Tech Science Press

A Novel Segment White Matter Hyperintensities Approach for Detecting Alzheimer

Antonitta Eileen Pious[1,*] and U. K. Sridevi[2]

[1]Department of Science and Humanities, Sri Krishna College of Engineering and Technology, Coimbatore, 641008, India
[2]Department of Applied Mathematics and Computational Sciences, PSG College of Technology, Coimbatore, 641008, India
*Corresponding Author: Antonitta Eileen Pious. Email: eileenpious@gmail.com

Abstract: Segmentation has been an effective step that needs to be done before the classification or detection of an anomaly like Alzheimer's on a brain scan. Segmentation helps detect pixels of the same intensity or volume and group them together as one class or region, where in that particular region of interest (ROI) can be concentrated on, rather than focusing on the entire image. In this paper White Matter Hyperintensities (WMH) is taken as a strong biomarker that supports and determines the presence of Alzheimer's. As the first step a proper segmentation of the lesions has to be carried out. As pointed out in various other research papers, when the WMH area is very small or in places like the Septum Pellucidum the detection of the lesion is hard to find. To overcome such problem areas a very optimized and accurate Threshold would be required to have a precise segmentation to detect the area of localization. This would help in proper detection and classification of the Anomaly. In this paper an elaborate comparison of various thresholding techniques has been done for segmentation. A novel idea for detection of Alzheimer's has been presented in this paper, which encompasses the effectiveness of an optimized and adaptive technique. The Unet architecture has been taken as the baseline model with an adaptive kernel model embedded within the architecture. Various state-of-the-art technologies have been used with the dataset and a comparative study has been presented with the current architecture used in the paper. The lesion segmentation in narrow areas has accurately been detected compared to the other models and the number of false positives has been reduced to a great extent.

Keywords: Alzheimer's; adaptive threshold; deep learning; segmentation

1 Introduction

Thresholding has been a technique used for a very long time to separate parts of an image based on some value, this helps segment the image and allows the Region of Interest (ROI) to be concentrated upon [1]. It also equips us from clearly differentiating the fore ground of an image from the background. In the detection of Alzheimer's this is a very crucial part as the area where the lesion lies is segmented and helps in tracing the level of abnormality present. Thresholding techniques play a vital part of any segmenting process and are

usually considered as the first step of segmenting. When we consider WMH as a strong biomarker in detecting Alzhemier's [2], a proper threshold value to correctly segment the area of lesion occurrence is mandatory. Thresholding can be performed in the following ways depending on what kind of data they are dealing with or working on. The Tab. 1 below gives us a broad classification of the various thresholding techniques used.

Table 1: Thresholding techniques

Categories	Data they deal with
1. Histogram shaped based methods	1. Selecting an optimal threshold value to divide the histogram into two classes [3].
2. Clustering based methods	2. Dividing an image into clusters based on an optimal threshold value [4].
3. Entropy based methods	3. Makes use of the entropy to differentiate the foreground and background [5].
4. Object based Attribute methods	4. They use attributes like shape, edge etc to make the differentiation [6].
5. Spatial methods	5. Makes use of Probability distribution to distinguish between pixels [7].
6. Local methods	6. Makes us of the local values of the pixels to differentiate [8].

The purpose of segmentation is to choose characteristics of our choice, as in this case the hyperintensity level to be more localized which in turn aids in the proper classification and detection of Alzheimer's. Automated segmentation in an aging brain becomes very complex to process because of the heterogeneous nature of WMH [9].

