

Brain Tumor Segmentation and Classification Using ACGAN with U-Net and Independent CNN-Based Validation

Satish Wagle
Department of Computational and Data Science
Major Advisor: Dr. Khem Poudel, PhD

Introduction

- Brain tumors pose significant challenges in medical field, ranging from aggressive malignancies to gliomas
- Neuroimaging technique like Magnetic Resonance Imaging (MRI) provides high resolution images for diagnosis
- Manual tumor detection by the doctors and radiologist often expose to human error, Inconsistency, time consuming and limited detection
- These threats has always inspired the tumor detection and diagnosis to be automated using AI

Motivation and Objectives

Motivation:

- Lack of annotated data for training and validation of automated diagnostic systems
- The need for diverse and high-quality data

Objectives:

- **Synthetic Image Generation:** Using ACGAN to generate diverse synthetic images
- **Precise Segmentation:** Employing U-Net for accurate segmentation maps
- **Validation:** Using an independent CNN-based classifier to validate synthetic images

Previous Research and Works

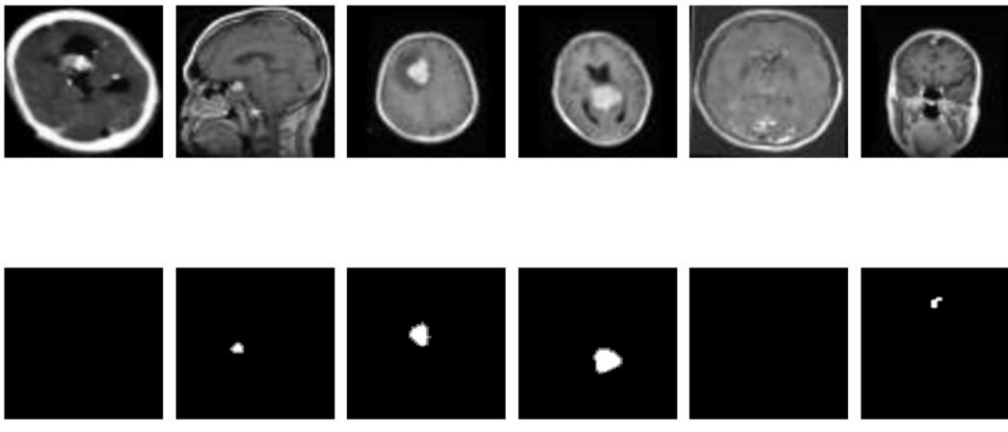
Table 1. Comparison of the model with other state-of-the-art models

Author	Approach	Result
Bowles et al., (2018) [10]	GAN Augmentation for Segmentation	DSC improved by 1-5%.
Han et al., (2018) [11]	DCGAN and WGAN for Synthetic Brain MR Images	Visual Turing Test: 54-77%.
Deepak et al., (2020) [12]	MSG-GAN for Brain MRI Synthesis	Balanced Accuracy: 90.3% to 93.1% with GAN; 86.4% to 88.7% with 35% data + GAN.
Ronneberger et al., (2015) [13]	U-Net for Biomedical Image Segmentation	DSC: 92% (PhC-U373), 77.5% (DIC-HeLa).
Isensee et al., (2018) [16]	nnU-Net	Mean DSC: 84.3%, Highest DSC: 90.5%.
Chen et al., (2021) [17]	Generative Adversarial U-Net	Mean DSC: 0.83, Highest DSC: 0.85.
Kazeminia et al., (2020) [18]	Pix2Pix with U-Net	DSC: 0.84 (Whole), 0.70 (Core), 0.65 (Enhancing); Sensitivity: 0.83 (Whole), 0.74 (Core), 0.72 (Enhancing).
Jiang et al., (2020) [19]	Conditional GAN for COVID-19 CT Image Synthesis	Mean DSC: 0.91.
Wang et al., (2020) [20]	Semi-supervised with ADC maps and U-Net	Balanced Accuracy: 90.3% without GAN, 93.1% with GAN, 86.4% with 35% data, 88.7% with 35% data + GAN.
Our Study	ACGAN with U-Net and CNN-Based Validation	ACGAN Classifier: Accuracy: 84%, U-Net: Accuracy: 99.54%, MeanIOU: 92.91%, Dice Coefficient: 76.43%, Precision: 99.61%, Sensitivity: 99.49%, Specificity: 99.87%..

