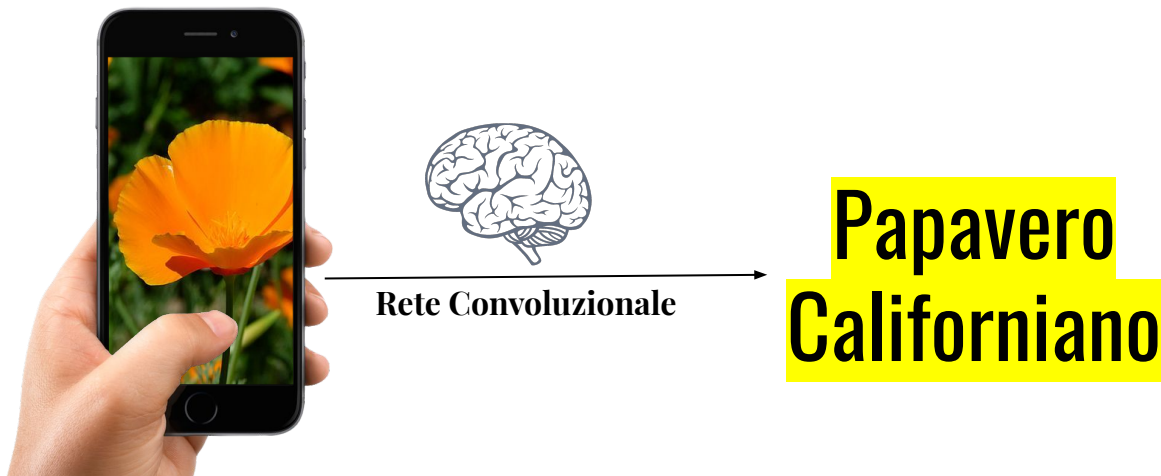


Classificazione di immagini di 102 specie di fiori

Baldi - Olivadese - Vitali

Obiettivi del progetto

- **Classificazione** supervisionata di **immagini** rappresentanti **102 specie di fiori**
- **Confronto** di 4 categorie di **modelli** basati su **reti neurali**
- **Ottimizzazione** degli **iperparametri** e della **struttura** dei modelli

