Deep Convolutional Neural Networks for Brain Tumour Detection and Analysis

V Vishnu Laxmi Manasa

Computer Science, RMK College of Engineering and Technology, Chennai, India,
vjvlakshmimanasa@gmail.com

Poluru Manasi

Electronics & Communication, BML Munjal University, Haryana, India,

manasipoluru@gmail.com

D Sree Nikitha Reddy

Computer Science, RMK College of Engineering and Technology, Chennai, India,
sreenikithareddy@gmail.com

Abstract: Analysing brain tumour with no human intervention is considered as a vital area of research. However, this has been successfully achieved using deep convolutional neural networks (DCNNs). They have performed exceptionally well in solving computer vision problems and many others such as visual object recognition, detection and segmentation. It is used in detecting the brain tumour by optimising the brain images using segmentation algorithms which are highly resilient towards noise and cluster size sensitivity problems with automatic region of Interest (ROI) detection. One of the main reasons choosing DCNNs is due to its high accuracy and it is not necessary to perform manual feature extraction in these networks. In this research paper we present an extensive review on how convolutional neural networks (CNNs) techniques are applied in brain magnetic resonance imaging (MRI) analysis, their architectures, pre-processing, data-preparation and post-processing strategies. Analysing how different CNNs architectures have evolved, discuss their strategies and examine their outputs is the primary aim of this paper. Finally, we present the future of CNNs in which we mention few research directions in the coming years.

Keywords: Deep Convolutional Neural Network; Brain MRI; Segmentation; Review; Magnetic Resonance

I. INTRODUCTION

Today we live in an era where diseases are increasing day by day and there is a necessity to develop the quality of treatment. The tumour is an irregular lump on any body part and is considered one of the dangerous diseases. Out of all the tumours, brain tumour is the fatal one that can occur in any part of the brain. It is mainly defined as abnormal growth of cells within the brain. These abnormal cells can affect healthy brain cells which in turn results in malfunctioning of the brain. Brain tumours are mainly classified into benign and malignant which are further classified into primary and secondary. A benign tumour is of the primary type and is not cancerous. A malignant tumour is of the second type and stores cancerous cells in

it. It is not an easy task to detect the brain tumour and accurately identify the type. DCCNs performance is better than others because of its wide usage in recognising images. It is basically a group of neurons and has learnable weights. Besides this, they are known for high accuracy and performance. For the above requirements Tensor Flow library is the best choice to use. It is developed by Google and supports DCNNs and other related neural network models. It shows mathematical computation as a graph with edges and nodes where edges represent data which flows from one node to another node. To import the libraries, we use anaconda framework. Subsequently, for training the data a tool of neural networks is implemented with tensor flow.

II. METHODOLOGY

In this article convolutional neural networks are used for segmenting MRI brain tumour using tensor flow. Traditionally, the segmentation was done using several tools like MATLAB, LABVIEW etc. This research article makes use of tensor flow dependent MRI tumour segmentation to improve segmentation accuracy, sensitivity and speed. Python programming is used to segment MRI brain tumour. Python programming is chosen for this research work because Python code is compact and understandable than MATLAB. Its data structure is superior than that of MATLAB. It also provides more graphic packages. There are additional python packages for the implementation process. The flow chart of the work is shown in Fig 1.

III. PROCESS

It involves Data Preparation that is input Dataset and Data Augmentation, Data Pre-processing and Data Splitting, Training the model and Validation and Testing.

A. Data Preparation

Data preparation refers to the operations performed before feeding the data into the network. Data preparation involves taking the data set as the input and augmenting the data. A strategy which helps to increase the diversity of training model's data is Data augmentation. It doesn't need

to collect any new data. For example, in the research we carried out, before data augmentation, the dataset consisted of 155 positive and 98 negative examples, which is 253 example images in total whereas after data augmentation, it has 2065 example images in which there are 1085 positive and 980 examples.
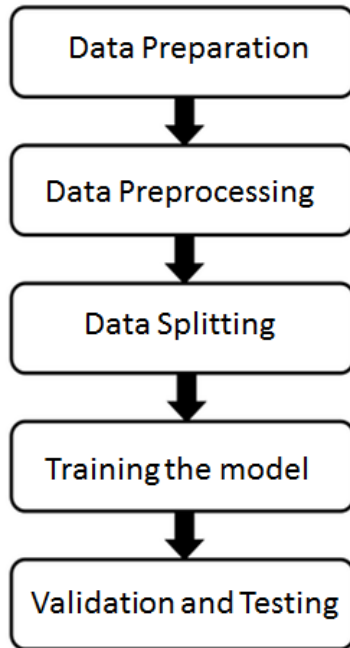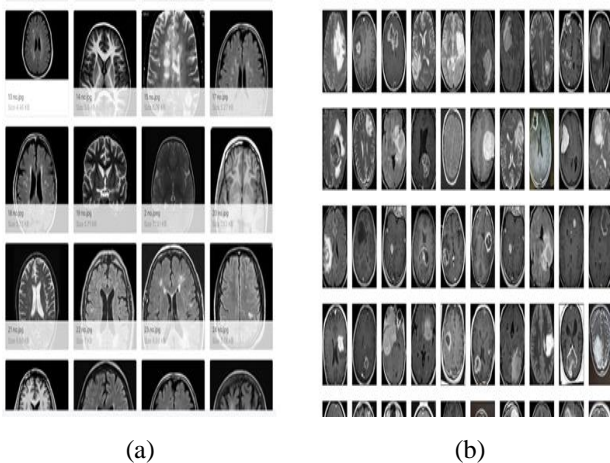

