CT Image Hemorrhage Classification and Segmentation by ML Methods

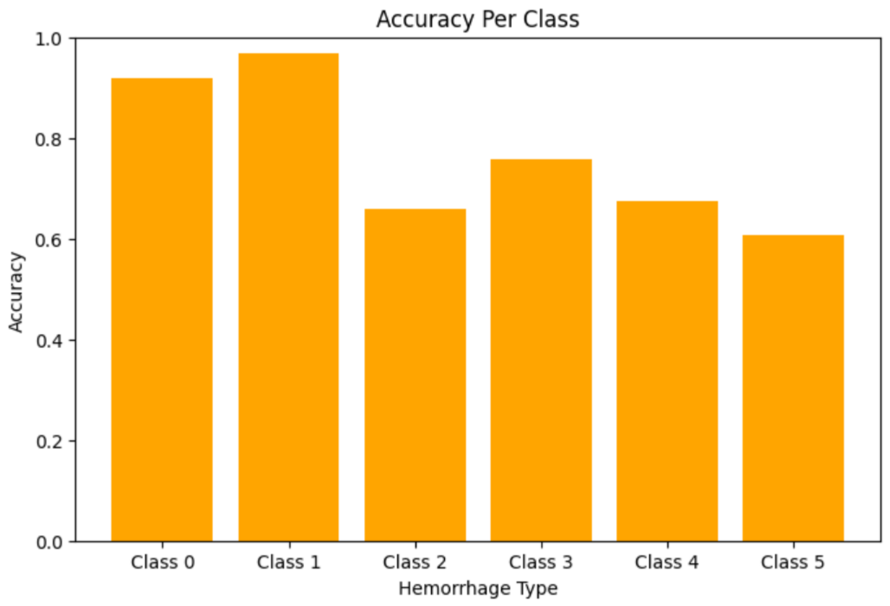
Broderick Kelly, Mitch Whelan, Minoo Mohebbifar

Project Background and Goal

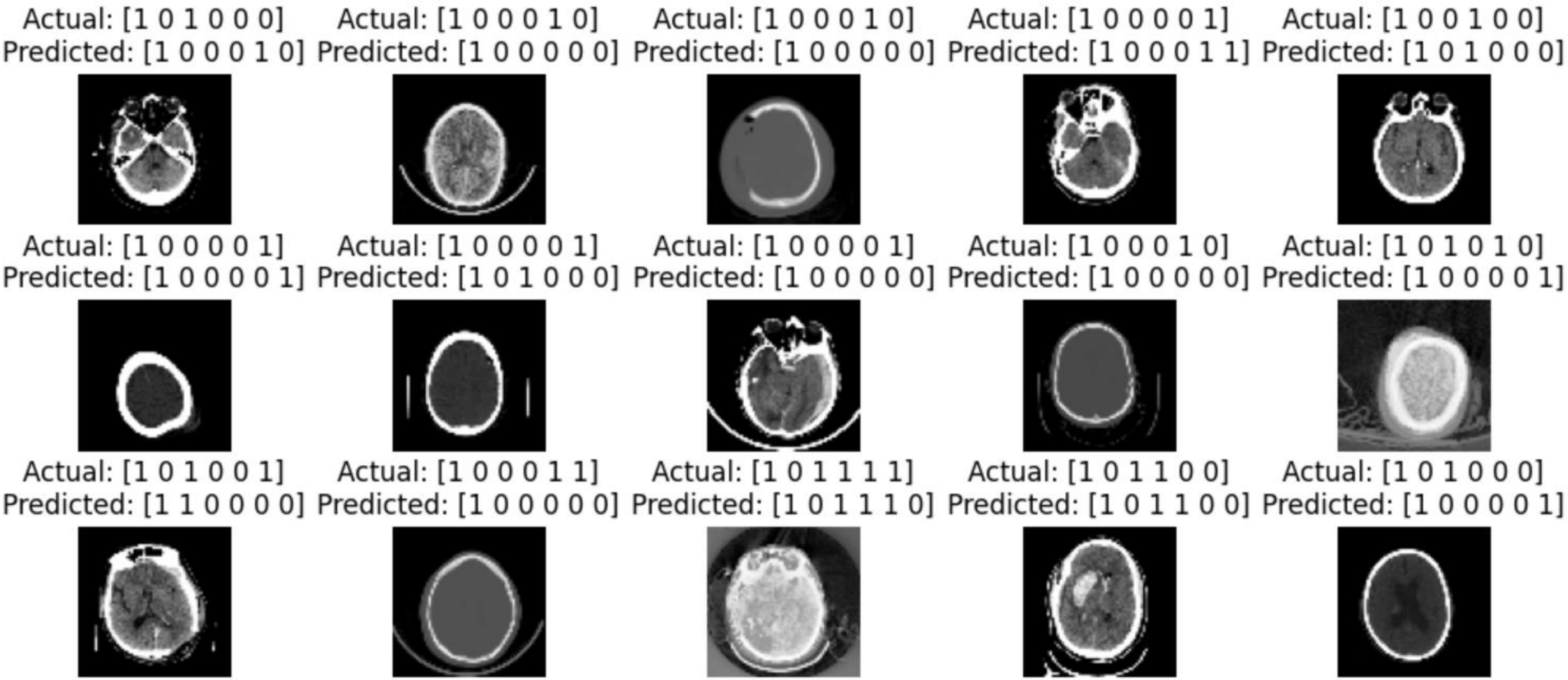
- Zeta Surgical
  - XN partner
  - Improve access to image guidance for emergency and bedside procedures
- Dataset
  - Brain CT scans with different types of hemorrhages (intraparenchymal, intraventricular, subarachnoid, epidural, multi and normal control)
  - Masking data for some images, highlighting hemorrhaged regions
- Goal – apply techniques in ML to perform classification and segmentation of these CT images

