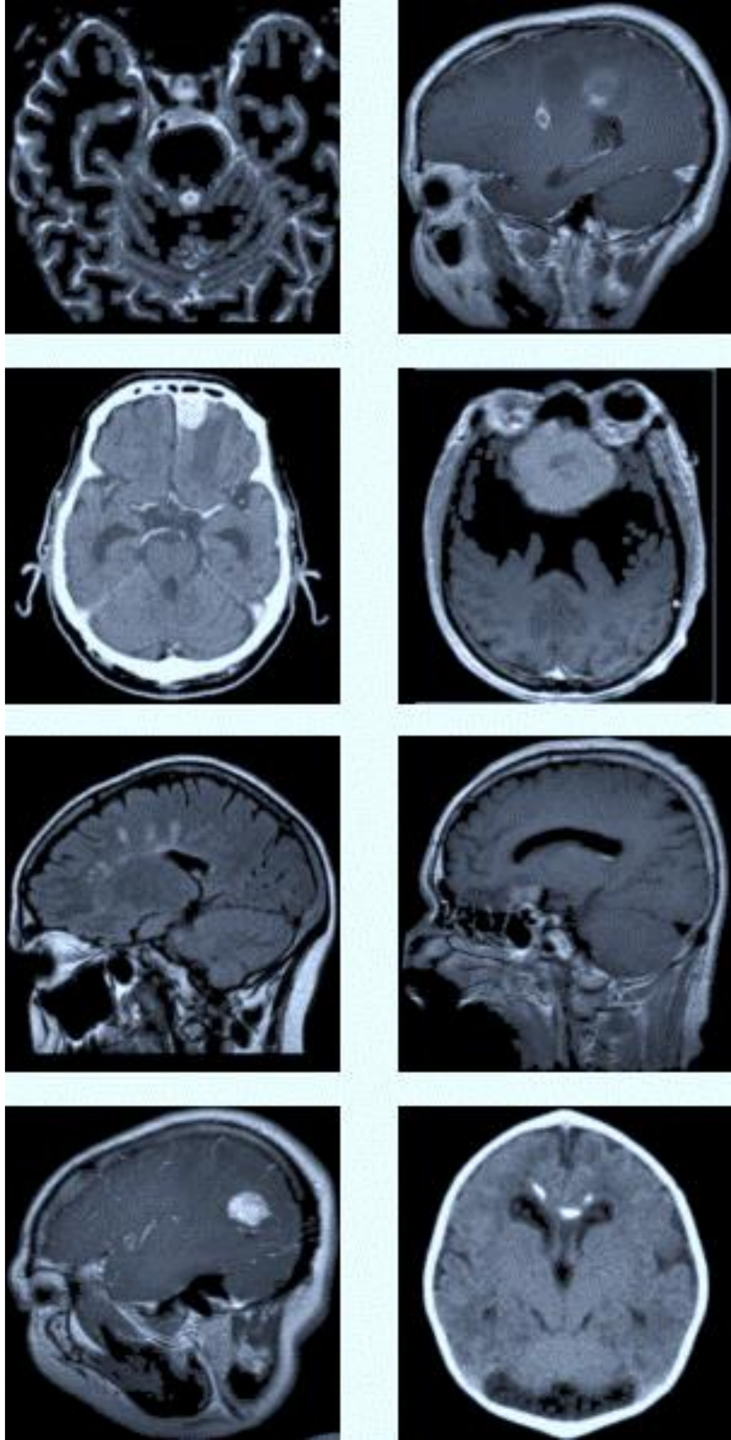
Image Classification of Brain Tumors

Corso di Pattern Recognition e Machine Learning

- **Simone Catenacci**
- **Marco Salvatori**
- **Simone Sorgonà**

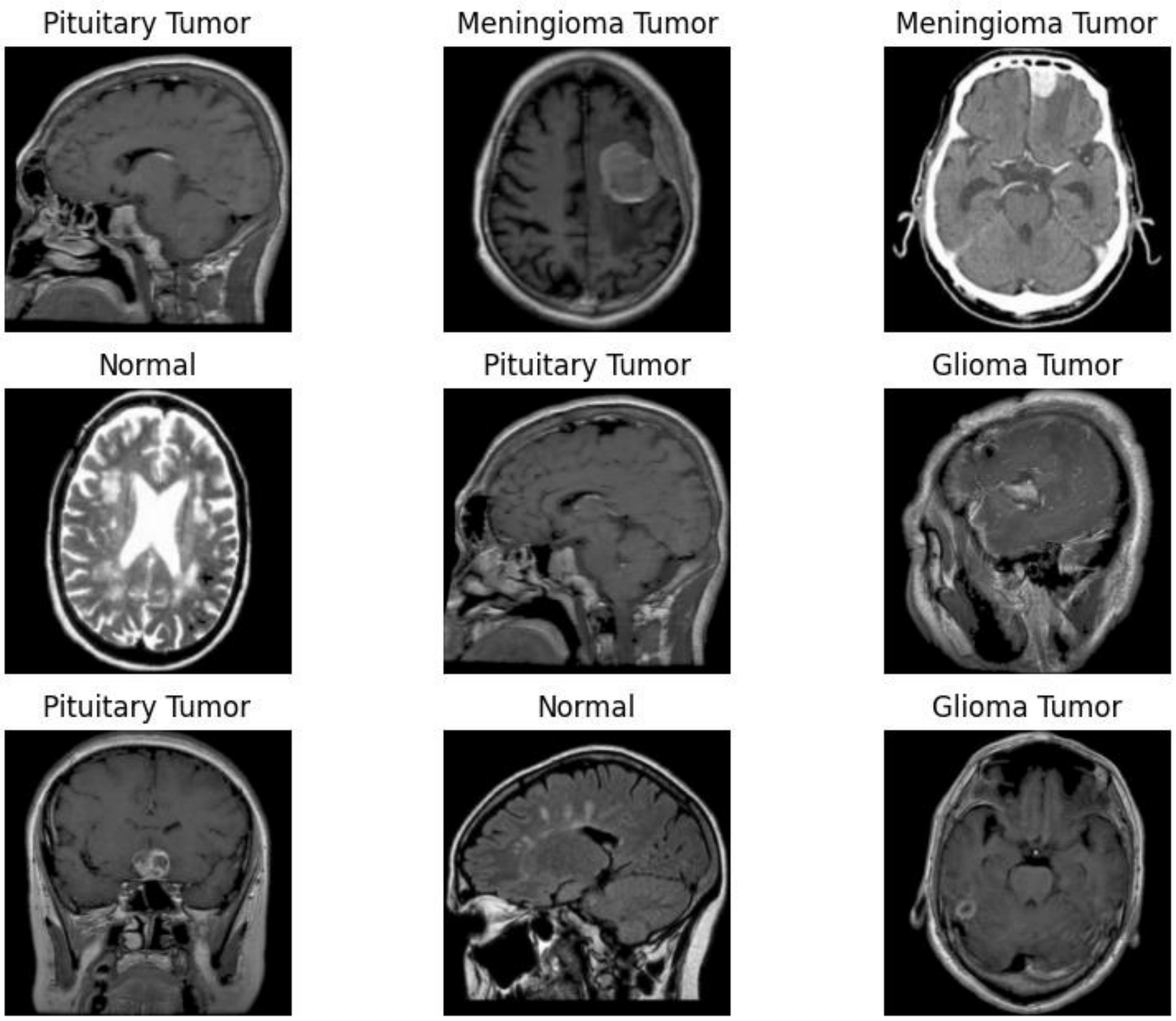
A.A. 2023/2024



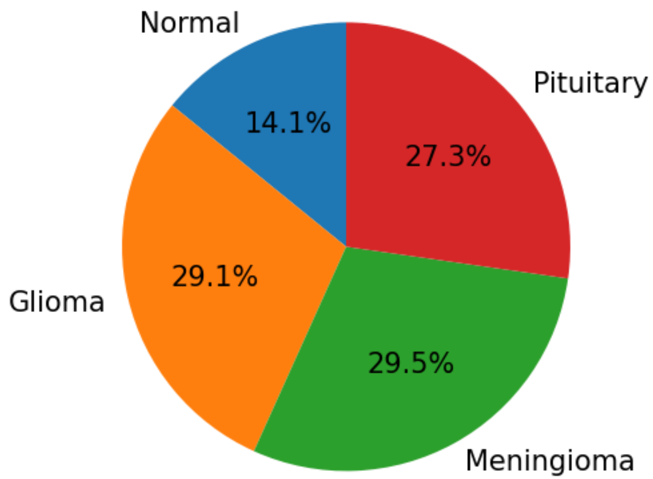
Sommario

- Analisi Dataset
- Processing Dataset
- Implementazione
- Analisi dei Risultati
- Conclusioni

