**Deep Learning:**

Deep Learning (DL) is a subfield of machine learning (ML) that focuses on learning representations of data and hierarchical feature learning. For feature extraction, DL algorithms employ an arrangement of several layers of nonlinear processing identities. As we progress further into the network, the output of each sequential layer becomes the input of the next one, which aids in data abstraction [26]. For semantic segmentation especially in the medical field, U-net models are used. So we are mainly focused on the U-net model. The Architecture section describes the network's detailed design. Following its architecture, we experimented with the network by changing the number of layers in the network. We have also implemented the generic U-net model and evaluated its result and compared it with the modified U-net model.

**U-net Architecture:**

The generic U-net architecture is shown in the fig [2]. U-network design (example for 32x32 pixels in the lowest resolution). A multi-channel function map is represented by each blue box. The number of channels is shown on the box's end. The x-y size is given on the box's lower left side. White boxes show feature maps that have been copied. The arrows represent the various operations [2[7](#_ENREF_32)].

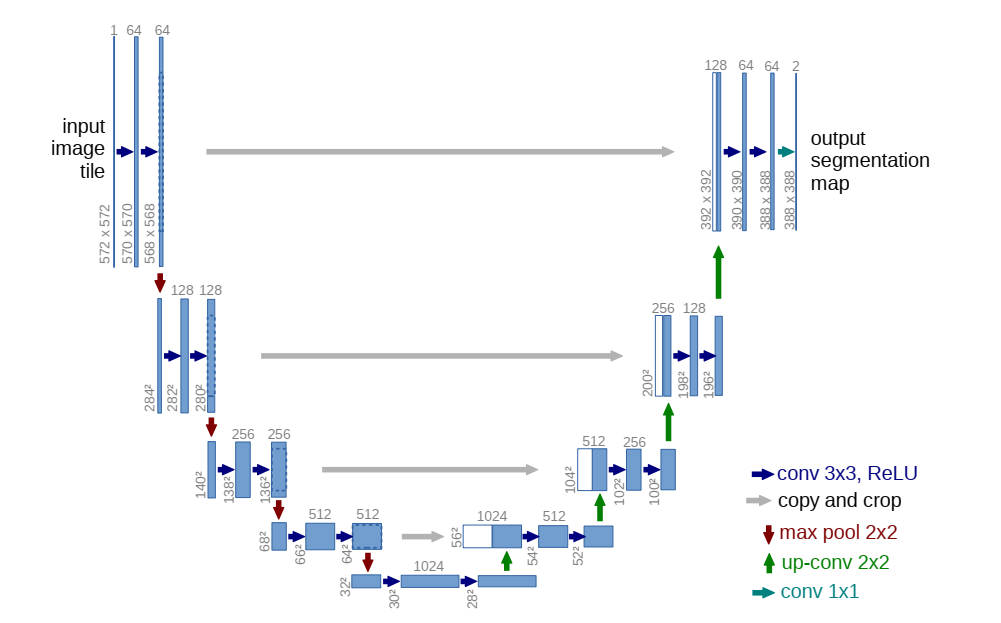


Figure 2 Generic U-net Model

**Modified U-net Model:**

Figure [2] depicts a generic model of the Unit. The implementation of the modified model can be seen in figure [3].