Logistic Regression for Classification

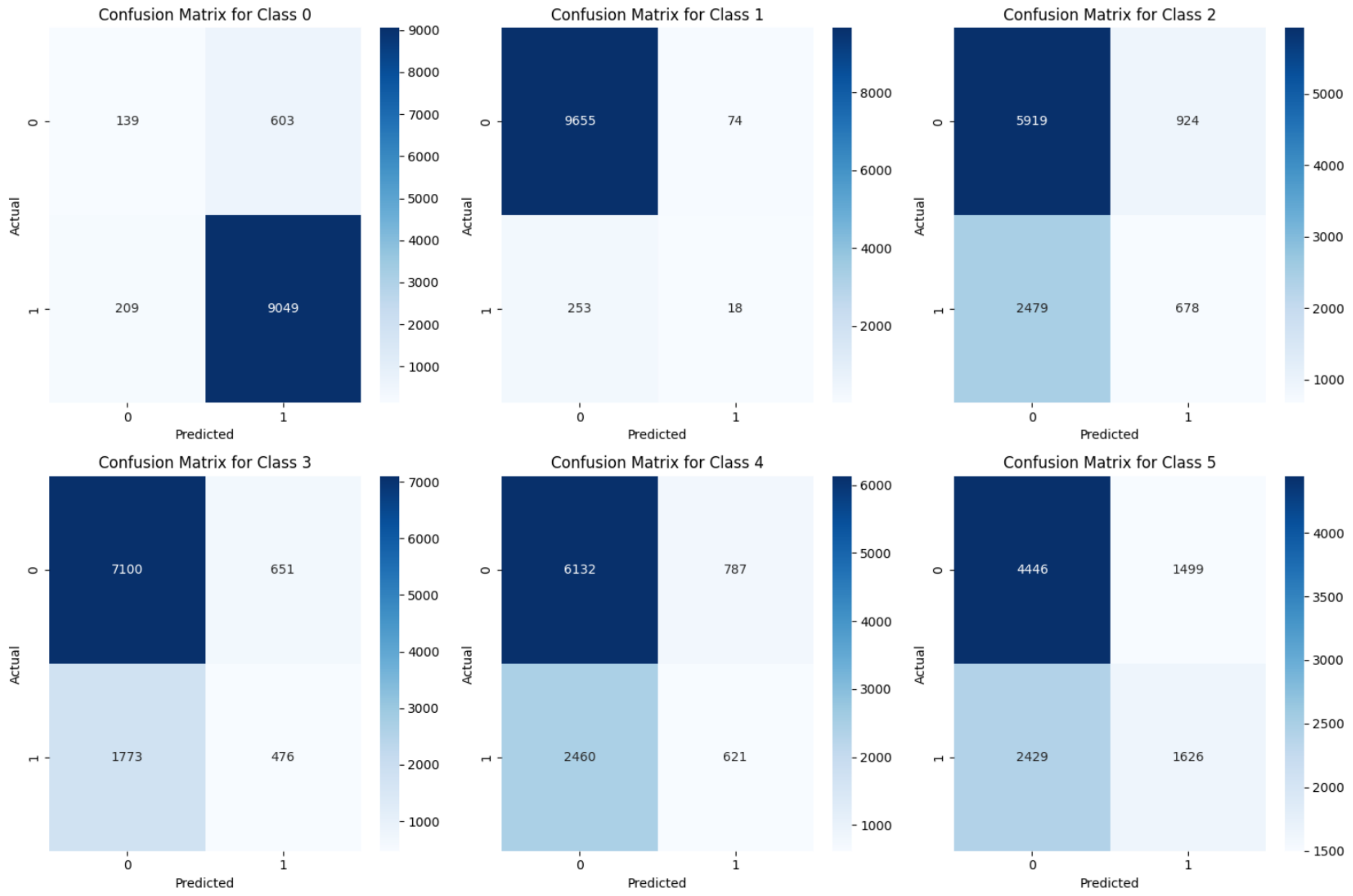
- Used the hemorrhage-labels excel sheet that labeled images from the classes any, epidural, intraparenchymal, intraventricular, subarachnoid, subdural
- Used a random subset of 50,000 images that I loaded in from all windows and classes
- I decided to create an individual binary classification model for each class and combine them
- Total combined model accuracy: ~15%
- Individual model accuracy was good though