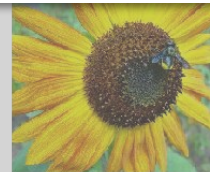
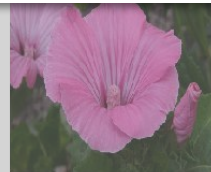
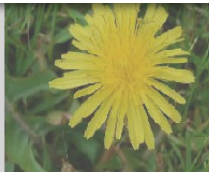
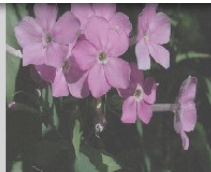
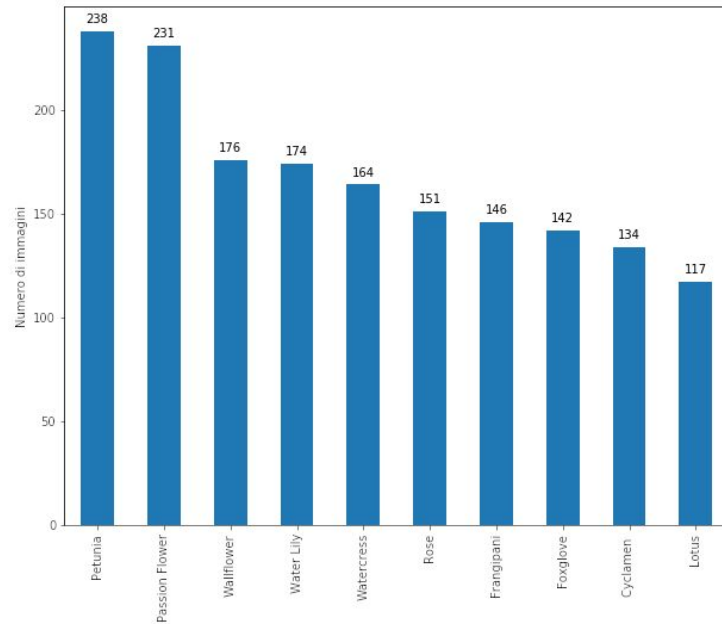
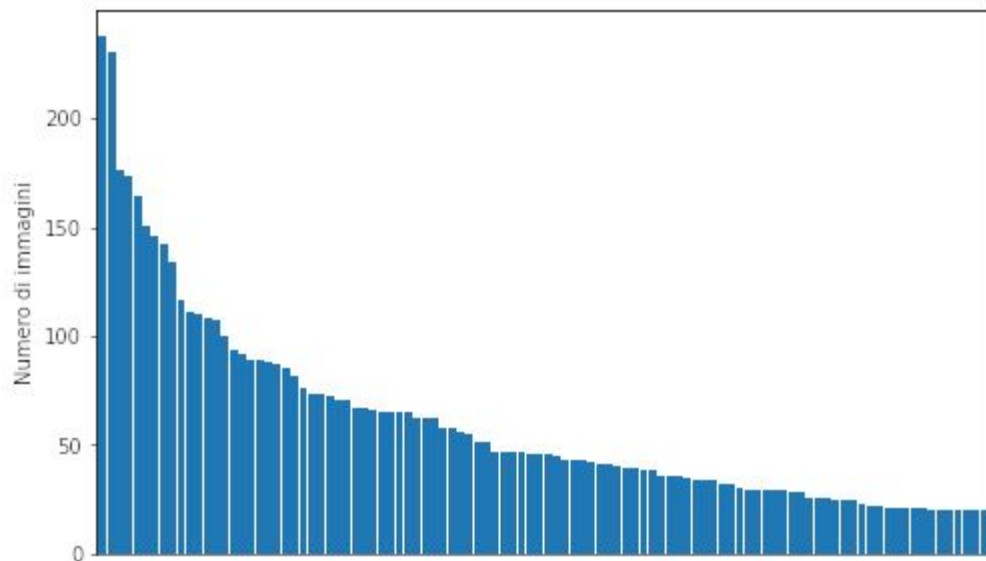


Dataset

- **102 Flowers Dataset** – Università di Oxford
 - 102 specie di fiori
 - **1020** immagini di **training** (10 per classe)
 - **1020** immagini di **validazione** (10 per classe)
 - **6149** immagini di **test**



Distribuzione Dataset di Test



Dataset

Alta similarità tra specie diverse

English Marigold



Barbeton Daisy



Tree Poppy



Japanese Anemone



Hibiscus



Azalea



Alta variabilità nella stessa specie

Snapdragon



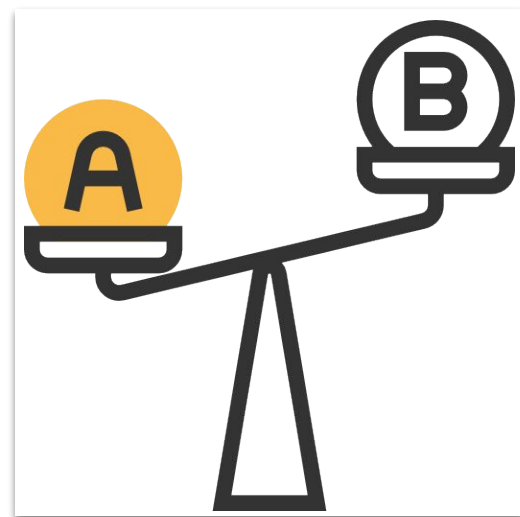
Dataset - Data Augmentation

- Solo 10 immagini per classe nel training set
- **Data Augmentation** per ogni epoca della fase di training, uno o più tra:
 - **Ribaltamento** verticale o orizzontale
 - **Zoom-in** e **Zoom out** del 20%
 - **Rotazione** tra -30° e $+30^\circ$
 - Variazione della **luminosità** del 20%



