3D MRI Image Segmentation using 3D UNet Architectures: Technical Review

Vijaya Kamble¹a and Rohin Daruwala²b
¹Research Scholar Department of Electronics Engineering, Veermata Jijabai Technological Institute (VJTI), Mumbai, Maharashtra, India
²Department of Electronics Engineering, Veermata Jijabai Technological Institute (VJTI), Mumbai, Maharashtra, India

Keywords: 3D UNet, MRI Images, Segmentation, Brain Tumors.

Abstract: From last few decades machine learning & deep convolutional neural networks (CNNs) used extensively and have shown remarkable performance in almost all fields including medical diagnostics. It is used in medical domain for automatic tissue, lesion detection, segmentation, anatomical or structure segmentation classification & survival predictions. In this paper we presented an extensive technical literature review on 3D CNN U-Net architectures applied for 3D brain magnetic resonance imaging (MRI) analysis. We mainly focused on the architectures, its modifications, pre-processing techniques, types datasets, data preparation, methodology, GPU, tumor disease types and per architectures evaluation measures in this works. Our primary goal for this extensive technical review is to report how different 3D U-Net architectures or CNN architectures have been used to differentiate between state-of-the-art strategies, compare their results obtained using public/clinical datasets and examine their effectiveness. This paper is intended to present detailed reference for further research activity or plan of strategy to use 3D U-Nets for brain MRI automated tumor diseases detection, segmentation & survival prediction analysis. Finally, we are presenting a novel perspective to assist research directions on the future of CNNs & 3D U-Net architectures to explore in subsequent years to help doctors & radiologist.

1 INTRODUCTION

Over last few decades the use of machine learning and deep learning techniques revolutionized medical imaging field for tumor or disease segmentation, detection & survival predication. It is helping physicians to diagnose brain cancers quickly to boost prognosis. A patient’s MRI is the three-dimensional brain anatomy (Oday Ali Hassen, et al., 2021). MRI images of different modalities such as T1-weighted, T2-weighted, T1c, and Flair as T1c has precise data such as tumor form, location, and scale. Different MRI modalities are used in brain tumor extraction and segmentation. Among all types of brain tumors Gliomas are most common tumors in brain cancer with high mortality rate. These brain tumors originating from the glial cells in the central nervous system. Gliomas are 70% of all brain tumors. The survival duration of patients with high grade gliomas (HGG) lead less than 2 years if prognosis is poor. Compared with HGG, prognosis of low grade gliomas (LGG) are more effective (Chandan Yogananda, et al., 2020).

Different architectures of CNNs used in medical imaging and other applications from year 1990s. Medical Image data is sensitive patients data and not available easily. Earlier limitations on on performance of CNN networks years as less labeled medical data available. But now large annotated medical public & clinical data sets available online & on demand and more powerful graphics processing units (GPUs) available for data processing so this is enabling researchers to continue working in the area to help doctors (Chandan Yogananda, et al., 2020).

Automated or semi automated segmentation methods saving physicians time and provide an accurate reproducible solutions for 3D brain tumor analysis and patient monitoring. Convolutional neural
networks (CNN) able to learn from examples so they demonstrate state-of-the-art segmentation accuracy both in 2D natural images (Andriy Myronenko, 2019) and in 3D medical image modalities. Its difficult to differentiate brain tumors from normal tissues because tumor boundaries are ambiguous and there is a high degree of variability in the shape, location, intensity in homogeneity, or different intensity ranges between the same sequences and acquisition scanners and extent of the patient (Li Sun, et al., 2019). This can influence the segmentation accuracy and correct detection of tumor. Different hospitals show different gray-scale values for the same tumorous cells may when they are scanned differently. Although advance automatic algorithms used for brain tumor segmentation, the problem is still remains a challenging task.

To address issues in this research area we have done extensive comparative review of most cited research papers based on 3D U-Net architectures & 3D medical imaging modalities, with different processing techniques use of powerful GPUs different software’s with various high grade tumors classification segmentation & survival prediction of patients.

Summary of this extensive most cited research is mentioned in table no 1 with reference to paper, few prominent U-Net model parameter & methodology discussed in short with figure. Different imaging modalities, preprocessing techniques datasets, evaluation parameters advantages & limitations also mentioned. Most of the reviewed content got dice scores above 0.75 to 0.89 range for whole tumor core tumors & enhancing tumors. Some of the papers got excellent accuracy, sensitivity & specificity.

