

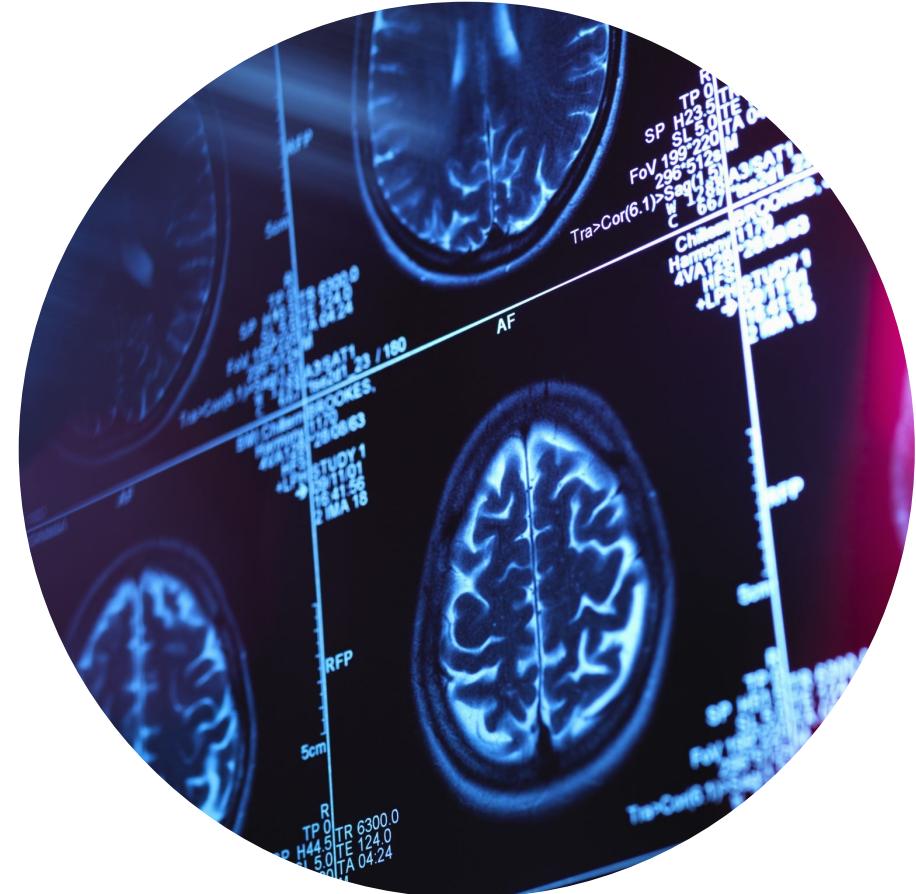
# Brain CT image Hemorrhage Prediction

MATH7243 XN Project

Hsiao-Yu Peng  
Northeastern University

# Outline

- Introduction
- Dataset
- Model Training & Evaluation
- Discussion



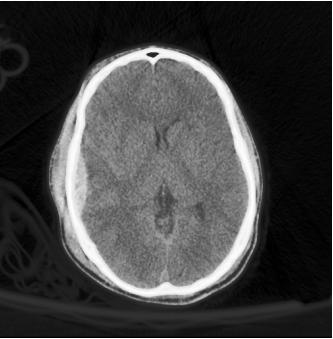
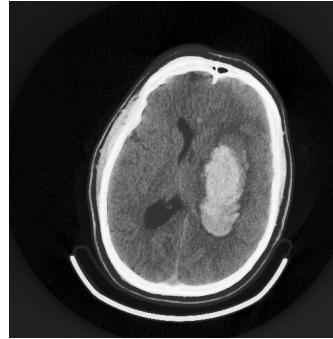
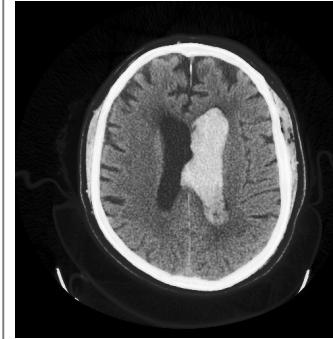
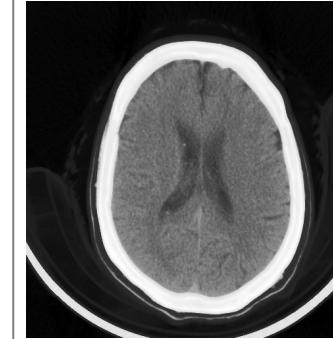
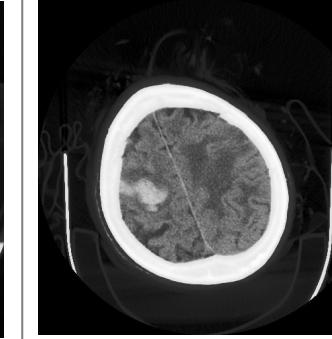
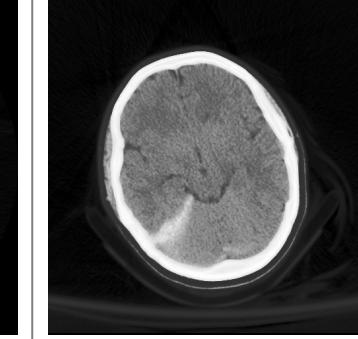
# Introduction

- The early detection of brain hemorrhages is crucial for timely and effective medical intervention.
- Challenges include limited access to fast, accurate image guidance and constraints in utilizing it directly at the point-of-care, impacting treatments for brain hemorrhages.
- This project focuses on leveraging the power of machine learning to develop models for classification, regression, and segmentation of hemorrhages in brain CT images.