The lesser the threshold, the more homogeneous the segments will be. A larger threshold value will cause a more heterogeneous and generalized segmentation result. Every image is made up of a collective group of pixels values. In image segmentation the task is to classify the image into various segments depending on the classes assigned to the pixel value. Processing the entire image doesn't serve any fruitful purpose, So segmenting pixels on an apt threshold value helps us in detecting WMH areas that are usually missed out. Thresholding is a type of segmentation where we change the pixels of an image to make the image easier to analyze. The school old methods of thresholding like Otsu's Thresholding [10], adaptive thresholding, and grab cut thresholding methods were used with the dataset to help gain a better understanding of the optimized threshold value. The drawback of Otsu's thresholding of WMH was the inability of the method to detect the number of occurrences of the false positives because of the heterogenous nature of the pixels [11] Otsu's thresholding technique would have been suitable for a bimodal image. The histograms that were displayed for the noisy images fed, did not serve the purpose and the peak of thresholding wasn't visible enough. Next the adaptive thresholding technique was utilized with multiple thresholds this was apt when dealing with the varying intensities of WMH. But the disadvantage of adaptive thresholding was that the size of the neighborhood had to be large enough to cover sufficient foreground and background pixels, otherwise a poor threshold was chosen. In the case of WMH, the pattern of occurrence is very hazy and hard to identify [12]. The grabcut segmentation worked fairly well but required the ROI to be specified. This study as shown in Fig. 1 helped us gain an insight that deep learning models would be well suited for the automatic detection of features and to increase the accuracy of segmentation and speed.
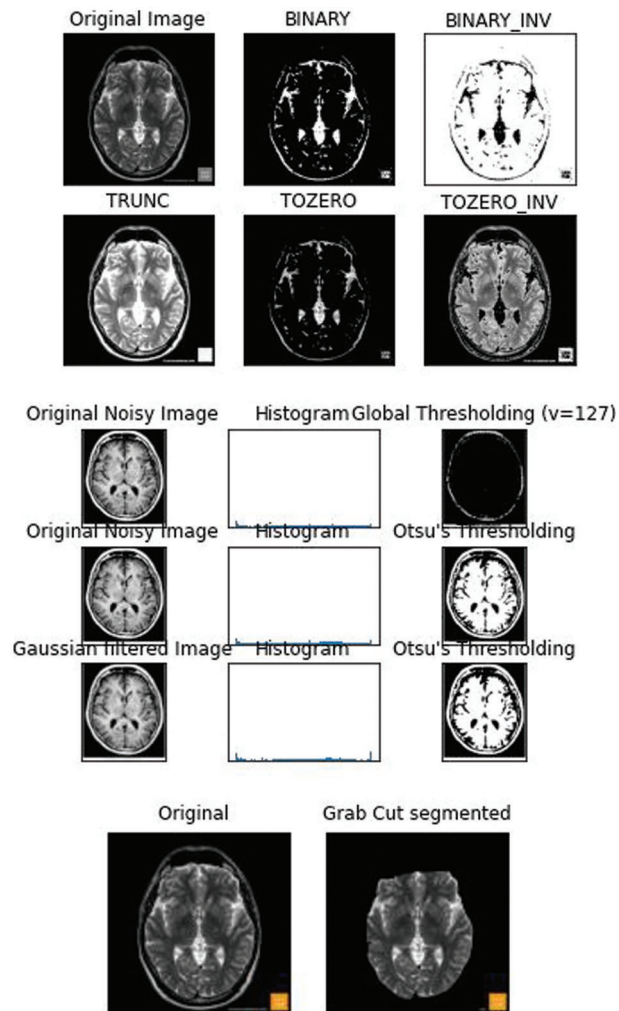
Figure 1: 2D FLAIR image from the WMH dataset was fed and the basic thresholding methods were performed on it

2 Materials and Methods

2.1 Materials Used

The dataset used in this paper was got from the WMH segmentation challenge. The medical images used were collected from five different scanners from hospitals in Netherland and Singapore. Both 3DT1 and FLAIR images were used. Elastix 4.8 a toolbox for medical image registration was used for the alignment of 3DT1 with FLAIR images. The images were preprocessed using SPM12r6685 to correct the bias field inhomogenities using statistical parameter mapping.

2.2 Preprocessing

We further pre processed the raw data to serve our purpose of removing false positives, the segmentation of the non brain tissues from the actual brain tissues is a tremendous task for clinicians too. The available dataset from the WMH segmentation challenge is 3D T1-weighted sequence consisting of 192 slices, each slice having a voxel size: $1.00 \times 1.00 \times 1.00$ mm^3. The dataset also has a 3D FLAIR sequence of 321 sagittal slices, voxel size: $1.04 \times 1.04 \times 0.56$ mm^3. Each folder contains 3 types of images in Nifti