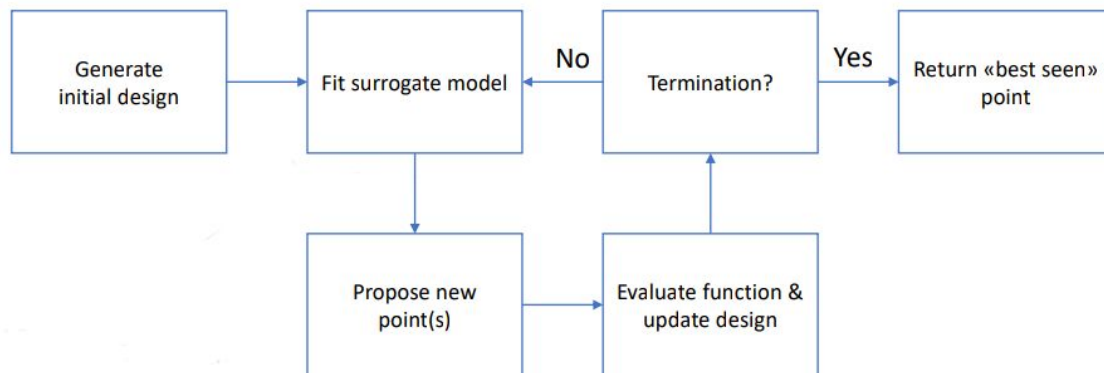
Approccio Metodologico

- Confronto di 4 modelli
 - **Convolutionale “semplice”**
 - Basato su **Autoencoder**
 - Basato su **Feature Extraction VGG16**
 - Basato su **Fine Tuning VGG16**
- Per tutti i modelli
 - Funzione di Loss: **Categorical Cross Entropy**
 - Funzione di attivazione dell'ultimo layer: **Softmax**
 - 100 epoche massime di training, **Early Stop** con patience a 5 sulla validation loss
 - Aggiunti layer di dropout per regolarizzazione
 - Ottimizzatore **Adam**
 - Learning rate deciso con AutoML



Ottimizzazione iperparametri e struttura

- **Difficile** individuare la **struttura ottima** dei modelli e i loro **iperparametri**
- Utilizzo di **SMAC** per **SMBO**
 - Ottimizzazione Bayesiana
 - Modello surrogato **Random Forest**
 - Funzione di acquisizione **Expected Improvement**
- **30 campionamenti** dello spazio di ricerca + 6 di inizializzazione
- Valutazione della bontà del modello: **loss** sul validation set



Classificatore Convolutionale “Semplice”

- **Layer di Input** (244 x 244 x 3)
- 5 coppie di layer:
 - 64/128/256/512/512 **Kernel Convoluzionali** 3 x 3
 - Layer di **Max Pooling** 2 x 2
- **Layer di Flatten** (25088 valori)
- Da 1 a 4 coppie di layer:
 - **Layer di Dropout** con rate a 0.1
 - **Layer Denso** (dimensione tra 32 e 4096)
- **Layer di Dropout** con rate a 0.1
- Layer di Output con attivazione SoftMax

Decisi con
AutoML

Classificatore basato su Autoencoder

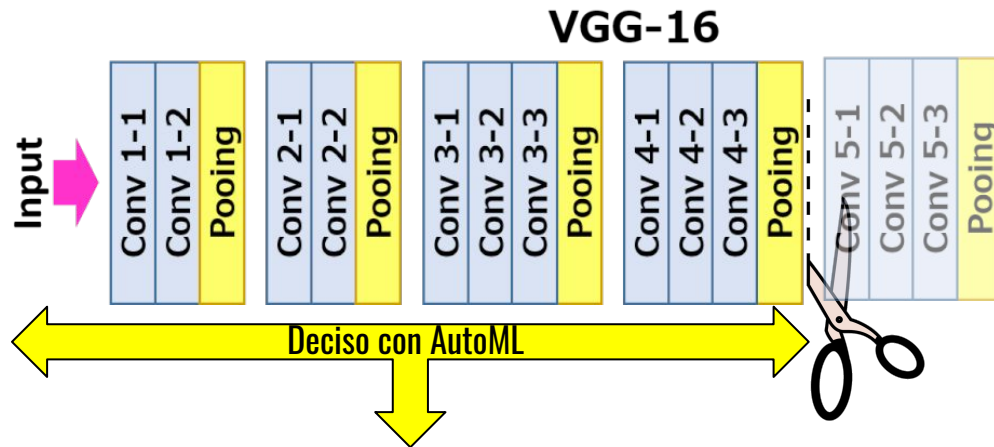
- **Encoder Convoluzionale** (Output $7 \times 7 \times 512$)
- **Layer di Flatten** (25,048 valori di output)
- Da 1 a 4 coppie di layer:
 - **Layer di Dropout** con rate a 0.1
 - **Layer Denso** (dimensione tra 32 e 4096)
- **Layer di Dropout** con rate a 0.1
- **Layer di Output** con attivazione SoftMax