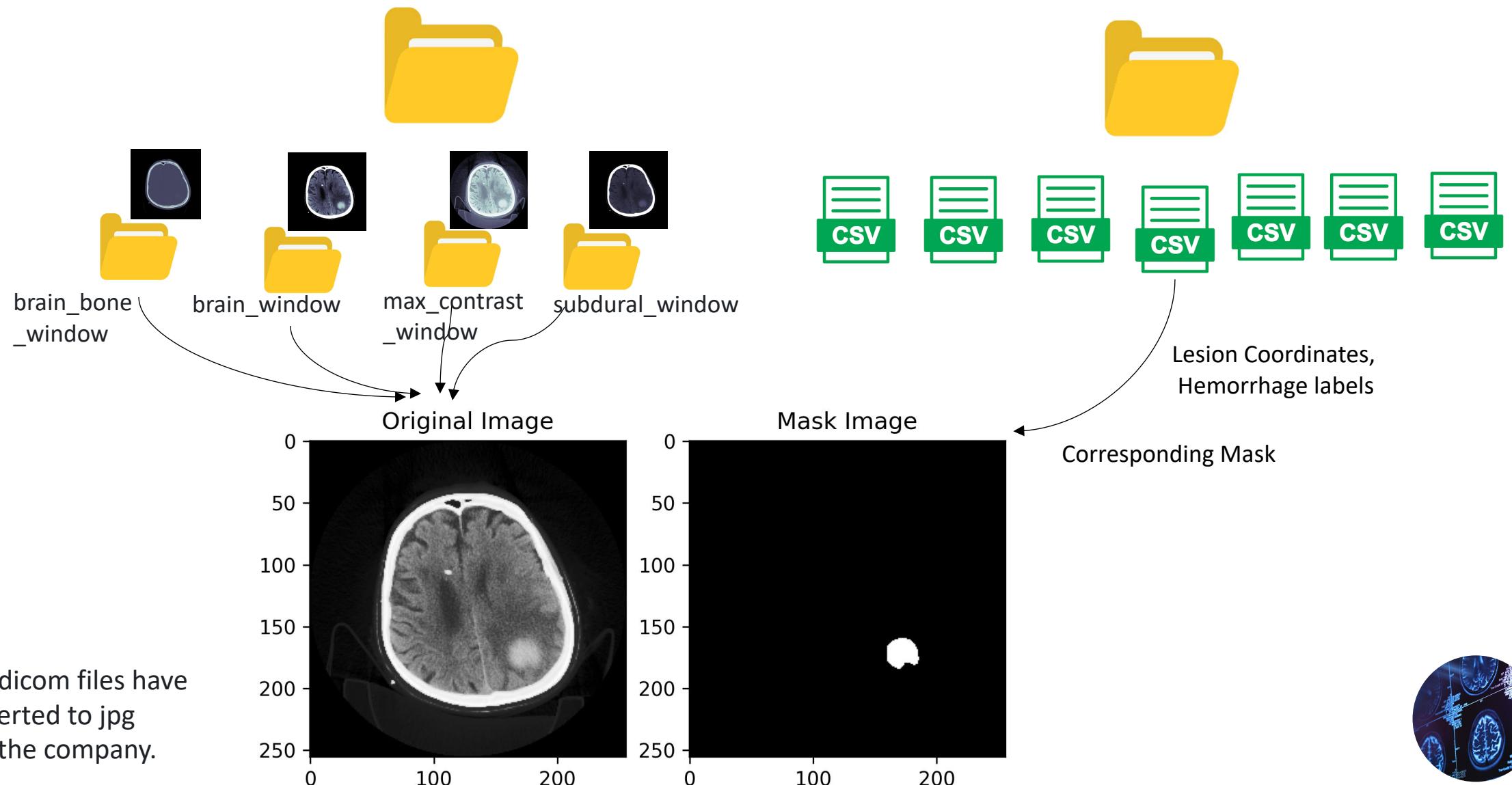
# Dataset

- Experimental Network (XN) Project on MATH 7243
- Partner Company: Zeta Surgical
- Brain Hemorrhage Type:

epidural: 0	Intra-parenchymal: 1	Intra-ventricular: 2	normal: 3	subarachnoid: 4	subdural: 5
					



# Data Preparation



# Traditional Machine Learning Classification Report

## ML Over Sampling

	Precision	Recall	F1-score	Accuracy
Logistic Regression	0.28	0.30	0.28	0.303
LDA	0.63	0.64	0.63	0.635
QDA	0.90	0.74	0.77	0.741
Random Forest	0.76	0.76	0.76	0.762
SVM	0.54	0.55	0.54	0.552

## ML Down Sampling

	Precision	Recall	F1-score	Accuracy
Logistic Regression	0.20	0.25	0.20	0.233
LDA	0.32	0.32	0.31	0.313
QDA	0.32	0.19	0.17	0.189
Random Forest	0.33	0.34	0.32	0.329
SVM	0.28	0.29	0.28	0.284



