

Enhancing Brain Tumor Detection in MRI Images Through Explainable AI

DSAI 305, Spring 2025

Project Overview



01

Problem Statement

MRI tumour detection is vital yet error-prone. Deep nets are accurate but opaque, limiting clinical trust.

02

Dataset:

Kaggle Brain-Tumour MRI: 7 023 axial T1-Gd slices (glioma, meningioma, no-tumour, pituitary).

03

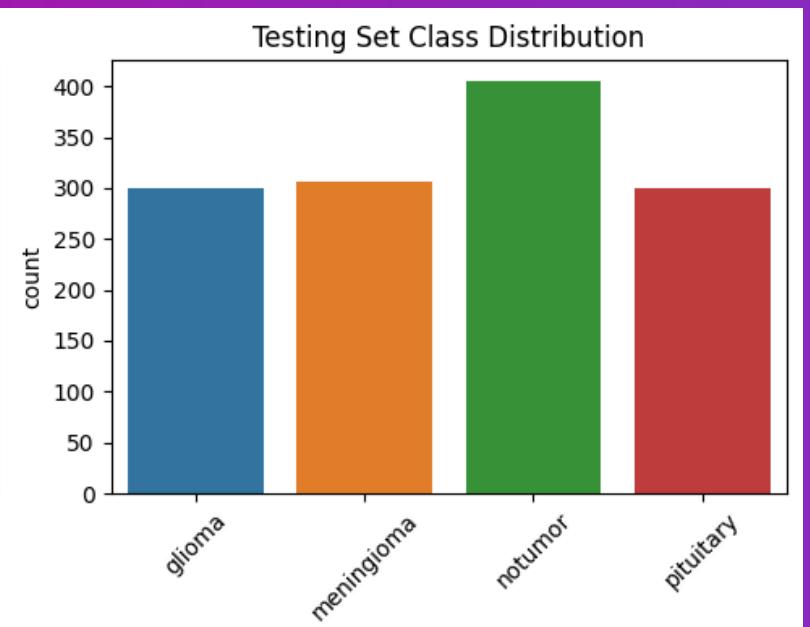
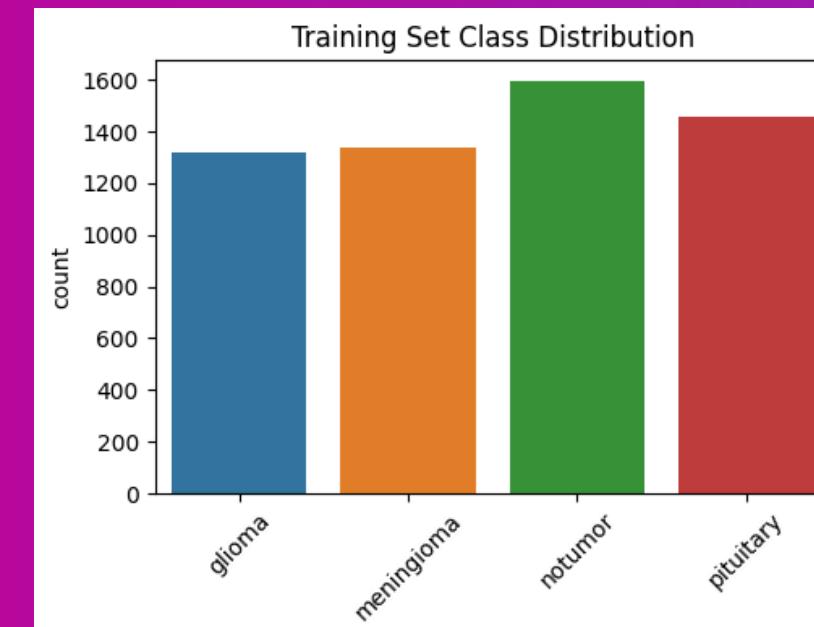
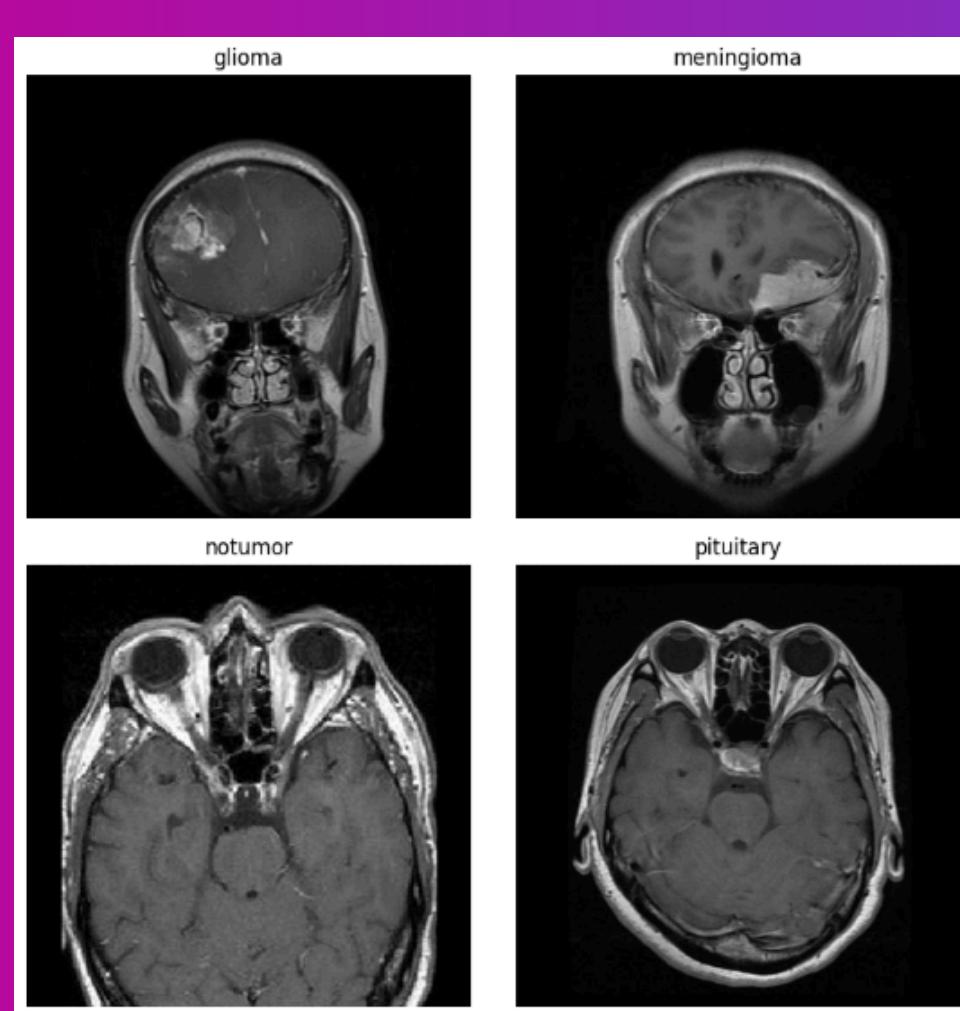
Goal

Build multiple hybrid models and compare four XAI tools to raise both accuracy and interpretability.

Problem & Data

Key EDA points

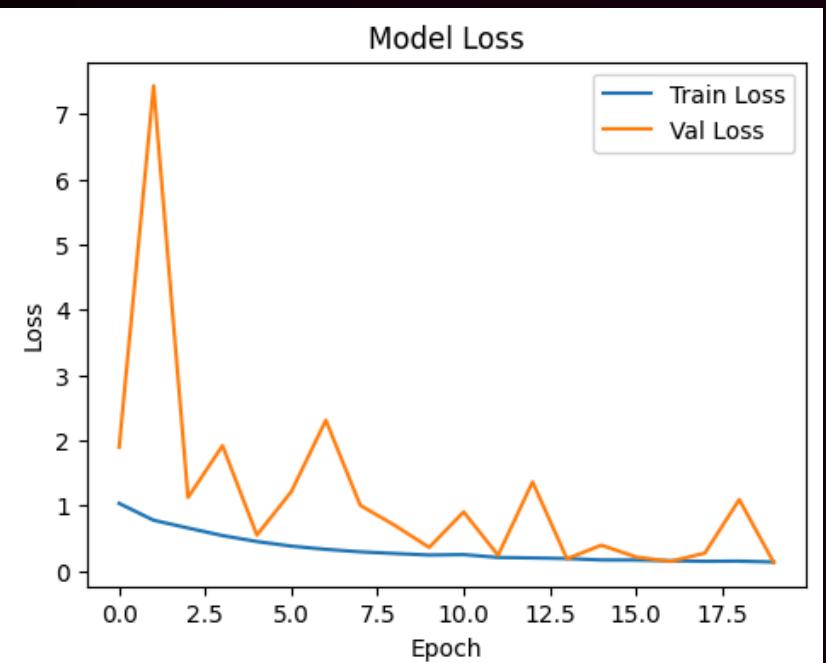
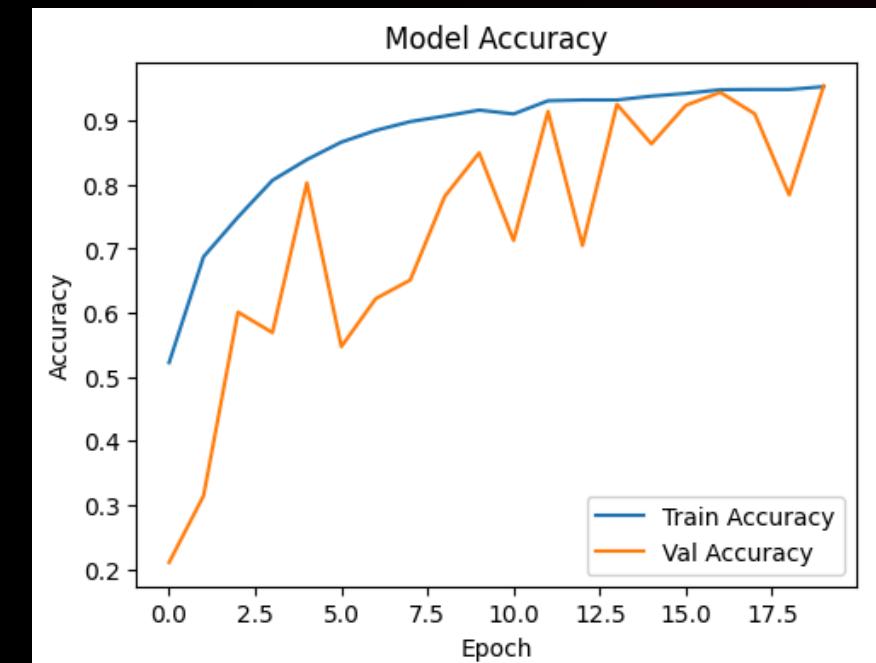
- Slight class imbalance (no-tumour largest).
- Images vary in brightness; normalised to ImageNet stats.
- Single axial slice per case.



Mazen Khaled

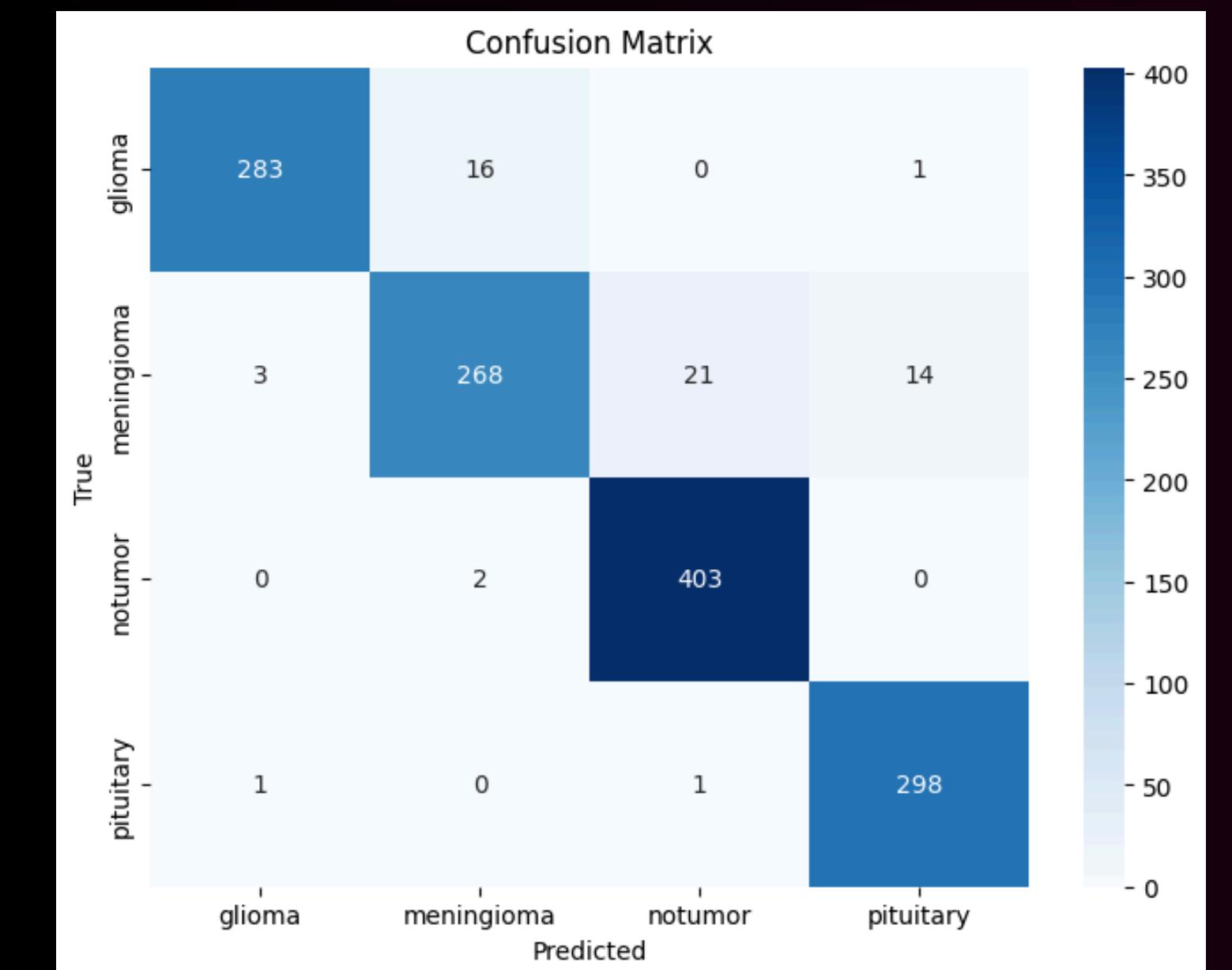
Model 1 - Custom CNN – Settings

Blocks : Conv-BN-Pool $\times 4$,
Dense 512, Dropout 0.4
Epochs : 25 (EarlyStop best
@ 20)
Augment : $\pm 15^\circ$ rotate, zoom 0.1,
flips
Best val-acc : 95.47 %, Test :
95.0 %

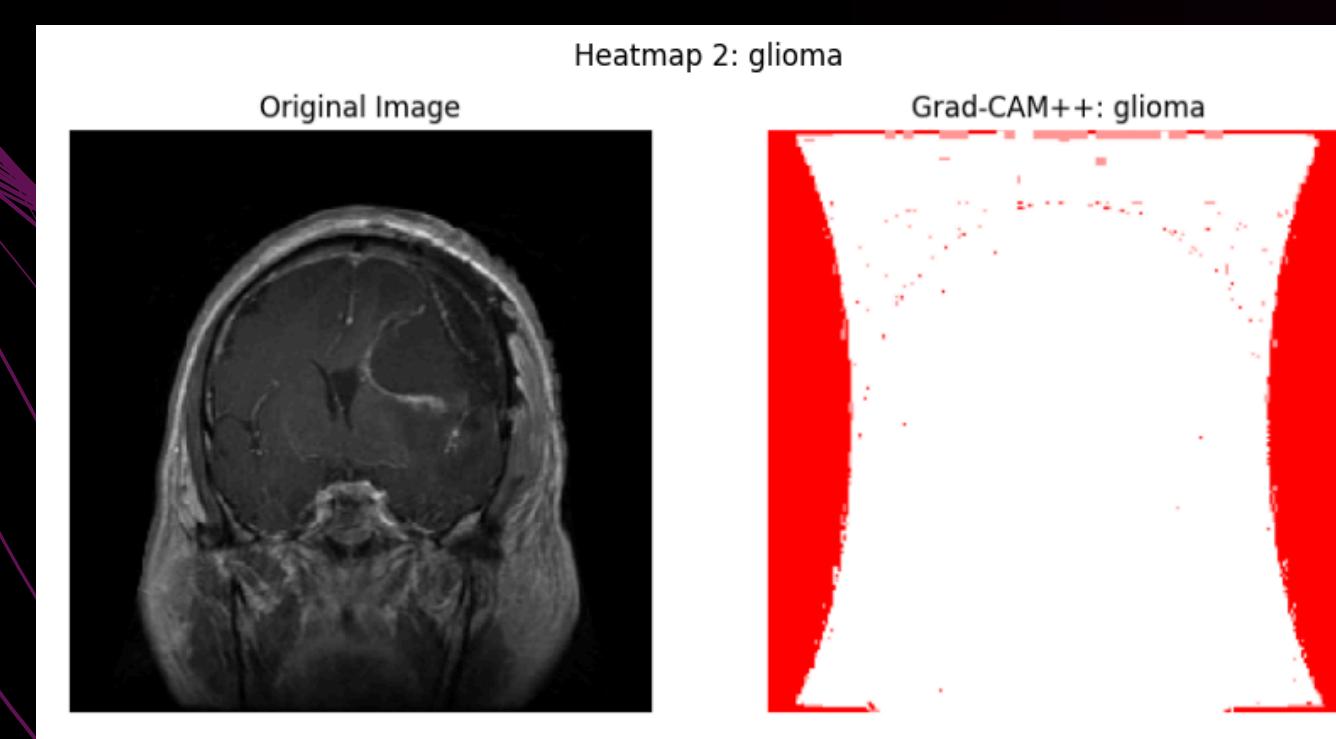
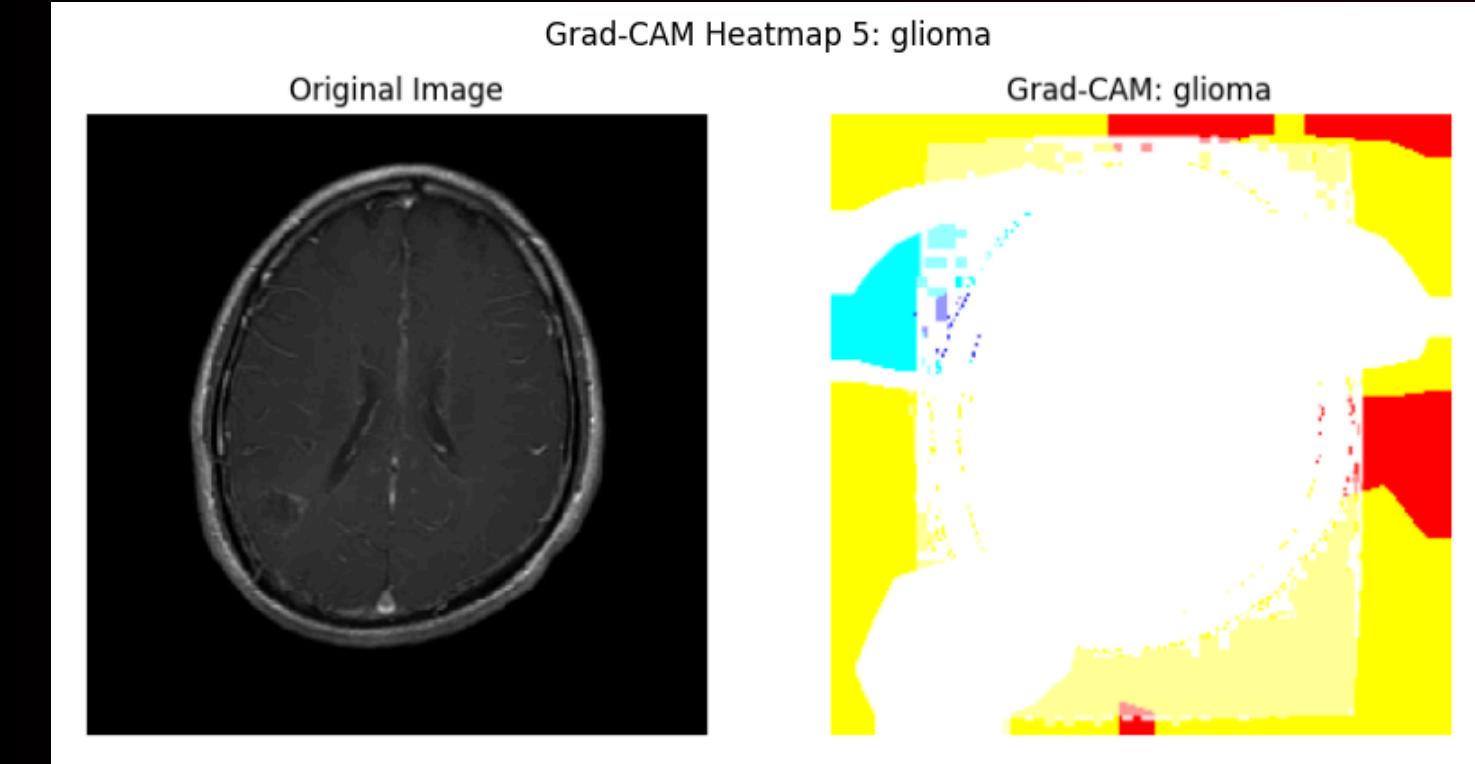
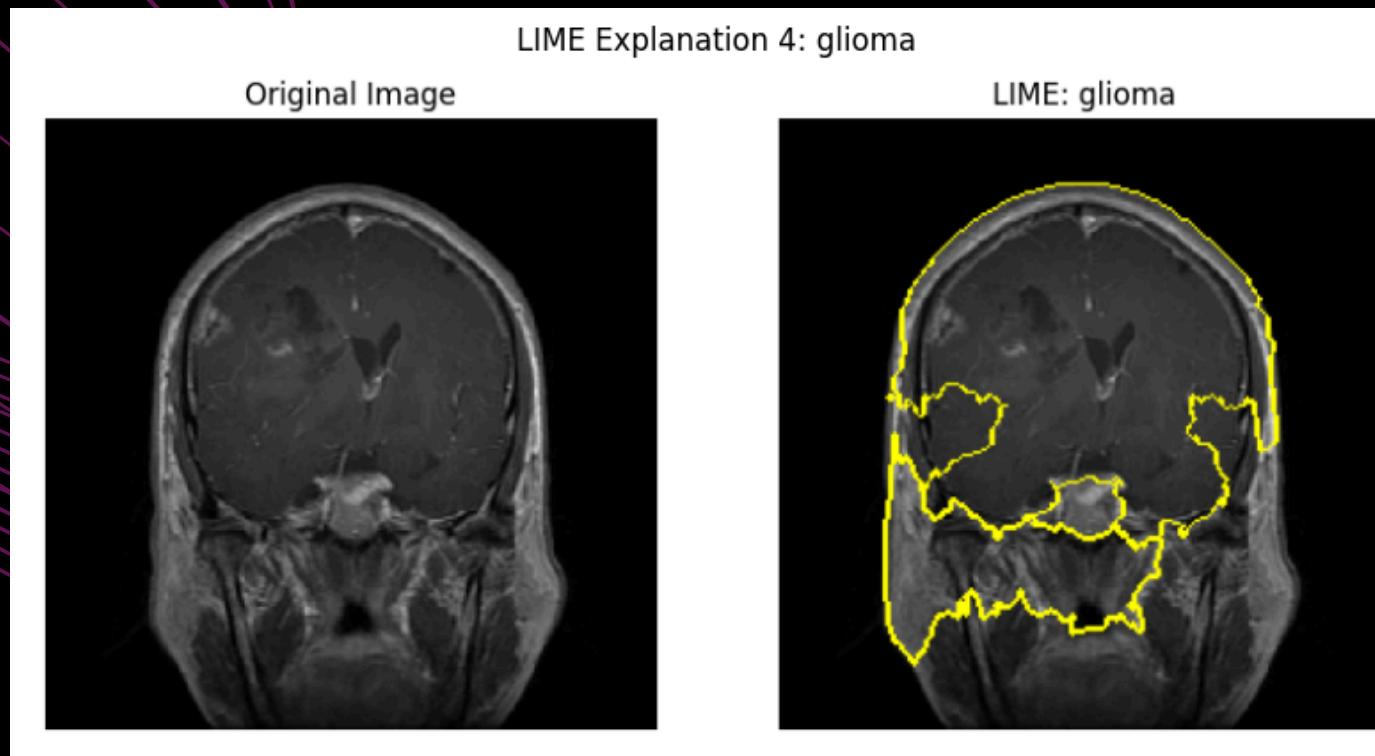


Model 1 - Custom CNN – Performance

- F1 (macro) 0.95; meningioma recall lowest 0.88
- Minor over-fit spikes managed by callbacks
- Confusion reveals 16 glioma → meningioma switches



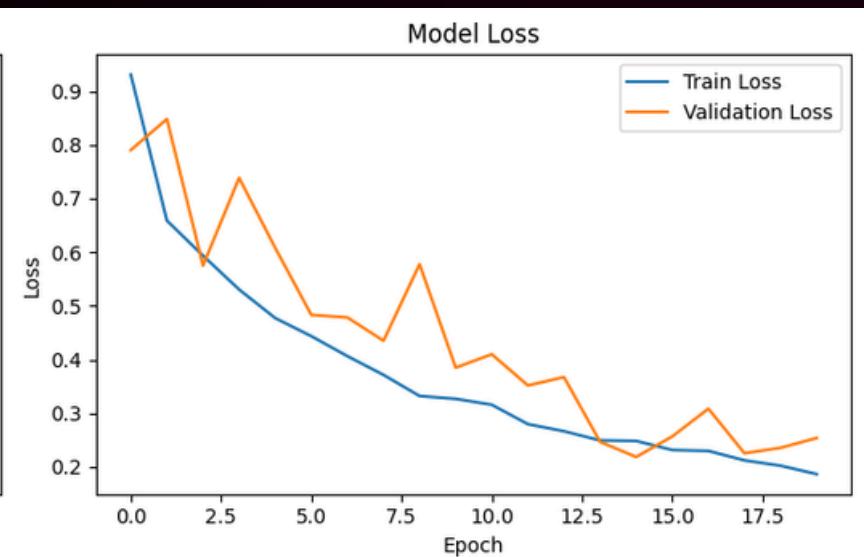
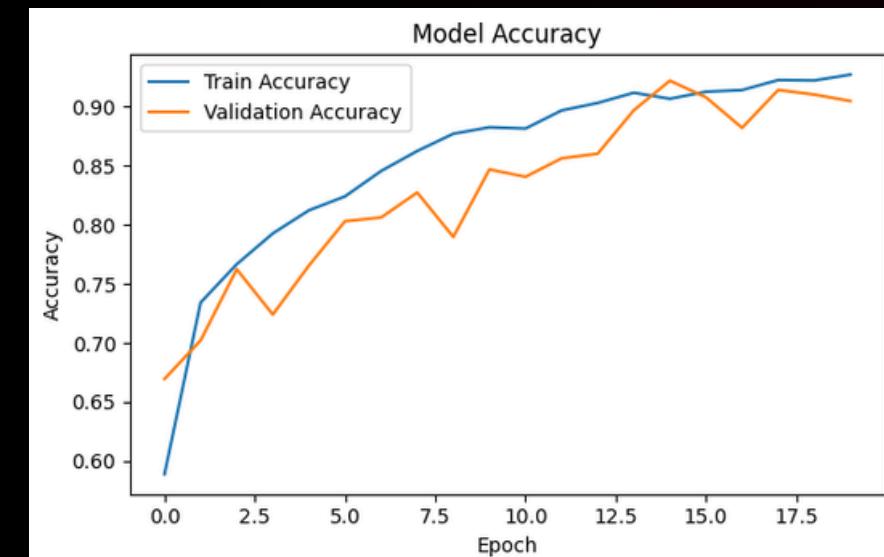
Model 1 - Custom CNN – XAI



Mazen Khaled

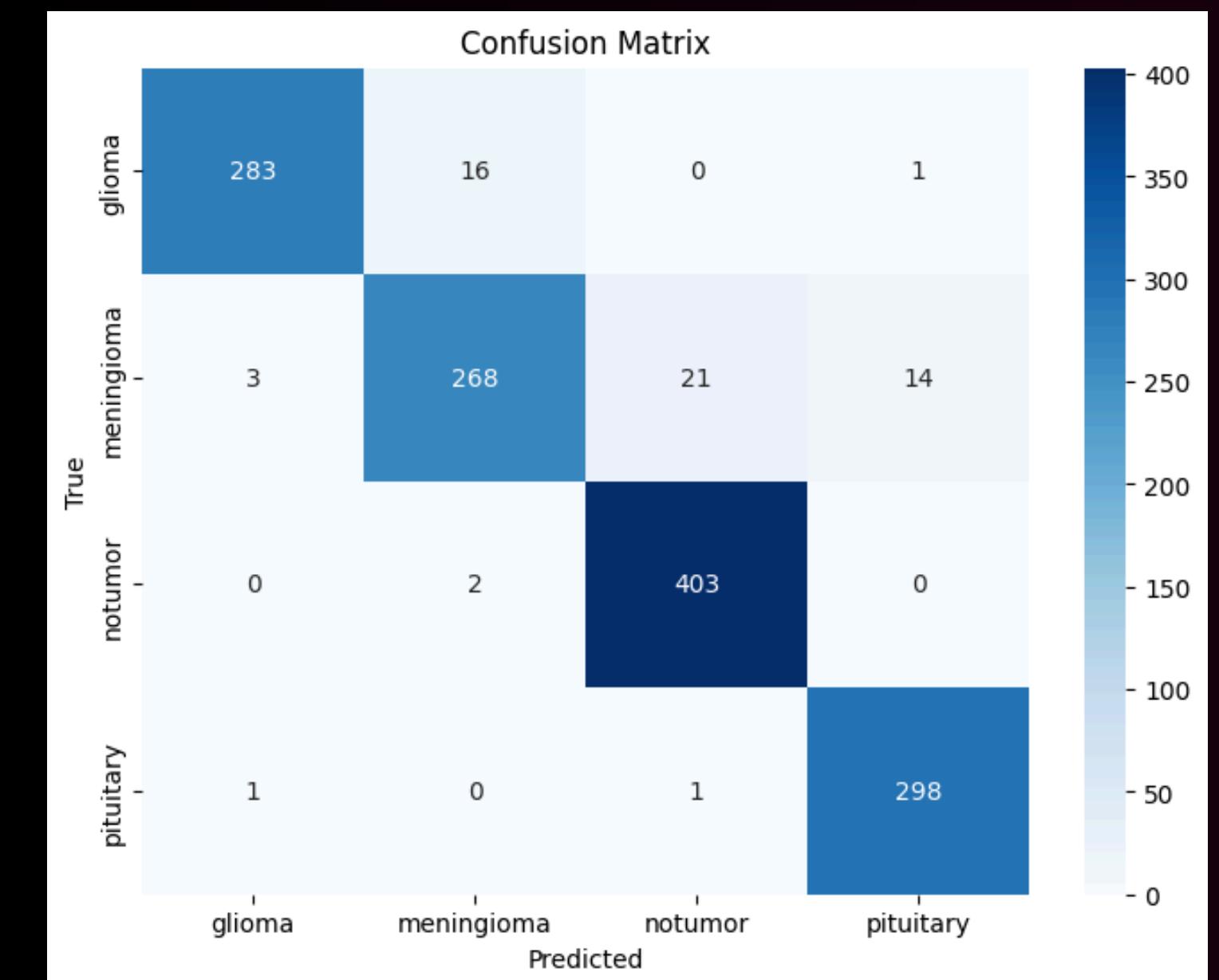
Model 2 - DSCC-Net Transfer – Settings

**Backbone : DSCC-Net
dermatology model
Epochs 20, final val-acc 92.18 %
→ Test 90.4 %
Weak spot: Glioma recall 0.80
(FN risk)**

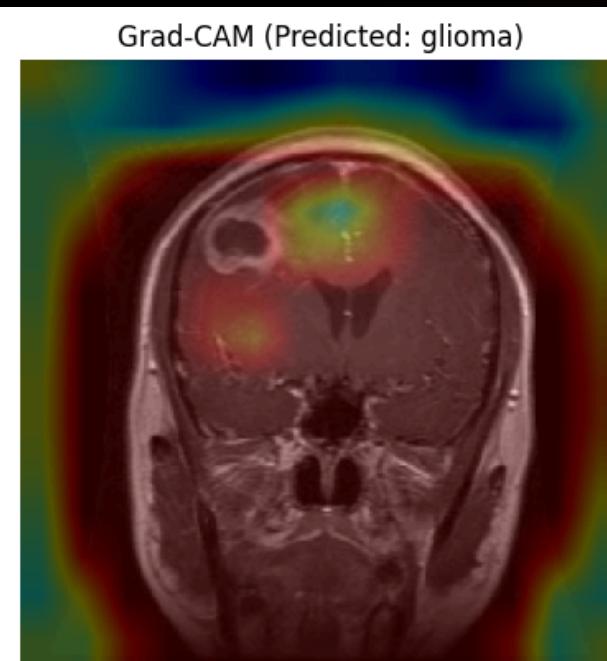
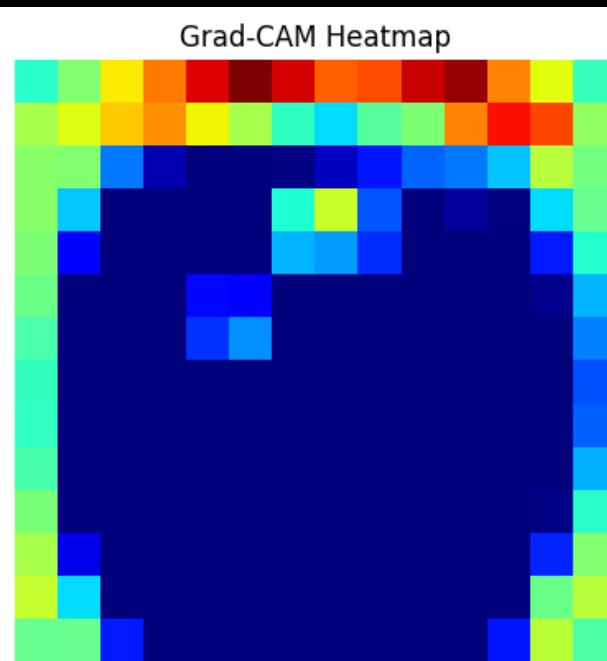
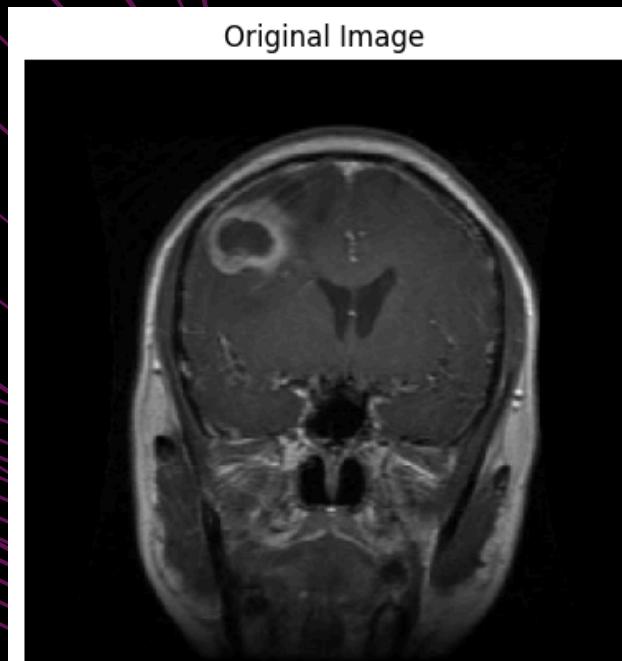


