**INTRODUCTION:**

Many medical applications require high resolution 3D images for early and accurate diagnosis. But required resolution cannot always acquired (achieved) due to acquisition constraints such as insufficient acquisition time, moving subjects, Signal to Noise Ratio (SNR). Acquiring higher resolution 3D images also require improved hardware which highly increases the cost of equipment. Hence, several super resolution methods are being applied to Magnetic Resonance Images (MRI) as a post-acquisition technique to achieve higher resolution. Super resolution (SR) is a process of estimating high resolution (HR) Image by upscaling or enhancing low resolution (LR) image. This method is proven advantageous in representing high resolution and offers possibility of reconstructing high-SNR[2]*.*

The resolution of MRI is defined by size of the voxel. A voxel is a three-dimensional volume element which are frequently non-uniform in three different directions (Anisotropic). The size of voxel/ resolution depends on matrix size, field of view (FoV), slice thickness. The matrix size is defined by the number of frequency encoding steps in one direction and the number of phase encoding steps in other direction. Field of view is the size of area covered by the matrix of frequency and phase encoding, in plane voxel size is calculated by dividing field of view by the matrix size and is referred as in plane resolution. Slice thickness determines the depth of voxel and is considered as the through plane resolution.

**Goal of Task:**

This work mainly focuses on implementing a deep neural network (UNet) for improvising through plane resolution of brain MRI images.

**Data Preparation:**

From publicly available IXI, T2-weighted brain MRI dataset, 22 images with equal sizes are selected for implementing the task. The neural network used requires low resolution images as training input and high-resolution images as the training target. Images selected from dataset are anisotropic (non-uniform), so these images are made isotropic (uniform) to treat them as training target (high resolution images).These isotropic images are down sampled two times (2X) to reduce through plane resolution and used them as training input (low resolution images)[3].

**UNet Architure:**

UNet justifies name with its U-shaped architecture. This architecture is first proposed for medical image segmentation, showed such good practical results that it is used in several other deep learning tasks like single image super resolution, 3D segmentations, image colouring. In this task a basic UNet model is referred and modified as an attempt for improving the through plane resolution of brain MRI images. This architecture consists of an up sampling, contraction, middle and expansion sections. In up sampling section, the inputs are up sampled to the size of target images using trilinear interpolation method. Contraction section has a series of blocks with each block underlying two 3D convolutional layers following max pooling layer. In this section the feature maps are doubled after each block for learning complex structures effectively. Middle section has two 3D convolutional layers with doubled feature mapping, treated as bottom layer. Expansion section can be considered as heart of this architecture. Like in contraction section this section also has a series of blocks, each block consisting of a up sampling layer, concatenation layer following two convolution layers. After each block, the feature maps get halved to get symmetry with the contraction and middle sections, ensures that the features learned in contraction section are used to reconstruct the image. Later the resulting mappings are passed through the single 3D convolutional layer to output feature maps equal to that of first layer in contraction section.

**Training setup for the network:**

To find optimal parameters of model the loss between the predicted output and corresponding ground truth must be minimized. Research works shown that the Mean Squared Error (MSE) has provided reasonable solution for super resolution tasks. Hence, MSE is used in this model as a loss function for evaluating the image restoration quality. This function is defined as the mean of squared difference between the original value () and the predicted value (). Considering “n” as the number of training samples MSE can be formulated as below.

Mean Squared Error =

MSE loss function largely favours high quality measurement between original and reconstructed image known as Peak Signal to Noise Ratio (PSNR). Higher the PSNR, higher the quality of reconstructed image. For optimization Adam optimizer with learning rate of 0.001 is used to minimize loss.

**Further plans of work:**

* Evaluate the performance of model by increasing data for training.
* Increase the number of blocks in contraction and expansion sections to check if it effects the performance of the model.
* Verify the model performance for different down-sampling factors.
* Despite of MSE favouring high PSNR, various literature works show that MSE loss function does not fully exploit the potential of networks. So SSIM, MSSIM, L1 loss functions can also be used as alternative loss functions and look for the best fit.
* Check how model perform while repeating the above steps with reduced in plane resolution.

**References:**

[1] C. Dong, C. C. Loy, K. He, and X. Tang, “Image Super-Resolution Using Deep Convolutional Networks,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 2, pp. 295–307, 2016, doi: 10.1109/TPAMI.2015.2439281.

[2] E. Van Reeth, I. W. K. Tham, C. H. Tan, and C. L. Poh, “Super-resolution in magnetic resonance imaging: A review,” *Concepts Magn. Reson. Part A Bridg. Educ. Res.*, vol. 40 A, no. 6, pp. 306–325, 2012, doi: 10.1002/cmr.a.21249.

[3] J. Kim, J. K. Lee, and K. M. Lee, “Accurate image super-resolution using very deep convolutional networks,” *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-Decem, pp. 1646–1654, 2016, doi: 10.1109/CVPR.2016.182.