Fig 1.   Flowchart of the work carried

(a)   (b)

Fig 2.   Dataset, (a) Before augmentation, (b) After augmentation

B. Data Pre-processing

The primary goal of pre-processing is to improve the quality of the MR images and make it suitable for further processing by human or machine vision systems. There are improvements in certain MR imaging parameters like visual image enhancing, unnecessary noise removal, edge preserving, improvement in signal to noise ratio and inner part smoothing. It also improves the signal-to-noise ratio

and thus we obtain the clarity of the raw MR images that we fed to the system. Based on sigmoid function, adaptive contrast enhancement is applied.

Skull Stripping

One of the most important procedures in image analysis skull stripping. In this process all the non-brain tissues are discarded in the brain images. It is used to discard other cerebral tissues such as skin, fat and skull in the images. There are few methods to apply this process. This study uses the skull stripping technique that involves a threshold operation to discard skull tissues. Initially, in this process, the pre-processed brain MR image is converted into a binary image with a threshold for the cut-off of 128 being selected. The pixels greater than this threshold are then mapped to white colour, whereas others are coloured as black giving rise to two different regions around the effected tumour tissues. In the next step, morphology is applied which is one of the erosion operations. Finally, the brain MR image mask is obtained by partitioning the image that is eroded and the original image from which the black region is obtained from.
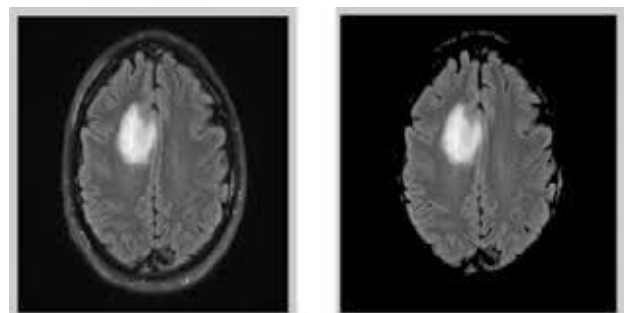

(a)   (b)

Fig 3.   (a) Original image, (b) After skull stripping

C. Data Splitting

The data is split as follows - 70% of the data is allotted as training set, 15% for validation set and rest is allotted for test set.

D. Training the model

As we have finished data pre-processing and splitting, now the model (DCCNs) is trained with the data allotted for training purposes. Once the model has been trained with the training data, test data is used to test the output of the system and the validation data to validate the results. Since DCNNs are used, here is a brief about DCNNs.

CNNs is the most widely used approach of deep neural networks. It can deal array-like data like images and video sequences. The idea behind CNNs is to find compositional hierarchy features, low-level features form patterns. These patterns form intrinsic structures such as shapes, textures etc. Further information about the building blocks of CNN is as follows.

Building blocks of CNNs

a) Convolutional layer

CNNs have the capability to learn the relationships among the pixels of input images by extracting features using convolution and pooling operations. The features that are extracted at each layer vary with respect to their complexities, with the first layer extracting simple features, like edges and the other layers extracting more complex and high- level features. The convolution layer mainly serves three purposes. Initially, the weight-sharing mechanism helps in dealing with the high dimensional data such as 2D images or 3D data. Second, using 2D or 3D kernels, local connectivity of the input topology can be exploited. Finally, using a pooling layer slight shift invariance is achieved. In order to substitute the old conventional layer, very deep advanced CNNs are created in such a way that they use very little computational resources and give highly powerful presentations.

b) Non-linearity layer

Convolutional layer is succeeded by non-linearity operations. Non-linearity is obtained using activation functions. These activation functions help that the representation is mapped to a sparse one. It helps in achieving invariance to data variability and computationally efficient representation. Traditionally, sigmoid and hyperbolic tangent functions were used for this purpose. Recently, for large-scale image recognition, novel activation functions are being suggested. We can describe the frequently used activation functions as follows.