Logistic Regression Results

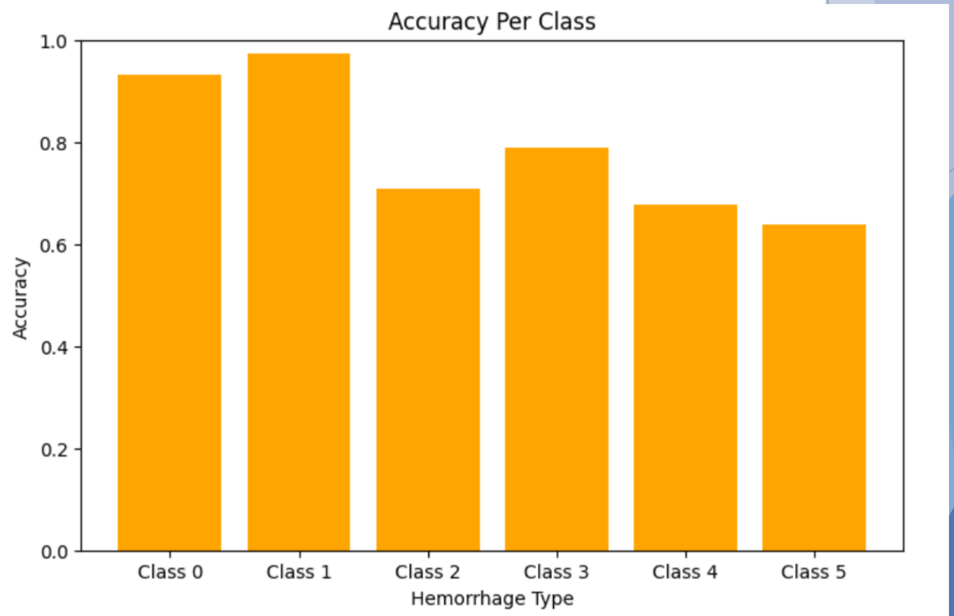


Logistic Regression Results

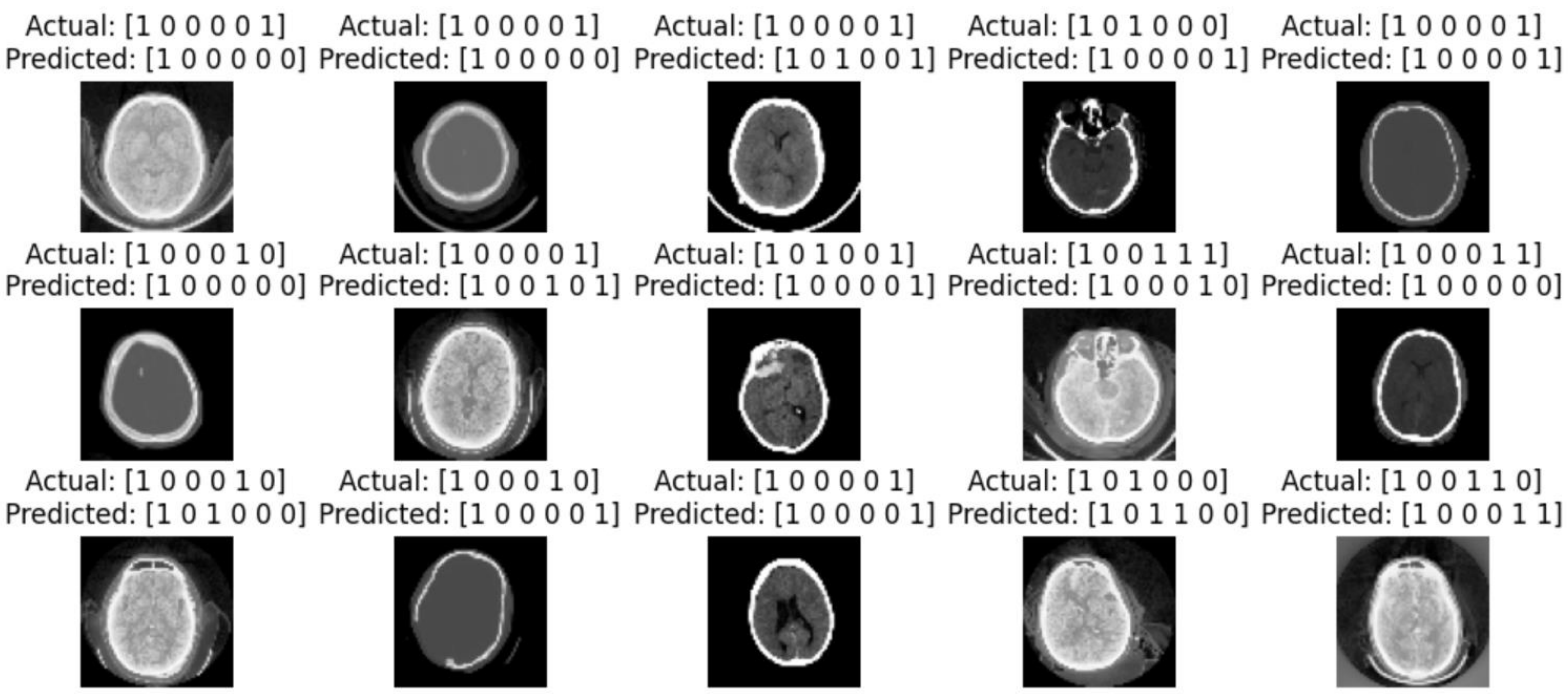


KNN for Classification

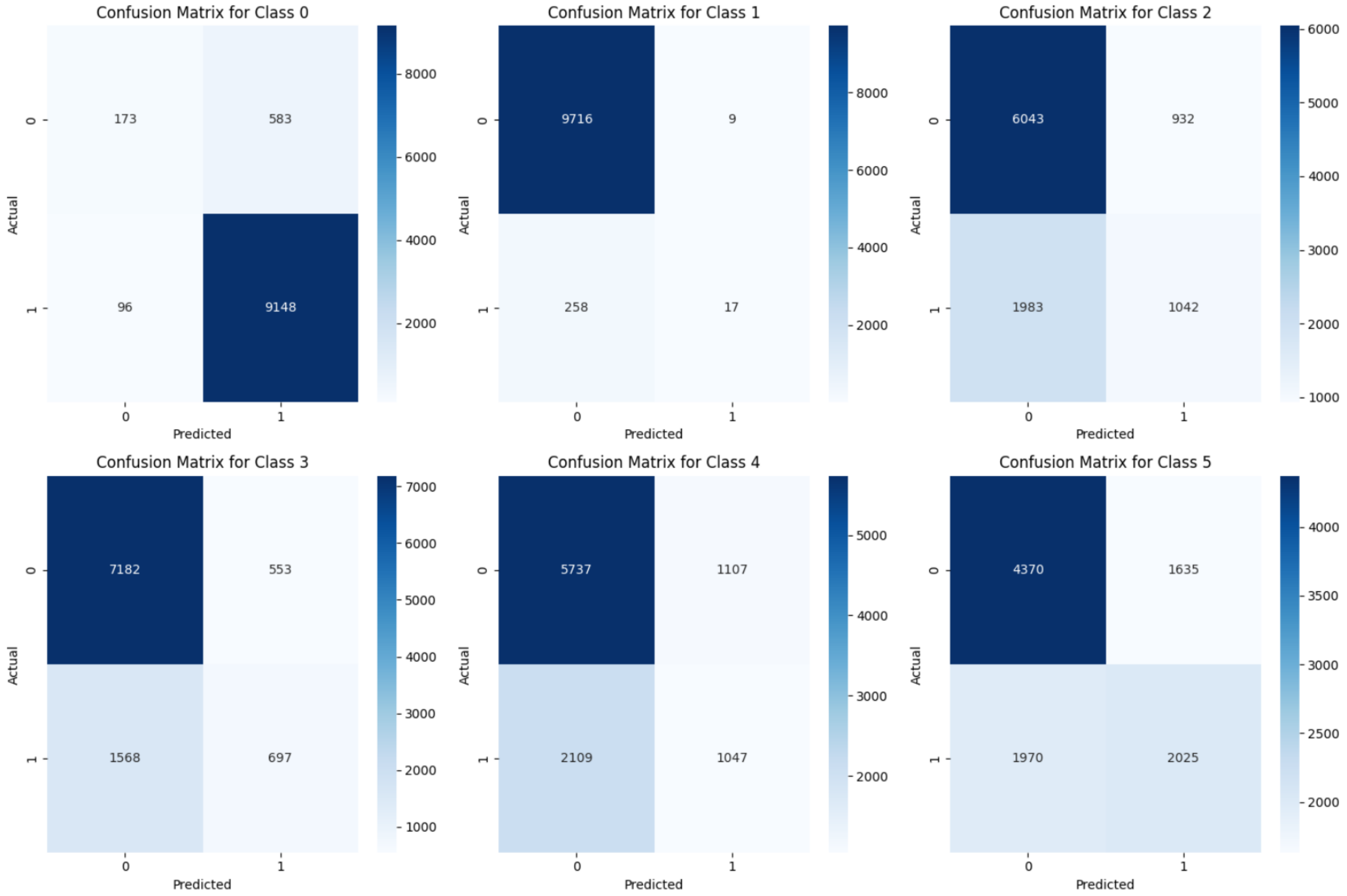
- Used the hemorrhage-labels excel sheet that labeled images from the classes any, epidural, intraparenchymal, intraventricular, subarachnoid, subdural
- Used a random subset of 50,000 images that I loaded in from all windows and classes
- I decided to create an individual binary classification model for each class and combine them
- Total combined model accuracy: ~25%
- Individual model accuracy was still very good



