riguarda invece la segmentazione delle chiome, diverse ricerche come [4], propongono di eseguire una trasformata di Hough circolare poiché le chiome hanno una forma simile ad una circonferenza, oppure l’algoritmo di K-Means (dove le chiome saranno i clusters); Da osservare che nel nostro caso non è stato necessario effettuare questo ulteriore step poiché ci sono state già fornite immagini con la segmentazione delle chiome da usare come maschera binaria sulle immagini satellitari.

Infine, [5] e [6] mostrano come utilizzare (e con quali risultati), gli algoritmi di Machine Learning utili alla classificazione delle specie di ulivi; di seguito una lista dei più utilizzati in letteratura:

* Support Vector Machines (SVM)
* Random Forest
* Linear Discriminant Analysis (LDA)
* Neural Networks per la classificazione

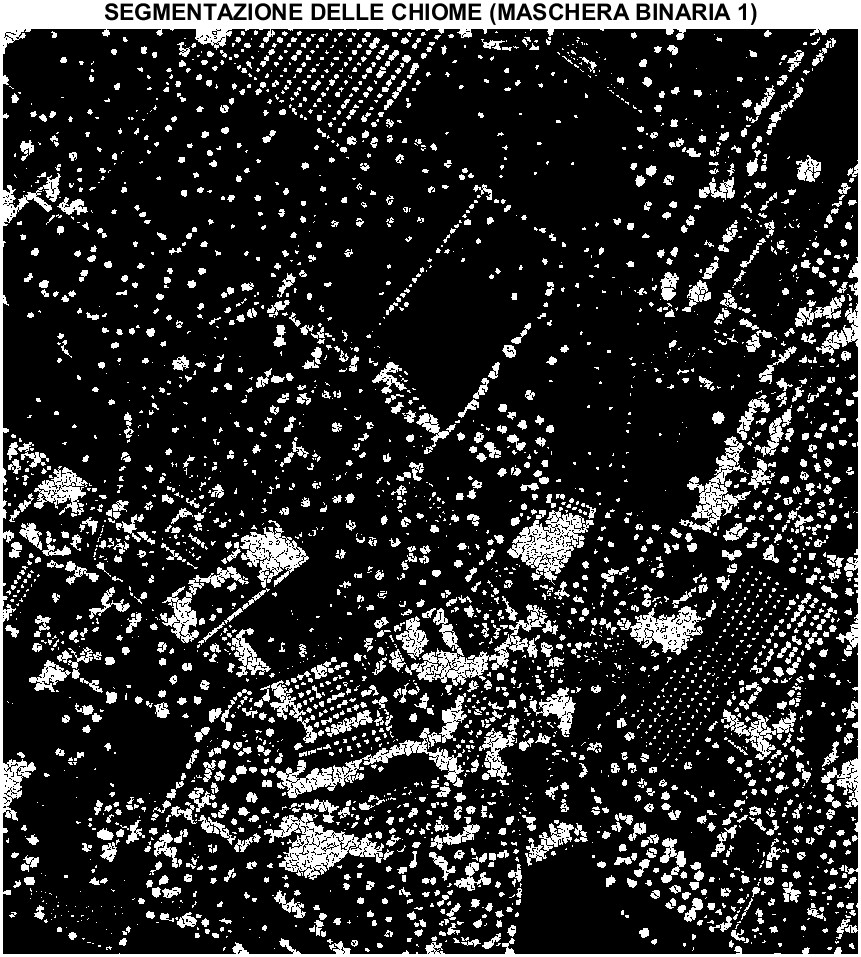
**SEZIONE 1 – CREAZIONE DEL DATABASE**

Il primo step è quello di ottenere un database da usare per training e test; esso dovrà contenere, per ogni chioma, informazioni sul valore medio dei suoi pixel, per ognuna delle 47 bande; questo vuol dire avere uno spettro medio per ogni albero e poterlo usare in seguito, per la classificazione.

