Journal of Intelligent & Fuzzy Systems

Brain tumor MRI Swin Transformer Classification using Laplacian Generated Super Resolution Image

--Manuscript Draft--

Manuscript Number:	IFS-237556
Full Title:	Brain tumor MRI Swin Transformer Classification using Laplacian Generated Super Resolution Image
Article Type:	Research Article
Section/Category:	Regular Papers Section
Keywords:	Laplacian Super Resolution Network (LAPSRN), Laplacian Generative Adversarial Network(LapGAN), Swin Transformer, super resolution
Abstract:	To better classify and identify brain malignancies, we want to understand how super resolution improves medical brain MRI image discrimination. To test this, a class-based super-resolution pipeline is introduced. An picture is created from multiple areas, and each region must be reconstructed differently. The processing network for region-wise extracted characteristics may vary. Thus, each deconstructed picture may be handled with a distinct super-resolution network based on its restoration capacity. LAPSRN, a novel super-resolution pipeline based on classes, is presented using this approach. Consider super resolution as a preprocessing step for picture categorization. This class-based super-resolution uses a convolutional layer and the LAPGAN network. Two new loss functions guide the LAPGAN and convolution layers. This pipeline improved tumor classification with 99% accuracy using the IXI and brain tumor datasets.

Journal of Intelligent & Fuzzy Systems
IOS Press

Brain tumor MRI Swin Transformer Classification using Laplacian Generated Super Resolution Image

Sangeetha G[a,*],Vadivu G[b]
[a] *Department of Data Science and Business Systems, SRM Institute of Science and Technology, Kattankulathur, Chennai, India*
[B] *Department of Data Science and Business Systems, SRM Institute of Science and Technology, Kattankulathur ,Chennai, India*

Abstract. To better classify and identify brain malignancies, we want to understand how super resolution improves medical brain MRI image discrimination. To test this, a class-based super-resolution pipeline is introduced. An picture is created from multiple areas, and each region must be reconstructed differently. The processing network for region-wise extracted characteristics may vary. Thus, each deconstructed picture may be handled with a distinct super-resolution network based on its restoration capacity. LAPSRN, a novel super-resolution pipeline based on classes, is presented using this approach. Consider super resolution as a preprocessing step for picture categorization. This class-based super-resolution uses a convolutional layer and the LAPGAN network. Two new loss functions guide the LAPGAN and convolution layers. This pipeline improved tumor classification with 99% accuracy using the IXI and brain tumor datasets.

Keywords: Laplacian Super Resolution Network (LAPSRN), Laplacian Generative Adversarial Network(LapGAN), Swin Transformer, super resolution.

1. Introduction

The technological advancements made in medical imaging play a significant role in identifying the most complicated diseases associated with vital organs of the body. From X-rays to MRI, identifying tumors in brain images has evolved; however, the quality of the images produced is highly dependent on various factors that are contributed by environment, equipment, and budget [1]. Under certain conditions, where the human is expected to get exposed to lower radiation, this always results in images of lower resolution. These low-resolution images are not effective in clinical diagnosis or classifying brain tumors [2]. Meningiomas, for example, one of the primary brain tumors, are difficult to detect as they are tiny. In such situations, with a low-resolution image, some parts of the tumor are not visible [3]. Though using a high-resolution image is an ideal choice for accurate classification, this is not always possible in practice. Therefore, transforming the low-resolution image to a higher quality may become useful in detecting the less salient objects and thus contribute to better classification of the brain tumors.

There are various methods available for quality enhancement, like denoising, colorization, and super-resolution. In our study, we focused on understanding the relationship between the quality enhancement method and its contribution to detecting and classifying brain tumors. So, every method's recent work using deep learning is reviewed to understand its contribution to the classification of brain tumors, and its results are presented in Figure 1. As shown in Figure 1, denoising methods using Generative Adversarial Network (GAN) [6], diffusion model [5], and denoising autoencoders [4] have shown significant results in the detection of tumors. However, due to its highest computation cost, ringing artifacts and, most importantly, classification done based on the denoising procedure don't leverage the local visual information. Thus, though the detection results are good with the use of denoising procedures, their contribution to the classification is limited due to the non-consideration of local information. Similar outcomes are observed with the colorization method as well, for example, based on the results shown in [7 and 8]. Though colorization aided in enhancing the tumors, its inherent semantic information was not considered for the classification to happen correctly.

The final technique that obtains higher-quality images from low-quality images is termed super-resolution, and

some of the recent results exhibit the importance of obtaining high-quality images for medical diagnosis [9, 10]. Though various methods are proposed for improving the aesthetic quality of the images [11], super-resolution focused on class separation is minimal, and this frames our problem statement.e technological advancements made in medical imaging play a significant role in identifying the most complicated diseases associated with vital organs of the body. From X-rays to MRI, identifying tumors in brain images has evolved; however, the quality of the images produced is highly dependent on various factors that are contributed by environment, equipment, and budget [1]. Under certain conditions, where the human is expected to get exposed to lower radiation, this always results in images of lower resolution. These low-resolution images are not effective in clinical diagnosis or classifying brain tumors [2]. Meningiomas, for example, one of the primary brain tumors, are difficult to detect as they are tiny. In such situations, with a low-resolution image, some parts of the tumor are not visible [3]. Though using a high-resolution image is an ideal choice for accurate classification, this is not always possible in practice. Therefore, transforming the low-resolution image to a higher quality may become useful in detecting the less salient objects and thus contribute to better classification of the brain tumors.

There are various methods available for quality enhancement, like denoising, colorization, and super-resolution. In our study, we focused on understanding the relationship between the quality enhancement method and its contribution to detecting and classifying brain tumors. So, every method's recent work using deep learning is reviewed to understand its contribution to the classification of brain tumors, and its results are presented in Figure 1. As shown in Figure 1, denoising methods using Generative Adversarial Network (GAN) [6], diffusion model [5], and denoising autoencoders [4] have shown significant results in the detection of tumors. However, due to its highest computation cost, ringing artifacts and, most importantly, classification done based on the denoising procedure don't leverage the local visual information. Thus, though the detection results are good with the use of denoising procedures, their contribution to the classification is limited due to the non-consideration of local information. Similar outcomes are observed with the colorization method as well, for example, based on the results shown in [7 and 8]. Though colorization aided in enhancing the tumors, its inherent semantic information was not considered for the classification to happen correctly.