Model 2 - DSCC-Net Transfer – Performance

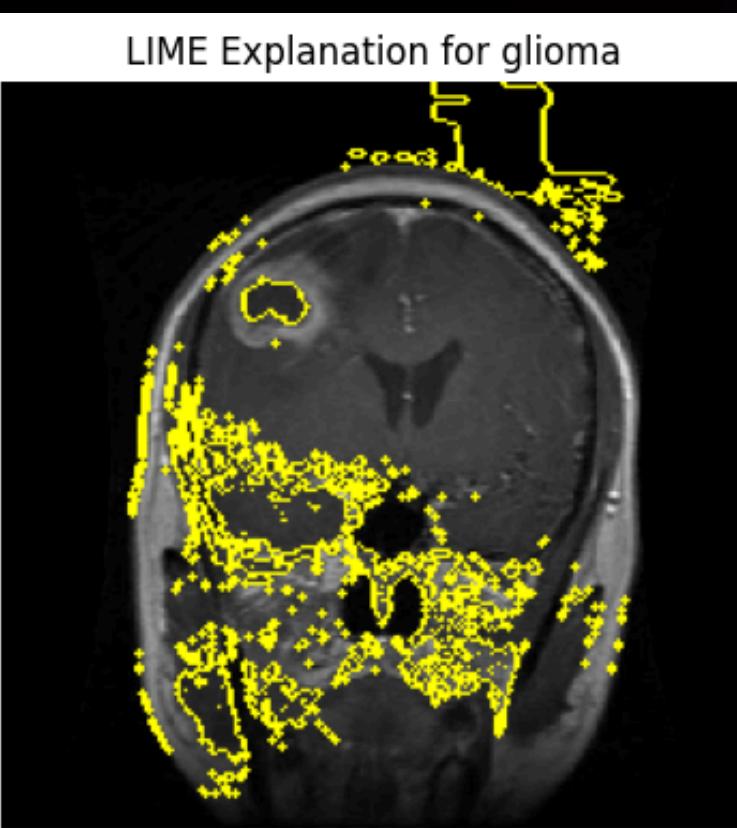
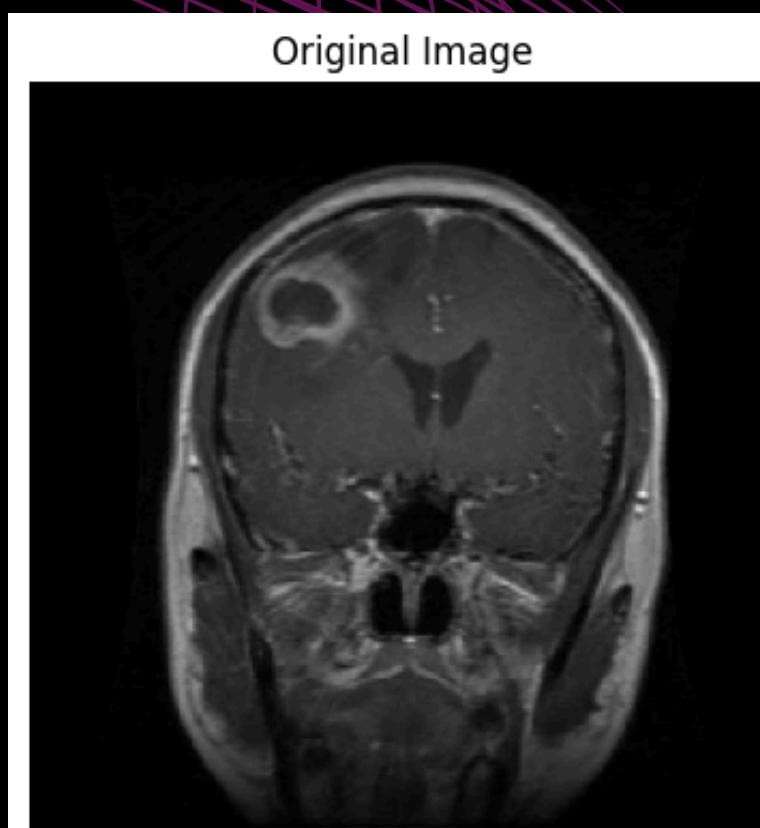
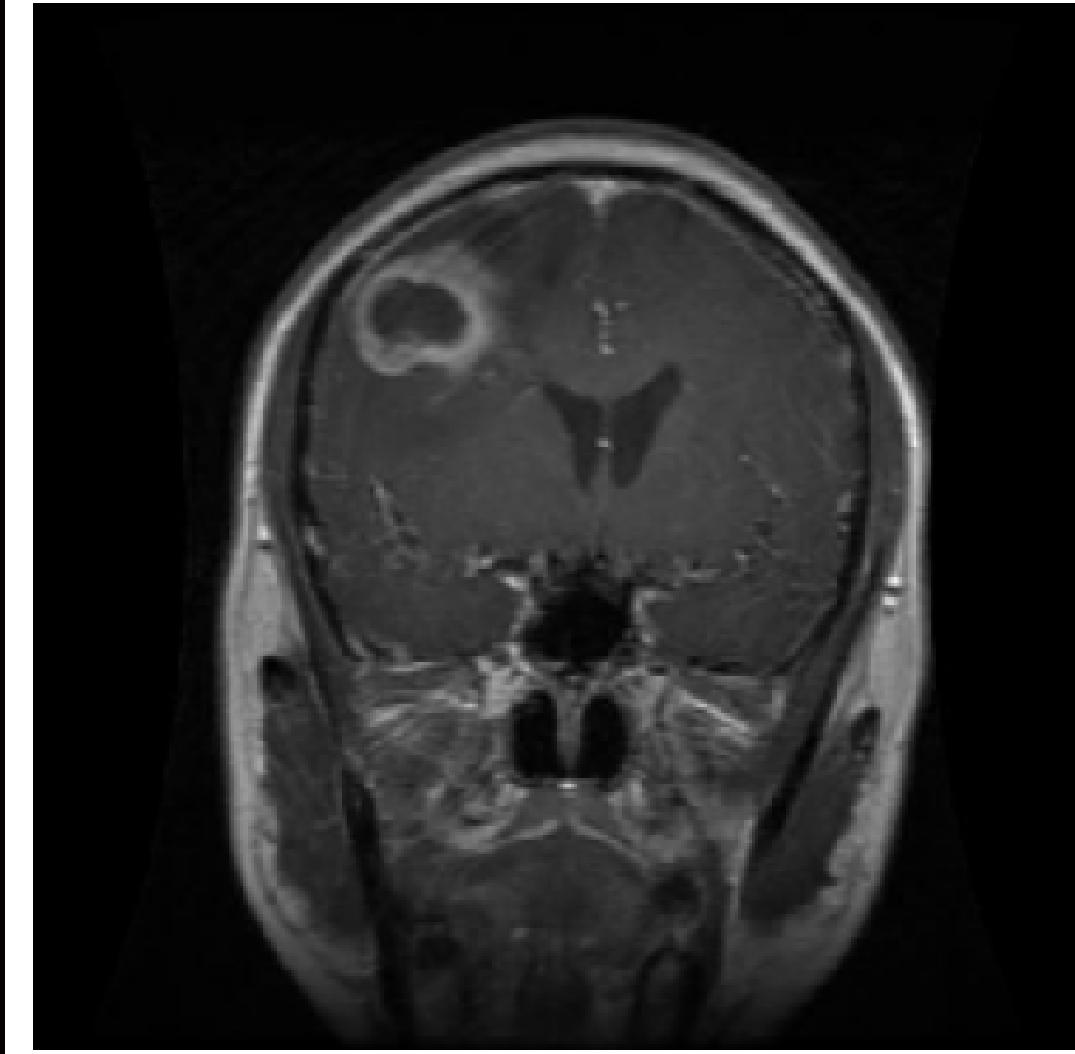
- Overall test Acc 90.4%; loss 0.26 → still solid but ↓ 5 pts vs Custom-CNN
- Strength: 'notumor' & 'pituitary' near-perfect (recall 1.0 / 0.99)
- Weak spot: glioma recall 0.80 → 61 FN (45→meningioma, 16→pituitary)
- Likely cause: DSCC-Net pre-trained on dermoscopic texture; struggles with diffuse glioma edges.
- Action: feed T2/FLAIR channels or fine-tune deeper to lift glioma sensitivity.



Model 2 - DSCC-Net Transfer – XAI



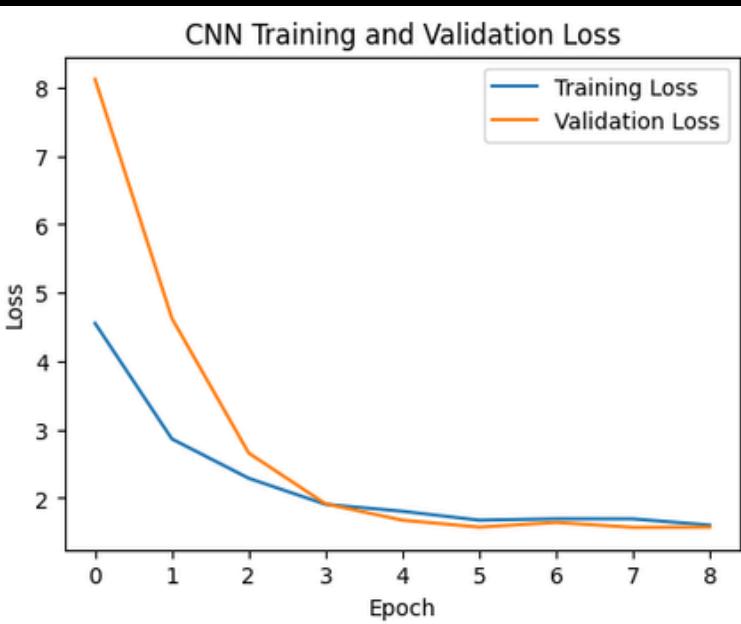
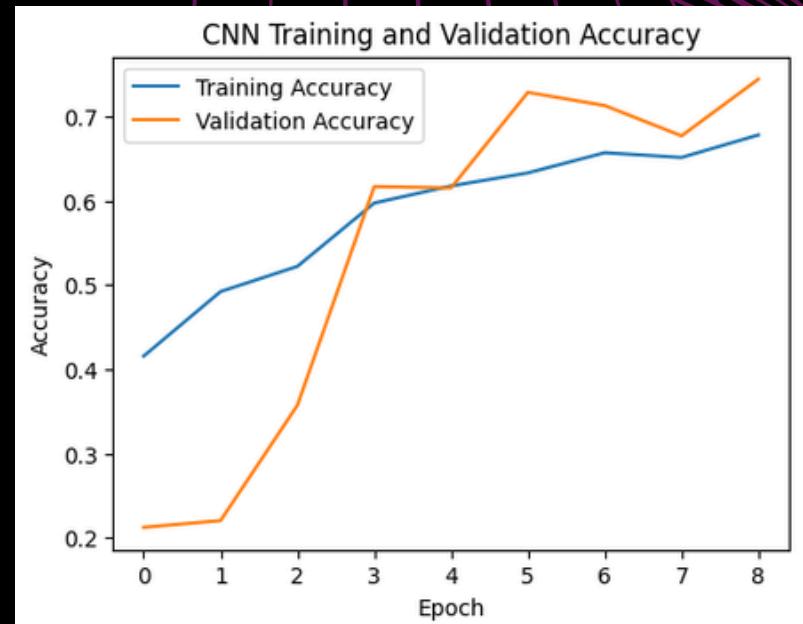
Sample Image (Predicted: glioma)



Mazen Khaled

Model 3 - comparison CNN + SVM

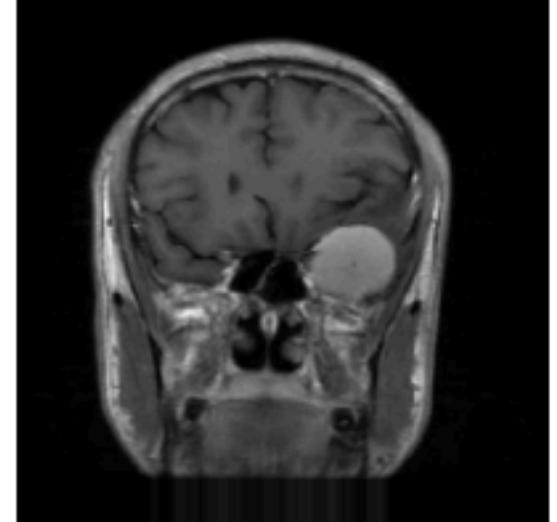
- Custom CNN with two kernel branches (3×3 & 5×5), Dropout 0.4
- Final test accuracy: CNN = 72.8 %, SVM-RBF = 69.9 %, SVM-Linear = 64.3 %
- PCA (50 comps) before SVM, with permutation-importance
- Confusion reveals frequent glioma \leftrightarrow meningioma switches
- CNN beats both SVM baselines on overall and macro-F1



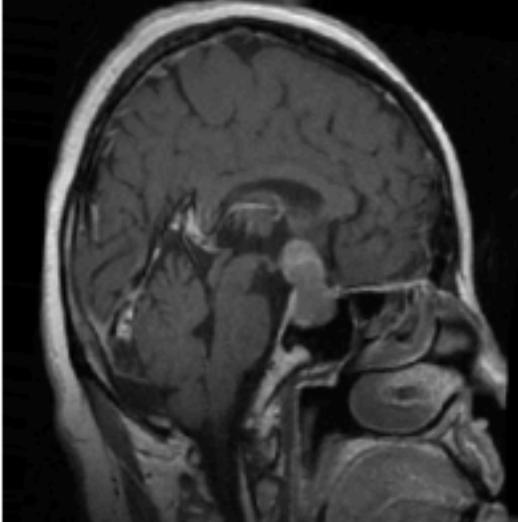
Model 3 - comparison CNN + SVM – Performance

- CNN macro-F1 = 72%; no_tumor best class (0.90 F1)
- Worst recall: meningioma (0.45), glioma ↔ meningioma confusion
- Overall test acc = 72.79 %
- Misclassifications visually evident in predictions

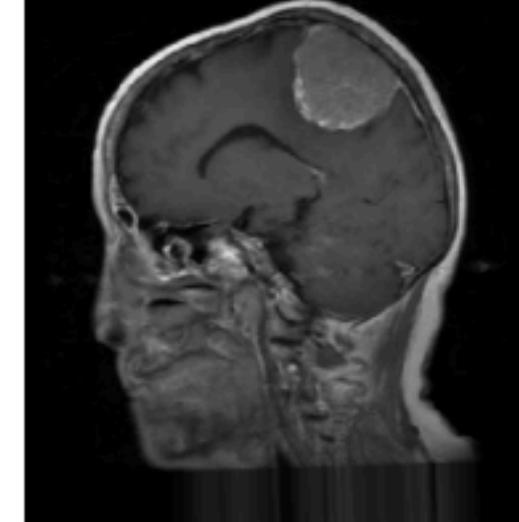
True: meningioma
Pred: meningioma



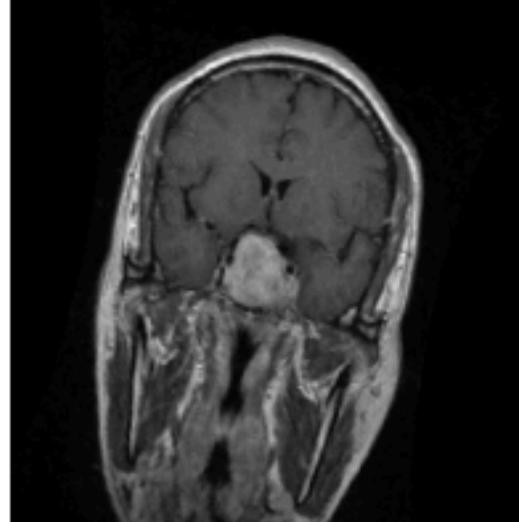
True: pituitary
Pred: pituitary



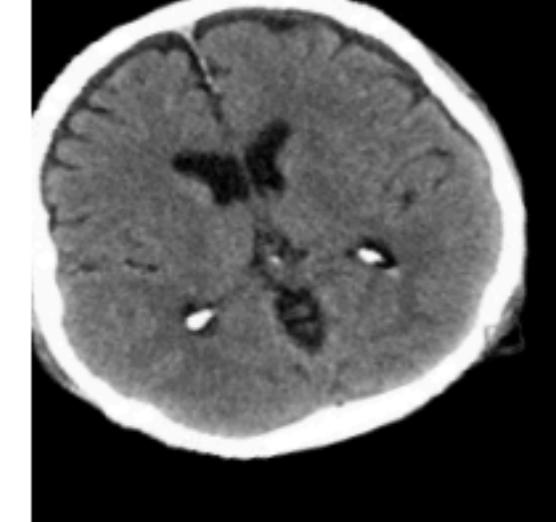
True: meningioma
Pred: glioma



True: pituitary
Pred: glioma

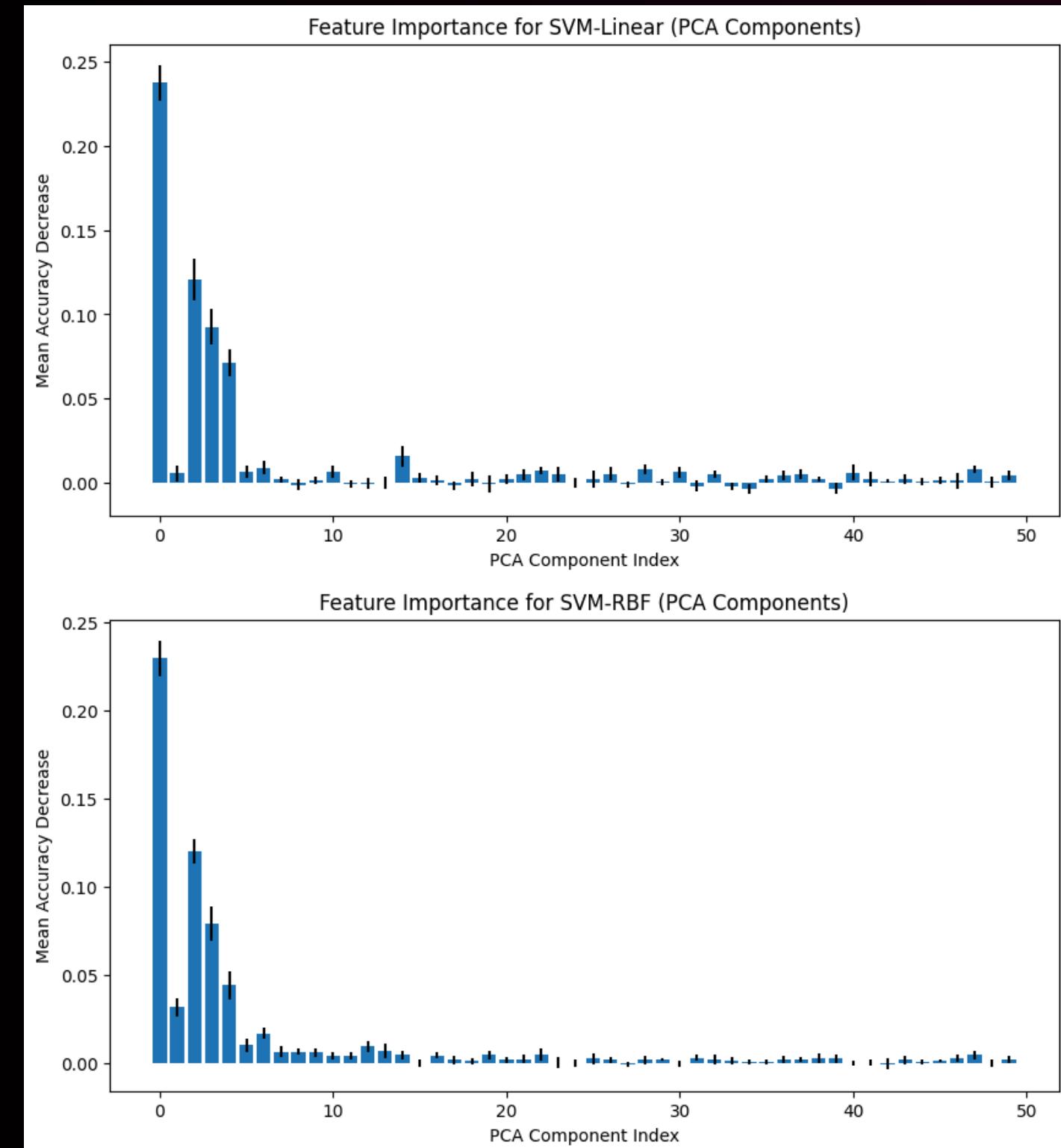


True: no_tumor
Pred: no_tumor

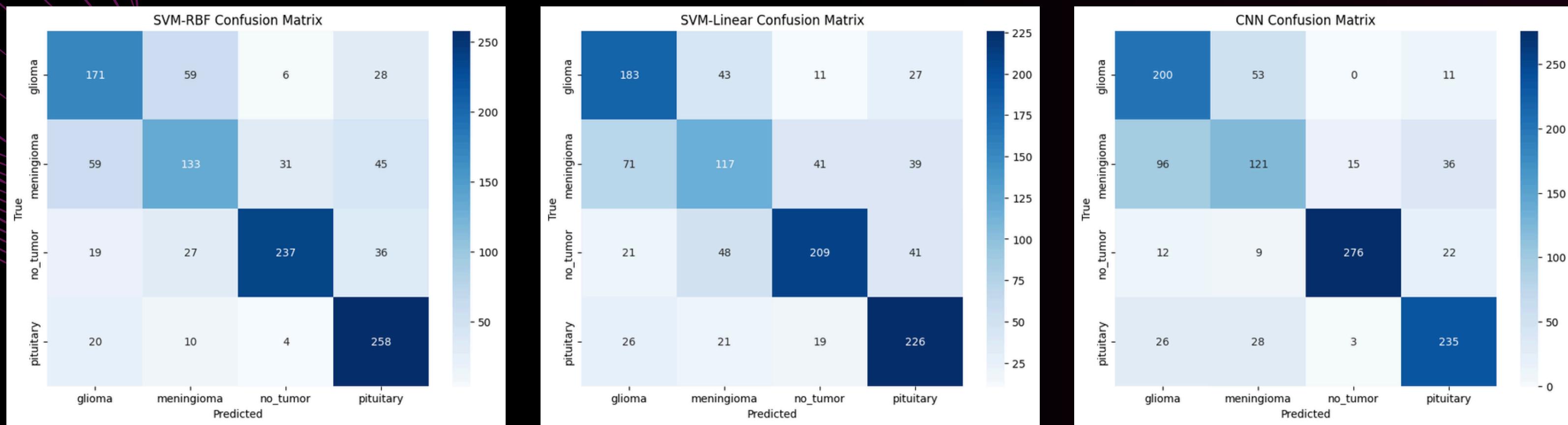


Model 3 – comparison CNN + SVM – Feature Importance (SVM)

PCA₀ and PCA₂ dominate both kernels; importance drops off after top 5.

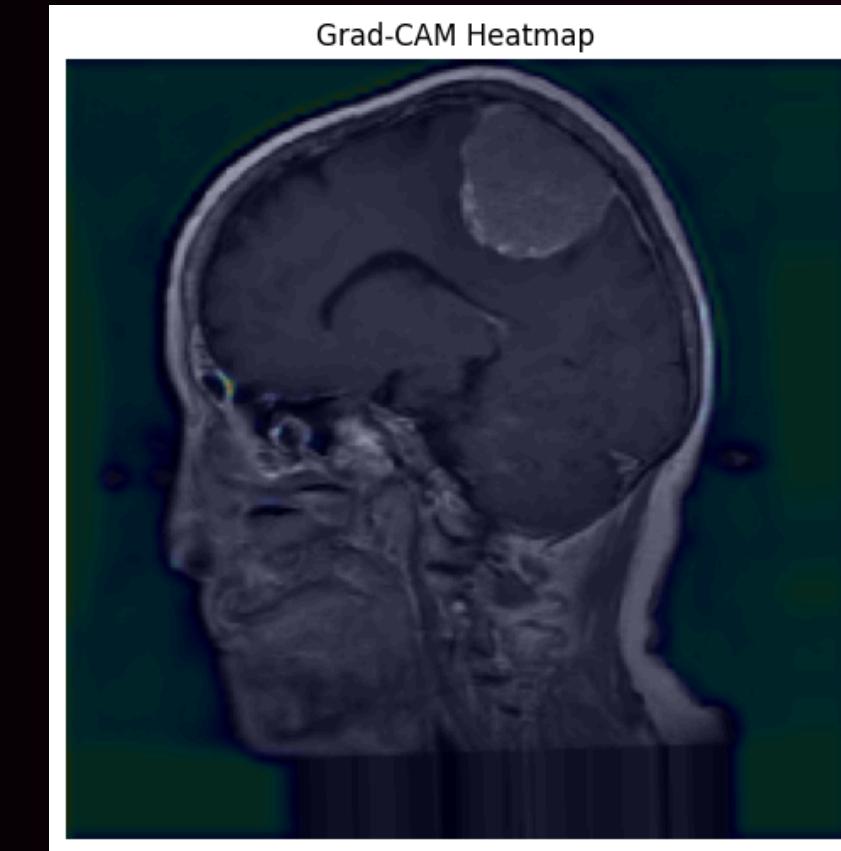
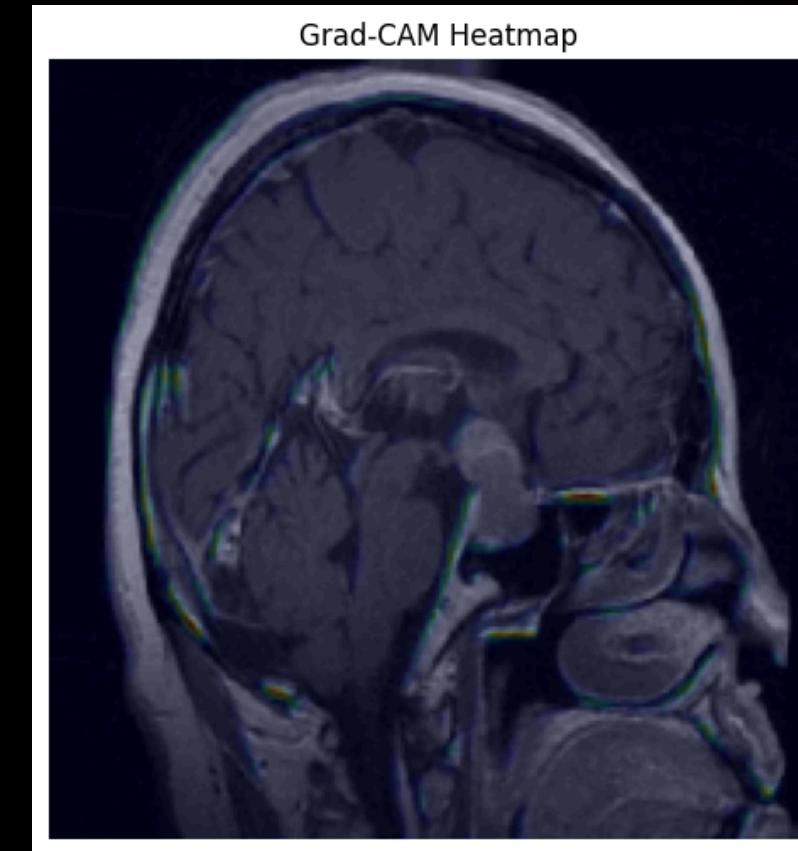
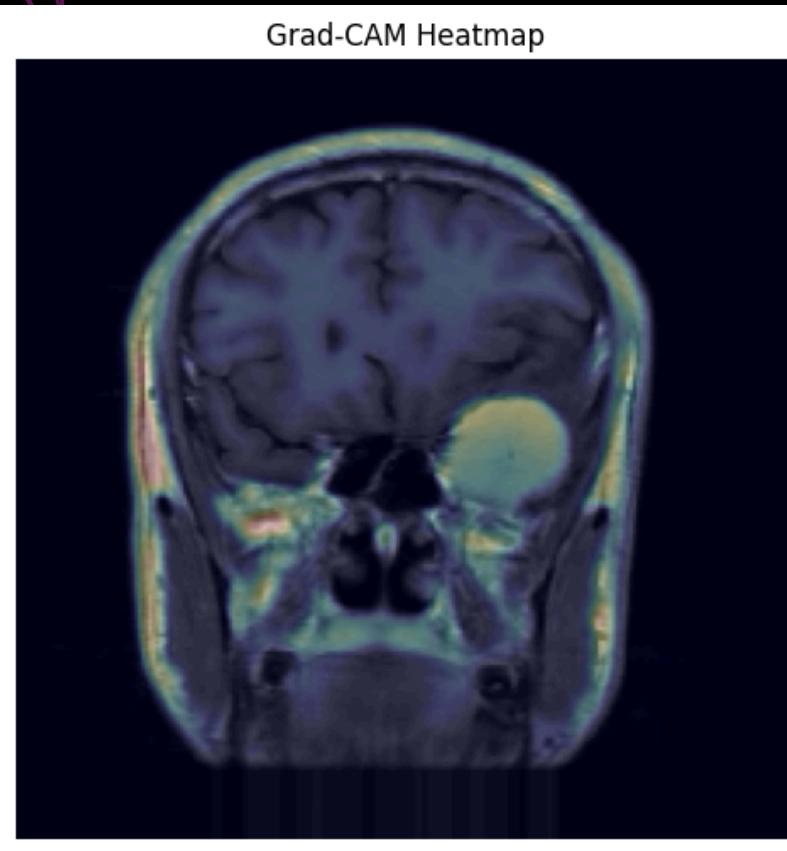


Model 3 - comparison CNN + SVM – Confusion Matrices



- CNN captures no_tumor best (87% recall)
- SVM-RBF better than linear, esp. on pituitary
- All models struggle on meningioma (class imbalance effect)

Model 3 - CNN XAI (Grad-CAM)

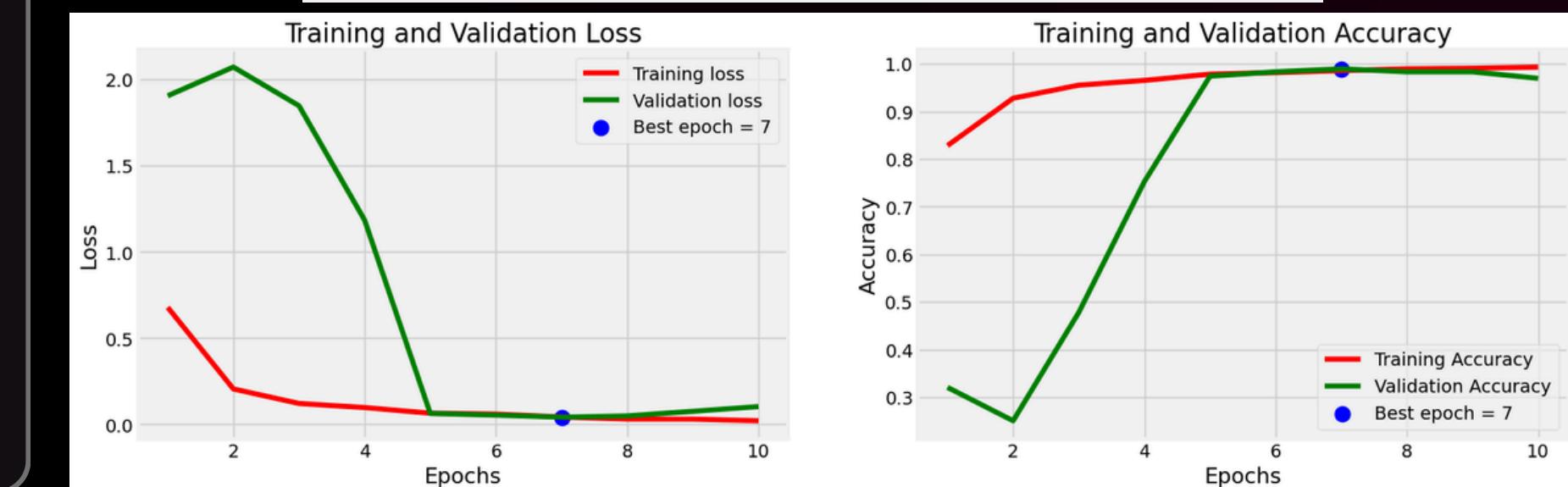
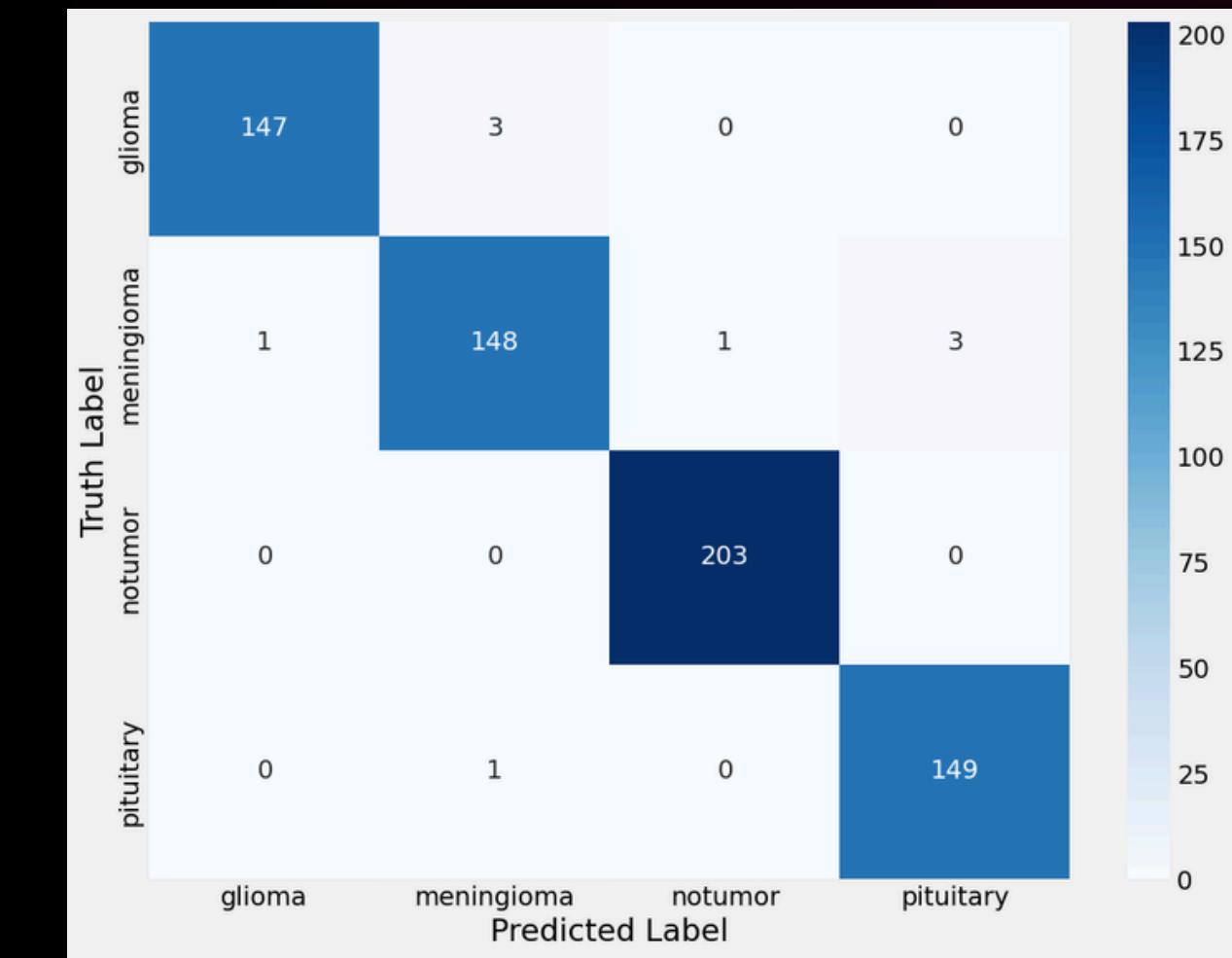


Heatmaps focus on tumour core for true positives. Error case shifts heat away from lesion.