Decisi con
AutoML



Fattore di compressione
7.13

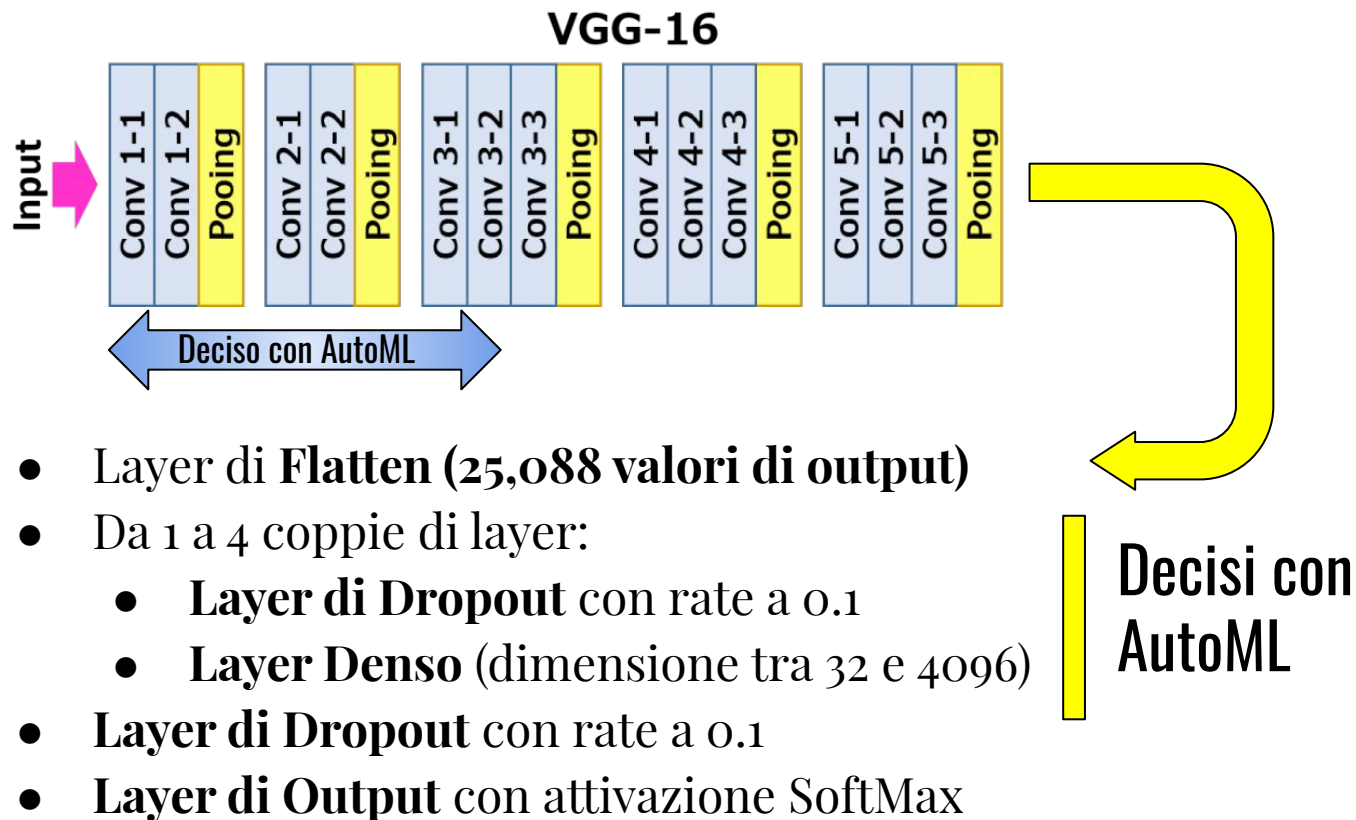
Classificatore Convoluzionale basato su Feature Extractor