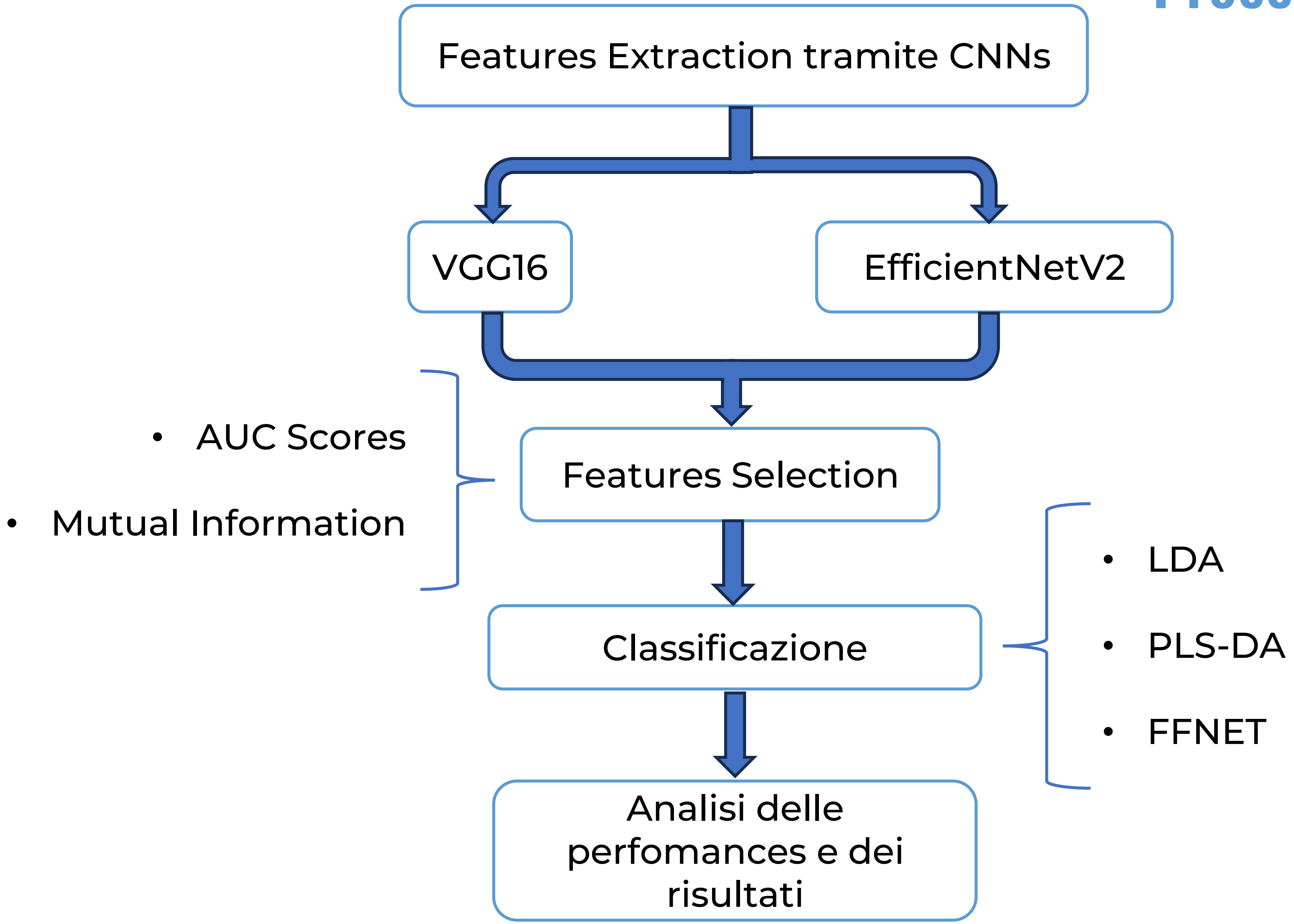
Struttura Dataset



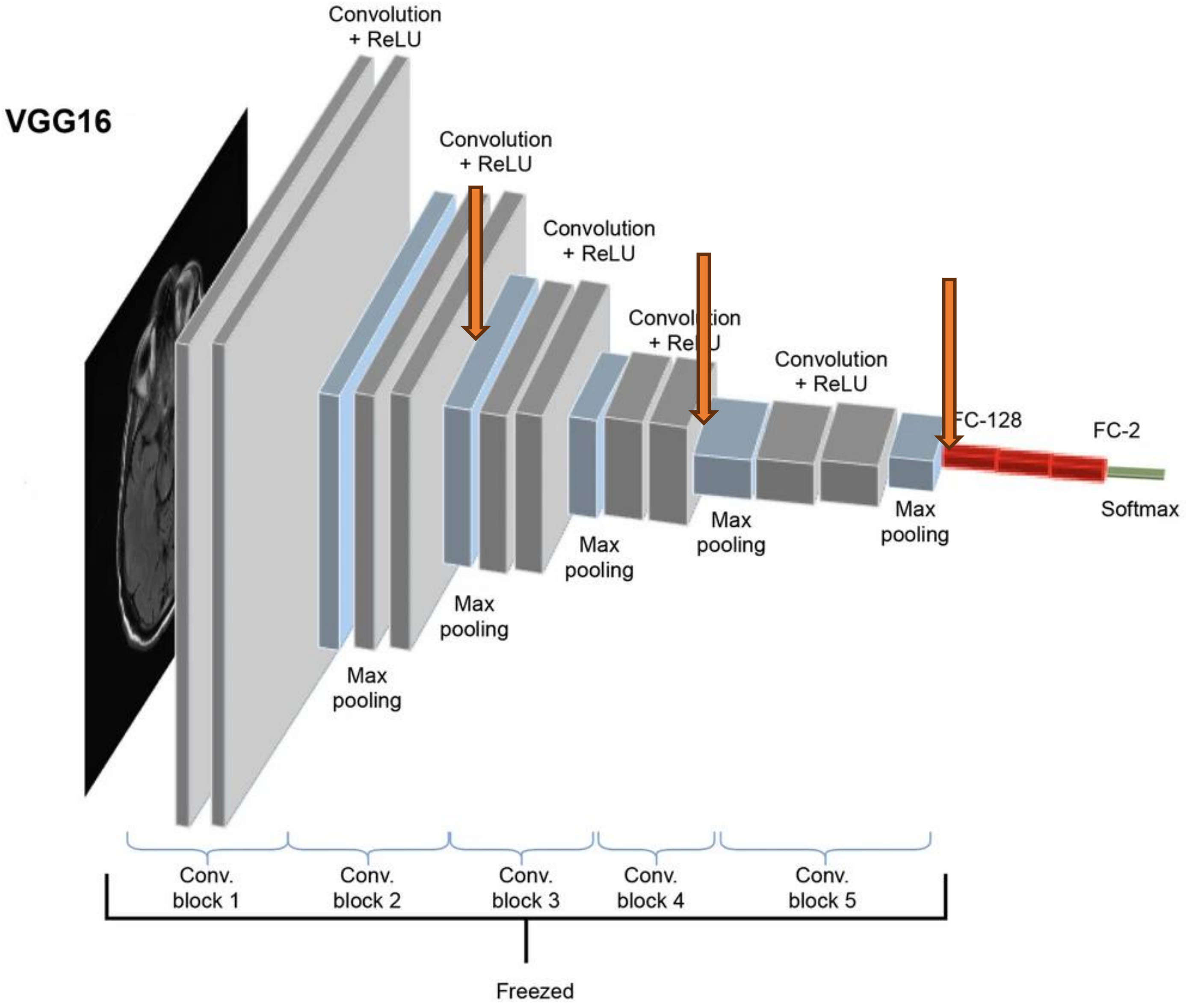
- 3096 immagini totali
- 4 classi di tumori :
  - Normal
  - Glioma Tumor
  - Meningioma Tumor
  - Pituitary Tumor
- Model Validation: K-Fold Crossvalidation



Processing dei dati



Features Extraction: VGG16

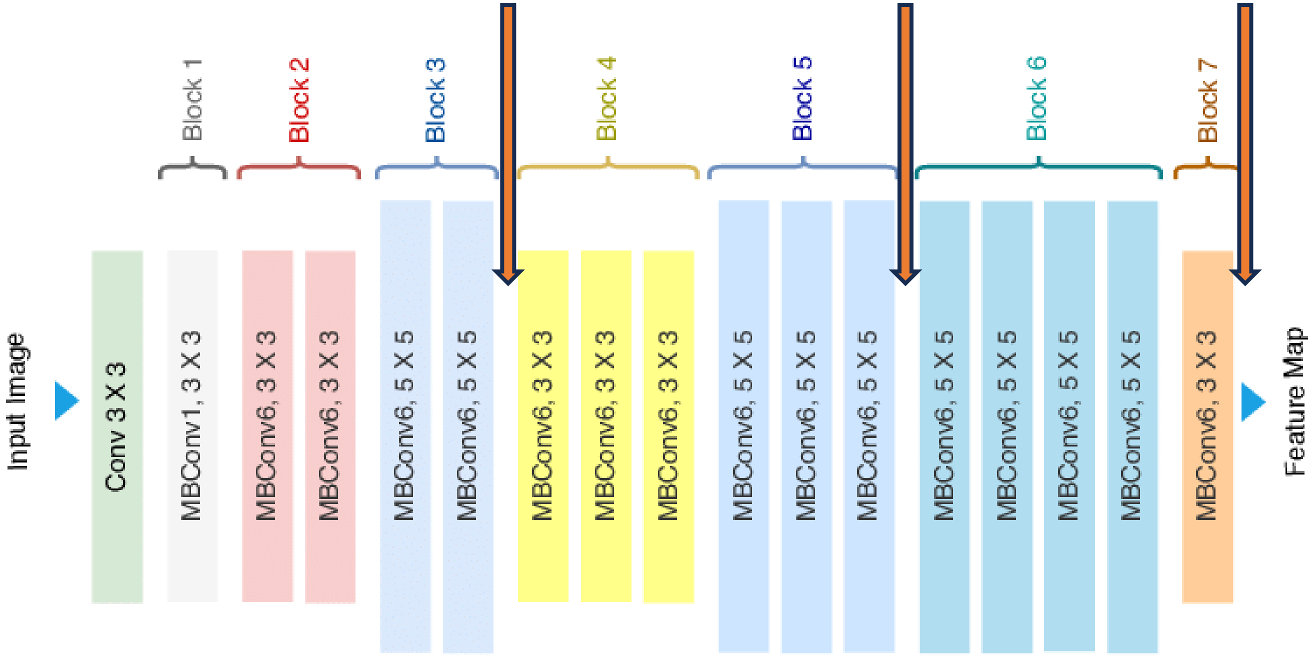


Parametri: 138.4M
Top 1 Accuracy (Imagenet): 71.3%

Uscite della CNN considerate nei layer di:

- Max pooling di conv 2 block
- Max pooling di conv 4 block
- Max pooling di conv 5 block

Features Extraction EfficientNetV2L

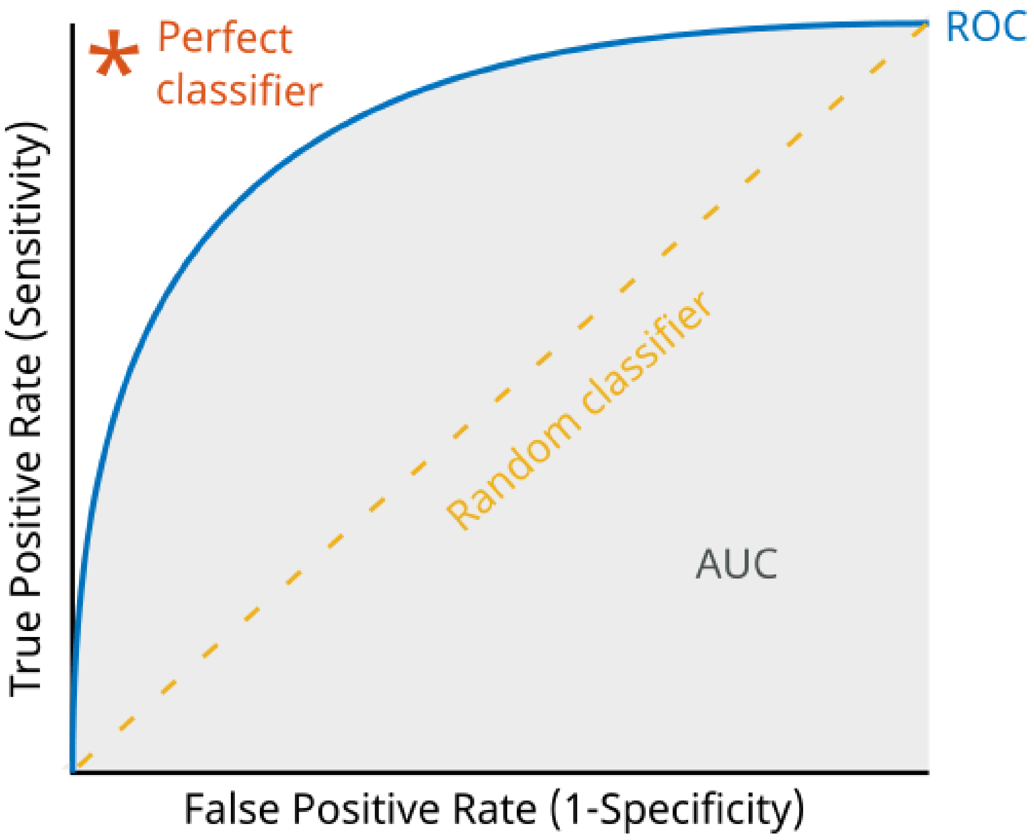


Parametri: 119M
Top 1 Accuracy (Imagenet): 85.3%

Uscite della CNN considerate nei layer di:

- Output Layer di Block 3
- Output Layer di Block 5
- Output Layer di Block 7

Features Selection



- **Area Under the ROC Curve** fornisce una misurazione aggregata del rendimento in tutte le possibili soglie di classificazione