Materials and Methods

Semantic Annotation for Image Segmentation:

- Total Dataset after Data augmentation in the real dataset: 5490
- 75% of the total dataset as Training sets: 4117
- Remaining 25% split equally as validation and test sets: 686
- Each MRI image in the dataset includes a segmentation mask with annotated and labeled tumor regions



Materials and Methods

- Four different classes of tumor as represented in Table 2:

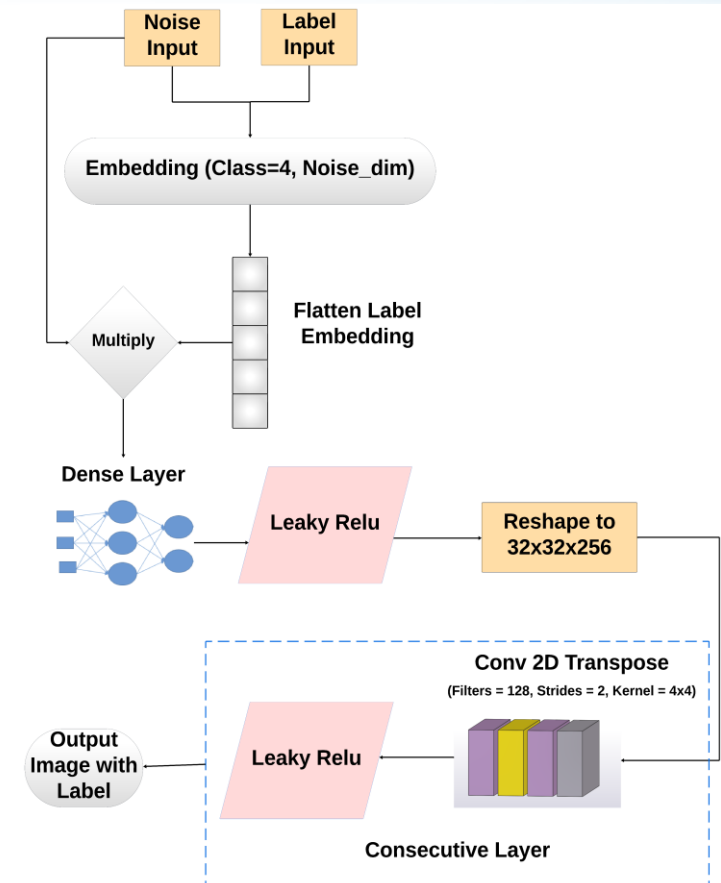
Table 2: Tumor Classes and respective Numerical Encoding

Tumor	Numerical Encoding
Non-Tumor	0
Glioma	1
Meningioma	2
Pituitary	3

Materials and Methods

CNN-Based Generator in ACGAN:

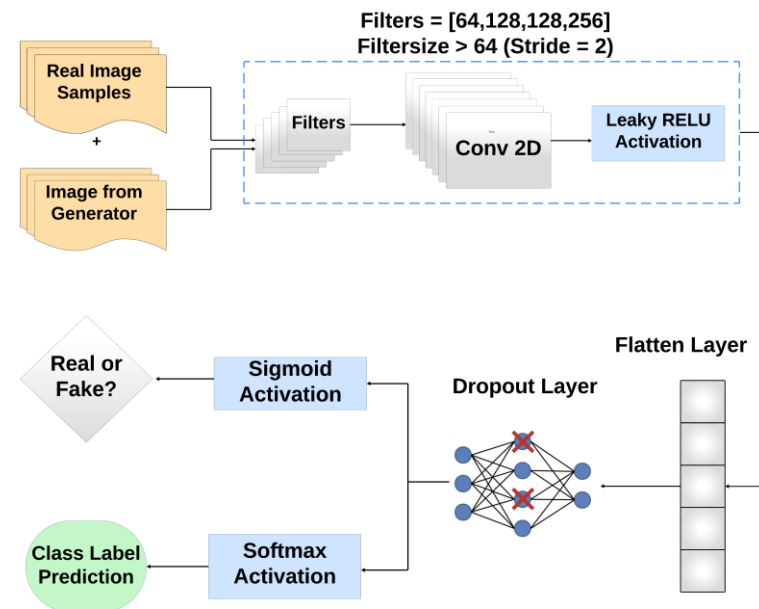
- Generator input: Noise , class label
- Embedding layer translates class label into a dense vector representation
- The dense vector or embedding captures the semantic meaning of the class allowing generator to generate images conditioned on the class label
- The flattened embeddings and noise vector are multiplied and passed to dense layer to expand the dimensionality of the input representation
- Dense layer size: $32 \times 32 \times 256$ with leaky RELU activation
- Reshaped to form 3D tensors, passed to conv2d Transpose to increase the spatial dimension to 128×128



Materials and Methods

CNN-Based Discriminator in ACGAN:

- Distinguishes real vs. synthetic data and performs classification.
- Batch of real and fake images fed during training.
- Images pass through convolutional layers to extract features.
- Each convolutional layer is accompanied by LeakyReLU activation to introduce nonlinearity in the model
- Final conv2d layer output is flattened to a 1D vector.
- 1D vector processed through a dense layer with 0.4 dropout to prevent overfitting.
- Dense layer output passes through a layer with 1 neuron (sigmoid) and 4 neurons (SoftMax).



Materials and Methods

Activation Function:

- Uses Leaky RELU activation in both generator and discriminator
- Introduces nonlinearity in the model to learn complex pattern by approximating continuous function
- Overcomes dying RELU problem by introducing small nonzero gradient when input is negative

$$\text{LeakyReLU}(x) = \begin{cases} x & \text{if } x > 0 \\ \alpha x & \text{otherwise} \end{cases}$$

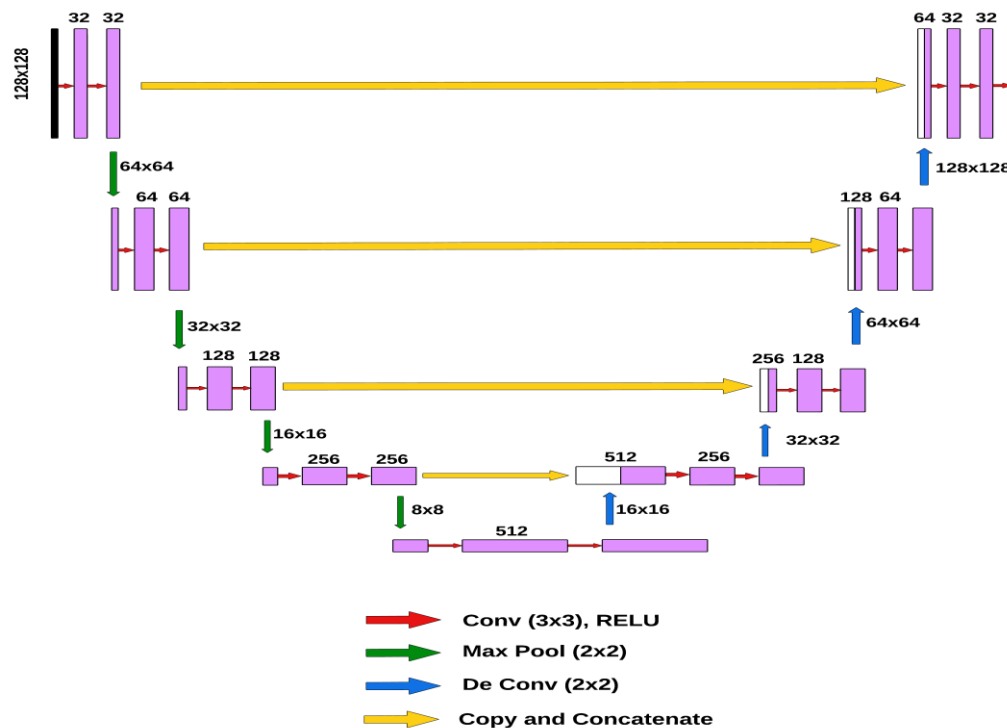
Loss Function:

- Adversarial loss : Binary cross entropy
- Classification loss: Categorical cross entropy

Materials and Methods

U-Net Architecture for Brain Tumor Segmentation:

- Designed to segment the tumor region in the MRI



Materials and Methods

Segmentation performance evaluation metrics:

- Both Dice coefficient and IOU measure the overlap between predicted and ground truth masks in image segmentation
- They differ in their calculation and sensitivity to the overlap region

Dice Coefficient:

- Formula: $2 \times \frac{\text{Area of Overlap}}{\text{Area of Predicted Mask} + \text{Area of Ground Truth Mask}}$
- More sensitive to overlap size.
- Higher scores for larger overlaps.

Intersection over Union (IoU):

- Formula: $\frac{\text{Area of Overlap}}{\text{Area of Union}}$
- Less sensitive to overlap size.
- Focuses on overall alignment.

Materials and Methods

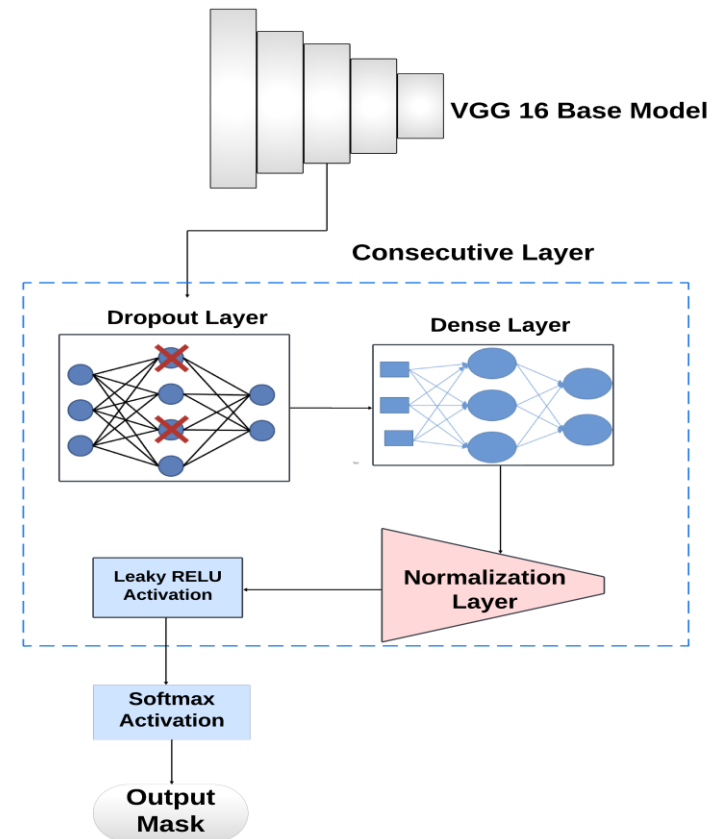
CNN-based Classifier for Evaluating Synthetic Image :

- Model first trained on real data
- The saved model further used to classify the real and synthetic data combined
- Comparison made between two classification report
- If the classifier performs similarly on both real and synthetic images, it indicates that the synthetic images have successfully captured the key features of the real images