2 CNN ARCHITECTURES

CNN architectures used in medical imaging for segmentation detection & predictions of disease diagnosis prognosis. CNN architectures can be grouped around five sub types:

I) Based on interconnected operating modules,
II) Selection of types of input MRI modalities,
III) Selection of input patch dimension,
IV) Number of Predictions at a time
V) Based on implicit and explicit contextual information.

In this summary of the literature review methods distinguish with the different CNN architectures mostly on types of U-Nets, pre processing, post-processing and target of the segmentation & tumor types.

2.1 UNet Architectures Literature Survey

In medical imaging for brain tumor disease diagnosis prognosis for image semantic segmentation and classifications mostly U-Net ResNet architectures are used.

The U-Net is one of the most popular convolutional neural network end-to-end architectures in the field of semantic segmentation.a that is designed for fast and precise segmentation of images. In several challenges U-Net has performed extremely well.

U-Net Architecture split the network into two parts:

Encoder: The encoder path is the backbone. The encoder captures features at different scales of the images by using a traditional stack of convolutional and max pooling layers. A block in the encoder consists of the repeated use of two convolutional layers (k=3, s=1), each followed by a non-linearity layer, and a max-pooling layer (k=2, s=2). For every convolution block and its associated max pooling operation, the number of feature maps is doubled to ensure that the network can learn the complex structures effectively.

Decoder: The decoder path is a symmetric expanding counterpart that uses transposed convolutions. This type of convolutional layer is an up-sampling method with trainable parameters and performs the reverse of (down)pooling layers such as the max pool. Similar to the encoder, each convolution block is followed by such an up-convolutional layer. The number of feature maps is halved in every block. Because recreating a segmentation mask from a small feature map is a rather difficult task for the network, the output after every up-convolutional layer is appended by the feature maps of the corresponding encoder block. The feature maps of the encoder layer are cropped if the
dimensions exceed the one of the corresponding decoder layers.

In the end, the output passes another convolution layer \( (k=1, s=1) \) with the number of feature maps being equal to the number of defined labels. The result is a U-shaped convolutional network that offers an elegant solution for good localization and use of context. Let’s take a look at the code. max pooling operations (in each dimension) are the most appropriate.

In these section the from most cited literature review best architecture discussed. Researchers proposed common U-nets, cascaded U-Nets, modified type of Unet architectures for brain tumor detection & survival predictions.

Xue Feng et al. explained generic 3D U-Net structure with different hyper-parameters, deployment of each model is for full volume prediction and final ensemble modeling. Model fitting done for the survival task feature extraction (Xue Feng, et al., 2019).

---

Yan Hu et al proposed algorithm for intra-tumor structure segmentation using three cascaded U-Net models as shown in figure 2. They are concatenated and further processed by two convolutional layers to detect tumor region. The feature maps generated by three cascaded U-Net models using T1, T1c, T2 and FLAIR modalities. Patches are cropped within tumor region detected for classification model (Yan Hu , et al., 2018).

---

3 EVALUATION MEASUREMENTS

Final Deep CNN models performance majorly depend on the types of dataset, modalities, types of tumors regions, sub-regions & model parameters. In this extensive research survey of 3D UNet & MRI imaging following evaluation metrics were used for segmentation & classification of tumors:

i) Global accuracy,

ii) Dice coefficient,

iii) Recall,

iv) Precision and

v) Hausdorff distance measure.
Following are the 6 Equations for evaluation parameters with the well-known terms False Positives (FP), False Negatives (FN), True Negative(TN):

\[
\text{Global Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}
\]

(1)

\[
\text{Precision} = \frac{TP}{TP + FP}
\]

(2)

\[
\text{Recall} = \frac{TP}{TP + FN}
\]

(3)

\[
F1 = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}
\]

(4)

\[
\text{Dice} = \frac{2 \times TP}{2 \times TP + FP + FN}
\]

(5)

\[
\text{Specificity} = \frac{TN}{TN + FP}
\]

(6)

The main evaluation measures for the challenges mentioned previously are DSC, specificity, sensitivity, positive predictive value (precision), average surface distance (ASD), average volumetric difference (AVD) and modified Hausdorff distance (MHD).

Table 1 summarizes the types of architectures, databases, numbers of samples, MRI modalities considered, tumor diseases types, GPU types, Software's used, evaluation measurements applied and corresponding results reported of the extensive technical surveyed work (Chandan Yogananda, et al., 2020 - http://www.tomography.org/).

4 MEDICAL CNN-BASED SOFTWARES

Now a days most of researchers release their winning competitions or challenges source codes to the public it helps for research in the medical & other fields.

There are few free deep learning libraries for MRI segmentation as listed below:

i) Tensorflow,

ii) Theano,

iii) Caffe,

iv) Keras

v) PyTorch.

