Image of Brain Stroke Lesion segmented with X-Net, FSM and Quaternion Neural Network

Manan Parihar
200101063

Ripudaman Singh
200101090

Kartikey Gupta
200101055

Vansh Rathore
200101102

Keerat Kaur Jassal
200107036

***Abstract*—The accurate segmentation of brain stroke lesions in medical imaging is of paramount importance, facilitating early diagnosis and treatment. In this study, we present a novel approach leveraging the power of X-Net architecture and quaternions, two emerging paradigms in the field of deep learning. Our method automates the identification of stroke lesions within brain images, ultimately aiding medical professionals in their diagnostic tasks. Our findings underscore the efficacy of DSC, particularly in capturing long-range dependencies within medical imaging data. We provide insights into the dataset used for training and the technical aspects of our methodology, including data pre-processing and splitting strategies. The core of our approach involves two key networks, X-Net and X-Net with quaternions, which enable precise lesion segmentation in brain images. In the conclusion, we present a comprehensive analysis of our results, comparing them with current methods. Our study offers a promising solution for automating stroke lesion identification in medical imaging.**

I. Introduction

A. Background

The introduction sets the stage for our investigation into advanced convolutional neural network architectures. In recent years, Depthwise Separable Convolutions (DSC) have emerged as a promising approach in the field of computer vision, particularly for medical image analysis. In this forum, we delve into the intricacies of standard convolutions and DSC, exploring their implications in terms of computational power and model complexity. Our primary aim is to leverage these insights to advance the capabilities of image segmentation, with a focus on detecting intracranial hemorrhages in brain CT scans.

B. Challenges and Motivation

We encounters two pivotal challenges in the context of X-Net and medical image segmentation. Firstly, X-Net, when used without the feature similarity module (FSM), faces the risk of erroneously identifying non-lesion features as stroke lesions. For instance, it may wrongly detect bone layers as lesions, which can have critical clinical implications. To address this issue, we have integrated X-Net with FSM, which is essential for accurate feature discrimination. Secondly, the sheer number of trainable parameters in X-Net with FSM poses a computational challenge, especially when dealing with extensive medical imaging datasets. In response, we have leveraged quaternions to streamline the model's complexity while retaining its effectiveness. These challenges motivate our research, aiming to enhance the precision of medical image segmentation and ensure the reliable identification of brain stroke lesions, a critical need in the healthcare domain.

Identify applicable funding agency here. If none, delete this.

II. Related Works

A. Prior work on X-Net and Quanternion

- *X-Net: Brain Stroke Lesion Segmentation Based on Depthwise Separable Convolution and Long-range Dependencies:* This work lays the foundation for our research, introducing X-Net as a novel architecture for brain stroke lesion segmentation. It explores the use of Depthwise Separable Convolutions and their implications for medical image analysis.
- *A Survey of Quaternion Neural Networks:* This comprehensive survey provides insights into the landscape of quaternion neural networks, a crucial reference for our approach's integration of quaternions. It offers an overview of the state of the art in this domain.

B. Prior work on DSC

Deep Learning with Depthwise Seperable Convulation: The paper envisions that depthwise separable convolutions will assume a pivotal role in the future design of convolutional neural network architectures, offering Inception-like characteristics with the ease of use akin to regular convolution layers.

C. Our Contribution

Our research centers around the implementation of X-Nets, a cutting-edge convolutional neural network architecture, enriched by the integration of a feature similarity module (FSM). This extension allows us to significantly improve the accuracy of our model by ensuring that features closely related to stroke lesions are accurately identified, while minimizing false positives. In addition, we harness the power of quaternions, a specialized mathematical construct, to reduce the number of trainable parameters in our X-Nets with FSM. This approach not only streamlines the model's computational demands but also enhances its efficiency. By reducing the model's complexity without sacrificing performance, we aim to develop a more computationally efficient and reliable tool for medical image segmentation, particularly in the precise identification of brain stroke lesions

III. Proposed Work

A. Notations

I : Input Feature map
O : Output Feature map
R : Set of real numbers
HXW : Dimension of filter
C_i : number of input channels
C_o : number of output channels

B. Network Architecture

In order to have a more effective method for our objectives, we build an encoder-decoder architecture, that is called X-Net based on X-block and FSM as we can see in Figure 1

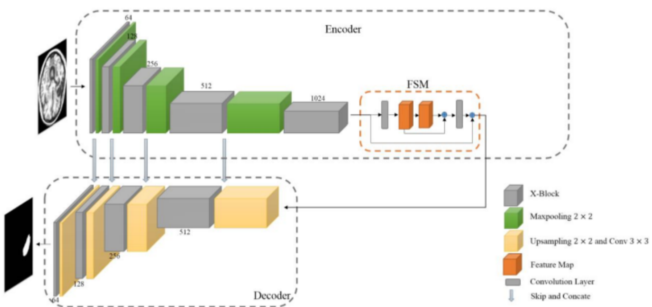

Fig. 1. Encoder-decoder architecture

- The encoder includes X-blocks and Maxpooling layers. The decoder incorporates Upsampling and convolution layers for spatial resolution recovery. The Feature Similarity Module (FSM) captures long-range spatial information to aid lesion segmentation.
  FSM is integrated into the encoder of the neural network, connecting it to the decoder.