# ML Over Sampling

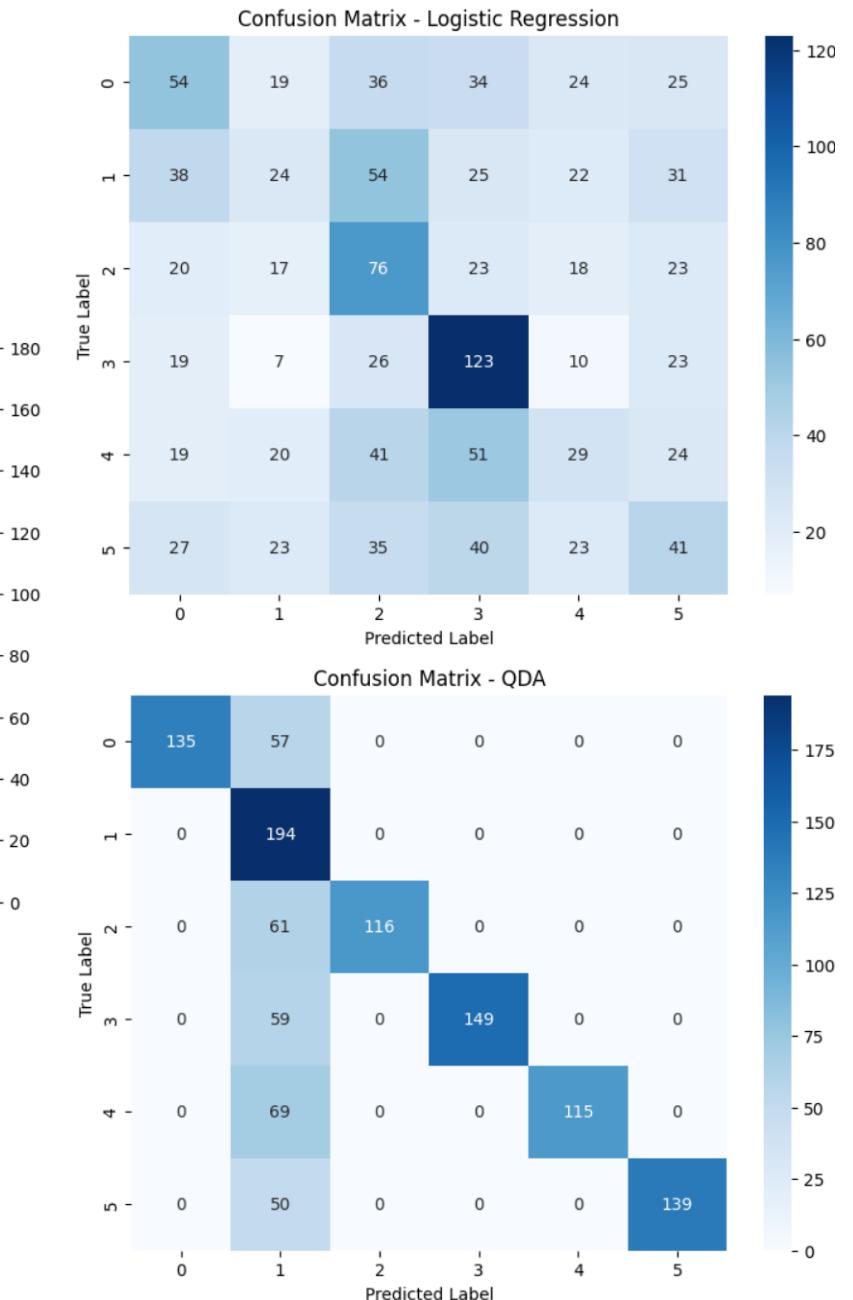
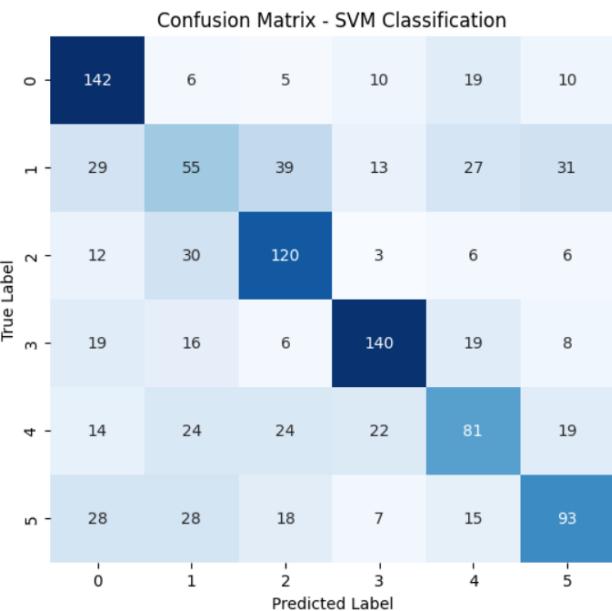
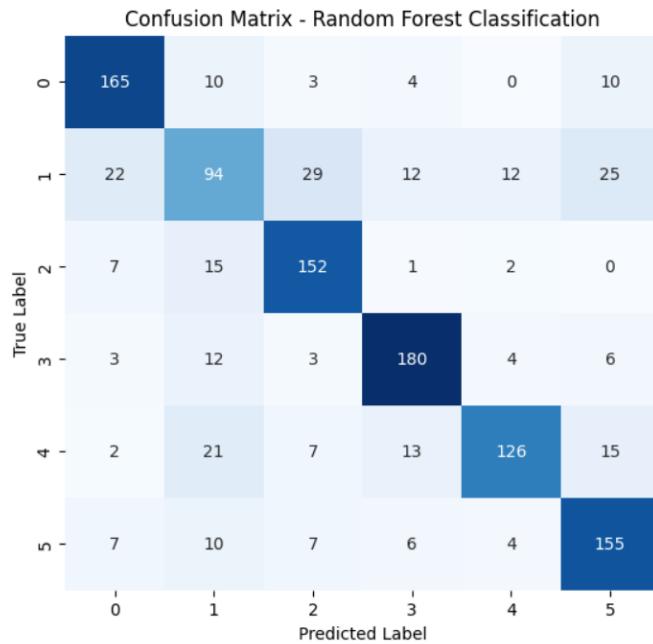
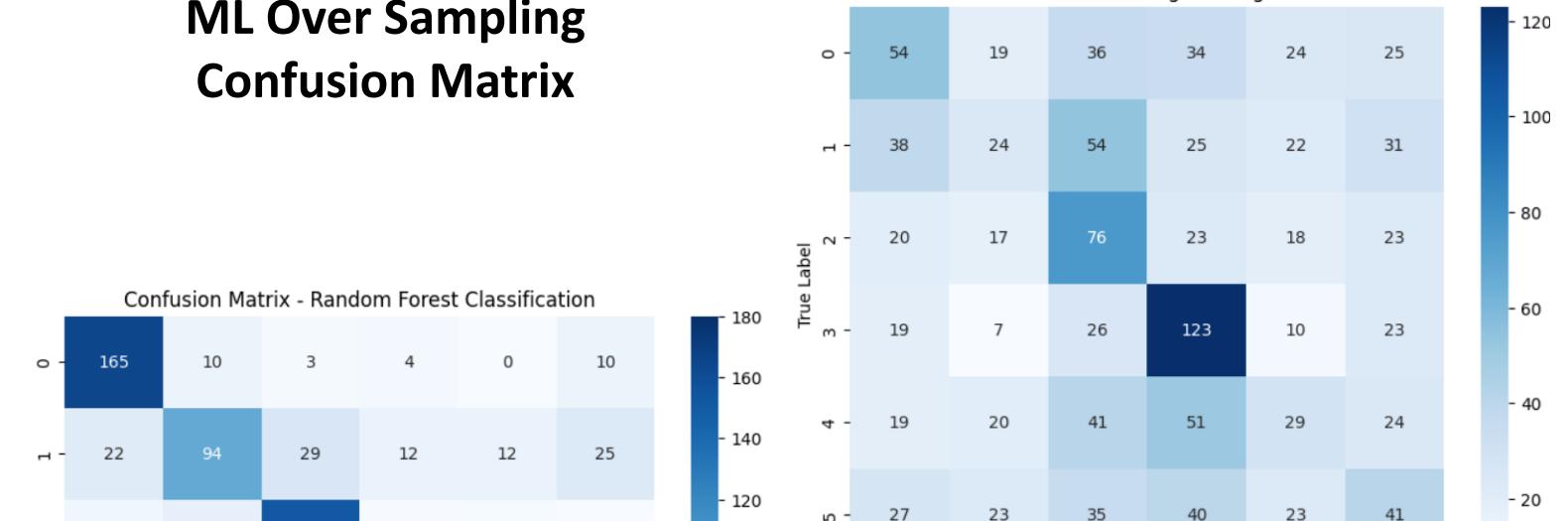
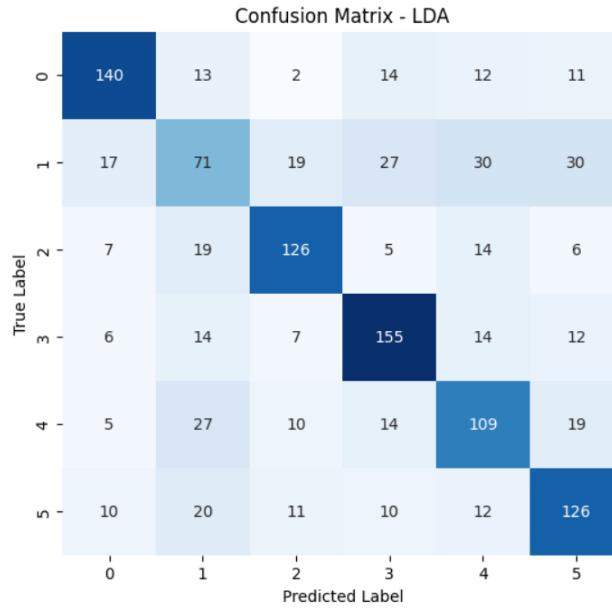
	Precision					Recall					F1-score				
	Logistic	LDA	QDA	RF	SVM	Logistic	LDA	QDA	RF	SVM	Logistic	LDA	QDA	RF	SVM
0	0.31	0.76	1.00	0.80	0.58	0.28	0.73	0.70	0.86	0.74	0.29	0.74	0.83	0.83	0.65
1	0.22	0.43	0.4	0.58	0.35	0.12	0.37	1.00	0.48	0.28	0.16	0.40	0.57	0.53	0.31
2	0.28	0.72	1.00	0.76	0.57	0.43	0.71	0.66	0.86	0.68	0.34	0.72	0.79	0.80	0.62
3	0.42	0.69	1.00	0.83	0.72	0.59	0.75	0.72	0.87	0.67	0.49	0.72	0.83	0.85	0.69
4	0.23	0.57	1.00	0.85	0.49	0.16	0.59	0.62	0.68	0.44	0.19	0.58	0.77	0.76	0.46
5	0.25	0.62	1.00	0.73	0.56	0.22	0.69	0.74	0.82	0.49	0.23	0.64	0.85	0.78	0.52