format. The NiBabel package was used to load the images into the Python software. Three variables were used to differentiate the images as-T1weighted: This is the raw MRI image with a single channel. The Images in the dataset are 3D and can be imagined as multiple 2D images stacked together. T1w_brainmask: It is the image mask of the brain or can be called as the ground truth. It is obtained using Extractor tool from the deepbrain 0.1 library. T1w_brain: This can be thought of as part of brain stripped from above T1weighted image. This is similar to overlaying mask to actual images. A Skull extraction was performed to help remove the non brain areas from the brain areas [13] as shown in Fig. 2. The GPU environment used was the Google Colaboratory.
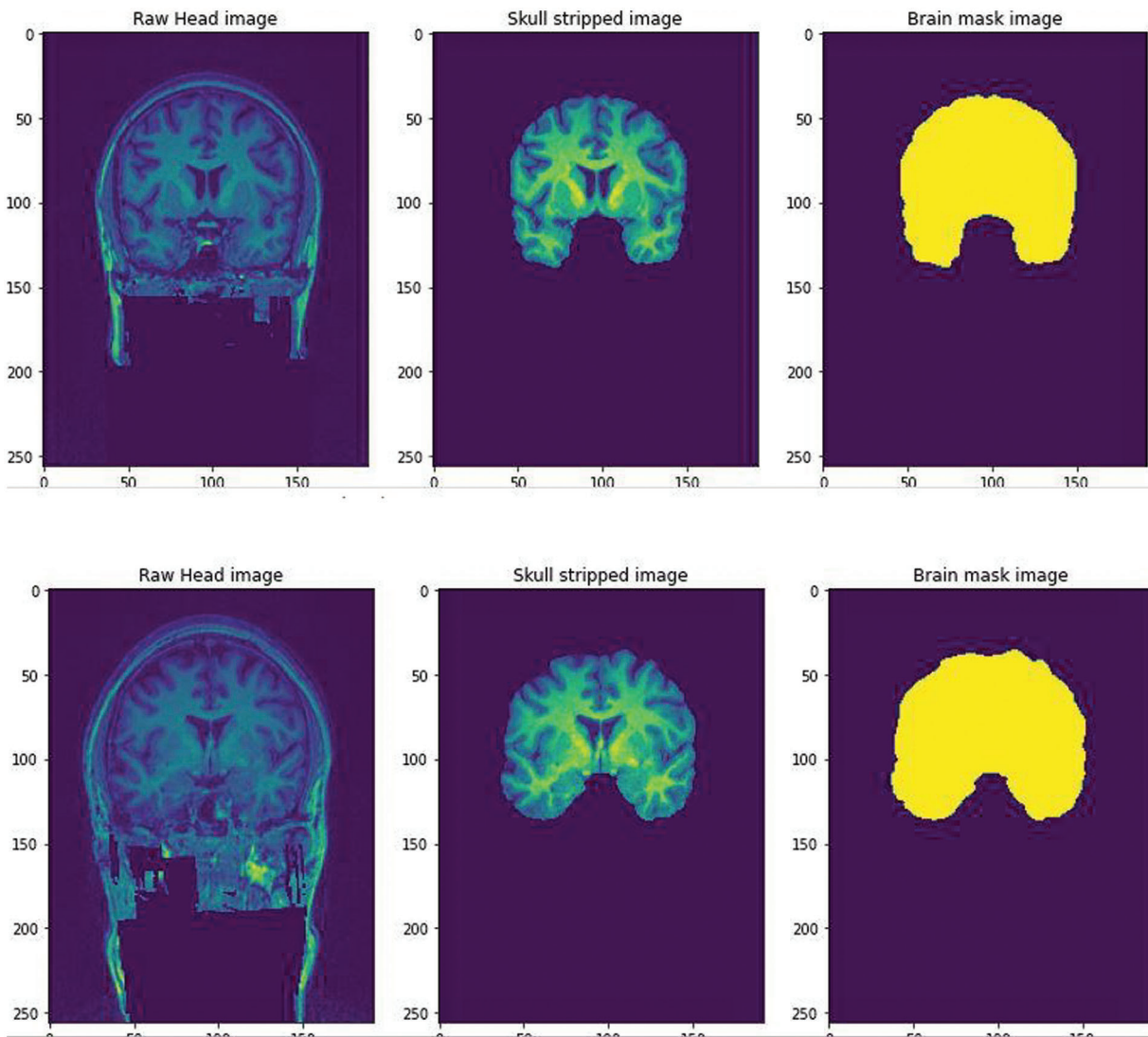

Figure 2: The Raw image from the WMH dataset was fed and skull stripped using the Extractor tool from deepbrain library. A brain mask image was created to help train the CNN for a predicted output

A simple CNN architecture [14] was used to create a mask to train the model in predicting the skull stripped image. After training for 60 epochs we got a validation Intersection over Union (iou_score) of

0.88, the result obtained is shown in Fig. 3. The iou_score helped in identifying the quantification of overlap between the target mask and the prediction image.
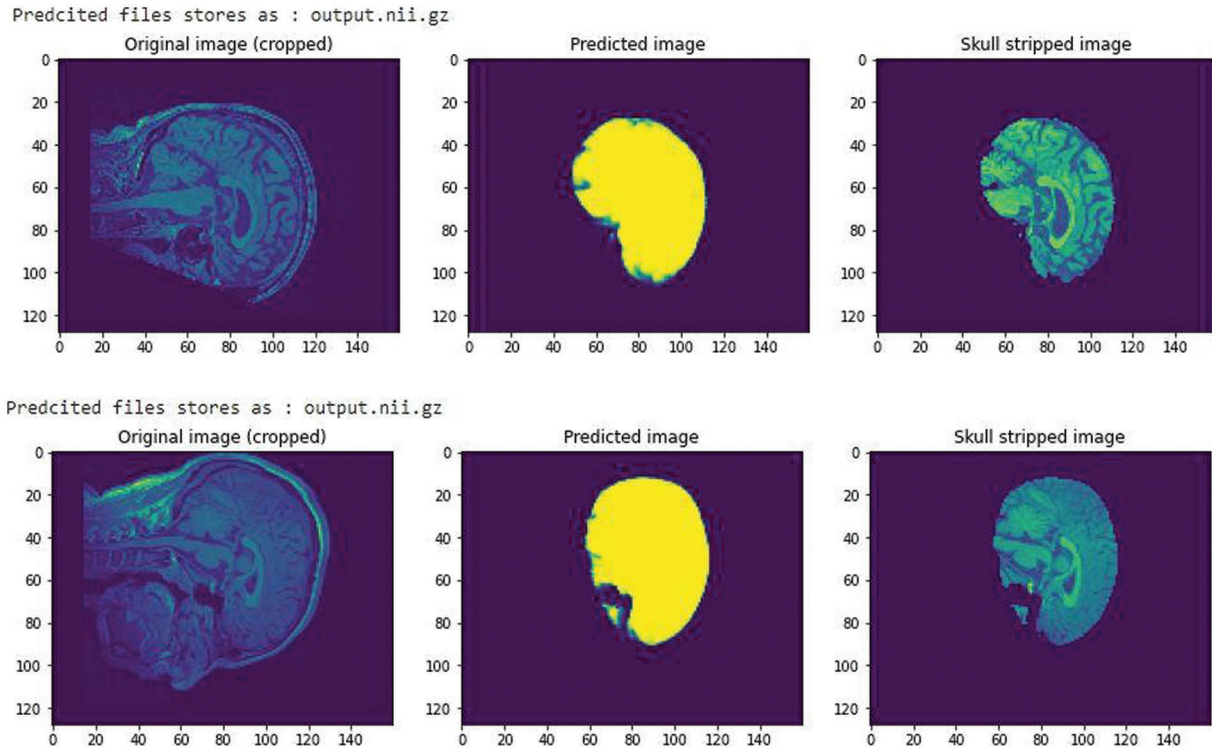