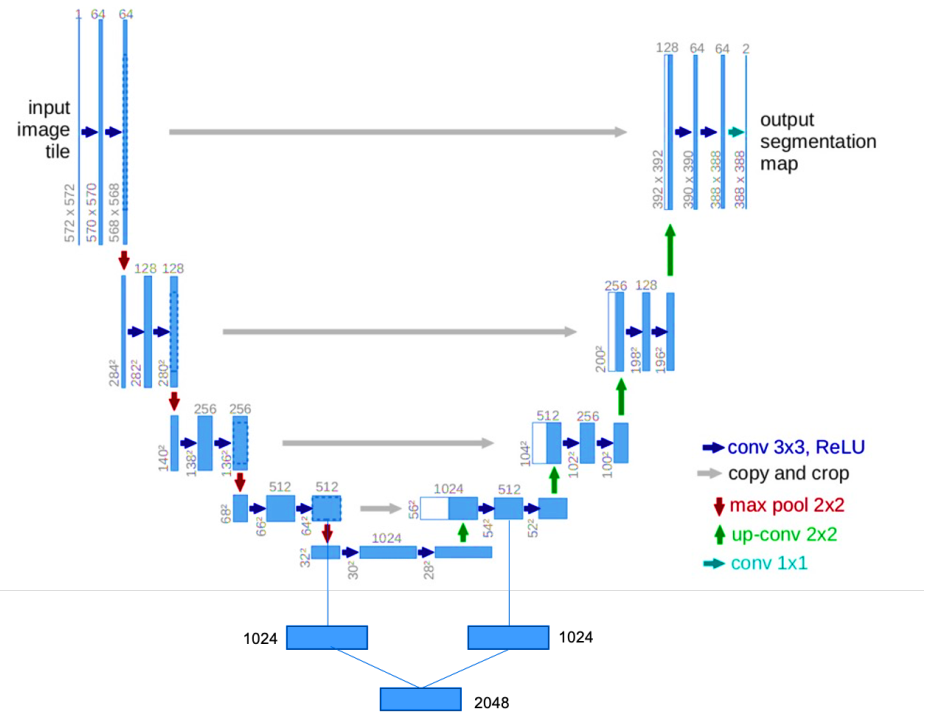


Figure 3 Modified U-net Model

If we divide the architecture into two parts, the left side will be the contracting path and the right side will be the expanding path. The number of features would be doubled every time the contracting direction was followed, a method is known as downsampling. Following downsampling, the expansive direction of upsampling will have several features. Downsampling is accomplished by repeating two 3x3 convolutions, followed by a rectified linear unit (ReLU) activation function and a max-pooling with a kernel size of 2x2 and a stretch of 2. In the process of upsampling, we repeat the same procedure that we did in the contracting path, which is two 3x3 convolutions followed by a ReLU activation function. In addition, the layers in the expansive path are concatenated with the corresponding layers in the contracting path. A 1x1 convolution is used at the final layer to map into the desired class.

**Training:**

In this model we train our model on 322 images with their ground truth images. For training we use the relu activation function and we use kernel he normal. We set padding same to resize image after upsampling as same as input image.

**Activation Function:**

The activation function is used to decide whether a neural network's performance is yes or no. It converts the values to a range of 0 to 1 or -1 to 1. In this model for input layers we use ReLU activation function and at output layer we use sigmoid function.

**ReLU Function:**

ReLU stands for rectified linear unit it give value linearly when threshold value is more than zero.

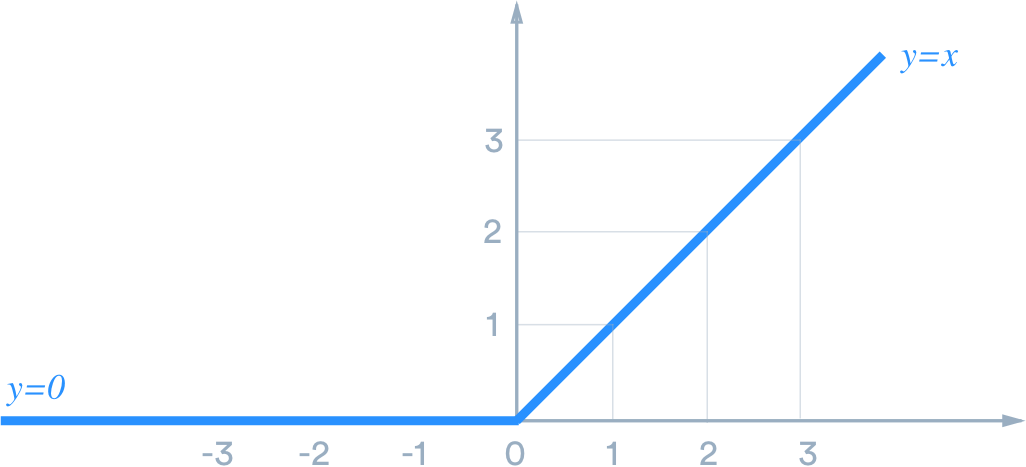


Figure 4 ReLU Function

**Equation:**

Equation 1 ReLU Function Equation

**Sigmoid Function:**

The logistic function, also known as the sigmoid activation function, has long been a common activation function for neural networks. The function's input is converted to a value between 0.0 and 1.0.