Class Mapping:

0: epidural, 1: intraparenchymal, 2: intraventricular, 3: normal, 4: subarachnoid, 5: subdural



# ML Over Sampling Confusion Matrix



# Image Segmentation

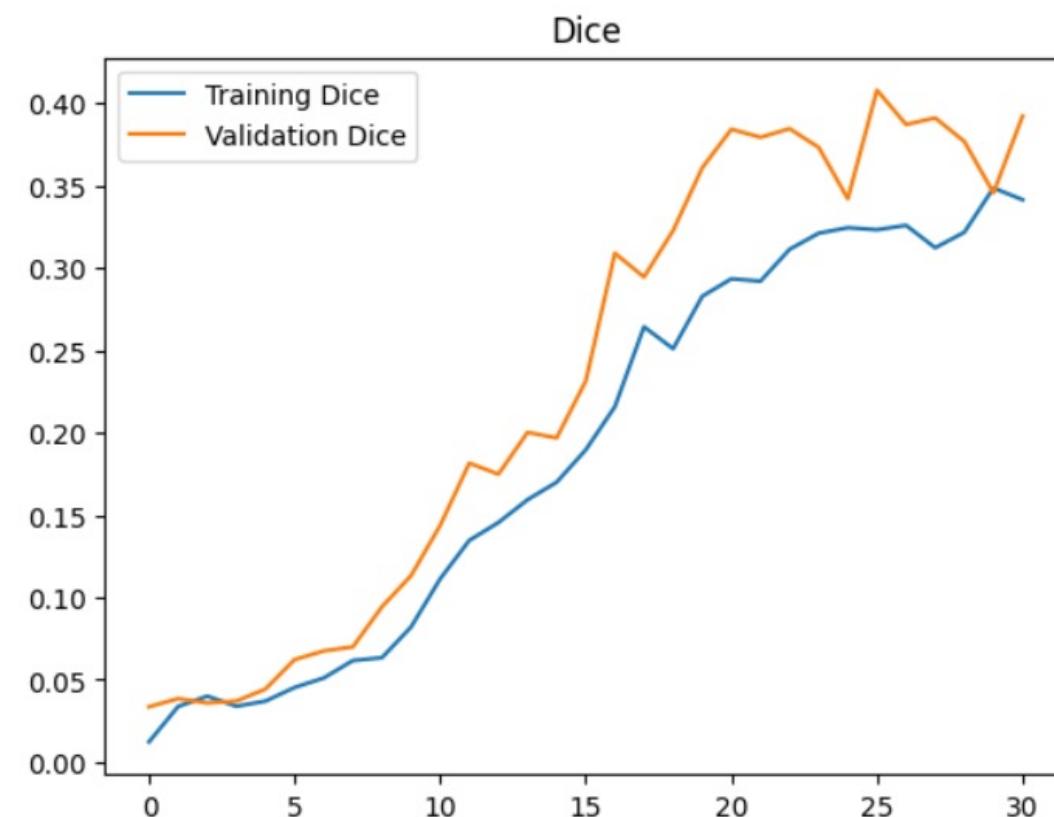
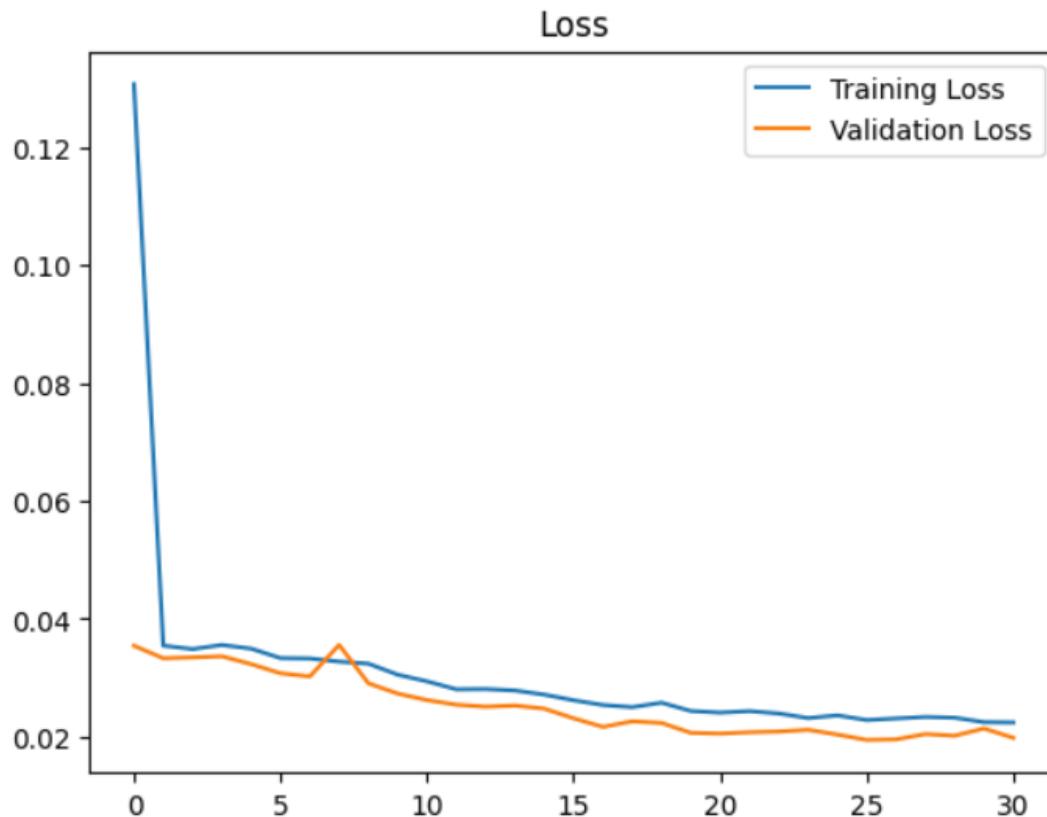
- U-net
- Pilot Run: Utilized 4 types of image data sets. "Epidural", "Intraparenchymal", "Subdural", "Intraventricular"
- Final Run
  - a. Incorporated additional images, subarachnoid and normal types, for training the U-net model.
  - b. Data Augmentation Strategies
- Dice coefficient (from 0 to 1)
  - measures the similarity between two sets of data