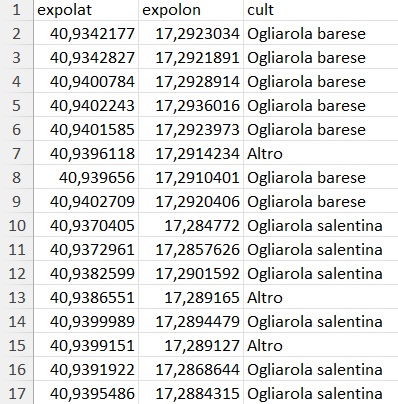
**Materiale fornito**

Per raggiungere il nostro scopo, a cui abbiamo accennato nell’introduzione, possiamo contare sul seguente materiale:

* due immagini satellitari multispettrali ad alta risoluzione con 47 differenti bande in formato .TIF di campagne Pugliesi contenenti ciascuna migliaia di chiome.
* Le rispettive maschere binarie contenenti tutte le chiome arboree già segmentate (da utilizzare per isolare l’ulivo dallo sfondo).



* I rispettivi database in formato Excel che includono alcune centinaia di alberi con la loro posizione in coordinate geografiche, già classificati con alcune delle tipologie di ulivo presenti sul territorio Pugliese (come “Leccino”, “Ogliarola Barese”, ecc.); i dati sono raccolti da operatori direttamente sul campo.



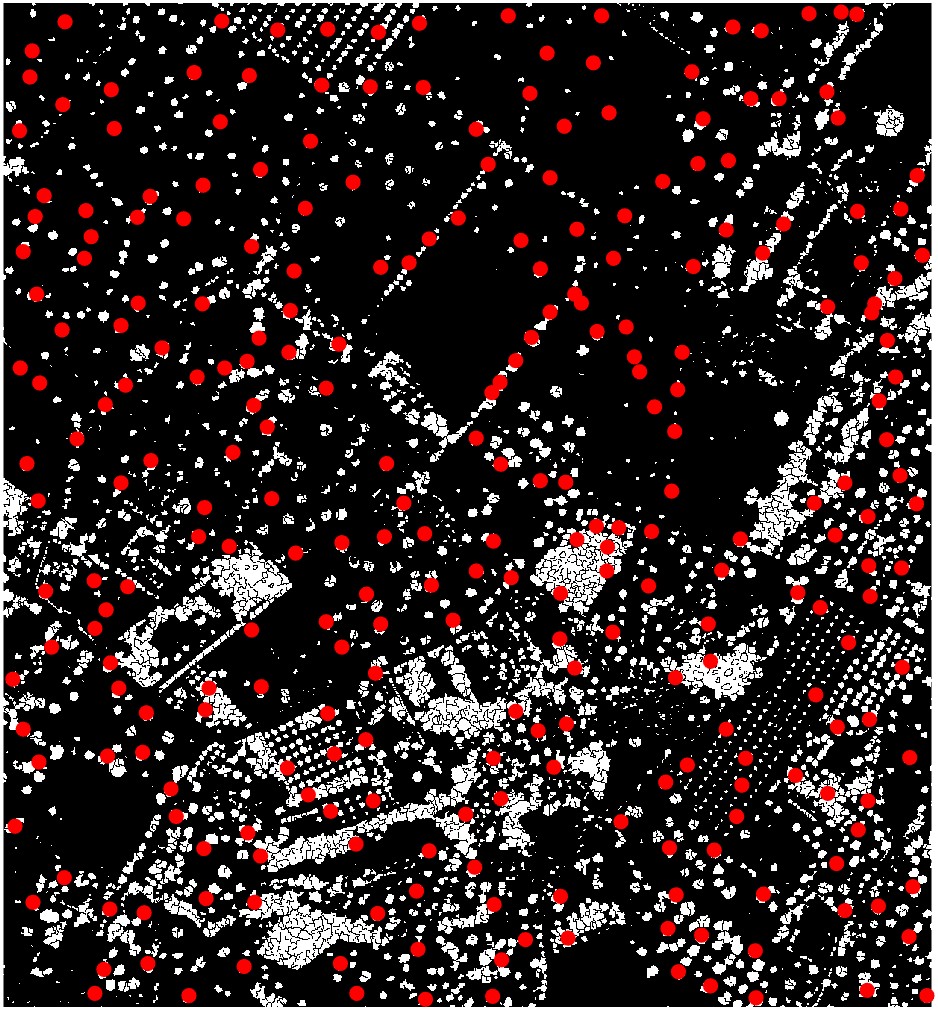
**Importazione del Database**

Per prima cosa utilizzo il comando “readtable” di Matlab per leggere il database in formato Excel già fornito ed elimino le righe (o samples) che contengono dati mancanti, successivamente assegno a delle variabili i valori di latitudine, longitudine e coltivazione.