KNN Results


Actual: [1 0 0 0 0 1]
Predicted: [1 0 0 0 0 0]
Actual: [1 0 0 0 0 1]
Predicted: [1 0 0 0 0 0]
Actual: [1 0 0 0 0 1]
Predicted: [1 0 1 0 0 1]
Actual: [1 0 1 0 0 0]
Predicted: [1 0 0 0 0 1]
Actual: [1 0 0 0 0 1]
Predicted: [1 0 0 0 0 1]
Actual: [1 0 0 0 1 0]
Predicted: [1 0 0 0 0 0]
Actual: [1 0 0 0 0 1]
Predicted: [1 0 0 1 0 1]
Actual: [1 0 1 0 0 1]
Predicted: [1 0 0 0 0 1]
Actual: [1 0 0 1 1 1]
Predicted: [1 0 0 0 1 0]
Actual: [1 0 0 0 1 1]
Predicted: [1 0 0 0 0 0]
Actual: [1 0 0 0 1 0]
Predicted: [1 0 1 0 0 0]
Actual: [1 0 0 0 1 0]
Predicted: [1 0 0 0 0 1]
Actual: [1 0 0 0 0 1]
Predicted: [1 0 0 0 0 1]
Actual: [1 0 1 0 0 0]
Predicted: [1 0 1 1 0 0]
Actual: [1 0 0 1 1 0]
Predicted: [1 0 0 0 1 1]

KNN Results



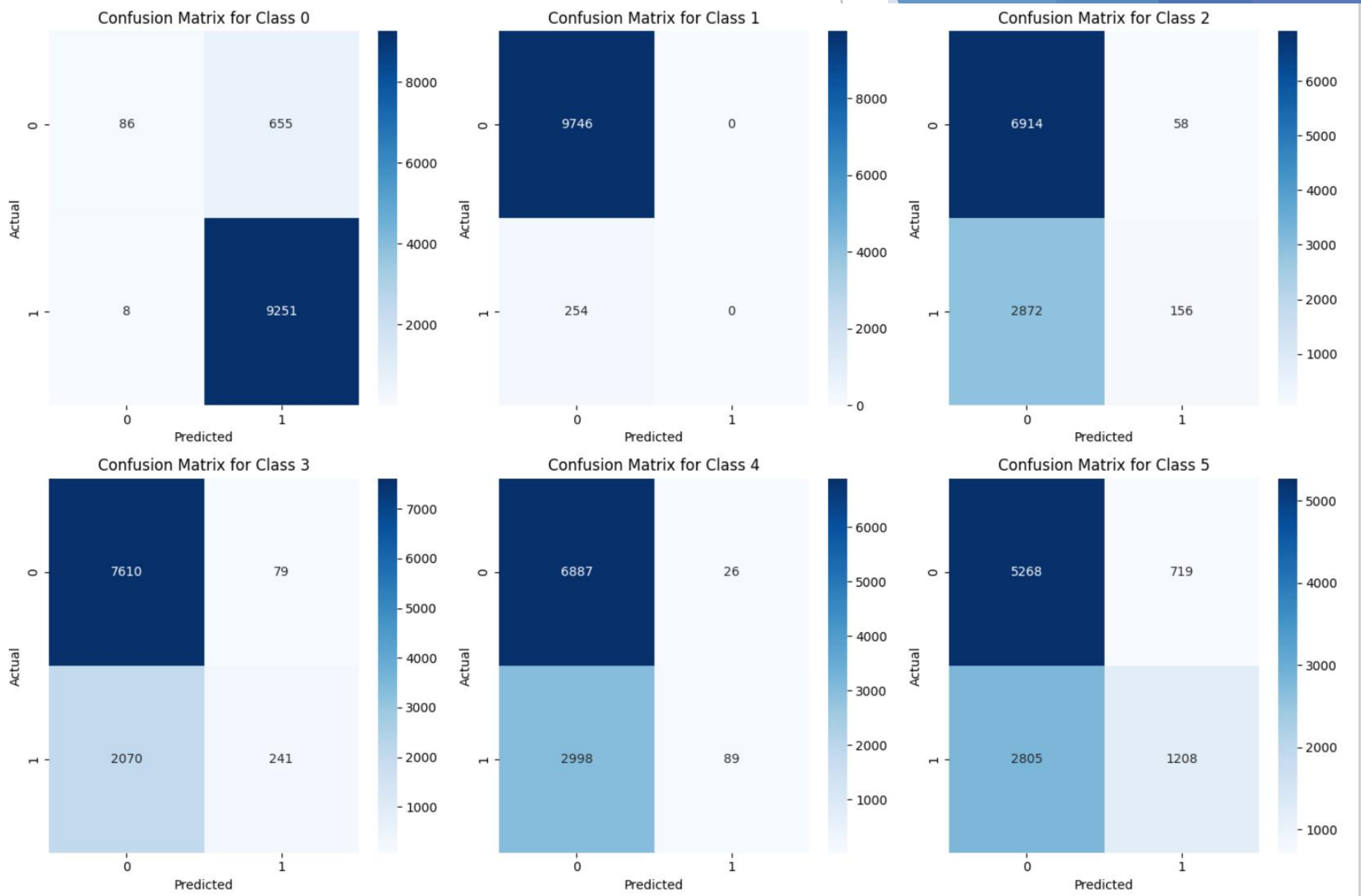
Confusion Matrix for Class 0
173
583
96
9148
Confusion Matrix for Class 1
9716
9
258
17
Confusion Matrix for Class 2
6043
932
1983
1042
Confusion Matrix for Class 3
7182
553
1568
697
Confusion Matrix for Class 4
5737
1107
2109
1047
Confusion Matrix for Class 5
4370
1635
1970
2025
Actual
Predicted

Tree Methods – Decision Tree

Accuracy (overall): 0.1977

Detailed classification report:

	precision	recall	f1-score	support
0	0.94	0.94	0.94	9259
1	0.05	0.06	0.05	254
2	0.36	0.36	0.36	3028
3	0.36	0.36	0.36	2311
4	0.35	0.37	0.36	3087
5	0.46	0.47	0.47	4013
micro avg	0.61	0.62	0.62	21952
macro avg	0.42	0.43	0.42	21952
weighted avg	0.62	0.62	0.62	21952
samples avg	0.58	0.59	0.57	21952

