A REPORT ON

SEMANTIC SEGMENTATION OF THz IMAGES USING AI – BASED METHODS

BY

Mohit Singh

120AD0001, B. TECH 4th YEAR STUDENT, Artificial Intelligence and Data Science,

INDIAN INSTITUTE OF INFORMATION TECHNOLOGY, DESIGN & MANUFACTURING, KURNOOL, ANDHRA PRADESH

AT

CEERI CHENNAI

CSIR Madras Complex, Tharamani, Chennai

NDIAN INSTITUTE OF INFORMATION TECHNOLOGY, DESIGN & MANUFACTURING, KURNOOL

(ANDHRA PRADESH)

ACKNOWLEDGEMENT

I am grateful to Dr. P C Panchariya, Director (CSIR-CEERI) for giving me this opportunity to work as an intern under the guidance of renowned people in the field of AI.

I would also like to express my endless thanks to our mentor Dr. Mercy Latha, Principal scientist at CSIR-CEERI, Chennai for her sincere and dedicated guidance throughout the project development.

Also, our hearty gratitude to our faculty mentor Dr. Eswaramoorthy K. V., Assistant Professor (Grade-I) at IIITDM Kurnool, for giving us encouragement and technical support on the project.

MOHIT SINGH

INDIAN INSTITUTE OF INFORMATION TECHNOLOGY, DESIGN & MANUFACTURING, KURNOOL

(ANDHRA PRADESH)

Station: CEERI Chennai
Duration: 8 weeks

Centre: CEERI Chennai
Date of Start: 12th June 2023

Date of Submission: 31st July 2023

Title of the Project: SEMANTIC SEGMENTATION OF THz IMAGES USING AI – BASED METHODS

Name of the Student	ID. No.	Discipline
Mohit Singh	**120AD0001**	**AI&DS**

Name(s) and designation(s) of the expert(s): **Dr. Mercy Latha**, Principal Scientist at CSIR-CEERI

Name(s) of the PS Faculty: Dr. Eswaramoorthy K. V., Assistant Professor (Grade-I) at IIITDM Kurnool

Key Words: **THz images, CNN, UNET, VGG-19, Transfer Learning**

Project Areas: **THz Imaging, Semantic Segmentation, Transfer Learning**

Signature of Student

Date: 31-07-2023

Signature of PS Faculty

Date: 31-07-2023

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

THz (THz) imaging technology is an evolving area of research. The potential application of THz imaging ranges from biomedical diagnosis to quality control. THz radiation is electromagnetic waves corresponding to the transition region from far infrared to microwave waves. This gap between infrared waves and microwaves is known as the “THz gap” the “gap” is used to tell that the application of THz is still in infancy. This region corresponds to 0.3 to 3 THz (1012 hertz).

THz radiation can penetrate thin layers of materials but are blocked by thicker objects. Thus, they can penetrate through skin or bags and detect thicker materials such as tissues or metals. They are non-ionizing and of lower energy than X-rays which are generally used for imaging. Therefore, THz radiation is seen as a plausible alternative to X-rays. The following report shows the application of THz imaging for threat identification. The current security scans at airports utilize radio wave type or X-ray type screening systems. Radio waves fail to detect objects at a large distance and have low penetrating power, while X-rays are ionizing and can cause harm to the object being scanned. THz scanning systems are therefore seen as a better alternative to the other two.

In this study, we present a comprehensive investigation into tumor semantic segmentation using deep learning techniques, initially applied to magnetic resonance imaging (MRI) of the brain, and later extended to the domain of Terahertz (THz) images.

The accurate segmentation of brain tumors in MRI is of paramount importance due to its relevance in neurosurgery and oncology. For this purpose, highly effective U-Net architecture is employed, which has demonstrated remarkable performance in medical image segmentation tasks. By leveraging transfer learning with pre-trained VGG-16 and VGG-19 models.

Building upon the success of the approach in MRI-based brain tumor segmentation, research is expanded to explore the application of deep learning techniques in Terahertz medical imaging. Terahertz imaging offers unique advantages, including non-invasiveness and the ability to visualize biological tissues without ionizing radiation. However, it also poses specific challenges, such as noise characteristics and the scarcity of annotated datasets.

2. METHODOLOGY

The main objectives of this research were twofold: firstly, to establish a reliable and accurate tumor segmentation framework for brain MRI, and secondly, to extend this successful methodology to Terahertz images, thereby advancing the field of THz medical imaging for tumor detection and characterization.

Following methodology divided into sections is used to build a robust model of semantic segmentation in both approaches:

Dataset Acquisition

Data Preprocessing

Model Architecture - U-Net

Transfer Learning with VGG-16 and VGG-19

Model Training

Model Evaluation

Visualizing Segmentation Results

2.1 SEMANTIC SEGMENTATION OF MRI IMAGES

Semantic segmentation is a computer vision task that involves dividing an image into meaningful and semantically coherent regions, where each pixel is assigned a specific class label. In the context of MRI (Magnetic Resonance Imaging) images, semantic segmentation is used to automatically identify and segment different structures or regions of interest within the images, such as organs, tissues, lesions, or abnormalities.

2.1.1 DATASET ACQUISITION

For building a segmentation model, BRATS dataset (BraTS) is used to acquire Brain MRI Images with ground truth masks.

The BRATS dataset (BraTS) stands for the Multimodal Brain Tumor Segmentation Challenge. It is a widely used dataset in the medical imaging and machine learning research community, specifically for tasks related to brain tumor segmentation. The dataset contains magnetic resonance imaging (MRI) scans of the brain with different modalities, such as T1-weighted, T2-weighted, fluid-attenuated inversion recovery (FLAIR), and post-contrast T1-weighted images.

The main objective of the BRATS challenge is to develop algorithms that can automatically segment and classify different types of brain tumors, including gliomas, into their respective regions in MRI scans. This segmentation process is crucial for diagnosis, treatment planning, and assessing the progression of brain tumors.