i. Rectified linear units (Re-LUs) and variants

$$f(z\ xy\ lk) = \begin{cases} \max(z\ xy\ lk)\, , 0\ if\ z\ xy\ lk > 0 \\ \min(\alpha \cdot z\ xy\ lk), 0\ if\ z\ xy\ lk \leq 0 \end{cases} \quad (1)$$

where, $z\ xy\ lk$ is the input in the position (x, y) on k^{th} feature map at the lth layer and α is the slope of the negative linear function as shown in (1).

ii. Max out and variants

Max out is used in making optimization better and model averaging performance of dropout training. It is also a generalization of Re-LUs. It is calculated by considering the maximum across K affine feature maps as shown in (2).

$$f(z\ xy\ lk) = \max k \in [1, K] z\ xy\ lk \quad (2)$$

iii. Exponential Linear Units (ELU) and variants

Unlike Re-LUs variants, ELU provides saturated outputs for non-positive inputs. For improving the performance of CNNs, the negative regions with saturation are more useful. It is shown in (3).

$$f(z\ xy\ lk) = \max(z\ xy\ lk)\, , 0 + \alpha \cdot \min(e\ z\ xy\ lk - 1, 0) \quad (3)$$

Different processes of the ELU are monitored by the parameters and are understood using gradient descent. Speed and performance are improved by this ELU parameter in training the deep neural networks.

c) Pooling and subsampling layer

To extract the statistics of the neighborhoods that do not overlap, pooling and subsampling layers are considered. These are used to minimise the number of parameters. Out of all the pooling procedures, max pooling and average pooling are most commonly used. They are similar except that the former uses the maximum activations and the latter uses the average of them. Max pooling is widely used due to its empirical performance.

d) Fully connected layer

In contrast to the convolutional layer, the fully connected (FC) layer has a connection with all the units in the previous layer. This alters the previous layer's 2D structure features into a predefined one-dimensional feature vector. The main purpose of the FC layer is to mine the incoming features to obtain the information about the content of the input image. The process usually involves flattening the feature maps from the convolutional layers. Implementing FC layers requires a large number of parameters than the other layers. Besides, there are two drawbacks in using these kinds of layers. Initially, a single output is produced if it is used as an output layer and accepts fixed-size inputs. The second issue is with extending the network. Either the input images are scaled to fit the requirements of the network or the network is re-factored to be able to process the new data. Solution to these problems lies in the fact that FC layers can be converted to convolutional layers of 1×1 kernels and $1 \times 1 \times 1$ kernels in case of 3D. In this way, the model keeps up with the connected functionality as well as accepting arbitrary input size images.

E. Validation and testing

Validation and testing are carried out to determine the accuracy and performance of the system. They are done using validation and testing datasets. Results obtained by DCNNs are shown in Table 1. To neglect the imbalance of classes of tumours in the database, we have represented mean average precision, recall and F1-score in Table 1.

Therefore, as shown in Table 1 our network obtained better results. The best result is achieved for the record-wise method and for the augmented dataset, and the accuracy is 96.34%. For the subject-wise approach, accuracy was about 88.36% for the augmented dataset. Per image the average test execution was less than 15 milliseconds. Thus, the results obtained show that the system has a good generalization capability and execution speed. Therefore, it works well for medical diagnostics.

Division/ Dataset	Testing Approach	Test accuracy	Average Precision [%]	Average Recall [%]	Average F1-Score [%]
Record-wise/the	10-fold	95.04	94.22	95.25	94.33
Original dataset	one test	97.42	95.12	96.45	96.25
Record-wise/the	10-fold	96.34	95.44	96.84	96.46
Augmented dataset	one test	97.33	97.12	97.23	97.15
Subject-wise/the	10-fold	84.22	81.23	82.45	81.20
Original dataset	one test	90.22	85.69	85.46	85.13
Subject-wise/the	10-fold	88.36	86.12	87.94	86.62
Augmented dataset	one test	91.02	83.12	82.66	81.36

Table 1. Results after testing the system

IV. FUTURE DIRECTIONS

Out of all the application domains, DCNNs have made a large impact. Today, they are in the first place to solve many problems in computer vision, speech recognition and natural language processing (NLP). Due to its advantage of working with 2D and 3D patches, rather than the entire slice or volume, researchers in brain MR image analysis train deep CNNs to obtain accurate segmentation algorithms. This success has a high rate of acceptance by the community. Currently, most CNN architectures have many additional normalisation layers, such as batch normalisation. Additionally, each architecture is becoming increasingly more sophisticated, employing ideas from optimisation and probabilistic models. Although there is a large rate of success with CNNs, they are not fully utilised in brain MRI analysis. When they are trained and tested on images that are of same acquisition characters, performance improves even more. However, the progress in this domain adaptation techniques is not up to the mark and more research in this sense is needed before adapting them in medical centres. Thus, recommended practice to avoid such issues is validating the model on different datasets. This practice helps in designing a robust model that can be applied to datasets of similar applications with different types of MRI scanners, imaging modalities and numbers of training cases. With insufficient training data, the poor spatial resolution of MR images and the need for a reduced prediction time, it has been impossible to train considerably deep CNNs. To train these networks, a considerable effort is required in designing faster methods. The new algorithms that are yet to come are expected not only to improve the performance of deep CNNs but also to be highly optimised, with less hyper-parameters that constitute one of the major bottlenecks for most of the users.