The final technique that obtains higher-quality images from low-quality images is termed super-resolution, and some of the recent results exhibit the importance of obtaining high-quality images for medical diagnosis [9, 10]. Though various methods are proposed for improving the aesthetic quality of the images [11], super-resolution focused on class separation is minimal, and this frames our problem statement.

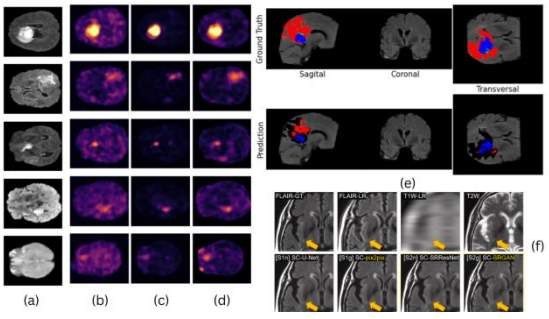


Fig 1:Demonstration of various approaches that are used as preprocessing before the classification: (a) represents the true image; (b) GAN-based denoising; (c) diffusion-based denoising; (d) variational auto-encoder-based denoising; (e) represents the colorization approach that shows variation between the truth and prediction image; (f) super-resolution-based reconstruction. Images are inferred from [6], [8], and [11].

Motivation of this work:

The major motivation for this work comes from understanding the problems associated with denoising and colorization to make the classification problem perform better. Since well-resulting classification networks have already been developed [11–18], our main goal here is to understand the contribution of the super-resolution method when used as a preprocessing step when attempting to train and test the MRI brain images. The general loss functions used for performing the super-resolution are per-pixel loss or perceptual loss. During the training phase, per-pixel loss aims to minimize the per-pixel difference between the prediction and the original image. Perceptual loss in contrast gets the difference between the prediction and the original image from a pretrained convolutional neural network. Both methods possess their pitfalls, as they concentrate on reconstructing an image that holds aesthetic importance. Irrespective of this pitfall, we frame our motivation from the questions below that frame our hypothesis.

1.What is the contribution of the superresolution technique in classification if it is used as a preprocessing phase in brain MRI image classification?

2.What changes does the super-resolution technique make if the aesthetic concentrated super-resolution network is modified to use class separation?

Our contribution and Novelty:

Our contribution and novelty come when we try to answer both questions that are put forward under the motivation section. This work proposes a LAPGAN-based transformer network after examining the state of

the art and understanding the common concerns associated with the super-resolution technique. It draws influence from other similar works mentioned in [20, 21], and after careful examination of its limitations, which mention that the high resolution did not contribute to better classification, we aimed for a framework that could perform the super-resolution based on class separation. So here we have used LAPGAN, which trains the discriminator in the process of generating fooling samples. These samples can become efficient for training the classification network. Thus, the major contributions of this work are as follows:

•To address the issue of the aesthetic concentrated super-resolution, a novel method that can perform super-resolution based on class separation is proposed.

•The LAPGAN network is modified to take the down-sampled input images instead of the random seed so that the super-resolution is done based on class separation.

•Class separation-based losses are introduced that guides the network for utilizing every branch of the network equally.

•The SWIN transformer network has been used as the base framework for performing the brain MRI classification. This is the first time that this network has been used for this dataset.

•The evaluation is performed using the IXI, brain tumourdatasets, and a comparative analysis is done to understand the effect of the LAPGAN technique and its contribution to the classification.

The rest of the paper is organized as follows:

Section 2 reviews some of the super-resolution methods adopted in the literature, followed by the proposed methodology in Section 3. Section 4 presents the experimentation and results, and Section 5 arrives at the final discussion and future directions.

2. Related Work:

Since our focus is to understand how well the super-resolution technique contributes to the classification results when used as a pre-processing step, it is ideal to understand the state of the art in super-resolution techniques. Thus, here we try to understand the various methods that are used for obtaining super-resolution images. The two major classifications of superresolution are the single-image superresolution method (SISR) and multi-image superresolution (MISR). As their names indicate, the basic variation between them lies in the number of input images for performing the super-resolution construction. In SISR, super-resolution reconstruction is obtained from a single image, whereas in MISR, super-resolution reconstruction is done using multiple low-resolution images of the same scene [20]. As far as brain MRI superresolution is considered, possessing multiple images is practically impossible here; the focus is only on SISR-based reconstruction. SISR-based reconstruction is further classified into three categories: interpolation-based SR, reconstruction/edge-based SR, and statistical/learning-based SR.

Interpolation based Super resolution:

This sampling-theory-based technique seeks to rebuild the unknown pixels from the nearby values. Although there are extensions of bicubic interpolation, as indicated in [22], the results of base bicubic are used for the meta-analysis. Bicubic image interpolation is a popular technique used in brain MRI scans. This bicubic strategy, which can yield sharper images than the other interpolation algorithms, takes into account the 16 pixels that are closest to the unknown pixels. The issue here, though, is the aliasing at the margins caused by the blurring effect it creates at high-frequency details. Bicubic is employed in our experimentation to make comparisons anytime there is upscaling because it is the best illustration of the interpolation-based superresolution technique.

Edge based Super resolution:

With the application of prior knowledge on the edges to eliminate edge artifacts, the goal, in this case, is to attain super-resolution in the upsampled images. It is crucial that the upsampled image closely reflects that of the low-resolution input image when applying the prior knowledge [23]. Backpropagation makes sure of this, and the kind of priors utilized affects the sharpness of the image. A smooth contour prior, which accomplishes reconstruction by interpolating along the contours of strong spatial edges, is one such prior.

Learning based Super resolution:

The most popular type of super-resolution technique nowadays is learning-based, which learns the differences between high- and low-resolution images. The dataset of high-resolution photos provides the knowledge of high-resolution images, which is then applied to the reconstruction of low-resolution images. The reconstruction will be better if there is a larger resemblance between the trained high-resolution image and the tested low-resolution image. Convolutional neural networks (CNN) and their derivatives have shown good results in the state-of-the-art. To determine if our framed hypothesis is addressed in the recent studies, some of the most recent publications are studied.

