AN EXPLAINABLE AI FRAMEWORK USING ATTENTION ENCODER-DECODER ARCHITECTURE

Dr. PL. Chithra[1]*, Ms. Dhivya.S.D[2]

Abstract:
Explainable Artificial Intelligence, widely known as interpretable artificial intelligence, is the key component of artificial intelligence. It helps to characterize model accuracy, fairness, transparency, and outcomes of a model in decision-making. The core idea behind LIME XAI leads to new insight into how machine learning models work and validates model efficiency. Especially, in the medical field diagnosing a disease with an explanation about the prediction is a milestone in recent years. This work proposes an attention encoder-decoder, skip connection that combines spatial information of MRI images from the down-sampling path with the up-sampling path to retain good spatial information. Soft attention implemented at the skip connections will actively suppress activations at irrelevant regions. The bilinear interpolation is used in the upsampling process to retain even small details of an image that is lost during encoding and decoding.
Keywords: attention mechanism, dimensionality, interpretable artificial intelligence, LIME UNet, semantic features, transposed convolution, superpixel

[1]*Department of Computer Science, University of Madras, Chennai - 600 005, Tamil Nadu, India.
Email: chitrasp2001@yahoo.com
[2]Department of Computer Science, University of Madras, Chennai-600 005 Tamil Nadu, India.
Email: dhivyadharmalingam@gmail.com

***Correspondence Author:** Dr. PL. Chithra
*Department of Computer Science, University of Madras, Chennai - 600 005, Tamil Nadu, India.
Email: chitrasp2001@yahoo.com

DOI: - 10.48047/ecb/2023.12.si5a.0152

1.Introduction:

In recent years automated techniques for brain tumor classification using deep learning approaches have gained significant research interest. Yet, the difficulties in extracting and classifying the tumor regions from the 3D images do not have a definite solution. To address these problems, this paper proposes an encoder–decoder-based model. LIME UNET aims to retain even small-scale images. LIME (explainable artificial intelligence) explains the classification results efficiently. The key feature of LIME its accessibility and simplicity. The core idea behind LIME is Model agnosticism and Local explanations

Model agnosticism refers to the property of LIME using which it can explain any supervised learning model by treating it as a 'black box' separately. This means that LIME can handle almost any model that exists. Local explanations mean that LIME provides locally faithful explanations within the surroundings or vicinity of the observation sample being explained.

Gliomas are the most common type of tumor of the central nervous system (brain and spinal cord). They originate from glial cells that grow uncontrollably. Gliomas are classified according to the World Health Organization (WHO) by Grades (I-IV), considering their growth pattern and infiltrating capacity. The WHO further classifies gliomas into two grades namely low-grade gliomas (LGG grade I, II) and high-grade gliomas (HGG grade III, IV). Low-Grade Gliomas, Grade I, and Grade II gliomas are considered slow-growing tumors that on rare occasions they can transform into aggressive tumors. High-Grade Gliomas, Grade III, and Grade IV gliomas can appear spontaneously, but in many cases, they arise from low-grade tumors. Grade IV lesions are the most aggressive, fast-growing, and invasive. At present, the diagnosis of glioma grade is determined by surgical biopsy or histopathological analysis. However, the diagnostic method is invasive and limited for patients who are not suitable for surgery. Therefore, it is necessary to design a non-invasive and highly accurate glioma grading system

2. Literature Survey:

DL techniques enhanced the performance of medical imaging diagnostic frameworks, and XAI provides the reason behind the specific classification and prediction results. XAI's primary goal is to investigate and develop methods for explaining the individual predictions of DL systems. That a clear explanation of the reached decision is critical in medical applications depending on images. In an ideal case, the system makes a decision based on the input data and justifies which image part led to a certain classification result. some of the recent related works are Towards explainable artificial intelligence, in Explainable AI: Interpreting, Explaining and Visualizing Deep Learning [1], Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI) [2]. MRI segmentation are threshold-based [3], region based[4], deformation-based[5], fuzzy-based[6], and neural network based[7].

Ankit Rajpal and Kountay Dwived proposed their work "An explainable AI-driven biomarker discovery framework for Non-Small Cell Lung Cancer classification", in February 2023[8]About framework comprises three modules — an autoencoder to shrink the input feature space, a feed-forward neural network to classify NSCLC instances into lung Adenocarcinoma (LUAD), Lung Squamous Cell Carcinoma (LUSC), a biomarker discovery module that leverages the combined network comprising the autoencoder and the feed-forward neural network using a small scale dataset.

Ahmad Chaddad, Jihao Peng, and Jian Xu work on the "Survey of Explainable AI Technique in HealthCare", in January 2023[9] Which provides a survey of the most recent XAI techniques used in healthcare and related medical imaging applications. It categorizes the XAI types, to increase interpretability in medical imaging with this seed, the proposed work is focused on the challenging XAI problems in medical applications and provides guidelines to develop better interpretations of deep learning models using XAI concepts in medical image and text analysis.