- Layer di **Global Average Pooling**
- Da 1 a 4 coppie di layer:
 - **Layer di Dropout** con rate a 0.1
 - **Layer Denso** (dimensione tra 32 e 4096)
- **Layer di Dropout** con rate a 0.1
- **Layer di Output** con attivazione SoftMax

**Decisi con
AutoML**

Classificatore Convoluzionale basato su Fine Tuning

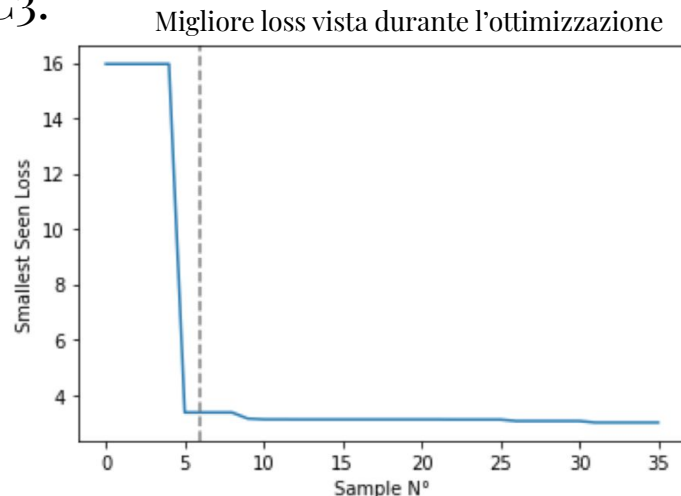
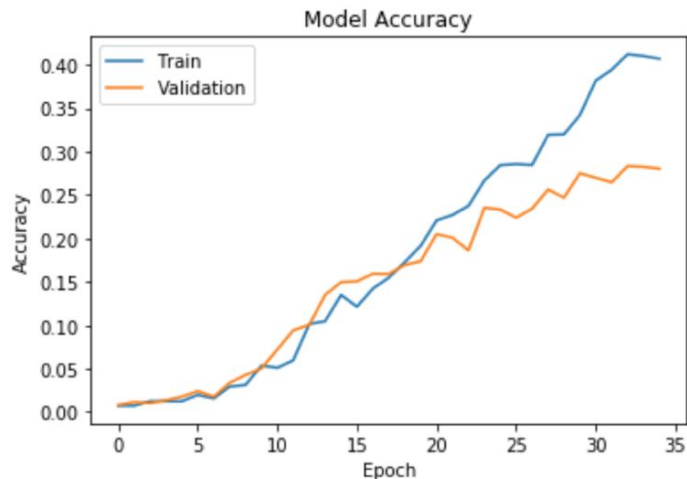


Risultati - Classificatore Convolutionale Standard

- **Configurazione ottimale** individuata da SMAC3:

- N. Layer Densi: 3
- Dimensione layer: 395 / 74 / 414
- Learning Rate Adam: 0.000306