Anyone can get access to BraTS dataset from following links:

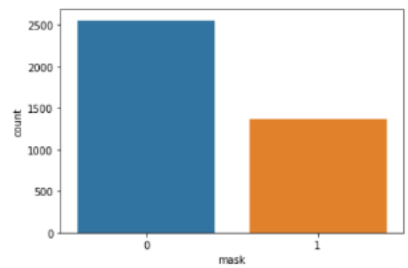
https://www.med.upenn.edu/cbica/brats2020/data.html

https://www.kaggle.com/datasets/awsaf49/brats2020-training-data

2.1.2 DATA PREPROCESSING

BraTS dataset consists of total 3929 instances of brain MRI images with ground truth masks, for the segmentation task we will be using 1373 instances of images with class 1(brain MRI images with tumors)

```
0    2556
1    1373
Name: mask, dtype: int64
```



` Figure 1

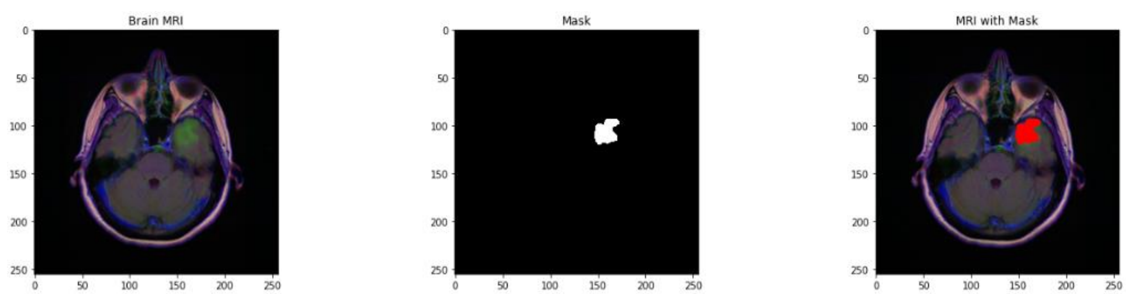


Figure 2

In preparation for the segmentation of brain MRI images, a crucial preprocessing pipeline has been implemented to ensure optimal results. The first step involved resizing the original images to a standardized resolution, effectively harmonizing the data and facilitating uniformity across the dataset. This process helps prevent any potential discrepancies arising from variations in image dimensions. Subsequently, the images were standardized, transforming their pixel intensity values to adhere to a common scale. Standardization eliminates the influence of differing brightness and contrast levels, thereby enhancing the performance of segmentation algorithms. By carefully applying resizing and standardization techniques, the preprocessed images are now ready for accurate and reliable brain MRI segmentation, a fundamental task in medical image analysis that can aid in the diagnosis and treatment of neurological conditions.

2.1.3 MODEL ARCHITECTURE

VGG – 16 UNet

VGG-16 UNet is a convolutional neural network architecture that combines the popular VGG-16 and the UNet models. VGG-16 is a deep convolutional network introduced by the Visual Geometry Group (VGG) at the University of Oxford. It is primarily designed for image classification tasks and is known for its simple architecture with 16 layers, consisting of 13 convolutional layers followed by 3 fully connected layers. The UNet architecture, on the other hand, is widely used for image segmentation tasks. It features a U-shaped encoder-decoder design, where the encoder captures the spatial features of the input image, and the decoder reconstructs the segmentation map from the encoded features.

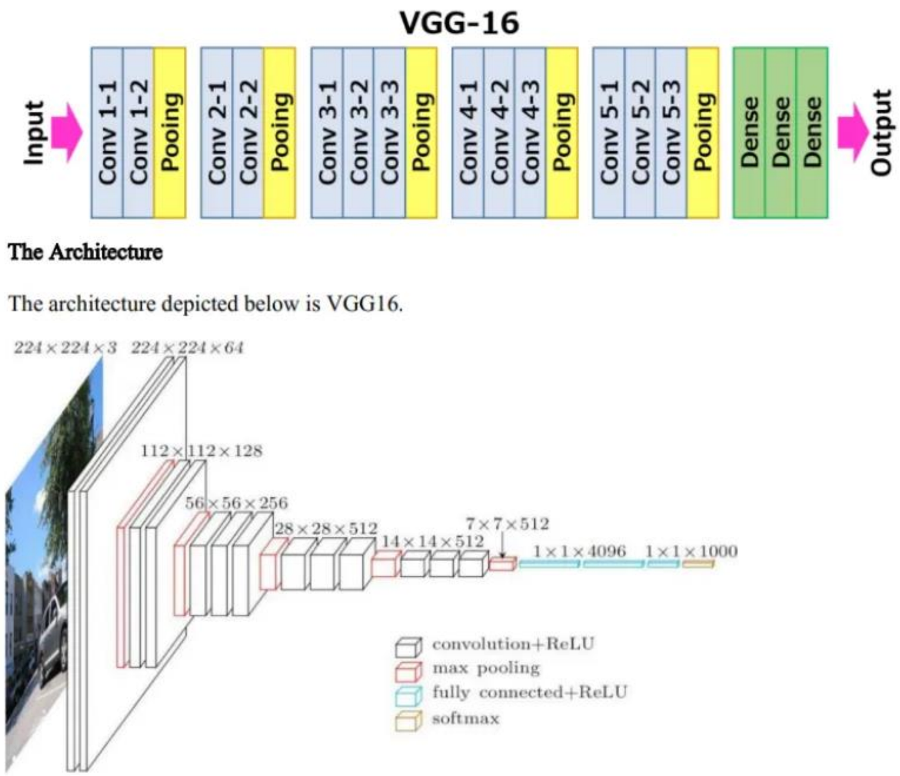


Figure 3

VGG – 19 UNet

Like VGG-16 UNet, VGG-19 UNet is a fusion of the VGG-19 and UNet architectures. VGG-19 is an extension of VGG-16, incorporating 19 layers, which makes it deeper and potentially more powerful for image representation.

The UNet component in VGG-19 UNet helps in preserving finer details during the segmentation process, which can be crucial for tasks like medical image segmentation or fine-grained object localization.

Overall, VGG-19 UNet is suitable for more challenging segmentation tasks that require a combination of high-level feature extraction and precise localization, thanks to the deep VGG-19 encoder and the U-shaped UNet decoder.