- Valori di soglia:
    - VGG16 → ≥ 0.8
    - ENetV2 → ≥ 0.8 ∧ 0.85

- Rimozione delle features più dipendenti tramite **Mutual Information**:

$$I(X,Y) = \sum_{X} \sum_{Y} p(x,y) \log\left(\frac{p(x,y)}{p(x)p(y)}\right)$$

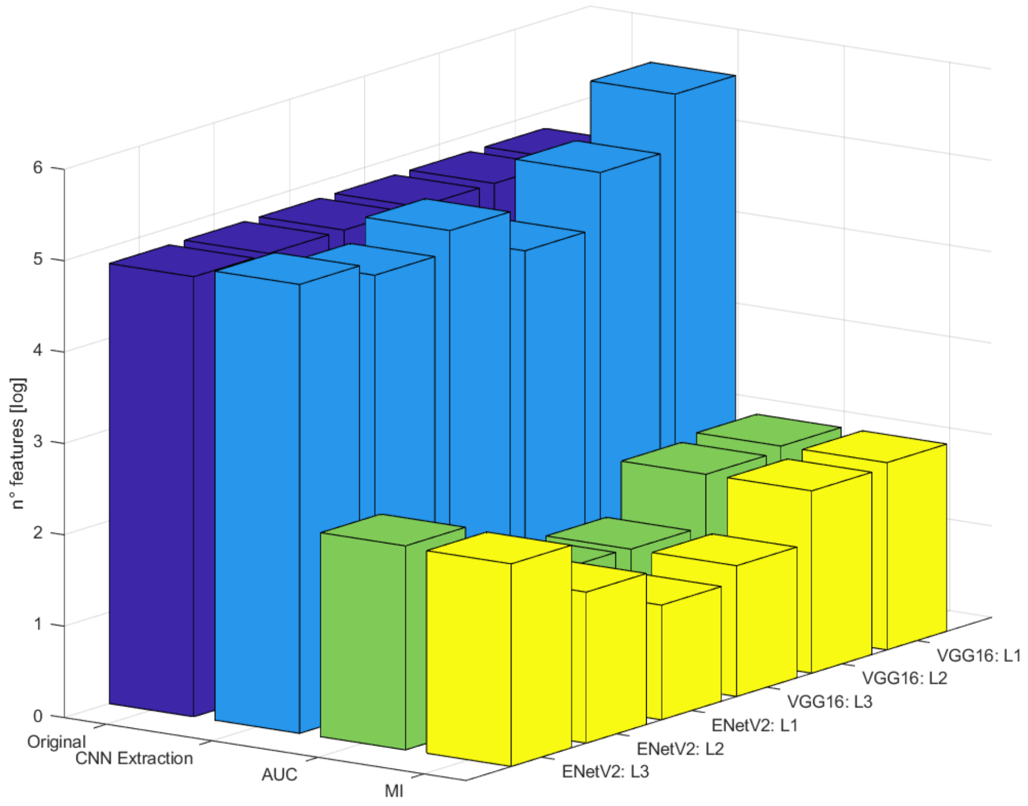
- **MI** è uguale a 0 se X e Y sono totalmente indipendenti (valori alti significa alta dipendenza)

Features Selection

Numero di features di un'immagine: 65 536

Numero di features per le diverse feature extraction e selection:

	Layer	CNN Extraction	AUC-ROC	MI
VGG16	Layer 1	524 288	112	112
	Layer 2	131 072	99	99
	Layer 3	32 768	27	27
ENetV2	Layer 1	98 304	28	18
	Layer 2	57 344	46	45
	Layer 3	81 920	171	165



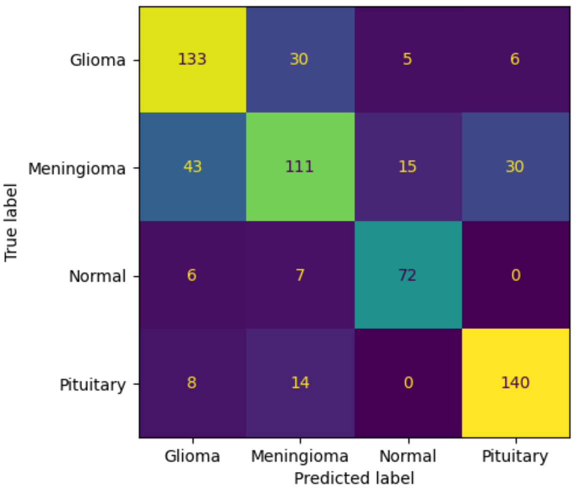
- **LDA** (**Linear Discriminant Analysis**) è una tecnica di classificazione lineare o di dimensionality reduction.

- LDA può essere descritta da modelli probabilistici che sfruttano la probabilità condizionale. La classe predetta è ottenuta utilizzando le regole di Bayes per ogni osservazione *x*.
  Si seleziona la classe *k* che massimizza la seguente probabilità a posteriori:

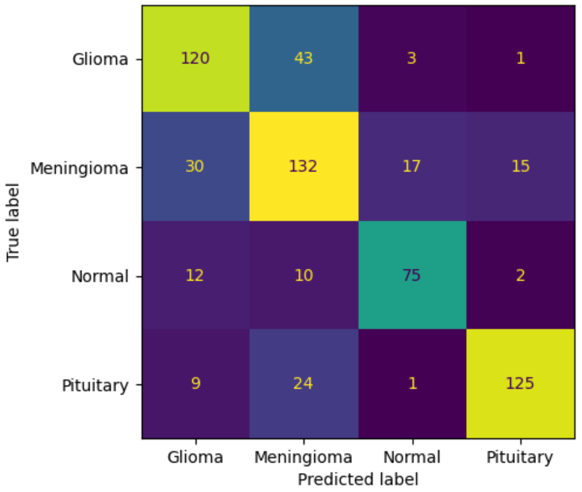
$$P(y = k|x) = \frac{P(x|y = k)P(y = k)}{P(x)} = \frac{P(x|y = k)P(y = k)}{\sum_l P(x|y = l)P(y = l)}$$