Figure 3: A CNN trained model with an iou_score of 0.88 after training for 60 epochs

All images were pre-processed with N4ITKBiasFieldCorrection [15] to correct bias field inhomogeneities.
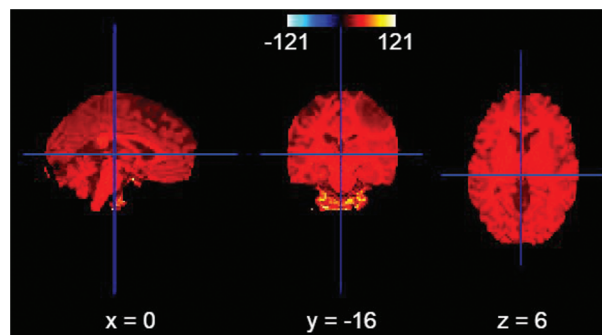

Figure 4: Bias field correction

3 Network Architecture

3.1 Baseline Model

In the current study, we put forth our work to combine an adaptive thresholding strategy with a 3D Unet architecture [16]. In the previous research studies as we have seen the Smaller WMH were difficult to trace

out and the detection of false positives in areas like the Septum Pellucidum was really very difficult to comprehend. According to our method used here, we make use of the Unet architecture which holds well in downsampling an image concentrating on the spatial dimensions and localization. Unlike CNN architecture which is quite suitable to only detect the presence or absence of a morbidity, the Unet architecture helps to classify the morbidity as well as trace the area where they are present by segmenting each and every pixel in the ROI (Region of Interest). The second reason of choosing a Unet is the availability of images of arbitrary sizes, which requires an architecture which would accept a heterogeneous kind. A general Unet architecture makes use of skip connections between the contracting and the expanding path in order to recover spatial information lost during down sampling. Or in other words it helps in removing the vanishing gradient problem. Here, we further advance 3D Unet for WMHs segmentation (Fig. 5) with two original components: A use of different details for the network architecture and inclusion of a threshold of pixilated values that are concatenated with the up sampled values.
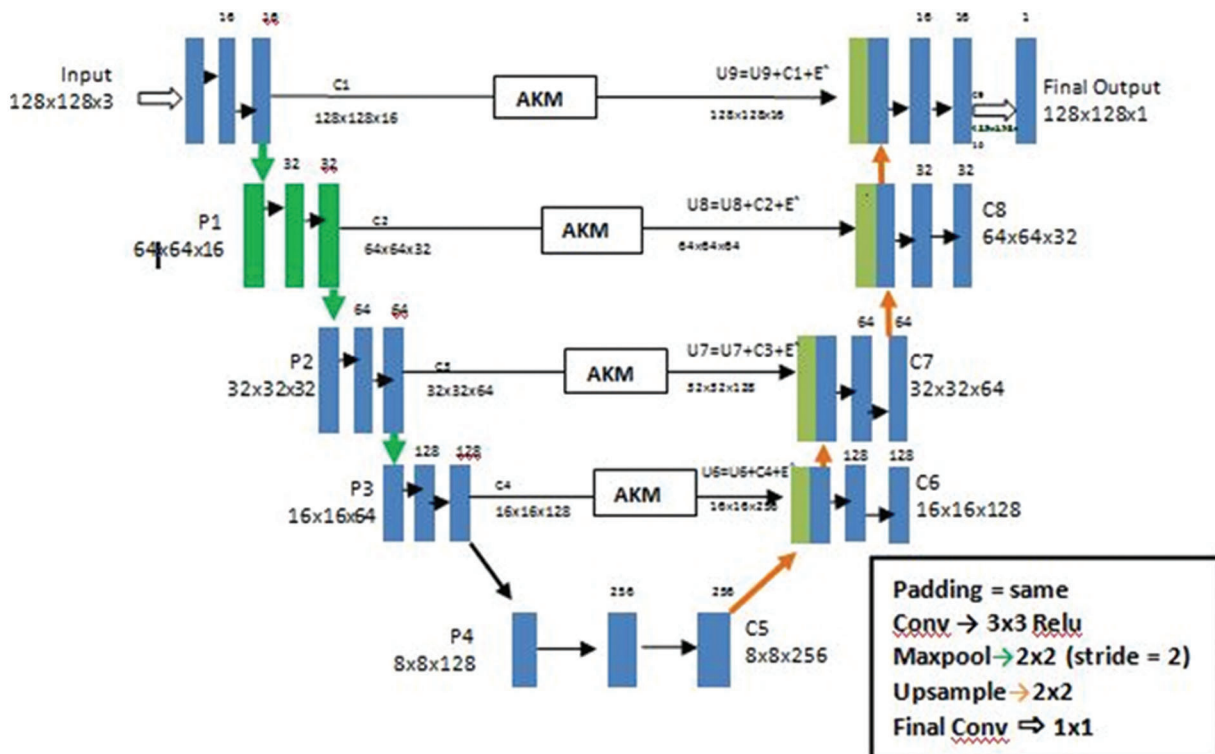

