Brain Tumor Segmentation and Surveillance with Deep Artificial Neural Networks

Asim Waqas, Dimah Dera, Ghulam Rasool, Nidhal Carla Bouaynaya, and Hassan M. Fathallah-Shaykh

Abstract

Brain tumor segmentation refers to the process of pixel-level delineation of brain tumor structures in medical images, including *Magnetic Resonance Imaging* (MRI). Brain segmentation is required for radiotherapy treatment planning and can improve tumor surveillance. Automatic segmentation of brain tumors is a challenging problem due to the complex topology of anatomical structures, noise from image acquisition, heterogeneity of signals and spatial/structural variations of tumors. *Machine Learning* (ML) techniques, including *Deep Artificial Neural Networks* (DNNs), have shown significant improvement in classification and segmentation tasks. This chapter provides a comprehensive review of supervised learning models and architectures for image segmentation. A particular emphasis will be placed on *U-Net* and *U-Net* with *Inception* and *dilated Inception* modules for brain tumor segmentation. The performance of the proposed models is evaluated using the multi-modal *BRAin Tumor Segmentation* (BRATS) benchmark dataset. Furthermore, we present a new Bayesian deep learning framework, called *extended Variational Density Propagation* (exVDP), for quantifying uncertainty in the decision of DNNs. In particular, exVDP provides a pixel-level uncertainty map associated with the network's segmentation output. Finally, we present clinical retrospective studies in tumor surveillance using MRI data from patients with glioma and show the advantages accrued from these methods.

Key words: tumor segmentation, tumor surveillance, computer vision, deep supervised learning, *U-Net*, *Inception*, uncertainty propagation.

Asim Waqas, e-mail: waqasa8@students.rowan.edu, Dimah Dera, e-mail: derad6@rowan.edu, Ghulam Rasool, e-mail: rasool@rowan.edu, Nidhal Bouaynaya, e-mail: bouaynaya@rowan.edu,
Rowan University, Henry M. Rowan College of Engineering,

Hassan Fathallah-Shaykh, e-mail: hfshaykh@uabmc.edu,
The University of Alabama at Birmingham, School of Medicine.

1 Introduction

The task of *brain tumor segmentation*, presented in this chapter, is the confluence of multiple techniques usually employed in diverse fields of science such as *Digital Image Processing* (DIP), *Computer Vision* (CV), and *Machine Learning* (ML). ML algorithms, specifically *Deep Artificial Neural Networks* (DNNs), have achieved state-of-the-art accuracy in CV related tasks, including image segmentation. DNNs are built using large stacks of individual artificial neurons, each of which performs mathematical operations of multiplication, summation, and non-linear operations. One of the key reasons for the success of DNNs is the ability to learn useful features automatically from the data as opposite to manual selection by expert humans [1]. Various architectures of DNNs employed for brain tumor segmentation have been discussed along with a case study of one of those architectures in detail. We discuss a new technique for quantifying uncertainty in the output decision of DNNs. We also present different techniques for tumor surveillance.

The rest of the chapter is organized as follows: In section 2, we touch upon the relevant theoretical background of the techniques involved in solving the tumor segmentation problem. We discuss image segmentation in general and medical image segmentation, particularly, and the concept of surveillance in the medical sphere. Section 3 demonstrates brain tumor segmentation through DNNs. Section 4 presents the particularly suited *Inception* modules in *Deep Learning* (DL) for brain tumor segmentation. Section 5 explains the concept of uncertainty estimation in the decision made by DNNs. Finally in Section 6, we discuss tumor surveillance techniques supported by a case study followed by conclusion in Section 7.

2 Theoretical background of the problem

The task of *brain tumor segmentation* using DNNs inherently involves various tasks. Therefore, it is imperative to imbibe some basic theoretical background about these concepts, which leads up to the task-at-hand.