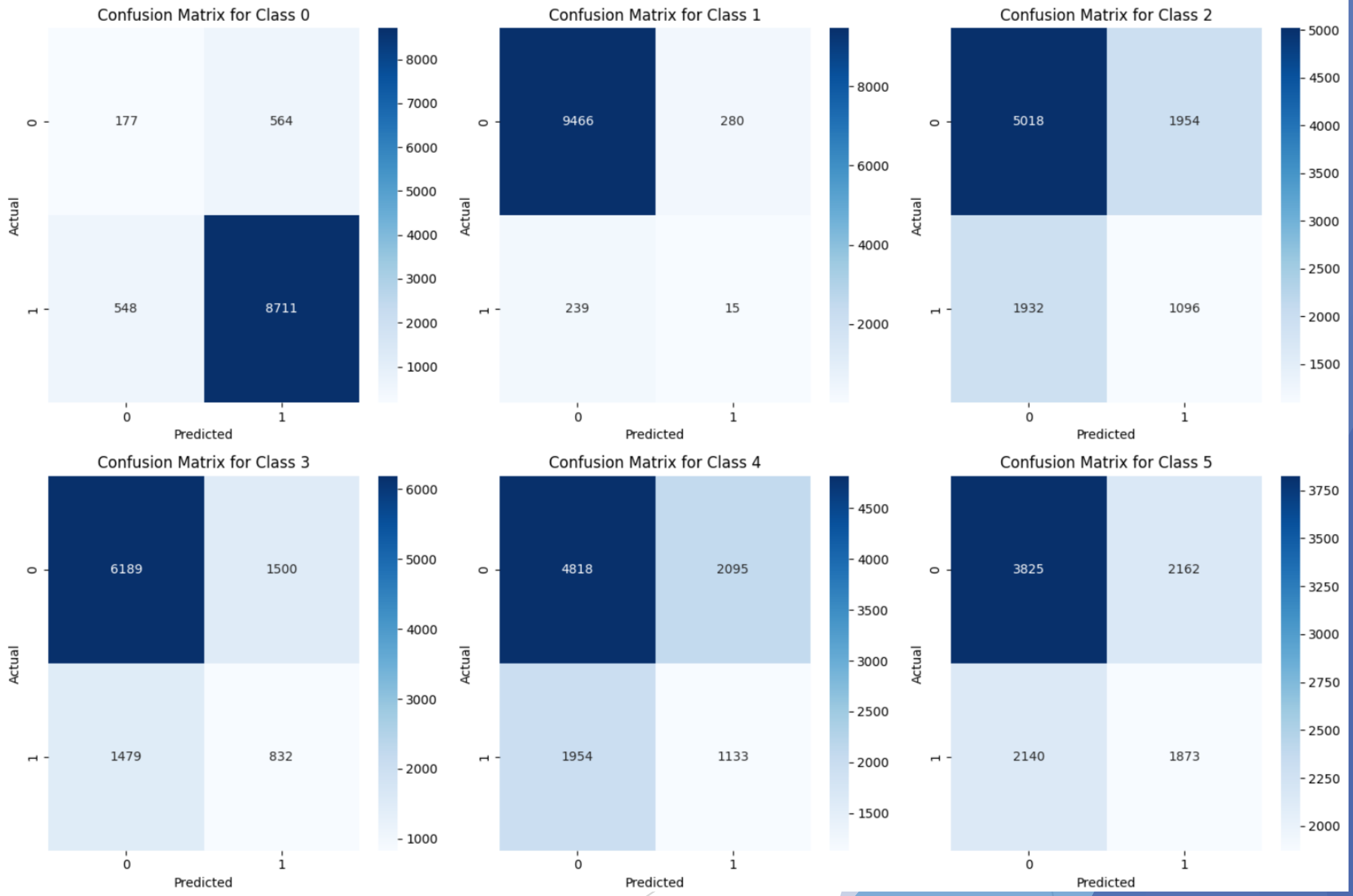


Tree Methods – Random Forest

Accuracy (overall): 0.1194

Detailed classification report:

	precision	recall	f1-score	support
0	0.93	1.00	0.97	9259
1	0.00	0.00	0.00	254
2	0.73	0.05	0.10	3028
3	0.75	0.10	0.18	2311
4	0.77	0.03	0.06	3087
5	0.63	0.30	0.41	4013
micro avg	0.88	0.50	0.64	21952
macro avg	0.64	0.25	0.28	21952
weighted avg	0.80	0.50	0.52	21952
samples avg	0.89	0.48	0.61	21952



QDA Classification

- From 116211 images in dataset 8278 are non-hemorrhoid and the rest have hemorrhoid
- First, we tried to classify images into 7 classes, but QDA didn't converge
- We simplified the problem to detecting hemorrhoid and non-hemorrhoid cases and to balance the number of cases for the two classes we choose randomly 8278 hemorrhoid and 8278 non-hemorrhoid cases.
- We chose only subdural-window to train QDA
- The result accuracy after these simplifications was still very low : %50

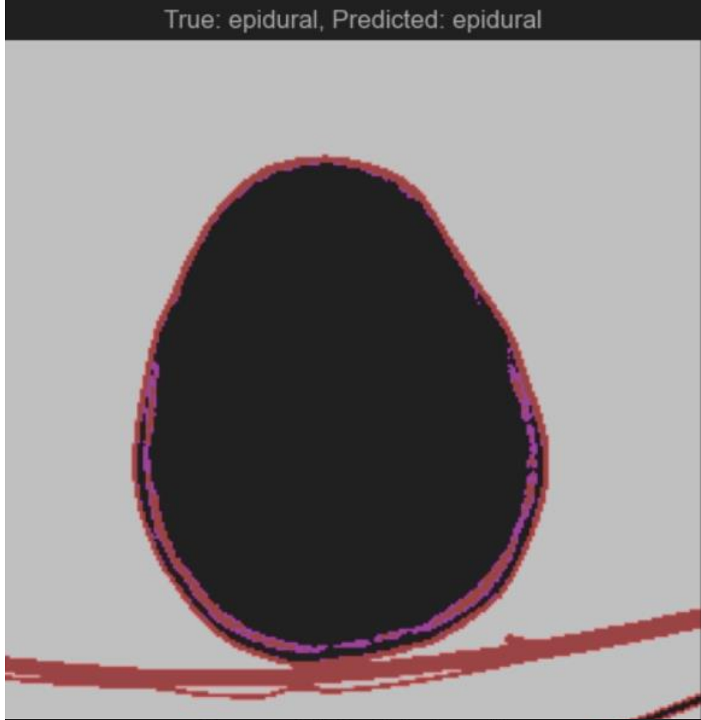
SVM Classification with (Radius Basis Function) RBF Kernel