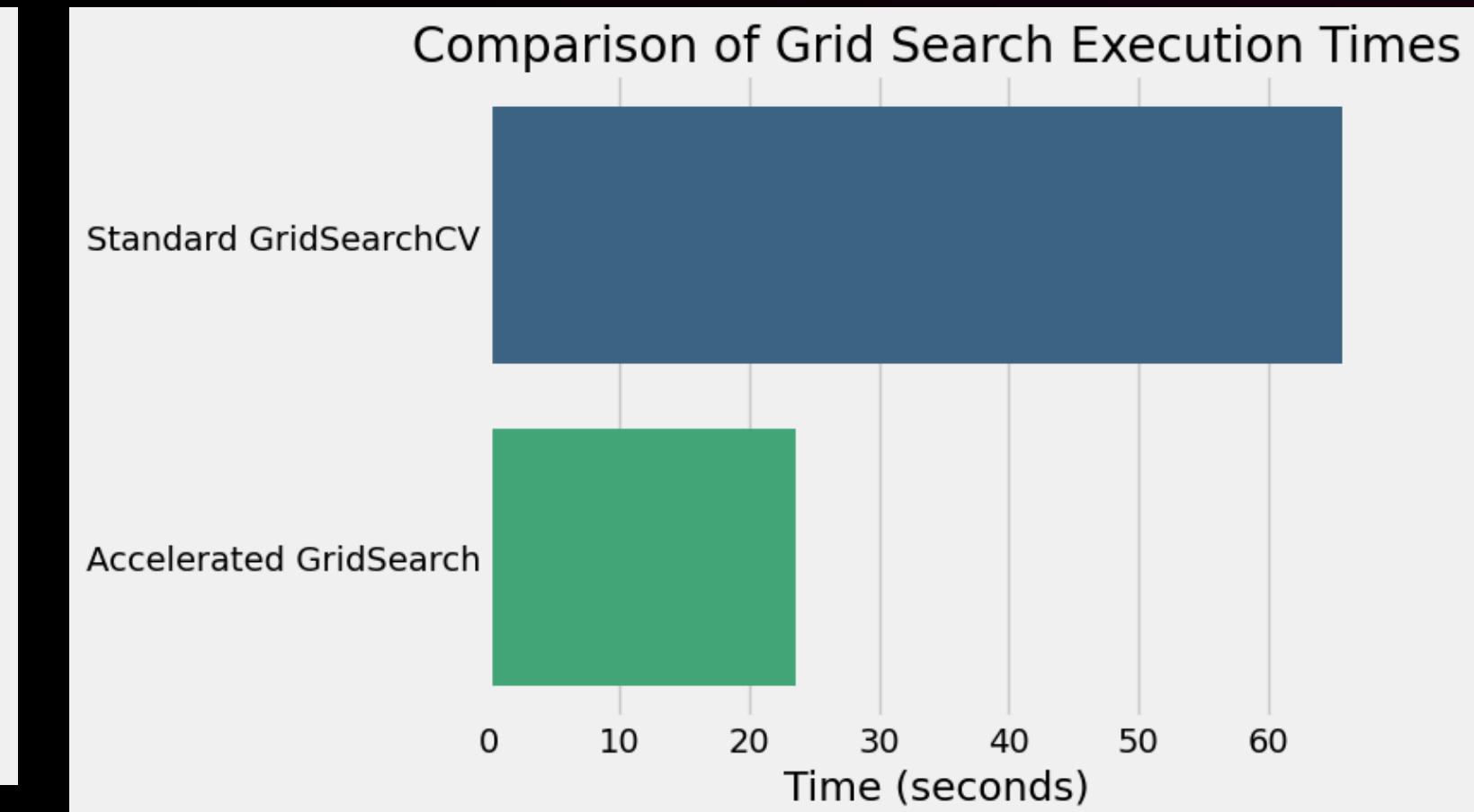
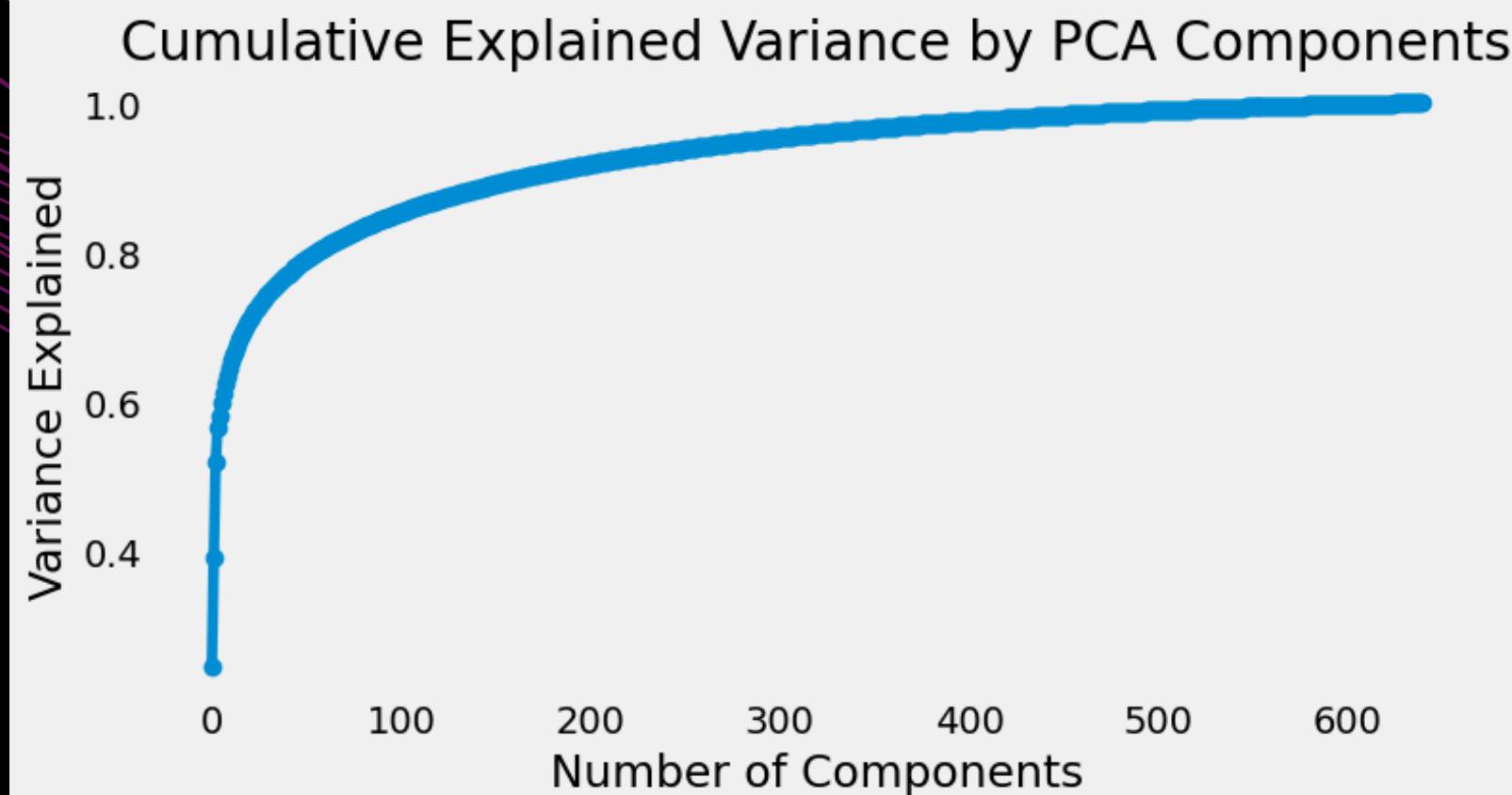
Ahmed Barakat

Model1 – RF feature-select → PCA 288 → SVM

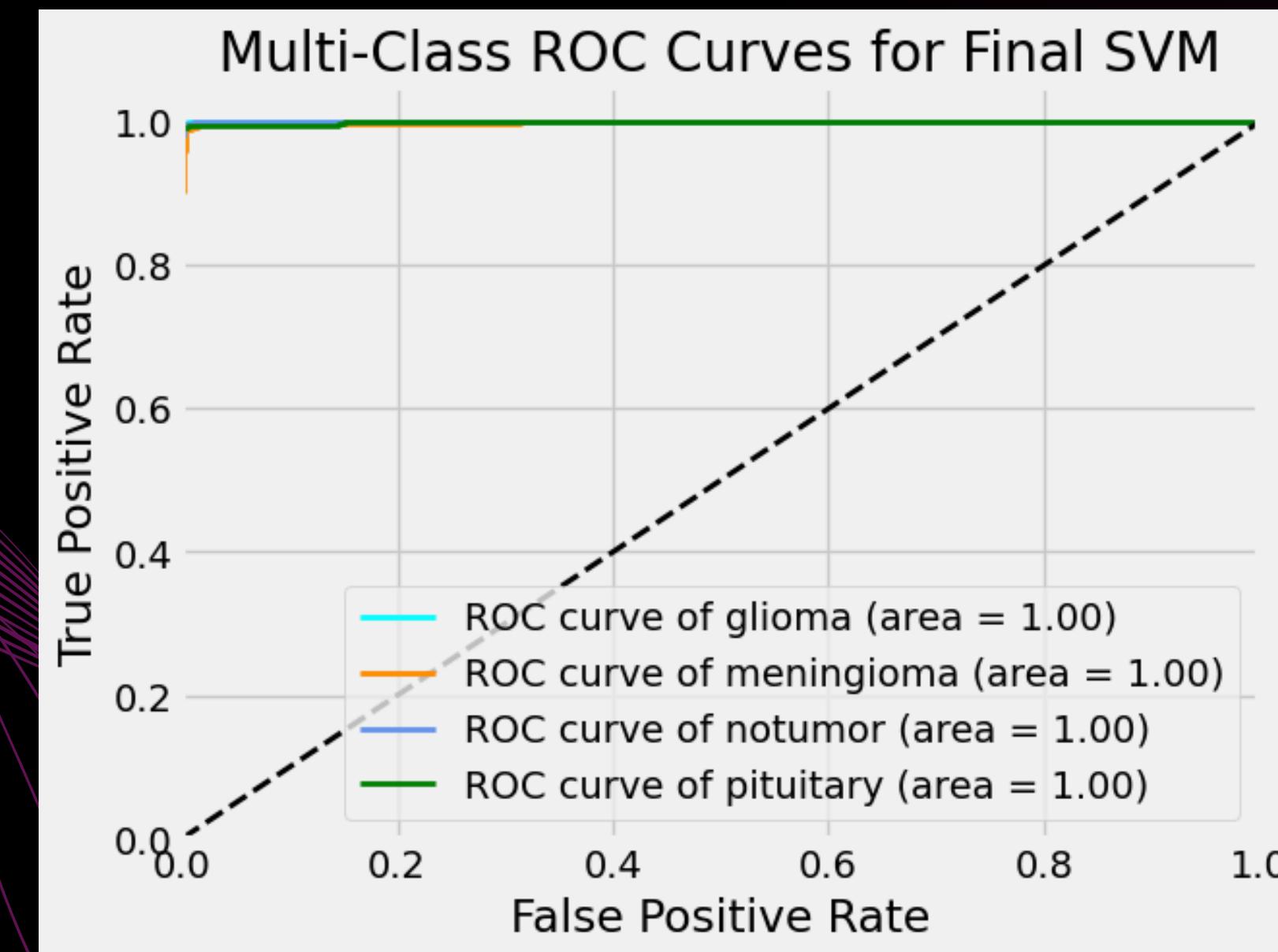
- RF selects top features → PCA 288 captures 95% variance
- SVM RBF ($C=10, \gamma=0.001$), GridSearch: 65s
- Accuracy: 98.63% on test, 99.22% (accel. search, same result)
- $\text{PCA}_1/2/3$ dominate all XAI tools



Model 1 - XAI



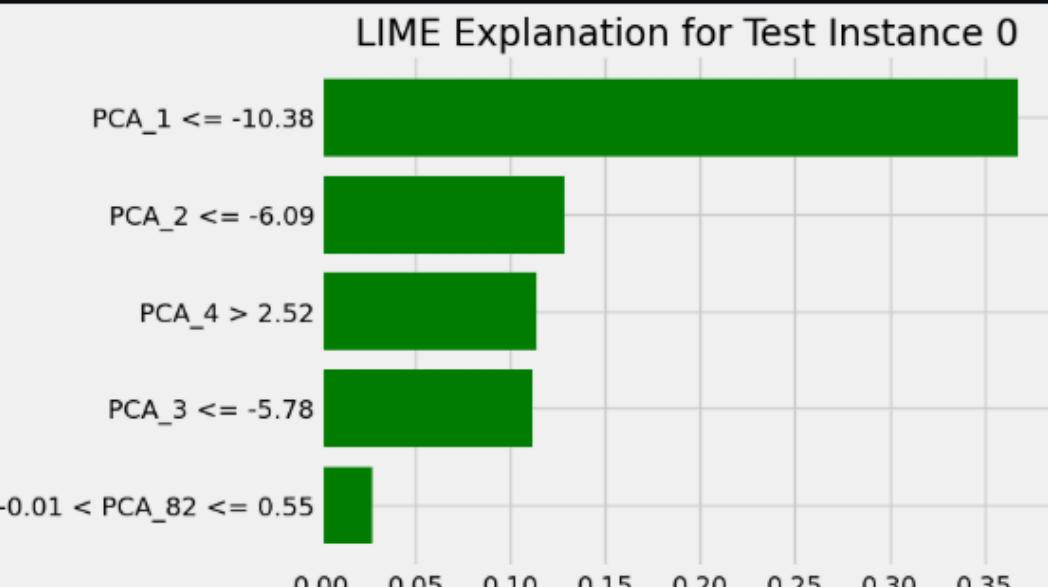
Model 1 - XAI



Model 1 - XAI - Lime

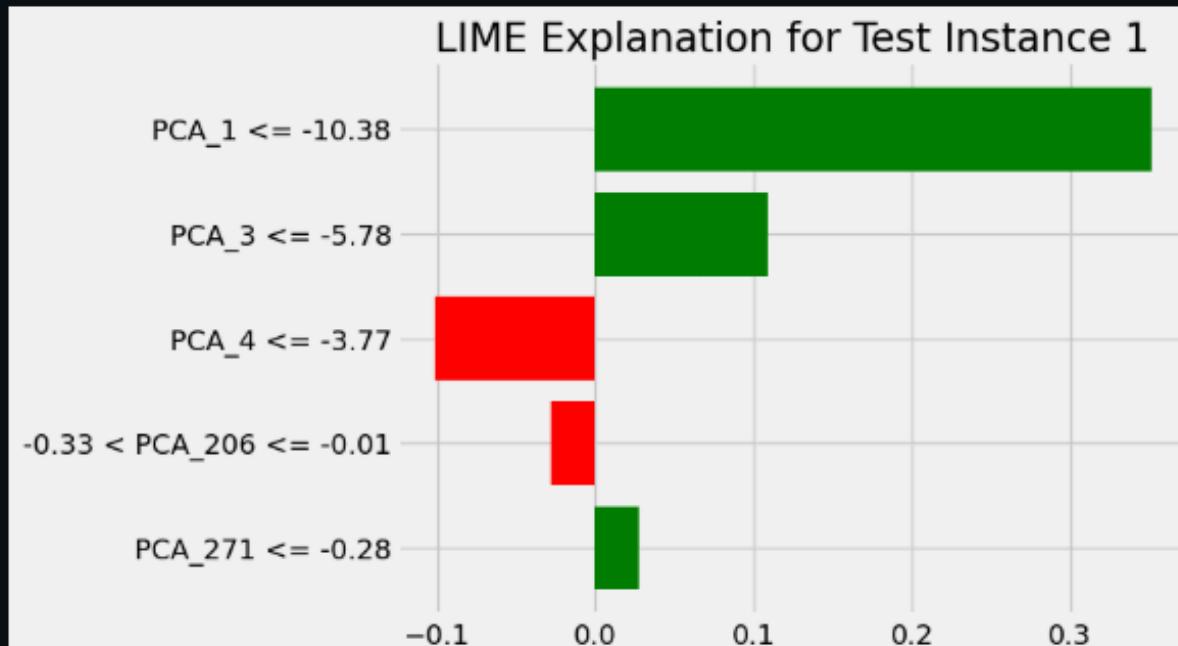
LIME Explanations for Selected Test Instances:

Explanation for test instance 0 (True label: meningioma, Predicted: meningioma)
PCA_1 \leq -10.38: 0.3679
PCA_2 \leq -6.09: 0.1280
PCA_4 > 2.52: 0.1139
PCA_3 \leq -5.78: 0.1117
-0.01 < PCA_82 \leq 0.55: 0.0270



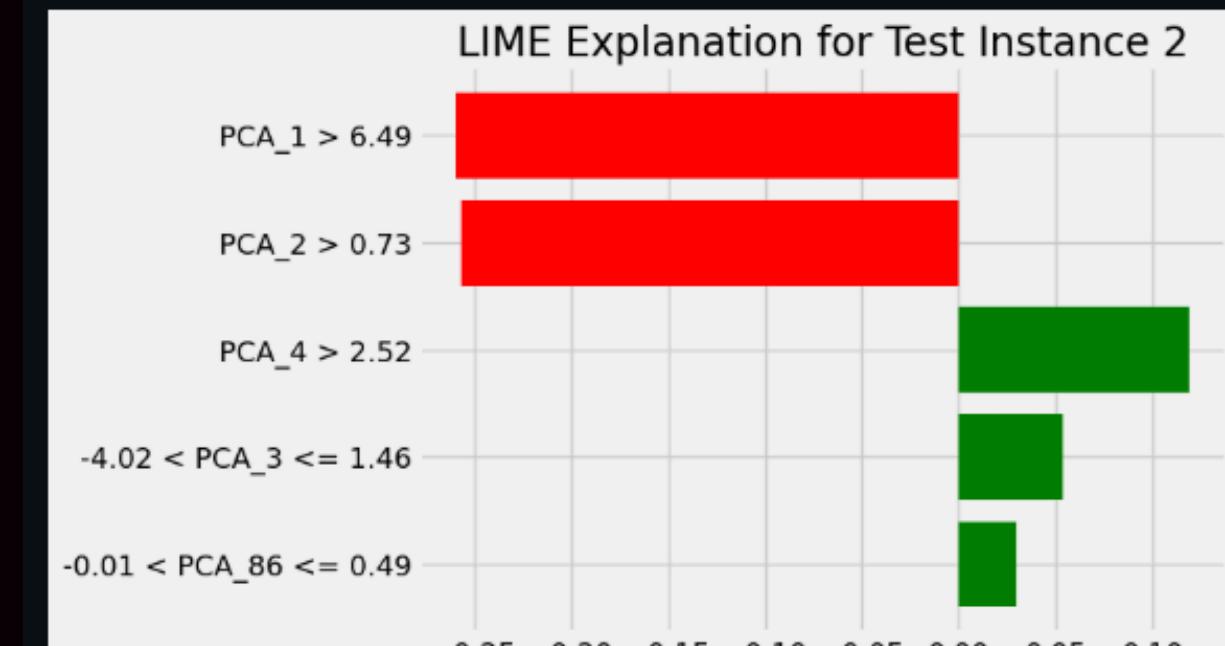
Explanation for test instance 1 (True label: meningioma, Predicted: meningioma)

PCA_1 \leq -10.38: 0.3511
PCA_3 \leq -5.78: 0.1087
PCA_4 \leq -3.77: -0.1018
-0.33 < PCA_206 \leq -0.01: -0.0276
PCA_271 \leq -0.28: 0.0275

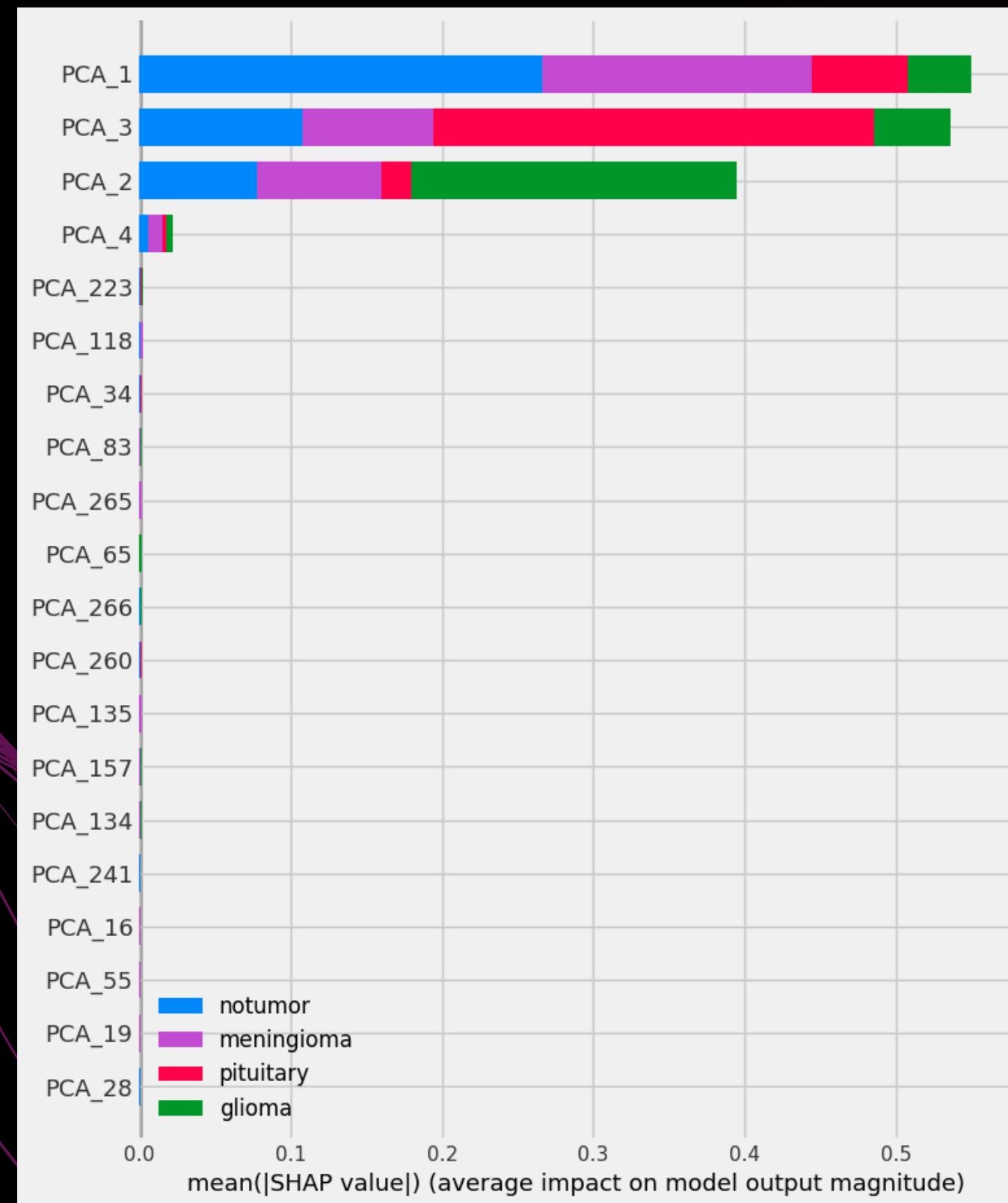


Explanation for test instance 2 (True label: glioma, Predicted: glioma)

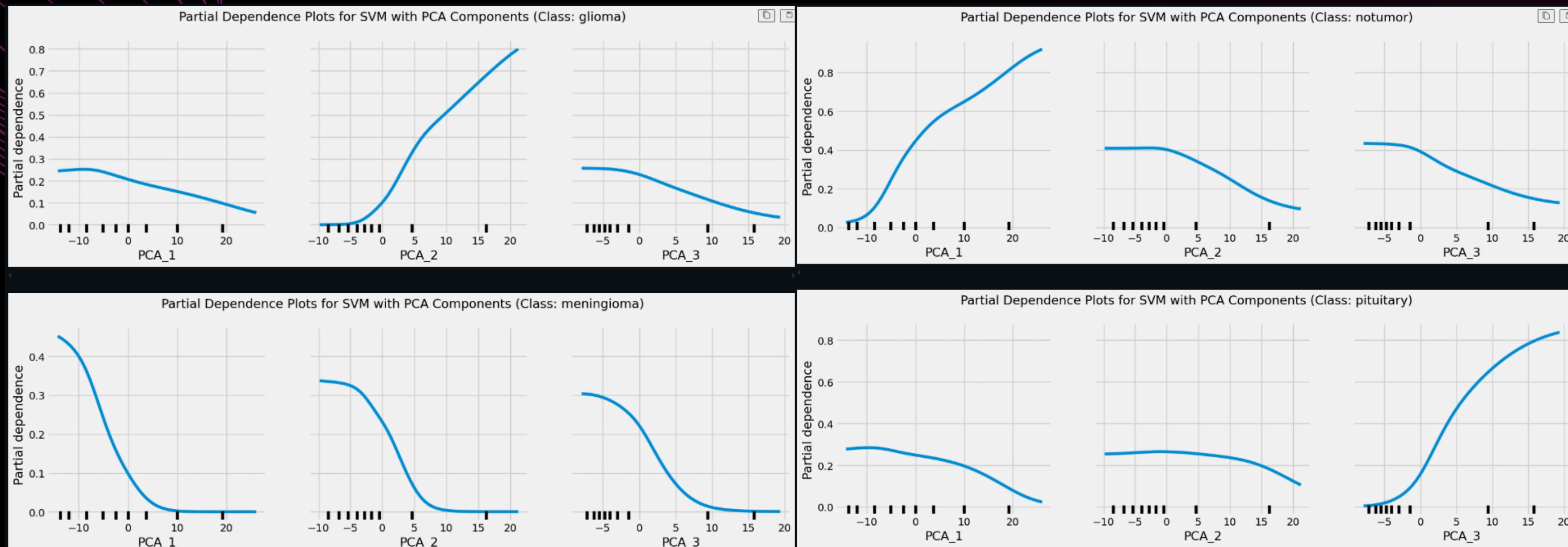
PCA_1 > 6.49: -0.2595
PCA_2 > 0.73: -0.2570
PCA_4 > 2.52: 0.1189
-4.02 < PCA_3 \leq 1.46: 0.0539
-0.01 < PCA_86 \leq 0.49: 0.0296



Model 1 - XAI - SHAP



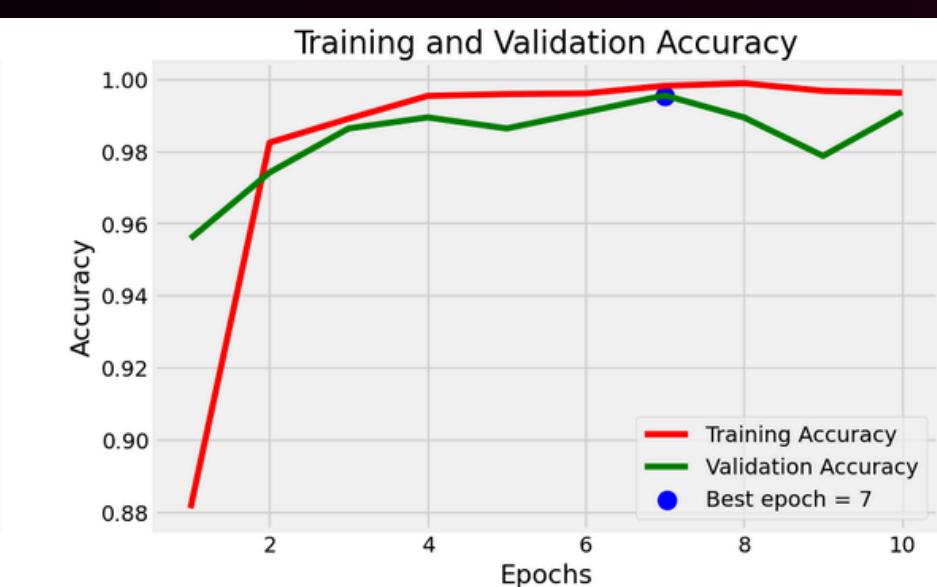
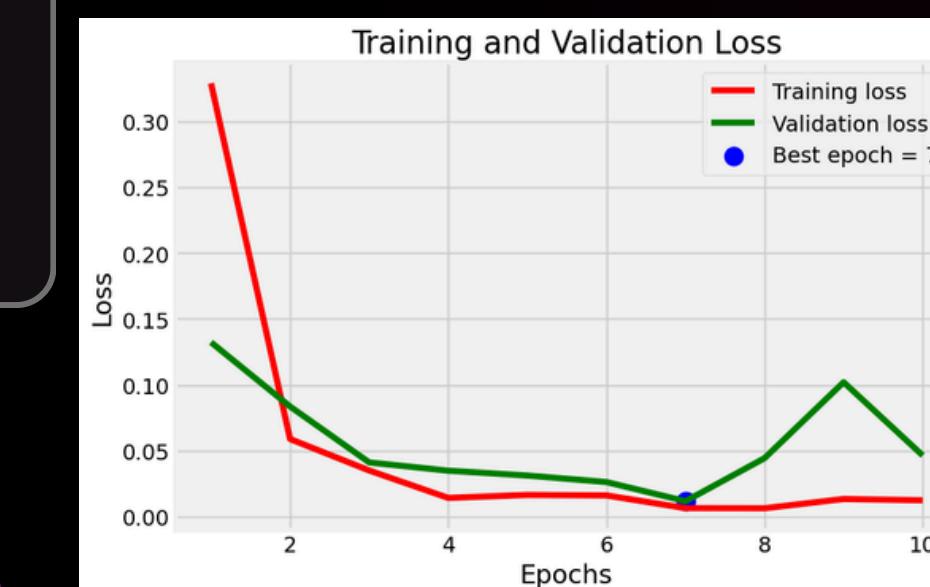
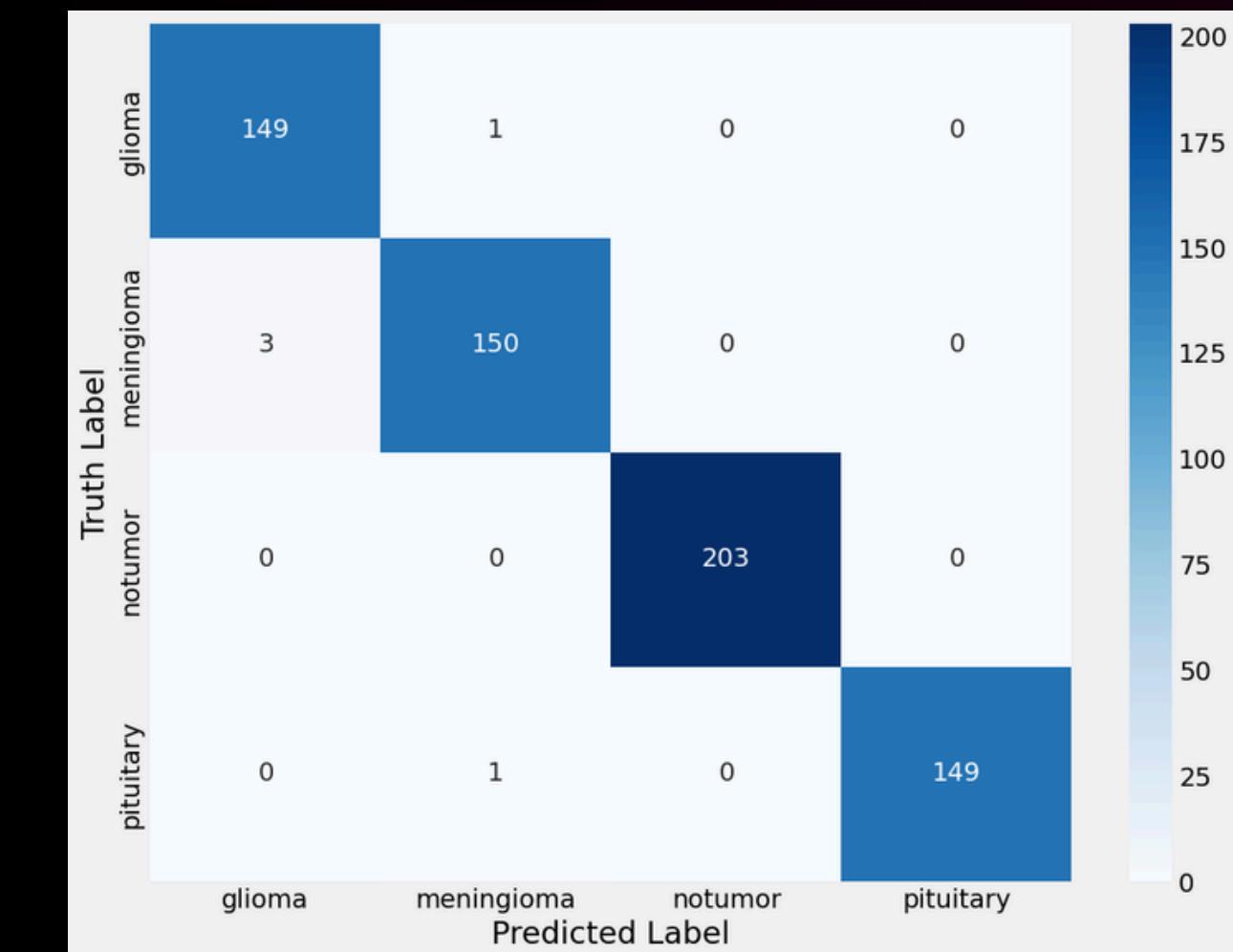
Model 1 - XAI - PCA