Figure 5: The architecture of a U-Net AKM model

3.2 Adaptive Kernel Model

Our primary goal is to make the architecture less location sensitive that is it should not reject or not detect a White Matter Hyperintensity in a low probability area like the Septum Pellucidum. Our second goal is ultimately to reduce the false positives for this we have used an Adaptive kernel model in between the skip connections of the encoder and decoder. The Adaptive Kernel Model (AKM) basically gets the feature map with a large receptive field from the encoder and performs an adaptive thresholding on it using a 3×3 and 5×5 kernel achieving the best result. The strategy used here was to give the integral image as input and figure the left and above pixel values, as shown in Fig. 6.

$$I(x,y) = f(x,y) + I(x-1,y) + I(x,y-1) - I(x-1,y-1) \tag{1}$$

Once we have the innate image, the sum of the function for any rectangle with upper left area (x1, y1) and lower right area (x2, y2) can be computed in constant time using the following equation

$$\sum_{x=x1}^{x2}\sum_{y=y1}^{y2} f(x,y) = I(x_2 y_2) - I(x_2, y_1-1) - I(x_1-1,y_2) + I(x_1-1,y_1-1) \tag{2}$$
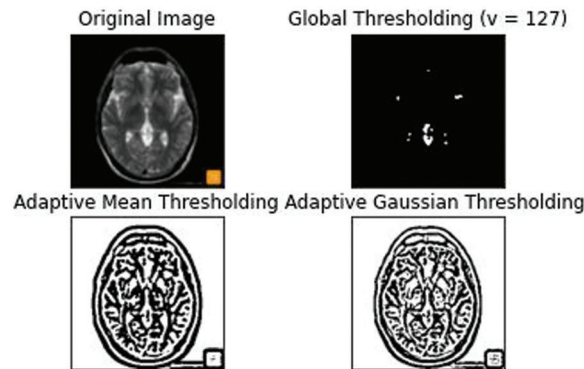

Figure 6: An Adaptive mean and Gaussian thresholding applied with an approximate threshold value of around 127

This thresholding technique helps in discriminating the higher and lower intensity areas of white matter. The output of A3 and A5 are fed into an element wise sum function, followed by a fully connected layer Fmp to carry out an accurate and localized selection. The resulting feature map contains the determined weights for White Matter of different intensities. This being a binary classification problem a sigmoid function is made use of, the parameters are then applied to each feature map and all of them are summed up as follows. The output of values is concatenated alone with the down sampled values. The inclusion of the AKM unit as shown in Fig. 7 has helped discern the difference in intensities to a larger extent with a focus on areas with meager existence of White matter like the Septum Pellucidum.

$$\widehat{E} = \widehat{A_3} + \widehat{A_5} = a.A_3 + b.A_5 \tag{3}$$
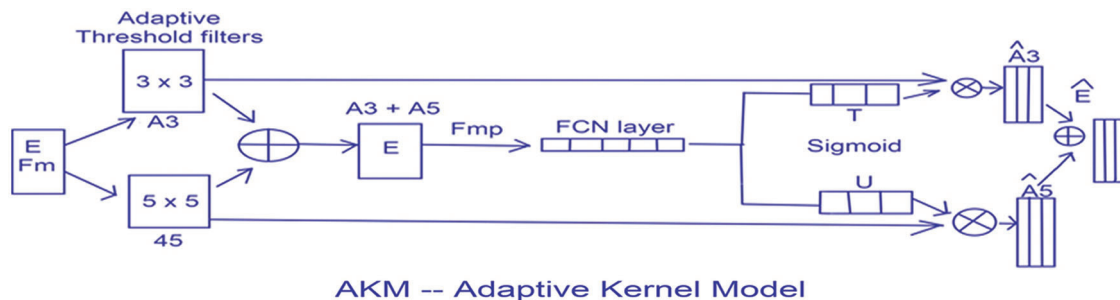
Figure 7: The AKM model architecture

4 Experiments and Results

The metrics to assess our proposed method were selected from the suggestion of the MICCAI 2017 challenge, which included the Dice Similarity Coefficient, Hausdroff distance, Absolute percentage Volume difference, the sensitivity for detecting individual lesions (recall) and F1 score for individual lesions (F1)

They are defined in the Equations given below to suit our requirement of tracing WMH.