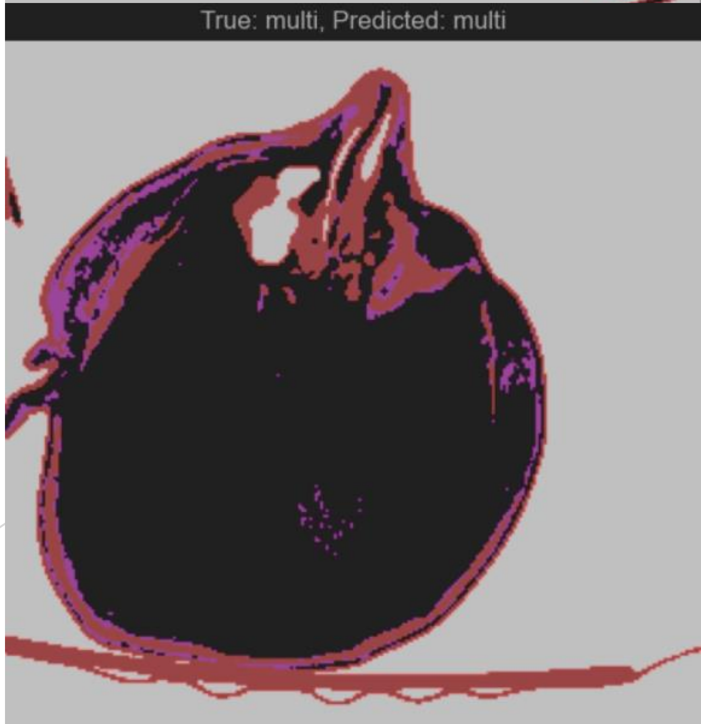
- We start with previous setup we used for QDA where we just classify subdural-window images to hemorrhoid and non-hemorrhoid cases:

Class	Precision	F1-score
Non-Hemorrhoid (0)	%81	0.71
Hemorrhoid (1)	%70	0.77

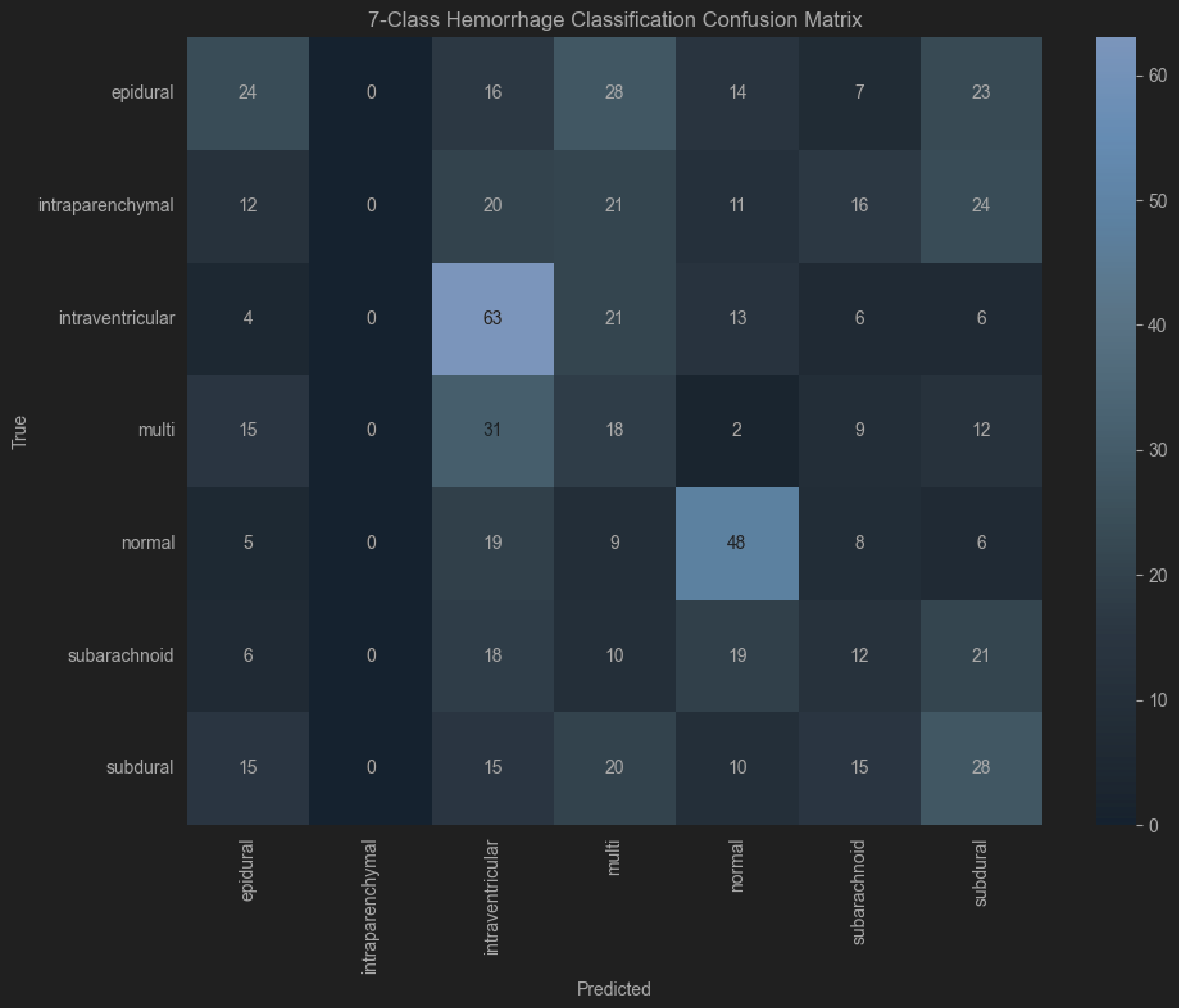
Neural Net Techniques for Classification