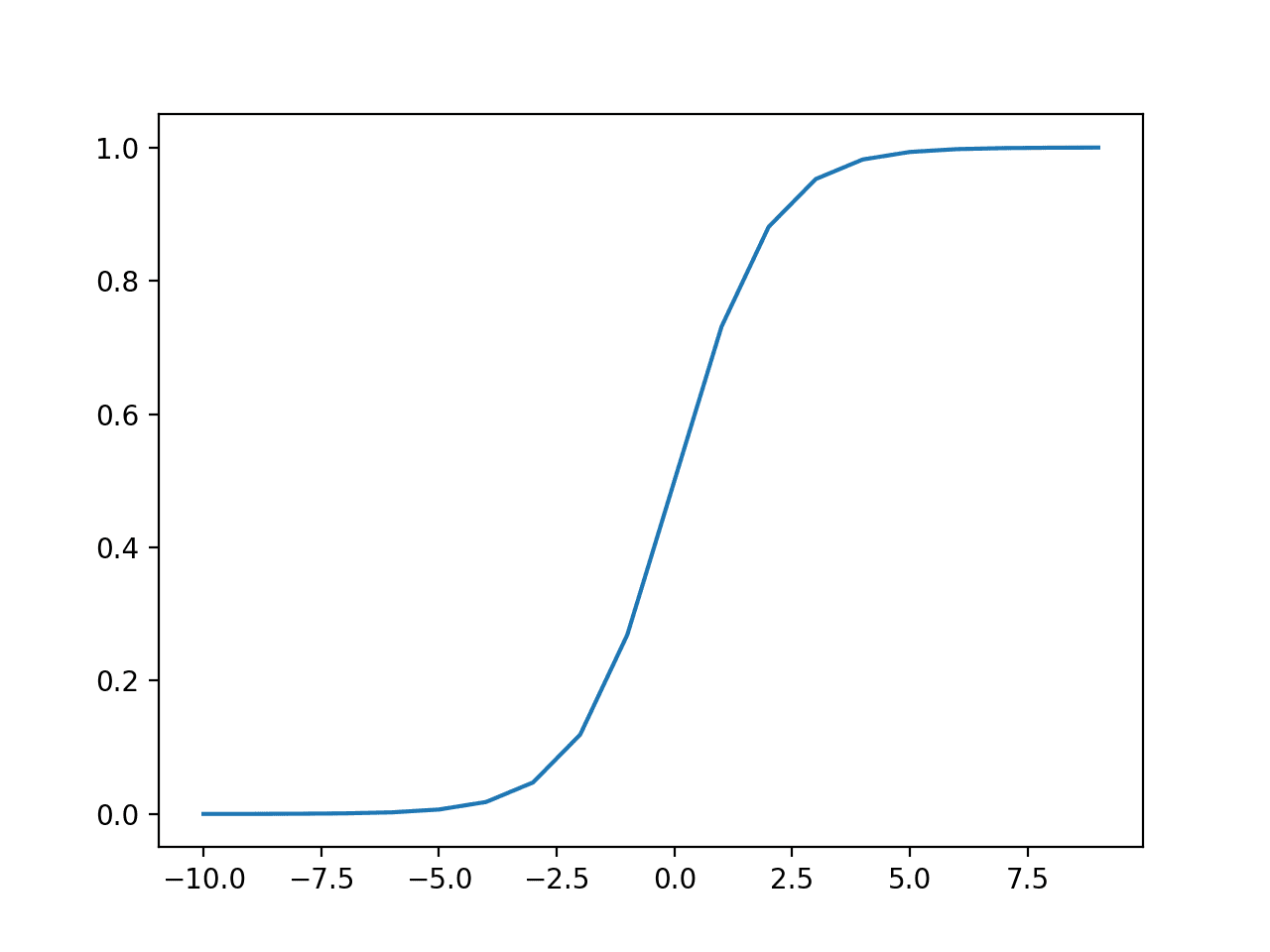


Figure 5 Sigmoid Function

**Equation:**

Equation 2 Sigmoid Function Equation

**Downsampling:**

We decide to downsample class 0 because the majority of the pixels belong to it and do little to boost distinguishing ability. The approach is straightforward: we only select slices that contain tumor pixels.in downsampling we keep shrinking image and applying filters on image in each step we shrink image by half and increase filter by double till we reach number of filters to 2048.

**Upsampling:**

In upsampling we reduce the number of filter and stretch image size by double and also add the cross ponding downsampling values with that so we can obtain features. At last layer we convert final output image into binary image and use activation function sigmoid to get more accurate value of prediction.

**Validation:**

We have also used some of the dataset as validation. We make a validation on 22 images. The validation model train accuracy is 96.226% which is batter then generic U-net.

**Evaluation Parameters:**

To evaluate the performance of our method, we use Dice Score. Dice Score Coefficient is used to check the accuracy between the ground truth images and segmented images. The dice coefficient is a fixed similarity measure function that is typically used to determine the similarity of two samples, with a value range of [0,1]. Dice Loss is 1 − Dice coefficient so if two samples more similar, the DL will be smaller.

(1)

Equation 3 **Dice Loss Equation**

The images in the dataset have also been tested and results are obtained in the form of specificity, sensitivity and accuracy of the following parameters.

* TP (True Positive)
* TN (True Negative)
* FP (False Positive)
* FN (False Negative)

#### Specificity

Sensitivity is the measure of correct analysis of person not having tumor. Sensitivity is given by:

(2)

Equation 4 **Specificity**

#### Sensitivity

Sensitivity is the measure of correct analysis of person having tumor, and is given by:

(3)

Equation 5 **Sensitivity**

#### Accuracy

Accuracy shows the efficiency of the proposed methodology in term of accurate classification.

(4)

Equation 6 **Accuracy**