Dice Similarity coefficient (DSC) [17] can be calculated as follows

$$DSC = \frac{2 * TP}{2 * TP + FP + FN} \tag{4}$$

Hausdroff distance (HD) helps to calculate the distance of a point (m) in the model image to the point in the tested image (s). A HD of values closer to zero indicates a partial matching while a value equal to zero tells us the ground image and segmented image have totally matched. We have used the following formula in calculating the Hausdroff distance here

$$M = \{m_1, m_2 \text{---------------} m_p\} \tag{5}$$

$$S = \{s_1, s_2 \text{--------------------} s_3\} \tag{6}$$

$$H(M, S) = \max(h(M, S), h(S, M) \tag{7}$$

where

$$h(S, M) = max_{m \in M} min_{s \in S} \|m - s\| \tag{8}$$

Absolute percentage Volume difference (AVD) [18] can be calculated as the absolute log value of the segmented volume divided by the true volume.

$$AVD = \left| log \frac{Segmented\ volume\ of\ image}{True\ volume\ of\ image} \right| \tag{9}$$

The recall or sensitivity can be defined as the true positive rate (TPR) [18], it measures the number of positive voxels in the expected image that are also identified as positive in the segmented image.

$$Recall = TPR = \frac{TP}{TP + FN} \tag{10}$$

And the final metric of assessment we use here is the precision or F1 score [18] which can be calculated as

$$F1 = \frac{TP}{TP + \dfrac{1(FP + FN)}{2}} \tag{11}$$

From the above equations TP-True Positive

TN - True Negative

FP - False Positive

FN - False Negative

5 Implementation Details

Our proposed architecture was implemented with the Python and Keras framework. We make use of the U-Net architecture as the base structure of the auto encoder-decoder framework. Since the images in the dataset have a heterogeneous resolution, we perform the pre processing techniques like skull stripping and bias correction. The parameters vary depending on the heterogeneity of the images available. The input taken was 128×128 matrix for an 3 channel input, with a convolutional layer of 3×3, a maxpooling of 2×2 and a stride of 2. We make use of the ADAM optimizing algorithm with a batch size of 32 and epochs of 30. The initial learning rate is set at 0.00003 and is dropped by 10 percent after each 30 epochs.

6 Results

According to the observations White Matter intensities in areas like the Septum pellucidum were hard to trace out because of the variability in intensities and volume. Here in our proposed method a 3D multislice attainment is made with relatively few slices for training the Unet architecture. As per the requirement the number of false positives of the White Matter Hyper intensities in smaller regions had to be segmented perfectly well. The combined approach of using an Adaptive thresholding method to the feature map received from the encoder to the decoder helped in bringing the recall and F1 rate to 0.86 and 0.80 respectively. This change in values implies the decrease in the false positive pixel values and the overall accuracy has improved extensively as shown in Fig. 8.
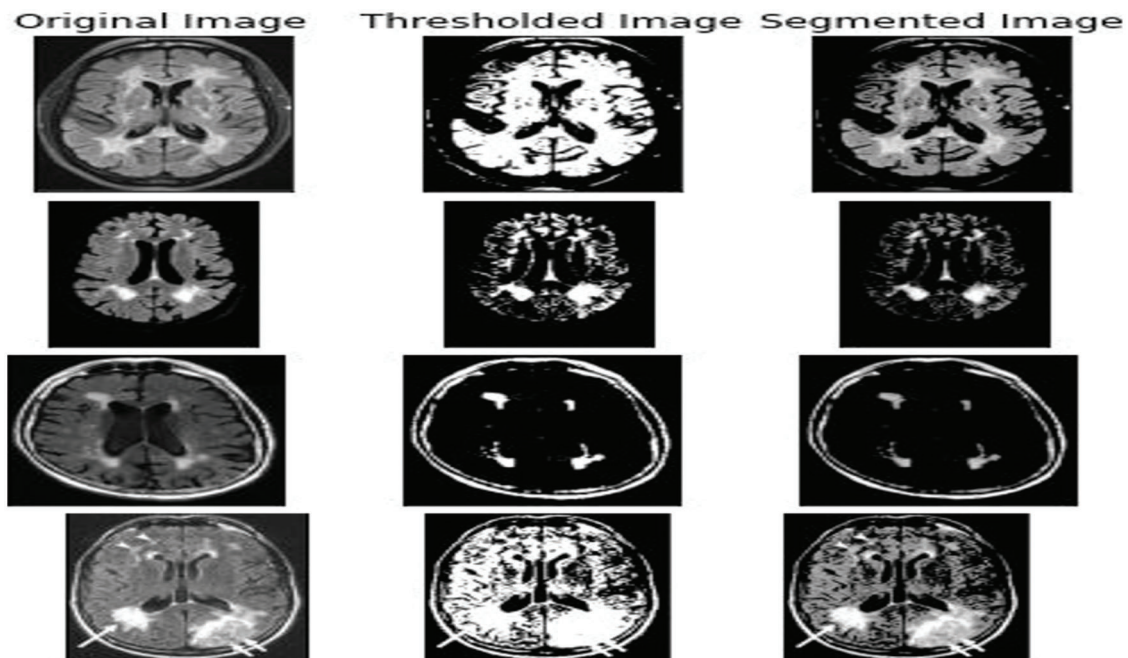