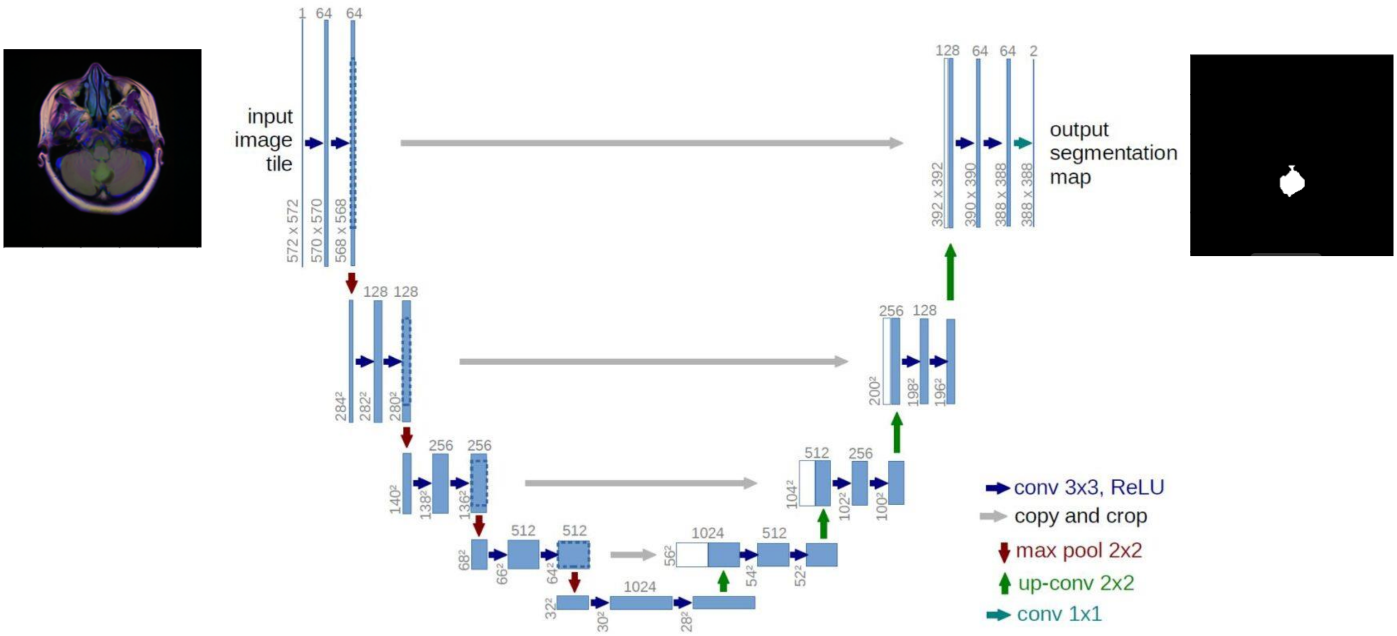


Figure 4

2.1.4 TRANSFER LEARNING WITH VGG-16 AND VGG-19

Transfer learning is a machine learning technique that involves leveraging knowledge gained from training a model on one task and applying it to a different but related task. In the context of computer vision, pre-trained convolutional neural networks (CNNs) like VGG-16 and VGG-19 have been widely used for transfer learning.

To perform transfer learning with VGG-16 or VGG-19 on an image segmentation task like semantic segmentation, the typical approach is to use the pre-trained VGG model as a feature extractor. The convolutional layers of the pre-trained VGG model are frozen, meaning their weights are not updated during training, so the learned features are preserved.

The training process involves optimizing the decoder part to adapt the model to the specific segmentation task using a labeled dataset of segmented images. The loss function used during training typically measures the difference between the predicted segmentation maps and the ground truth segmentation maps.

By employing transfer learning in this way, the model benefits from the powerful feature extraction capabilities of VGG-16 or VGG-19 while tailoring the decoder part to the specific segmentation task, even with a smaller dataset. This approach can lead to faster convergence and better performance compared to training a segmentation model from scratch.

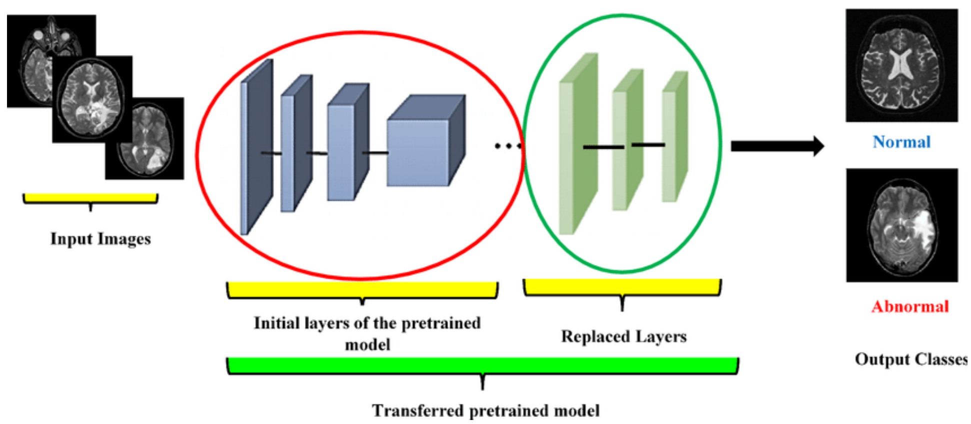


Figure 5

2.1.4 MODEL TRAINING

The VGG-16 and VGG-19 models were trained using the Adam optimizer, a popular optimization algorithm in deep learning, along with Intersection over Union (IOU) as the evaluation metric.

Adam Optimizer: Adam (Adaptive Moment Estimation) is an optimization algorithm that combines the benefits of both the AdaGrad and RMSProp optimizers. It is widely used due to its ability to adapt the learning rates of individual parameters based on their past gradients.

IOU Metric: IOU, short for Intersection over Union, is a performance metric commonly used in tasks like object detection and semantic segmentation. It measures the overlap between the predicted bounding boxes or segmentation masks and the ground truth. The IOU value is computed by dividing the area of intersection between the predicted and ground truth regions by the area of their union. Higher IOU values indicate better model performance.