One work mentioned in [26] proposes a feedback-based network called FNSAM that uses the recurrent neural network principle for building the feedback network. The feedback here basically inputs the high-level information to the lower layers so that the optimization happens at the shallow. The extraction of relevant features from this feedback is obtained by the self-attention module. CNN is modified to include varied sizes of kernels that can capture the different view ranges and different types of convolution operators are

used so that varied features are extracted [27]. A new denoising diffusion probabilistic model (DDPM) for achieving super-resolution in brain MRI images is exhibited in [28], where the architecture is modified to include the self-attention mechanism in all resolution layers instead of giving only 16 × 16 layers. Generative adversarial network (GAN) has also been used for super resolution [29].

As far as the learning-based super resolution methods are considered, a wide variety of networks ranging from CNN, RNN, GAN, diffusion, and transformer models are used. Like CNN, GAN also plays a significant role in obtaining super-resolution images, and one of the seminal works is proposed in [39], where the super-resolution GAN is modified in terms of its loss and architecture. The variation in resolution happened because of the residual block without batch normalization. The texture reconstruction and brightness were done by the discriminator using real values, and this has achieved state-of-the-art status while performing super-resolution reconstruction on real images. Similar work, BSRGAN [40], uses a degradation model that incorporates blurring at multiple random places, downsampling, and noise degradation to try and solve the diverse and dispersed blurring artifacts. The noise from the Gaussian noise and the Gaussian kernels were used to create the blurring. Consequently, the superresolution produced by this technology was better. SwinIR transformer [41] performs the super resolution based on residual swin tansformer blocks have also achieved good super resolution results.

In most of these methods, the input used is extremely low-resolution images; however, our interest is to consider a decent-resolution image and perform reconstruction on it as a first step. Though our hypothesis is to understand the contribution of super-resolution to that of the classification, the considered works gave us the idea of various super-resolution techniques that exist with brain MRI images, and separate context regarding the classification is presented in [11–18], and here also there is the involvement of CNN, GAN, and transformers. As will be discussed in the further sections, our approach falls under the learning-based super-resolution that uses the SWIN transformer network for classification and performs the class based super resolution using the LAPGAN.

3.Proposed System:

The overall objective of this work is to determine whether the super resolution methods help in the improvement of brain MRI classification. Experimentation is carried out using public brain classification datasets like IXI and the brain tumor dataset. The hypothesis framed is to determine the relationship between super-resolution and tumor classification, and this understanding is achieved with the novel pipeline utilizing the LAPGAN and Swin transformer networks. Figure 2 presents the overall architecture of the proposed methodology.

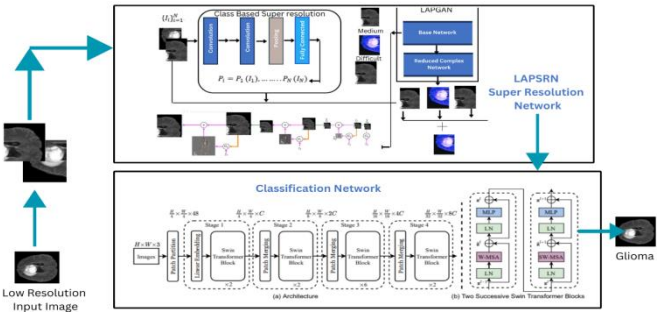


Figure 2: The overview of the proposed class-based separation of the images using the Laplacian Generative Adversarial Network (LAPGAN) and this SR image is handled by the SWIN Transformer network for the classification. The architecture of LAPGAN [31] and Swin transformer are used without modification, however the class-based separation is performed prior the sub images are fed into the LAPGAN.

As shown in Figure 2, this process of classifying the brain tumor images is carried out as a two-stage process, in which the first stage includes the super-resolution network named Laplacian Super Resolution Network (LAPSRN) and the second stage includes the Swin IR transformer network for performing the classification.

Laplacian Super Resolution Network (LAPSRN):

Class based super resolution is the new solution for the problems associated with aesthetic concentrated super-resolution. This resolution network takes the decomposed low-resolution image as input that if the entire image is defined I_LR then the decomposed images are the sub samples of the I_LR. This sub samples are considered as different classes as they represent different parts of the overlapping images. Thus, the input image I_LR after overlapping can be represented as shown in equation (1)

I_LR={I_(1),I_2……..I_N }……(1)

This equation (1) can also be represented as {⟦I_i}⟧_(i=1)^N. Each of the sub image is passed into the class-based network that includes 4 convolution layers, an average pooling layer and a fully connected layer that generates a probability vector as mentioned in equation (2)

P_i=P_(1)(I_1),……..P_(N)(I_N ……(2)

The one with the maximum probability is chosen to be indexed along with the Laplacian network for the formation of Laplacian pyramid. This LAPGAN super resolution network is designed as possessing

independent branches for each sub image. In this manner, the super resolution network consists of M branches and this can be denoted as {⟦lap⟧_sr^j }_(j=1)^M. The LAPGAN network is adopted as the base network and this considered as the most complex branch and the other branches that are obtained is through the reduction of the layers. In this work, the base network was with 16 layers and then the complex less networks are with 8,4. The minimum network complexity by reducing the number of layers was 4.

The class-based network instead of relying on the labelling of the images, they attempt to perform its classification based on the reconstruction capability. Every sub image is processed by the branches of the LAPGAN network to understand the difficulty in reconstruction. Hence this reconstruction becomes class based super resolution instead of just focusing on the aesthetic component. Figure 3 shows the class based Laplacian super resolution network. As shown in figure 2, the sub-images reconstruction and the probability values are multiplied for the super resolution image. Thus, the final super resolution image is obtained using the equation (3)

$$I\_super=\sum\_{(i=1)}^{MP\_i} ⟦lap⟧\_sr^j \quad \text{.....(3)}$$

The image loss is used as a constrain for I_super from which the probability values are obtained.