Figure 8: The output got from various input images like the 3DT1 and 3D Flair images that were fed into the Unet AKM model

The overall accuracy of the proposed model has turned out to be around 0.75 iou_score for 120 epochs. From the plot of accuracy Fig. 9 the model could most likely be trained a little more as the trend for accuracy on both datasets is still rising for the last few epochs. It can be noticed that the model has not yet over-learned the training dataset, showing similar ability on both datasets using for training and testing. From the plot of

loss, the model has an optimal performance on both train and validation datasets (labeled test). If these parallel plots start to depart consistently, it might be an indication to stop training at an earlier epoch.
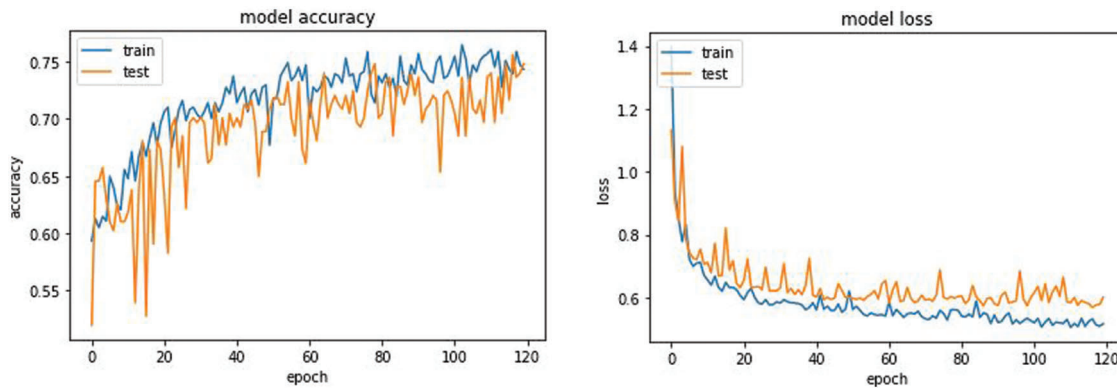

Figure 9: Model accuracy and model loss of the Unet AKM model

7 Comparison with Other Published Methods

A comparative study was done with three frameworks as shown in Tab. 2 like the fully convolutional neural network with an ensemble of three different networks and three different initializations of weight. A data normalization and augmentation was applied to remove the false positives. The second framework that was used as a comparative model was a multi dimensional gated recurrent unit. And the third framework that was used for a comparative study was a multi scale deep network with annealed training with a patch based architecture, where patches of sizes 128×128, 64×64 and 32×32 were extracted and fed into a deep convolutional network [19] as shown in Fig. 10.

Table 2: Performance of various models

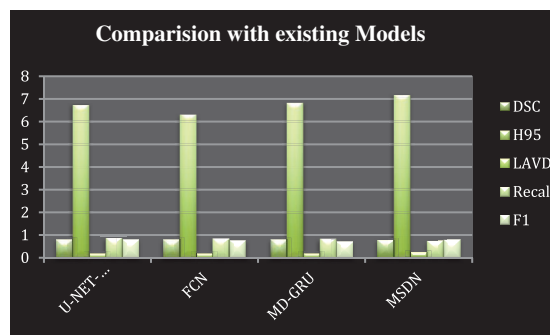| Number	Model	DSC	H95	LAVD	Recall	F1
1	Unet AKM (Proposed method)	0.80	6.72	0.193	**0.86**	**0.80**
2	FCN	**0.80**	**6.30**	0.193	**0.84**	0.76
3	MD-GRU	0.78	6.82	**0.193**	0.83	0.70
4	MSDN	0.77	7.16	0.219	0.73	**0.78**

Figure 10: Comparison with other existing models

8 Conclusion

Our proposed work has helped reduce the occurrence of false positives and improve the accuracy rate compared to the already existing Deep learning framework. It has also helped in detecting the presence of WMH in smaller regions like the Septum Pellucidum. Further work can involve the inclusion of volume and shape of the WMH along with the intensity changes. This would help to unravel the actual presence of WMH and the extent of the spread of the morbidity. This would help clinicians with a more clear and precise view in the detection of Alzheimer's.