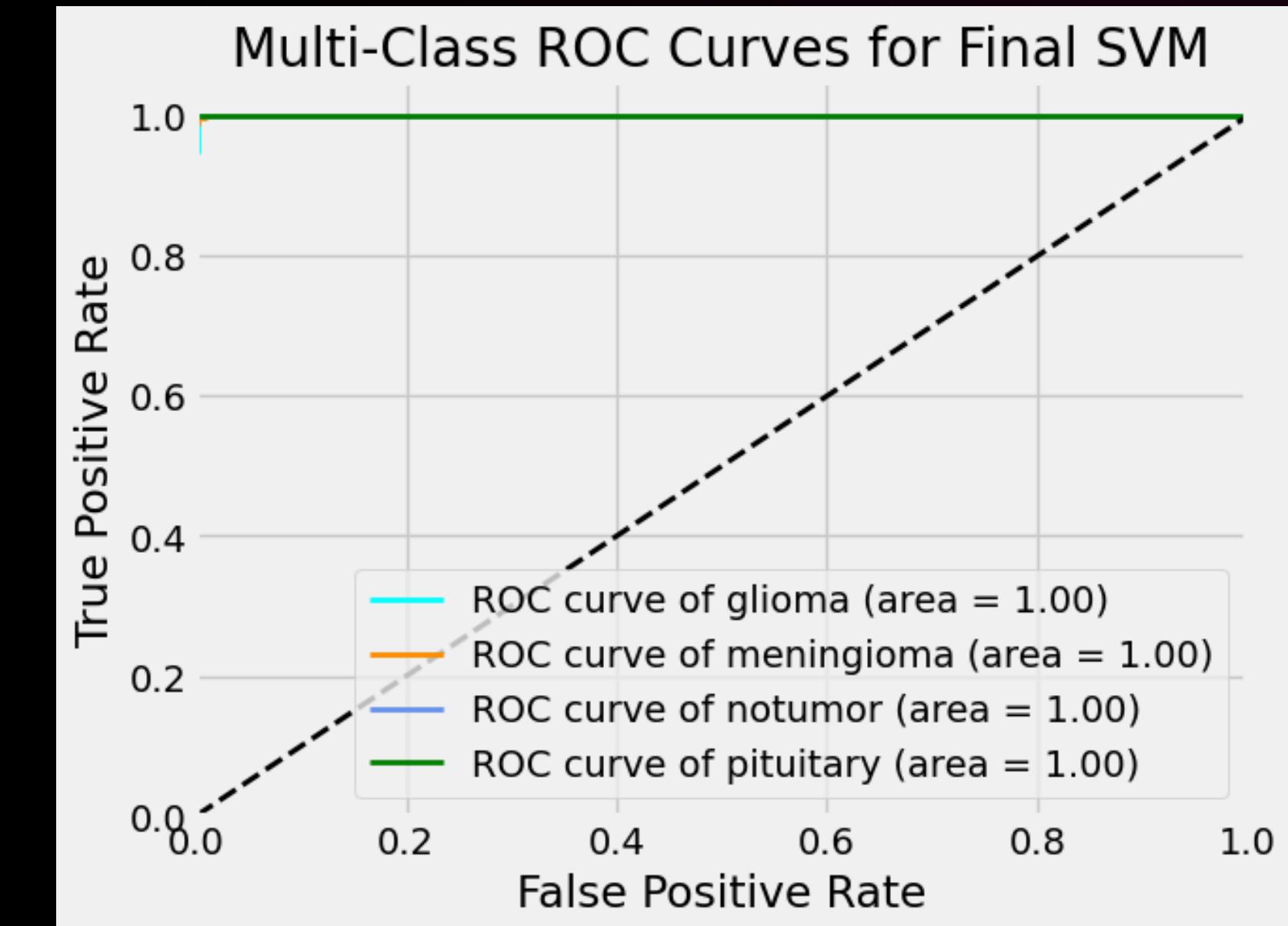
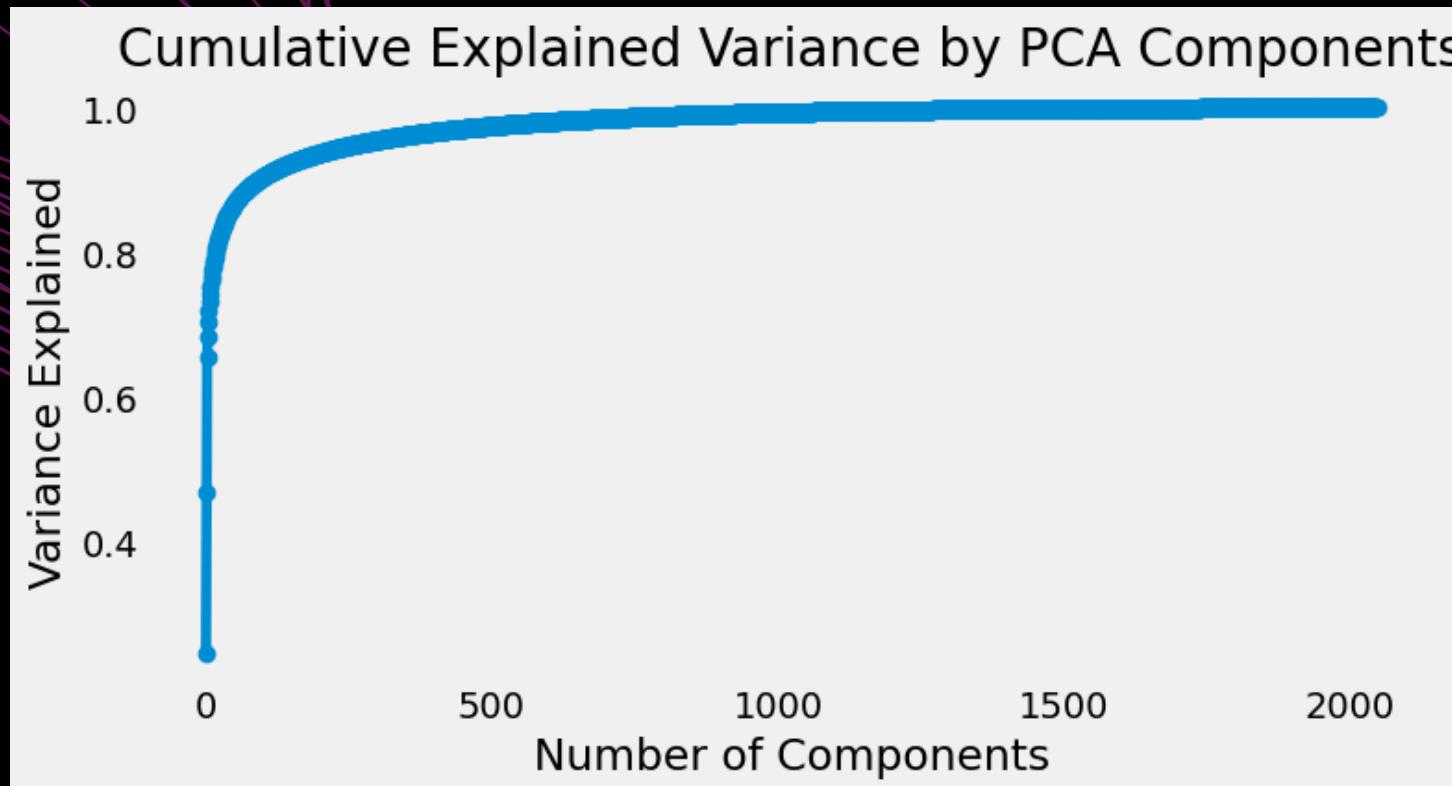
Ahmed Barakat

Model 2 - Model 2 - PCA 266 → SVM (Best Accuracy)

- RF select → PCA 266 dims (optimal)
- SVM RBF ($C=10$, $\gamma=\text{scale}$), Accel. GridSearch in 23s
- Test accuracy 99.24% (best in team)
- LIME & SHAP show PCA₂ (meningioma), PCA₃ (pituitary), PCA₁ (global)

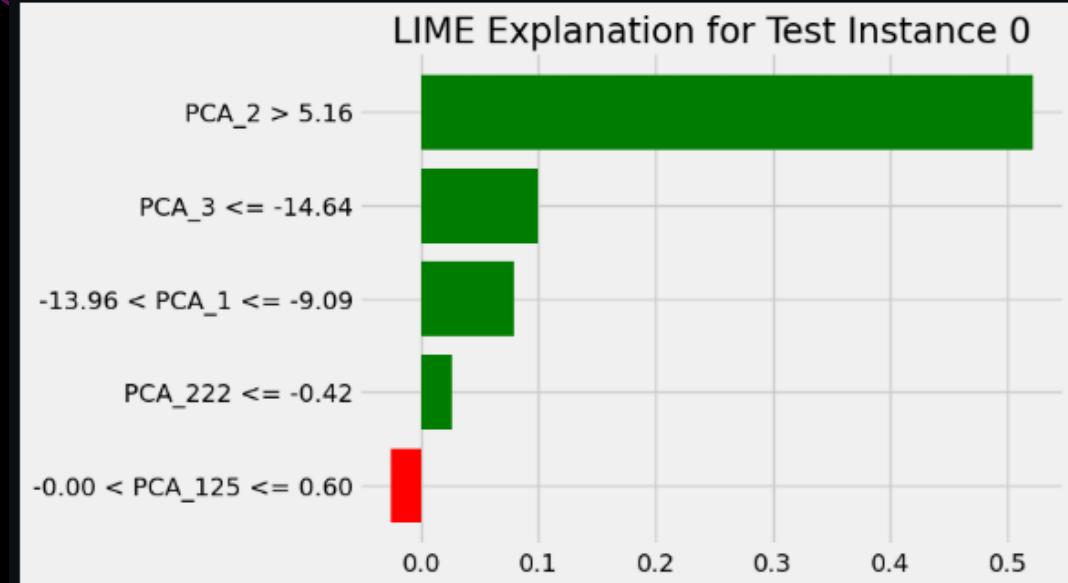


Model 2 – XAI

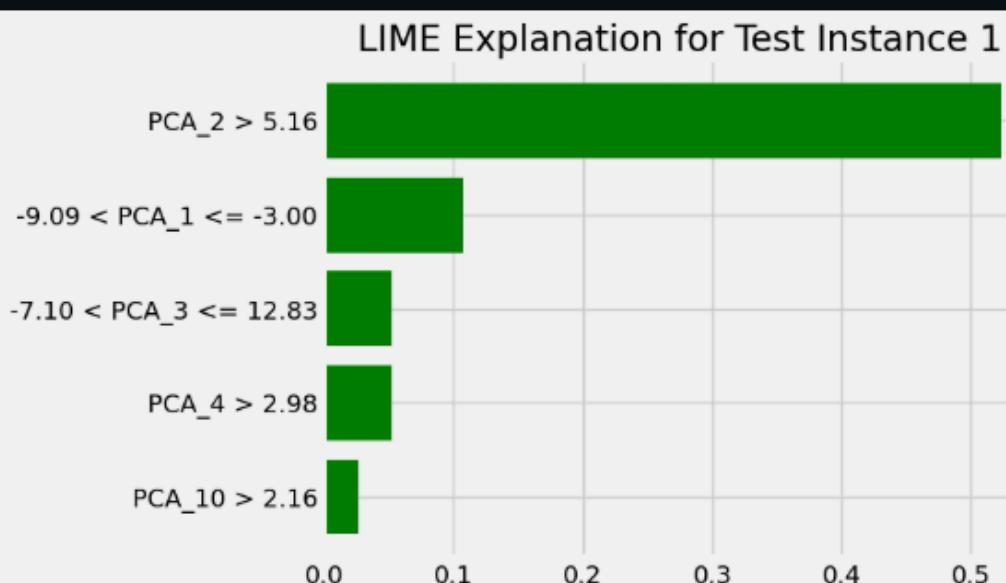


Model 2 – XAI

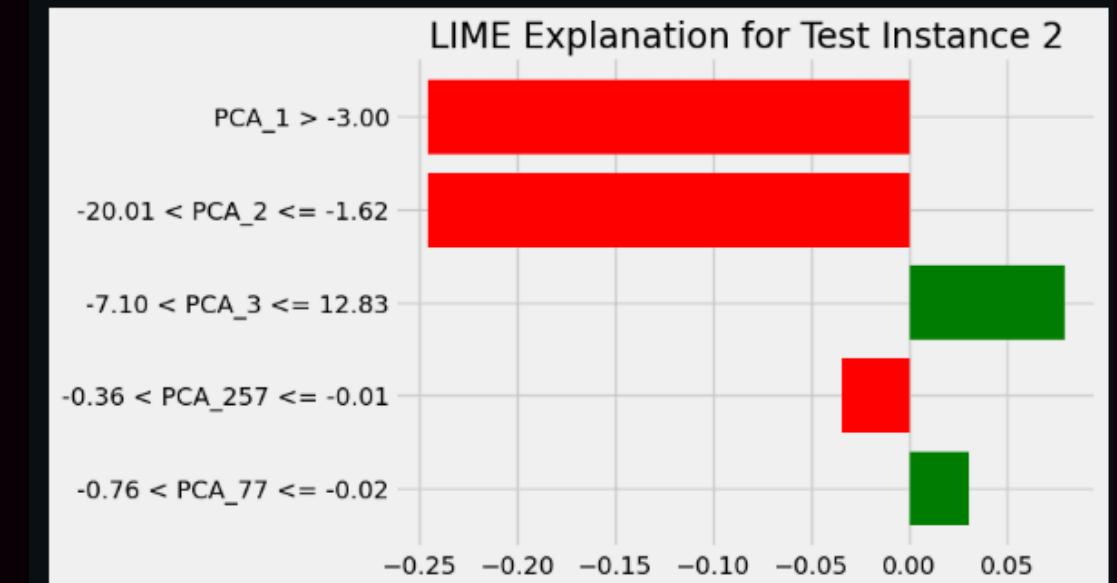
Explanation for test instance 0 (True label: meningioma, Predicted: meningioma)
PCA_2 > 5.16: 0.5214
PCA_3 <= -14.64: 0.0999
-13.96 < PCA_1 <= -9.09: 0.0796
PCA_222 <= -0.42: 0.0261
-0.00 < PCA_125 <= 0.60: -0.0252



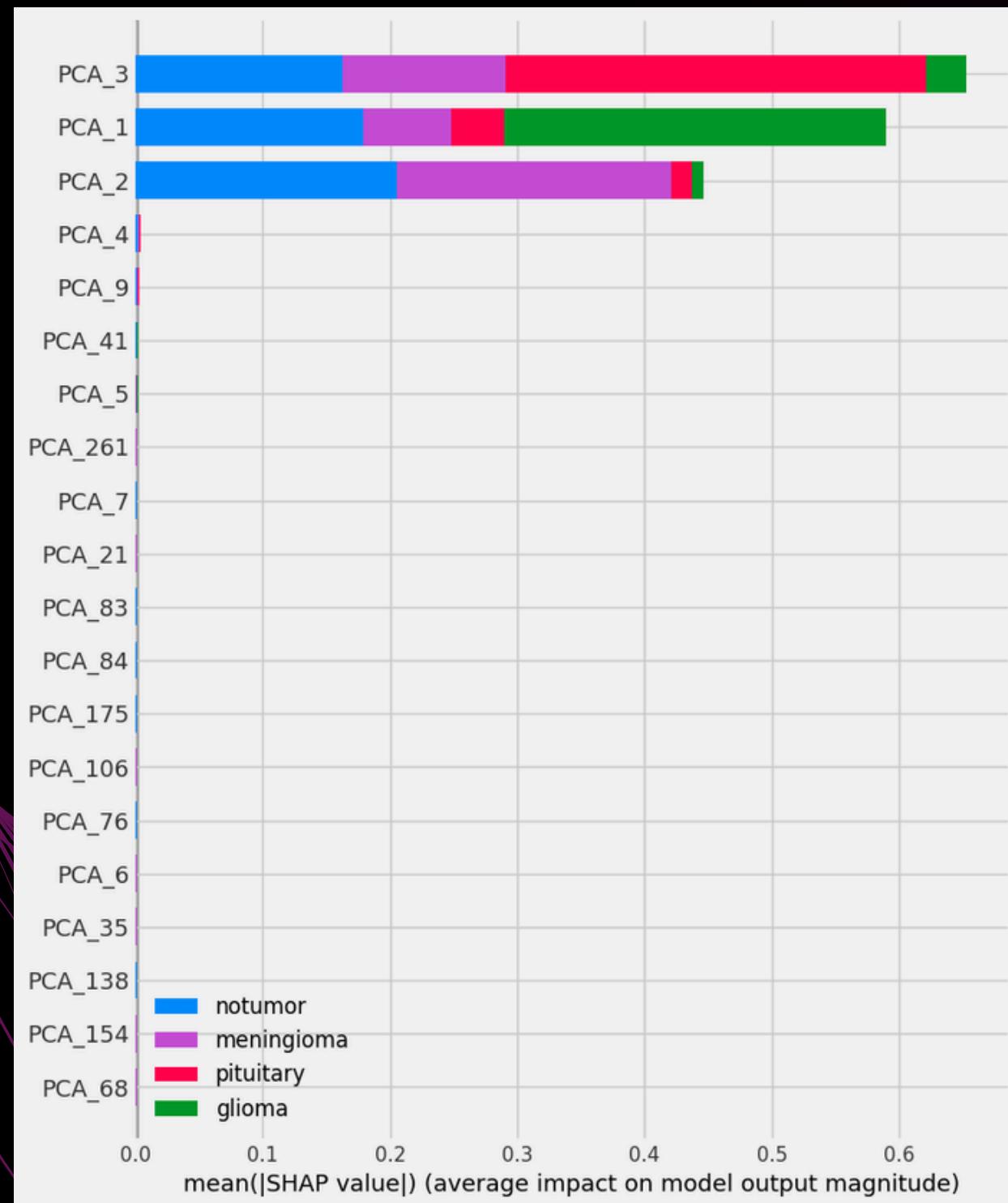
Explanation for test instance 1 (True label: meningioma, Predicted: meningioma)
PCA_2 > 5.16: 0.5237
-9.09 < PCA_1 <= -3.00: 0.1075
-7.10 < PCA_3 <= 12.83: 0.0522
PCA_4 > 2.98: 0.0518
PCA_10 > 2.16: 0.0265



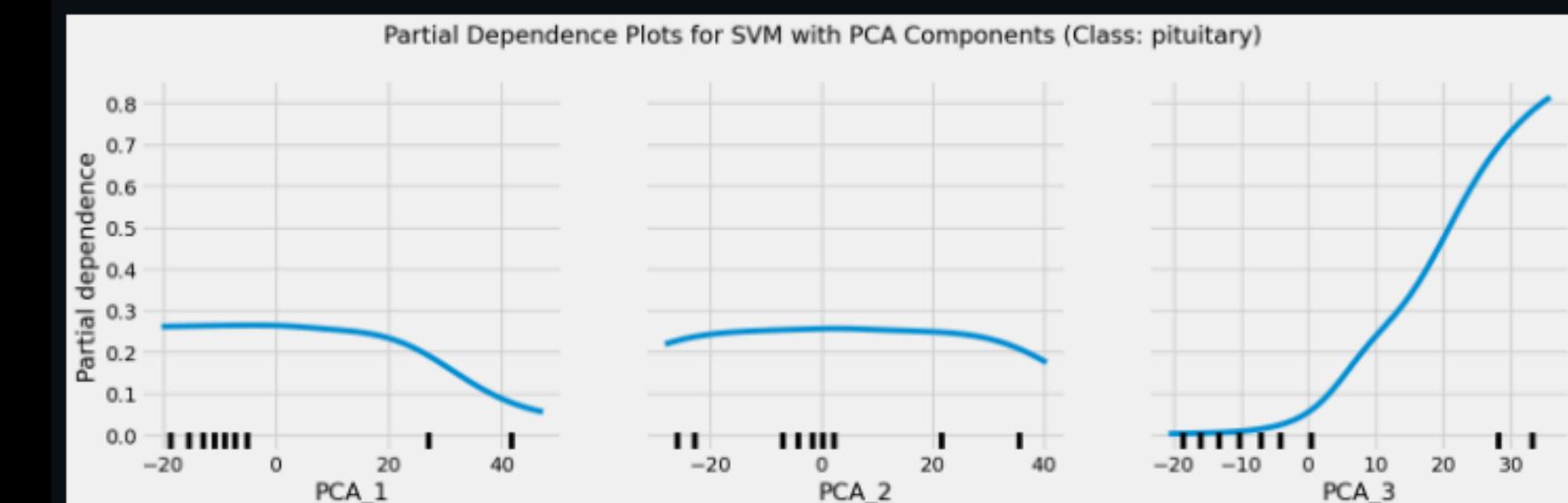
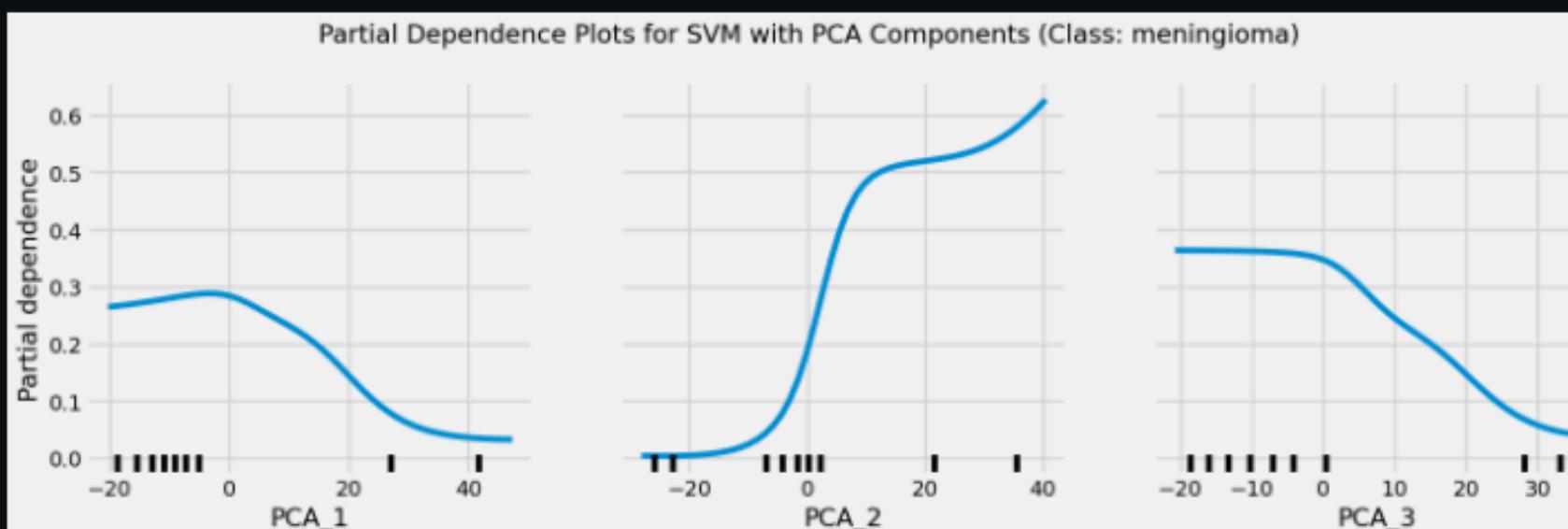
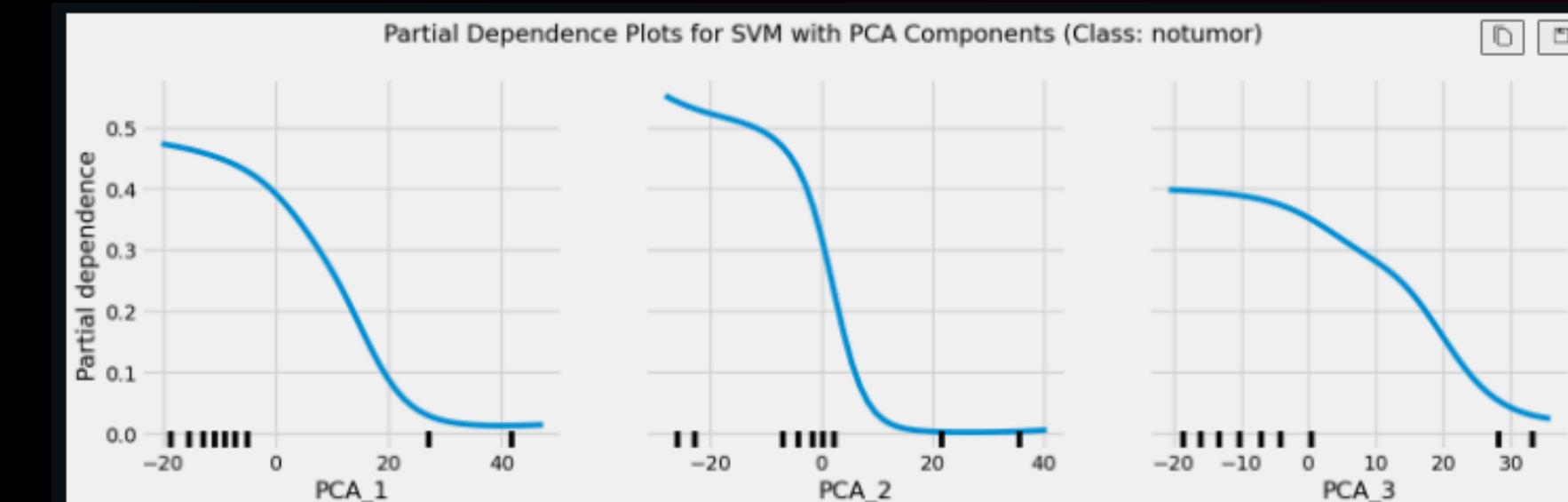
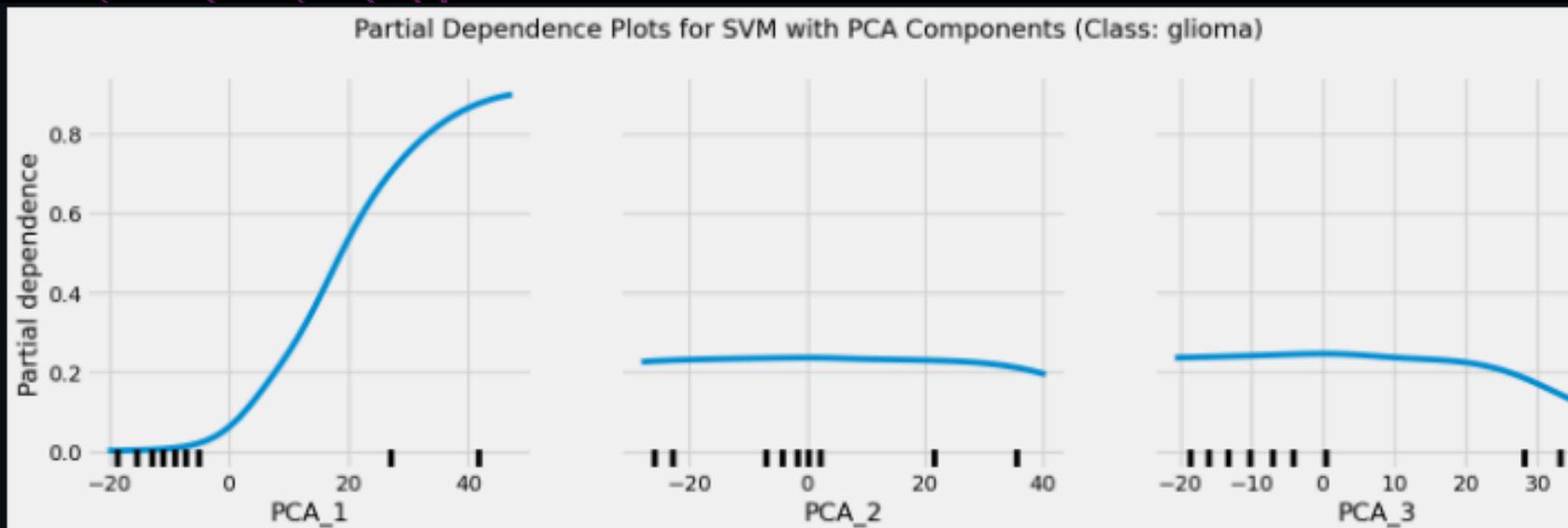
Explanation for test instance 2 (True label: glioma, Predicted: glioma)
PCA_1 > -3.00: -0.2459
-20.01 < PCA_2 <= -1.62: -0.2457
-7.10 < PCA_3 <= 12.83: 0.0792
-0.36 < PCA_257 <= -0.01: -0.0346
-0.76 < PCA_77 <= -0.02: 0.0303



Model 2 – XAI



Model 2 – XAI



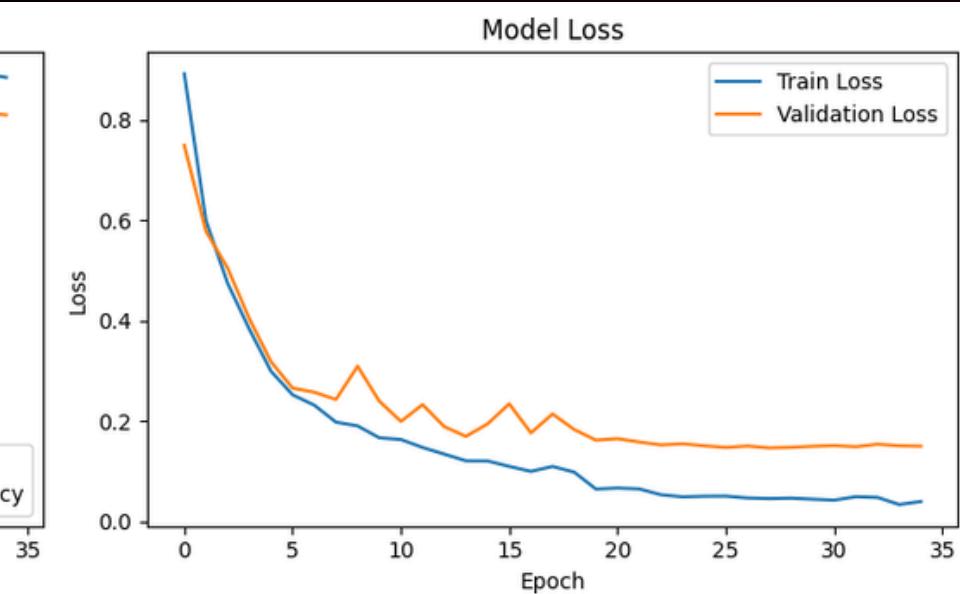
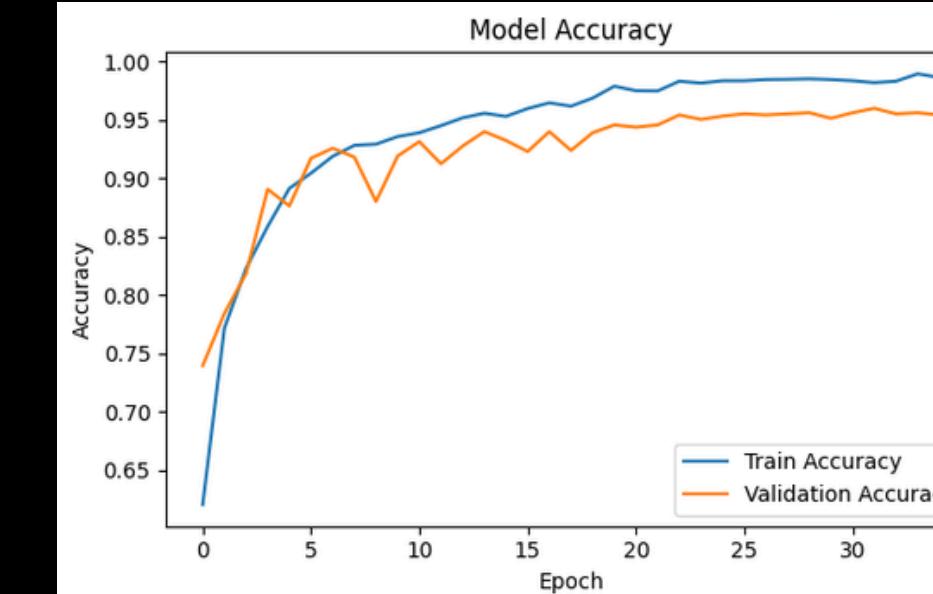
Ahmed Barakat

Model 3 – CNN → PCA → SVM

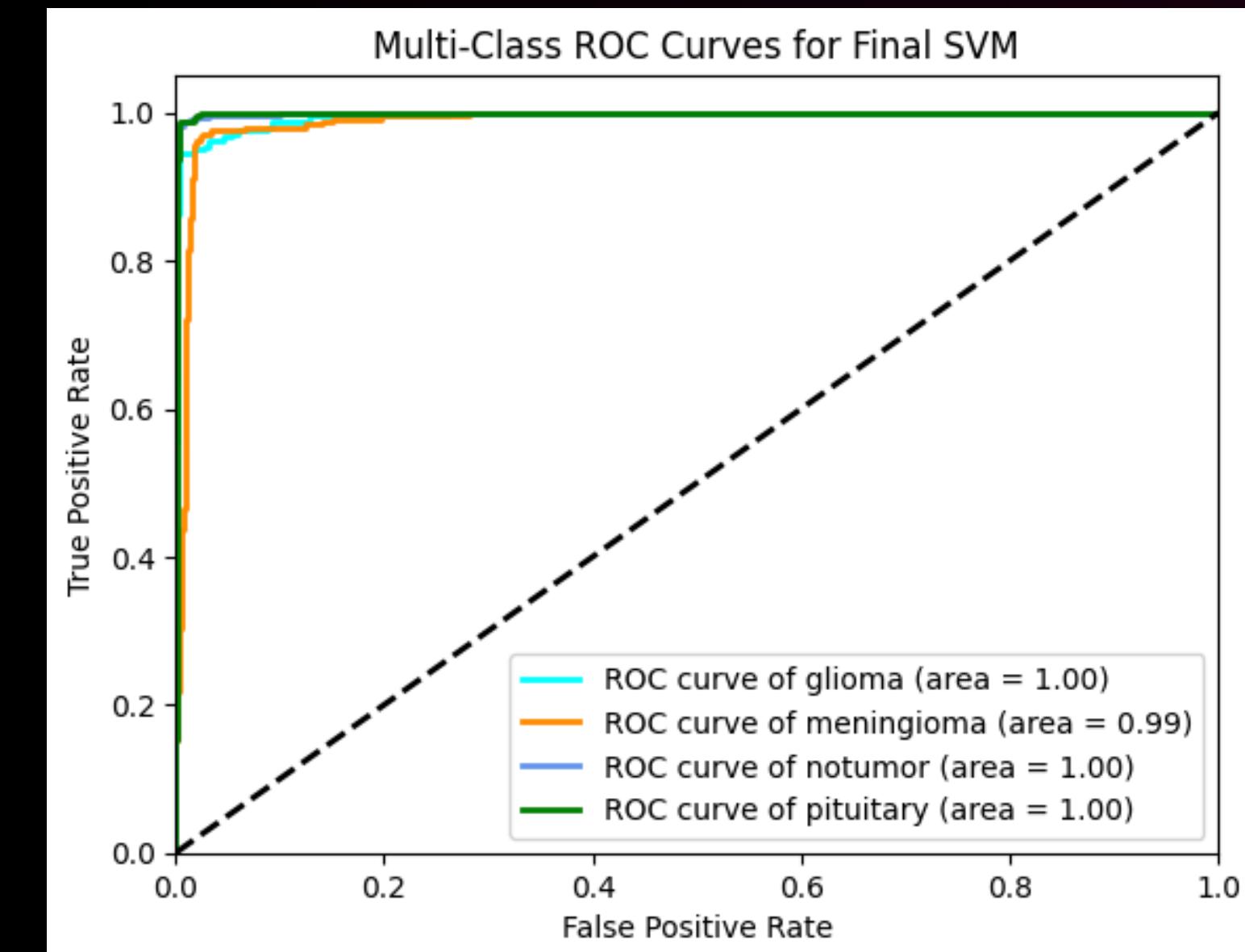
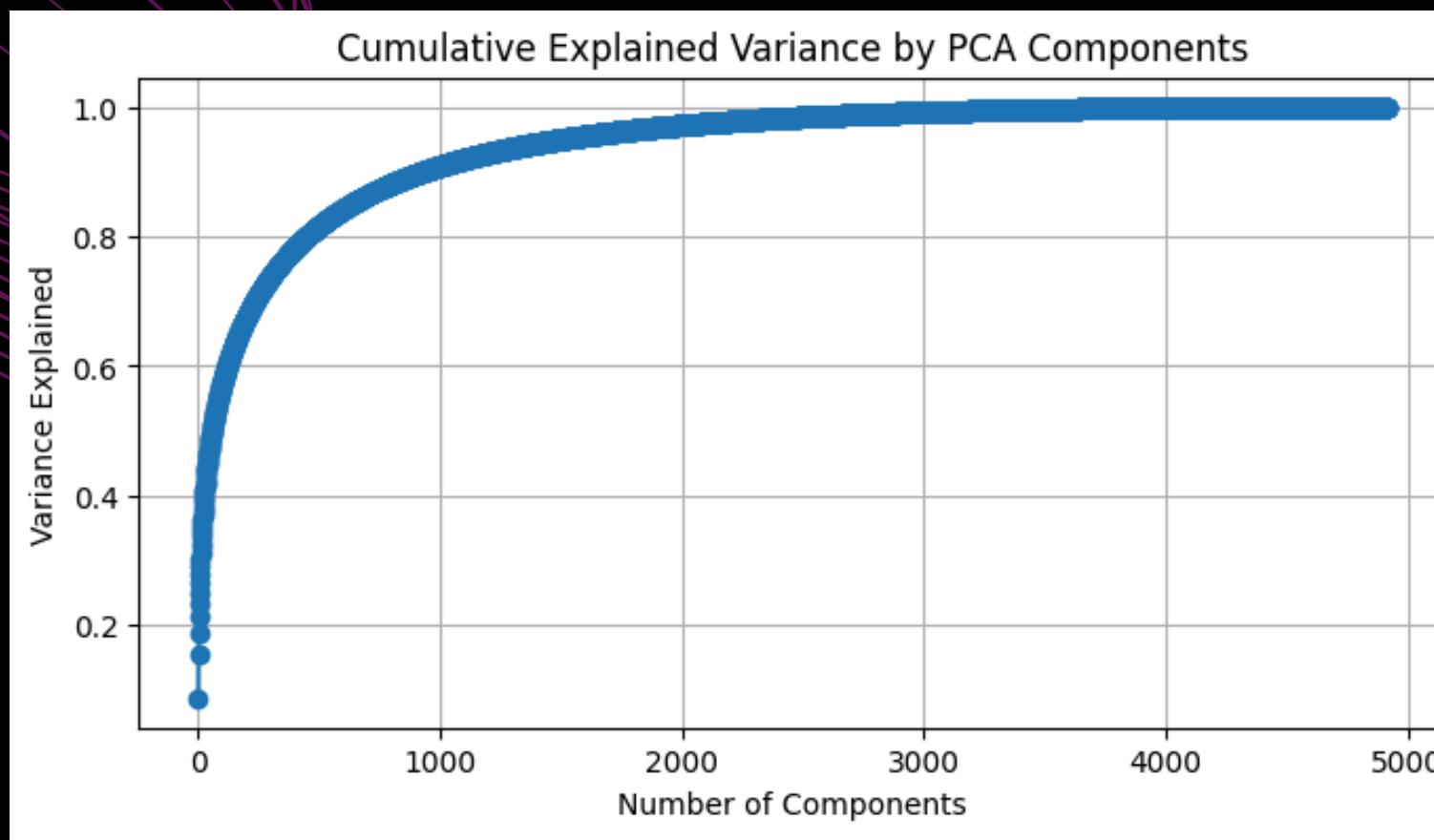
- CNN trained 35 epochs, Acc 94.3%
- CNN-features → PCA → SVM lifts to 97.2%
- Top 3 PCs explain >85% decision power
- ROC AUC: 0.99+ for all classes