SWIN transformer classification:

The SWIN transformer network [32] is used as the classification network. As shown in Figure 1, the output of the super resolution image was supplied as an input to this transformer block for performing the classification. Like another transformer, here the image is basically split into patches and they are used for performing the feature extraction. Here in our implementation, we have used the image with three channels and size 224 × 224. The dimension of each patch 4 × 4 and the number of features in one patch is 4 × 4 × 3 =48. The total number of patches in the whole image is 224/4×224/4 = 3136. The projection of this feature is performed using the embedding layer. Then the construction of stage 2,3 and 4 of the patches are obtained from stage 1 and each subsequent stage are formed by performing the merging of the patches. As the network grew deeper, the number of patches are reduced by the patch merging layers. So at stage 1, the number of patches was 3136 and at stage 2 would be 224/8×224/8 and at stages 3 and 4 the number of patches was reduced to 224/16×224/16 and 224/64×224/64 . In this manner, the number of patches is reduced as the network grew deeper. All the stages jointly contribute for the better classification of the image with appropriate features. The multihead attention here is the shifted window mechanism that is followed by the 2-layer multiperceptron with GELU non-linearity. Residual connection applies after each multihead self-attention. Since we have used the same architecture mentioned in [32] a detailed explanation of the same is not mentioned here. However, the ideas about the embedding layer are only mentioned here.

Loss Functions:

For the proposed system, the loss function includes the derivative of the L1 loss Charbonnier penalty function, and specific to the system we have defined additional two losses that are specific to the class-based super-resolution. The new losses that are defined include the class loss and average loss. These three losses are designed in such a way to enhance the reconstruction quality based on the classes and also since we have fed this in multiple branches of the LAPGAN an average loss is introduced to ensure that at each branch the sub-images are treated equally. Thus, the overall loss function can be defined as shown in equation (4).

$$⟦Loss⟧\_total= ⟦loss⟧\_Char+ ⟦loss⟧\_class+⟦loss⟧\_avg \quad \text{.....(4)}$$

Where ⟦loss⟧_(char)is the Charbonnier loss function , ⟦loss⟧_class is the class loss function and the ⟦loss⟧_avg is the average loss function. These losses are detailed below

Charbonnier loss function (⟦loss⟧_(char)) :

The image loss considered here is the Charbonnier loss function instead of the MSE as they are faster and it is a differentiable variant of the L1 loss and thus can handle outliers well. Since our objective is to generate multi class images, the prediction will involve the sub-band residual images hence a deeper supervision is required to guide the model appropriately and hence ⟦loss⟧_(char)is considered. The ⟦loss⟧_(char)is defined using the equation mentioned as (5)

$$⟦loss⟧\_(char)=1/N \sum\_{(i=1)}^{N} \sum\_{(l=1)}^{L} \rho(y\_l^i-y\_l^{(i')}) \quad \text{.....(5)}$$

Where N denotes the number of samples and l represents the number of levels.

Class based Loss:

As mentioned in figure 3, here we aim to perform a class based super resolution and that generates a probability distribution of the class images. This loss is included in such a way that the gap between the different classes is maximum so that it can be as close to 1. And hence this loss is formulated as shown in equation (6)

$$⟦loss⟧\_class=- \sum\_{(i=1)}^{(M-1)} \sum\_{(j=i+1)}^{M}|P\_i(I)-P\_j(I)| \quad \text{.....(6)}$$

Average Loss:

This loss is included because the LAPGAN network is

designed to be of multiple branches with a reduction in complexity. Irrespective of the class of the image, every image must get equal opportunity in the network without getting restricted to the base/ complicated network. Since the base network can give results at a faster rate every image is not restricted to the base network and if this is done so, there is no concept of class-based super-resolution. To address this, it is important to ensure that every branch in the LAPGAN super-resolution network gets an equal chance to be selected. Thus the average loss defined here will constrain the class results and this loss is formulated using equation (7)

$[\![loss]\!]\_avg= \sum\_(i=1)^M \sum\_(j=1)^B P\_(i) (I\_j)- B/M$……..(7)

As mentioned in equation 7, the number of sub-images getting passed through this LAPGAN during the training phase should be the same. To ensure this, the probability $\sum\_(j=1)^B P\_(i) (I\_j)$ of identifying the number of sub- images and using B/M the average number of samples for each class is identified ensuring equal opportunity for the branches for super resolution.

4. Experimentation:

This experiment is carried out to determine the validity of the framed hypothesis and the contribution of the suggested system to the validity of the hypothesis. Therefore, the experiment is carried out through several steps, which are as follows:

The Swintransformer is used to classify the actual dataset without utilizing any super-resolution techniques, and the results are seen.

After using the suggested superresolution technique, the classification outcomes are seen via the swin transformer.

The state-of-the-art comparison of different super-resolution strategies and how they affect classification results is completed.

In this section, we'll go through the fundamental preparations made for each stage, the training method employed, the dataset, and the qualitative and quantitative outcomes.

Training and Testing Data:

We have used the IXI dataset [33] and the brain tumor dataset [34] for training. For using this data to perform super-resolution, the training data is constructed considering the classes. For this, the actual images are downsampled by using bicubic downsampling with a scaling factor of 0.8 and 0.9 for the input low-resolution images. Dense cropping is carried out for the creation of the sub-images, and this is done because we do not want the LAPGAN to be fed with a random feed; instead, the feeding will be from the image itself. These sub-images are classified into three classes: hard, medium, and easy, according to the reconstruction complexity, which is identified using the PSNR values. A total of 1311 images were cropped into 1.39 million sub images with a size of 32×32. A similar approach is used for the IXI data as well. The testing set is formed with a combination of the images from the IXI and brain tumor dataset and the choosing of the samples was random and a total of 500 images was used as the test set. While performing testing, the low-resolution images are cropped and finally the super resolution images are joined together based on the PSNR values.

Training Setup:

For the classification task, the training did not involve any pretrained models. The model was trained using the train set with an Adam Optimizer for 100 epochs using the cosine decay learning rate schedular and a weight decay of 0.03. The batch size and initial learning rate are 1024 and 0.001 respectively.

The training setup for the LAPSRN is carried out with the pretraining of the sub-sample images on the LAPGAN network and for this, the optimizer used is Adam with a learning rate of 0.001. The learning rate is adjusted using cosine decay learning. Then the class-based super-resolution network is trained for the loss functions. The batch size of 100 is used for training this for 200 epochs. Then both the network's class-based network and the LAPGAN network are trained jointly without making any changes to the setting. Now that both the super-resolution and the classification network are trained the implementation is built using Python and training using NVIDIA 2080Ti GPUs.

