

Brain Image Segmentation using Machine Learning

Shobhit Verma¹, Chirag Rana²

Department of Computer Engineering, Sardar Patel Institute Of Technology, Mumbai, India

shobhit.verma@spit.ac.in, chirag.rana@spit.ac.in

Abstract—Deep Neural Networks have shown to be extremely effective in the areas of object recognition, discovery, and division in computer vision. These methods have been applied in the field of clinical image examination. Convolutional neural networks (CNNs), a crucial aspect of deep learning applications for visual reasons, have gotten a lot of attention in recent years due to their advanced capabilities in computer vision applications. They've made significant progress in the areas of object identification, recognition challenges, and division challenges. Our focus is on the models that are being used, information pre-handling and preparedness, and appropriately preparing the following information or picture. The U – Nets are a type of CNN with accuracy comparable to that of humans. We developed and applied the U-Net CNN architecture, as well as image segmentation for brain magnetic resonance images (MRI).

The goal of our research is to focus on pre-processing MRI images, performing Skull Stripping with a Deep CNN architecture U-Net, and performing image segmentation.

Index Terms— Machine Learning, MRI Analysis

1. INTRODUCTION

The most advanced imaging techniques, which are employed as an effective and efficient tool in diagnoses, treatment, and therapy, are largely responsible for recent advances in medical science.

Artificial intelligence, machine learning, deep learning, and computer vision advancements have created huge prospects for developing intelligent decision support systems with higher accuracy, lower error rates, and the discovery of new knowledge about disease and therapy. The use of artificial intelligence in computers to interpret picture data and forecast anomalies has improved diagnostic confidence and accuracy in the image analysis process.

Image segmentation is a key step in medical image processing that allows you to divide an image into separate parts and analyse them in greater depth pathologically with varying degrees of accuracy and complication.

Various segmentation techniques have been proposed and utilised on brain pictures throughout the last decade.

With the use of magnetic resonance scans, rapid progress has been made in understanding the structure of the brain (MRI).

To aid doctors in qualitative diagnosis, computerised methods for MRI image segmentation, registration, and visualisation have been widely used.

Because the brain tissues contain many irregularities and aberrant tissues such as tumours, brain MRI image segmentation is a very complex and demanding process.

Artificial intelligence makes brain imaging analysis easier and more practical when dealing with vast amounts of data.

2. BACKGROUND AND RELATED WORK

Brain MRI image analysis has long been a popular topic of study, with tasks including lesion identification and segmentation, tissue segmentation, and brain parcellation performed on neonates, infants, and adults. CNNs have achieved accuracy comparable to that of humans for 2D images and are a strong tool for extracting feature hierarchy. As a result, we've progressed to using 3D CNNs in biomedical data analysis.

Bharath Hariharan et al. [1] execute fine-grained localization using hyper columns at each pixel for vector activation of CNN units above the pixel.

For semantic segmentation, Jonathan Long et al. [2] developed a Fully connected Convnet, or FCN. They claim Convnets are based on translation invariance. Their fundamental components (convolution, pooling, and activation functions) work on local input regions and are solely affected by relative spatial coordinates.

Alex Krizhevsk et al. [3] suggested an eight-layered network, with five layers of convolution and three layers of fully-connected networks with weights. With a million parameters, this network maximised multinomial logistic regression and reduced the problem of overfitting. As a result, with supervised learning, a deep CNN can provide the greatest results. However, when a single convolution layer was removed, the network's performance suffered

For 3D segmentation, Kleesiek et al. [4] developed an end-to-end 3D CNN technique. However, because their network was not very deep and only had one max-pooling after the first convolutions, many scale structures were not investigated.

3. METHODOLOGY

3.1. Data Preprocessing

3.1.1 Normalization

Normalisation function is applied to the image. Since a neural network works better with values in range [0,1], the default min and max values are 0 and 1 in the functionality. We can also use z-score normalisation, but it doesn't help in better classification of the images. Normalisation of the image content does not affect the way the image will be interpreted and doesn't do any radical changes to the pixel details. It is to Scale down the range of the intensity values of pixels in an image. It also helps view an image in proper contrast.