Materials and Methods

CNN based Classifier

- VGG16 used as a base model
- Input image of shape 128,128,3
- Two consecutive layer of dropout, dense, normalization and activation
- Final dense layer with 4 neurons and SoftMax activation assuming four classes of classification
- Sparse categorical cross entropy (loss function)



Results

Realism and Accuracy of Synthetic images

- Accessed through CNN-based classifier
- Classification on Real Images
- Classification on Real and Synthetic Images combined
- Classification report compared based on Sensitivity, Specificity, Precision, Recall and F1 Score
- Sensitivity:
$$\frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Negatives (FN)}}$$
- Specificity:
$$\frac{\text{True Negatives (TN)}}{\text{True Negatives (TN)} + \text{False Positives (FP)}}$$

Results

Classification with Real VS combined Real and Synthetic images

Classification with Real Images

- Overall Accuracy: 0.98

Table 3. Classification Report for Real Images

Class	Sensitivity	Specificity	Precision	Recall	F1 Score
Healthy	0.99	0.998	0.99	0.99	0.99
Glioma	0.96	0.99	0.97	0.96	0.96
Meningitis	0.97	0.98	0.96	0.97	0.96
Pituitary	0.99	0.996	0.99	0.99	0.99

Classification with combined Real and Synthetic Images

- Performed over 5-fold cross validation
- Overall Accuracy: 0.84

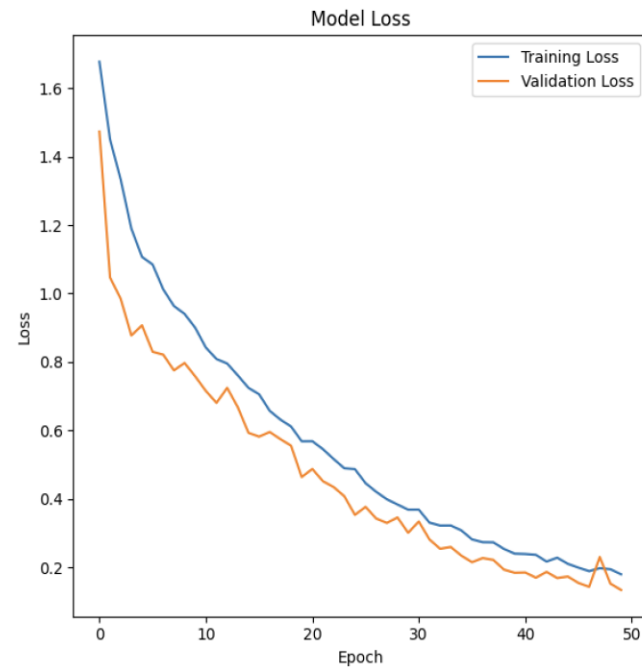
Table 4. Classification Report for Combined Real and Synthetic Images

Class	Sensitivity	Specificity	Precision	Recall	F1 Score
Healthy	0.77	0.92	0.73	0.77	0.75
Glioma	0.81	0.99	0.95	0.81	0.87
Meningitis	0.93	0.93	0.80	0.93	0.86
Pituitary	0.82	0.97	0.89	0.82	0.85