Result Analysis without using Super resolution:

As mentioned earlier, since this work concentrates in the verification of the framed hypothesis, it is important to understand the classification results without using the super resolution technique. While trained with swin transformer from scratch using the brain tumour dataset images, the training and testing loss along with the obtained accuracy is plotted for 10 epochs in figure 4.

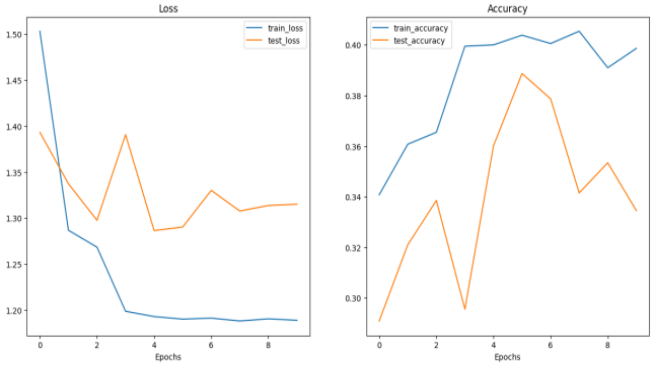


Fig 4: Plots showing the training and testing loss and accuracy for the brain tumour dataset without pre trained models.

According to Figure 4, both the training and testing accuracy were found to be between 50% and 60%, making the prediction accuracy quite poor. As a result, the pre-trained ResNet 50 is utilized to increase prediction accuracy, and Figure 5 shows the related training and testing loss and accuracy. The loss and accuracy have significantly improved, as shown in Figure 5, and without applying any super-resolution techniques, the prediction accuracy was 72%. Therefore, it is clear from these result analyses that the quality enhancement technique is crucial for better classification outcomes.

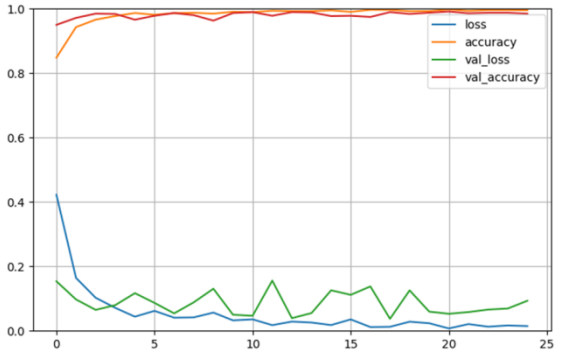


Figure 5: Plots showing the training and testing loss and accuracy for the brain tumour dataset with ResNet50.

The different forms of the Swin transformer is also verified for its accuracy value with the Resnet 50 pretrained model and the results are shown in table 1.

BackBone Network	Image Size	Number of parameter	FLOPS	Throughput	Accuracy
Swin-T	224×224	20M	3.7G	892.1	74.1%
Swin-S	224×224	46M	6.2G	402.1	70.5%
Swin-B	224×224	80M	17.2G	262.4	73.6%
Swin-B	384×384	80M	32.4G	56.2	75.2%

Table 1: Comparison of different swin transformer backbone on ResNet 50 classification

SwinLap with other super resolution methods:

The SwinLap framework formed is not specific to the brain MRI images. So, this framework can be incorporated for any deep learning network and for any image dataset. Though it is applicable to every dataset, here we are not making a comparison with other image dataset, the quantitative and qualitative comparison is done with the brain MRI image dataset. To demonstrate the effectiveness of this approach, some of the best performing super resolution networks like ESRGAN [35], BSRGAN [36] and SWIN [37]. The qualitative results obtained in comparision with that of these models are shown in the figure 6.

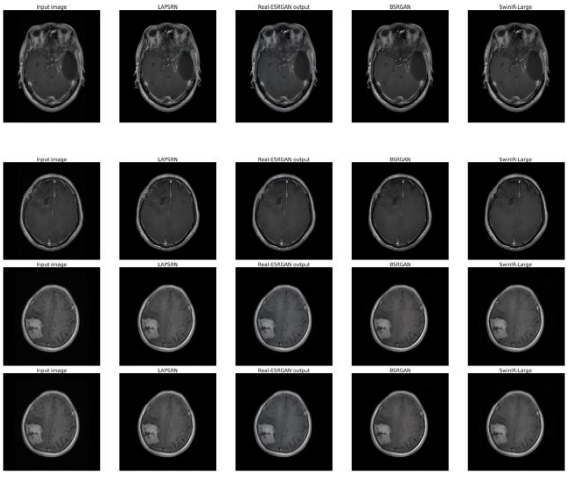


Figure 6: Demonstration of the input image after it is up sampled using the ESRGAN [35], BSRGAN [36] and SWIN [37] and the proposed LAPSRN method

From the visualization results, it is clear that the proposed super-resolution technique can perform super-resolution better than the state-of-the-art. However, achieving a better resolution is not the focus of this study, so this super-resolution image when it is used for training the classification network, its effectiveness in handling the classification is the focus of this study to justify the framed hypothesis.

A quantitative evaluation is also done in this section to exhibit the superiority of the class based super resolution method. Table 2 presents the quantitative PSNR and SSIM values obtained for the brain tumor and IXI dataset.

Dataset	**IXI**		**Brain Tumour**	
Scale	**4**		**4**	
Metric	**PSNR**	**SSIM**	**PSNR**	**SSIM**
Bicubic	24.32	0.7642	29.27	0.8945
ESRGAN [35]	28.32	0.8231	28.32	0.9212
BSRGAN [36]	28.76	0.8324	26.43	0.9321
SWIN [37]	29.45	0.8562	27.82	0.9342
Guided Diffusion [38]	30.21	0.8876	33.45	0.9561
LAPSRN(Ours)	**37.82**	**0.9621**	**39.67**	**0.9721**

Table 2: Quantitative results on IXI dataset and brain tumour dataset with scale 4 in terms of the mean PSNR and SSIM. These results achieved by our model is shown in bold.

From the qualitative and quantitative results obtained the proposed methodology is exhibiting its superiority in performing the super resolution. Now it is important to understand this super resolution image is really helpful in improving the classification results and this is analysed in the next section.

LAPSRN role in better classification:

This analysis is done with the pre-trained models ResNet50, and ImageNet 22K along with the LAPSRN-generated super-resolution images. The classification results are evaluated quantitatively using the plots of training and testing accuracy. Along with this, the confusion matrix plotted is also presented in this section as Figures 7,8 and 9.