IV. Experimental Details

A. Dataset

Dataset 1: Brain CT images with Intracranial Haemorrhage Masks We use the open-source dataset which is composed by 82 patients. For each patient there are two folders: bone folder and brain folder. For the studies that we have done, the only folder that is important is brain folder which contains approximately 30 image slices per patient, named with a number (i.e., "1.jpg"). In this folder there are some images taken in a different part of the brain. If the image taken into account is an image that illustrate a part of the brain that has a stroke lesion, the following image remark only the problem encountered that is called with the same name of the previous image following by "$_HGE_Seg.jpg$".

B. Training Details

1. X-Net : We have implemented X-Net as in Figure 1. The framework is built in which:

• we add a depthwise separable convolution layer in order to reduce the number of trainable parameters (that is less than

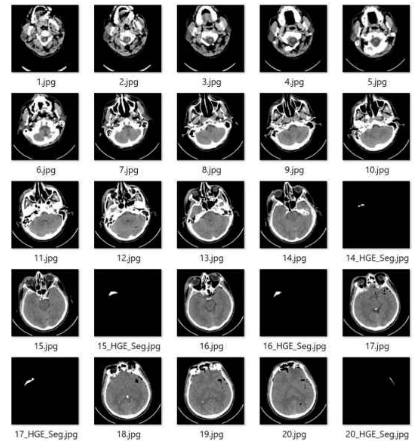

Fig. 2. Example of Dataset

what we have in the other methods) and it ensures the strength of feature extraction and representation;

• the long-range dependencies are effectively explored for brain stroke lesion segmentation.

So, U-Net is replaced by DSC (Depthwise Separable Convolution) in order to reduce the convolution kernel parameters. Specifically, the DSC is a convolution that works independently over each channel of the input feature map.

At each X-block we have:

• Input feature map: $I \in \mathbb{R}^{H \times W \times C_i}$, where C_i is the number of input channels.

• Two paths:

1) First path: The residual connection consists in 1×1 convolution layer in order to guarantee that the number of channels in output is equal to C_0.

2) Second path: Three depthwise separable convolution layers in cascade with kernel size 3×3. Convolution of size 1×1.

• Output feature map: $O \in \mathbb{R}^{H \times W \times C_0}$, where C_0 is the number of output channels. It is obtained by summing the output of the first and the second path.

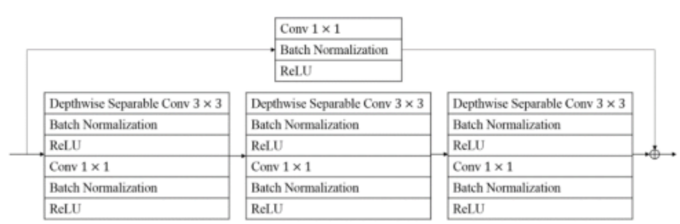

Fig. 3. X-Block

2. FSM : FSM (Feature Similarity Module) is a non-local operation. It is used to capture long-range spatial information, which contributes to the segmentation of lesions with different scales and shapes. We consider FSM as a network model that can be linked with other fully convolutional neural networks.

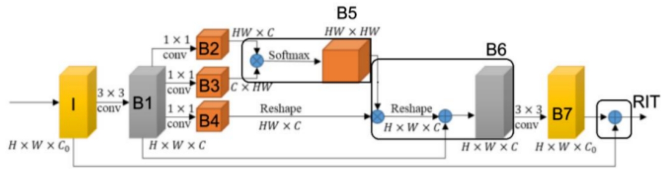

Fig. 4. : FSM: Feature Similarity Module

FSM is part of the Encoder of our neural network and in particular, allow us to connect the encoder to the decoder, as we can see in the Figure 1.

3. X-Net with quaternions Neural Network (QNN) :

We want to find a representation for the multidimensional data in order to decode the relations between the properties of observable entities.

Furthermore, in real-valued NN, since we have that the estimation of the parameters may fail to associate all the appropriate relations to all the pixels that compose a picture, we conclude that, in the latent space, the latent relation between RGB components of a pixel may not be suitably encoded.

This is where Quaternions come into picture, which are fourth dimensional and the aim objective is to process and build entities that are composed by four elements (at maximum). The black and white image were represented in the quaternion domain in the three complex channels and the real channel was set to zero. We have done this choice because we had images in black and white and not with colors. If we have had the color images, we would have put the primary color component for each channel. To implement this network, we have used PyTorch.