C. Ledig et.al, "Efficient multi-scale 3D CNN with fully connected a dual pathway of an 11-layer deep, three-dimensional Convolutional Neural Network for brain lesion segmentation" [10], which results in an in-depth analysis of the limitations of current networks. It has overcome the computational burden of processing 3D medical scans and devised an efficient and effective dense training scheme that joins the processing of adjacent image patches into one pass through the network while automatically adapting to the inherent class imbalance present in the data. LIME UNET intends to predict the attention encoder-decoder 's working methodology [11][12]

using LIME XAI and brings out the transparency of the architecture with better accuracy. LIME UNET intends to predict the attention encoder-decoder's working methodology using LIME XAI and brings out the transparency of the architecture with better accuracy. In general, Deep Learning (DL) techniques [13][14] enhanced the performance of medical imaging diagnostic frameworks, specifically, XAI provides the classification and prediction interactively

This paper proposes a methodology in section 3, dataset and preprocessing in section 4, followed by result conclusion in section 5.

3. Methodology:

The Lime Unet architecture shown in Fig. 1 includes encoder-decoder, attention mechanism, 3D Slice-overlay, upsampling transposed convolution, the loss is calculated by combining the dice coefficient and binary cross-entropy, and finally LIME superpixel.

3.1. Encoder:

Five coding units, E_i shown in Fig.2 are used to extract high-level features from the training slices of an image. Suppose the size of the input training slices X is C_i x W_i x H_i where C_i is the number of channels, W_i, H_i is the width and height of the feature training slices. C_0= 5, W_0= H_0 = 5. E_i includes a residual part, LIME, and Segmentation layer.

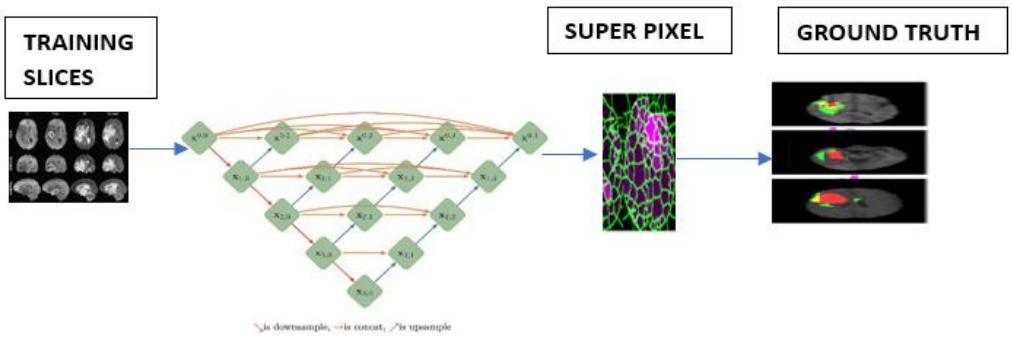

Fig. 1 Architecture of LIME UNET

The residual structure consists of the tensor X passing through a convolutional layer with the 1 x 1 kernel and a batch normal (BN) layer to obtain a tensor X1 (2 C_i x W_i x H_i) mentioned in equation 1

$$X1 = BN\left(Conv_{1x1}(X)\right) \in \mathbb{R}^{2C_i x W_i x H_i} \quad (1)$$

The residual structure consists of, the tensor X1 passing through a convolutional layer with the 3 x 3 kernel, batch normal (BN) layer, and a Relu layer to obtain a tensor X2 (2 C_i x W_i x H_i) mentioned in equation 2

$$X2 = \left(Relu\left(BN\left(Conv_{3x3}(X1)\right)\right)\right) \in \mathbb{R}^{2C_i x W_i x H_i} \quad (2)$$

The residual structure consists of, the tensor X2 passing through a convolutional layer with the 3 x 3 kernel batch normal (BN) layer, and a Relu layer to obtain a tensor X3 (2 C_i x W_i x H_i) mentioned in equation 3

$$X3 = \left(Relu\left(BN\left(Conv_{3x3}(X1)\right)\right)\right) \in \mathbb{R}^{2C_i x W_i x H_i} \quad (3)$$

Then, the skip connection [13]is applied from X1 to X3, giving the output of the residual structure to X_{Res} (2 C_i x W_i x H_i) as mentioned in equation 4.

The residual structure helps the coding unit E_i to overcome the degradation problem of the multilayer downsampling network.

$$X_{Res} = X1 + X3 \in \mathbb{R}^{2C_i x W_i x H_i} \quad (4)$$

Later, X_{Res} is passed to the channel attention mechanism, to obtain X_c, it is then passed to the 2 x 2 max pooling layer and finally to E_{i+1}. Then X_c is also passed to the upper decoding unit $V_{i-1,1}$ is obtained as $2C_i$ x $2W_i$ x $2H_i$ as the outcome of 4 x 4 transposed convolution.