CNN Classification Report:

	precision	recall	f1-score	support
glioma	0.93	0.93	0.93	243
meningioma	0.87	0.91	0.89	247
notumor	0.98	0.96	0.97	300
pituitary	0.98	0.96	0.97	264
accuracy			0.94	1054
macro avg	0.94	0.94	0.94	1054
weighted avg	0.94	0.94	0.94	1054



Model 3 – XAI

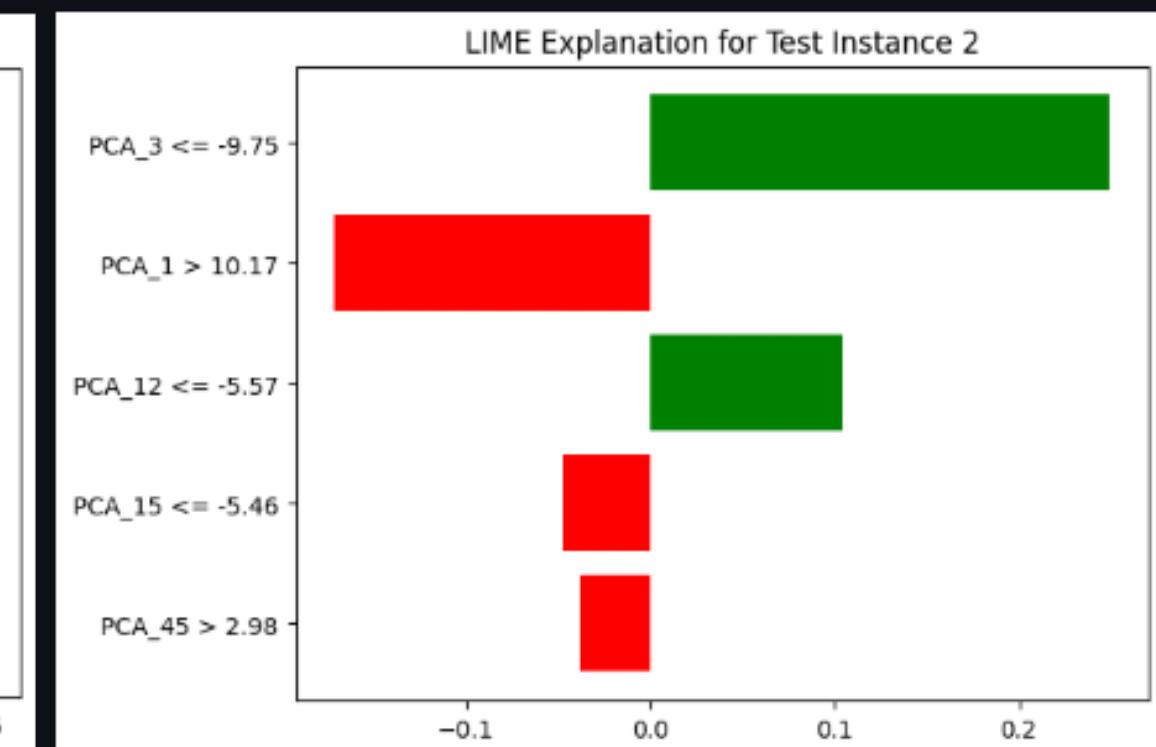
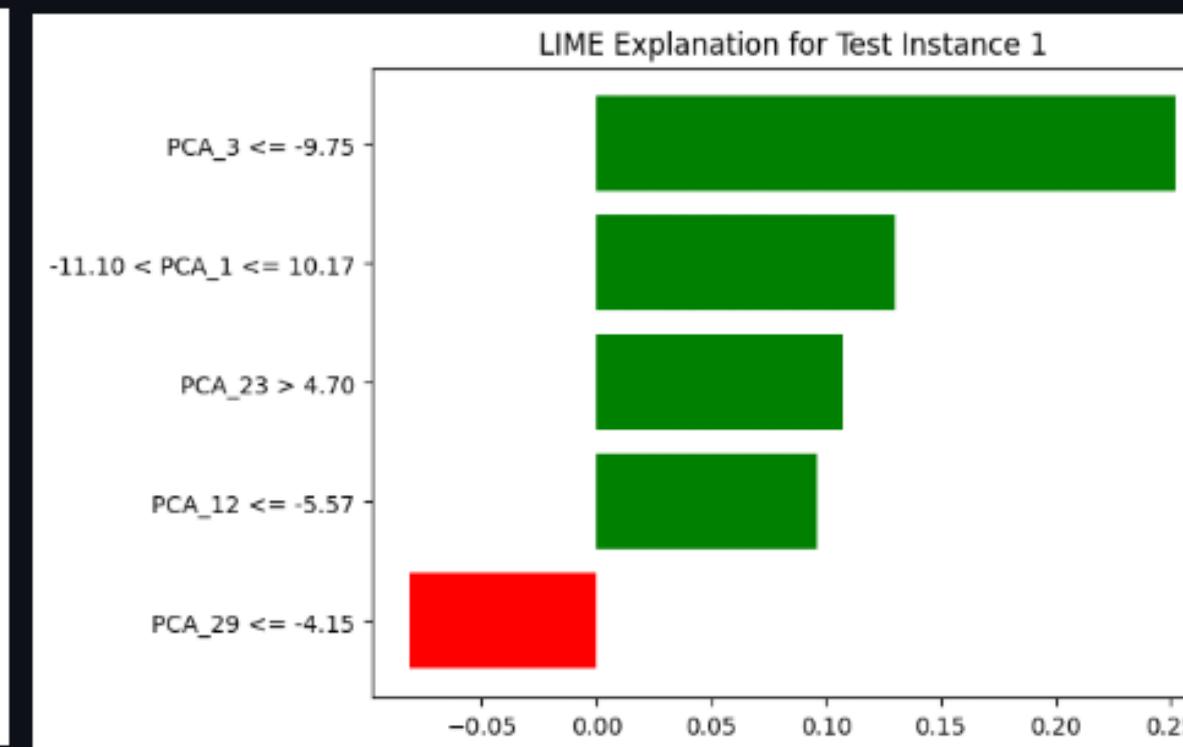
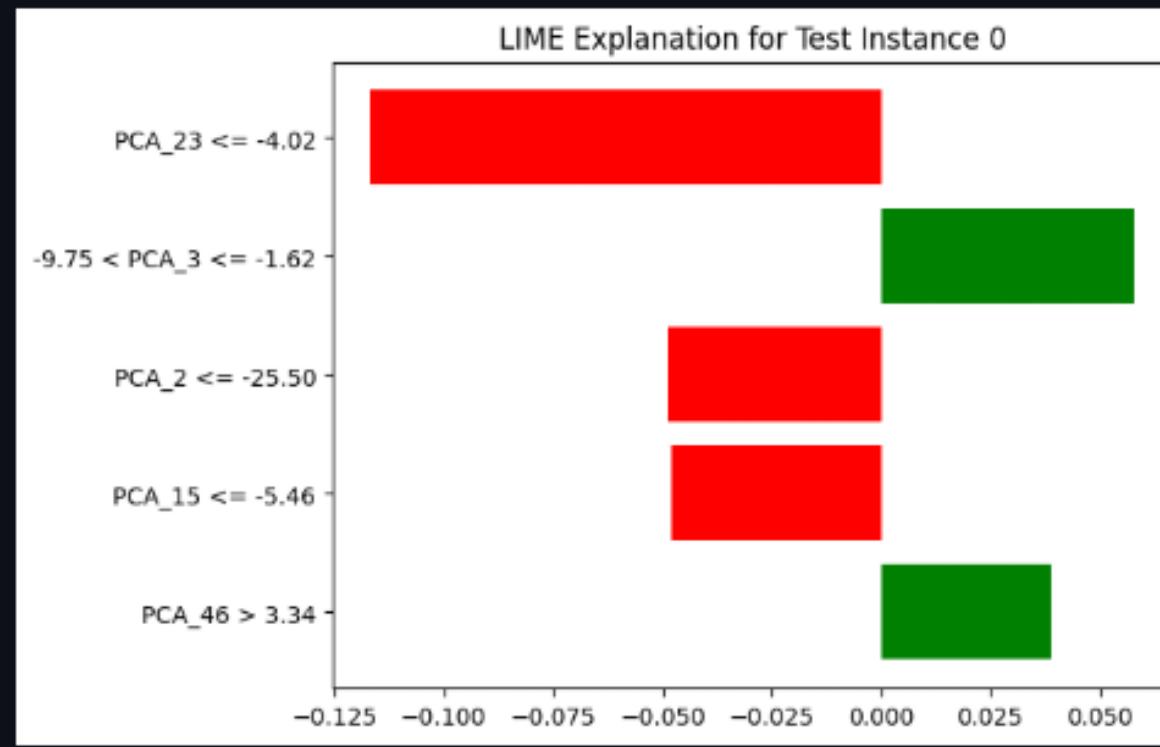


Model 3 – XAI

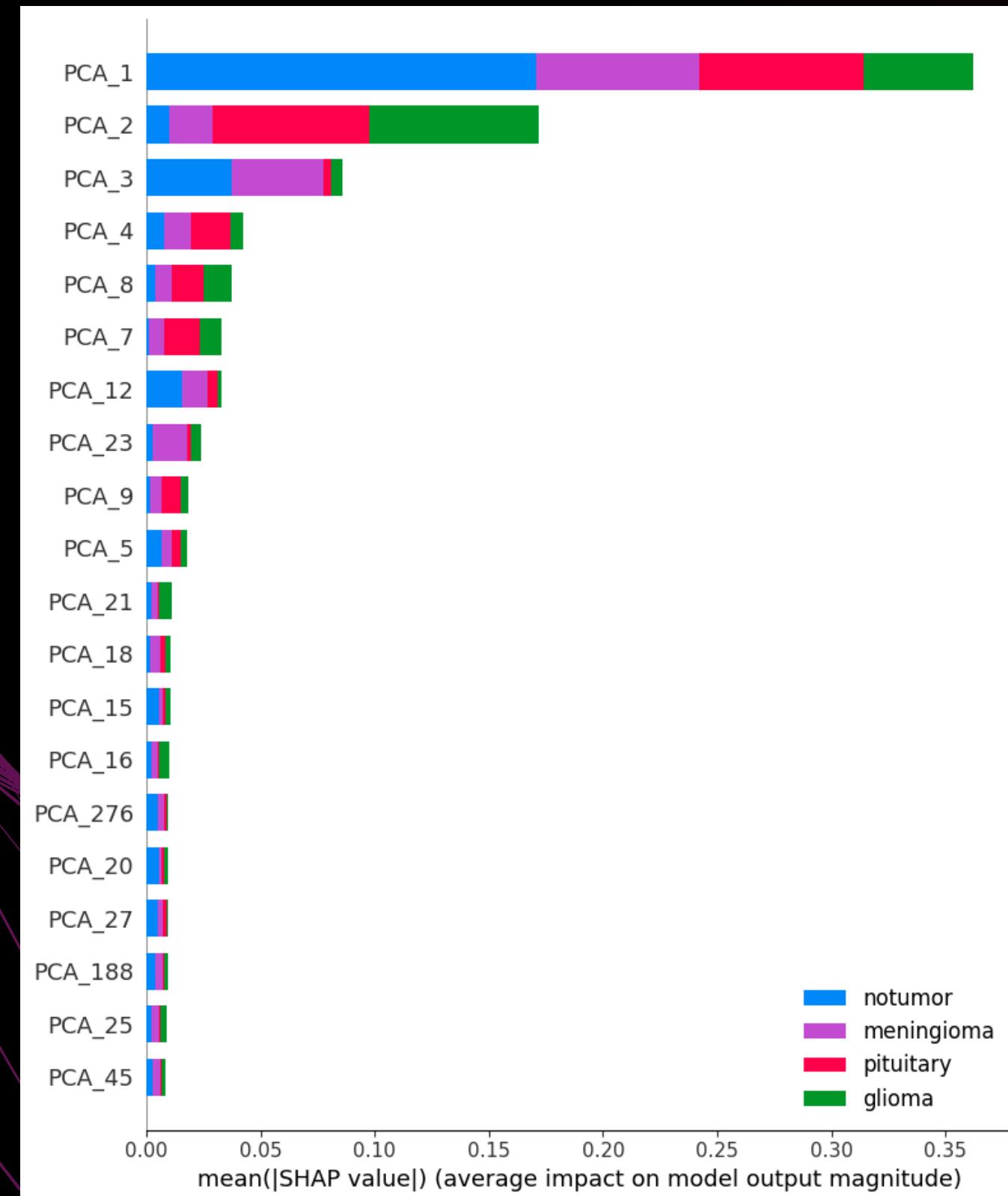
Explanation for test instance 0 (True label: glioma, Predicted: glioma)
PCA_23 \leq -4.02: -0.1167
-9.75 < PCA_3 \leq -1.62: 0.0580
PCA_2 \leq -25.50: -0.0485
PCA_15 \leq -5.46: -0.0481
PCA_46 > 3.34: 0.0391

Explanation for test instance 1 (True label: meningioma, Predicted: meningioma)
PCA_3 \leq -9.75: 0.2522
-11.10 < PCA_1 \leq 10.17: 0.1298
PCA_23 > 4.70: 0.1073
PCA_12 \leq -5.57: 0.0964
PCA_29 \leq -4.15: -0.0810

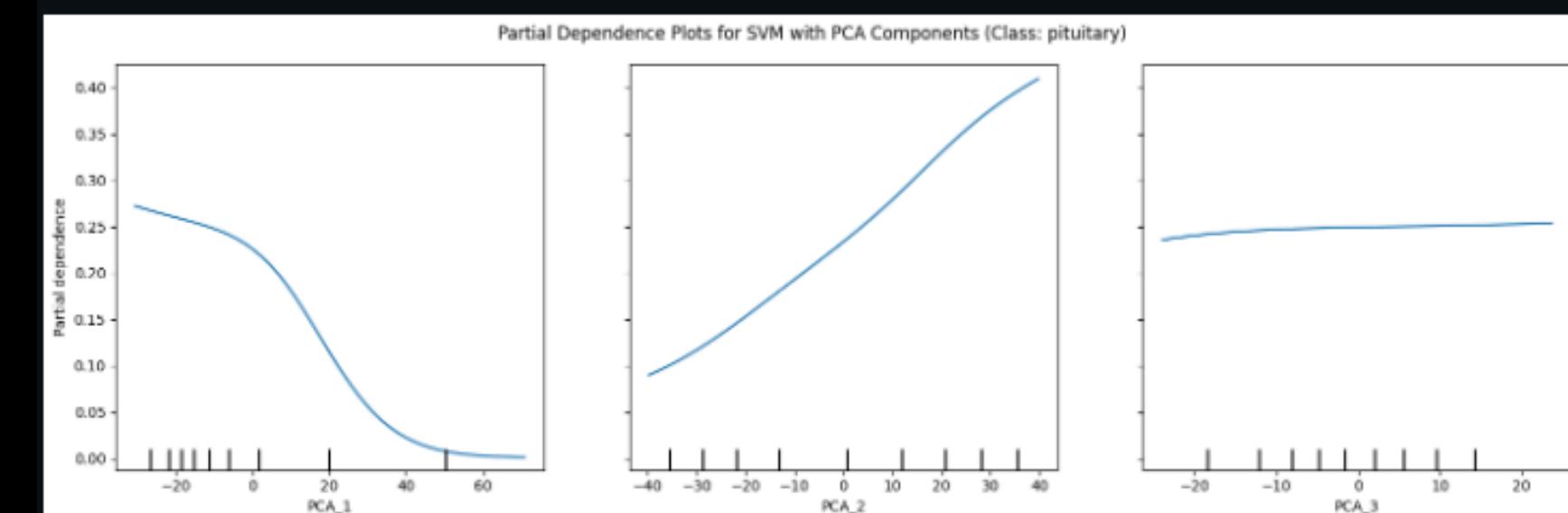
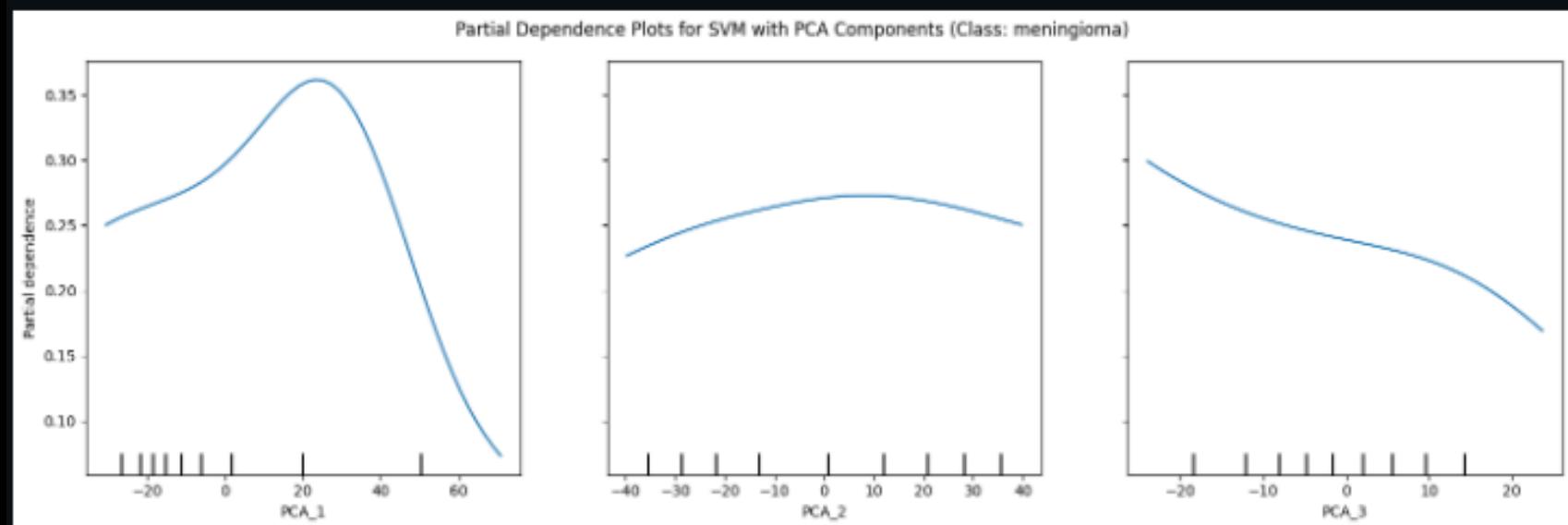
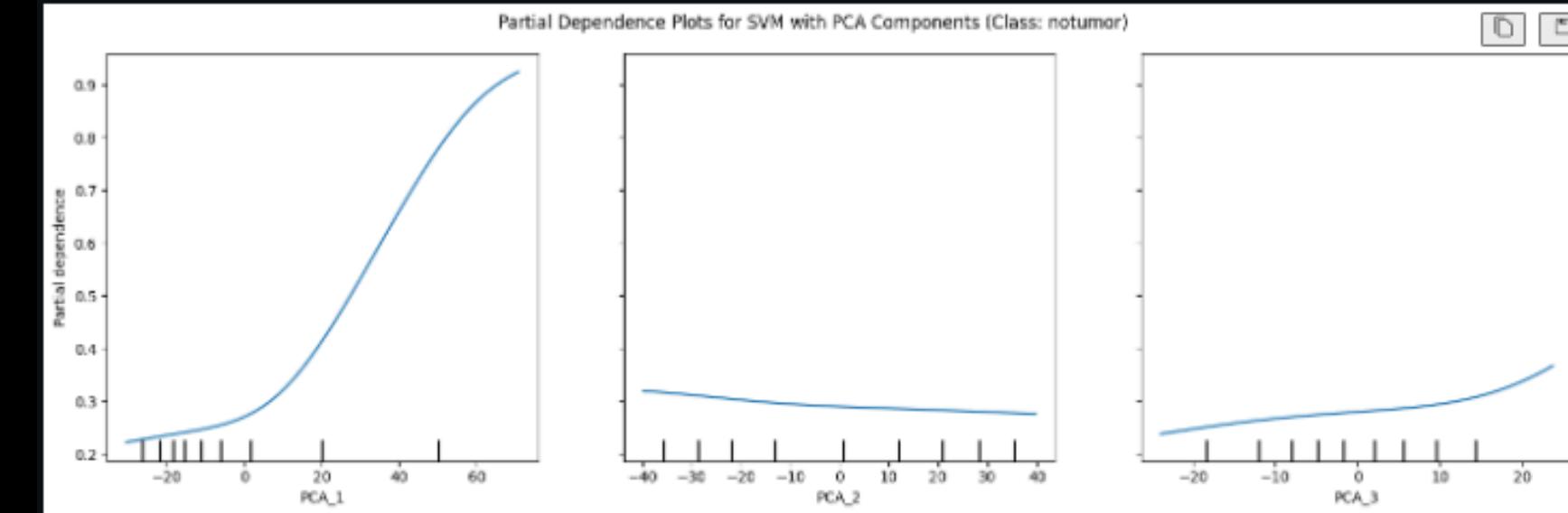
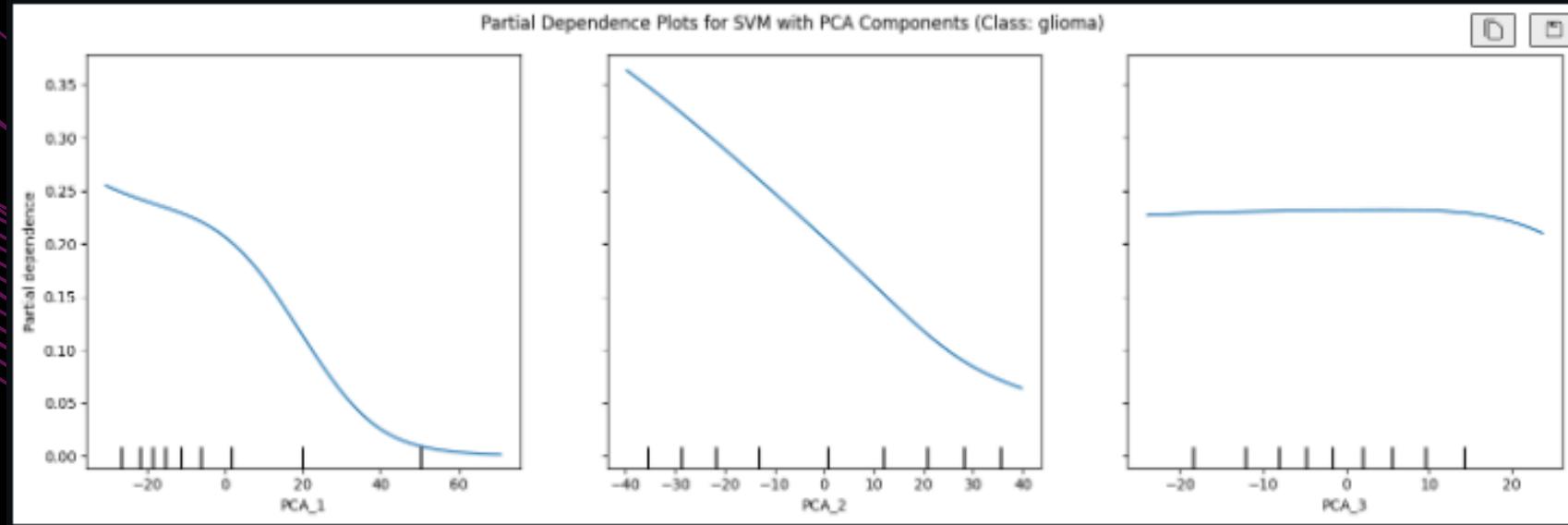
Explanation for test instance 2 (True label: notumor, Predicted: notumor)
PCA_3 \leq -9.75: 0.2489
PCA_1 > 10.17: -0.1711
PCA_12 \leq -5.57: 0.1048
PCA_15 \leq -5.46: -0.0477
PCA_45 > 2.98: -0.0383



Model 3 – XAI



Model 3 – XAI



Toka Mokhtar

Model1 - CNN → RF → PCA(100) → SVM

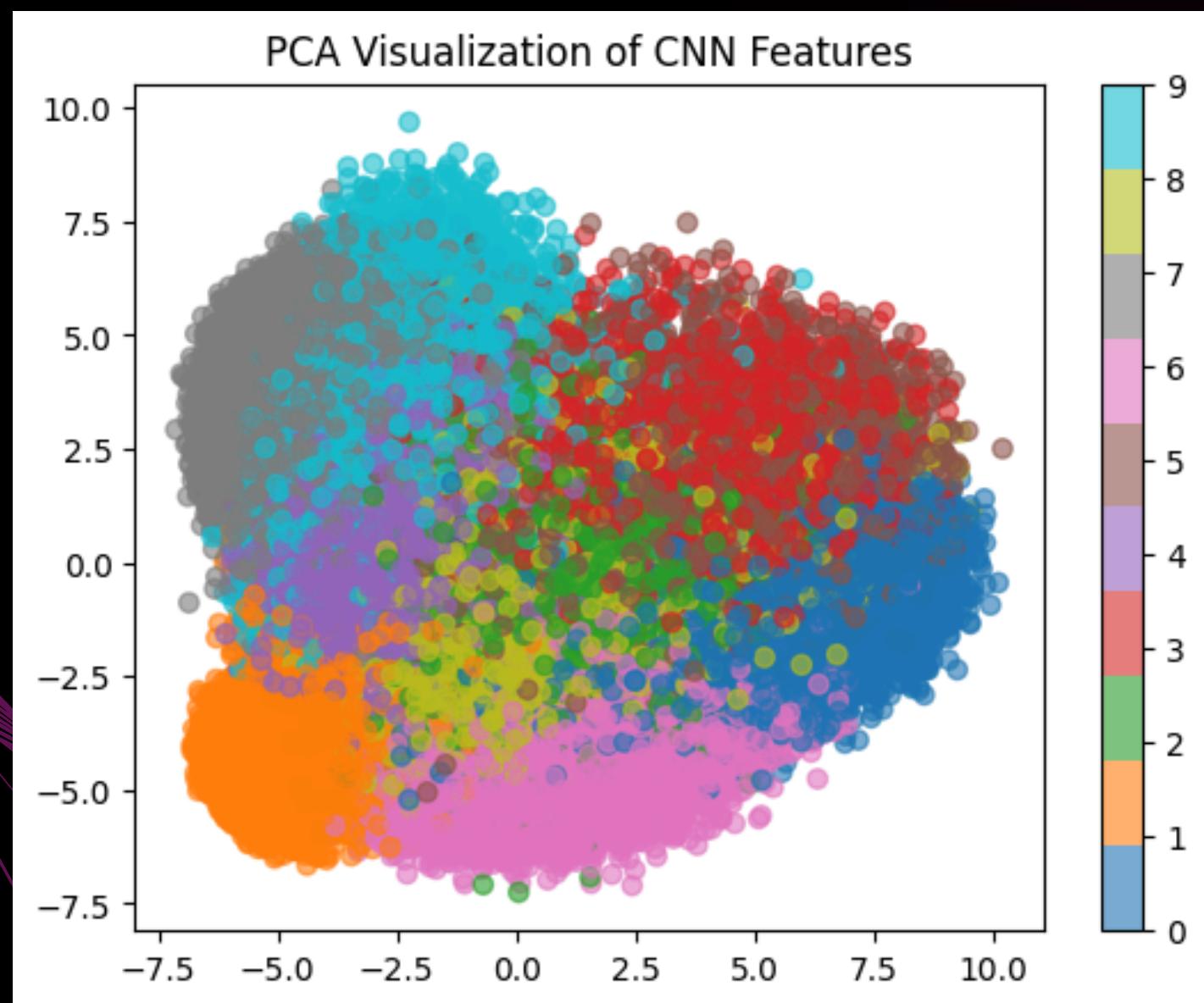
- CNN 315k params → Dense(128)
→ RF top 1152 features
- PCA(100) preserves cluster shapes, SVM RBF trained
- Accuracy 99.27%
- All classes: precision & recall ≥ 0.99
- Explainers: LIME + SHAP

SVC Accuracy: 0.9927

SVC Classification Report:

	precision	recall	f1-score	support
0	0.99	1.00	1.00	980
1	0.99	1.00	1.00	1135
2	0.99	0.99	0.99	1032
3	0.99	1.00	0.99	1010
4	0.99	0.99	0.99	982
5	1.00	0.99	0.99	892
6	1.00	0.99	0.99	958
7	0.99	0.99	0.99	1028
8	0.99	0.99	0.99	974
9	0.99	0.99	0.99	1009
accuracy			0.99	10000
macro avg	0.99	0.99	0.99	10000
weighted avg	0.99	0.99	0.99	10000

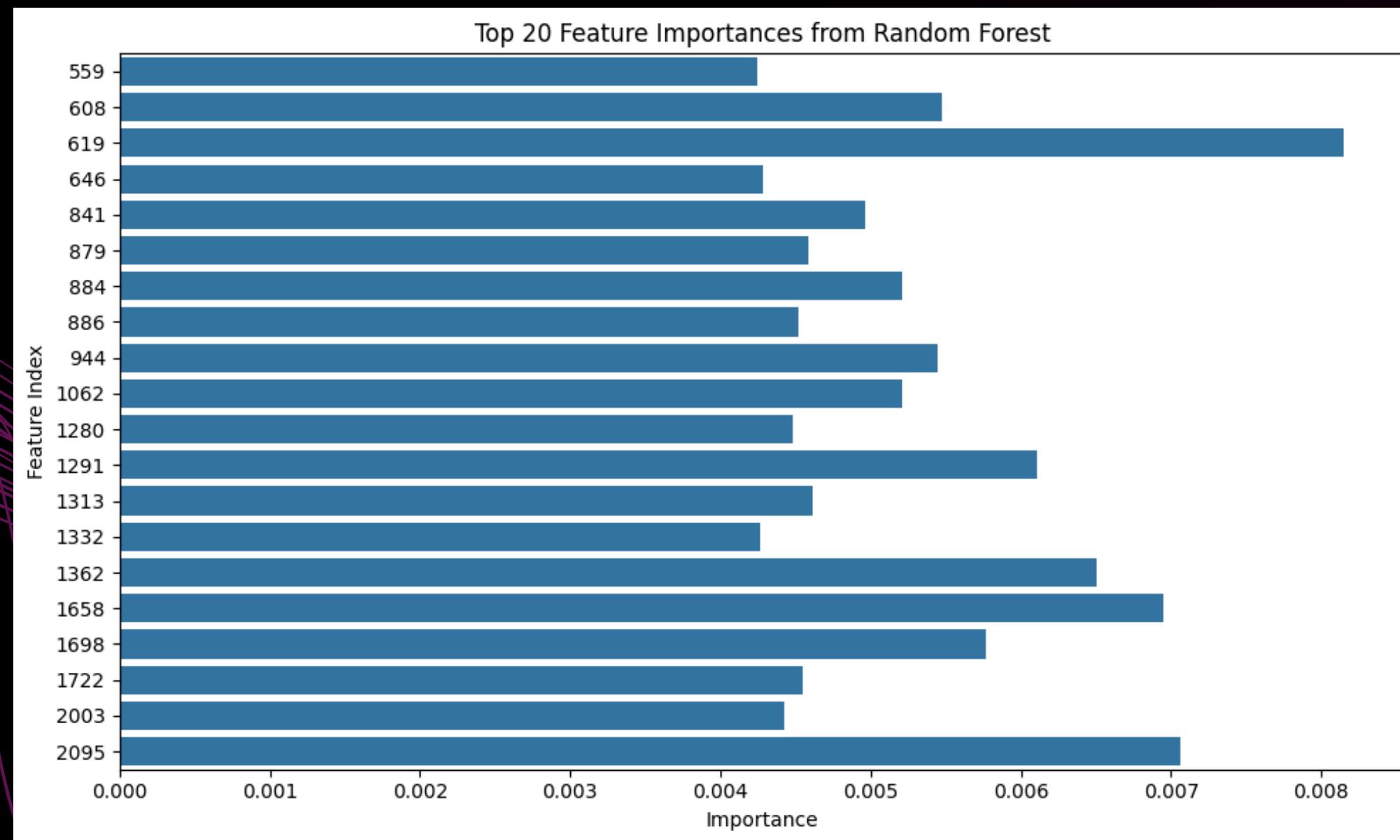
Model 1 – XAI



Model 1 – XAI



Model 1 - XAI



Toka Mokhtar

Model2 – VGG19 (fixed) → SVM

- VGG19 weights='imagenet', no top; flatten to 25088 features
- Optional: SelectKBest(k=300)
- SVM (RBF or linear) tested
- Full set: Accuracy 98.5 % on 2-class tumour / no-tumour
- Subset: 95 % accuracy with just 300 features