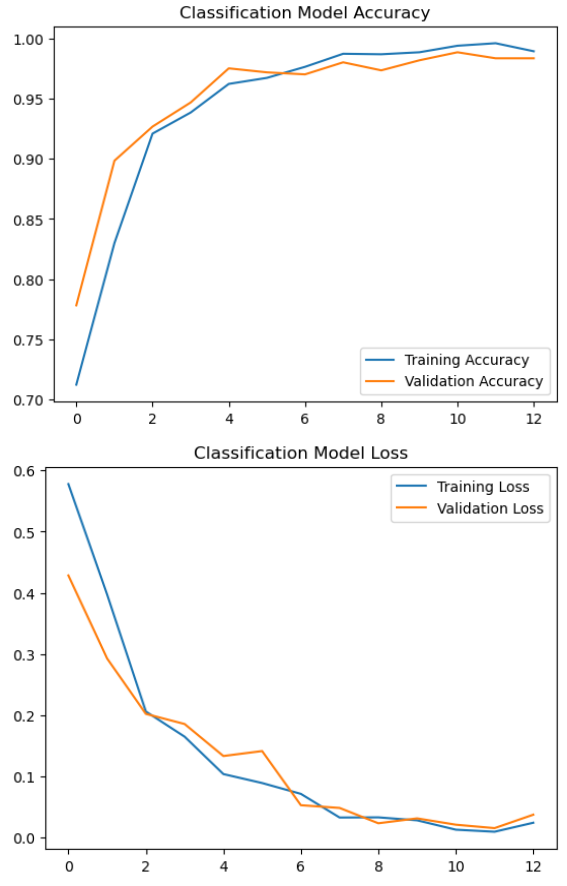
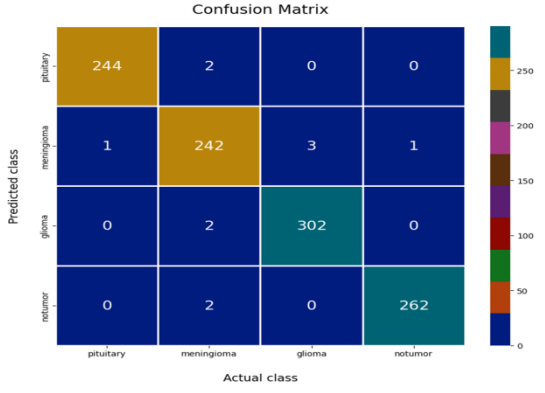


Figure 7: Plots showing the training and testing loss and accuracy for the brain tumour dataset with ResNet50 after.

As seen in figure 7, the training as well as the testing loss and accuracy is improved showcasing the improvement with prediction, and the classification accuracy is shown in figure 8.



	precision	recall	f1-score	support
glioma_tumor	1.00	0.99	0.99	246
meningioma_tumor	0.98	0.98	0.98	247
no_tumor	0.99	0.99	0.99	304
pituitary_tumor	1.00	0.99	0.99	264
accuracy			0.99	1061
macro avg	0.99	0.99	0.99	1061
weighted avg	0.99	0.99	0.99	1061

Table 3: Classification results obtained after training the transformer network with the super resolution images obtained from the LAPSRN.

	precision	recall	f1-score	support
glioma_tumor	0.86	0.19	0.30	100
meningioma_tumor	0.75	0.86	0.80	115
no_tumor	0.56	0.93	0.80	105
pituitary_tumor	0.86	0.76	0.81	74
accuracy			0.69	394
macro avg	0.76	0.68	0.65	394
weighted avg	0.75	0.69	0.65	394

Table 4: Classification results obtained without training the transformer network with the super resolution images obtained from the LAPSRN.

As per the classification results shown in Figures 8 and 9, there is nearly a 25% increment in the classifier's efficiency in classifying the tumor images. This quantitative analysis thus exhibits the inclusion of the super-resolution images that are obtained from the proposed LAPSRN contributes to the enhancement of the classification results. With this analysis, we can give justifications regarding the framed hypothesis.

Ablation Study:

This ablation is done to understand whether the inclusion of loss functions contributes to the PSNR and Flops during the training. So the class loss is removed first from the loss function equation and then the training process is carried out. With the removal of class loss, the model was not able to converge because the class loss is constrained by the average loss and since this is removed, the average loss made the probability vector have the same value for all the classes that made the model not to converge. Since the class loss was introduced to manage the equal opportunity for the images in the LAPGAN network since it was removed, the LAPGAN was fed with random fed which is one of the problems that we wanted to address. This denotes the inclusion of class loss did contribute to the LR input image being fed into the LAPGAN instead of a random seed. This signifies the importance of the class loss in our model and Figure 10 shows the nonconvergence scenario without the class loss function.

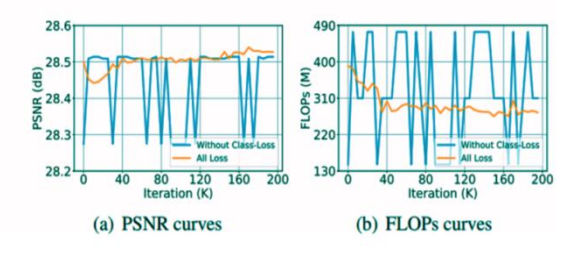
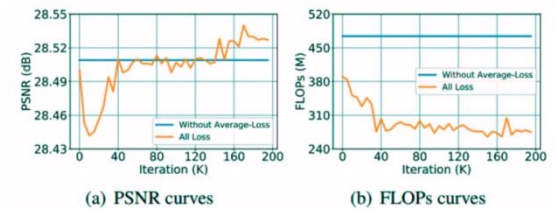


Figure 8: Training curves comparison of LAPSRN with and without the class loss function.

Next, the average loss function is removed from the loss function to understand its effect on the PSNR and FLOP. This loss is mainly introduced in LAPSRN to ensure the branches are treated equally. Thus, the removal had an impact and all the images were assigned to the most complex network which is our base network. The effect of average loss removal is shown in Figure 9.