First, we have adapted the Depthwise Separable Convolution class in order to work with quaternions, but then the implementation of the X-Net is the same that we have done for the first model. So, we have the same structure of the X-block, but also the same structure of the Encoder, FSM and Decoder.

The crucial change is that in all this implementation we have added the forward function and we have modified the dimension of all the element in the network to fit the quaternion.

C. Baseline Methods

We have used two models for final evaluation and results comparision:

• **Model 1:** We have compared the results obtained with X-Net by using our dataset, with those in the paper;

• **Model 2:** We have compared the results obtained with X-Net model and X-Net with quaternions.

D. Evaluation Metrics

We decide to use the metrics written in the Table 4. In particular, we use:

• Dice score: $(2 \cdot Area\ of\ Overlap)/total\ pixels\ combined$

• Loss function: *BinaryCrossEntropy* + (*1* - *Dice*)

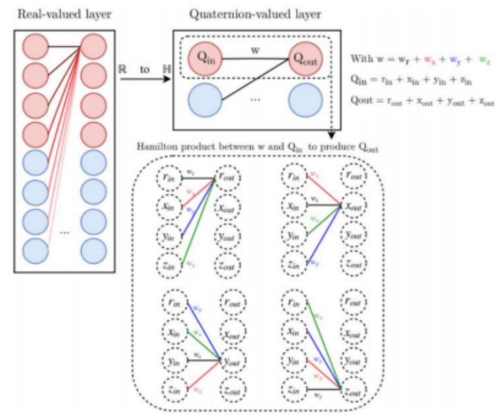

Fig. 5. Real-valued layer compared to quaternion-valued layer
“

V. Results

A. Comparison with State-of-the-art Methods

1. Quantitative Results:

To show the results, we have decided to use the loss function and the dice score. Furthermore, we have done the same training of the data by using early stopping, that is a regularization method used to avoid overfitting.

- **Model** 1 :
  We think that is interesting to compare the performance of our network with FSM and without it because with this comparison we understand the effectiveness of FSM.

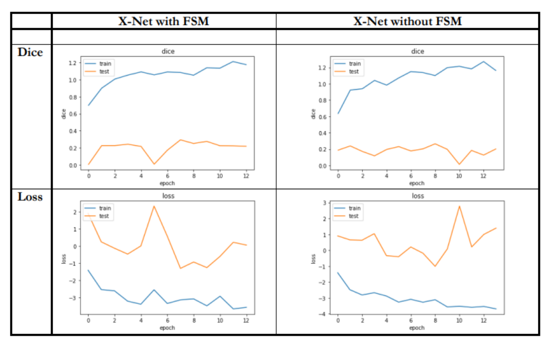

Fig. 6. Dice and loss for X-Net with FSM and without it
“

We can see from Fig 6 that the result that we have obtained are very similar, and it is a little better in the model with FSM. But it is possible that this result is the consequence of the small dimension of our dataset. We think that if the dataset is bigger, the power of FSM is much visible.

- **Model 2**:
  We have used the loss function given by: *BinaryCrossEntropy* + (*1* - *Dice*)
  For Dice metric, we have not the good results, but we can see that it tends to increase, that is positive. We can

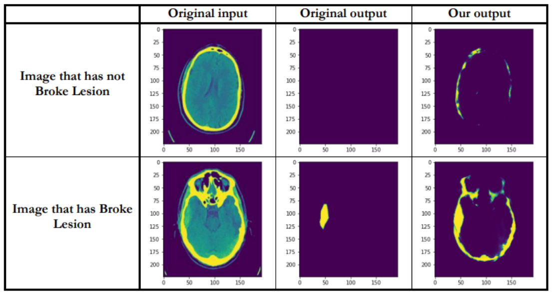

Fig. 7. Example of random samples using X-Net (with FSM)
“

have some different reasons to explain this trend. Firs, it is possible that the number of epochs that we have chosen is not sufficient for a quaternion network. The trend can be better or worse if we train it again and again. Secondly, it is probably that our dataset is not big enough to be trained with QNN. Finally, this structure of the network might not adapt to quaternions or the choices that we have done to represent the images in the quaternion domain is not adapt.