**Trasformazione in coordinate intrinseche**

Per mappare le coordinate fornite dal database in punti dell’immagine in cui vi è presente una chioma, trasformiamo le coordinate geografiche in coordinate intrinseche dell’immagine georeferenziata: viene fatto effettuando una proiezione tramite il comando “worldToIntrinsic”.

Ora abbiamo una matrice “points” contenente le coordinate delle chiome presenti e classificate nel nostro database; di seguito una visualizzazione della loro distribuzione rispetto al totale degli alberi presenti in un crop:



OSSERVAZIONE – Essendo un database costruito tramite dati ottenuti in modo manuale da operatori sul campo, le coordinate di alcuni alberi risultano non coincidenti con una specifica chioma nell’immagine, bensì sono sullo sfondo: in questi casi il data sample andrà perduto e la chioma non considerata nel training e nel test.

**ID chioma univoco**

Per trattare ogni chioma della nostra immagine in modo univoco, necessitiamo di un ID da conferire ad ognuna di esse; quindi, dopo aver importato la maschera binaria andiamo ad usare il comando Matlab “bwlabel” che va ad etichettare ogni cluster dell’immagine segmentata ottenendo una matrice “L” con valore ‘0’ per le celle dello sfondo e ‘ID\_chioma’ per le celle dove sono presenti ulivi; inoltre, in “num” avremo il numero totale di essi che è nell’ordine delle migliaia.

Successivamente eseguo un semplice ciclo for per tener conto, in un vettore chiamato “id\_chiome\_db”, degli ID dei soli ulivi georeferenziati nel Database (e quindi utili alla classificazione).

**Calcolo spettro medio**

A questo punto, il nostro intento è quello di creare una matrice con 47 colonne e tante righe quanti alberi, in modo da contenere per ogni riga (corrispondente alla singola chioma), lo spettro medio; per prima cosa, quindi, importiamo l’immagine multispettrale a 47 bande su cui effettuare le varie elaborazioni.

Successivamente creiamo un ciclo for che scorre tutte le chiome del nostro database in base al loro ID e, per ognuna delle 47 bande, andiamo per prima cosa a ‘mascherare’ l’immagine multispettrale eliminando lo sfondo; di seguito una visualizzazione del risultato di questo processo.

|  |
| --- |
| Immagine che contiene schermata, stella, costellazione, spazio  Descrizione generata automaticamente |

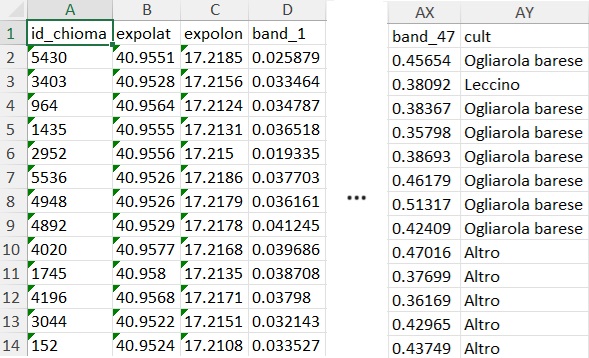
Poi, andiamo a calcolare l’array “Values” che contiene tutti i valori dei pixel di una singola chioma per una specifica banda: su questo andremo a ricavare il valore medio eliminando gli outliers.

OSSERVAZIONE – Consideriamo outliers tutti i valori che si allontanano dalla media più di una certa soglia (nel nostro caso la soglia è due volte la deviazione standard)

A questo punto utilizzeremo questo valore medio per riempire la cella (i, banda) della matrice “Firma\_spettrale\_media” dove “i” è la specifica chioma.