There are few CNN open-source frameworks namely NiftyNet 17 and DLTK.

Researchers work on clinical & publicly available datasets depending on their application.

5 CONCLUSION AND FUTURE SCOPE

Automatic brain tumor segmentation for cancer diagnosis & prediction is a challenging task. Most recent advancements in medical diagnostic research using Deep Convolution Neural network 3D Unet architectures discussed in this technical most cited literature review paper. This 3D Unet architectures & modified frameworks indicate significant potential to segment classify & predict the brain tumors lesions from the 3D MRI images. Even though MRI images are of different modalities intensities and categories still complex features from these MRI images can be automatically extracted from 3D Unet architectures it also segment tumor with subregions. There is always chance of improvements and modifications in CNN architectures, Unet architectures to improve the efficiency of segmentation, detection & predictions of cancerous brain tumors.

With this deep technical review we observed and analyzed that most of the proposed methods are based on specific 3D MRI modalities for high grade tumor segmentation so they have computational complexities as well as memory constraints & in need of specific GPU speed for software's. In most of papers deep learning software libraries are used to implement layers of deep CNNs. They are arranged either parallel or distributed or cascaded frameworks, which help researchers to train their models in multi-core architectures or GPUs. Mostly Nvidia GPUs & Intel GPUs used for training and implementation of 3D Unet CNN models. It is observed from evaluation measures that the training and validation for brain image analysis is significantly affected by the data imbalance problem. Lesions are smaller than the entire volume so it affects generalization & robust model. We observed that the full capacity of 3D Unet CNN architectures has not yet been fully leveraged in brain MRI analysis. More sophisticated dedicated softwares are available for Medical imaging or Brain MRI analysis. But there is always a challenge for domain adaptation techniques, more research in this sense is needed for permanent solutions for high grade and low grade tumors for correct diagnosis without experts interventions.
### Table 1: Comparison with different U-Net models.