$$\frac{2*|X \cap Y|}{|X|+|Y|}$$

(X is the predicted set of pixels and Y is the ground truth)



# Segmentation (Pilot Run)

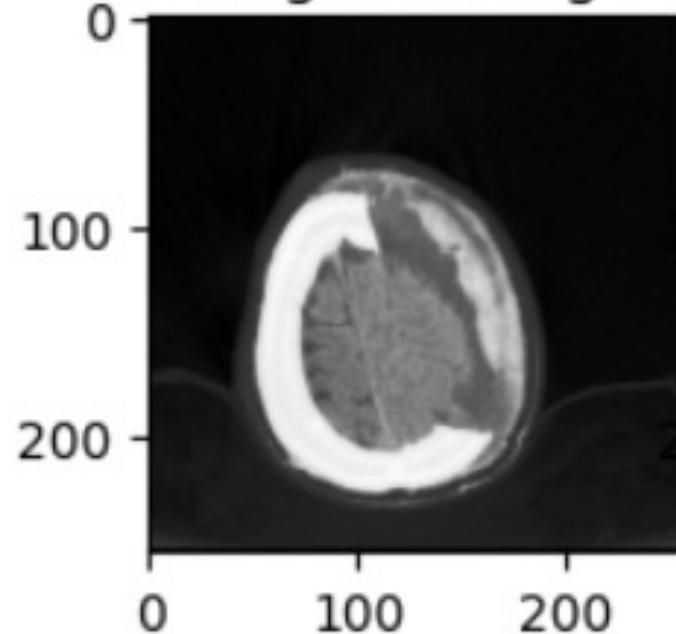


# Segmentation (Pilot Run)

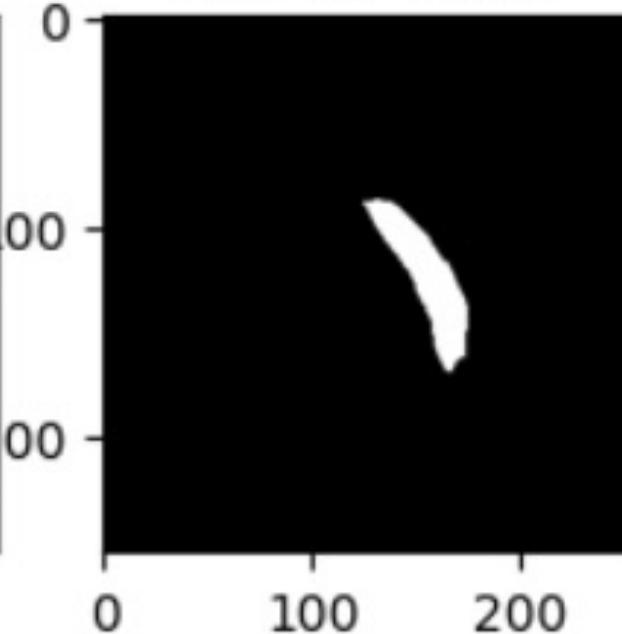