3.2. Decoder

In the decoding process, multiple cascaded decoding units are used ($V_{i,j}$, where, i = 0,1,.., 3 and j = 1,2,...,4 in Fig.3) . The input of $V_{i,j}$ consists of two-part in the first part the feature map obtained by the lower decoding unit $V_{i+1,\ j-1}$ after transposed convolution the size is 4 C_i x 4 H_i x 4 W_i and in the second part the feature maps output is obtained by the same layer decoding unit $V_{i,0}$, $V_{i,1}$ . . . . , $V_{i,j-1}$ shown in fig 3. The size of each feature map is $2C_i$ x $2W_i$ x $2H_i$ $V_{i,0} = E_i$. As $Y_{i,j}$ of $V_{i,j}$ is the same size as the feature map can be obtained by superimposing on the channel and size (4+2j) C_i x H_i x W_i. Each structure contains a

3 x 3 convolutional layer, a batch normal, and an activation layer.

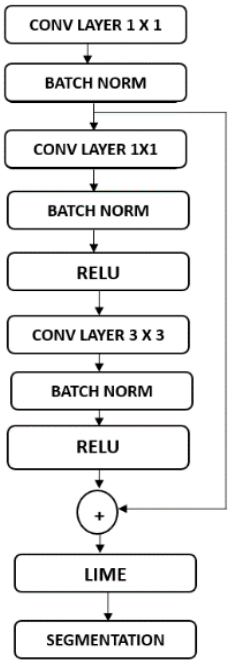

Fig. 2a E_i , Coding Unit

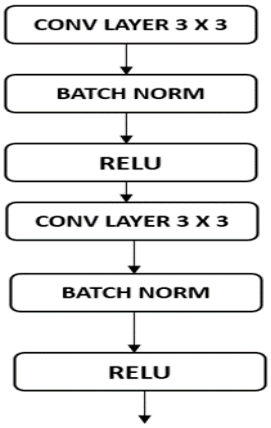

Fig. 2b $V_{i,j}$

At first $Y_{i,j}$ passes through the first VGG structure [15] to obtain the feature map of $Y1_{i,j}$ ($2C_i$ x H_i x W_i) is mentioned in equation 5

$$Y1_{i,j} = (\text{Relu}\,(\text{BN}\,(\text{Conv}_{3x3}\,(X1)))) \in \mathbb{R}^{2C_i x W_i x H_i} \quad (5)$$

$Y1_{i,j}$ passes through the first VGG structure to obtain the feature map of $Y2_{i,j}$ ($2C_i$ x H_i x W_i) is mentioned in equation 6

$$Y2_{i,j} = (\text{Relu}(\text{BN}\,(\text{Conv}_{3x3}\,(X1)))) \in \mathbb{R}^{2C_i x W_i x H_i} \quad (6)$$

3.3. Attention Mechanism:

The LIME UNet uses the AM layer consisting of the convolution layer and Relu in each encoding unit to retain and amplify the features of even small targets. In image segmentation, attention layers are used to highlight only the relevant activations during training. This reduces the computational resources wasted on irrelevant activations, providing the network with better generalization power so that the network can pay "attention" to the required parts in the image.

Due to the undifferentiated treatment of all channels, the edge information will be ablated in multiple down-sampling operations (max pooling). Therefore, this makes it easy for these models to miss some small-scale targets. In the AM layer module, we use global pooling and a fully connected layer (FC layer)[15] to continue exploring the feature map. Then, feed the result of feature extraction to different channels through the attention mechanism, to give higher weight to the feature map of the captured lesion. This process will retain the most useful information. The structure of the AM layer module is shown in Figure 3.

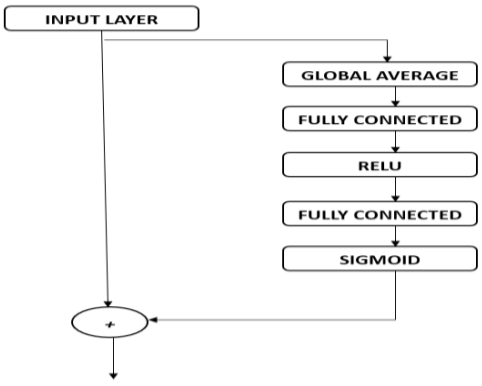

Fig. 3 Attention Mechanism Layer

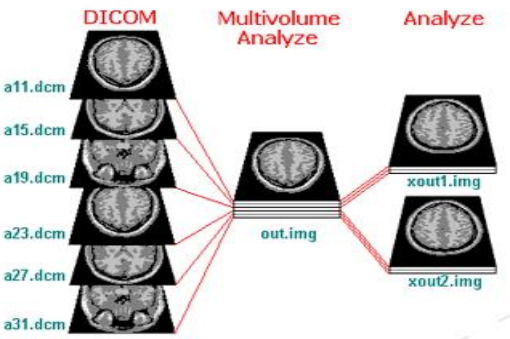

Fig. 4 3D Slice Overlay Method

Assume that the input of AM layer is a feature map X (C x W x H) and the module uses global average pooling to obtain the global features in each channel feature map to obtain output Z[K] mentioned in equation 7