**Aggiornamento del Database**

Per concludere il nostro lavoro di creazione del database da utilizzare per il successivo step di classificazione delle specie arboree, vogliamo esportare (utilizzando Matlab) un file Excel che contenga, oltre le informazioni di latitudine, longitudine e coltivazione, anche lo spettro medio delle chiome e un ID univoco; quello che vogliamo ottenere, quindi, è un file strutturato in questo modo:

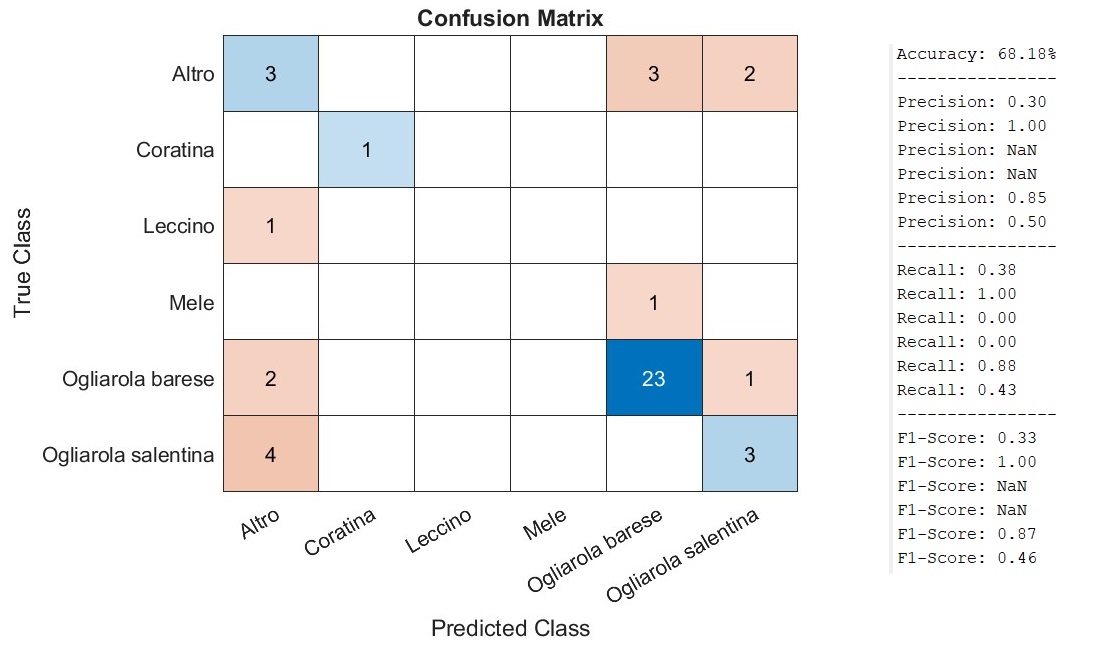
****

Lo facciamo concatenando orizzontalmente i vari vettori contenenti le informazioni che ci servono per costruirlo (rispettivamente ID, latitudine, longitudine, banda 1, banda 2, …, banda 47, coltivazione) e associando i dati alle rispettive labels con “array2table”; infine usiamo il comando “writetable” per esportare la tabella creata in un file Excel nella directory corrente.