LDA: VGG16

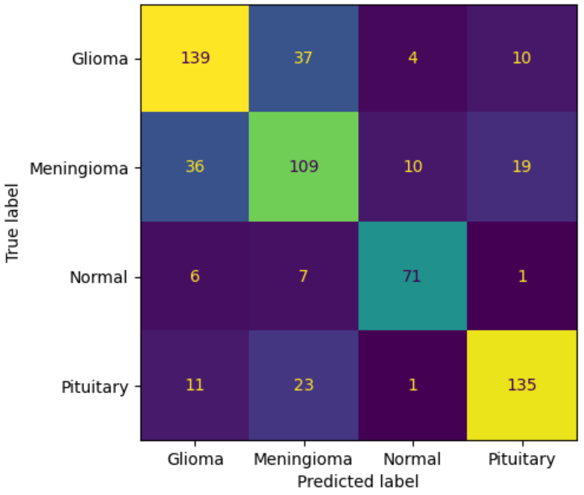
Val. Accuracy = 73.55%



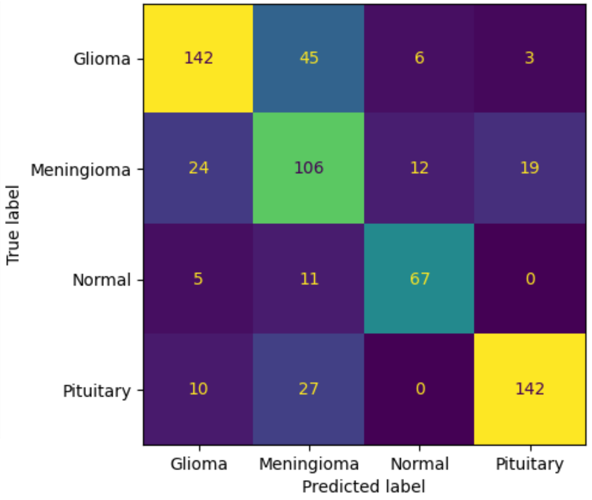
Val. Accuracy = 73.02%

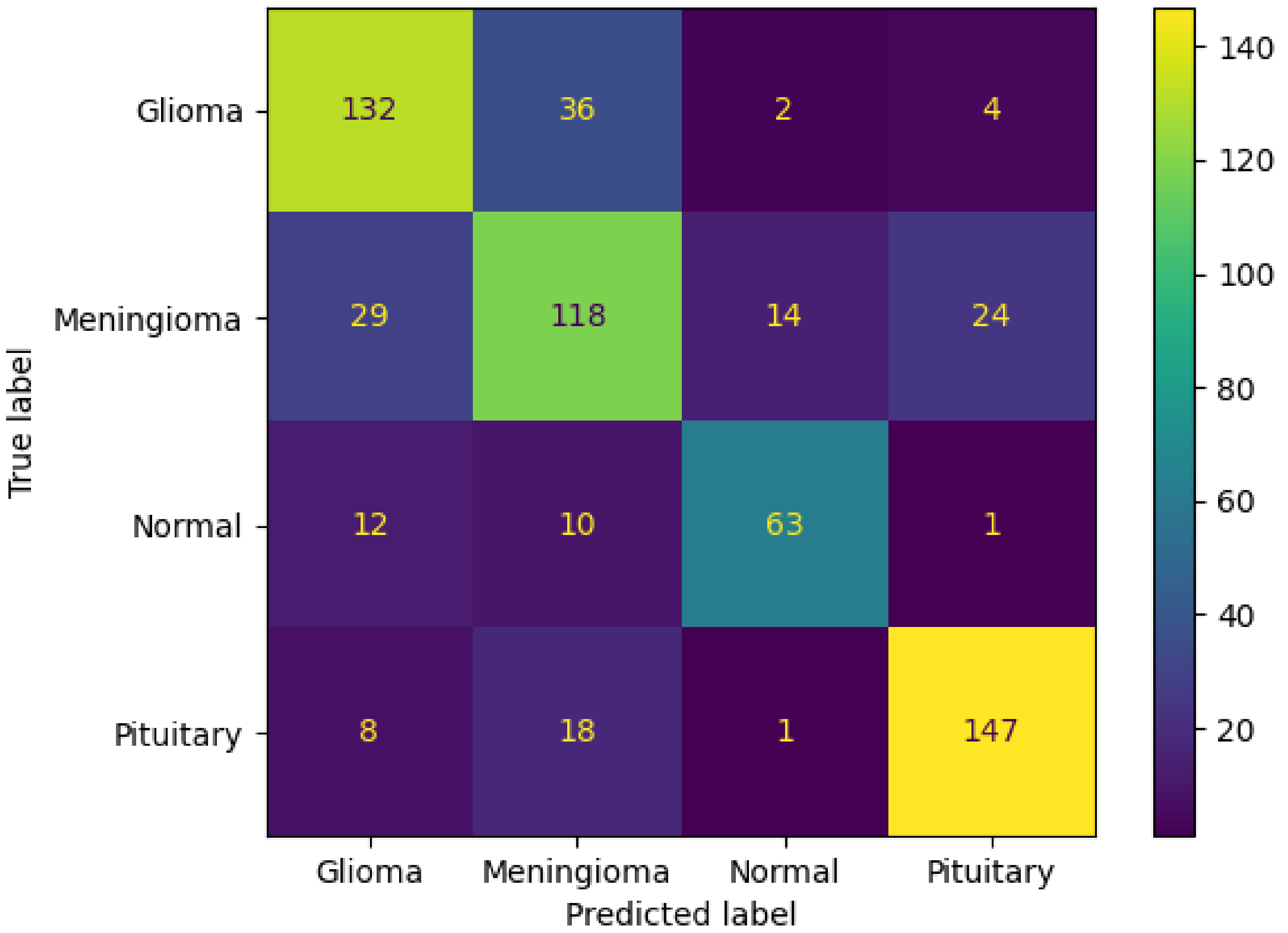


Val. Accuracy =73.34%



Val. Accuracy = 73.83%

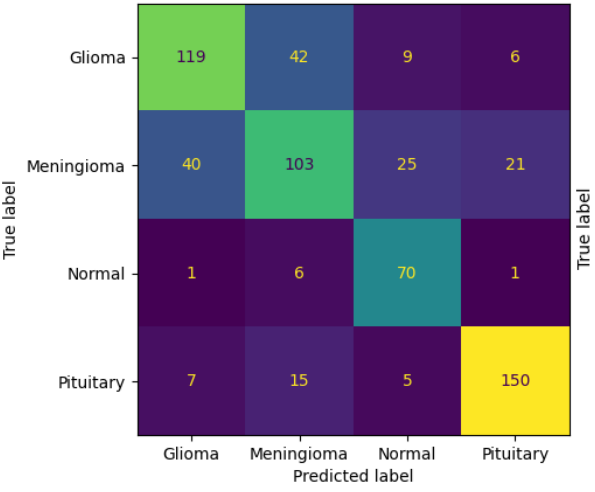




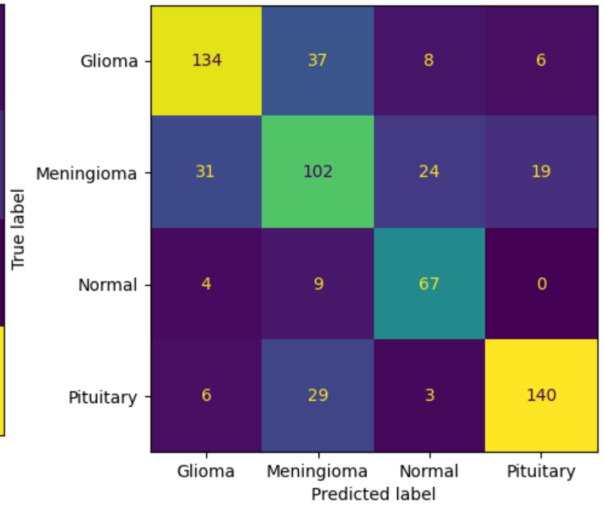
Val. Accuracy = 74.31%

LDA: EfficientNetV2

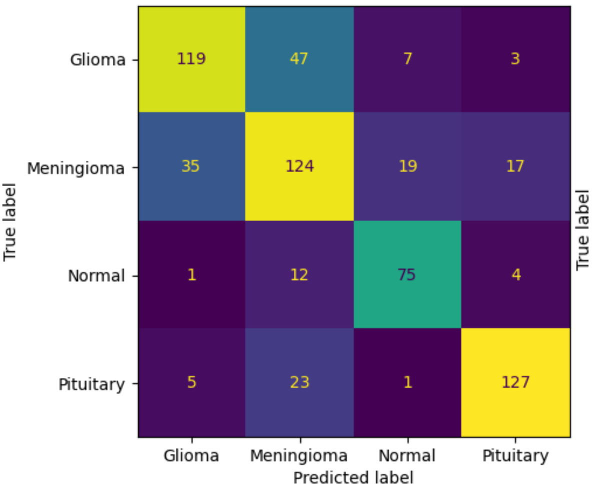
Val. Accuracy = 71.29%



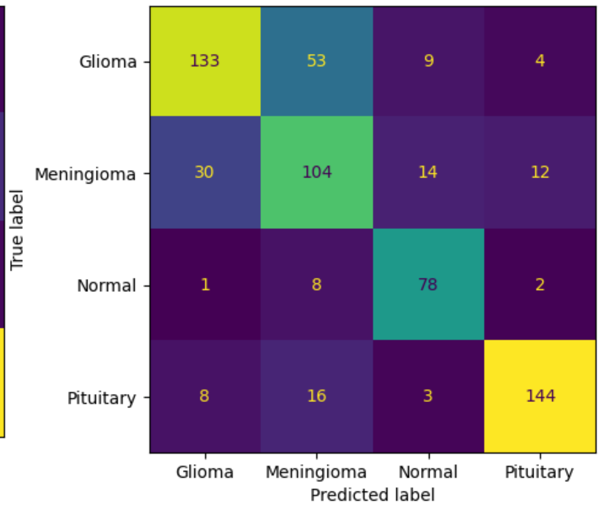
Val. Accuracy = 71.57%

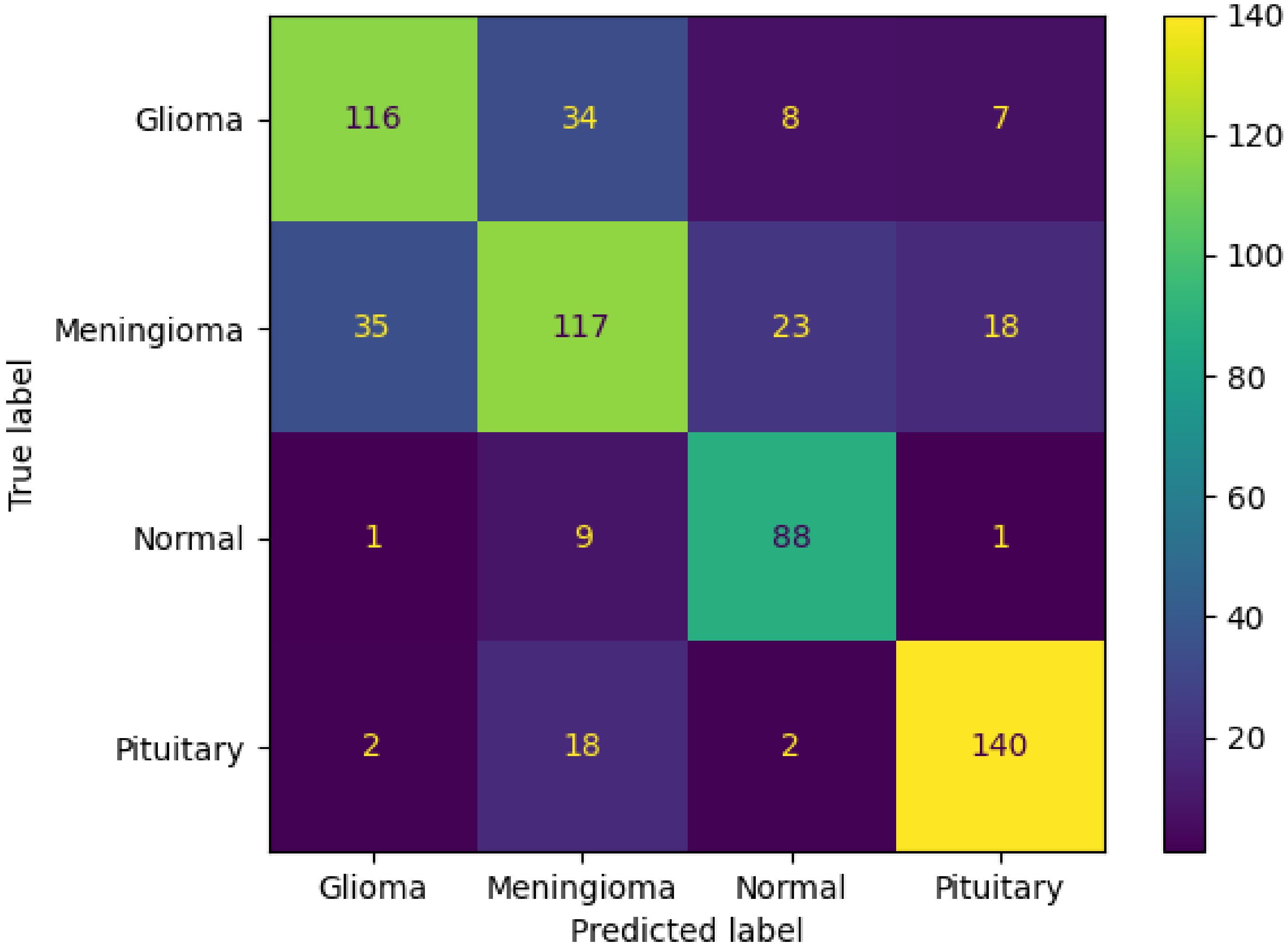


Val. Accuracy = 71.89%



Val. Accuracy = 74.15%





Val. Accuracy = 74.48%

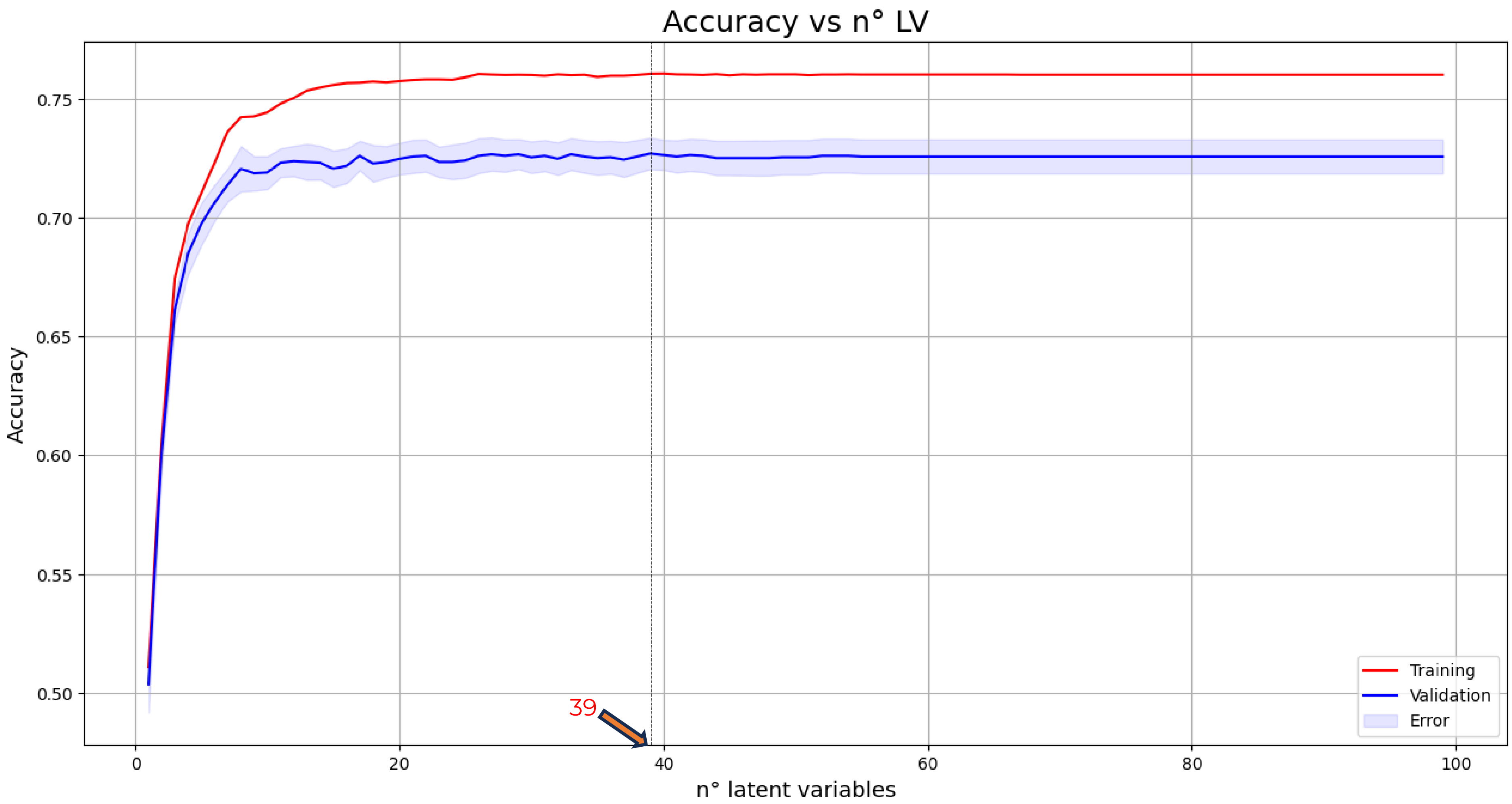
PLS-DA

- La **PLS Regression** trova un modello di regressione lineare, proiettando le variabili X e Y in un nuovo spazio, le nuove componenti sono denominate *variabili latenti* (LV) e sono determinate massimizzando la covarianza delle score matrix T e U, sottospazi di X e Y.

- **PLS-DA** (**Partial Least Square Discriminant Analysis**) è un metodo di classificazione basato su una PLS Regression in cui le uscite sono di tipo categorico.