$$Z[K] = F_{GP}(X[K]) = \sum_{i=1}^{w}\sum_{j=1}^{h} \frac{1}{WH} X[K](i,j) \quad (7)$$

Where X[K] (K=1,2, . . ., C) is the feature map of Kth channel[K](i, j) in X[K] and Z[K] is the calculation result of Kth channel. Later global average pooling is calculated for input feature map X is mentioned in equation 8

$$Z = F_{GP}(X) \in \mathbb{R}^{C} \quad (8)$$

Then Z is passed to the first fully connected layer where the parameter matrix is performed to reduce the dimension of Z, Relu activation function is used to obtain w1 and passed to obtain w2 Then here sigmoid activation is used. The attention coefficient is calculated using A is mentioned in equation 9

$$A = F_{Att}(X) = \sigma(W2 \text{ x } \text{Relu}(W1x\, F_{GP}((X)))\, \mathbb{R}^{C} \quad (9)$$

At last attention, the coefficient is multiplied with the feature map X to each corresponding channel to enlarge or reduce the information accordingly to obtain X1 is mentioned in equation 10, where $\odot$ is Hadamard product

$$X1 = F_{Att}(X) = A \odot X \in \mathbb{R}^{CxWxH} \quad (10)$$

3.4. Upsampling:

We use transposed convolution to achieve up-sampling because this process is learnable, and high-resolution information can be fully recovered during the parameter adjustment process, avoiding errors introduced in the bilinear interpolation process, and retaining the clear boundaries. and for edge slicing, a mirror image extension method is used.

3.5. 3DSlice Overlay method:

After preprocessing, training slices of a patient are arranged vertically into series of slices S_1, S_2, . . . , S_m, and each slice is adjusted to a single-channel slice with a size of 256 x 256 generates masks for each slice in turn to retain even small details of tumors [16] as shown in figure 4. to generate the mask of slice S_i , the input of the model S_{i-l}, S_{i-l+1} . . . , S_{i+1} and by using slice overlay method is mentioned in equation 11. The tumor positioning is in a Computed tomography (CT) slice with superimposed segmentation. Three-dimensional (3D) rendering of the segmented data set, structures can be removed or made transparent for exact visualization of tumor positioning and tumor relationship with blood vessels. Thus, feature map $F\boldsymbol{X_i}$ (2l +1) x 256x 256) is formed by an overlay of the slices and the number of overlayed channels are p = 2l+1 is one of the hyperparameters of the model

$$X_i = [S_{i-l}, S_{i-l+1}, \ldots, S_i \ldots, S_{i+l-1}, S_{i+l}] \quad (11)$$

3.6. LIME

LIME works by identifying features, such as superpixels (picture zones), that improve the likelihood of the expected class, that is areas that support the present model prediction. Since this model actively employs such regions to produce predictions, they might be considered important. LIME used MRI images as input and generated the sequence present in the image. After that, it generated the interpretable representations and generated N samples. Then, it matched each sample with the featured map of the input images and calculated the predicted label and distance from the predicted output. These labels and distance values were passed to a linear model that provided the explanations, and a specific result was calculated. LIME also highlighted the regions in the image, which can represent and explain the reason for classification results mentioned in Equation 12

$$\xi(x) = \underset{g \varepsilon G}{\text{argmin}}\, L(f, g, \pi_x) + \Omega(g) \quad (12)$$

Where f is a model classifier, G is the class of interpretable models, gG shows learning of local explainable model, g is used to measure model complexity, x is the proximity measure and L (f, g) is loss or distance function.

4. Dataset and preprocessing

4.1. Dataset acquisition:

BraTS 2020,2021 multimodal scans are available as NIfTI files (.nii.gz) commonly used in medical imaging format to store brain imaging data obtained using MRI and describe different MRI settings.

1. **T1**: T1-weighted, native image, sagittal or axial 2D acquisitions, with 1–6 mm slice thickness.
2. **T1c**: T1-weighted, contrast-enhanced (Gadolinium) image, with 3D acquisition and 1 mm isotropic voxel size for most patients.

3. **T2**: T2-weighted image, axial 2D acquisition, with 2–6 mm slice thickness.
4. **FLAIR**: T2-weighted FLAIR image, axial, coronal, or sagittal 2D acquisitions, 2–6 mm slice thickness.