Results

Classifier Training Process

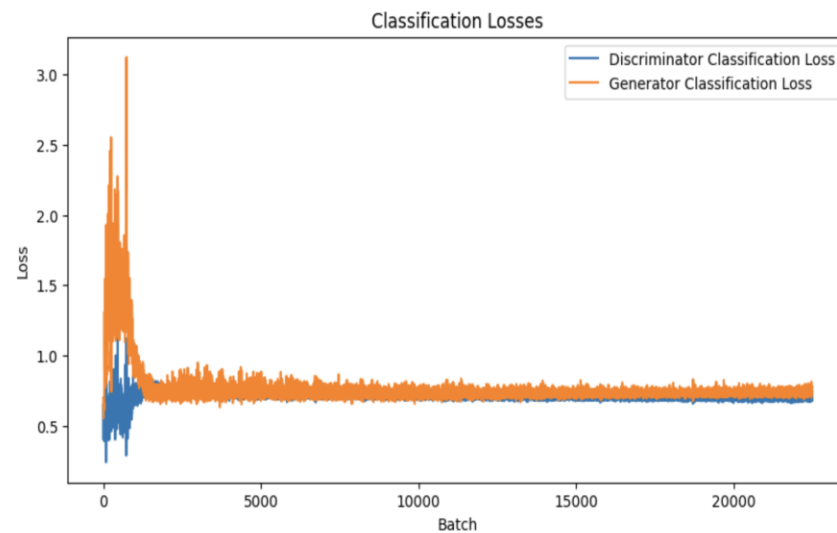
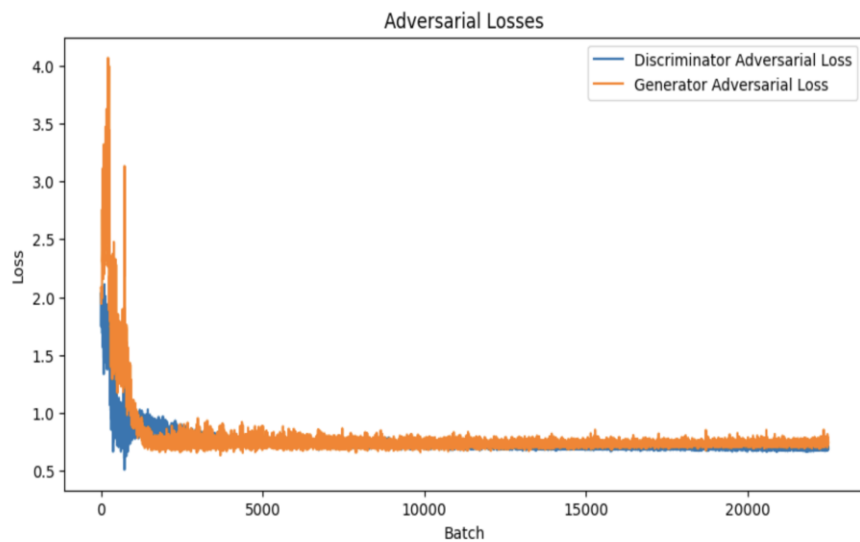
- Accessed through Loss and Accuracy graph



Results

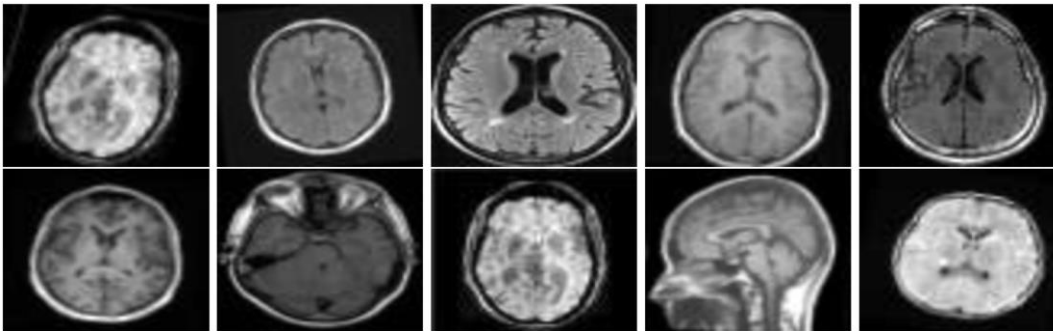
ACGAN Training Process

- Monitored for generator and discriminator for both adversarial and classification losses over the epoch