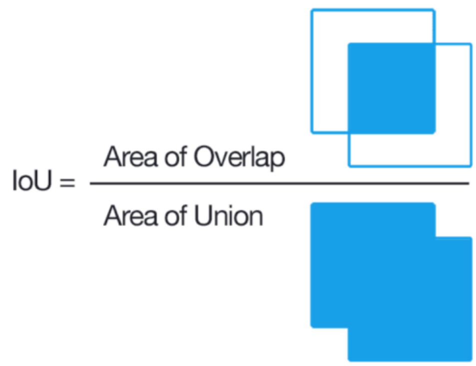


Figure 6

Other than that, while fitting the model with training data, callbacks like ReduceLROnPlateau and ModelCheckpoint were used to get a good IoU score from the segmentation model.

ReduceLROnPlateau is a learning rate scheduling technique commonly used during training to dynamically adjust the learning rate based on the model's performance. The "plateau" refers to a situation where the model's validation loss or a specified metric has stopped improving. When this happens, the learning rate is reduced, allowing the optimization process to take smaller steps and hopefully find a better minimum in the loss landscape.

ModelCheckpoint callback is used in conjunction with training using model.fit() to save a model or weights (in a checkpoint file) at some interval, so the model or weights can be loaded later to continue the training from the state saved.

2.1.5 MODEL EVALUATION

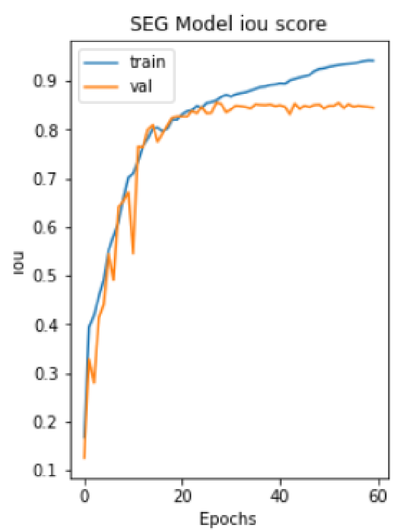
Both VGG – 16 and VGG – 19 were trained on many different resolutions of brain MRI images. The aim was to explore their capabilities in accurately identifying and segmenting tumors of different sizes and complexities. The results indicate that both models exhibit strong performance across a range of image resolutions, demonstrating their robustness and efficacy in tumor detection. The successful application of these models underscores their potential as valuable tools for medical professionals, aiding in early and precise tumor diagnosis, ultimately leading to improved patient outcomes.

The below given table shows the Mean IoU scores obtained on the Validation images after applying various resolutions to the training images.

Model	32x32	64x64	96x96	256x256
VGG 16	0.78	0.80	0.83	0.84
VGG 19	0.80	0.82	0.84	0.845

Table 1

IoU score vs Epochs for VGG-19 model trained on images resized to 256x256 with validation IoU of 84.5



IoU score vs epoch after training the model for 60 epochs

2.1.6 VISUALIZING SEGMENTATION RESULTS

Segmentation results are visualized for the brain MRI images that are in the validation set to analyze models' performance on segmenting tumors in brain MRI images.

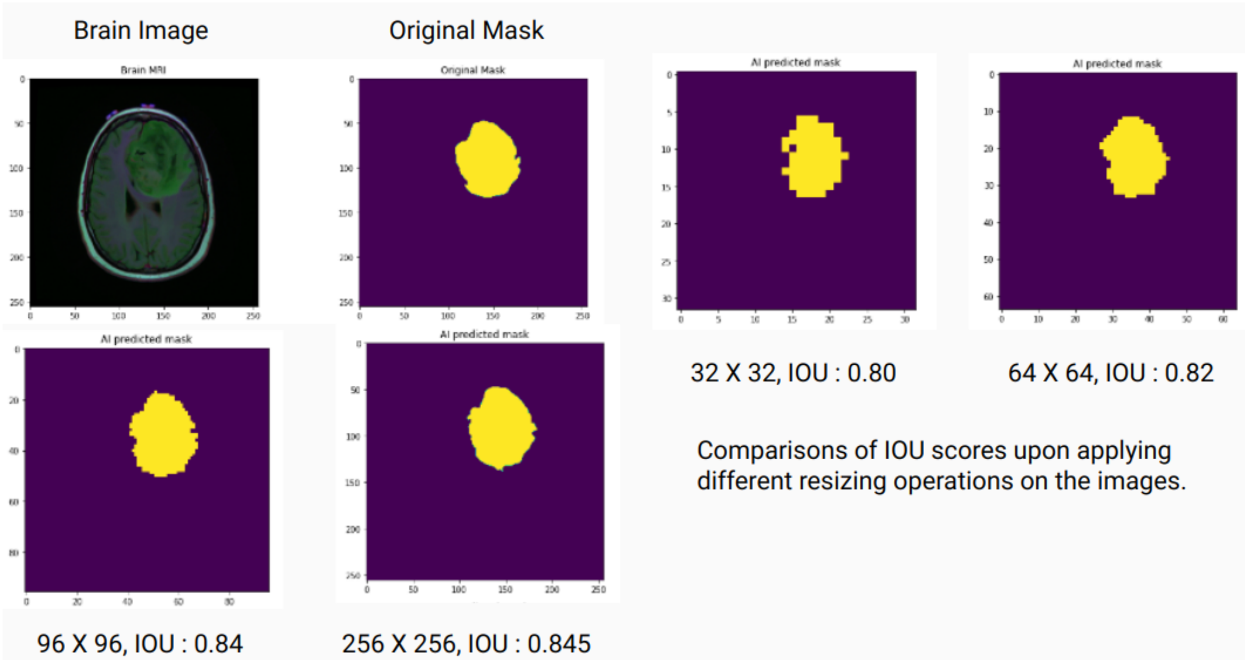


Figure 7

2.2 SEMANTIC SEGMENTATION FOR TERAHERTZ IMAGES

Building upon the results of brain MRI images segmentation, the present report embarks on a new frontier by exploring the applicability of VGG-19 in the segmentation of tumors in terahertz images. Terahertz imaging is a cutting-edge medical imaging modality that offers unique insights into tissue structures and characteristics, making it a promising avenue for diagnosing tumors. However, the precise localization and segmentation of tumors in terahertz images pose significant challenges due to their complex nature and subtle visual features.