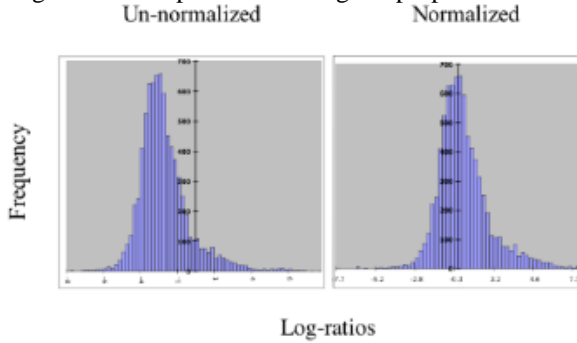


Fig. 1 Intensity Ranges

3.1.2 Image Registration

Image registration helps to organise and put the different sets of data into one coordinate system. Medical imaging always makes use of this technique for image analysis. Registration helps in providing an order to compare data obtained from different sources and methods. The basis for image registration was homography. Homography is a simple 3x3 matrix which maps collineation between two images. Homography is calculated on planar images. Let's say (x1, y1) is a point in first image and (x2, y2) is a point in another image. Then the homography relates them in the following way:

$$\begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = H \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = \begin{bmatrix} h_{00} & h_{01} & h_{02} \\ h_{10} & h_{11} & h_{12} \\ h_{20} & h_{21} & h_{22} \end{bmatrix} \cdot \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix}$$

The image registration process is not necessary if we are making use of a FCN (Fully convolutional neural network), since a convolutional network gradually down samples an image and focuses more on the context rather than the location of the relevant features in an image. But proper alignment of images helps in creation of a good data that might be useful to other people.

3.1.3 Skull Stripping

The way toward extricating the mind tissue from non cerebrum one is alluded to in the writing as skull stripping. We perform skull stripping and tissue segmentation using an FCN called as UNet.

3.2 CNN for Image Segmentation

UNet, developed from the customary convolutional neural system, was first planned, and applied in 2015 to process biomedical pictures. A general convolutional neural system concentrates its assignment on picture arrangement, where

info is a picture and yield are in one name, yet in biomedical cases, it requires us not exclusively to recognize whether there is an illness, yet in addition to restrict the territory of variation from the norm. The reason it can localise and distinguish borders is by doing classification on every pixel, so the input and output share the same size.

3.2.1 UNet – Convolution Operation

A convolutional operation takes two inputs – A 3d volume (nin x nin x channels) and a set of k filters (called as kernels) each with size (f x f x channels). A typical kernel is of size 3 x 3 x channels. The output of a convolutional operation is a 3d image of size (nout x nout x k).

The relation between nin and nout is given as: -

$$n_{out} = \left\lfloor \frac{n_{in} + 2p - k}{s} \right\rfloor + 1$$

nin: number of input features
nout: number of output features
k: convolution kernel size
p: convolution padding size
s: convolution stride size

The convolution operation helps find relevant features in the receptive field i.e. the area where the filter is looking at.

3.2.2 Max Pooling Operation

Max pooling operation helps reduce the size of the feature map so that we have fewer parameters in the network.

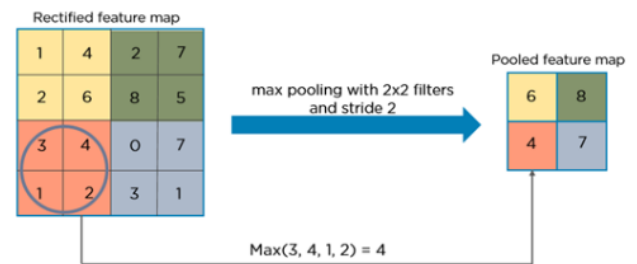


Fig.2 Max Pooling Operation

The idea behind max pooling operation is to retain only important features (max valued pixels) from each region and throw away irrelevant information. The information that best explains the context of the image will be retained. This operation, like convolutional operation, reduces the size of the image. Reduction of image size is known as down sampling. This basically means a high-resolution image is being converted to a low-resolution image. Both Max pooling and Convolutional operation down-sample a feature map of an image. By down sampling, the model better understands “WHAT” is present in the image, but it loses the information of “WHERE” it is present.

3.3.3 Deconvolution

As we have performed segmentation to perform skull stripping and tissue segmentation, we need a high-resolution image as an output with all our pixels classified. A regular convolutional network will lose all the information of the location of the information. Hence, we need to up-sample the image to obtain our high-res image. We make use of transposed convolution in UNet.

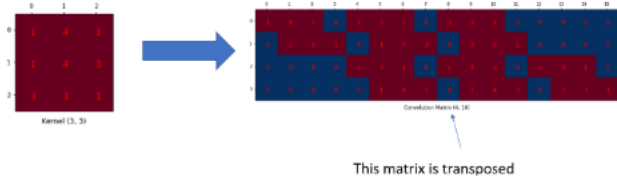


Fig. 3 Deconvolution Operation

Deconvolution is the exact opposite of the normal convolution operation. The 3×3 kernel is rearranged into a 6×16 matrix.