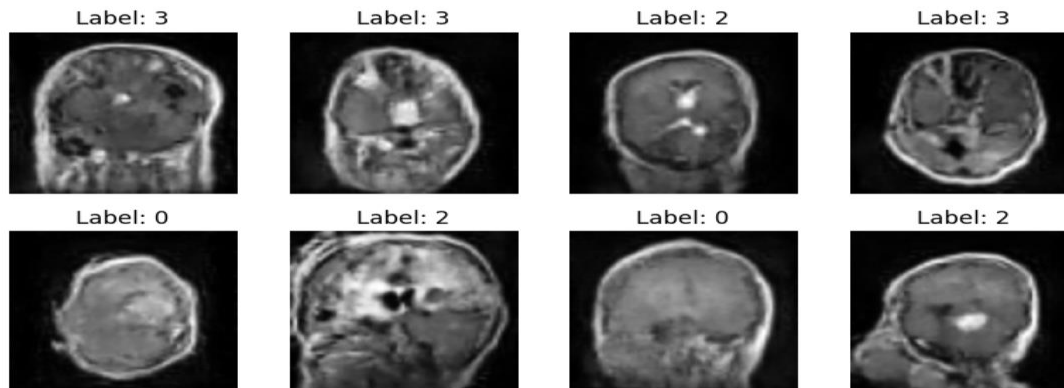


Results

Sample Real Images



Sample Synthetic Images



Results

Segmentation Accuracy

- Dice Coefficient and Intersection over Union (IoU) metrics were computed to evaluate the accuracy of the generated segmentation maps
- Beside Dice Coefficient and IoU, the model's performance on the test set was evaluated using several critical metrics

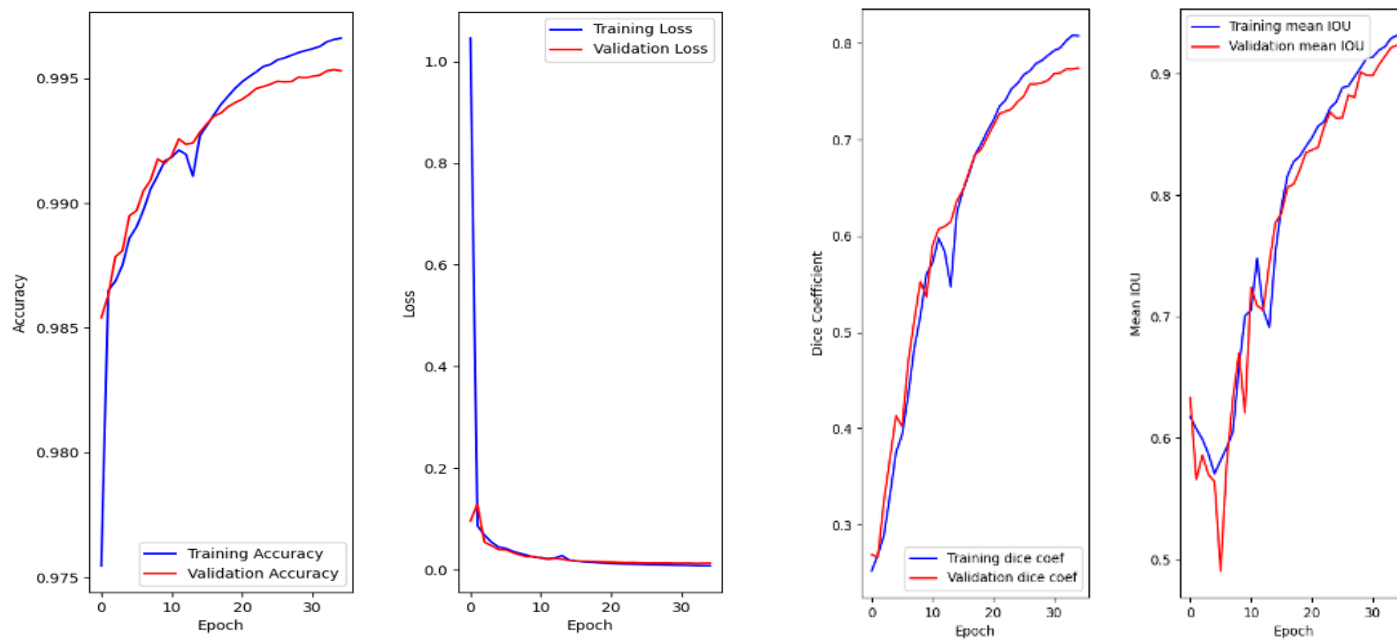
Table 5. Performance Metrics for U-Net Model on Test Set