- Bone, Brain and Subdural Windows (no max contrast) combined as RGB images (approx. 2000 of each type)
- 7-classes: epidural, intraparenchymal, intraventricular, multi, normal, subarachnoid, subdural
- Down-sampled to 256x256 images (original 512x512)
- Basic CNN - Test accuracy: 0.1502 (almost random)
- CNN with ResNet-50 - Test accuracy: 0.3029 (2x random)
  - “Dead-neuron” effect for intraparenchymal
- Drop multi – Test accuracy: 0.3140 (~2x random)
  - Class prediction disparities persist
- 2-class - Test accuracy: 0.7054
- CNN with xception base (7 class) - Test accuracy: 0.6586



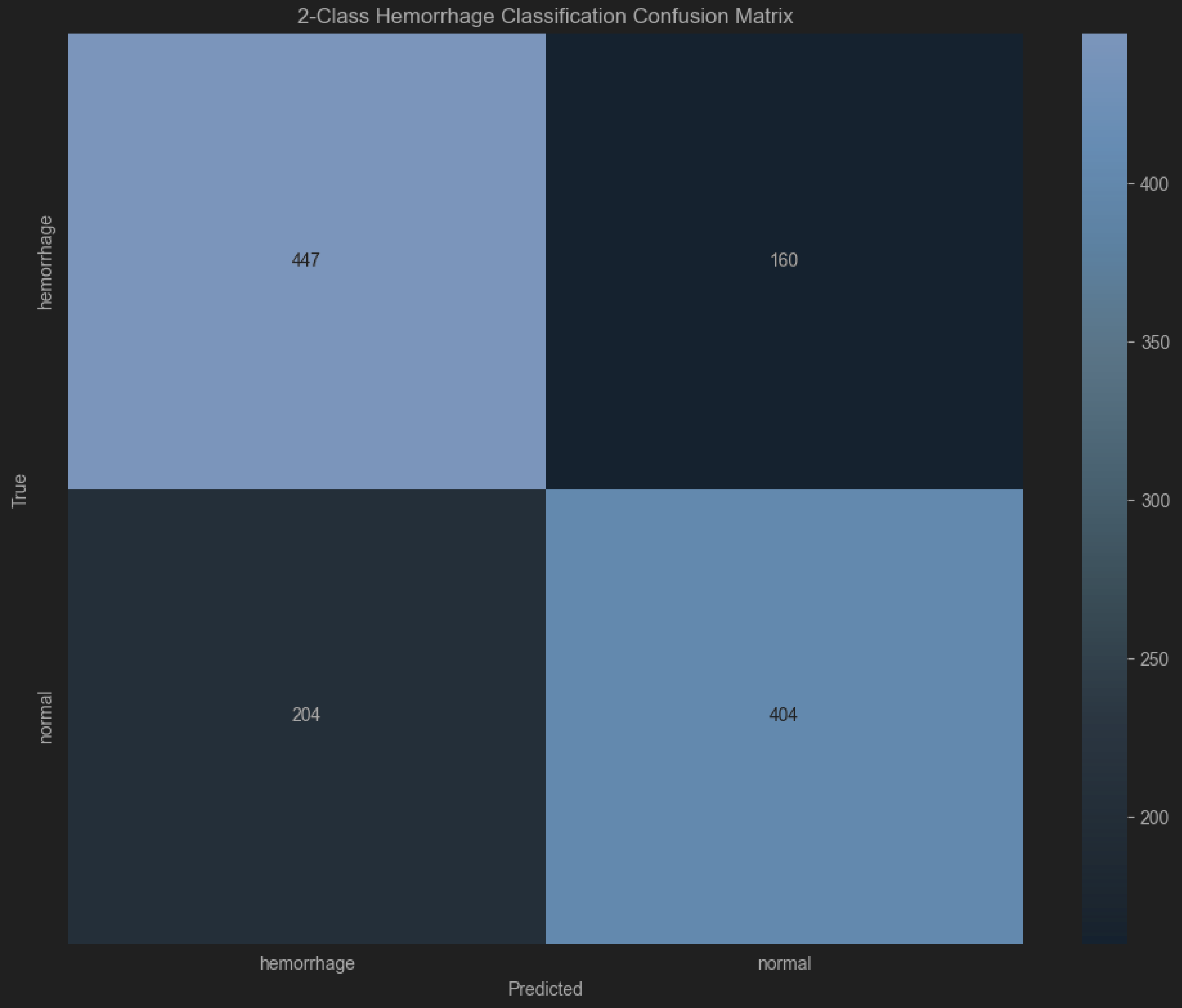


7-Class Confusion Matrix

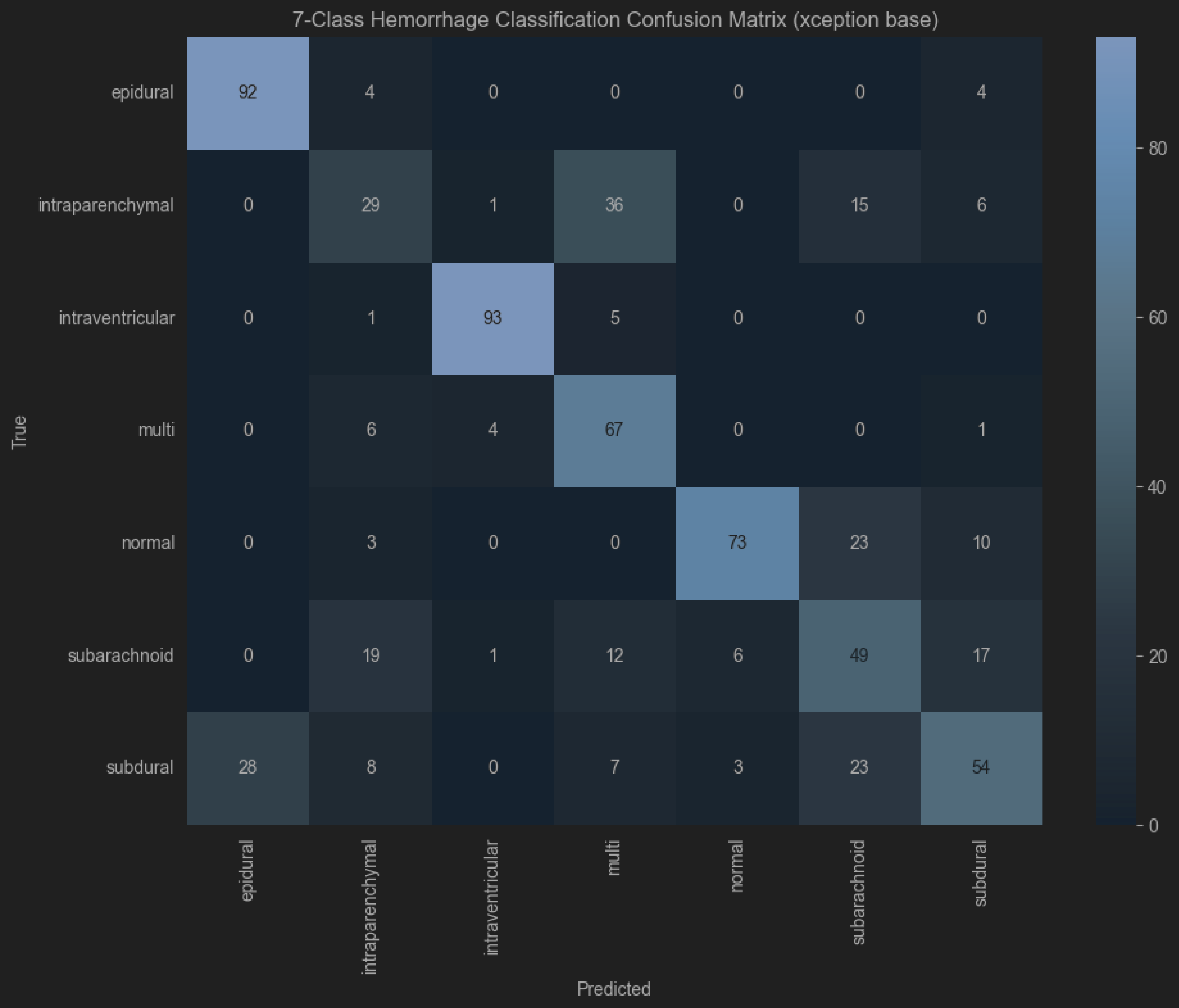


6-Class Confusion Matrix

2-Class Confusion Matrix