The main objective of this study is to assess the effectiveness of using the VGG-19 model for terahertz brain tumor segmentation. The approach involves employing transfer learning, where the VGG-19 model is fine-tuned. The goal is to leverage the knowledge gained from analyzing MRI images and apply it to improve tumor segmentation in the terahertz domain, thus achieving more robust results.

This report details the methodology, data acquisition, and pre-processing techniques adopted for the terahertz dataset. Additionally, the fine-tuning process on the VGG-19 model is elaborated and the specific modifications made to adapt the model for terahertz image segmentation is discussed. Furthermore, a comprehensive evaluation of the model's performance is given, comparing the results obtained from VGG-19.

2.2.1 DATA ACQUISITION

A total of 41 terahertz images were received from CEERI, Chennai. Out of those 41 images, 35 were images of tissues with cancer and 6 images were normal. There were further variations in the 35 tumor images as some were raw images and some of them were adjusted images (images with preprocessing applied).

The 35 images which had tumor present were used for the splitting into training and validation split. The split carried out was 80:20(training: validation).

The images had regions marked with different colors in RGB format to represent different parts like Tumor Area, Tumor Growing Area and Background. These ranges were used to create thresholds for creating masks for segmentation model training.

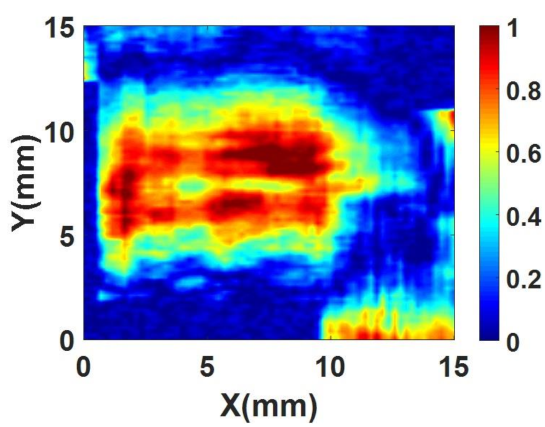


Figure 8

A two-step approach is adapted to accurately delineate tumor regions within medical images. Initially, a segmentation model is trained with images and masks with two classes: tumor and background. This initial model was essential for identifying the core tumor region, enabling the differentiation between the tumor and the surrounding healthy tissues. By doing so, the model successfully created a binary mask highlighting the tumor area while excluding the background.

However, to improve the precision of the tumor segmentation and gain a more comprehensive understanding of the tumor's growth pattern, the approach will be extended to incorporate three classes. The additional class represented the tumor growing area, which denotes the region around the core tumor that exhibits signs of progression or infiltration into adjacent healthy tissues.

2.2.2 SEGMENTATION MODEL FOR 2 CLASSES

2.2.2.1 DATA PREPROCESSING

In preparation for the segmentation of Terahertz images, a crucial preprocessing pipeline has been implemented to ensure optimal results. The first step involved resizing the original images to a standardized resolution, effectively harmonizing the data and facilitating uniformity across the dataset. This process helps prevent any potential discrepancies arising from variations in image dimensions.

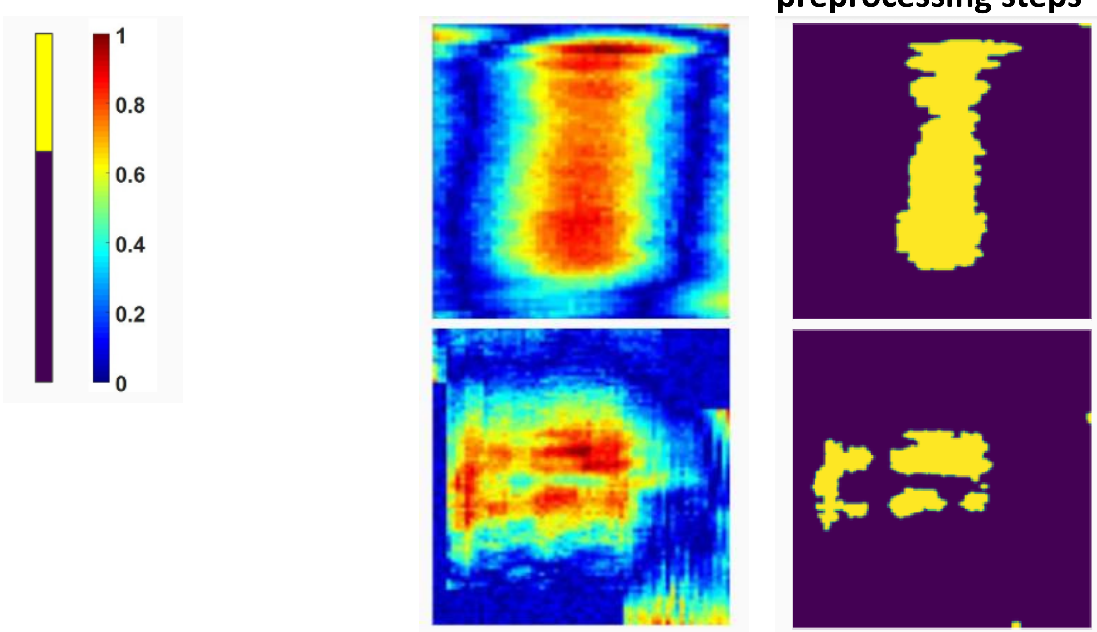
Subsequently, the images were standardized, transforming their pixel intensity values to adhere to a common scale. Standardization eliminates the influence of differing brightness and contrast levels, thereby enhancing the performance of segmentation algorithms. By carefully applying resizing and standardization techniques, the preprocessed images are now ready for accurate and reliable brain MRI segmentation, a fundamental task in medical image analysis that can aid in the diagnosis and treatment of neurological conditions.