Dice Coefficient on Test Set: 0.3565

Dice Loss on Test Set: 0.6435

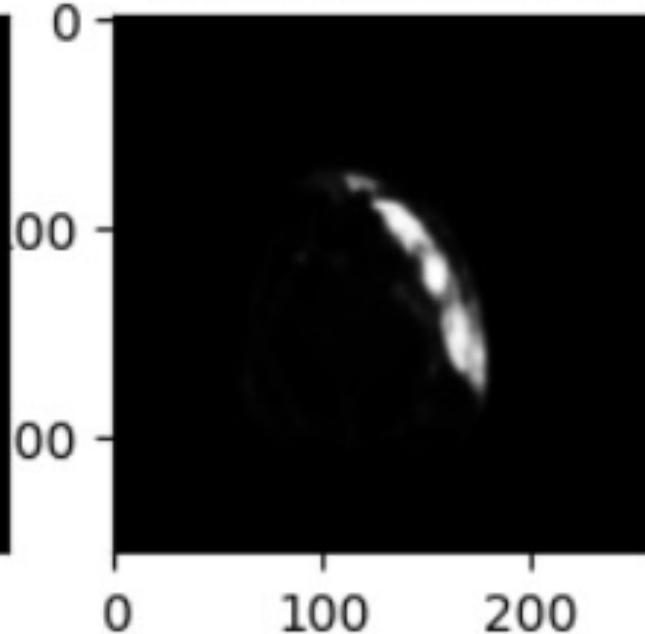
Original Image



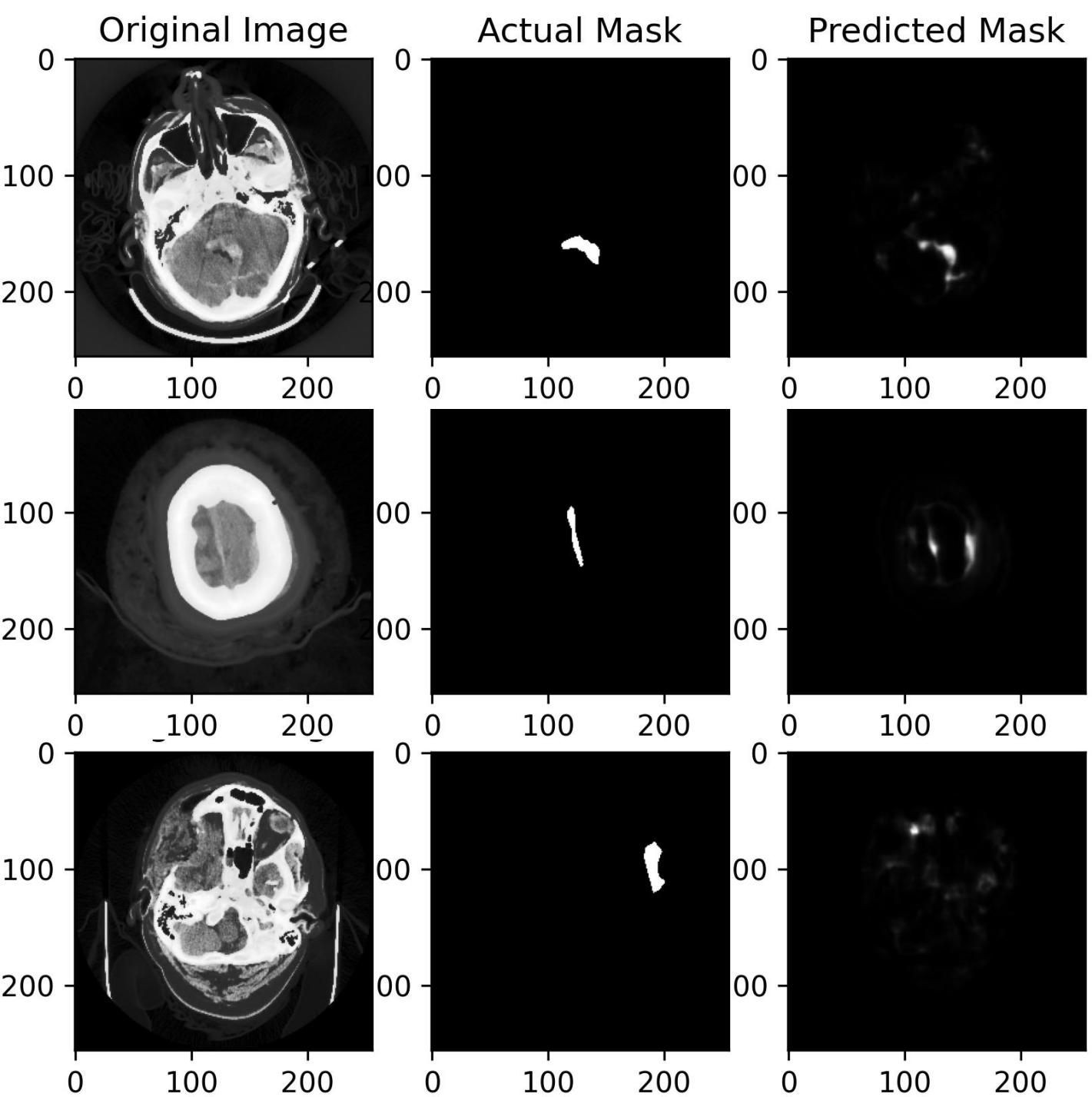
Actual Mask



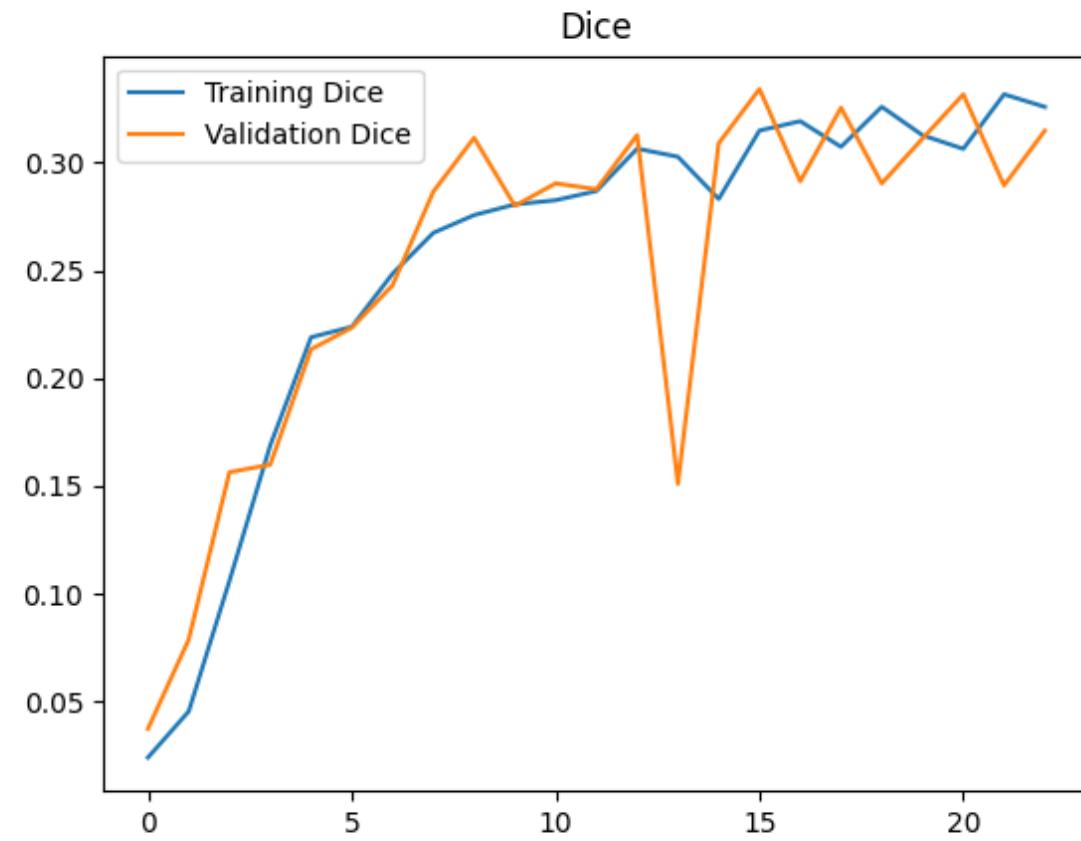