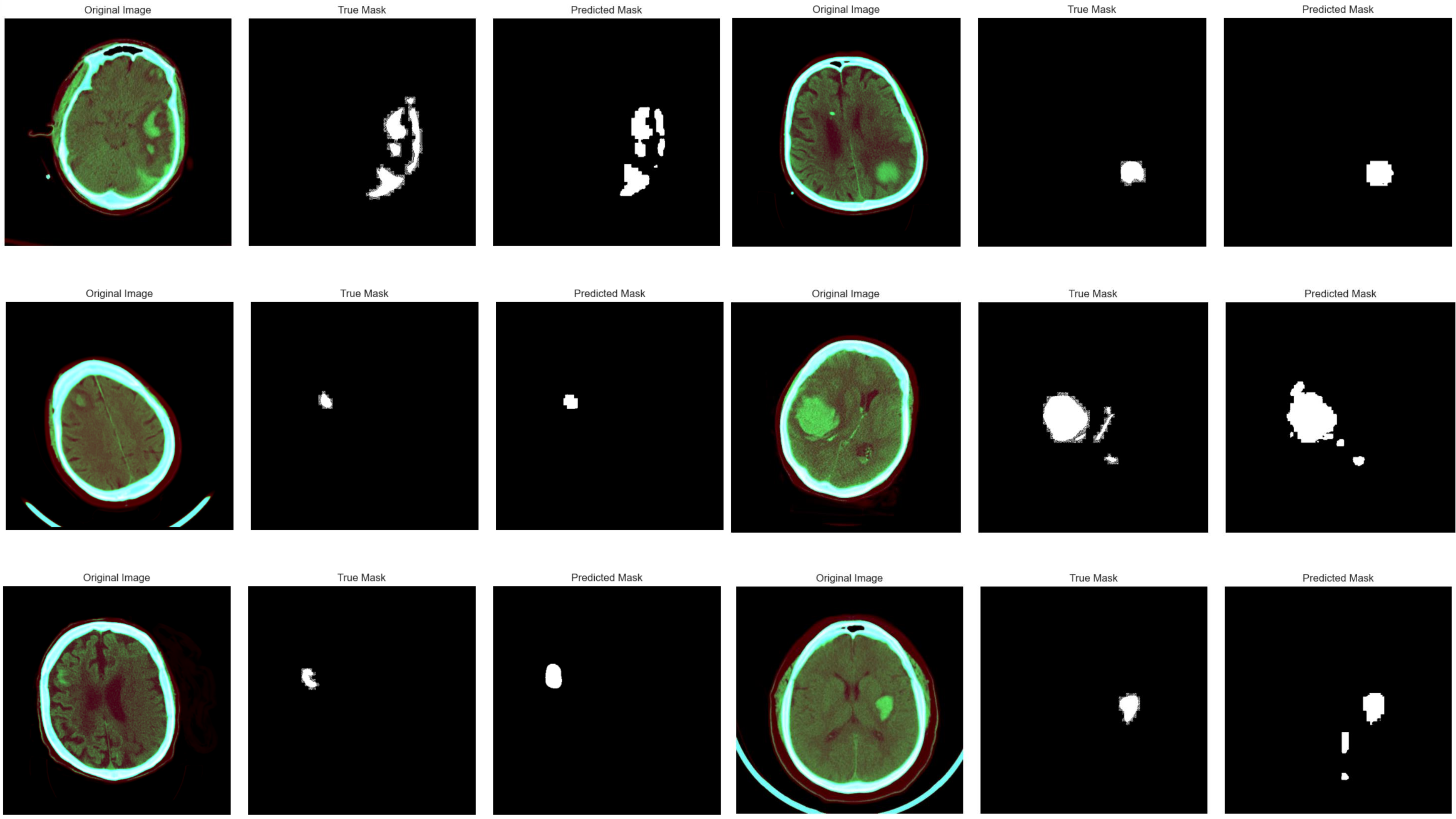


7-Class Confusion Matrix (xception)



Neural Net Techniques for Segmentation

- Mask Generation: Binary and Quaternary
- Bone, Brain and Subdural Windows combined as RGB images (max contrast omitted)
- General Difficulties:
  - CPU & GPU constraints and batch processing
  - Quaternary / non-binary masks and blank images
- Simple U-NET approach – struggled and predicted empty images or overly large regions
- Pretrained Base Model (ResNet-50)
  - Test Accuracy: 0.9894, Precision: 0.7732, Recall: 0.7165


Original Image
True Mask
Predicted Mask
Original Image
True Mask
Predicted Mask
Original Image
True Mask
Predicted Mask
Original Image
True Mask
Predicted Mask
Original Image
True Mask
Predicted Mask
Original Image
True Mask
Predicted Mask