(a) PSNR curves (b) FLOPs curves

5. Conclusion:

In this work, through our experimentation, we show that our original hypothesis is proved, for the method and scenario that we have used. The hypothesis framed was to understand the role of super-resolution images in the classification of brain tumors. To test this hypothesis, we proposed a new superresolution network based on the classes without performing superresolution based on aesthetic quality. The super-resolution network key idea is to feed the LAPGAN with low-resolution images instead of the random feed and these sub-images are identified as different classes of images based on its reconstruction ability. According to the reconstruction ability, the images are processed in different branches of the LAPGAN network. By training them with and without super-resolution images, the SWIN transformer is used to demonstrate how this super-resolution affects classification. With the addition of super-resolution images, the classification results significantly increase based on the training and testing loss values. Although the results are good and can support the posed hypothesis, further research is needed to confirm the usefulness of this SR technique with different models before we can confidently state that the hypothesis is true.

Acknowledgement: The authors are thankful to this work involved human subjects in its research. Approval of all ethical and experimental procedures and protocols was granted by SRM Medical College Hospital and Research Centre, Kattankulathur, Chennai-603203,India,Under Ethical Clearance No:8709/IEC/2023.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] F. Yang, H. Yang, J. Fu, et al.," Learning texture transformer network for image super-resolution", Proceedings of the Conference on Computer Vision and Pattern Recognition. IEEE, 2020: 5791-5800.

[2] X. Zhu, L. Zhang, L. Zhang, and et al.," GAN-based image super resolution with a novel quality loss", Mathematical Problems in Engineering, 2020, 2020.

[3] Shahin, Ahmed & Aly, Saleh & Aly, Walaa. (2023). A novel multi-class brain tumor classification method based on unsupervised PCANet features. Neural Computing and Applications. 35. 1-17. 10.1007/s00521-023-08281-x.

[4] Antanas Kascenas, Pedro Sanchez, Patrick Schrempf, Chaoyang Wang, William Clackett, Shadia S. Mikhael, Jeremy P. Voisey, Keith Goatman, Alexander Weir, Nicolas Pugeault, Sotirios A. Tsaftaris, Alison Q. O'Neil, The role of noise in denoising models for anomaly detection in medical images, Medical Image Analysis ,2023,102963, ISSN 1361-8415, https://doi.org/10.1016/j.media.2023.102963.

[5] Zhanxiong Wu, Xuanheng Chen, Sangma Xie, Jian Shen, Yu Zeng,Super-resolution of brain MRI images based on denoising diffusion probabilistic model,Biomedical Signal Processing and Control, Volume 85,2023,104901,ISSN 1746-8094,https://doi.org/10.1016/j.bspc.2023.104901.

[6] Q. Yang, P. Yan, Y. Zhang, H. Yu, Y. Shi, X. Mou, M. K. Kalra, Y. Zhang, L. Sun, and G. Wang. Low-Dose CT Image Denoising Using a Generative Adversarial Network With Wasserstein Distance and Perceptual Loss. IEEE Trans Med Imaging, 37(6):1348–1357, 06 2018. 1, 2, 3.

[7] Zou, K., Yuan, X., Shen, X., Wang, M., Fu, H. (2022). TBraTS: Trusted Brain Tumor Segmentation. In: Wang, L.,

[8] Dou, Q., Fletcher, P.T., Speidel, S., Li, S. (eds) Medical Image Computing and Computer Assisted Intervention – MICCAI 2022. MICCAI 2022. Lecture Notes in Computer Science, vol 13438. Springer, Cham. https://doi.org/10.1007/978-3-031-16452-1_48

[9] Mavra Mehmood, Nasser Alshammari, Saad Awadh Alanazi, Asma Basharat, Fahad Ahmad, Muhammad Sajjad, Kashaf Junaid, Improved colorization and classification of intracranial tumor expanse in MRI images via hybrid scheme of Pix2Pix-cGANs and NASNet-large, Journal of King Saud University - Computer and Information Sciences, Volume 34, Issue 7,2022,Pages 4358-4374, https://doi.org/10.1016/j.jksuci.2022.05.015.

[10] Zhaohong Jia, Hongxin Zhu, Junan Zhu, Ping Ma,Two-Branch network for brain tumor segmentation using attention mechanism and super-resolution reconstruction,Computers in Biology and Medicine,Volume157,2023 ,https://doi.org/10.1016/j.compbiomed.2023.106751.

[11] Zhaoyang Song, Defu Qiu, Xiaoqiang Zhao, Dongmei Lin, Yongyong Hui, Channel attention generative adversarial network for super-resolution of glioma magnetic resonance image, Computer Methods and Programs in Biomedicine, Volume 229,2023,107255,ISSN 0169-2607, https://doi.org/10.1016/j.cmpb.2022.107255.

[12] Xianqi Li, Bernhard Strasser, Ulf Neuberger, Philipp Vollmuth, Martin Bendszus, Wolfgang Wick, Jorg Dietrich, Tracy T Batchelor, Daniel P Cahill, Ovidiu C Andronesi, Deep learning super-resolution magnetic resonance spectroscopic imaging of brain metabolism and mutant isocitrate dehydrogenase glioma, *Neuro-Oncology Advances*, Volume 4, Issue 1, January-December 2022, vdac071, https://doi.org/10.1093/noajnl/vdac071

[13] Muezzinoglu, Taha & Baygın, Nursena & Tuncer, Ilknur & Barua, Prabal Datta & Baygin, Mehmet & Dogan, Sengul & Tuncer, Turker & Palmer, Elizabeth & Cheong, Kang Hao & Acharya, U Rajendra. (2023). PatchResNet: Multiple Patch

Division–Based Deep Feature Fusion Framework for Brain Tumor Classification Using MRI Images. Journal of Digital Imaging. 36. 10.1007/s10278-023-00789-x.

[14] Ahmed I. Shahin, Walaa Aly, Saleh Aly, MBTFCN: A novel modular full convolutional network for MRI brain tumor multi-classification,Expert Systems with Applications, Volume 212,2023,118776,ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2022.118776.

[15] Hossein Mehnatkesh, Seyed Mohammad Jafar Jalali, Abbas Khosravi, Saeid Nahavandi, An intelligent driven deep residual learning framework for brain tumor classification using MRI images, Expert Systems with Applications, Volume 213, Part C, 2023,119087,ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2022.119087.