<table>
<thead>
<tr>
<th>SrNo</th>
<th>Article</th>
<th>Dataset</th>
<th>Number of scans</th>
<th>Model details</th>
<th>GPU</th>
<th>Softwares</th>
<th>Segmentation tasks</th>
<th>Evaluation Measure</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Yue Zhao et al.</td>
<td>BraTS 2018</td>
<td>75 low-grade and 210 high-grade gliomas</td>
<td>3D recurrent multi-fiber network</td>
<td>NVIDIA Tesla V100 32 GB GPU</td>
<td>Pytorch</td>
<td>MS lesion Whole brain, tissue and sub-cortical structure</td>
<td>Dice scores of WT 89.62%, TC 83.65% and ET 78.72%</td>
</tr>
<tr>
<td>2</td>
<td>Chandan Ganesh Bangalore Yogananda et al.</td>
<td>BraTS2017, BraTS2018, Oslo data set 52 LGG and HGG (age &gt; 18 years) scanned from 2003 to 2012</td>
<td>200 cases: 150 HGG and 50 LGG, validation: 63 48 HGG and 17 LGG and 30% (20 12 HGG and 8 LGG)</td>
<td>3D-Dense-UNet Combination of WT-net, TC-net, ET-net</td>
<td>Tesla P100, P40 or K80 NVIDIA-GPUs</td>
<td>Tensorflow, Keras, python package, and Psychiam IDEs with (Adam)</td>
<td>MS lesion Whole brain, tissue and sub-cortical structure, Stroke</td>
<td>Dice-scores for WT 0.90, TC 0.84, and ET 0.80</td>
</tr>
<tr>
<td>3</td>
<td>Oday Ali Hassan et al.</td>
<td>BRATS 2019 and BRATS 2017 At BRATS 2019</td>
<td>3D-MRI of 336 heterogeneous gliomas patients, 259 HGG and 76 Low-Grade Gliomas LGG</td>
<td>Population-based Artificial Bee Colony Clustering (P-ABCC) methodology, K-mean</td>
<td>Intel(R), Core(TM) i3 CPU 2.80 GHz, NVIDIA(R) GTX 1080 Ti GPU, 32 GB RAM</td>
<td>MATLAB R2018a</td>
<td>Brain Tumour</td>
<td>Entire Tumor (WT), Tumor Center (TC), Improved (ET) by 0.03%, 0.03%, and 0.01% respectively. At BRATS 2017, an increase in precision for WT was reached by 5.27%</td>
</tr>
<tr>
<td>4</td>
<td>Jing Huang and Minhui Sun et al.</td>
<td>BRATS 2017</td>
<td>285 patients, 210 HGG images, 75 LGG images</td>
<td>Residual 3D U-net, Dense inception-like architecture with multiple dilated convolutional layers</td>
<td>Intel(R), Core(TM) i3 CPU with 8.00 GB of RAM</td>
<td>Keras</td>
<td>Brain Tumour</td>
<td>Dice scores Similarity WT 0.9089, TC 0.7165, and ET 0.8398</td>
</tr>
<tr>
<td>5</td>
<td>Parvez Ahmad et al.</td>
<td>BRATS 2017</td>
<td>228 training images 57 testing images out of data set is 285.</td>
<td>Clustering technique integrates k-means and the dropping algorithm</td>
<td>Intel(R), Core(TM) i3 CPU with 8.00 GB of RAM</td>
<td>Keras</td>
<td>Brain Tumour</td>
<td>CPU: 7.14 GHz, GPU: NVIDIA(R) GTX 1080 Ti GPU, 32 GB RAM</td>
</tr>
<tr>
<td>6</td>
<td>Hassan A. Kallid et al.</td>
<td>BRATS 2017</td>
<td>163 training subjects, 285 training subjects, 66 subjects were provided as a validation</td>
<td>All ensemble of 3D U-Nets with different hyper-parameters for brain tumor segmentation</td>
<td>Intel(R), Core(TM) i3 CPU with 8.00 GB of RAM</td>
<td>MATLAB software R2018a</td>
<td>Brain Tumour</td>
<td>NVIDIA(R) CUDA(TM) i3 CPU, 8.00 GB RAM, 4 x NVIDIA(R) TITAN Xp GPU, 8G memory</td>
</tr>
<tr>
<td>7</td>
<td>Xue Feng et al.</td>
<td>CBICA’s Image Processing Portal</td>
<td>285 Training cases validation (66 cases) and the testing sets (191 cases)</td>
<td>Encoder-decoder based CNN architecture asymmetrically larger encoded to smaller decoder</td>
<td>NVIDIA Titan Xp GPU with 12 Gb</td>
<td>TensorFlow framework was used with Adam optimizer</td>
<td>Brain Tumour</td>
<td>NVIDIA Titan Xp GPU, 32 GB GPU</td>
</tr>
<tr>
<td>8</td>
<td>Andrey Myroshenko et al.</td>
<td>BraTS 2018</td>
<td>285 subjects, of which 210 are BraTS2018 and 75 are LGG</td>
<td>Separable 3D U-Net</td>
<td>NVIDIA Tesla V100 32 GB GPU</td>
<td>Pytorch toolbox, Adami</td>
<td>Brain Tumour</td>
<td>Dice scores of WT 0.8, 0.8948, and TC of 0.8154</td>
</tr>
<tr>
<td>9</td>
<td>Wei Chen, Boqing Liu et al.</td>
<td>BraTS 2018</td>
<td>285 subjects, of which 210 are BraTS2018 and 75 are LGG</td>
<td>3D Brain SegNet</td>
<td>NVIDIA(R) Titan Xp GPU, 32 GB memory for each GPU</td>
<td>Pytorch</td>
<td>Brain Tumour</td>
<td>Dice Score 0.50 ± 0.22, 0.35 ± 0.27, 0.43 ± 0.27</td>
</tr>
<tr>
<td>10</td>
<td>Xiaojuan Hu et al.</td>