PLS-DA: VGG16

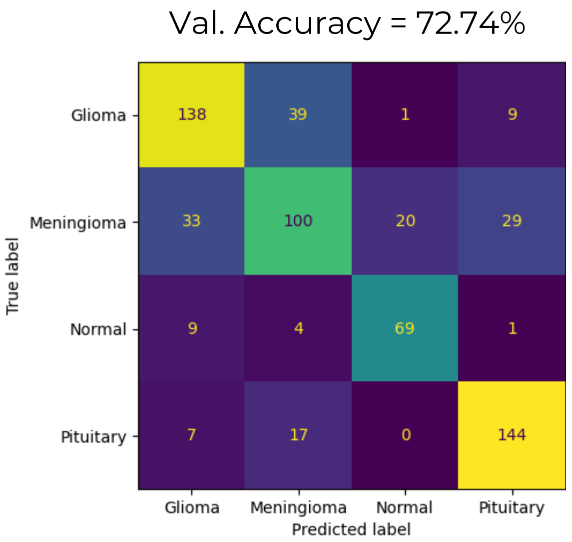
Accuracy media delle k crossvalidazione al variare del numero di variabili latenti
(Best case → VGG16 layer 2)

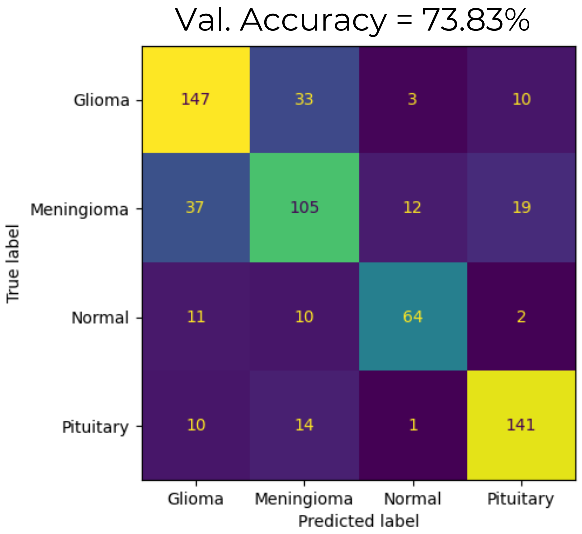


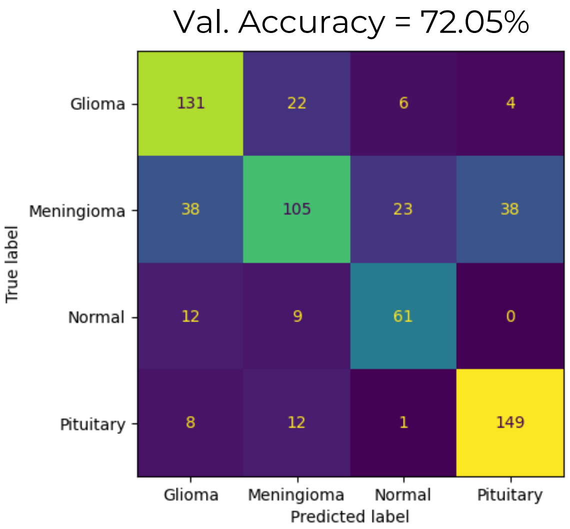
Best Acc. Mean = 72.71%

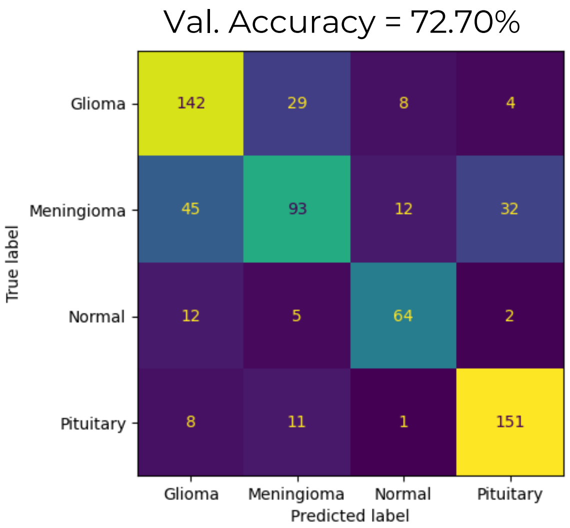
Std = 1.33%

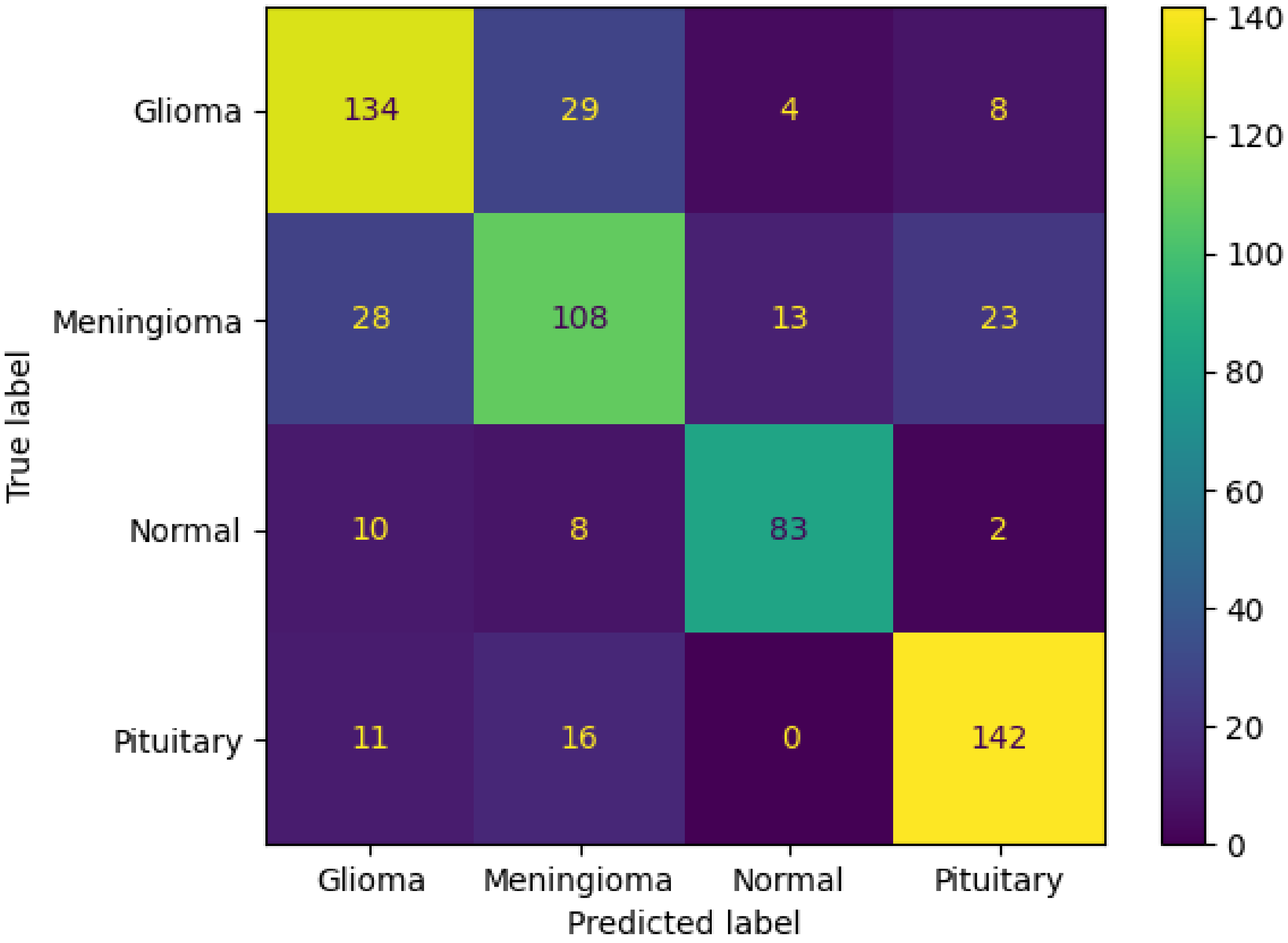
PLS-DA: VGG16







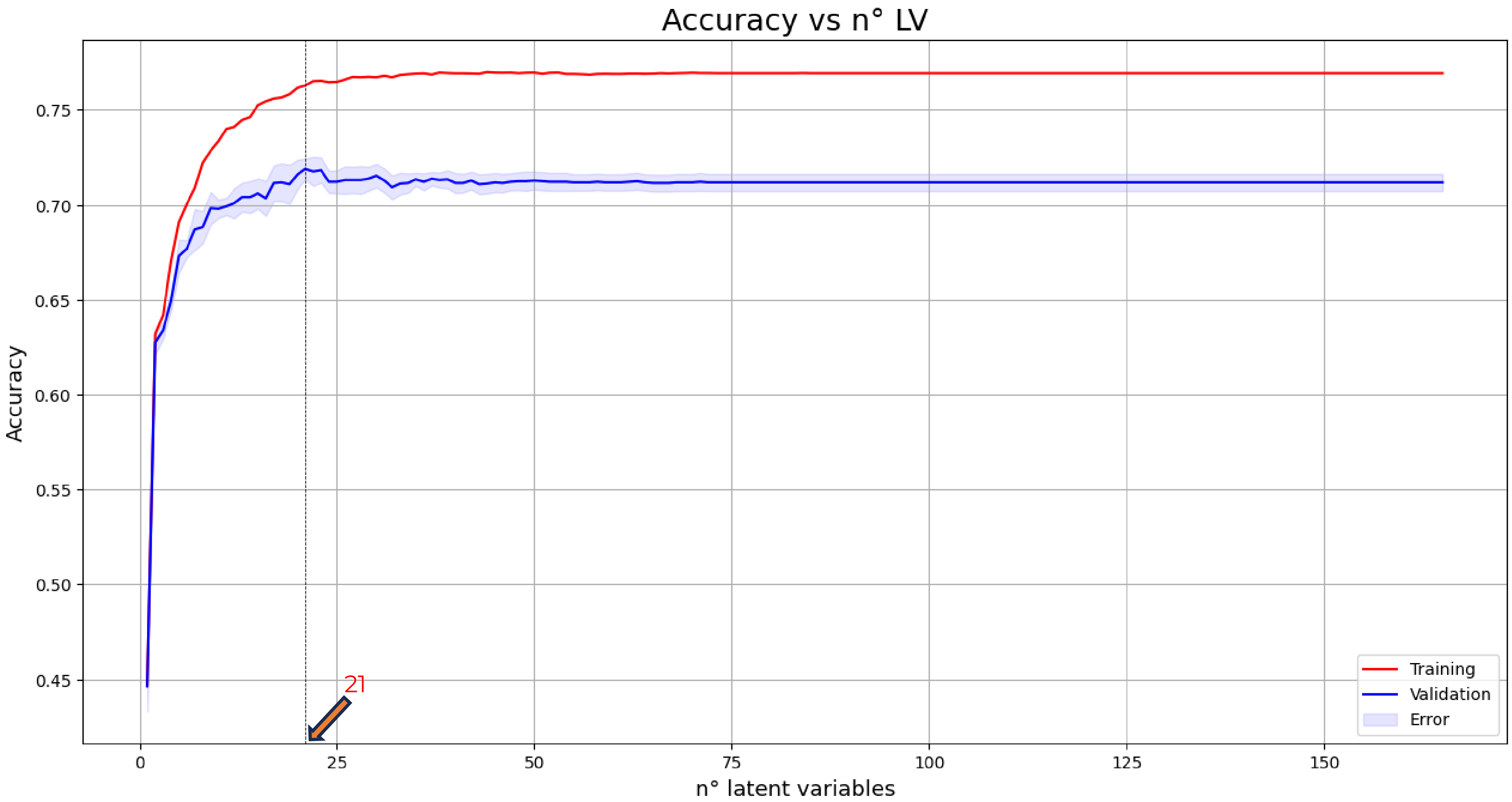




Val. Accuracy = 75.44%

PLS-DA: EfficientNetV2

Accuracy media delle k crossvalidazione al variare del numero di variabili latenti
(Best case → ENetV2 layer 3)