Accuracy: 0.985

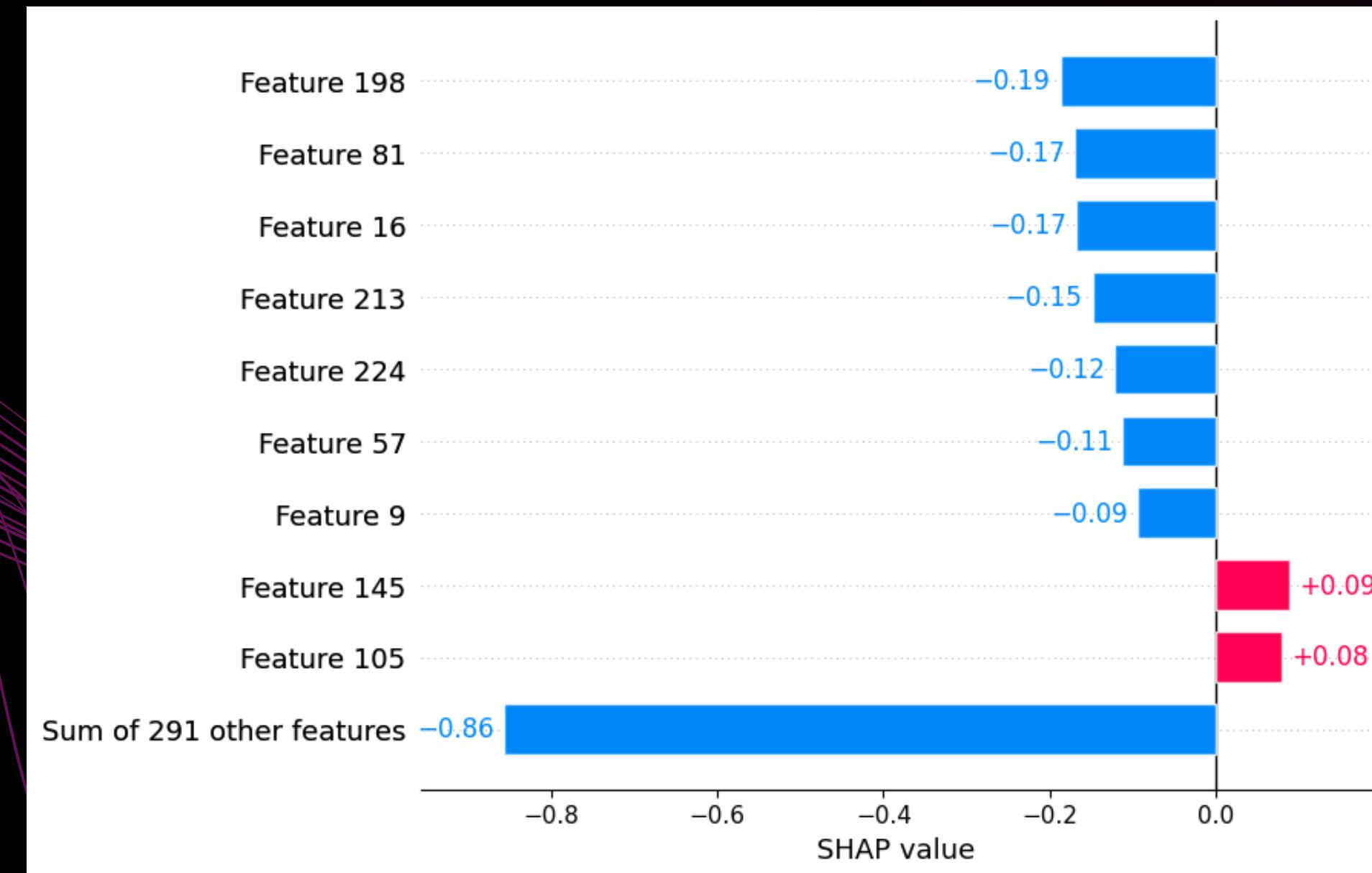
Classification Report:

	precision	recall	f1-score	support
0	0.98	0.99	0.99	619
1	0.99	0.98	0.98	581
accuracy			0.98	1200
macro avg	0.99	0.98	0.98	1200
weighted avg	0.99	0.98	0.98	1200

Confusion Matrix:

```
[[611  8]
 [10 571]]
```

Model 2 – XAI



Model 2 – XAI

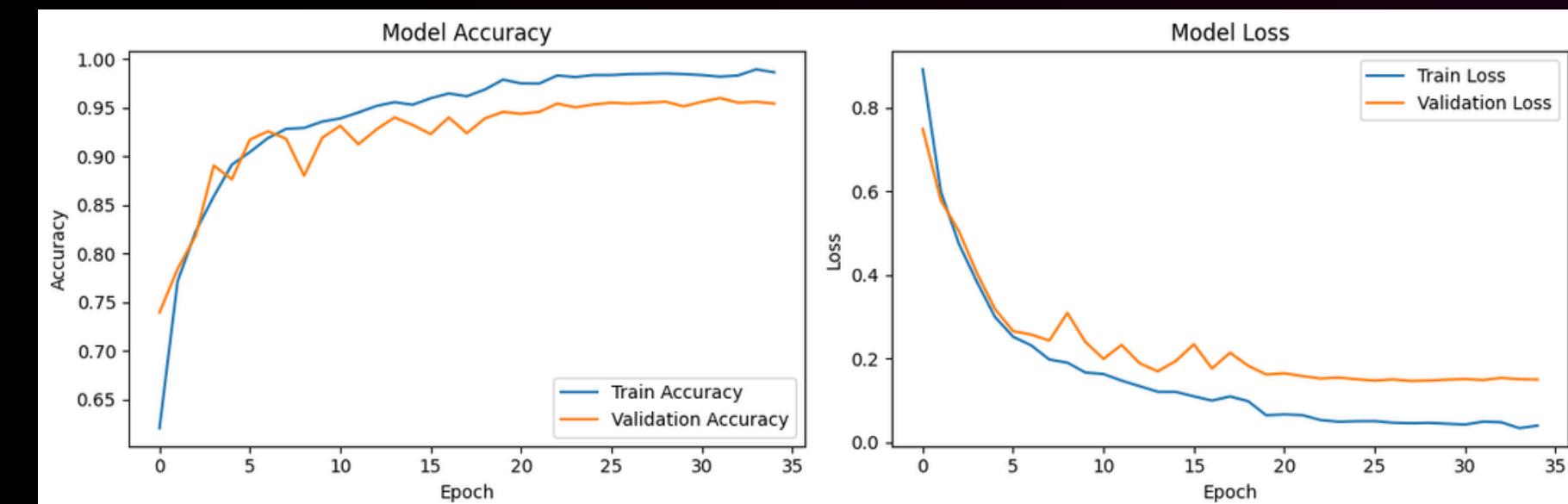


Toka Mokhtar

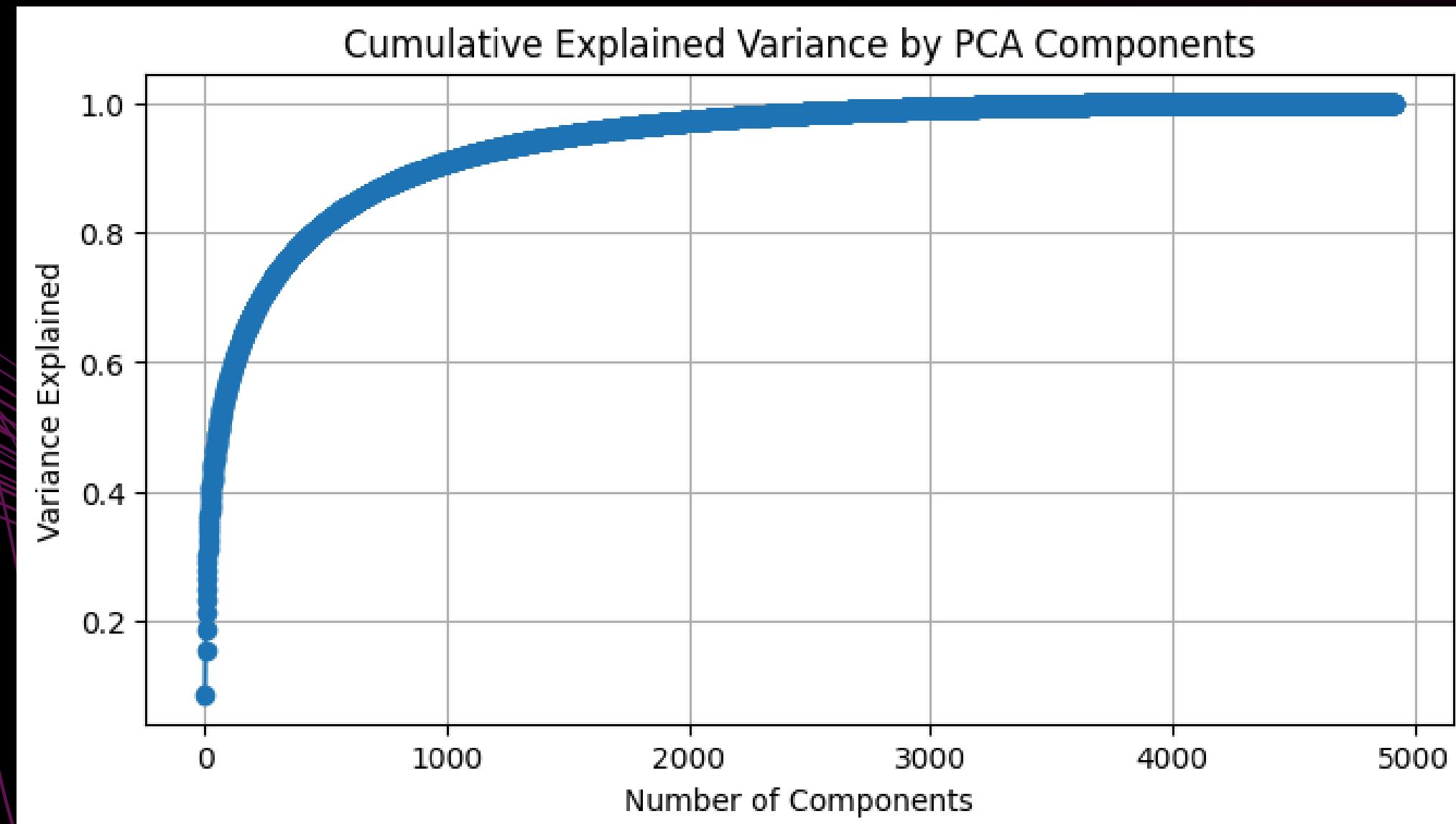
Model 3 – Brain Tumour CNN + PCA → SVM

- CNN (35 epochs) → Test accuracy 94.3%
- RF-selected CNN features → PCA → SVM accuracy 97.2%
- $\text{PCA}_1/2/3$: class-specific influence
- AUC ≈ 1.00 for 3/4 classes; meningioma AUC 0.99

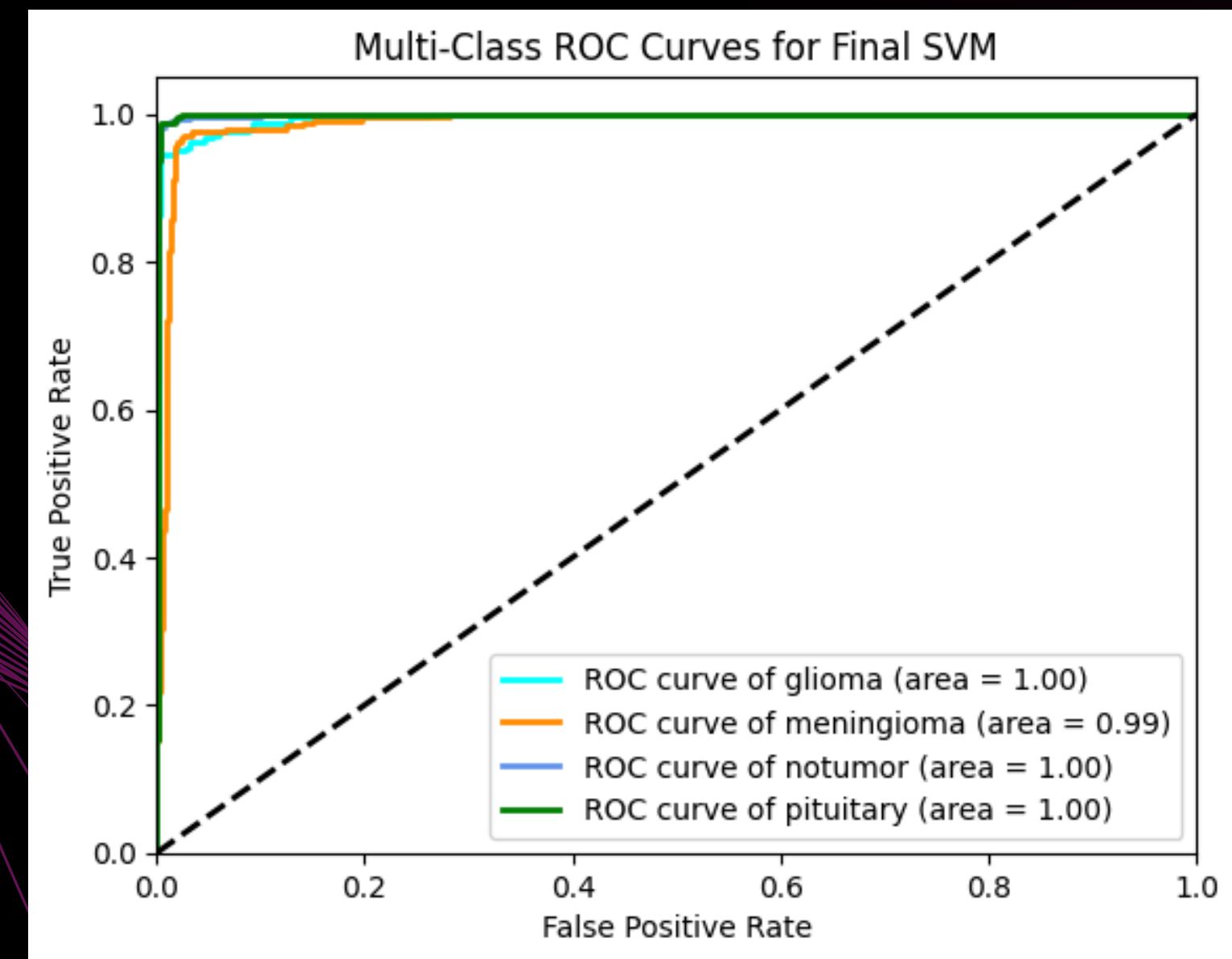
CNN Test Accuracy: 0.9430739879608154				
132/132	1s 7ms/step			
CNN Classification Report:				
	precision	recall	f1-score	support
glioma	0.93	0.93	0.93	243
meningioma	0.87	0.91	0.89	247
notumor	0.98	0.96	0.97	300
pituitary	0.98	0.96	0.97	264
accuracy			0.94	1054
macro avg	0.94	0.94	0.94	1054
weighted avg	0.94	0.94	0.94	1054



Model 3 – XAI

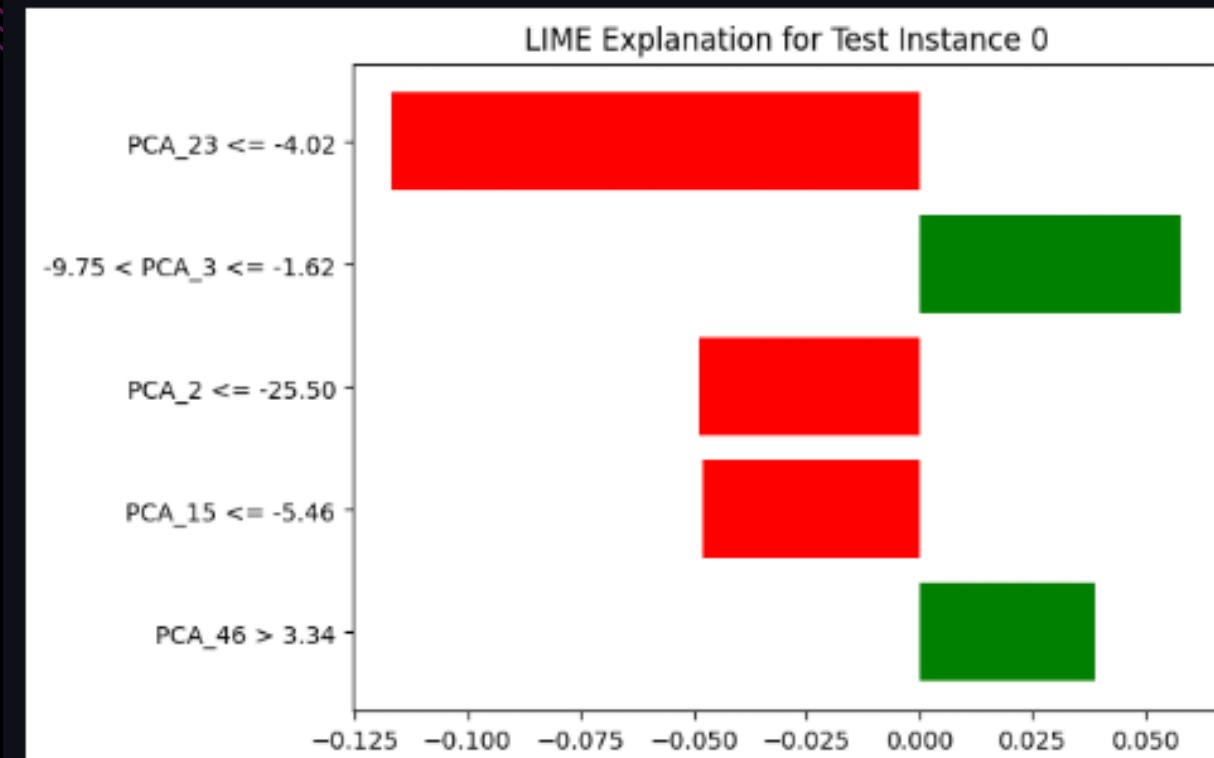


Model 3 – XAI

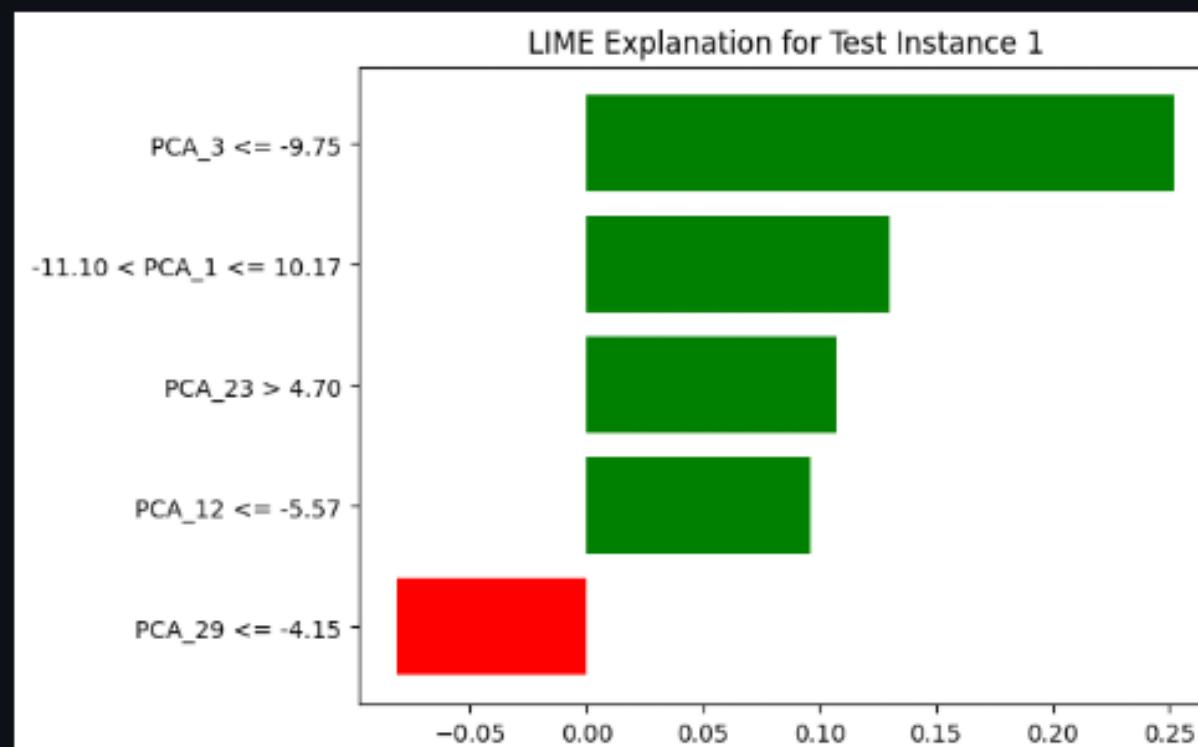


Model 3 – XAI

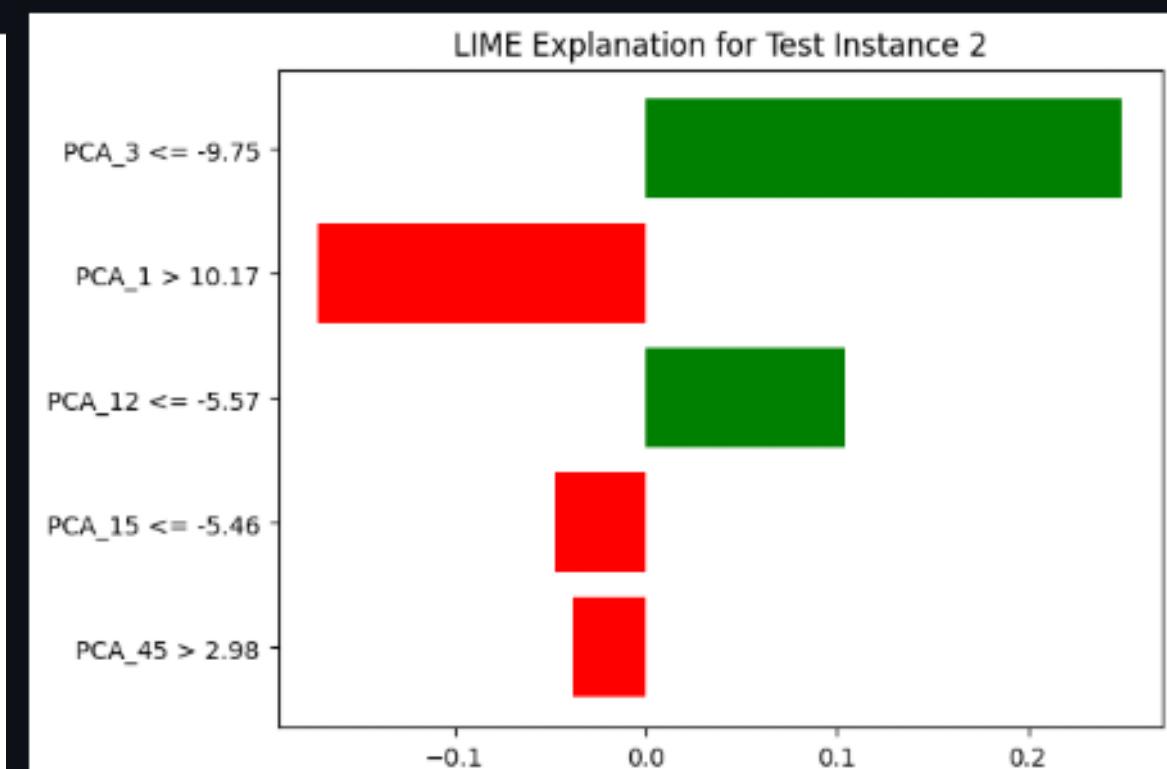
Explanation for test instance 0 (True label: glioma, Predicted: glioma)
PCA_23 ≤ -4.02 : -0.1167
-9.75 < PCA_3 ≤ -1.62 : 0.0580
PCA_2 ≤ -25.50 : -0.0485
PCA_15 ≤ -5.46 : -0.0481
PCA_46 > 3.34 : 0.0391



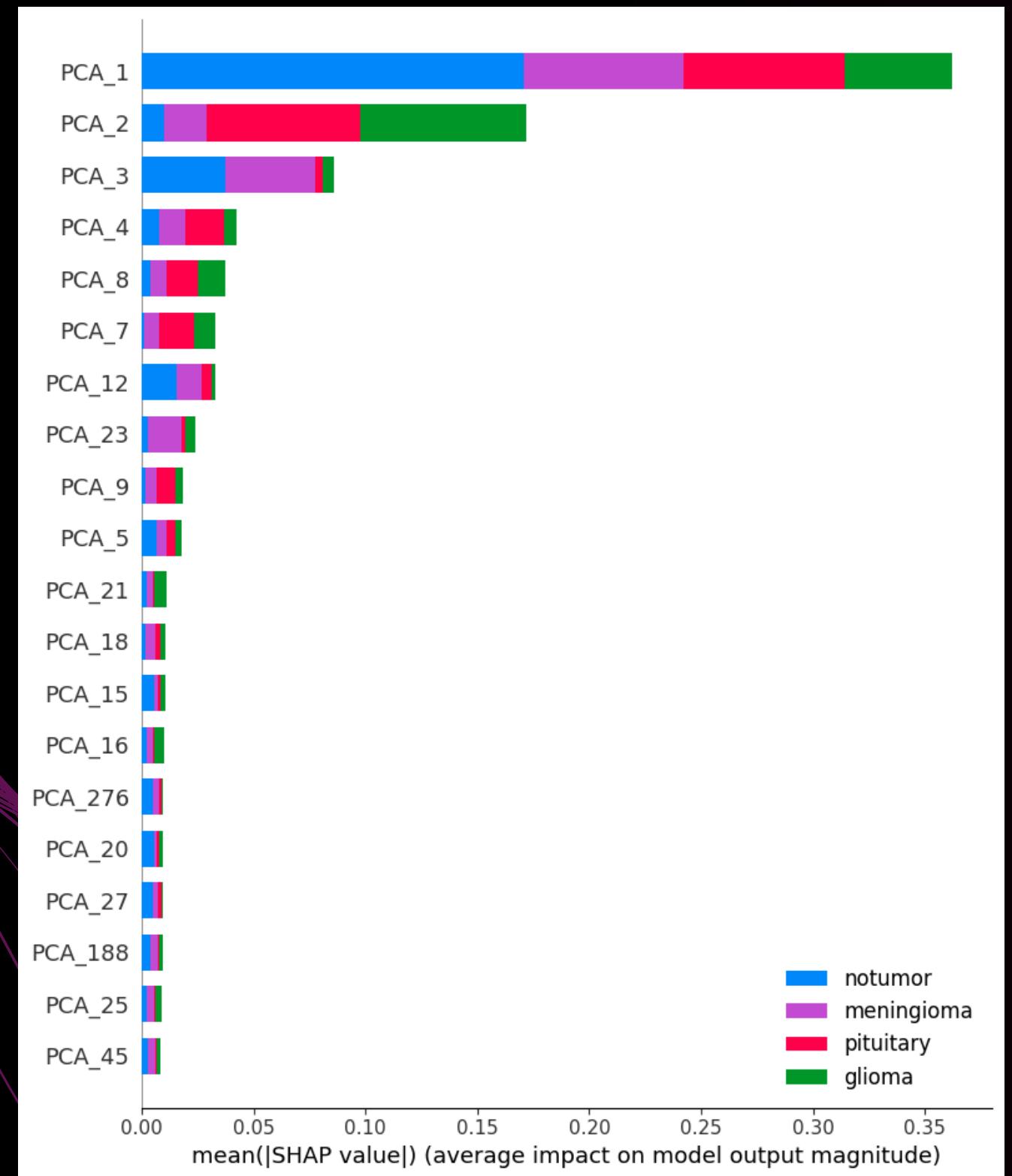
Explanation for test instance 1 (True label: meningioma, Predicted: meningioma)
PCA_3 ≤ -9.75 : 0.2522
-11.10 < PCA_1 ≤ 10.17 : 0.1298
PCA_23 > 4.70 : 0.1073
PCA_12 ≤ -5.57 : 0.0964
PCA_29 ≤ -4.15 : -0.0810



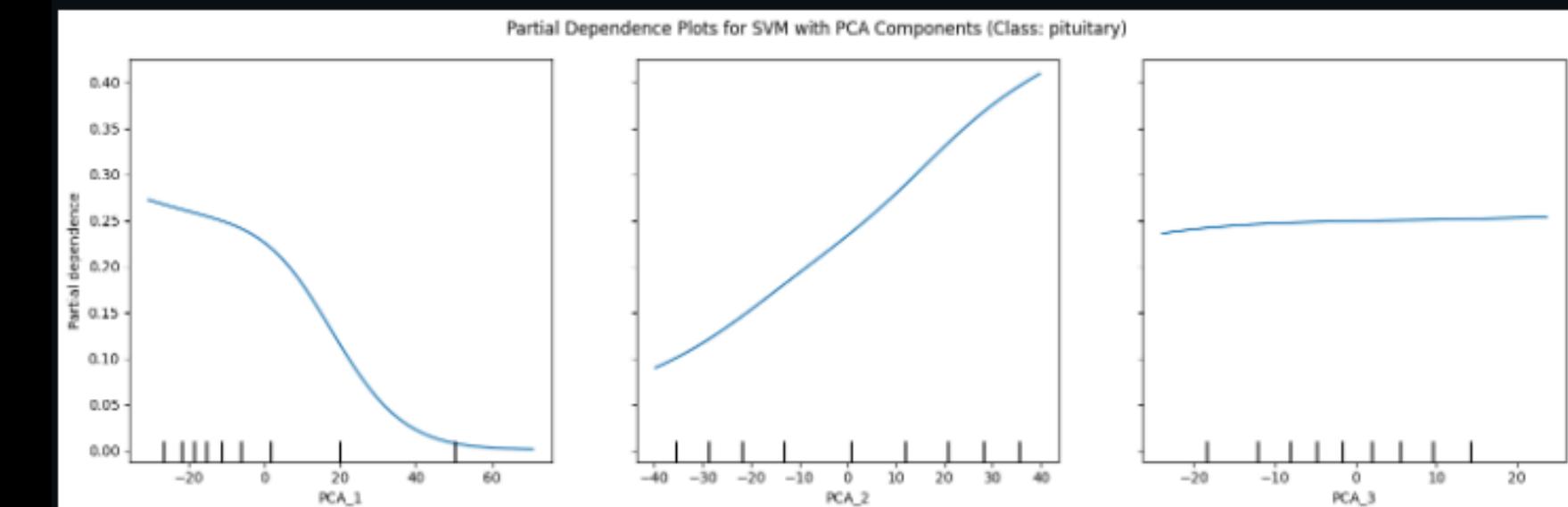
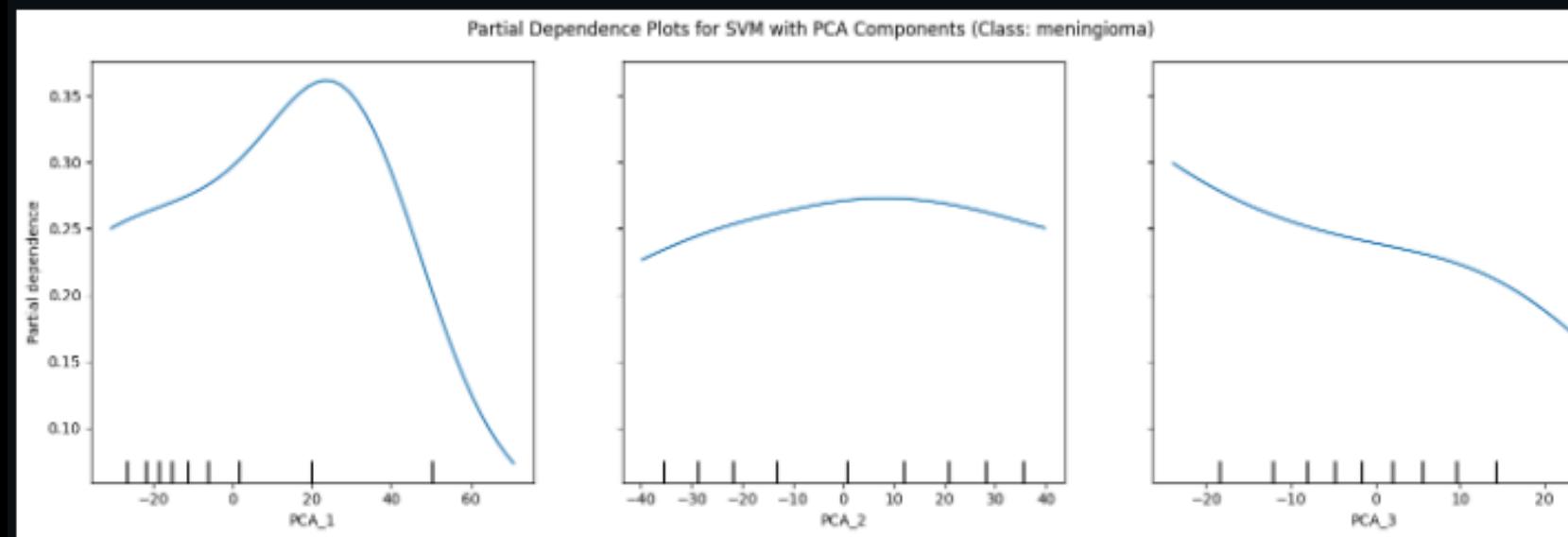
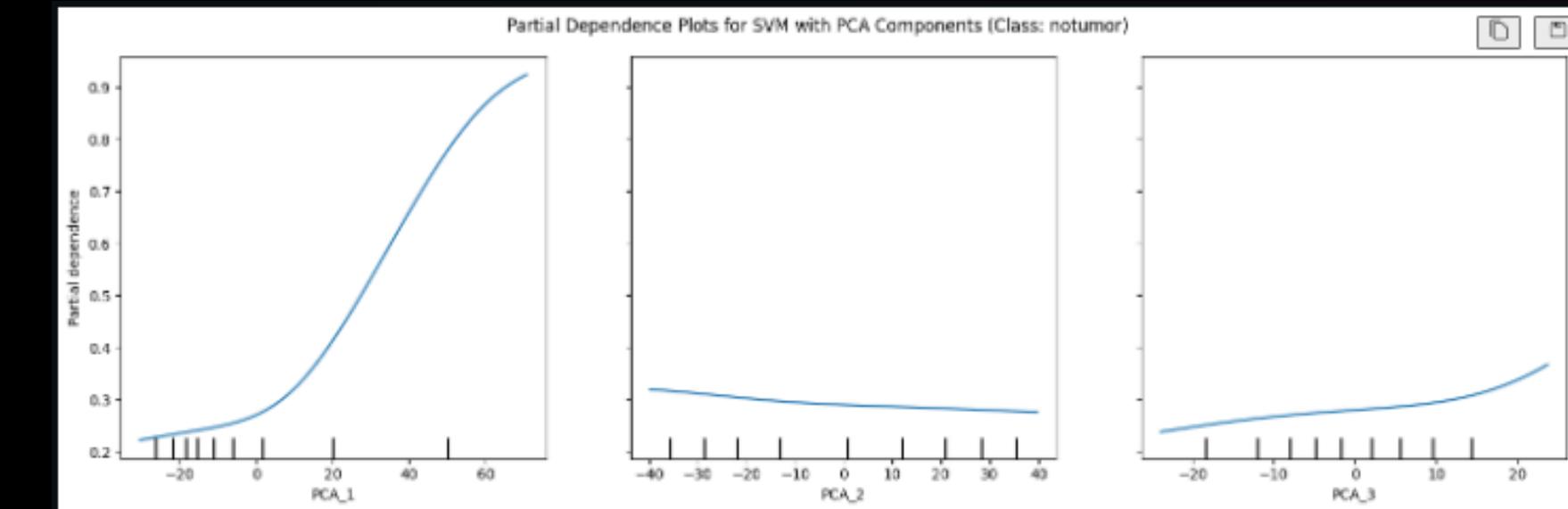
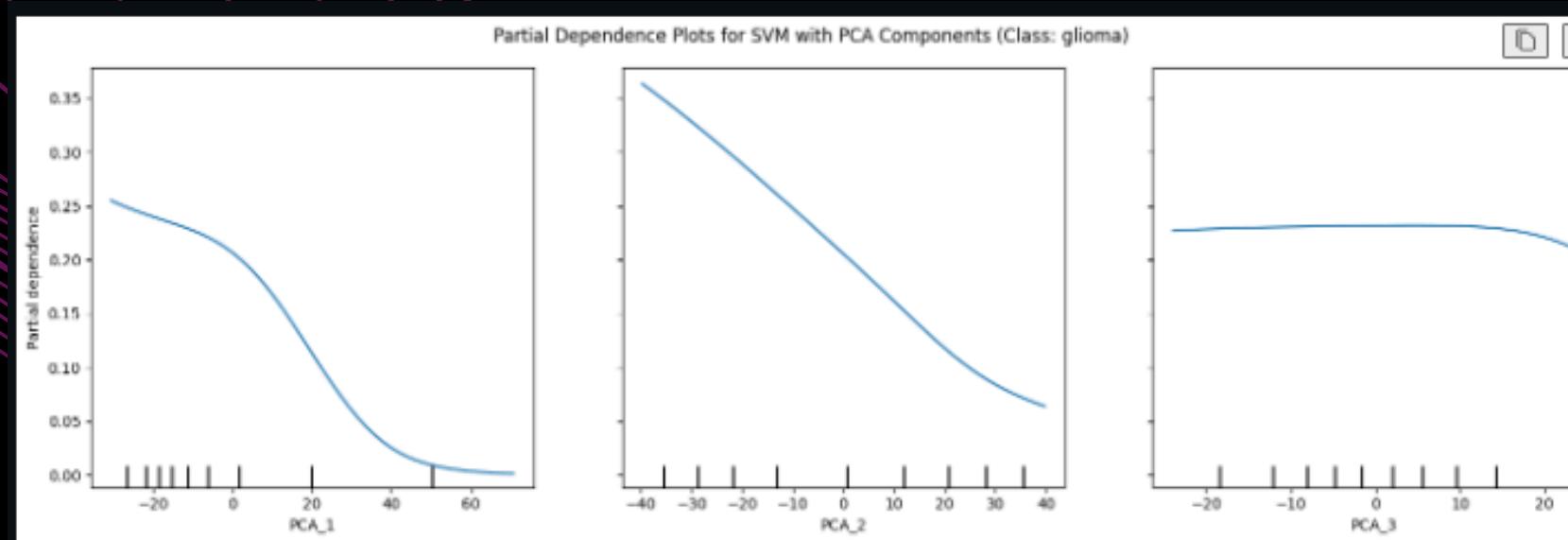
Explanation for test instance 2 (True label: notumor, Predicted: notumor)
PCA_3 ≤ -9.75 : 0.2489
PCA_1 > 10.17 : -0.1711
PCA_12 ≤ -5.57 : 0.1048
PCA_15 ≤ -5.46 : -0.0477
PCA_45 > 2.98 : -0.0383