**SEZIONE 2 – MODELLI DI CLASSIFICAZIONE**

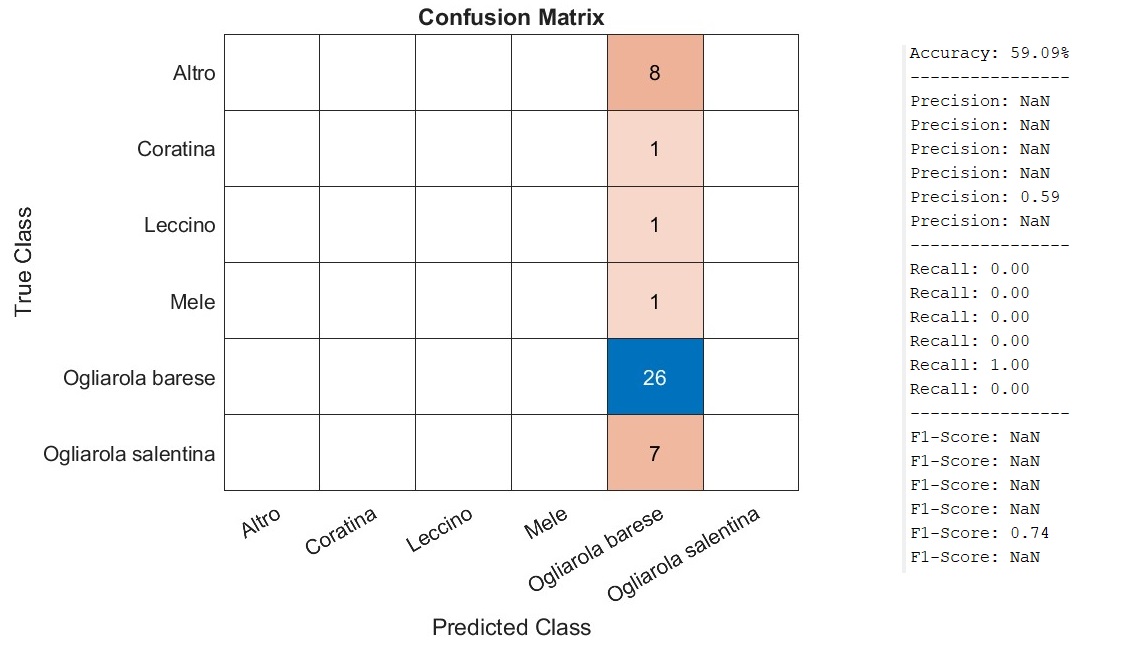
**RISULTATI**

**SVM con Kernel lineare**

****

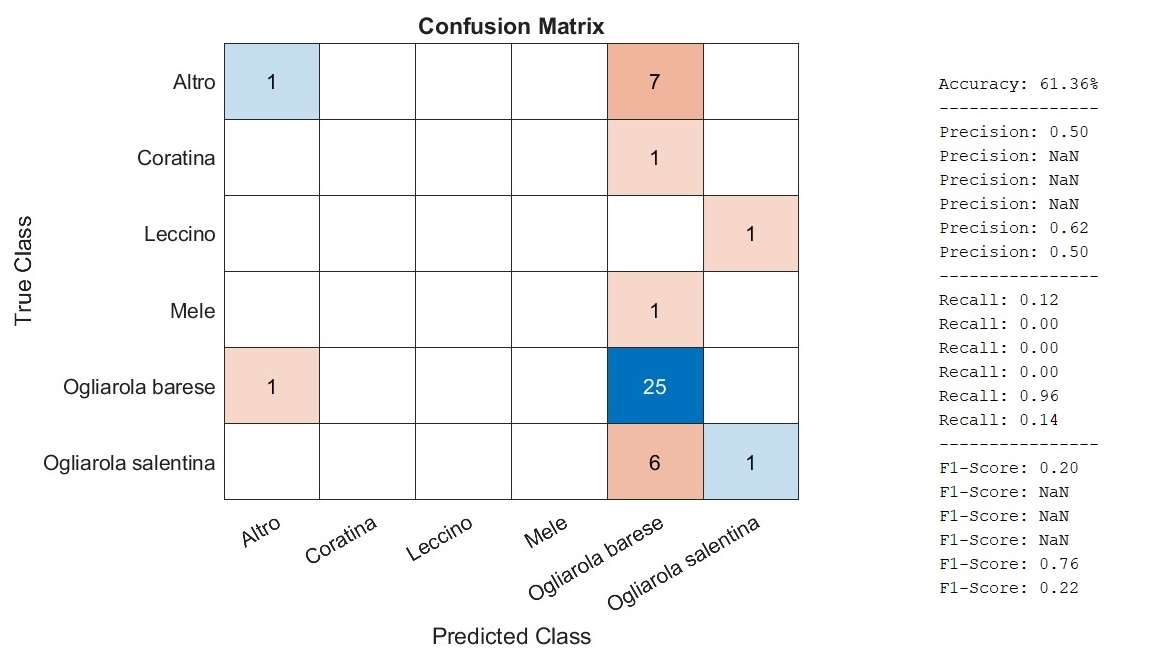
---------------------------------------------------------------------------------------------------------------