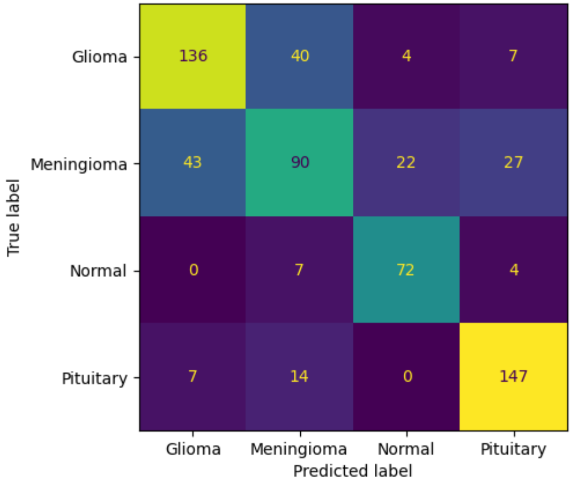


Best Acc. Mean = 71.90%

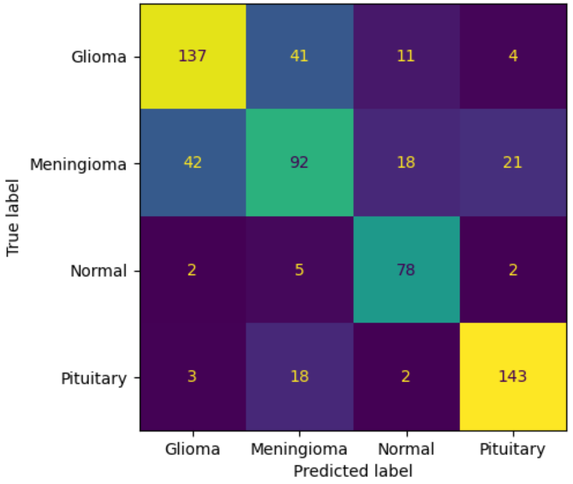
Std = 1.09%

PLS-DA: EfficientNetV2

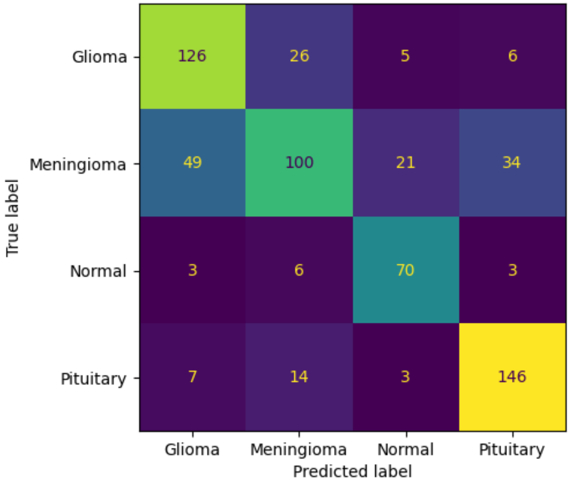
Val. Accuracy = 71.77%



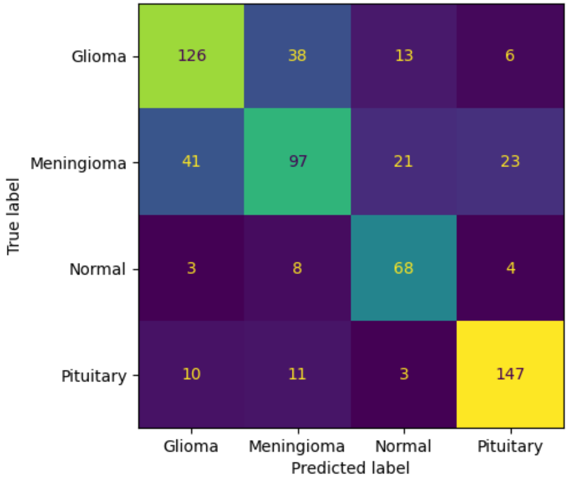
Val. Accuracy = 72.70%

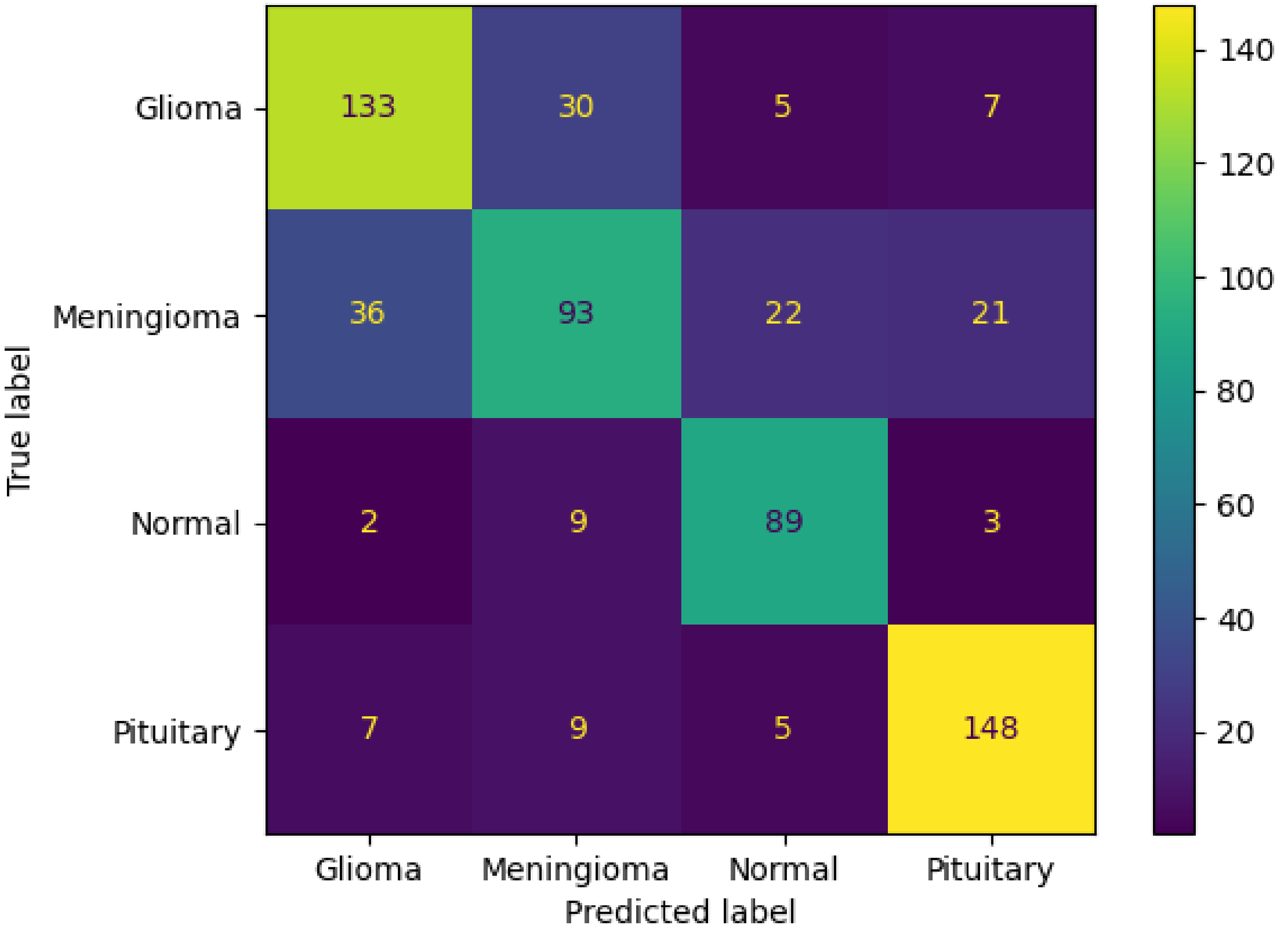


Val. Accuracy = 71.41%



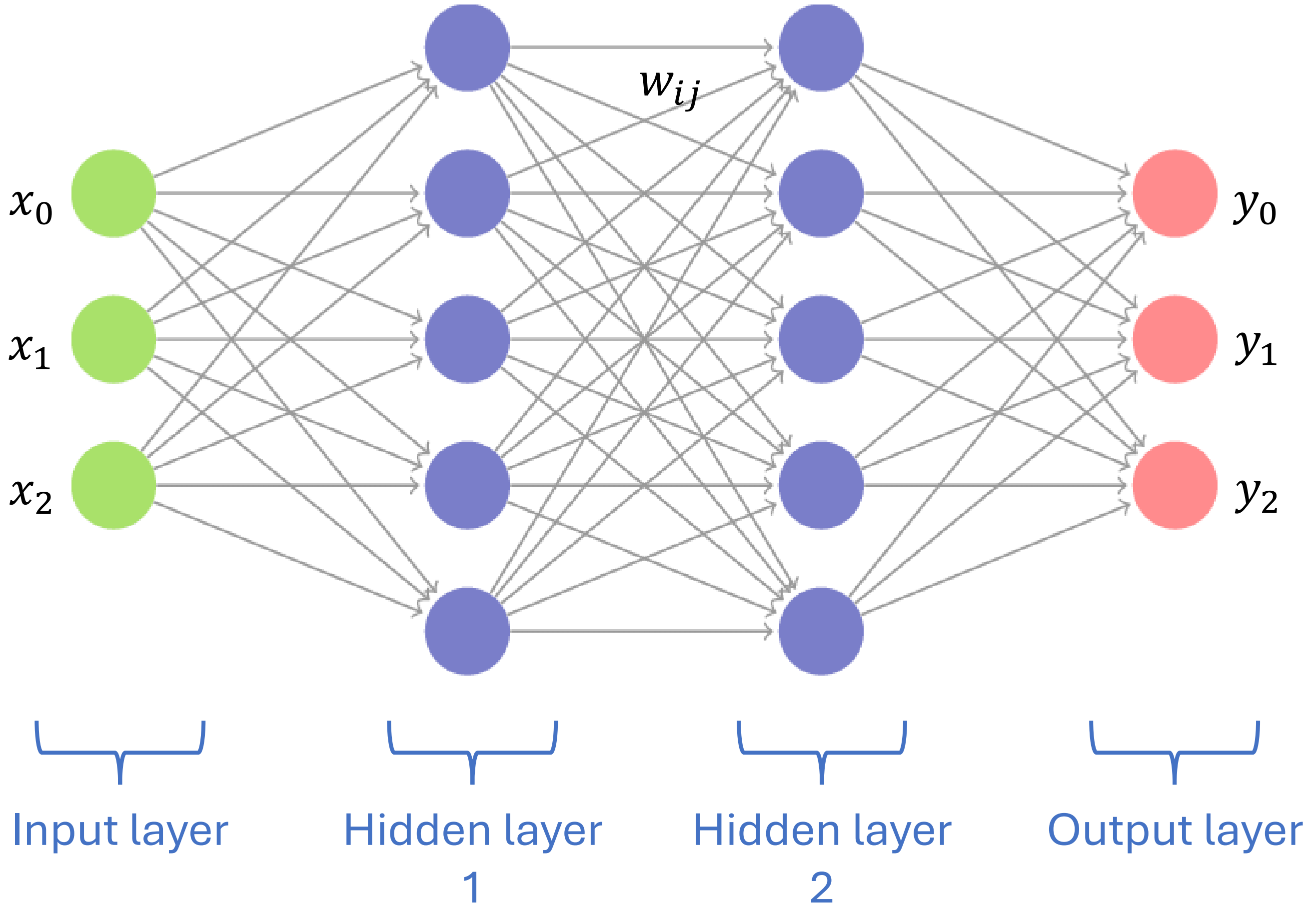
Val. Accuracy = 70.76%





Val. Accuracy = 74.80%

FeedForward Neural Network



- Nella **FeedForward Neural Network**, l'informazione si muove in avanti ed è composta dai seguenti layer:
  - Input layer
  - Hidden layer
  - Output layer