ACKNOWLEDGMENTS

REFERENCES

[1] H. Noh, S. Hong, B. Han, Learning deconvolution network for semantic segmentation, in: Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 1520–1528.

[2] H. Noh, S. Hong, B. Han, Learning deconvolution network for semantic segmentation, in: Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 1520–1528.

[3] M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe, S. Mougiakakou, Lung Pattern Classification for Interstitial Lung Diseases Using a Deep Convolutional Neural Network, IEEE transactions on medical imaging 35 (5) (2016) 1207–1216.

[4] T. Brosch, L. Y. Tang, Y. Yoo, D. K. Li, A. Traboulsee, R. Tam, Deep 3D Convolutional Encoder Networks With Shortcuts for Multiscale Feature Integration Applied to Multiple Sclerosis Lesion Segmentation, IEEE transactions on medical imaging 35 (5) (2016) 1229–1239.

[5] P. Moeskops, M. A. Viergever, A. M. Mendrik, L. S. de Vries, M. J. Benders, I. I˘sgum, Automatic segmentation of MR brain images with a convolutional neural network, IEEE transactions on medical imaging 35 (5) (2016) 1252–1261.

[6] Y. Bengio, A. Courville, P. Vincent, Representation learning: A review and new perspectives, IEEE transactions on pattern analysis and machine intelligence 35 (8) (2013) 1798–1828

[7] A. Hasegawa, S.-C. B. Lo, M. T. Freedman, S. K. Mun, Convolution neural-network-based detection of lung structures, in: Medical Imaging 1994, International Society for Optics and Photonics, 1994, pp. 654–662.

[8] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, in: Advances in neural information processing systems, 2012, pp. 1097–1105.

[9] Y. Sun, Y. Chen, X. Wang, X. Tang, Deep learning face representation by joint identification-verification, in: Advances in Neural Information Processing Systems, 2014, pp. 1988–1996

[10] K. Kamnitsas, C. Ledig, V. F. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon, D. Rueckert, B. Glocker, Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation, Medical Image Analysis 36 (2017) 61–78.

[11] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, et al., Imagenet large scale visual recognition challenge, International Journal of Computer Vision 115 (3) (2015) 211–252

[12] C. Szegedy, S. Ioffe, V. Vanhoucke, A. A. Alemi, Inception-v4, inception-resnet and the impact of residual connections on learning, in: Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, February 4-9, 2017, San Francisco, California, USA., 2017, pp. 4278–4284.

[13] S. Ioffe, C. Szegedy, Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, in: Proceedings of The 32nd International Conference on Machine Learning, 2015, pp. 448–456.

[14] W. Zhang, R. Li, H. Deng, L. Wang, W. Lin, S. Ji, D. Shen, Deep convolutional neural networks for multi-modality isointense infant brain image segmentation, Neuroimage 108 (2015) 214–224

[15] T. Brosch, Y. Yoo, L. Y. Tang, D. K. Li, A. Traboulsee, R. Tam, Deep convolutional encoder networks for multiple sclerosis lesion segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2015, pp. 3–11.

[16] W. Liu, E. CN, Y. Wen, M. SCUT, Z. Yu, M. Yang, Large-Margin Softmax Loss for Convolutional Neural Networks, in: Proceedings of The 33rd International Conference on Machine Learning, 2016, pp. 507–516

[17] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A. L. Yuille, Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs, in: International conference on learning representations, 2015, http://arxiv.org/abs/1412.7062.

[18] T. Paine, H. Jin, J. Yang, Z. Lin, T. S. Huang, GPU asynchronous stochastic gradient descent to speed up neural network training, CoRR abs/1312.6186, http://arxiv.org/abs/1312.6186

[19] J. Ashburner, G. Barnes, C. Chen, SPM8 Manual, http://www.fil.ion.ucl.ac.uk/, [Online; accessed 18 May 2017] (2012)

[20] M. Lai, Deep learning for medical image segmentation, CoRR abs/1505.02000, http://arxiv.org/abs/1505.02000

[21] M. D. Zeiler, ADADELTA: an adaptive learning rate method, CoRR abs/1212.5701, http://arxiv.org/abs/1212.5701.