**SVM con Kernel gaussiano**

****

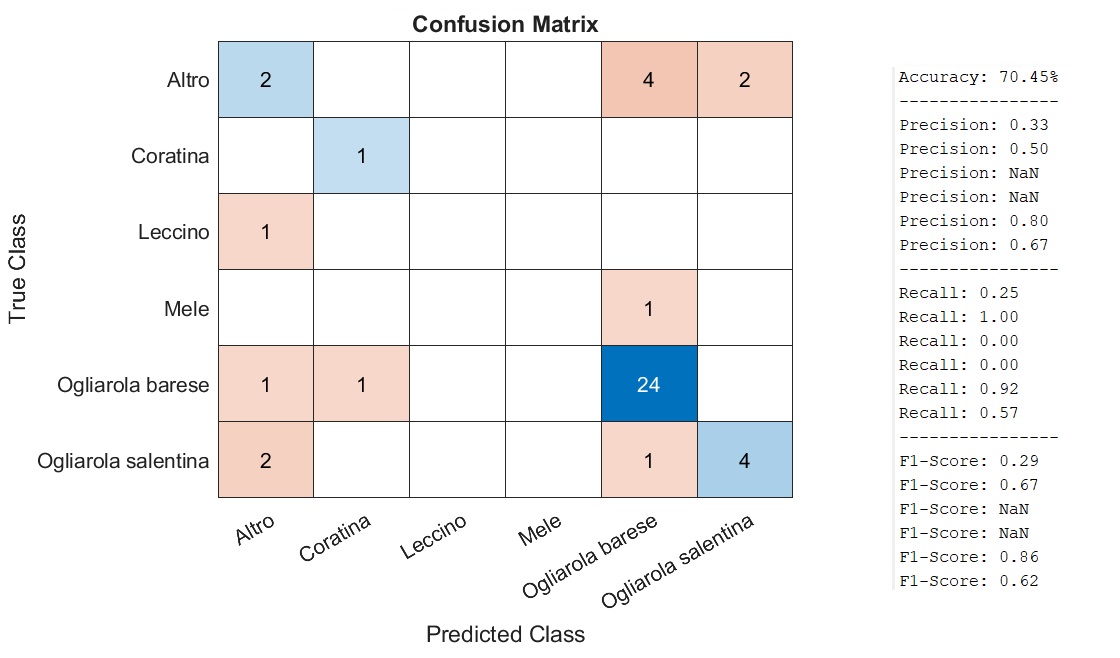
---------------------------------------------------------------------------------------------------------------

**Random Forest**

****

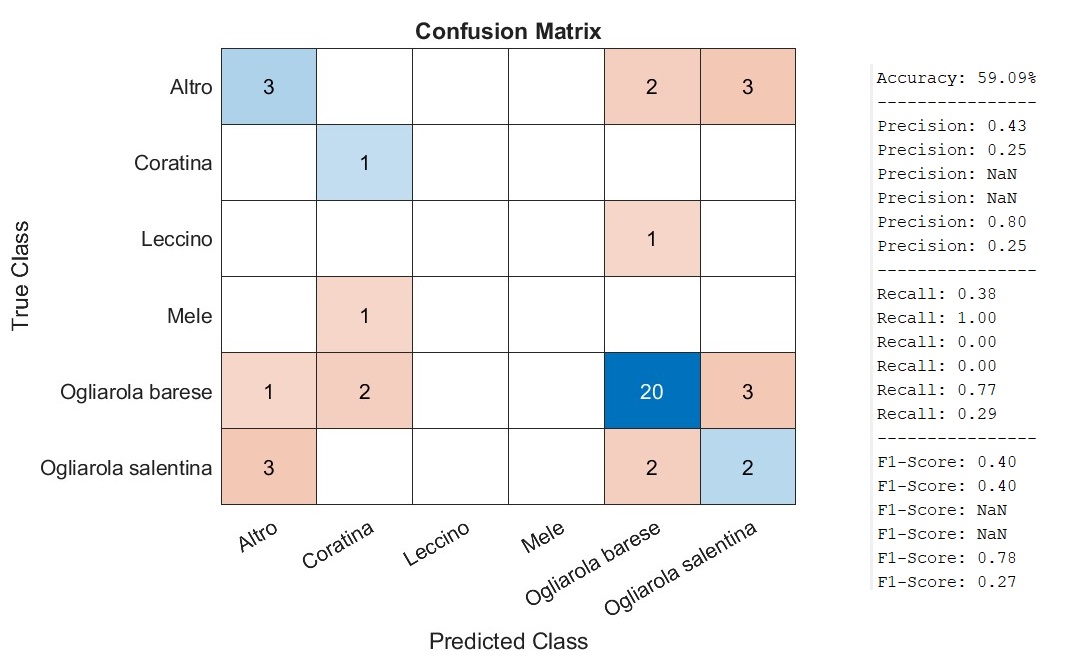
---------------------------------------------------------------------------------------------------------------