4. OBSERVATIONS AND RESULTS

Table 3 tabulates the DSC results of our cross-validated segmentation results for the HGG and LGG cases, respectively. In our current study, we only compared with three different deep learning-based studies that published recently. All these three studies currently ranked on the top of the BRATS challenge.

By using our method, the enhancing tumor segmentation for the LGG cases by using the T1c images only is not successful. This may be attributed to three reasons: (1) the BBB remains intact in most of these LGG cases and the tumor regions are rarely contrast enhanced; (2) the LGG cohort contains only 54 cases and the training datasets might be insufficient and (3) in these LGG cases, the borders between enhanced tumor and non-enhanced regions are more diffused and less visible that causes problems for both manual delineated ground truth and our fully automated segmentation model. Nevertheless, our method achieved 0.81 DSC for the enhancing tumour segmentation in the HGG cohort.

Fig. 4 displays the boxplots of the calculated Sensitivities and **Fig. 5** shows some exemplar qualitative overlaid segmentation results compared to the ground truth.

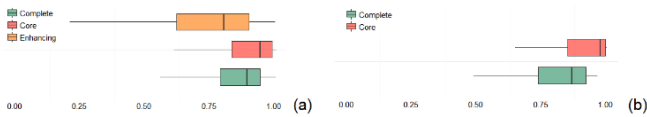


Fig. 4 Boxplots of the Sensitivity; (a) for HGG cases and (b) for LGG cases

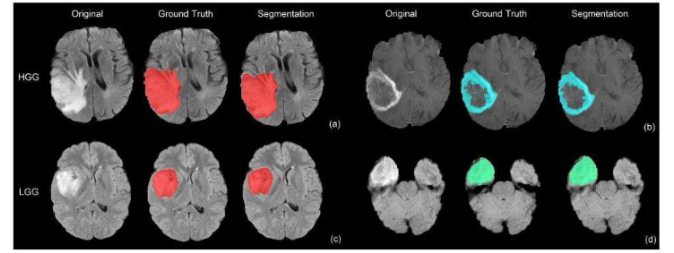


Fig. 5. Segmentations results for the exemplar HGG and LGG cases compared with manual delineated ground truth; (a) segmented complete tumor (red) of a HGG case overlaid on the FLAIR image; (b) segmented enhancing tumor (cyan) of a HGG case overlaid on the T1c im- age; (c) segmented complete tumor (red) of a LGG case overlaid on the FLAIR image; (d) segmented core tumor (green) of a LGG case overlaid on the FLAIR image.

5. CONCLUSION

In this paper, we presented a fully automatic brain tumor detection and segmentation method using the U-Net based deep convolution networks. Based on the experiments on a well-established benchmarking (TCIA) dataset, we have demonstrated that our method can provide both efficient and robust segmentation compared to the manual delineated ground truth. In addition, compared to other state-of-the-art methods, our U-Net based deep convolution networks can also achieve comparable results for the complete tumor regions, and superior results for the core tumor regions. We can envisage a straightforward application on an independent testing datasets and further applications for multi-institutional and longitudinal datasets. By using U-Net for image segmentation we received a DICE loss of 0.3066 on Test Data. The proposed method makes it possible to generate a patient-specific brain tumor segmentation model without manual interference, and this potentially enables objective lesion assessment for clinical tasks such as diagnosis, treatment planning and patient monitoring.

REFERENCES

- [1] B. Hariharan, P. Arbelaez, R. Girshick, and J. Malik, "Object Instance Segmentation and Fine-Grained Localization Using Hypercolumns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 4, pp. 627–639, Apr. 2017, doi: 10.1109/tpami.2016.2578328.
- [2] J. Long, E. Shelhamer, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," *arXiv.org*, 2014, <https://arxiv.org/abs/1411.4038>
- [3] A. Krizhevsky, Ilya Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Nips.cc*, pp. 1097–1105, 2012, [Online]. Available: <https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks>
- [4] "Hough-CNN: Deep Learning for Segmentation of Deep Brain Regions in MRI and Ultrasound. | PUMA," *puma.ub.uni-stuttgart.de*. <https://puma.ub.uni-stuttgart.de/bibtex/22e41e0afa317268511ee835d9b054253/dblp?lang=en> (accessed Apr. 25, 2022).
- [5] Ronneberger, Olaf, et al. "U-Net: Convolutional Networks for Biomedical Image Segmentation." *ArXiv.org*, 18 May 2015, arxiv.org/abs/1505.04597.