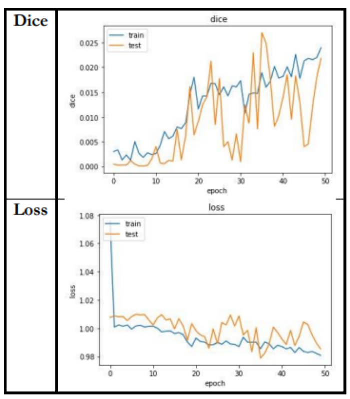

Fig. 8. Dice and loss for X-Net with quaternions
“

With the quaternions the network tends more to return black images. For this we have tried to put only the dice as a loss function to try to push the network not to return black images. We have a different behaviour because with the dice every time we use the colors, while with the loss function it does not.
The result that we have obtained (Fig 9) in the image that has broke lesion is good, because it highlights the problem that the patient has. But for the image that has not the broke lesion, the result is not good, because we have a white point so marked.

2. Qualitative and Quantitative Analysis:

- As we can see from Fig 6 and Fig 10, in this case we see a big difference. In fact, if we use X-Net without FSM, even if we cut the bone from the end image, we will not have a perfectly classification of the lesion because the problem is that we have a big area that are marked in

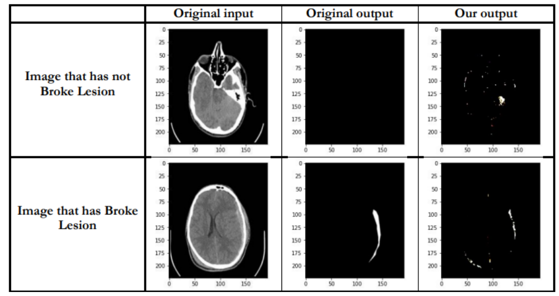

Fig. 9. Example of random samples using X-Net with quaternions
“

our output and not only those of the bones. So, we can conclude that the work of FSM is very important.

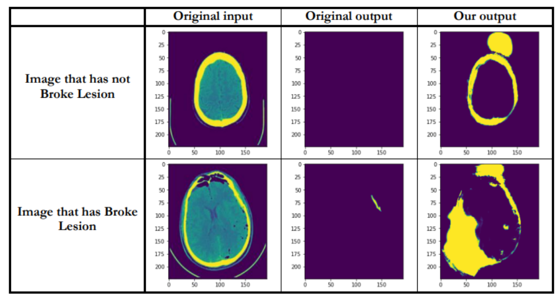

Fig. 10. Example of random samples using X-Net without FSM
“

- Comparisons between our results and paper results In the paper, the authors have used ATLAS dataset that contains 43281 images, instead our dataset contains 2501 images. So, if we compare the dimensions of the two datasets (ATLAS and ours) and given only the 6% of theimages that have ATLAS dataset, we have obtained the final dice = 0.2188 that is good compared to 0.4867 (paper).

Model	Dice	#Parameters
ResUNet	0.4702	33.2M
DeepLab v3+	0.4609	41.3M
2D Dense-UNet	0.4741	50.0M
PSPNet	0.3571	48.1M
SegNet	0.2767	29.5M
U-Net	0.4606	34.5M
X-Net (paper)	0.4867	15.1M
X-Net (ours)	0.2188	15.1M

Fig. 11. Comparison between our results and paper results
“

- Comparisons between Model 1 and Model 2 As we can see in Fig 12, we have that the number of parameters in Model 2 is almost double respect to which of the Model 1. This is a problem because usually the number of parameters should decrease in a quaternions network. Instead, for Dice value, we have that Model 1 is better. We have a so low dice value for Model 2 for the reasons that we have explained above.

Model	Dice	#Parameters
Model 1	0.2188	15.1M
Model 2	0.0230	26.5M

Fig. 12. Comparisons between Model 1 and Model 2
“

VI. Conclusion

In building this project we have encountered some obstacles, in particular with the Dataset. In fact, we were intent on using the ATLAS dataset that is the same that has been used in reference papers; but for some reasons we are not using this. We think that in the future, we can use the original dataset ATLAS and do more comparisons. Furthermore, we may also think to extend this project, for example by using the transformations (i.e., wavelet) in order to use the quaternions better and not in the simplest approach. Otherwise, we can also extend it by using a multimodal approach in order to analyse similar images linked to the same measure. In any case, we are satisfied of our work .

References

[1] Kehan Qi, Hao Yang, Cheng Li, Zaiyi Liu, Meiyun Wang, Qiegen Liu, Shanshan Wang:X-Net: Brain Stroke Lesion Segmentation Based on Depthwise Separable Convolution and Longrange Dependencies. (2019).

[2] Titouan Parcollet, Mohamed Morchid, Georges Linarès: A survey of quaternion neural networks (2019).

[3] François Chollet, Googl Inc.: Xception: Deep Learning with Depthwise Separable Convolutions (2017).

[4] https://www.kaggle.com/vbookshelf/computed-tomography-ct-images [for dataset 1]

[5] Sook-Lei Liew, Julia M. Anglin, [. . .] Alison Stroud A large, open source dataset of stroke anatomical brain images and manual lesion segmentations [for dataset 2: ATLAS]

[6] Keras.io