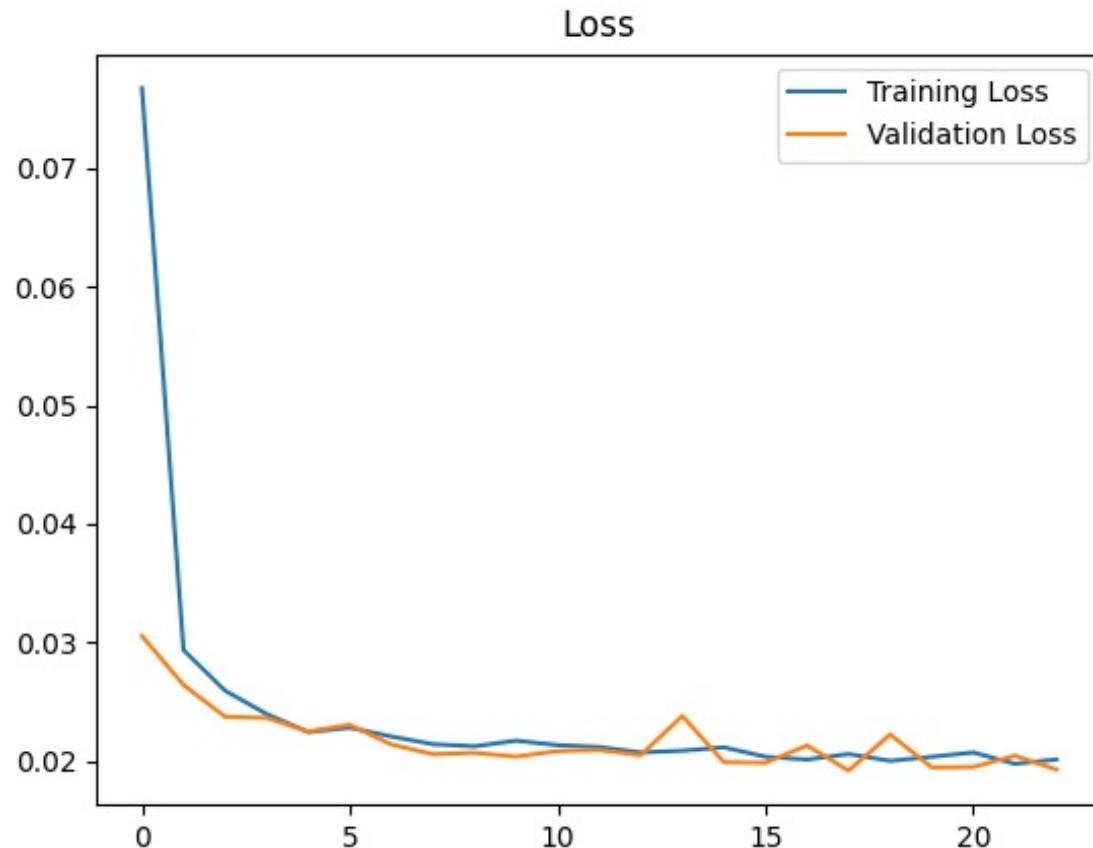
Predicted Mask



# Segmentation (Pilot Run)

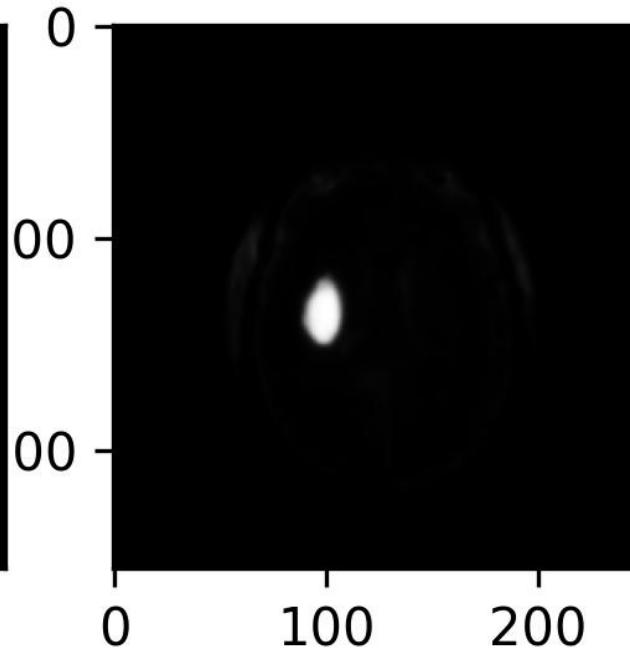
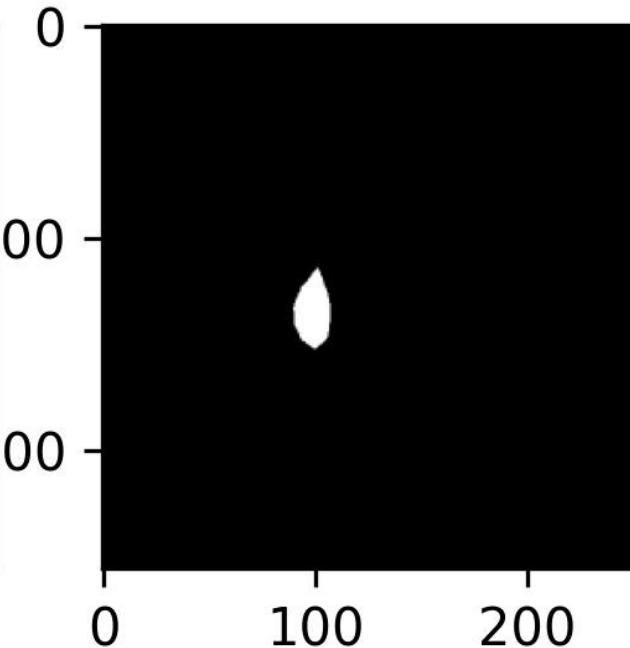
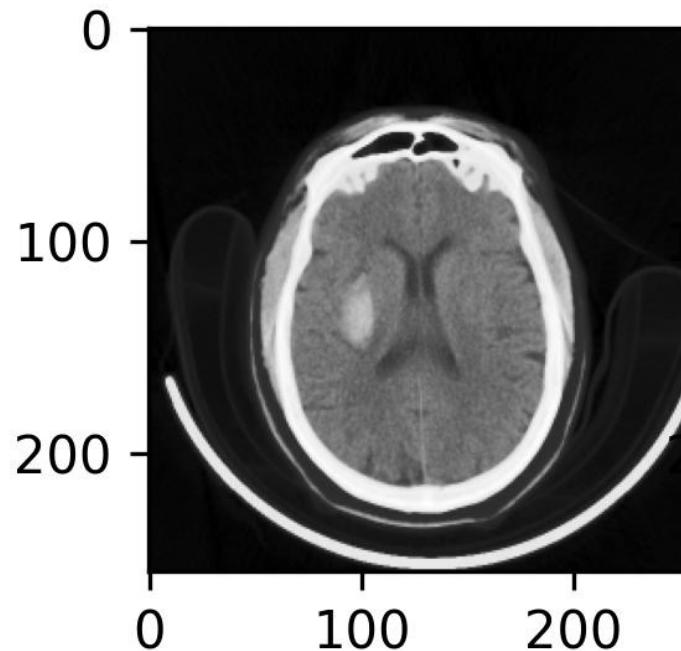