Images may contain noise, which can negatively impact segmentation accuracy. The common technique for noise reduction Gaussian blurring was applied to the terahertz images.

Mask Preprocessing steps were applied to get accurate ground truths for the corresponding tumor images. Mask preprocessing steps were as follow:

- Load the terahertz image
- Convert the image to the HSV color space
- Define the color thresholds for tumor and non-tumor regions
- Create the initial tumor mask based on the color thresholds
- Perform region filling to include missing parts in the tumor class
- Perform morphological operations to refine the mask

Thresholding Regions **Terahertz Images** **Masks produced after applying Mask preprocessing steps**



2.2.2.2 MODEL TRAINING

After analyzing the performance of VGG – 19 trained on “ImageNet” weights on the brain MRI images, the transfer learning VGG – 19 model is used on the Terahertz images as the segmentation results were accurate and benefit of the same model can be taken in segmentation of Terahertz images.

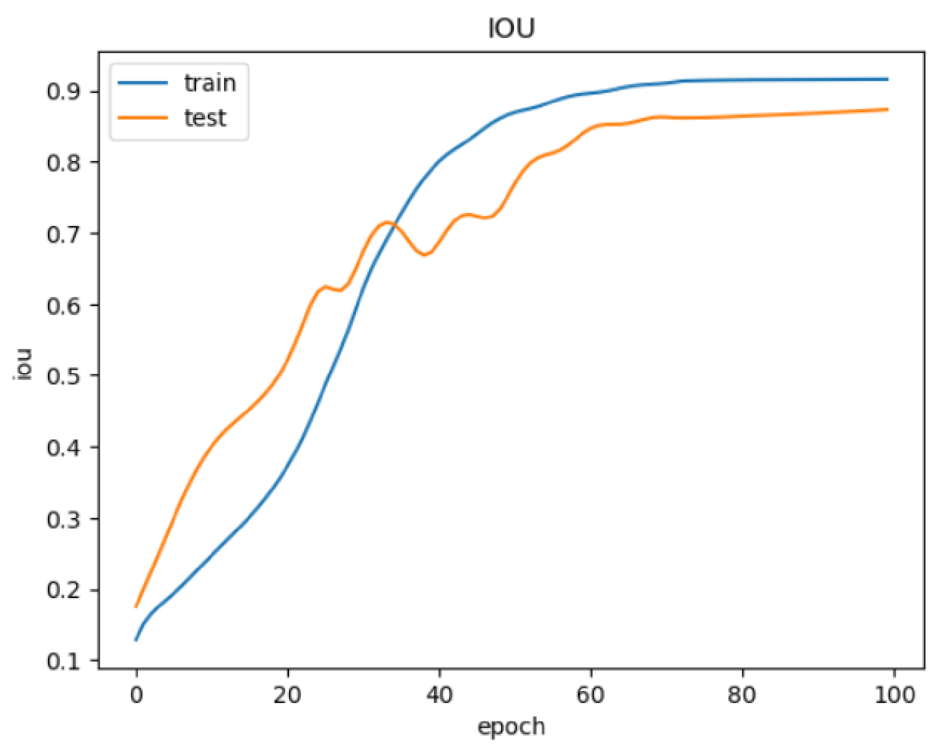
The optimizer used is Adam and callbacks used are ReduceLROnPlateau and ModelCheckpoint as they were used in segmentation model for brain MRI images. The same metrics (IoU score) is used for evaluating the segmentation results.

2.2.2.3 MODEL EVALUATION

The accuracies and IoU scores for training and validation sets after fitting the model are:

Training Accuracy	Validation Accuracy	Training IoU Score	Validation IoU Score
95.7%	92.3%	0.9160	0.8732

Table 2



IoU score vs epoch after training the model for 100 epochs

2.2.2.4 VISUALIZING SEGMENTATION RESULTS

Classwise IoU scores are introduced to get a better understanding of prediction of each class
A new metric, **The Dice coefficient** is introduced while visualizing the segmentation results to get a better overview for the segmentation results.
Dice coefficient also known as the Sørensen–Dice coefficient, is a metric commonly used to measure the similarity or overlap between two sets.
The Dice coefficient is calculated using the formula:

Dice coefficient = (2 * |A ∩ B|) / (|A| + |B|)
The Dice coefficient ranges from 0 to 1, where 0 indicates no overlap between the sets, and 1 indicates perfect overlap or complete similarity between the sets.

The Dice coefficient and Intersection over Union (IoU), also known as the Jaccard index, are closely related metrics used for evaluating the similarity or overlap between two sets, especially in the context of image segmentation, object detection, and binary classification tasks.
The Dice coefficient and IoU are related through the following formula:
Dice coefficient = 2 * IoU / (IoU + 1)


```
Image 1 :
PixelWise Accuracy : 0.99368286
{'IoU Background': 0.9934811362191617, 'IoU Tumor': 0.830466830126754}
Mean IoU : 0.91197395
Dice coefficient : 0.9520562299182095
```


```
Image 5 :
PixelWise Accuracy : 0.996994
{'IoU Background': 0.9969587501389873, 'IoU Tumor': 0.7941483795254458}
Mean IoU : 0.8955536
Dice coefficient : 0.9418710282051579
```


```
Image 7 :
PixelWise Accuracy : 0.9900665
{'IoU Background': 0.9866726717917851, 'IoU Tumor': 0.9624567473493393}
Mean IoU : 0.97456473
Dice coefficient : 0.987080446089961
```