**LDA con Discriminant Type linear**

****

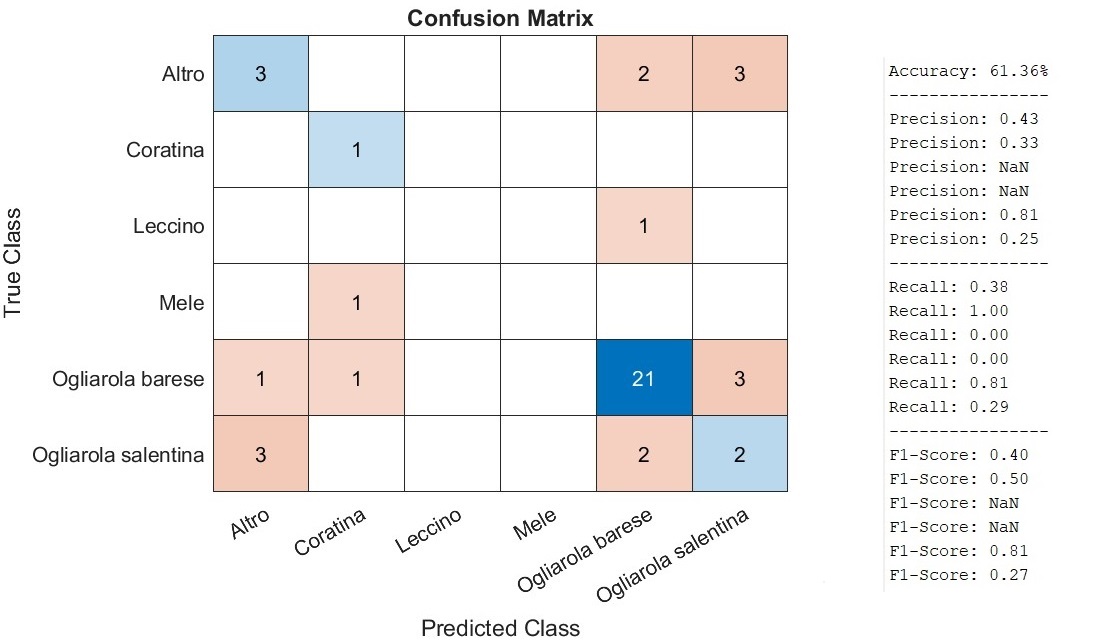
---------------------------------------------------------------------------------------------------------------

**LDA con Discriminant Type diaglinear**

****

---------------------------------------------------------------------------------------------------------------

**LDA con Discriminant Type pseudolinear**

****

**RIFERIMENTI**

[1] M. Waleed, T. -W. Um, A. Khan and Z. Ahmad, "An Automated Method for Detection and Enumeration of Olive Trees Through Remote Sensing," in IEEE Access, vol. 8, pp. 108592-108601, 2020, doi: 10.1109/ACCESS.2020.2999078.

[2] «Sensors | Free Full-Text | Fast Detection of Olive Trees Affected by Xylella Fastidiosa from UAVs Using Multispectral Imaging». Consultato: 8 marzo 2024. [Online]. Disponibile su: <https://www.mdpi.com/1424-8220/20/17/4915>

[3] F. Adamo, F. Attivissimo, A. Di Nisio, M. A. Ragolia and M. Scarpetta, "A New Processing Method to Segment Olive Trees and Detect Xylella Fastidiosa in UAVs Multispectral Images," 2021 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), Glasgow, United Kingdom, 2021, pp. 1-6, doi: 10.1109/I2MTC50364.2021.9459835.

[4] A. Khan et al., "Remote Sensing: An Automated Methodology for Olive Tree Detection and Counting in Satellite Images," in IEEE Access, vol. 6, pp. 77816-77828, 2018, doi: 10.1109/ACCESS.2018.2884199.

[5] M. S. Mandava, D. Jadhav and R. R. Naik, "Fault classification using SVM," 2015 IEEE International Circuits and Systems Symposium (ICSyS), Langkawi, Malaysia, 2015, pp. 17-21, doi: 10.1109/CircuitsAndSystems.2015.7394056.

[6] G. Ramat et al, "Mapping of olivetrees using Sentinel-2 and Sentinel-1 images: an assessation of pixel-based analyses," 2023 IEEE International Workshop on Metrology for Agriculture and Forestry (MetroAgriFor), Pisa, Italia, 2023, pp. 263-267, doi: 10.1109/MetroAgriFor58484.2023.10424313.