# Segmentation (Final Run)

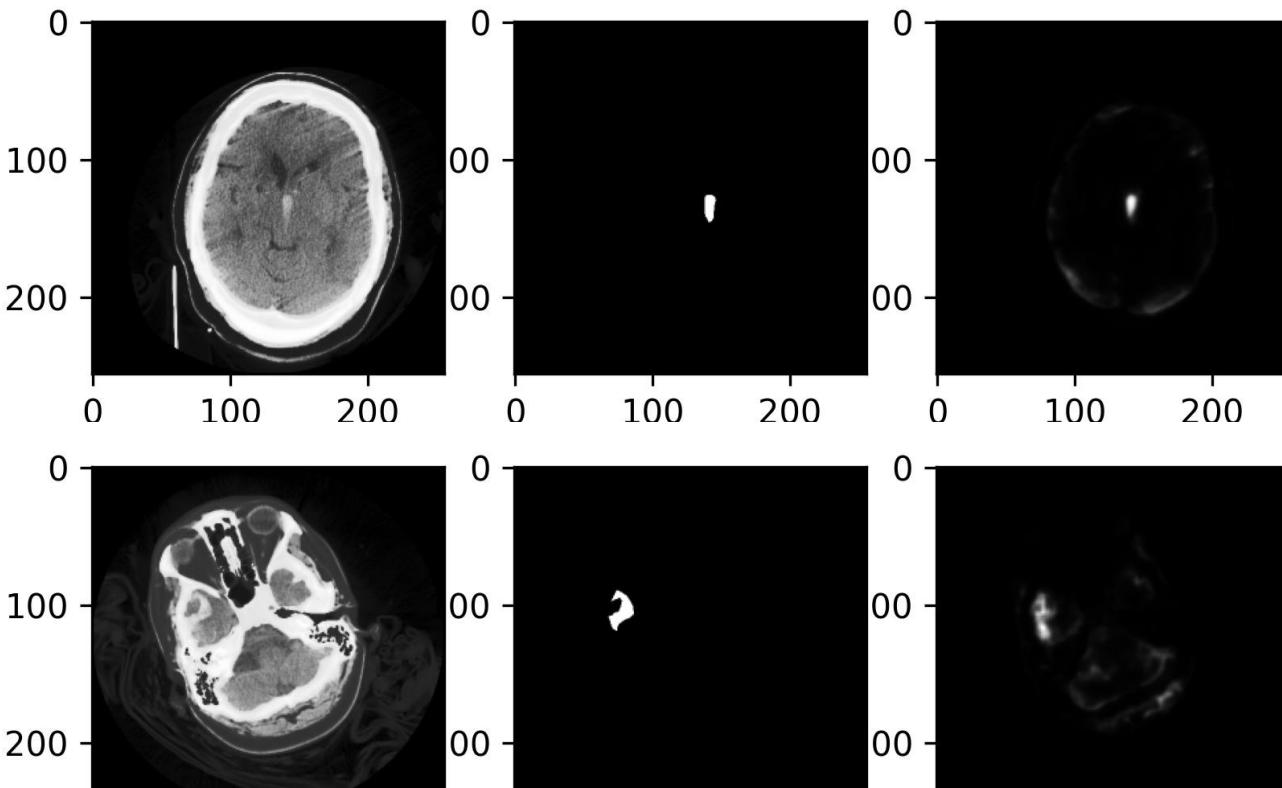


# Segmentation (Final Run)

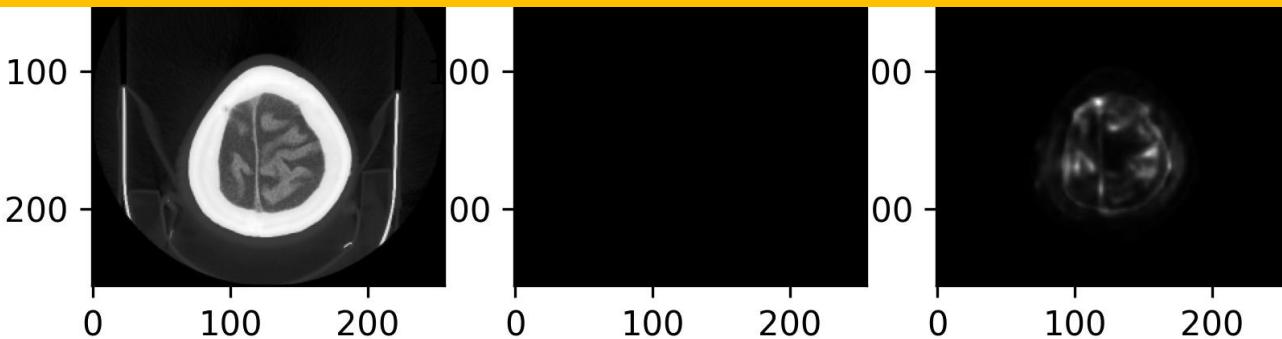
Dice Coefficient on Test Set: 0.3578  
Dice Loss on Test Set: 0.6422



# Challenges



- Dice coefficient did not show significant improvement despite dataset expansion and augmentation.
- Some additional images may lack the required quality for effective segmentation.
- Potential Unet Model Limitations.



# Discussion

- **Optimize Image Dataset:**
  - Improve dataset quality through enhanced preprocessing and augmentation.
  - Address biases and limitations within the dataset.
- **Explore Alternative Image Segmentation Models:**
  - Implement alternative models, such as Unet + ResNet architectures for precise segmentation.
  - Fine-tune parameters for optimal performance.
- **Simultaneous Classification and Segmentation:**
  - Develop a unified model for simultaneous classification and segmentation.
  - Explore multitasking approaches for efficiency and accuracy.