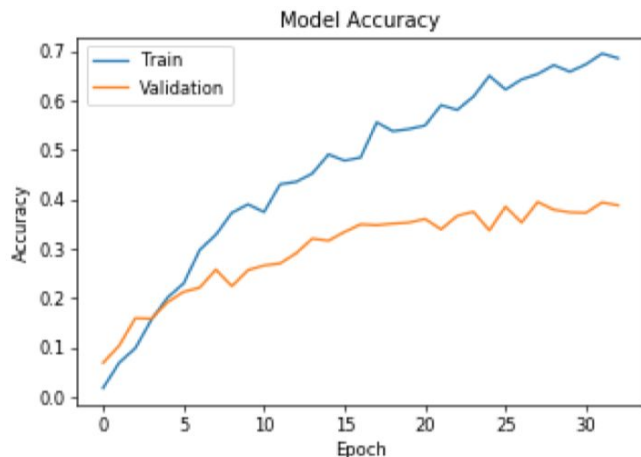
- **13,923,623** parametri addestrabili



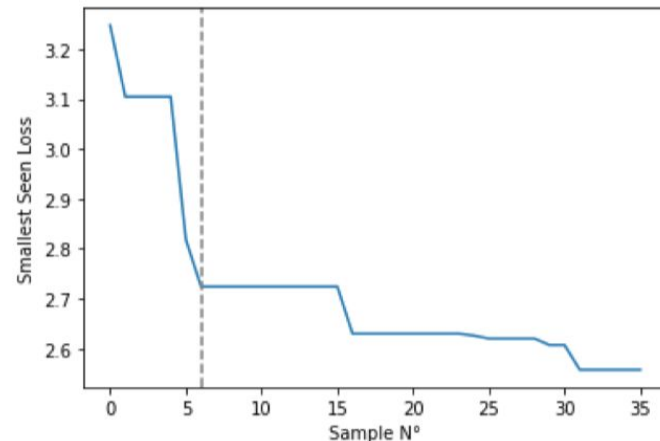
	Accuracy	Top-5 Accuracy	Loss
Training	0.407	0.7772	2.024
Validation	0.280	0.578	3.048
Testing	0.221	0.502	3.337

Risultati - Classificatore Conv. basato su Autoencoder

- **Configurazione ottimale** individuata da SMAC₃:
 - N. Layer Densi: **1**
 - Dimensione 1° layer: **1138**
 - Learning Rate Adam: **0.00101**
- **30,236,036** parametri (28, 667, 460 addestrabili)



Migliore loss vista durante l'ottimizzazione



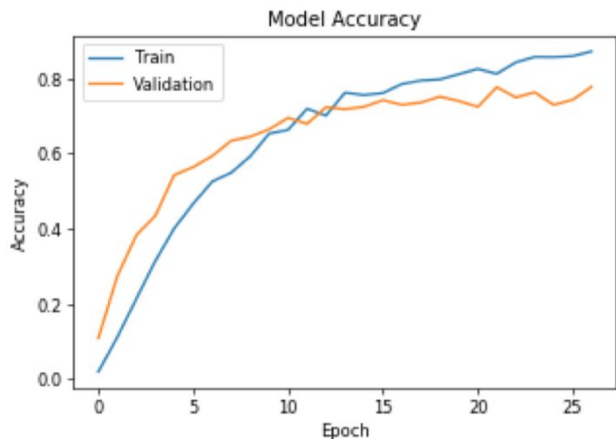
	Accuracy	Top-5 Accuracy	Loss
Training	0.686	0.929	1.192
Validation	0.388	0.670	2.578
Testing	0.331	0.619	2.889

Risultati - Classificatore Conv. basato su Feature Extractor

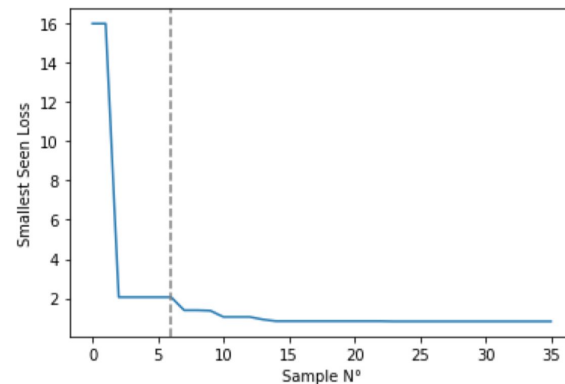
- **Configurazione ottimale** individuata da SMAC3:

- **Blocco convoluzionale VGG16:** 5°
- N. Layer Densi: 3
- Dimensione layer: **1128 / 3224 / 1994**
- Learning Rate Adam: **0.0000406**

- **28,261,765** parametri (13, 547, 077 addestrabili)