Metric	Value
Loss	0.0142
Accuracy	99.54%
Mean Intersection over Union (MeanIOU)	92.91%
Dice Coefficient	76.43%
Precision	99.61%
Sensitivity (Recall)	99.49%
Specificity	99.87%

Results

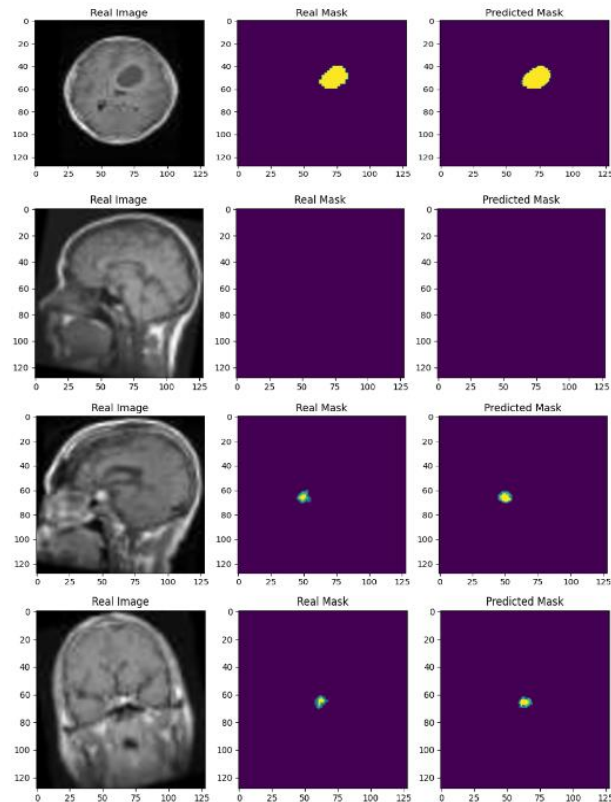
Training process evaluation of U-Net architecture

- Evaluated in terms of Accuracy, Loss, Dice Coefficient and MeanIOU for training and validation sets over the epochs



Results

Comparison of the predicted segmentation over ground truth for multiple tumor class



Discussion and Conclusion

- ACGAN generates satisfactory images compared to the real samples
- The quality of the synthetic images can be further improved to resemble the real sample by considering MRI views within the dataset
- Classifier accuracy on combined real and synthetic image: 0.84
- U-Net Segmentation shows good result
- Dice Coefficient: 76.43%, MeanIOU: 92.91%, Sensitivity: 99.49%, Specificity: 99.87%.
- Future Directions may include Dual Classification by incorporating both MRI views and tumor types, Multi-Modality Image synthesis, Semi-Supervised/Unsupervised Learning by utilizing unlabeled data and exploring different GANs for enhanced performance.

Acknowledgement

- Department of Computational and Data Science
- Major advisor Dr. Khem Poudel, PhD
- AIBigData research team members

References

- <https://www.kaggle.com/code/harshsingh2209/generating-brain-mri-images-with-dc-gan>
- <https://www.kaggle.com/code/abdallahwagih/brain-tumor-segmentation-unet-dice-coef-89-6>
- OpenAI. (2024). ChatGPT (August 6 version) [Large language model]. <https://www.openai.com/>