Data were acquired with different clinical protocols and various scanners from multiple (n=19) institutions. All the imaging datasets have been segmented manually, by using the Simple ITK tool, following the same annotation protocol, and their annotations were approved by experienced neuro-radiologists. Annotations comprise the GD-enhancing tumor (ET — label 4), the peritumoral edema (ED — label 2), and the necrotic and non-enhancing tumor core (NCR/NET — label 1), as described both in the BraTS 2020 as shown in figure 6 and the ages of the patient and the region of tumor such as necrotic region, edema and enhancing tumor region in table 1

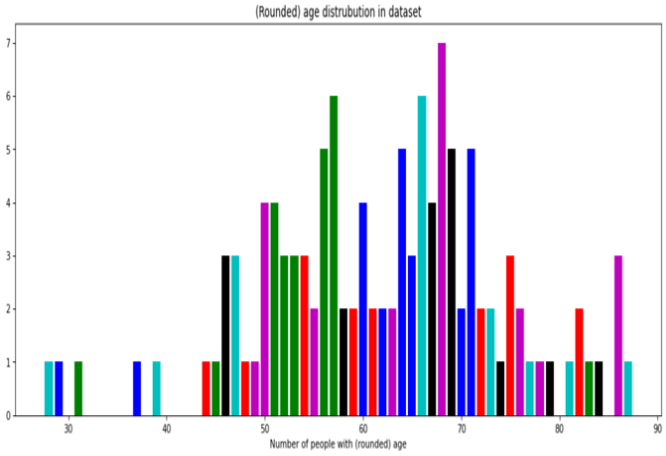

Fig. 6 Dataset of age and no of people

Table 1. Age and tumor regions of the patients

	age normalised	NECROTIC/CORE	EDEMA	ENHANCING
0	0.460631	0.060053	0.330630	0.105513
1	0.496066	0.374441	0.283556	0.406350
2	0.547051	0.113117	0.194263	0.633018
3	0.948964	0.973691	0.355460	0.520813
4	0.547373	0.011033	0.124241	0.147677
...	...	...	...	...
113	1.000000	0.337023	0.507990	0.673251
114	0.795687	0.207157	0.222702	0.605692
115	0.779440	0.037349	0.208887	0.115095
116	0.807277	0.383568	0.179196	0.338549
117	0.371000	0.069379	0.095903	0.263933

118 rows × 4 columns

4.2 Preprocessing:

All the imaging datasets have been segmented manually, by using the MATLAB OpenCV toolbox, for preprocessing bilinear interpolation for retaining the Region of Interest (ROI) detail of an image, z-axis interpolation for equal spacing between each slice, adjacent interpolation for labeling the position

5. Result and Conclusion:

5.1. Performance Matrices:

The LIME values for different super-pixels of the sampled images were computed using LIME value indicates the positive impact on the positive prediction, while the negative value means that the super-pixel contributes to the negative prediction as shown in Figure 7 and the dice similarity coefficient values are shown in Figure 8

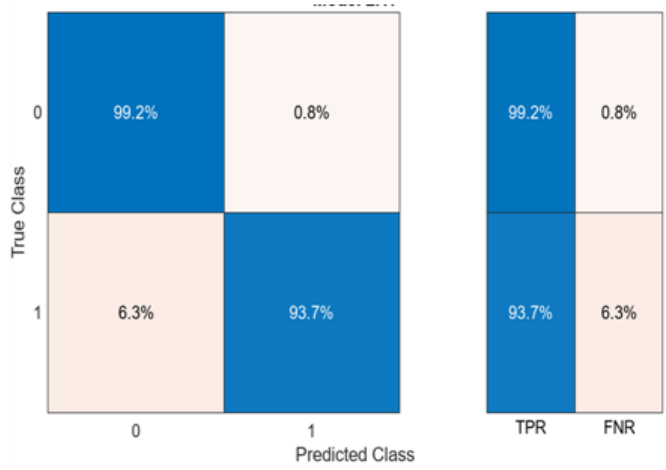

Fig. 7 TP and FN value

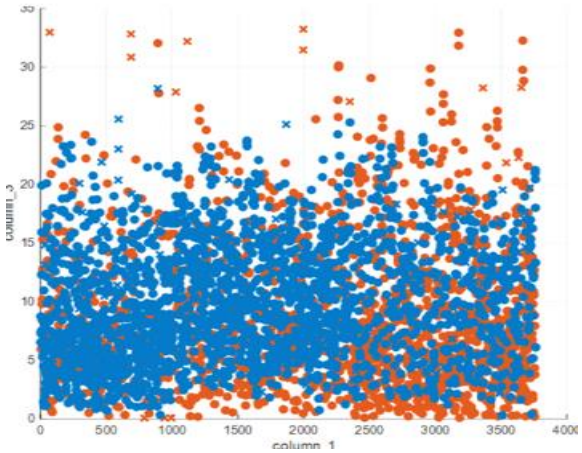

Figure 8: Dice Similarity Coefficient

The five parameters were used for measuring the performance of matrices as shown in Table 2 and as follows