2.1 Image Segmentation

A picture is worth a thousand words. This is because a picture contains far more information in a few pixels that the human brain can process simultaneously as compared to the numerous words that can express the same amount of information sequentially. Thus, understanding the image and extracting useful information from it bears the central role in the fields

of DIP and CV. Classification task, in particular, assigns a label or class to an input image. However, image classification does not provide pixel-level information, such as the location of objects in an image, objects' shapes and boundaries, information about which pixel belongs to which object, etc. For this purpose, images are segmented by assigning a specific label to pixels with similar characteristics in an image. *Segmentation* is a technique frequently used in DIP and CV fields for extracting useful information from images [2]. It is the process of partitioning an image into segments (having sets of pixels) representing various objects in the image. The purpose is to modify the representation of an image into a more elaborate format, which is easy to understand anatomically and helpful in extracting meaningful information for analysis. In usual practice, this process is used to locate objects of interest and draw boundaries/shapes conforming to these objects in an image(s). Image segmentation has contributed to many spheres of human life, ranging from the film-making industry to the field of medicine [3]. For example, the green screens used in *Marvel* [4] movies employed segmentation to extract the foreground objects and place them on different backgrounds depicting dangerous real-life scenes, Fig. 1(a). An example of medical image segmentation includes the identification of multiple organs in the abdomen and thorax, as shown in Fig. 1(b,c). Various techniques are used for classical image segmentation, e.g., *region-based* [5](*threshold segmentation*, *regional growth segmentation*), *edge-detection* [6](*Sobel operator, Laplacian operator*), *clustering-based* [6](*K-means*), and *weak-supervised learning* [7] methods in CNN. Further details on classical image segmentation techniques are beyond the scope of this chapter, so we will confine ourselves to medical image segmentation, in general, with particular focus on brain tumor [8] segmentation.

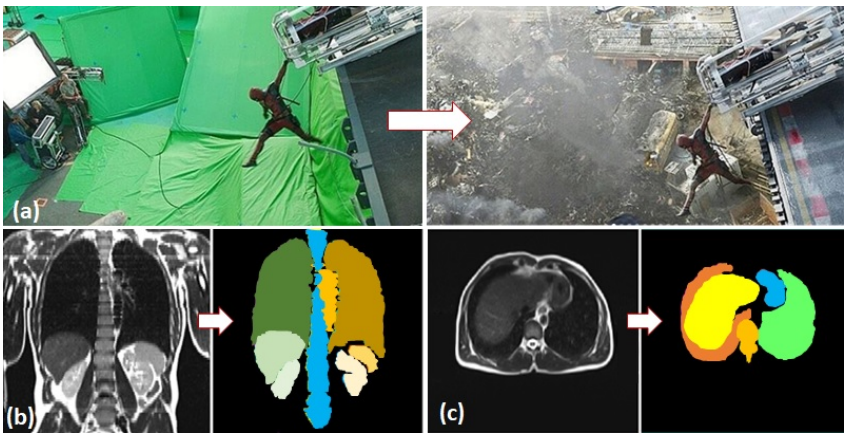

Fig. 1 : Use of Image Segmentation in (a) *Marvel* movies [9], (b) Medical imaging: Segmentation of representative organs in thorax and abdomen from CT images [10]

2.2 Brain Tumor Segmentation

Brain tumors [8] are masses or growths of abnormal cells in the brain, categorized into *primary* and *secondary* or *metastatic* types. Primary brain tumors originate from brain cells, whereas secondary tumors metastasize into the brain from other organs [11]. The most common type of primary brain tumors are *gliomas* [12], which arise from brain glial cells, and can be of *Low-Grade* (LGG) or *High-Grade* (HGG) sub-types. HGGs are aggressively-growing and infiltrative malignant brain tumors, which usually require surgery or radiotherapy and have poor survival prognosis with the highest mortality rate and prevalence [13]. *Magnetic Resonance Imaging* (MRI) is a crucial diagnostic tool for brain tumor analysis, monitoring and surgery planning. Several complimentary 3D MRI modalities such as T1, *T1 with gadolinium-enhancing Contrast* (T1C), T2-weighted (T2), and *FLuid-Attenuated Inversion Recovery* (FLAIR) are acquired to emphasize different tissue properties and areas of tumor spread. For example, in T1C MRI modality, the contrast agent (e.g., gadolinium) emphasizes hyper-active tumor sub-regions.