Migliore loss vista durante l'ottimizzazione



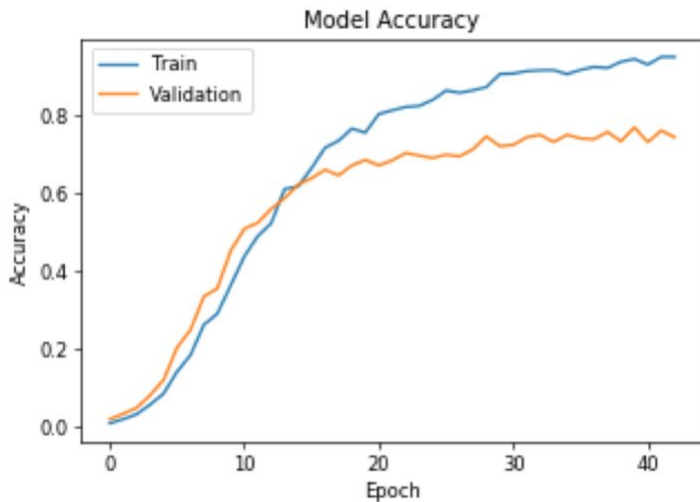
	Accuracy	Top-5 Accuracy	Loss
Training	0.871	0.985	0.512
Validation	0.777	0.934	0.877
Testing	0.727	0.903	1.088

Risultati - Classificatore Conv. basato su Fine Tuning

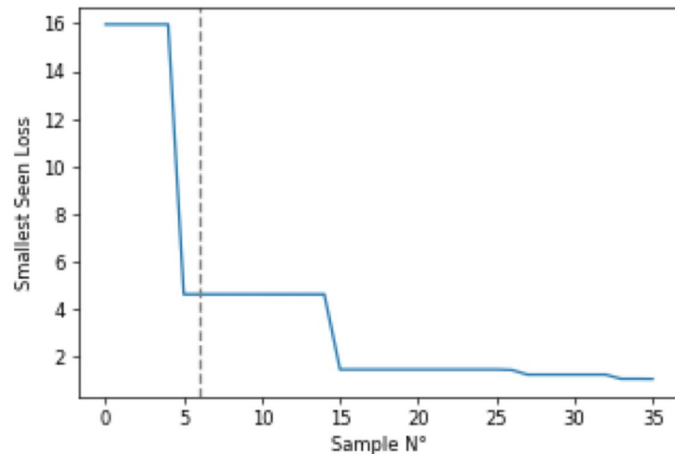
- **Configurazione ottimale** individuata da SMAC₃:

- N° Ultimi Layer Addestrabili VGG16: 2
- N. Layer Densi: 4
- Dimensione layer: 520, 1753, 54, 1344
- Learning Rate Adam: 0.0000679

- **28,980,107** parametri (16, 625, 227 addestrabili)



Migliore loss vista durante l'ottimizzazione



	Accuracy	Top-5 Accuracy	Loss
Training	0.949	0.996	0.166
Validation	0.743	0.918	1.184
Testing	0.727	0.901	1.245

Risultati - Modello Migliore

Modello basato su Feature Extraction da VGG16

- **Accuracy: 72.6%**
- **Top-5 Accuracy: 90.3%**
- **Loss: 1.088**
- **Precision: 75.6%** (76.4% pesata)
- **Recall: 73.4%** (72.6% pesata)
- **F1-score: 72.7%** (72.4% pesata)



Confusione tra specie

Snapdragon
Identified: 31 times (73.81%)



Mistaken as Petunia
5 times (11.90%)



Mistaken as Foxglove
3 times (7.14%)



Mistaken as Garden Phlox
2 times (4.76%)



Garden Phlox
Identified: 18 times (18.56%)



Mistaken as Azalea
11 times (11.34%)



Lotus
Identified: 65 times (79.27%)



Mistaken as Water Lily
7 times (8.54%)



Per concludere...

- Procedura di AutoML efficace per individuare punti ottimali
 - In genere **meno di 20 campionamenti** necessari
- Risultati soddisfacenti
 - Modello migliore basato su **Feature Extraction da VGG**
 - **Accuracy** migliore del **72%**
 - **Top-5 Accuracy** migliore del **90%**
- ...ma **migliorabili**
 - Con più dati di training
 - Con classificatori migliori
 - In letteratura performance migliori
- Principali **errori**
 - Confusione di specie simili tra loro

Fonte	Approccio	Accuracy	Top-5
Nilsback, 2008	SIFT/HOG/HSV	73% (*)	-
Sharif, 2014	Overfeat+SVM	87%	-
Gurnani, 2017	GoogleNet	47% (*)	69%
Nostro	VGG16 F.E.	72%	90%

Grazie per l'attenzione