[16] Choudhuri, R., Halder, A. Brain MRI tumour classification using quantum classical convolutional neural net architecture. *Neural Comput & Applic* **35**, 4467–4478 (2023). https://doi.org/10.1007/s00521-022-07939-2

[17] A. Mishra, R. Jha and V. Bhattacharjee, "SSCLNet: A Self-Supervised Contrastive Loss-Based Pre-Trained Network for Brain MRI Classification," in IEEE Access, vol. 11, pp. 6673-6681, 2023, doi: 10.1109/ACCESS.2023.3237542.

[18] Iyami, J., Rehman, A., Almutairi, F. *et al.* Tumor Localization and Classification from MRI of Brain using Deep Convolution Neural Network and Salp Swarm Algorithm. *Cogn Comput* (2023). https://doi.org/10.1007/s12559-022-10096-2

[19] S. Deepak, P.M. Ameer, Brain tumor categorization from imbalanced MRI dataset using weighted loss and deep feature fusion,Neurocomputing,Volume 520,2023,Pages 94-102,ISSN 0925-2312, https://doi.org/10.1016/j.neucom.2022.11.039 .

[20] Johnson, Justin, Alexandre Alahi, and Li Fei-Fei. "Perceptual losses for real-time style transfer and super-resolution." European Conference on Computer Vision. Springer International Publishing, 2016.

[21] L. Zhou, G. Chen, M. Feng and A. Knoll, "Improving Low-Resolution Image Classification by Super-Resolution with Enhancing High-Frequency Content," *2020 25th International Conference on Pattern Recognition (ICPR)*, Milan, Italy, 2021, pp. 1972-1978, doi: 10.1109/ICPR48806.2021.9412876.

[22] Li, Xuelong, et al. "A multi-frame image super-resolution method." Signal Processing 90.2 (2010): 405-414 .

[23] Jia Y, He Z, Gholipour A, Warfield SK. Single Anisotropic 3-D MR Image Upsampling via Overcomplete Dictionary Trained From In-Plane High Resolution Slices. IEEE J Biomed Health Inform. 2016 Nov;20(6):1552-1561. doi: 10.1109/JBHI.2015.2470682. Epub 2015 Aug 20. PMID: 26302522; PMCID: PMC5243920.

[24] Tai, Yu-Wing, et al. "Super resolution using edge prior and single image detail synthesis." Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on. IEEE, 2010.

[25] Rouf, Mushfiqur, et al. "Fast edge-directed single-image super-resolution." Electronic maging 2016.15 (2016): 1-8.

[26] Vanam, Rahul, Yan Ye, and Serhad Doken."Joint edge-directed interpolation and adaptive sharpening filter." Global Conference on Signal and Information Processing (GlobalSIP), 2013 IEEE. IEEE, 2013.

[27] Huang Y, Wang W, Li M. FNSAM: Image super-resolution using a feedback network with self-attention mechanism. Technol Health Care. 2023;31(S1):383-395. doi: 10.3233/THC-236033. PMID: 37066938; PMCID: PMC10200178.

[28] Xin Hua, Zhijiang Du, Hongjian Yu and Jixin Maa," Convolutional neural network based on sparse graph attention mechanism for MRI super-resolution",2023, 2305.17898.

[29] Zhanxiong Wu, Xuanheng Chen, Sangma Xie, Jian Shen, Yu Zeng,"Super-resolution of brain MRI images based on denoising diffusion probabilistic model",Biomedical Signal Processing and Control,Volume 85,2023.

[30] Sara Altun Güven, Muhammed Fatih Talu, "Brain MRI high resolution image creation and segmentation with the new GAN method", Biomedical Signal Processing and Control, Volume 80, Part 1,2023,104246, ISSN 1746-8094, https://doi.org/10.1016/j.bspc.2022.104246.

[31] Zhang, Bowen & Gu, Shuyang & Zhang, Bo & Bao, Jianmin & Chen, Dong & Wen, Fang & Wang, Yong & Guo, Baining. (2022). StyleSwin: Transformer-based GAN for High-resolution Image Generation. 11294-11304. 10.1109/CVPR52688.2022.01102.

[32] Full Citation: Tsang, S. H. (2019, March 20). Review: LapSRN & MS-LapSRN—Laplacian Pyramid Super-Resolution Network (Super Resolution). Medium. https://towardsdatascience.com/review-lapsrn-ms-lapsrn-laplacian-pyramid-super-resolution-network-super-resolution-c5fe2b65f5e8.

[33] Z. Liu, et al., "Swin Transformer: Hierarchical Vision Transformer using Shifted Windows," in 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada, 2021 pp. 9992-10002. doi: 10.1109/ICCV48922.2021.00986.

[34] Full Citation: IXI Dataset – Brain Development. (n.d.). IXI Dataset – Brain Development. https://brain-development.org/ixi-dataset/

[35] Msoud Nickparvar. (2021). Brain Tumor MRI Dataset [Data set]. Kaggle. https://doi.org/10.34740/KAGGLE/DSV/2645886

[36] X. Wang, L. Xie, C. Dong and Y. Shan, "Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data," *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, Montreal, BC, Canada, 2021, pp. 1905-1914, doi: 10.1109/ICCVW54120.2021.00217.

[37] Zhang, Kai and Liang, Jingyun and Van Gool, Luc and Timofte, Radu," Designing a Practical Degradation Model for Deep Blind Image Super-Resolution", IEEE International Conference on Computer Vision, 4791-4800,2021.

[38] Jingyun Liang, Jiezhang Cao, Guolei Sun, Kai Zhang, Luc Van Gool, Radu Timofte," SwinIR: Image Restoration Using Swin Transformer", arXiv:2108.10257 **,2021.**

[39] Dhariwal, P., Nichol, A.: Diffusion models beat gans on image synthesis. Advances in Neural Information Processing Systems 34, 8780–8794 (2021)

[40] Wang, X. *et al.* (2019). ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks. In: Leal-Taixé, L., Roth, S. (eds) Computer Vision – ECCV 2018 Workshops. ECCV 2018. Lecture Notes in Computer Science(), vol 11133. Springer, Cham. https://doi.org/10.1007/978-3-030-11021-5_5

[41] Jingyun Liang and Jiezhang Cao and Guolei Sun and Kai Zhang and Luc Van Gool and Radu Timofte," SwinIR: Image Restoration Using Swin Transformer", 2021, 2108.10257.