2.2.3 SEGMENTATION MODEL FOR 3 CLASSES

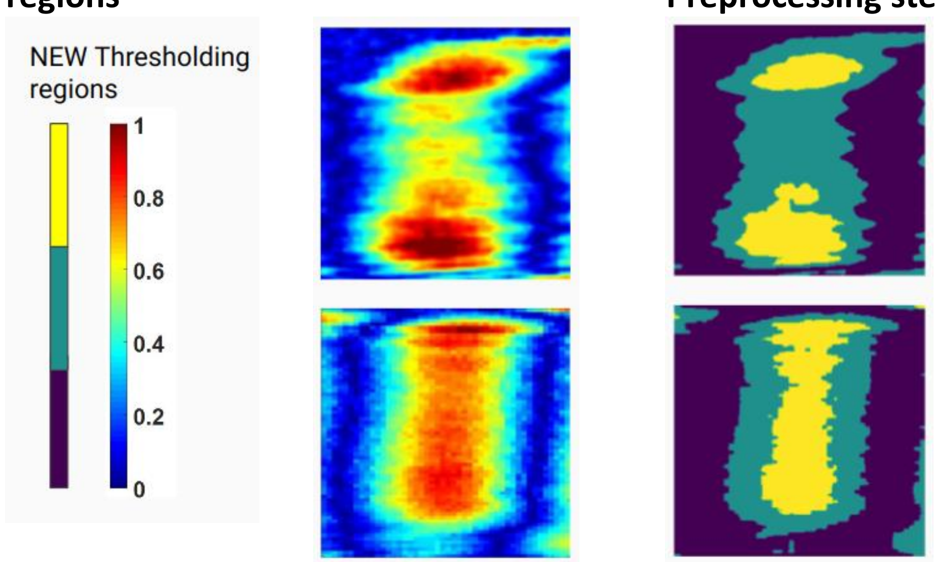
2.2.3.1 DATA PREPROCESSING

Similar preprocessing steps that were used for preprocessing the images in segmentation model for 2 classes are employed here.

Mask Preprocessing steps were applied to get accurate ground truths for the corresponding tumor images. Mask preprocessing steps were as follow:

- Load the terahertz image
- Convert the image to the HSV color space
- Define the color thresholds for tumor, non-tumor, and tumor growing area (3 classes)
- Create the initial tumor mask based on the new color thresholds
- Perform region filling to include missing parts in the tumor class
- Perform morphological operations to refine the mask
- Create the final mask with three classes (0, 1, 2)
- Apply One Hot Encoding to change the number of channels of the masks from 1 to 3(each channel will have binary representations of each class)

New Thresholding regions | **Terahertz Images** | **Masks produced after applying Mask Preprocessing steps**



2.2.3.2 MODEL TRAINING

After analyzing the performance of VGG – 19 trained on "ImageNet" weights on the segmentation model with 2 classes, the transfer learning VGG – 19 model can be used on the Terahertz images for 3 classes as the segmentation results on 2 classes were accurate and benefit of the same model can be taken in segmentation of 3 classes in Terahertz images.

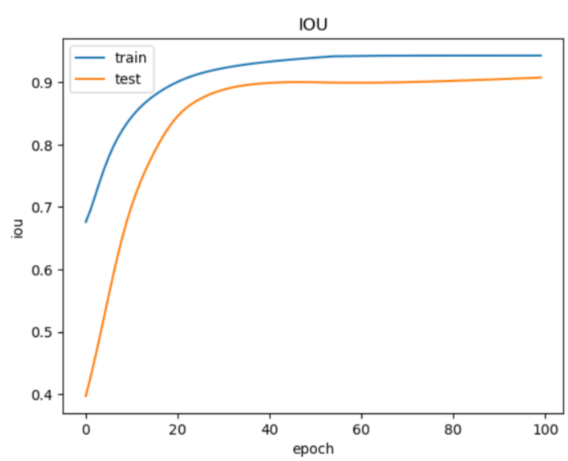
The last layer of the model is changed as per the need as the number of classes is increased from 2 to 3, now the output of the model will be image height*image width*3, the activation function for the last layer is also changed to "SoftMax" instead of "sigmoid" as now there are more than 2 classes.

2.2.3.3 MODEL EVALUATION

The accuracies and IoU scores after for training and validation sets after fitting the model are:

Training Accuracy	Validation Accuracy	Training IoU Score	Validation IoU Score
98.86%	98.06%	0.943	0.91

Table 3

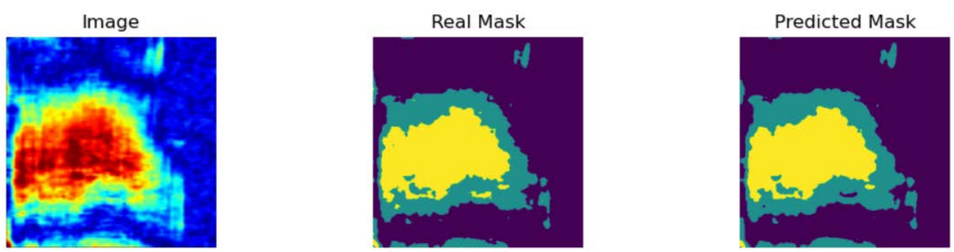


IoU score vs epoch after training the model for 100 epochs

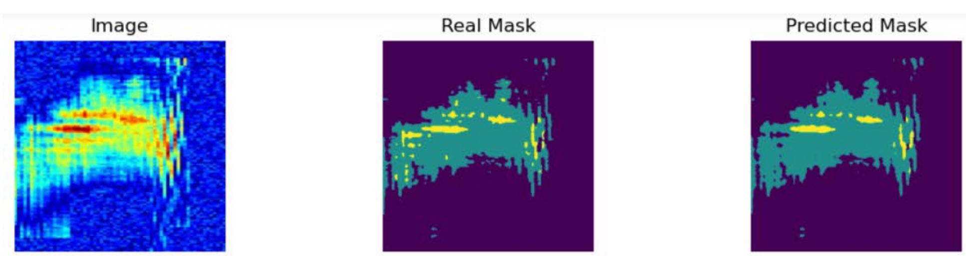
2.2.3.4 VISUALIZING SEGMENTATION RESULTS


```
Image 1 :
PixelWise Accuracy : 0.9884186
{'IoU Background': 0.9798341141895709, 'IoU Non Tumor': 0.9767506227211381, 'IoU Tumor': 0.9429223739986656}
Mean IoU : 0.96650237
Dice coefficient : 0.9828919137298646
```