5.1.1. Accuracy:
The correct predicted cases divided by the total number of cases gives us accuracy [17]. It is a sum of true positives and negatives which is TP + TN divided by the sum of TP (True positives), TN (True Negatives), FP (False positives), and FN (False negatives) as mentioned in equation 13

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (13)$$

5.1.2. Precision:

Precision is the number of correct results out of the predicted results. It is calculated by dividing true positives by the sum of true positives and false positives as mentioned in Equation 14

$$\text{Precision} = \frac{TN}{TP+FP} \quad (14)$$

5.1.3. Specificity:
A number of valid negative predictions divided by the total number of negative as mentioned in equation 15

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (15)$$

Table 2. Performance Metrics values

Classes / Scores	PE	NCR\NET	ET	NT
DSC (F1 score)	81%	67%	78%	98%
Mean IoU (JSC)	77%	75%	78%	97%
Precision	79%	83%	84%	99%
Sensitivity (Recall)	95%	87%	90%	97%
Mean BF score	76%	81%	87%	91%

5.1.4. Recall:
The recall is defined as a number of the positive predicted results out of the total positive cases, also known as Sensitivity and termed as the true positivity rate. It is measured by true positives which are divided by the sum of true predictions.

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (16)$$

Based on the dice similarity coefficient value and TP and FN accuracy and loss function were calculated as shown in equation 16

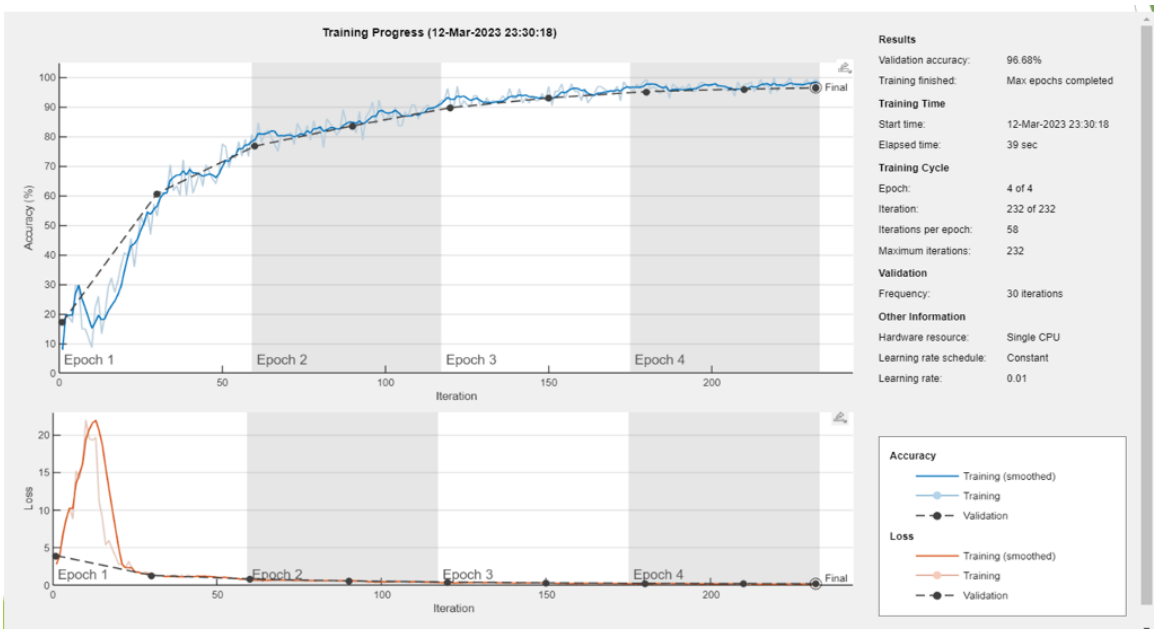

Fig. 9 Training dataset loss and accuracy

5.2. Comparative Analysis:

The Dice Similarity coefficient (DSC) is used for evaluating the segmentation process, are mainly used to measure ensemble similarity, which is usually used to calculate the similarity of two samples. The LIME UNET DSC is compared with UNET [18], UNET++[19], and nUNET[20] as shown in Figure 10, and the accuracy curve in Figure 11. Ground truth is shown in Figure 12