FFNET: Hyperparameter Tuning

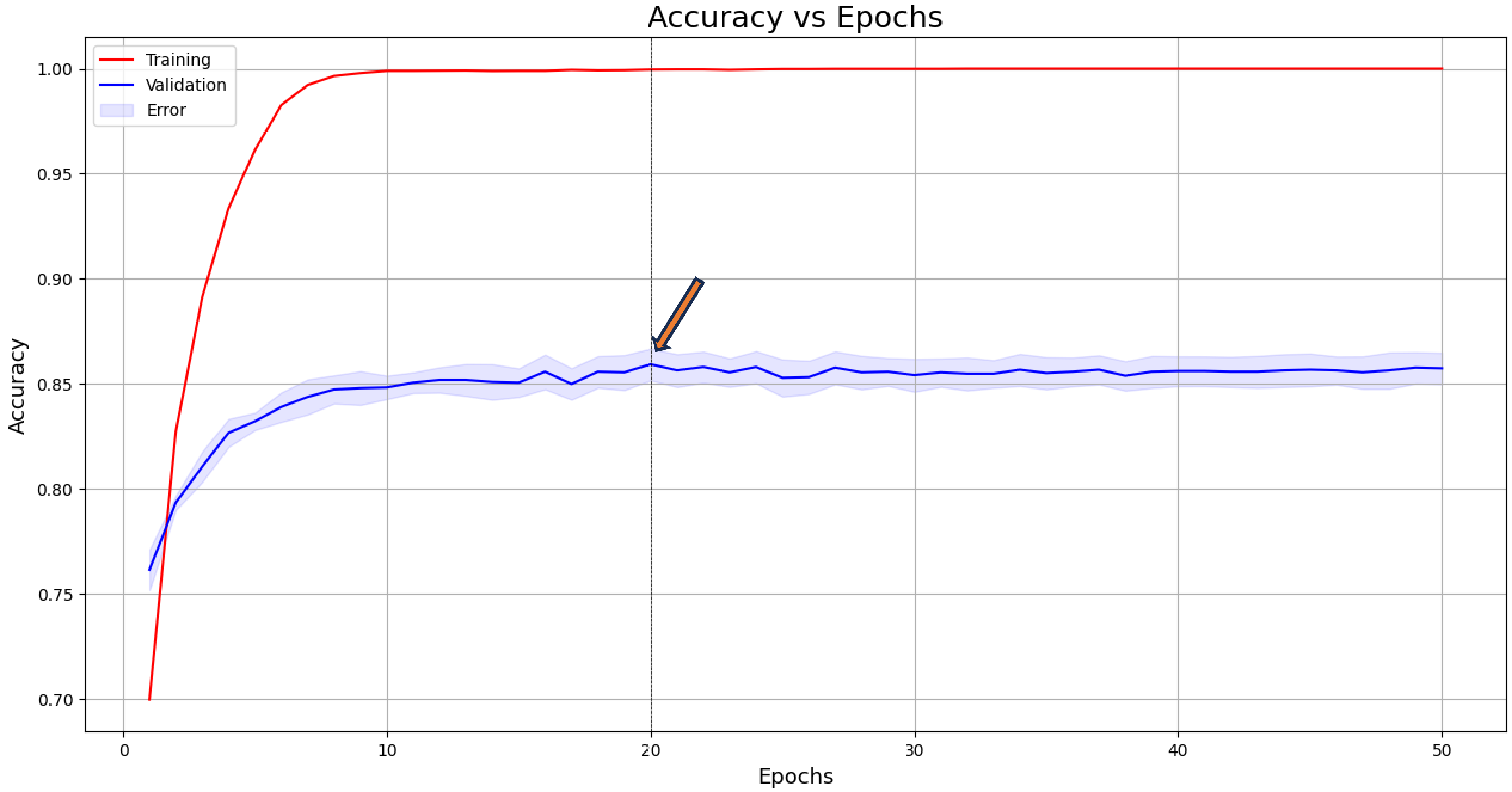
- Algoritmo di hyperparameter tuning → HyperBand

- Parametri ottimizzati:
  - Numero di Hidden Layers: 1 - 5
  - Neuroni per Hidden Layers: 50-500 con step 5
  - Learning Rate: 10^{-4} - 10^{-2} con log sampling

- Parametri fissi:
  - Neuroni Output Layer: 4
  - Funzione di attivazione: ReLU (hidden), softmax (output)
  - Optimizer: Adam

FFNET: VGG16

Accuracy media delle k crossvalidazione al variare delle epoche
(Best case → VGG16 layer 2)



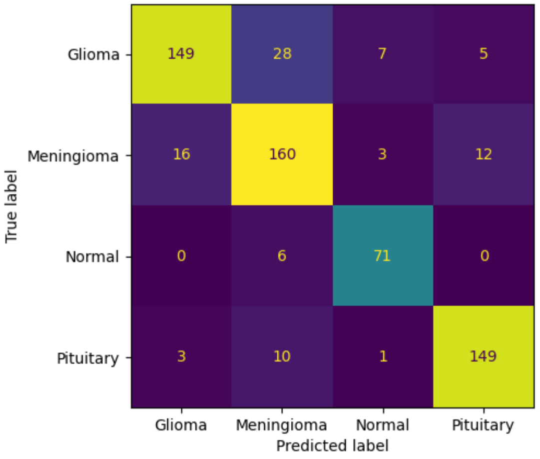
Best Acc. Mean = 85.92%

Std = 1.52%

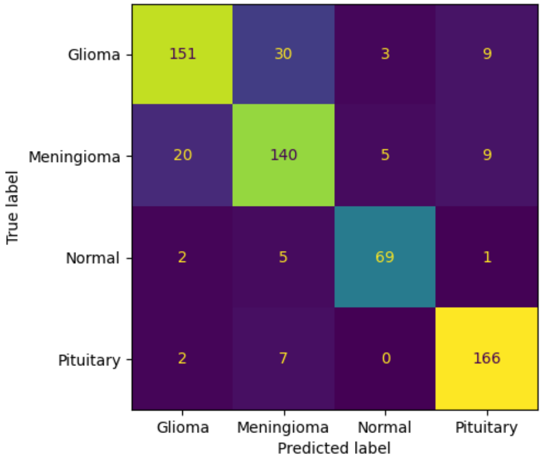
Hyperparameters:
- N° layer: 1
- Neuroni: 435
- Learn. Rate: $1.87 \cdot 10^{-3}$