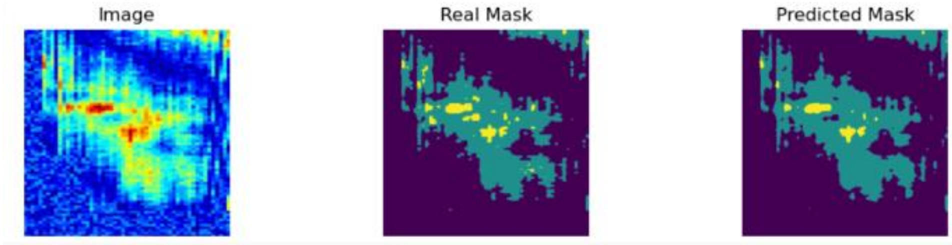

```
Image 2 :
PixelWise Accuracy : 0.98890686
{'IoU Background': 0.981422324194975, 'IoU Non Tumor': 0.9719120657971676, 'IoU Tumor': 0.9842693866309119}
Mean IoU : 0.97920126
Dice coefficient : 0.9894841336803654
```



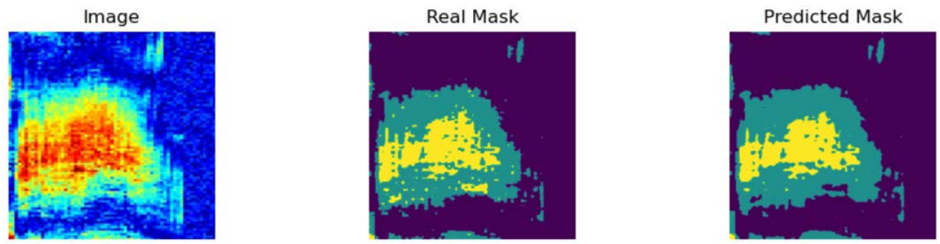

```
Image 3 :
PixelWise Accuracy : 0.9849243
{'IoU Background': 0.9819089207490478, 'IoU Non Tumor': 0.9390989988223651, 'IoU Tumor': 0.9689183588090411}
Mean IoU : 0.96330875
Dice coefficient : 0.9812262954757717
```



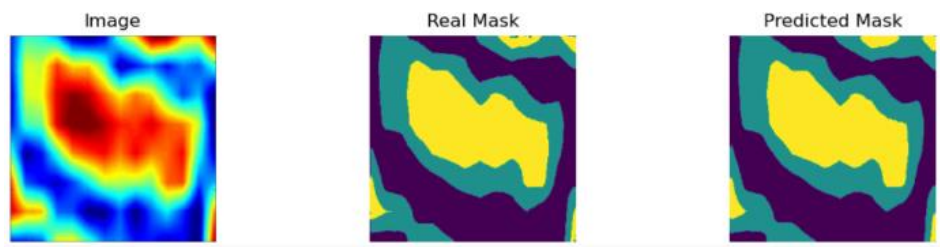

```
Image 4 :
PixelWise Accuracy : 0.9786377
{'IoU Background': 0.9801080695726677, 'IoU Non Tumor': 0.9247166755987104, 'IoU Tumor': 0.5726872243091962}
Mean IoU : 0.8258373
Dice coefficient : 0.8930438193864793
```




```
Image 5 :
PixelWise Accuracy : 0.97013855
{'IoU Background': 0.9674278596436074, 'IoU Non Tumor': 0.9041633935077902, 'IoU Tumor': 0.5756162554459585}
Mean IoU : 0.8157358
Dice coefficient : 0.8879232202715888
```




```
Image 6 :
PixelWise Accuracy : 0.9676819
{'IoU Background': 0.9705365204350488, 'IoU Non Tumor': 0.9017476788147598, 'IoU Tumor': 0.8350785339163573}
Mean IoU : 0.9024542
Dice coefficient : 0.947837384513269
```




```
Image 7 :
PixelWise Accuracy : 0.9856415
{'IoU Background': 0.9730161142834822, 'IoU Non Tumor': 0.9609321157408282, 'IoU Tumor': 0.9843648592256555}
Mean IoU : 0.972771
Dice coefficient : 0.9861737498726256
```

3. RESULTS

The models were tested as per the specified testing data in each model and their corresponding accuracies and IoU scores were recorded.

A summary of the validation set accuracy metric and IoU of the best performing models is shown in Table 4.

Data	Model	Metric	Type
MRI brain Images	VGG-19	Accuracy: 95.5% IoU: 0.845	Segmentation on 2 classes
Terahertz Image	VGG-19	Accuracy: 92.3% IoU: 0.873	Segmentation on 2 classes
Terahertz Image	VGG-19	Accuracy: 98.06% IoU: 0.91	Segmentation on 3 classes

Table 4

4. CONCLUSIONS

In conclusion, the development and successful implementation of the segmentation model for terahertz tumor images have proven to be a significant advancement in medical imaging and tumor analysis. **The model achieved an impressive Intersection over Union (IoU) of 90.7 even on a small dataset of terahertz images with lower resolutions, showcasing its capability to accurately delineate tumor boundaries and provide valuable insights for medical practitioners and researchers.**

The high IoU score indicates that the segmentation model has demonstrated exceptional precision in identifying tumor regions within the terahertz images. This level of accuracy is vital in assisting medical professionals with accurate diagnosis, treatment planning, and monitoring of tumor progression. The ability to precisely segment tumor regions using terahertz imaging can lead to early detection, more targeted therapies, and improved patient outcomes.

Additionally, the successful development of this segmentation model opens new possibilities for advancing the field of terahertz medical imaging. The application of deep learning techniques to terahertz imaging provides a promising avenue for enhancing the understanding of tumor behavior and interactions with surrounding tissues.

However, it is important to acknowledge that there are still challenges that need to be addressed. The model's performance may be influenced by factors such as data quality, class imbalance, or variations in tumor shapes and sizes. Therefore, further research and refinement of the model should be pursued to make it more robust and generalizable to different datasets.

5. REFERENCES

https://www.tensorflow.org/

https://paperswithcode.com/method/u-net

https://medium.com/analytics-vidhya/iou-intersection-over-union-705a39e7acef

https://idiotdeveloper.com/vgg16-unet-implementation-in-tensorflow/

https://idiotdeveloper.com/vgg19-unet-implementation-in-tensorflow/

https://www.kaggle.com/datasets/awsaf49/brats2020-training-data