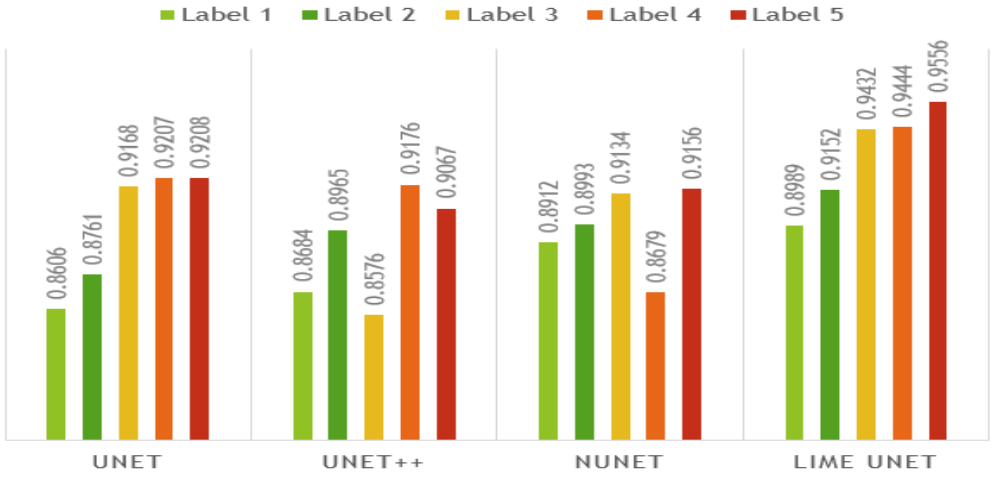

Fig. 10 Comparative analysis of Dice Similarity Coefficient

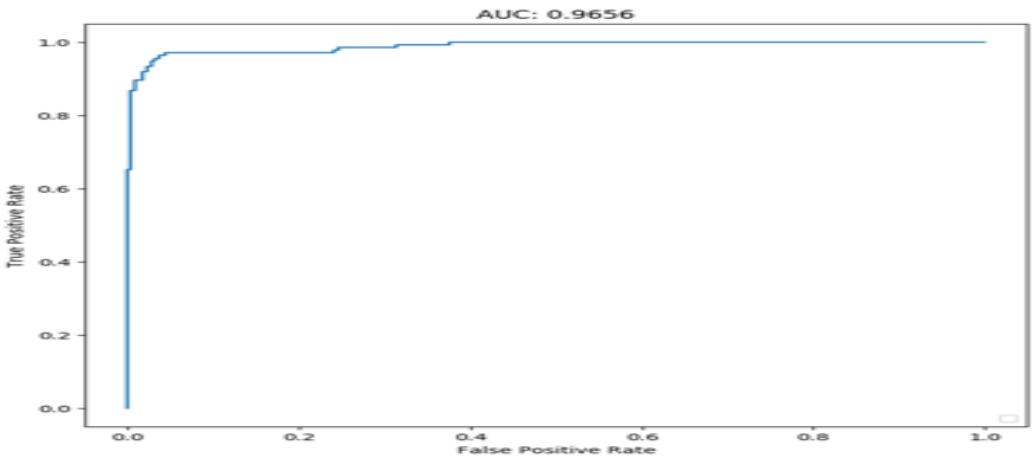

Fig. 11Accuracy Curve

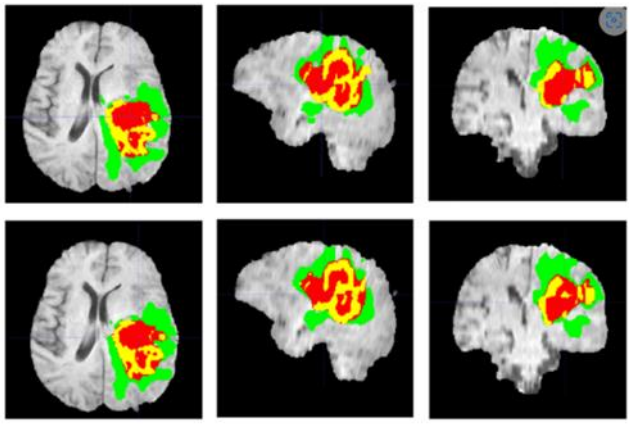

Fig . 12 Axia, Sagittal, Coronal views of the tumor region b. Ground Truth Visualization of Experimental Results

Conclusion:

An explainable AI-based framework was proposed in this research to address the challenge of classification result explainability in medical images. This research framework provides explainability This developed system achieved improved classification accuracies of 96% it further explains the classification results by using the interpretable LIME model. These highlighted regions are the important features that are used in the classification of diseases. Evaluation and testing show that this approach can explain the classification results using.

Acknowledgment:

Our Sincere thanks to the University of Madras, Department of Computer Science, Chennai, Tamil Nadu, India.