FFNET: VGG16

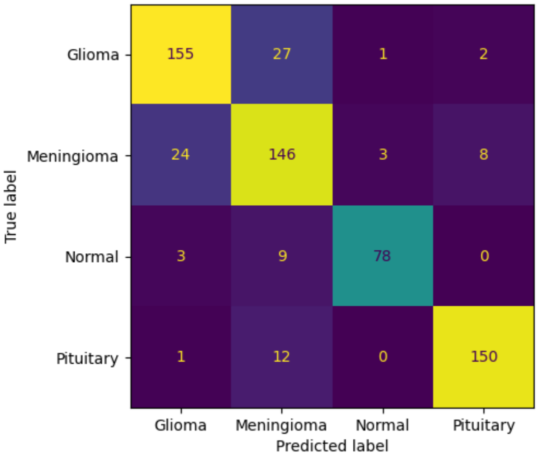
Val. Accuracy = 85.32%



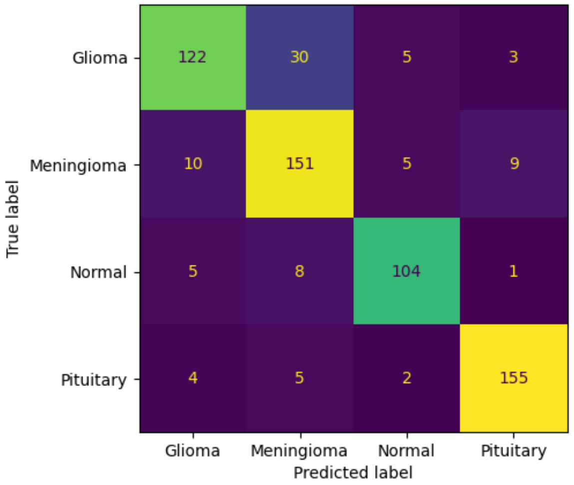
Val. Accuracy = 84.98%

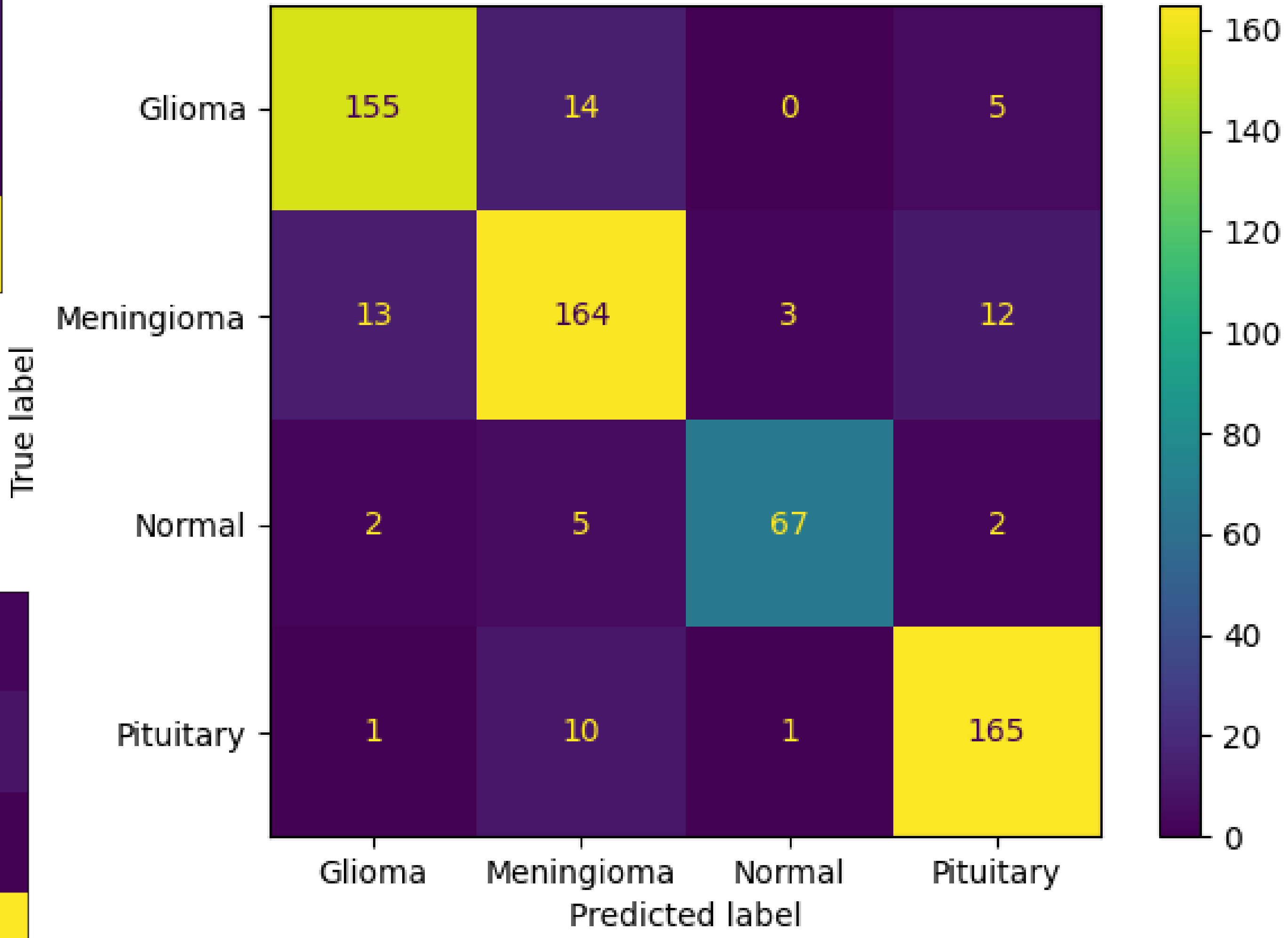


Val. Accuracy = 85.46%



Val. Accuracy = 85.95%

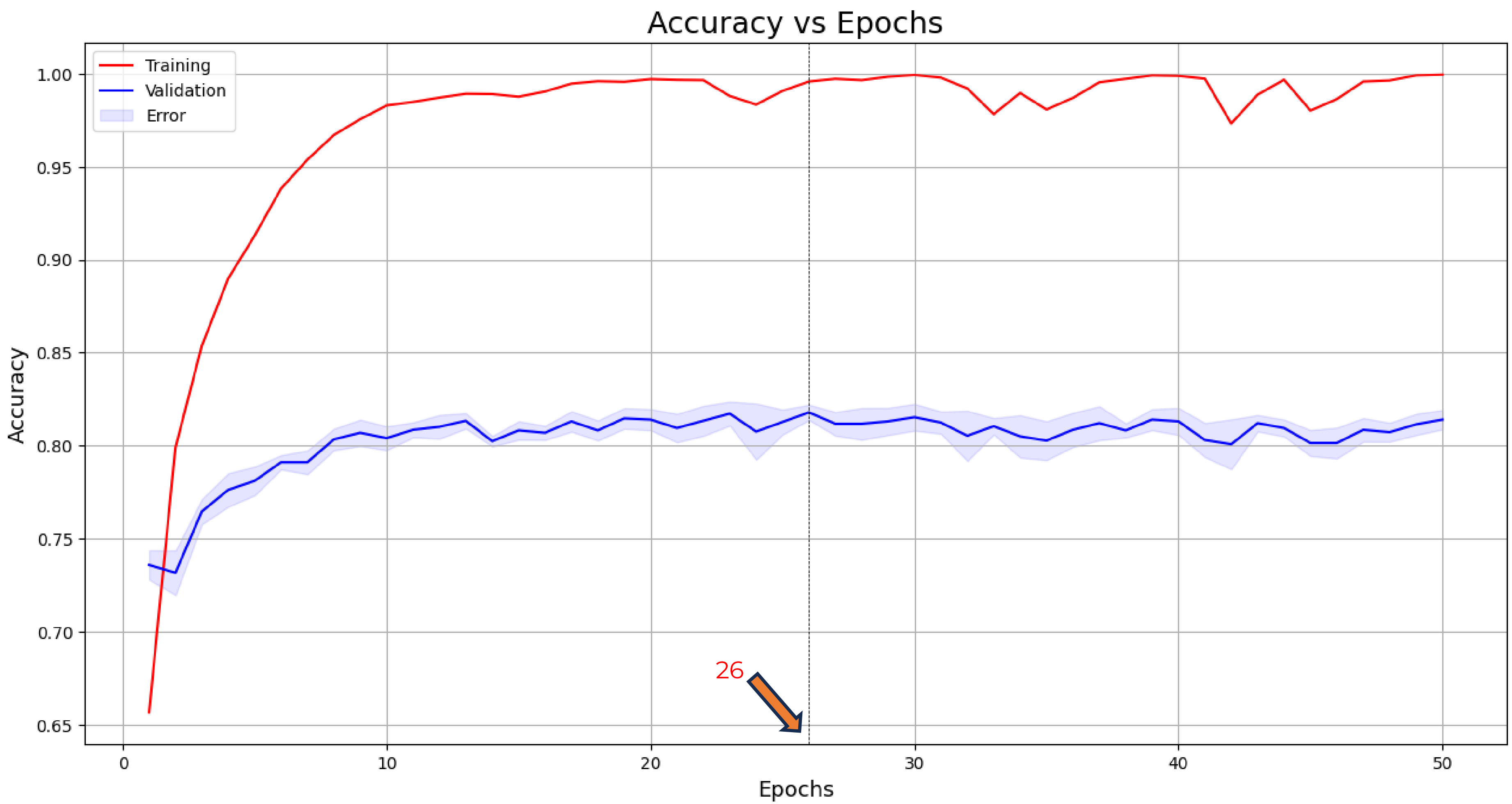




Val. Accuracy = 89.01%

FFNET: EfficientNetV2

Accuracy media delle k crossvalidazione al variare delle epoche
(Best case → ENetV2 layer 3)



Best Acc. Mean = 81.78%

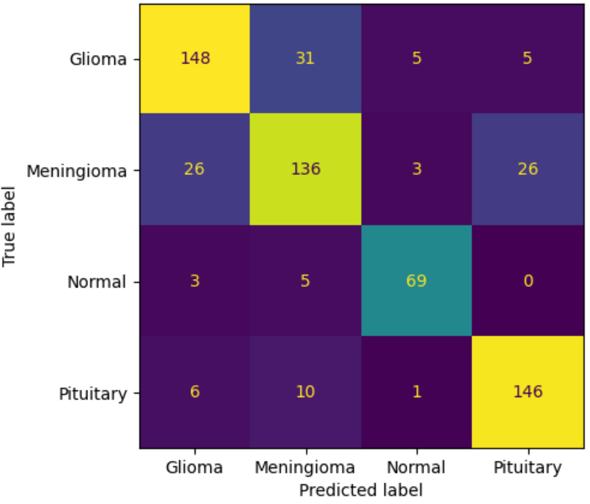
Std = 0.83%

Hyperparameters:

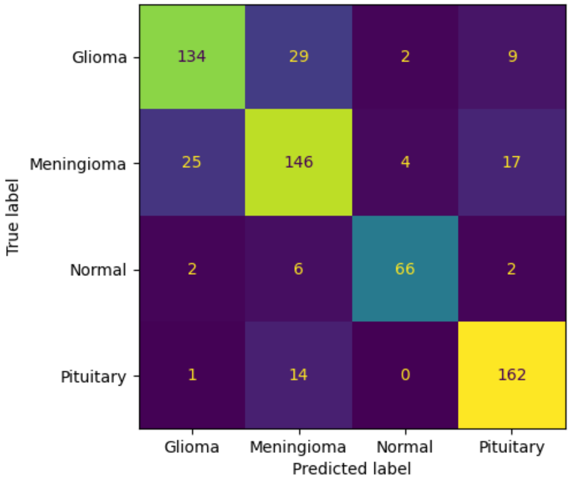
- N° layer: 1
- Neuroni: 395
- Learn. Rate: $1.6 \cdot 10^{-3}$

FFNET: EfficientNetV2

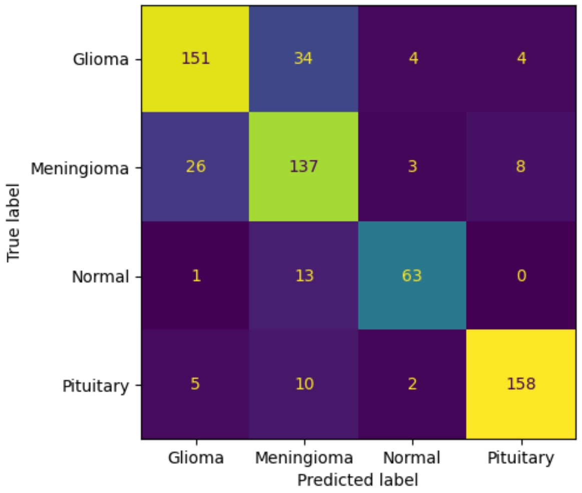
Val. Accuracy = 80.48%



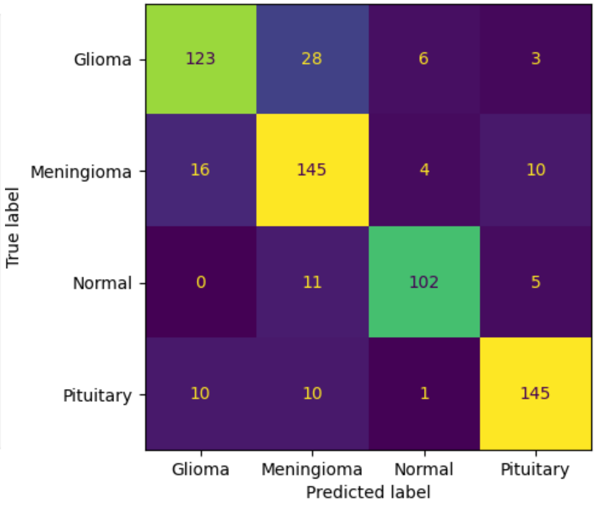
Val. Accuracy = 82.07%

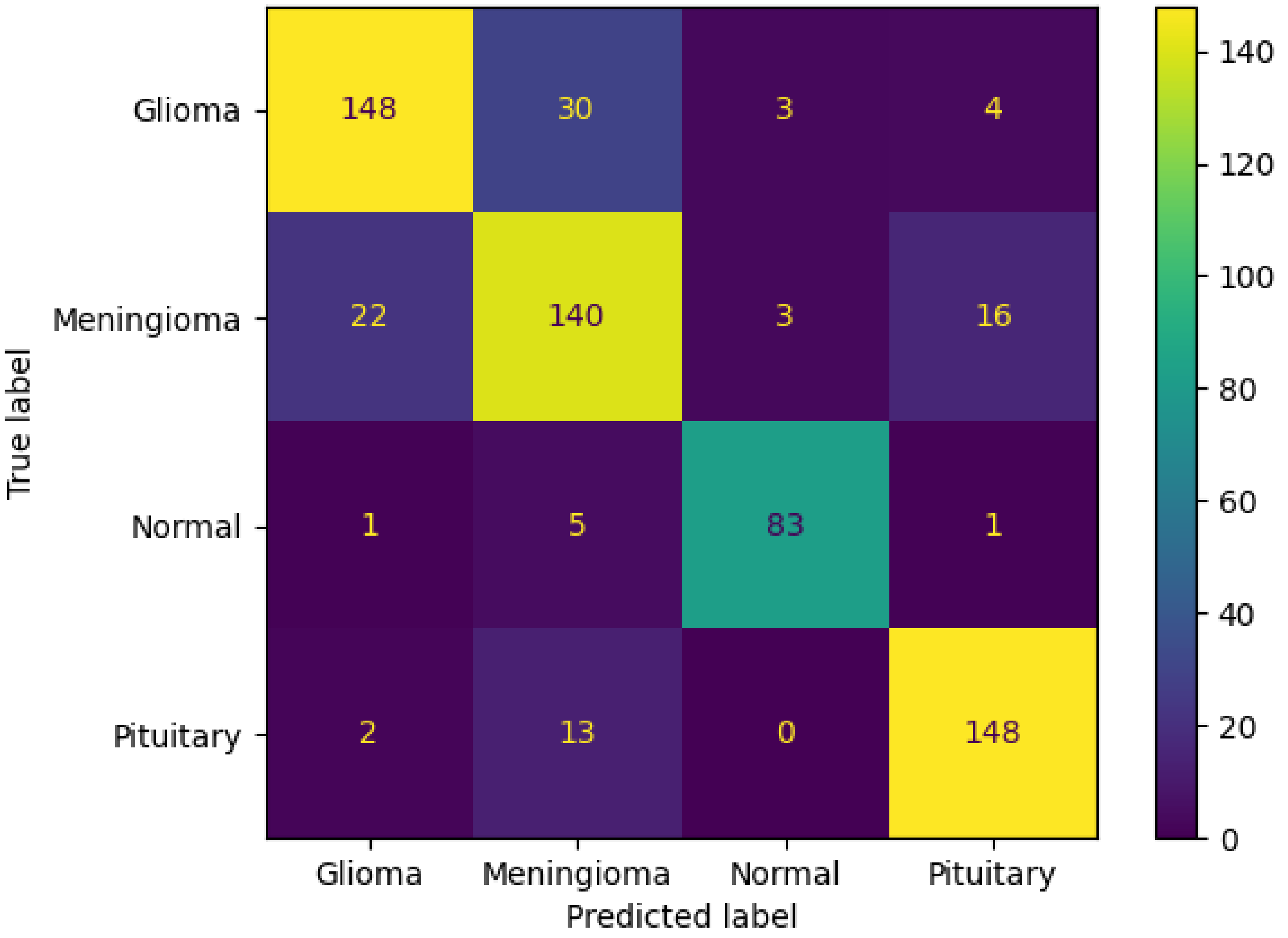


Val. Accuracy = 82.22%



Val. Accuracy = 83.20%





Val. Accuracy = 83.85%

Conclusioni e sviluppi futuri

Conclusioni

- Si osserva che le classificazioni migliori sono ottenute attraverso l'uso della CNN VGG16 considerando come uscita il layer 2
- Tra i vari classificatori quello che offre le migliori performance è la FFNET, raggiungendo accuracy medie del 86%

Sviluppi Futuri

Al fine di ottenere una migliore classificazione sarebbe importante:

- Avere a disposizione un numero maggiore di immagini
- Applicare metodi di Feature Selection migliori
- Eseguire Fine-Tuning

Grazie per l'attenzione