<td>ISLES 2017</td>
<td>3D-Brain SegNet architecture</td>
<td>4 Titan Xp GPUs, 32GB memory for each GPU</td>
<td>NVIDIA(R) Titan Xp GPU, 32 GB memory for each GPU</td>
<td>Pytorch</td>
<td>Brain Tumour, survival prediction</td>
<td>Dice score of WT 0.720, WT 0.878, TC 0.785</td>
</tr>
<tr>
<td>11</td>
<td>Li Sun et al.</td>
<td>BraTS 2018</td>
<td>210 HGG and 75 LGG</td>
<td>Three different CNN architectures (CPU,CNN, DFRKZ Net, 3D U-Net, Whole, Tissue, Edge)</td>
<td>NVIDIA(R) Titan Xp GPU, 32 GB memory for each GPU</td>
<td>Pytorch</td>
<td>Brain Tumour, survival prediction</td>
<td>Dice score of WT 0.918, TC 0.880 and ET 0.854</td>
</tr>
<tr>
<td>12</td>
<td>Dmitry Lachinov et al.</td>
<td>BraTS 2018</td>
<td>285 MRIs for training (210 high grade and 75 low grade glioma images), 67 validation and 192 testing MRIs.</td>
<td>Multiple Encoders Unet, Cascaded U-Net, and 3D U-Net</td>
<td>GTX 1080Ti</td>
<td>MNX framework</td>
<td>Brain Tumour, survival prediction</td>
<td>Dice of WT and TC of 0.8955 and 0.7828 from 0.8799 and 0.7693 of 3D U-Net, respectively</td>
</tr>
<tr>
<td>13</td>
<td>Ping Liu et al.</td>
<td>BraTS 2017</td>
<td>285 samples with manually annotated and confirmed ground truth labels</td>
<td>Deep supervised 3D Squeeze-and-Excitation V-Net</td>
<td>4 NVIDIA Titan Xp 1080 Ti 11GB GPUS</td>
<td>Pytorch</td>
<td>Brain Tumour, survival prediction</td>
<td>Dice scores of WT 0.918, TC 0.880 and ET 0.854</td>
</tr>
<tr>
<td>14</td>
<td>Pawel Mlynski et al.</td>
<td>BraTS 2017</td>
<td>285 scans (210 high grade and 75 low grade gliomas)</td>
<td>CNN-based model-heterogeneity tensor context and the 3D context</td>
<td>Keras, TensorFlow</td>
<td>Brain Tumour, survival prediction</td>
<td>Dice scores of WT 0.918, TC 0.880 and ET 0.854</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Suqing Peng et al.</td>
<td>BraTS 2015</td>
<td>220 HGG and 54 LGG</td>
<td>Multi-Stage 3D U-Net architecture</td>
<td>NVIDIA GeForce GTX 1080Ti GPU</td>
<td>Brain Tumour</td>
<td>Dice Similarity 0.85, 0.72, 0.61</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Mina Ghaffari et al.</td>
<td>BraTS 2018</td>
<td>230 cases for training, and the remaining 55 cases were reserved for testing.</td>
<td>Modified version of the well-known U-Net architecture</td>
<td>x NVIDIA Tesla Pascal P100</td>
<td>Brain Tumour</td>
<td>Dice Similarity 0.87, ET 0.79, WC 0.61</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Parvez Ahmad et al.</td>
<td>BraTS 2018</td>
<td>80% of subjects for training and 20% for validation</td>
<td>3D Dense Dilated Hierarchical Architecture</td>
<td>NVIDIA Tesla V100 GPU</td>
<td>Pytorch</td>
<td>Brain Tumour</td>
<td>Dice Similarity WT 0.8400, TC 0.8574 and ET 0.8219</td>
</tr>
<tr>
<td>18</td>
<td>Shangfeng Lu et al</td>
<td>BraTS2018</td>
<td>259 high-grade gliomas (HGG) and 76 low-grade gliomas (LGG)</td>
<td>Multiple feature extraction 3D CNN</td>
<td>NVIDIA Tesla 1080 Ti 11GB GPUs with 11G RAM</td>
<td>Pytorch</td>
<td>Brain Tumour</td>
<td>Dice Similarity WT 0.881, TC 0.837 ET 0.815</td>
</tr>
<tr>
<td>19</td>
<td>Saqib Qamar et al.</td>
<td>BraTS 2018</td>
<td>210 patients, to train and test our model</td>
<td>3D Hyper-dense Connected Convolutional Neural Network</td>
<td>NVIDIA(R) Tesla P40 or 1080 Ti GPU, 32 GB RAM</td>
<td>NVIDIA(R) Tesla P40 or 1080 Ti GPU, 32 GB RAM</td>
<td>Brain Tumour</td>
<td>Dice Similarity WT 0.81, ET 0.69 and FT 0.55</td>
</tr>
<tr>
<td>20</td>
<td>Yun Hu et al</td>
<td>BraTS 2017</td>
<td>285 training subjects, 46 validation subjects and 146 test subjects</td>
<td>3D Deep neural network</td>
<td>Intel(R), Core(TM) i7 2.10 GHz CPU, NVIDIA(R) GTX 1080 Ti GPU, 32 GB RAM</td>
<td>NVIDIA(R) Tesla P40 or 1080 Ti GPU, 32 GB RAM</td>
<td>Brain Tumour</td>
<td>Dice Similarity WT 0.81, ET 0.69 and FT 0.55</td>
</tr>
</tbody>
</table>
REFERENCES


Shangfeng Lu, Xutao Guo, Ting Ma, Chushu Yang, Tong Wang, Member, Pengzheng Zhou, “Effective Multipath Feature Extraction 3D CNN for Multimodal Brain Tumor Segmentation”, 2019 International Conference on Medical Imaging Physics and Engineering (ICMPE).