Model 3 – XAI



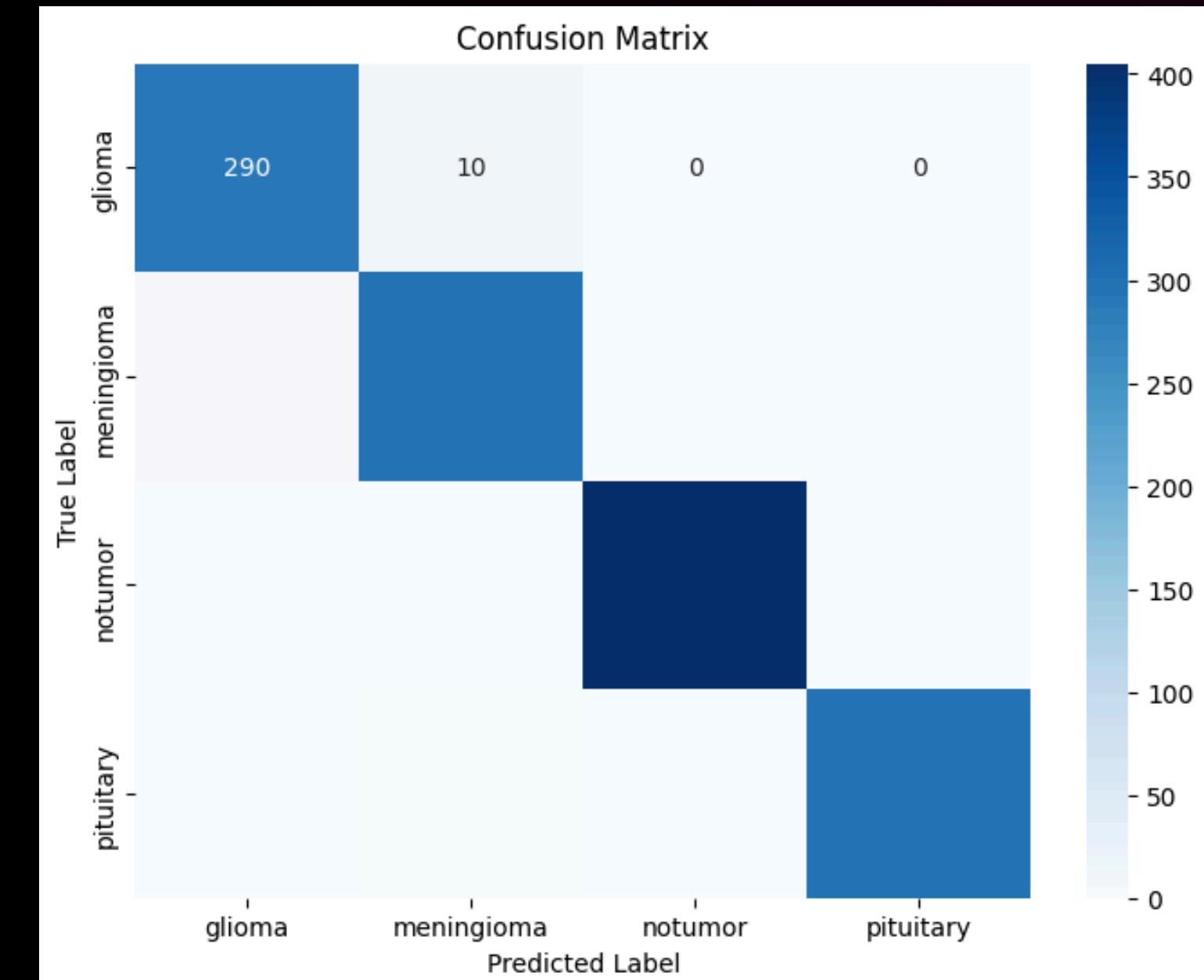
Model 3 – XAI



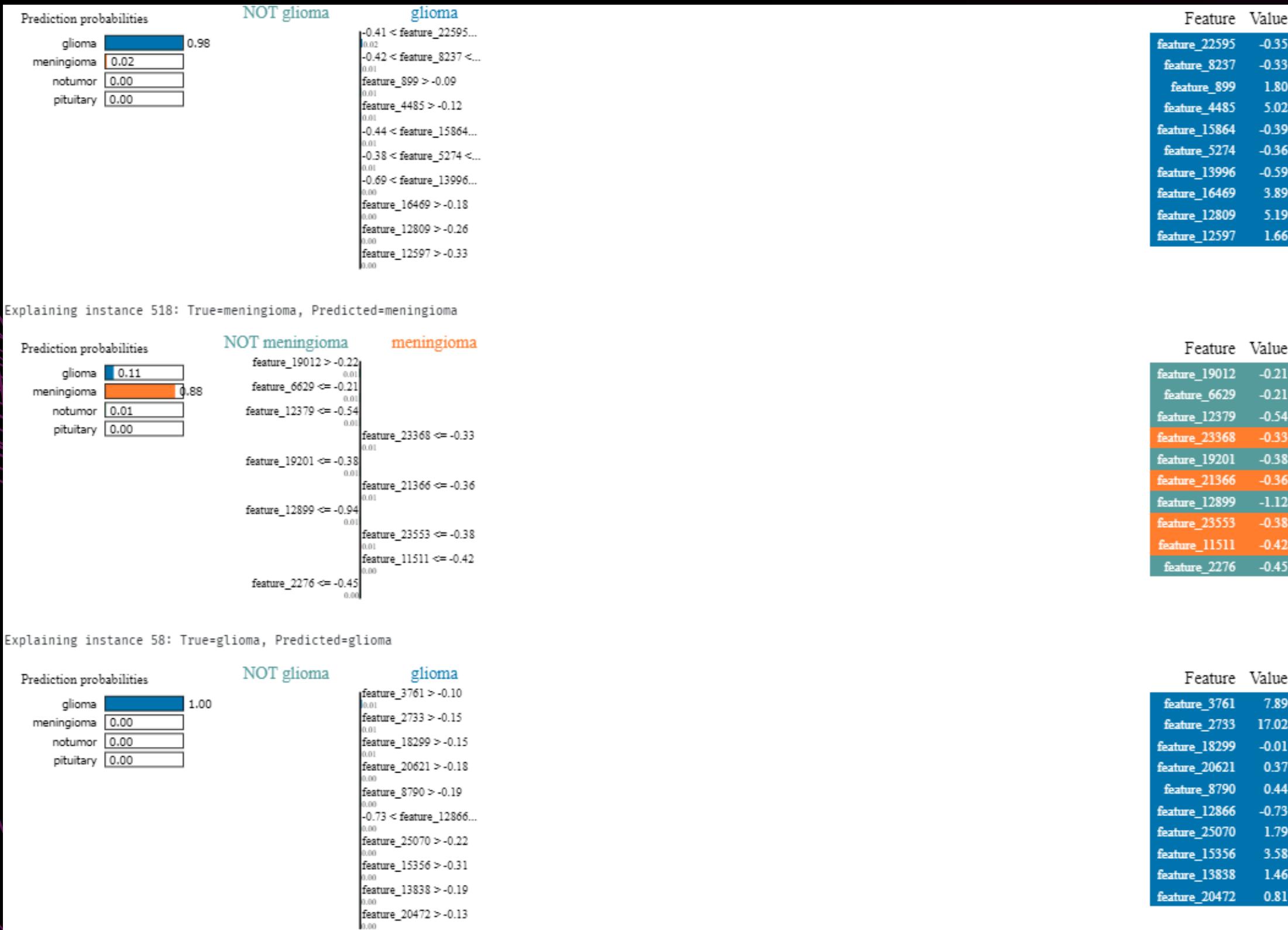
Hothifa Hamdan

Model 1 – VGG16 Features → SVM

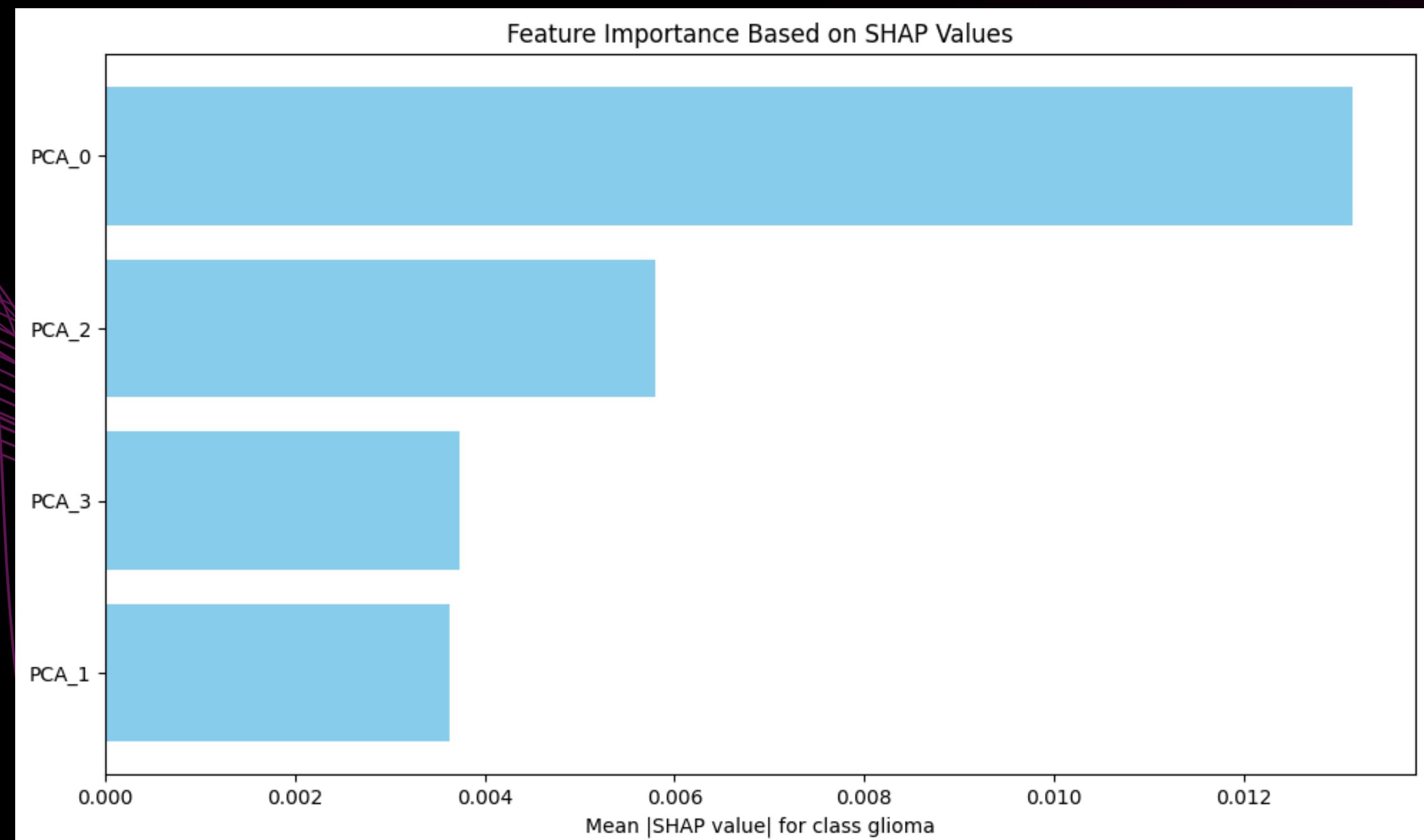
- Frozen VGG16 → Flatten(25 088)
- StandardScaler + Linear SVM (C=1.0)
- Support vectors: 2738
- Accuracy: 98.09 %
- Error: Only 21 (mostly glioma↔meningioma)



Model 1 – XAI



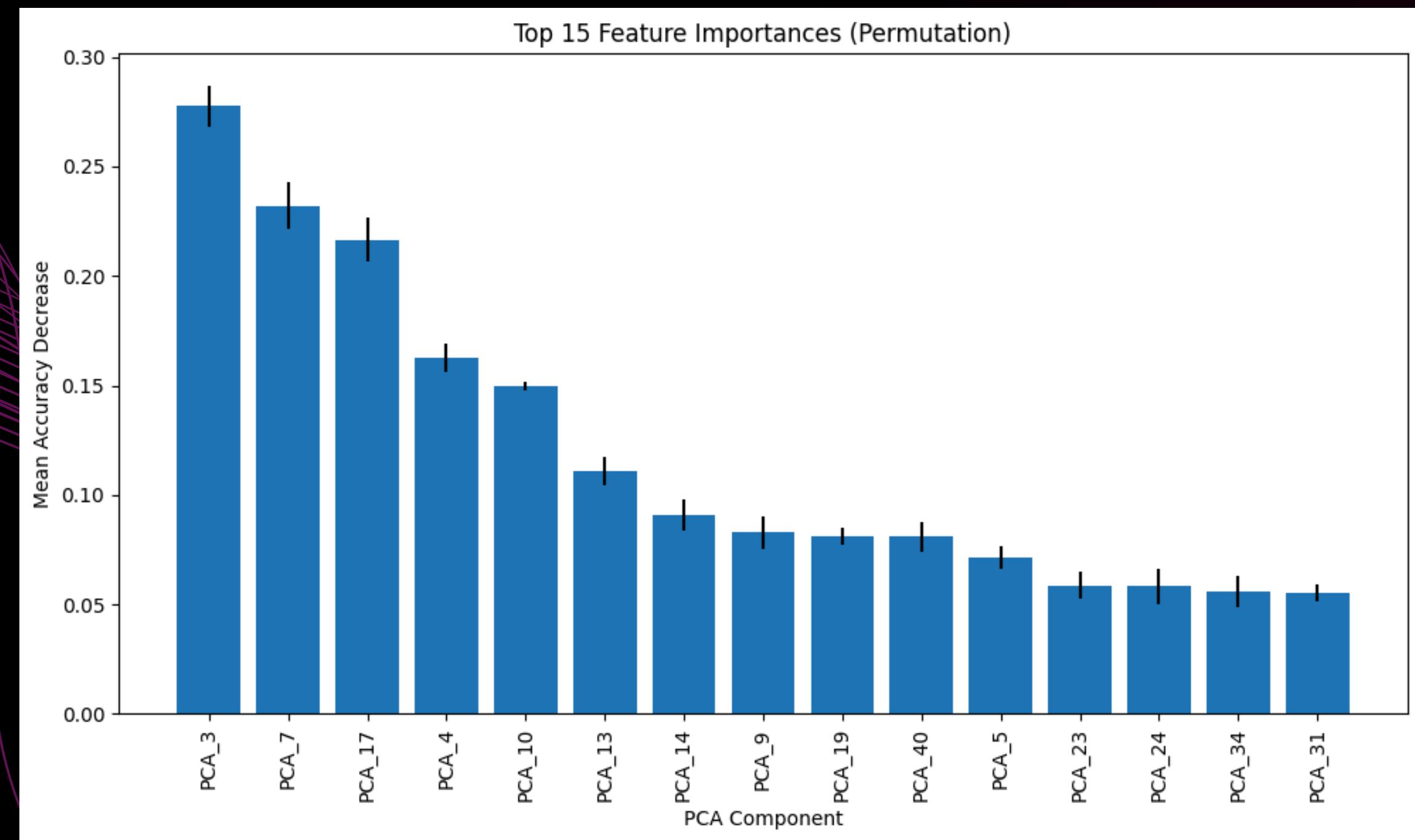
Model 1 – XAI



Model 1 – XAI

```
— XAI: SVM Intrinsic Analysis —  
SVM Coefficients shape: (6, 25088)  
Linear SVM (ovo): Showing top 10 feature coefficients for Class 0 vs 1  
Top 10 feature indices: [11322 15124 6565 20697 12326 16058 15739 4323 16359 10143]  
Corresponding coefficients: [ 0.01747708  0.01421918  0.01408765  0.01342165  0.01324259 -0.0124218  
 0.01231754 -0.0120121   0.01171734  0.0117003 ]  
  
Number of support vectors for each class: [657 939 448 694]  
Total support vectors: 2738  
Shape of support vectors: (2738, 25088)
```

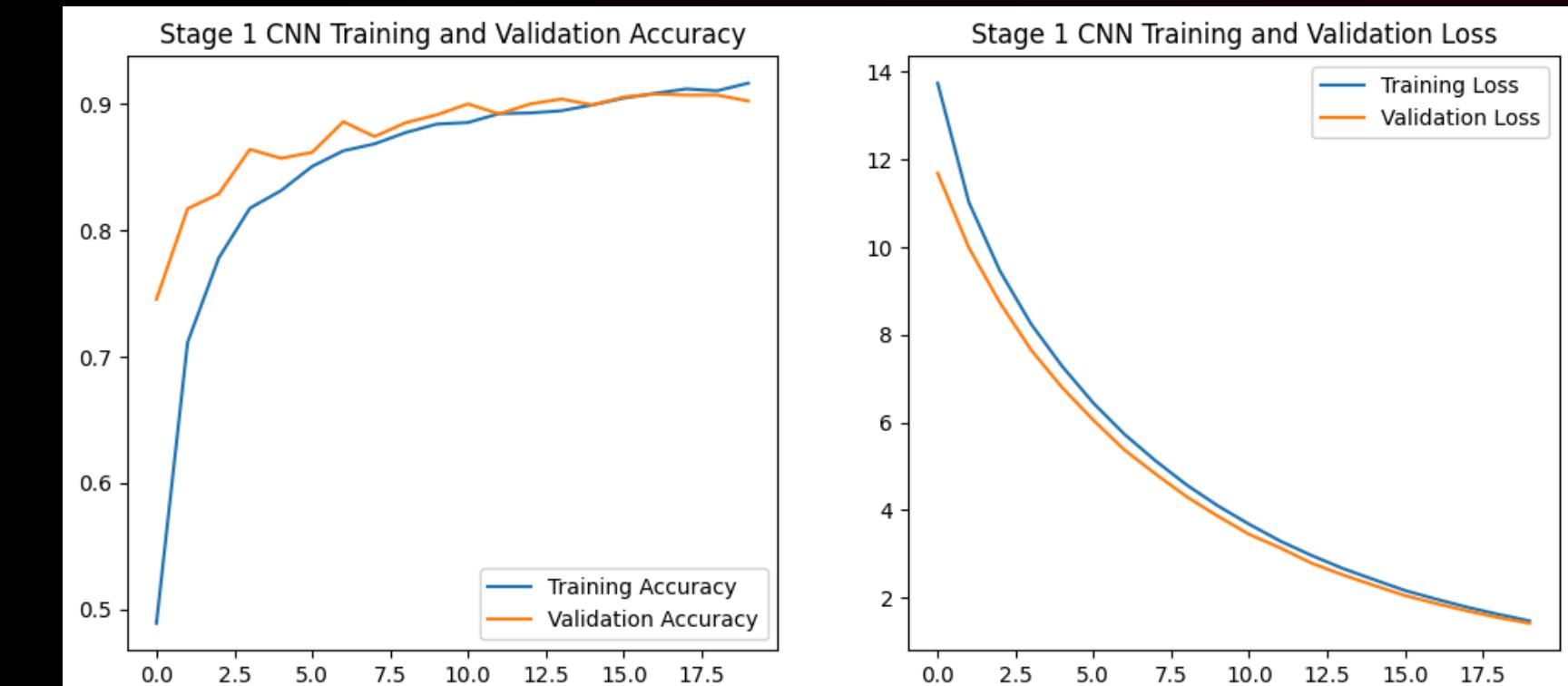
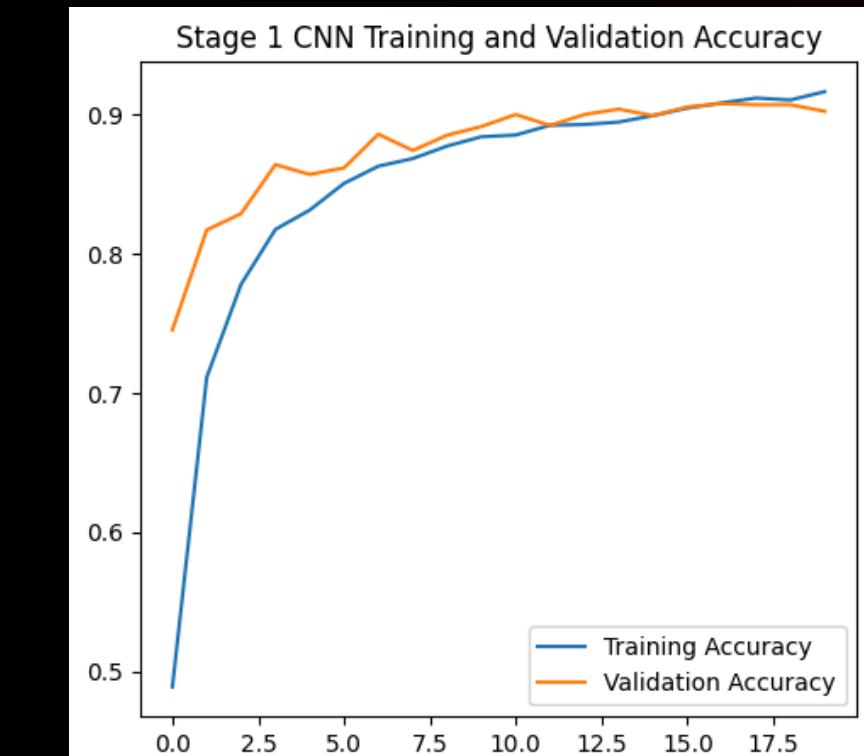
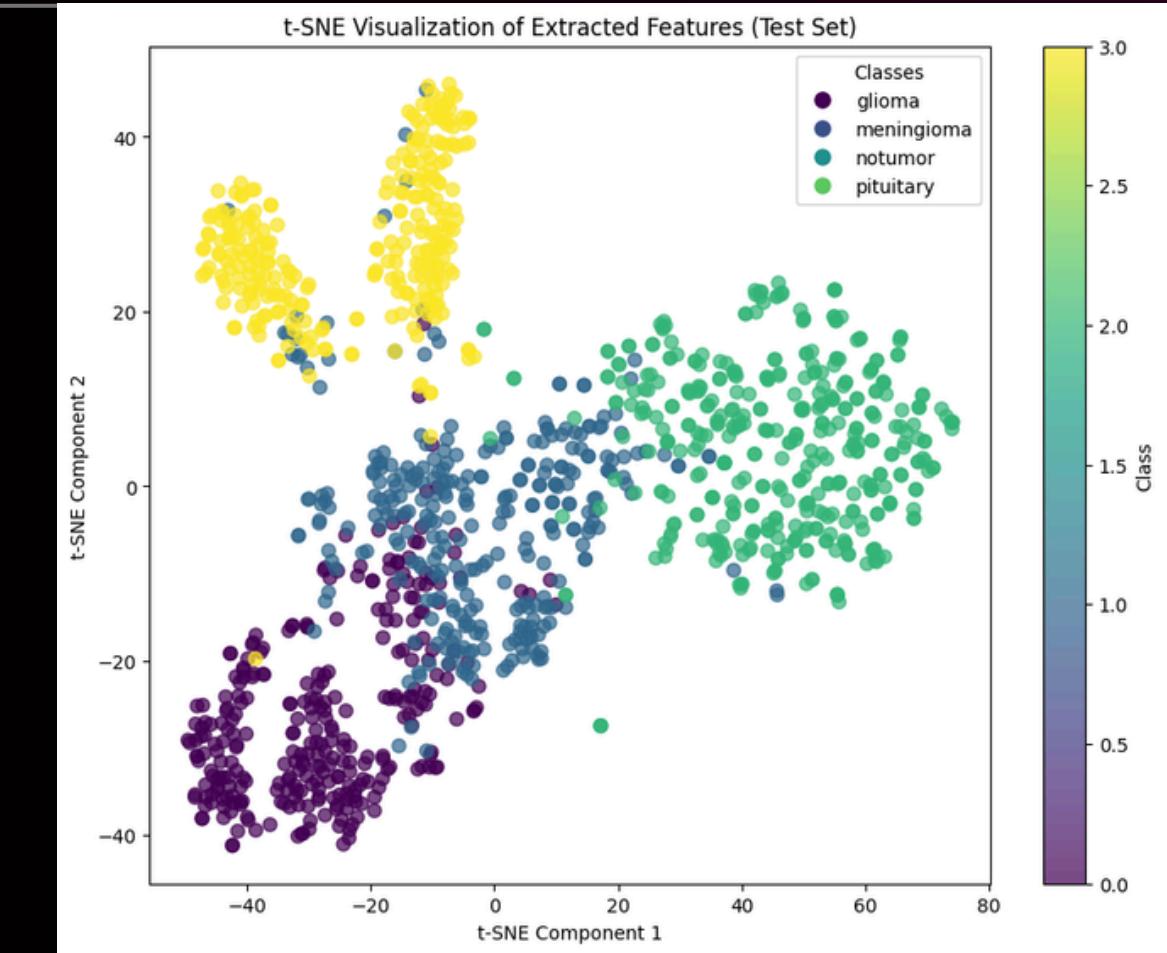
Model 1 - XAI

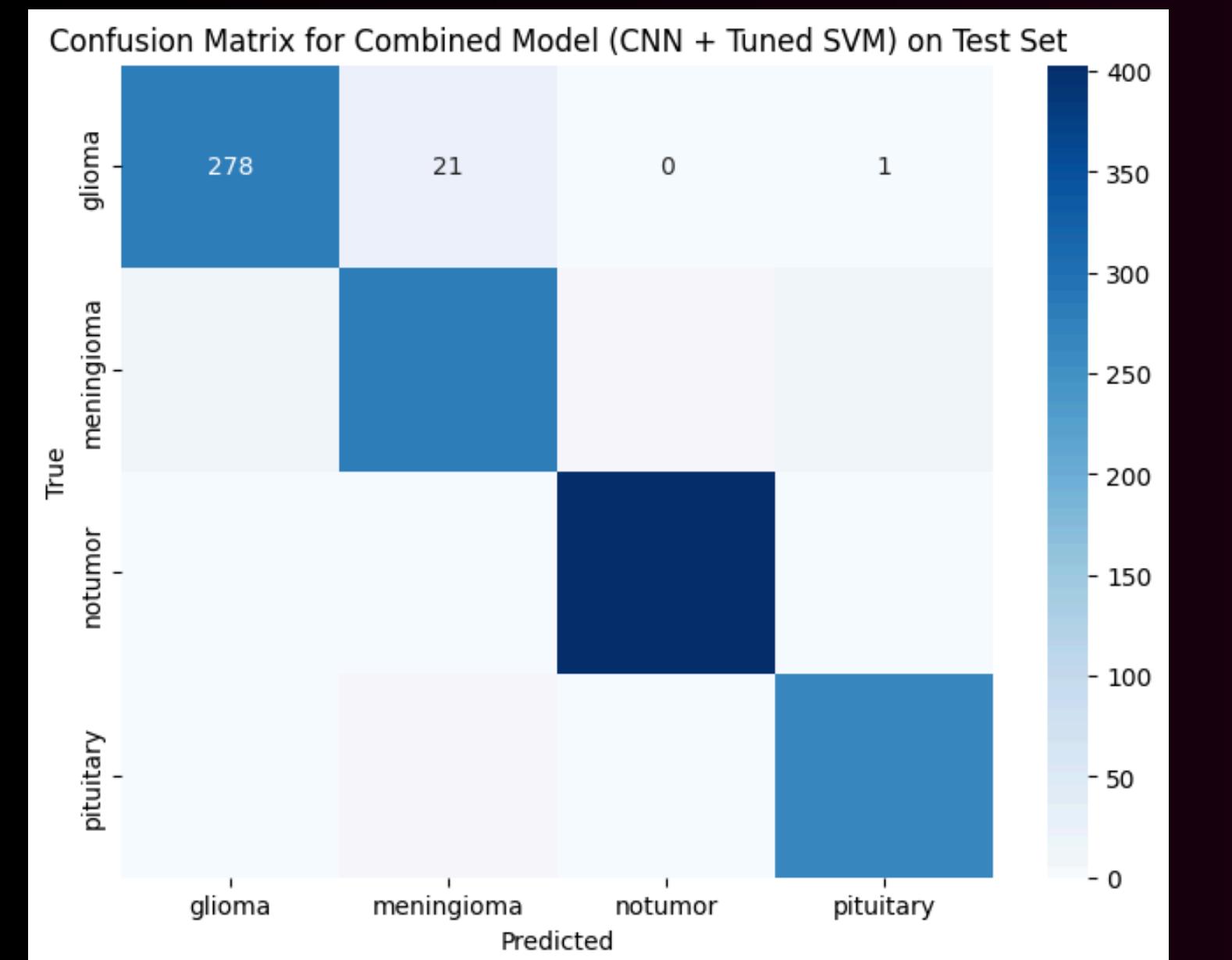
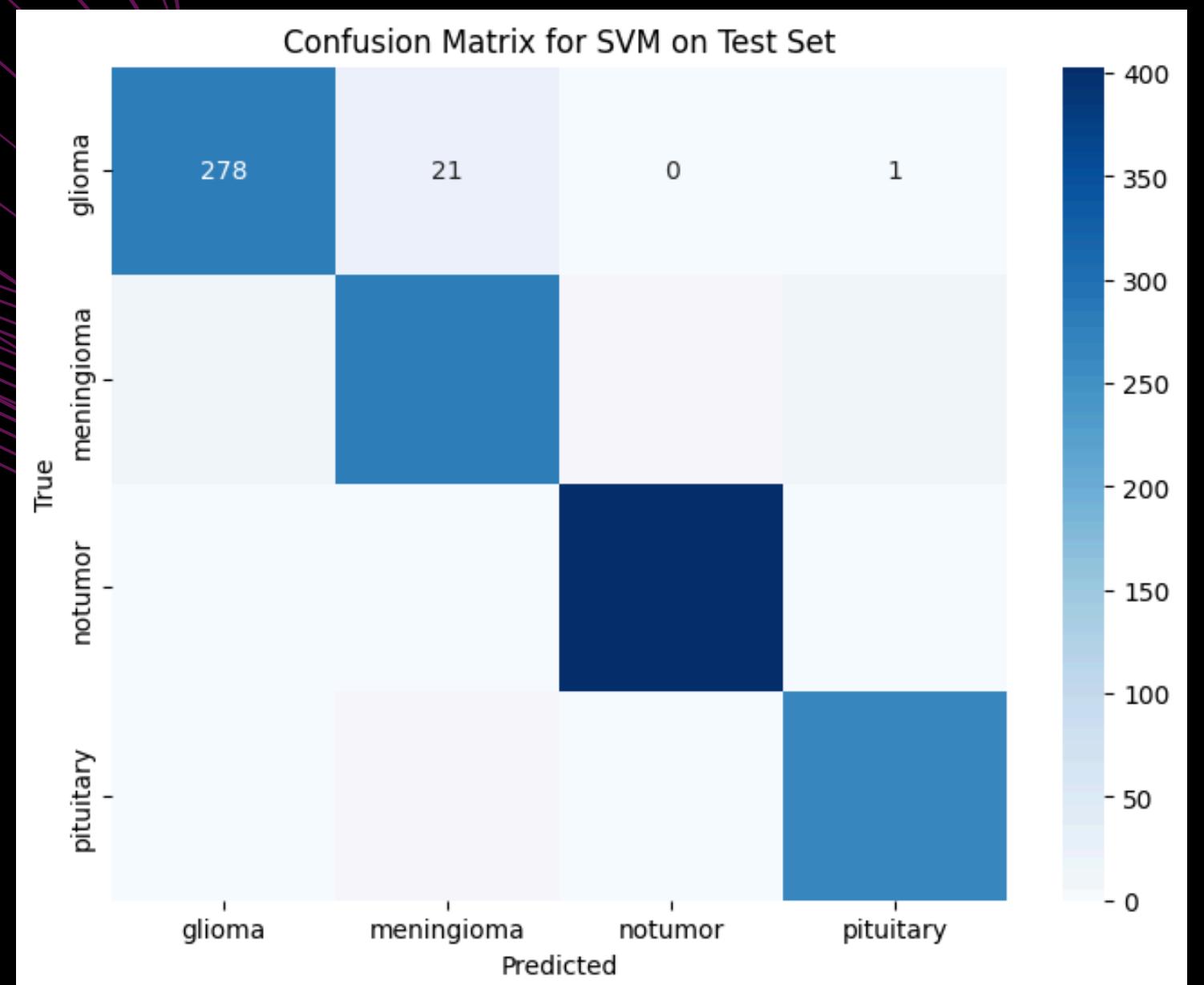