References:

1. W. Samek, K.-R. Müller, Towards explainable artificial intelligence, in Explainable AI: Interpreting, Explaining and Visualizing Deep Learning, Springer, 2019, pp. 5–22
2. A. Adadi, M. Berrada, Peeking inside the black box: A survey on explainable artificial intelligence (XAI), IEEE Access 6 (2018) 52138–52160.
3. A. Alansary, K. Kamnitsas, A. Davidson, R. Khlebnikov, M. Rajchl, C. Malamateniou, M. Rutherford, J. V. Hajnal, B. Glocker, D. Rueckert, et al., Fast fully automatic segmentation of the human placenta from motion corrupted MRI, in Proc. 19th Int. Conf. on Medical Image Computing and Computer-Assisted Intervention, Athens, Greece, 2016, pp. 589–597.
4. P. Moeskops, J. M. Wolterink, B. H. M. van der Velden, K. G. A. Gilhuijs, T. Leiner, M. A. Viergever, and I. Isgum, ˇ Deep learning for multi-task medical image segmentation in multiple modalities, in Proc. 19th Int. Conf. on Medical Image Computing and Computer-Assisted Intervention, Athens, Greece, 2016, pp. 478–486.
5. X. Zhou, T. Ito, R. Takayama, S. Wang, T. Hara, and H. Fujita, Three-dimensional CT image segmentation by combining 2D fully convolutional network with 3D majority voting, in Proc. 1st Int. Workshop on Deep Learning in Medical Image Analysis, Athens, Greece, 2016, pp. 111– 120.
6. R. Korez, B. Likar, F. Pernus, and T. Vrtovec, Model-based ˇ segmentation of vertebral bodies from MR images with 3D CNNs, in Proc. 19th Int. Conf. on Medical Image Computing and Computer-Assisted Intervention, Athens, Greece, 2016, pp. 433–441
7. Dosovitskiy, A., Springenberg, J.T., Riedmiller, M., Brox, T.: Discriminative unsupervised feature learning with convolutional neural networks. In: NIPS (2014)
8. Ankit Rajpal and Kountay Dwived proposed their work "An explainable AI-driven biomarker discovery framework for Non-Small Cell Lung Cancer classification", in February 2023.

9. Ahmad Chaddad, Jihao Peng, and Jian Xu work on the "Survey of Explainable AI Technique in HealthCare", in January 2023
10. O. ¨ C¸ ic¸ek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, 3D U-Net: Learning dense volumetric segmentation from sparse annotation, in Proc. 19th Int. Conf. on Medical Image Computing and Computer-Assisted Intervention, Athens, Greece, 2016, pp. 424–432.
11. C. Ledig, V. F. J. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon, and B. Glocker, "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation", in February 2017
12. Z. Zhou, M. R. Siddiquee, N. Tajbakhsh, and J. Liang, UNet++: A nested U-net architecture for medical image segmentation, in Proc. 4th Int. Workshop on Deep Learning in Medical Image Analysis, Granada, Spain, 2018, pp. 3–11
13. M. Drozdzal, E. Vorontsov, G. Chartrand, S. Kadoury, and C. Pal, The importance of skip connections in biomedical image segmentation, in Proc. 1st Int. Workshop on Deep Learning in Medical Image Analysis, Athens, Greece, 2016, pp. 179–187
14. K. Simanyan and A. Zisserman, Very deep convolutional network for large-scale image recognition, arXiv preprint arXiv: 1409.1556, 2015.
15. Ken Masamune, Yoshitaka Masutani, Three-Dimensional Slice Overlay system with accurate depth perception for surgery, arXiv.
16. Garg, P.K. Overview of Artificial Intelligence. In Artificial Intelligence; Chapman and Hall/CRC: London, UK, 2021; pp. 3–18. [CrossRef]
17. M. Shakeri, S. Tsogkas, E. Ferrante, S. Lippe, S. Kadoury, N. Paragios, and I. Kokkinos, Sub-cortical brain structure segmentation using F-CNN'S, in Proc. 13th Int. Sympos. on Biomedical Imaging (ISBI), Prague, Czech Republic, 2016, pp. 269–272
18. O. Ranneberger, P. Fischer, and T. Brox, U-Net: Convolutional networks for biomedical image segmentation, in Proc. 18th Int. Conf. on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 2015, pp. 234–241.
19. H. Huang, L. Lin, R. Tong, H. Hu, Q. Zhang, Y. Iwamoto, X. Han, Y. W. Chen, and J. Wu, UNet 3+: A full-scale connected UNet for medical image segmentation, in Proc. 2020 IEEE Int. Conf. on Acoustics, Speech and Signal Processing, Barcelona, Spain, 2020, pp. 1055–1059
20. F. Isersee, J. Petersen, A. Klein, D. Zimmerer, P. F. Jaeger, S, Kohl, J. Wasserthal, G. Koehler, T. Norajrtra, S. Wirkert, et al., nnU-Net: Self-adapting framework for U-Net-based medical image segmentation, arXiv preprint arXiv: 1809.10486, 2018