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

References

[1] P. K. Sahoo, S. Soltani and A. K. C. Wong, "A survey of thresholding techniques," *Computer Vision, Graphics, and Image Processing*, vol. 41, no. 2, pp. 233–260, 1988.

[2] A. M. Brickman, "Contemplating Alzheimer's disease and the contribution of white matter hyperintensities," *Current Neurology and Neuroscience Reports*, vol. 13, no. 2, pp. 415, 2013.

[3] W. J. Choi and T. S. Choi, "Automated pulmonary nodule detection based on three-dimensional shape-based feature descriptor," *Computer Methods and Programs in Biomedicine*, vol. 113, no. 1, pp. 37–54, 2014.

[4] C. W. Woo, A. Krishnan and T. D. Wager, "Cluster-extent based thresholding in fMRI analyses: Pitfalls and recommendations," *NeuroImage*, vol. 91, pp. 412–419, 2014.

[5] Y. Feng, H. Zhao, X. Li, X. Zhang, H. Li et al., "A multi-scale 3D Otsu thresholding algorithm for medical image segmentation," *Digital Signal Processing*, vol. 60, no. 9, pp. 186–199, 2017.

[6] X. Descombes, F. Kruggel, G. Wollny and H. J. Gertz, "An object-based approach for detecting small brain lesions: Application to Virchow-Robin spaces," *IEEE Transactions on Medical Imaging*, vol. 23, no. 2, pp. 246–255, 2004.

[7] S. Jansi and P. Subashini, "Optimized adaptive thresholding based edge detection method for MRI brain images," *International Journal of Computer Applications*, vol. 51, no. 20, pp. 1–8, 2012.

[8] Y. Feng, H. Zhao, X. Li, X. Zhang and H. Li, "An integrated method of adaptive enhancement for unsupervised segmentation of MRI brain images," *Pattern Recognition Letters*, vol. 24, no. 15, pp. 2549–2560, 2003.

[9] M. E. Caligiuri, P. Perrotta, A. Augimeri, F. Rocca, A. Quattrone et al., "Automatic detection of white matter hyperintensities in healthy aging and pathology using magnetic resonance imaging: A review," *Neuroinformatics*, vol. 13, no. 3, pp. 261–276, 2015.

[10] G. Sandhya, G. BabuKande and T. S. Savithri, "Multilevel thresholding method based on electromagnetism for accurate brain MRI segmentation to detect White Matter, Gray Matter, and CSF," *BioMed Research International*, vol. 2017, no. 19, pp. 1–17, 2017.

[11] M. F. Rachmadi, M. D. C. Valdés-Hernández, M. L. F. Agan, C. Di Perri and T. Komura, "Segmentation of white matter hyperintensities using convolutional neural networks with global spatial information in routine clinical brain MRI with none or mild vascular pathology," *Computerized Medical Imaging and Graphics*, vol. 66, no. 4, pp. 28–43, 2018.

[12] M. E. Caligiuri, P. Perrotta, A. Augimeri, F. Rocca, A. Quattrone et al., "Automatic detection of White Matter Hyperintensities in healthy aging and pathology using magnetic resonance imaging: A Review," *Neuroinformatics*, vol. 13, no. 3, pp. 261–276, 2015.

[13] A. Subudhi, J. Jena and S. Sabut, "Extraction of brain from MRI images by skull stripping using histogram partitioning with maximum entropy divergence," in *2016 Int. Conf. on Communication and Signal Processing (ICCSP)*, India, pp. 931–935, 2016.

[14] J. Kleesiek, G. Urban, A. Hubert, D. Schwarz, K. Maier-Hein et al., "Deep MRI brain extraction: A 3D convolutional neural network for skull stripping," *Neuroimage*, vol. 129, pp. 460–469, 2016.

[15] N. J. Tustison, B. B. Avants, P. A. Cook, Y. Zheng, A. Egan *et al.,* "N4ITK: Improved N3 Bias correction," *IEEE Transactions on Medical Imaging*, vol. 29, no. 6, pp. 1310–1320, 2010.

[16] C. Wang, T. MacGillivray, G. Macnaught, G. Yang, D. Newby *et al.,* "A two-stage 3D Unet framework for multi-class segmentation on full resolution image," *ArXiv abs/1804.04341*, 2018.

[17] K. H. Zou, S. K. Warfield, A. Bharatha, C. M. Tempany, M. R. Kaus *et al.,* "Statistical validation of image segmentation quality based on a spatial overlap index," *Academic Radiology*, vol. 11, no. 2, pp. 178–189, 2004.

[18] A. A. Taha and A. Hanbury, "Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool," *BMC Medical Imaging*, vol. 15, no. 1, pp. 29, 2015.

[19] R. J. Kay Oskal, Martin Risdal, E. Janssen, E. S. Undersrud and T. O. Gulsrud, "A U-net based approach to epidermal tissue segmentation in whole slide histopathological images," *SN Applied Sciences*, vol. 1, no. 7, pp. 672, 2019.