Hothifa Hamdan

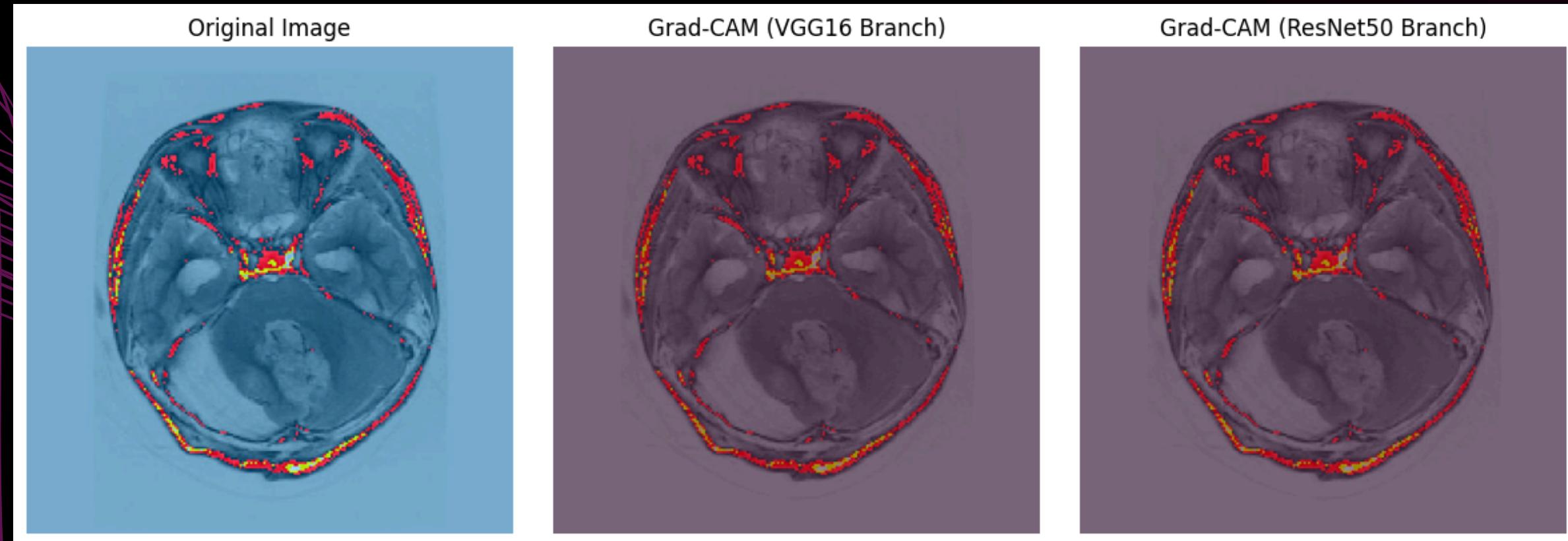
Model 2 – VGG16+ResNet50 CNN → SVM

- Parallel CNN (vGG+ResNet), frozen backbone
- CNN Val acc: 90.77% (best at Epoch 17)
- Extracted 512-D features → t-SNE shows clear class clusters
- Final SVM: GridSearch → C=1, γ=auto
- Test Accuracy: 95.78%

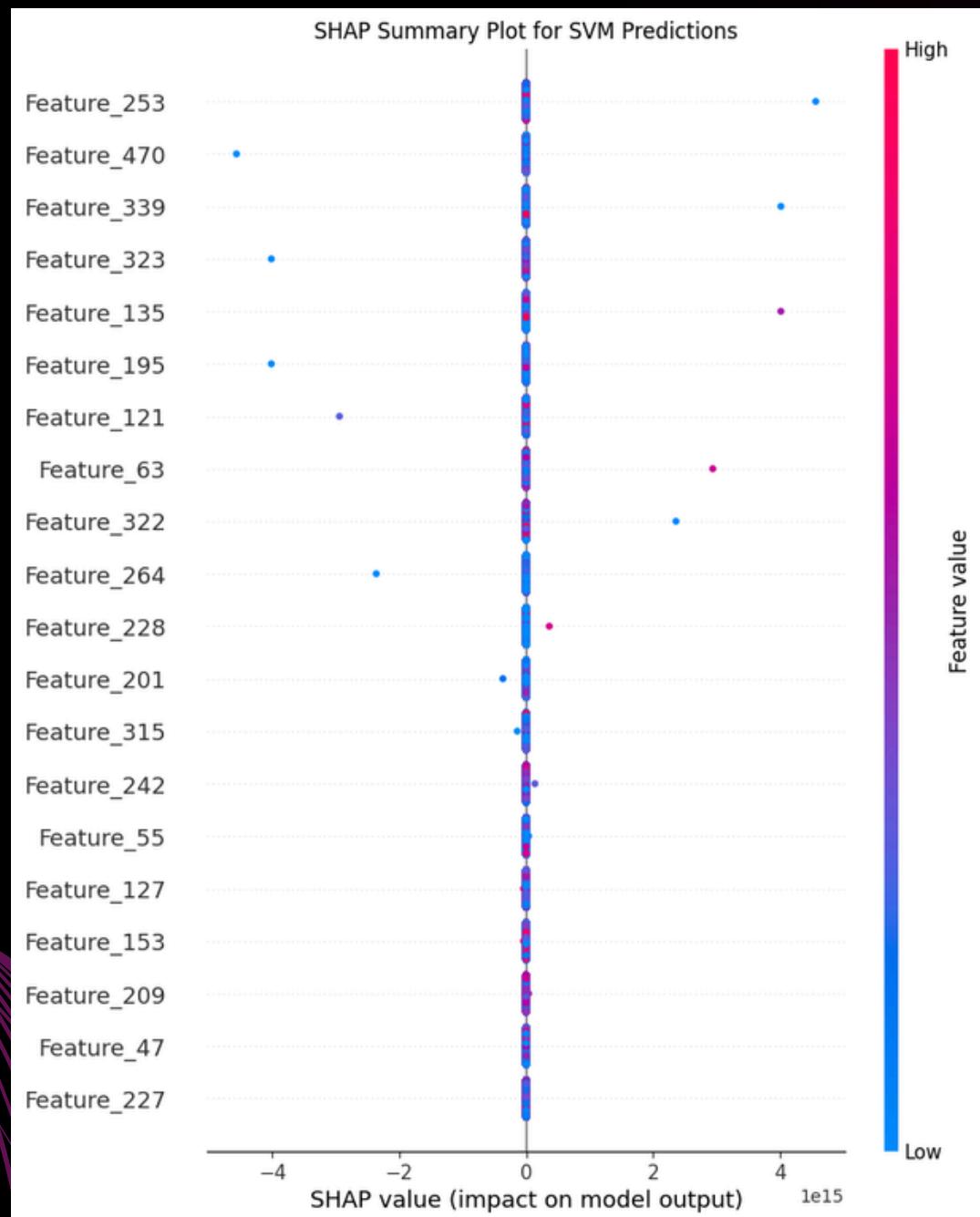




Model 2 – XAI



Model 2 – XAI



Hothifa Hamdan

Model 3 – ResNet50 + Softmax

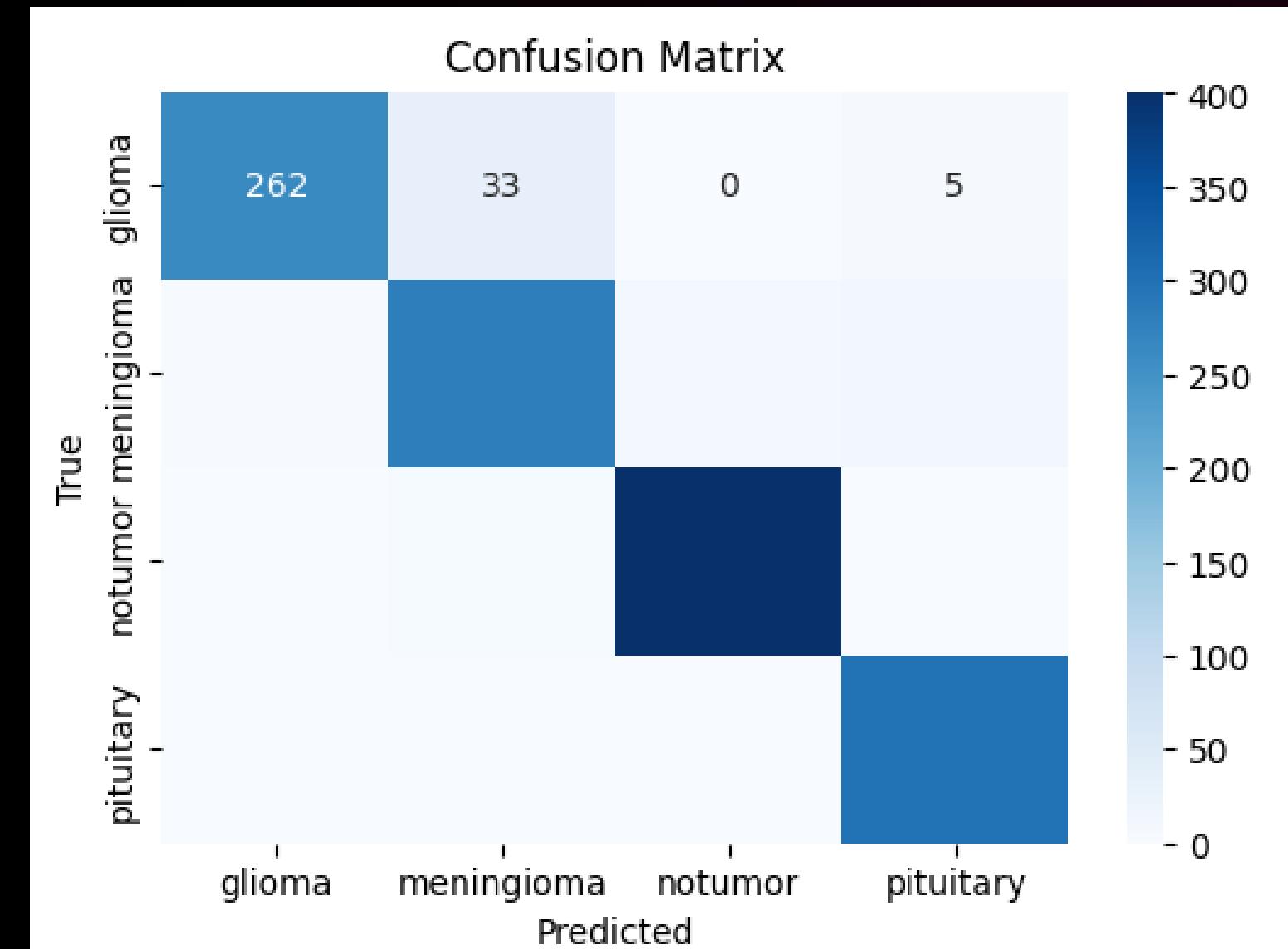
Backbone: ResNet50 (ImageNet,
frozen 10 ep)

Trainable: 8196 params

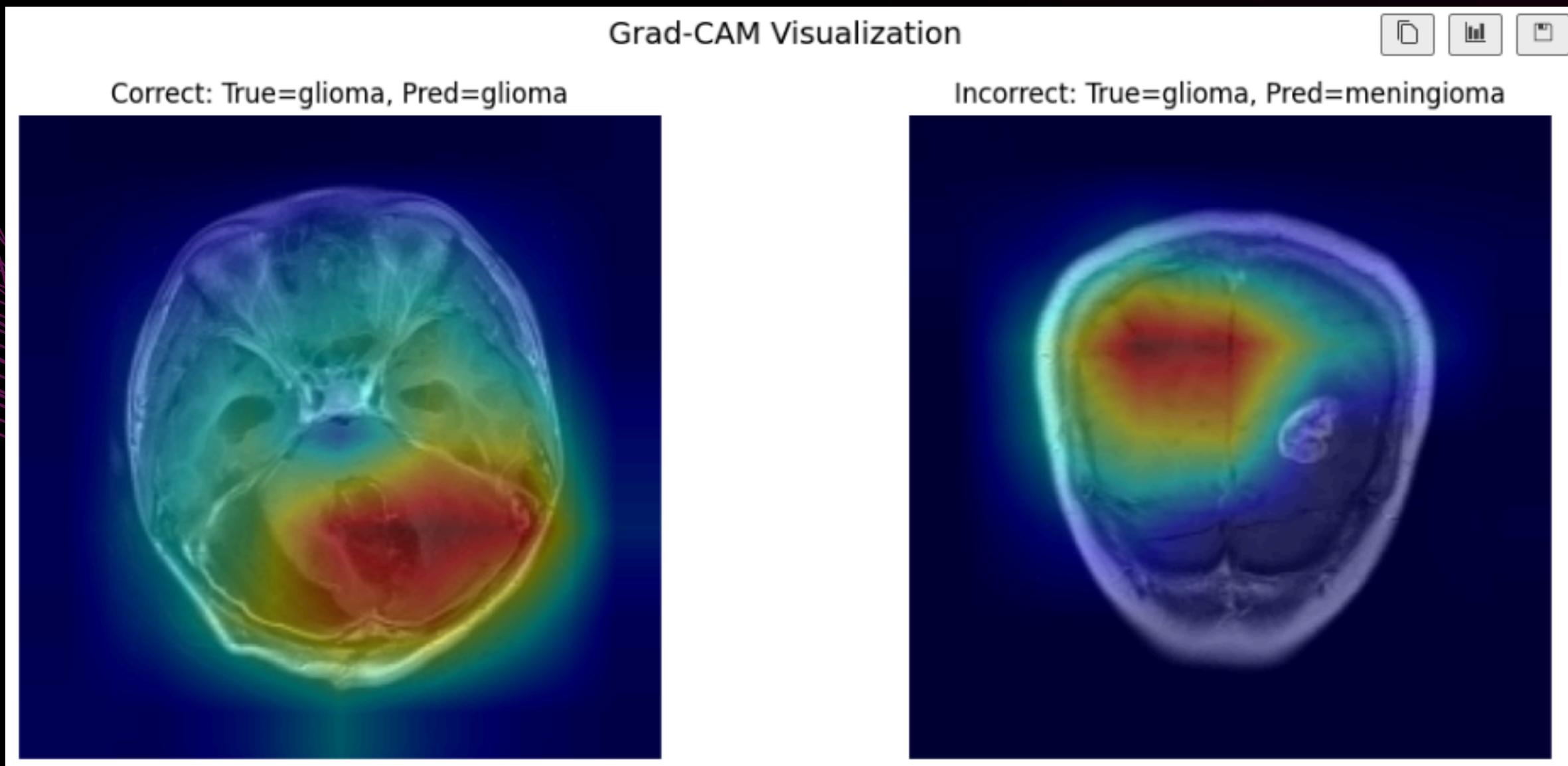
Top: GAP → Dropout → Dense 4
(softmax)

Val acc: ↑ 70% → 95% in 15 epochs

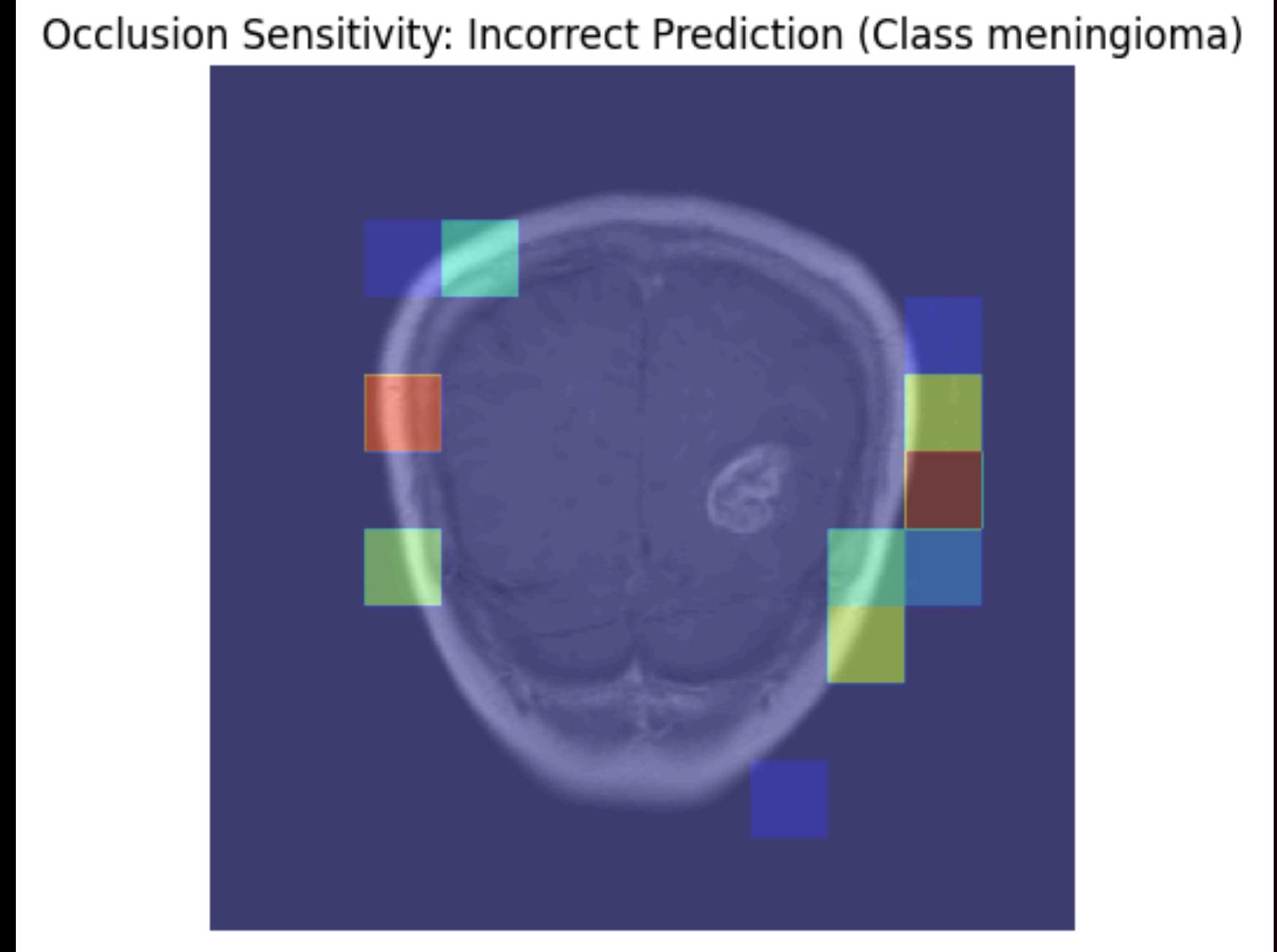
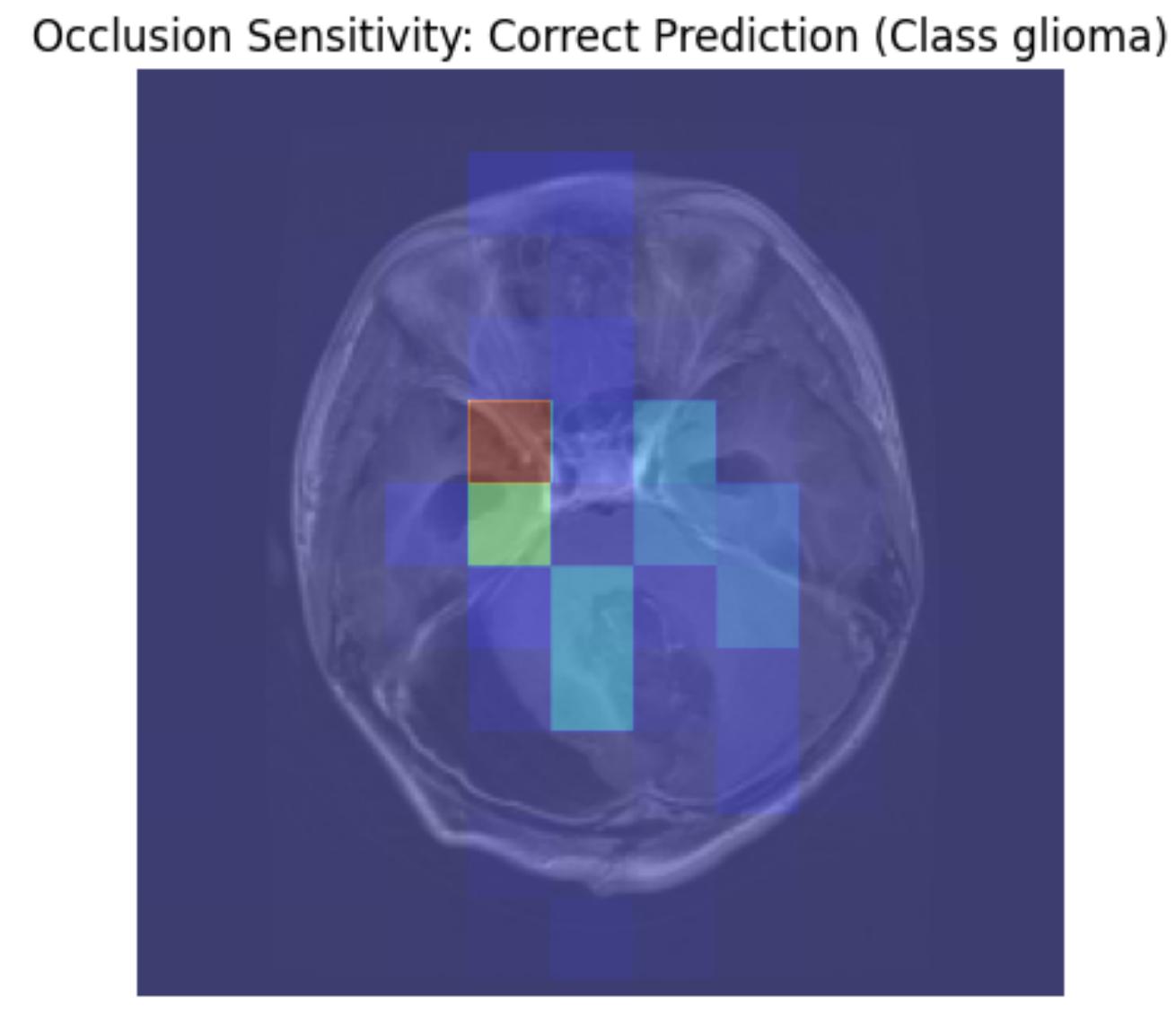
Loss ↓ 1.51 → 0.10



Model 3 – XAI

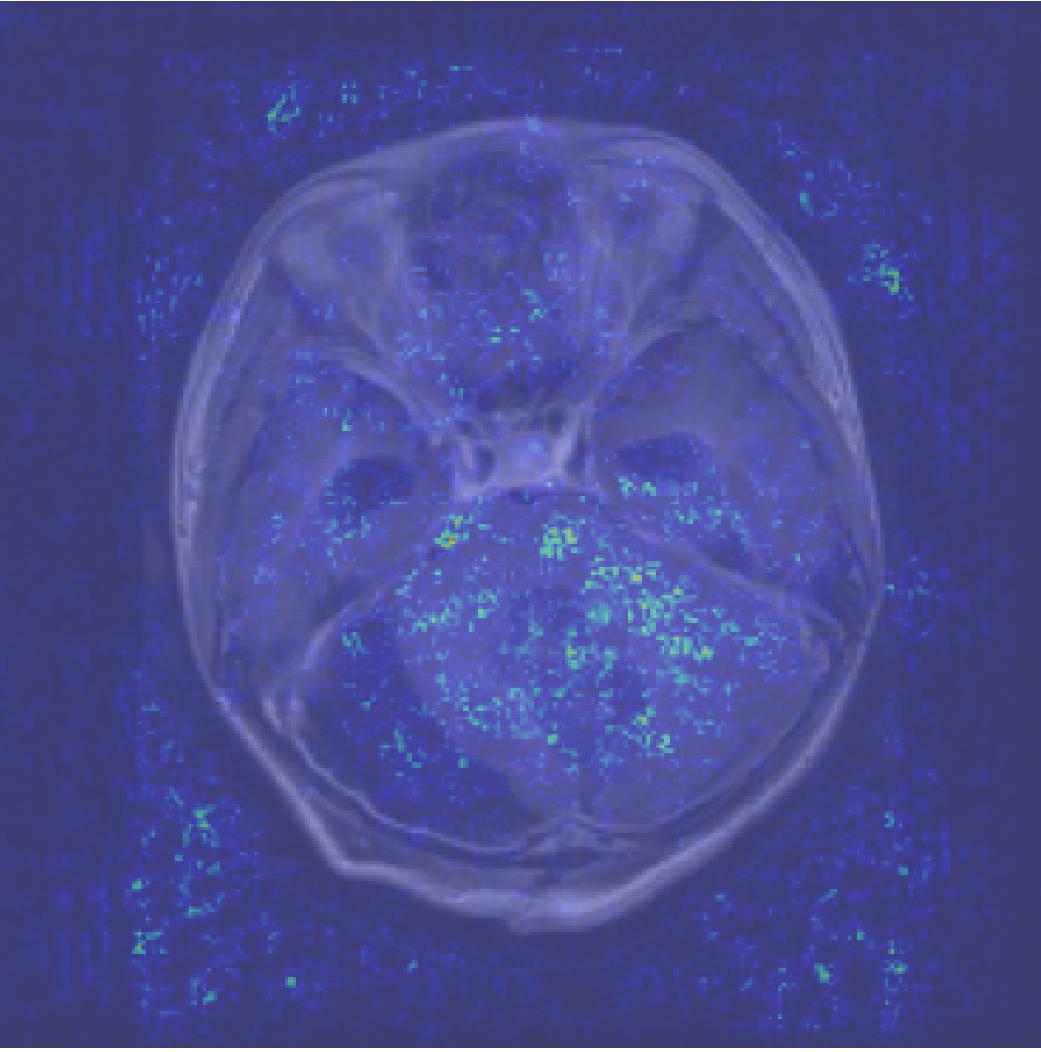


Model 3 – XAI

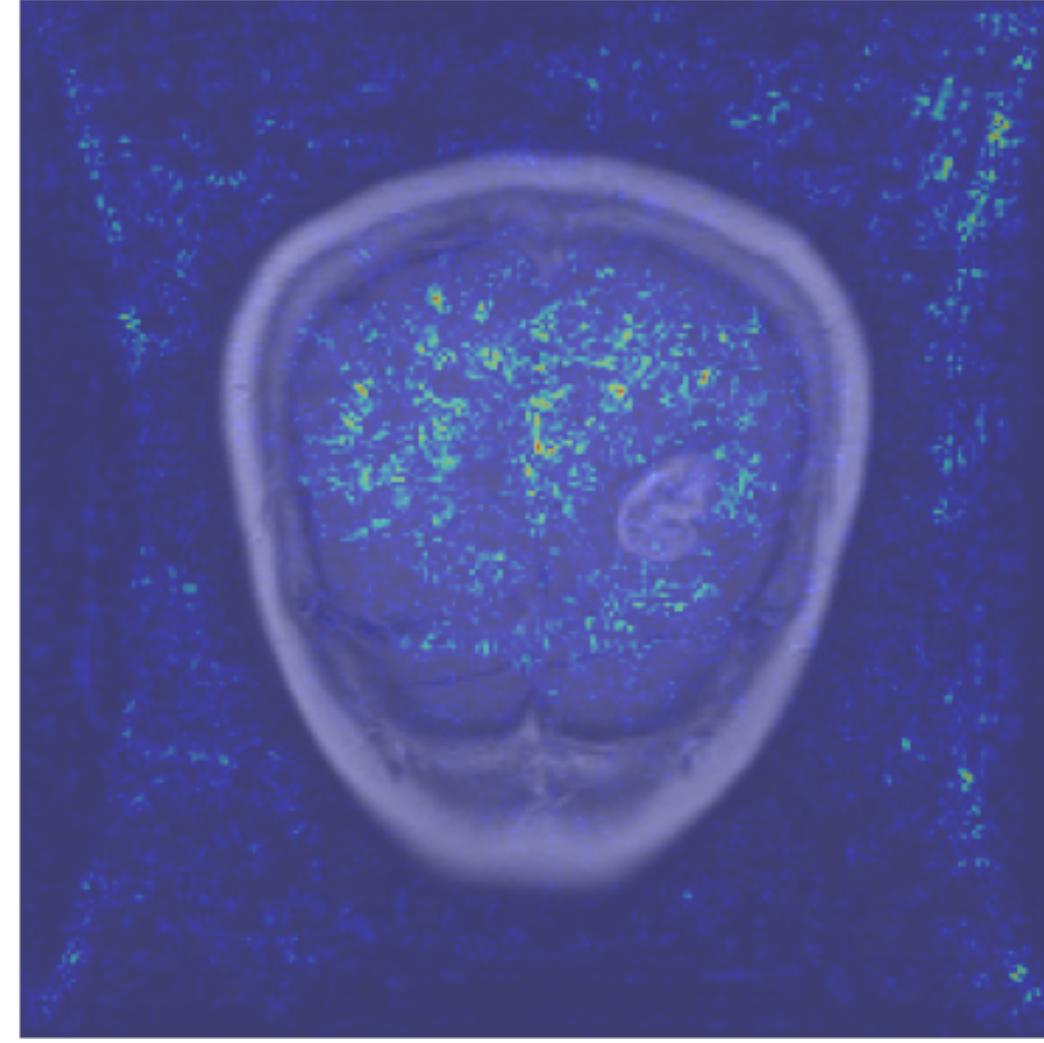


Model 3 – XAI

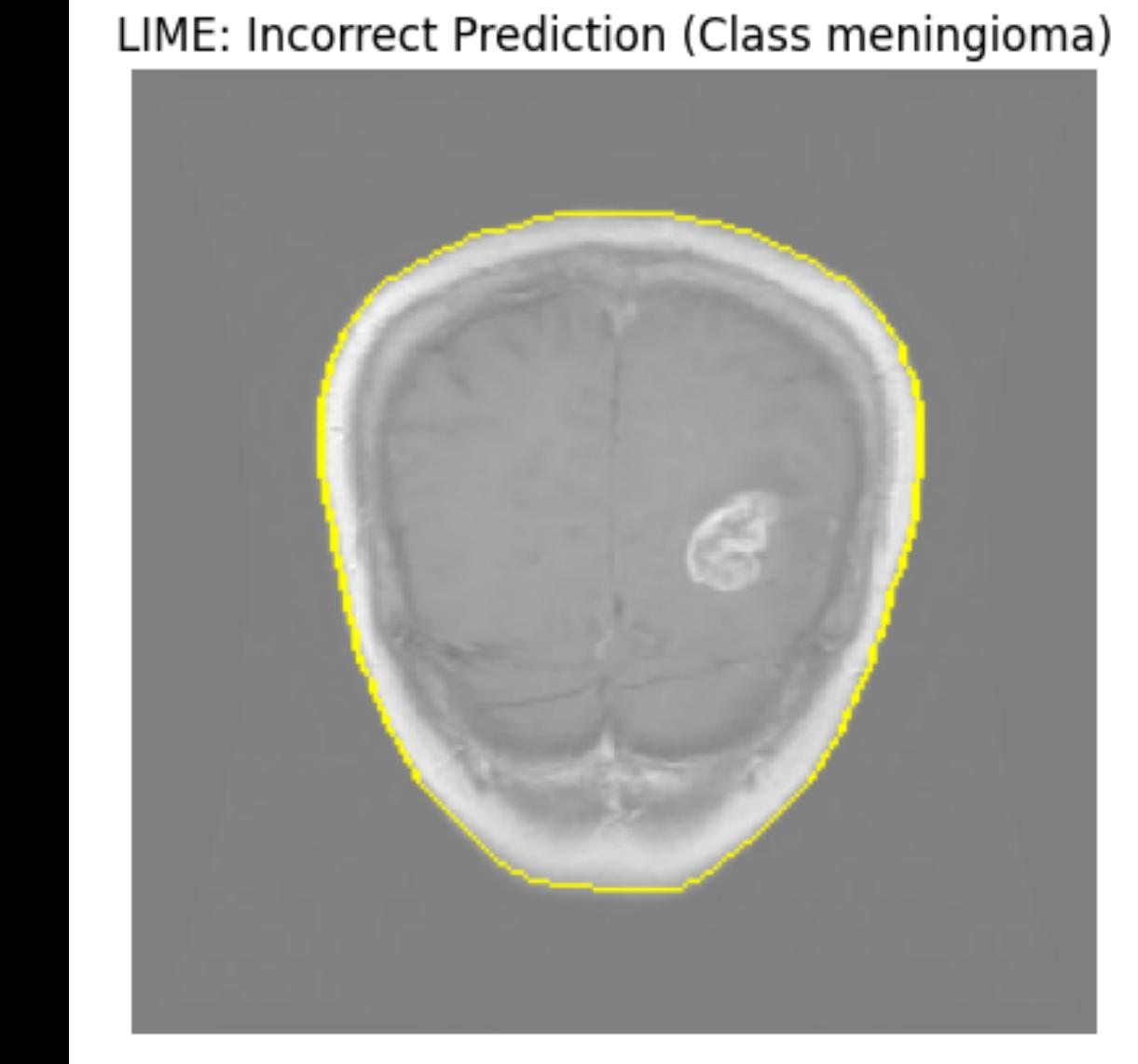
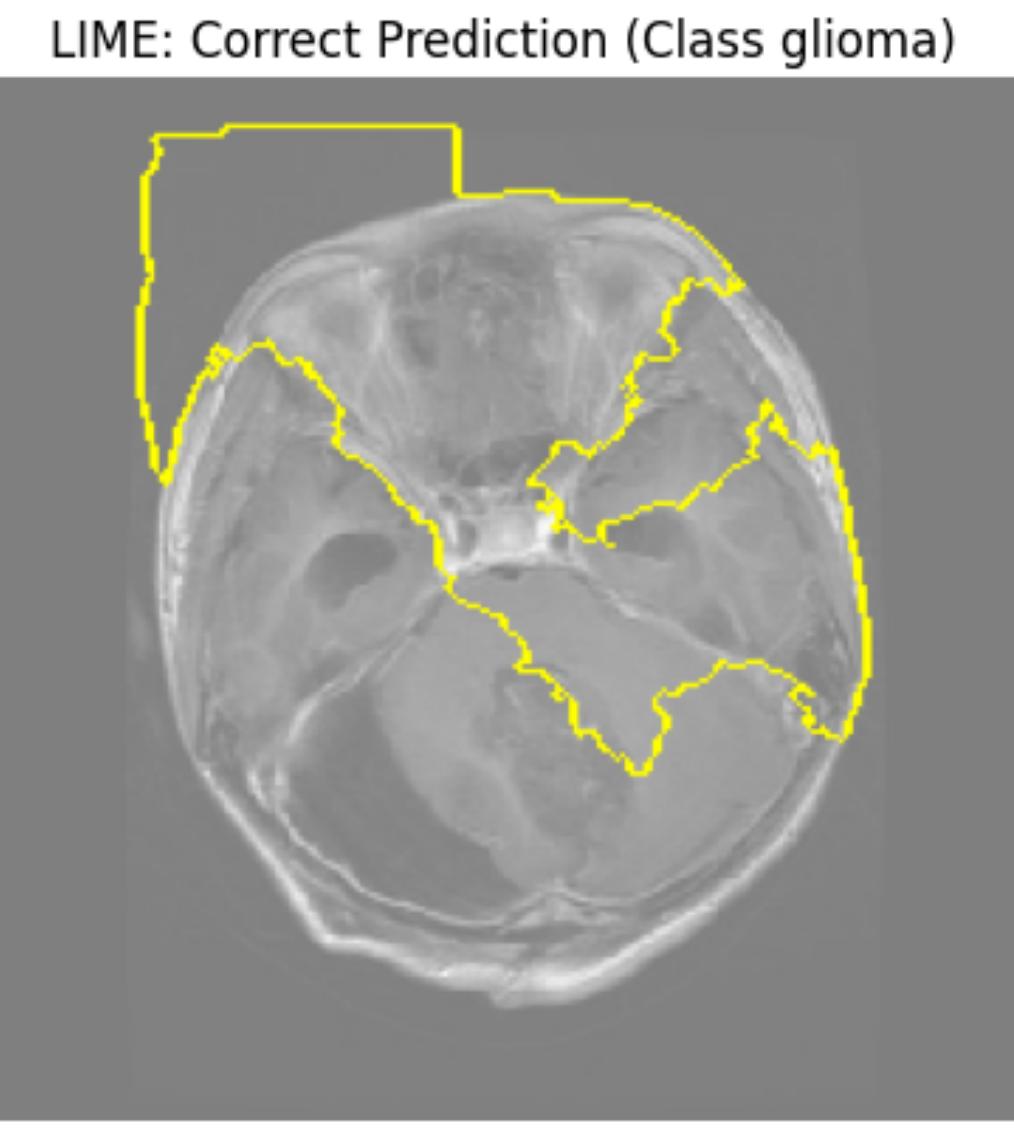
Saliency Map: Correct Prediction (Class glioma)



Saliency Map: Incorrect Prediction (Class meningioma)



Model 3 – XAI



Thank you