Brain tumor segmentation [14] is the technique of labeling tumor pixels in an MRI to distinguish them from normal brain tissues and artifacts. These MRI scans are the representation of the internal structure or function of the brain's anatomic region in the form of an array of picture elements called pixels or voxels. It is a discrete representation resulting from a sampling/reconstruction process that maps numerical values to positions of the space. The number of pixels used to describe the field-of-view of an acquisition modality is an expression of the detail with which the anatomy or function can be depicted. Depiction of the numerical value of pixel depends on the imaging modality, acquisition protocol, reconstruction, and post-processing. MRI scans come in various file formats and standards but the six commonly used are: *Analyze* [15], *Neuroimaging Informatics Technology Initiative* (NIfTI) [16], *Minc* [17], PAR/REC format used by Philips MRI scanners [18], *Nearly Raw Raster Data* (NRRD) [19], and *Digital imaging and communications in medicine* (Dicom) [20]. A comparison of characteristics of these formats is shown in Fig. 2. In clinical practice, the process of separating the tumor pixels from normal brain tissues provides useful information about existence, growth, diagnosis, surveillance and treatment planning. The process of manual delineation requires anatomical knowledge by specially trained persons, whereas such manual practices are expensive, time-consuming, and are prone to errors due to human limitations. The process of automated segmentation of brain tumors from 3D images facilitates in overcoming these shortcomings [21].

	Dicom	Nifti	Minc	Analyze	NRRD	PAR/REC format
Stands for	Digital Imaging and Communications in Medicine	Neuroimaging Informatics Technology Initiative	Medical Imaging NetCDF Toolkit	Developed by Biomedical Imaging Resource (BIR) at Mayo Clinic	Nearly Raw Raster Data	Philips MRI scanner format
File extension	.dcm	.nii	.mnc	.img and .hdr	.nrrd	.PAR
Data type	Signed, unsigned integer (8, 16-bit, 32-bit only allowed for radiotherapy dose)	Signed, unsigned integer (8-64 bit),float (32-128 bit), complex(64-256 bit)	Signed, unsigned integer (8-32 bit),float (32, 64 bit), complex(32, 64 bit)	Unsigned integer (8-bit), signed integer(16, 32 bit), float(32-64 bit), complex(64 bit)	-	8 or 16 bit unsigned integers
Header Length	Variable length binary format	Fixed-length 352 byte binary format (348 byte in case of data stored as .img and .hdr)	Extensible binary format	Fixed-length: 348 byte binary format	Extensible with attached and detached	8 or 16 bit unsigned integers

Fig. 2 : Summary of medical imaging file formats

2.3 Tumor Surveillance

National Cancer Institute (NCI), part of the U.S. *National Institutes of Health* (NIH), defines tumor surveillance as *closely watching a patient's condition but not treating it unless there are changes in test results. Surveillance is also used to find early signs that a disease has come back. It may also be used for a person who has an increased risk of a disease, such as cancer* [22]. The process of surveillance regularly involves (scheduled) medical tests and examinations to track the growth of the tumor. The term has also been used in the realm of public health, wherein collective information of a disease, such as cancer, is recorded in a group of people belonging to a specific category (ethnic, age, gender, regional, etc.) Active surveillance is extremely beneficial, especially for patients with low-risk cancer diagnoses. Apart from the routine biopsy, active surveillance is almost surgically noninvasive. It helps in the delay of more invasive treatments such as surgical removal of a tumor, sparing the patients from burdensome side effects and potential complications for as long as possible. Moreover, by deferring the invasive treatment to the point when the disease worsens, active surveillance enables cancer patients to maintain a quality of life. A case in point is the exceptionally beneficial surveillance of low-risk prostate cancer in men. The reason for this success is that almost 50% of prostate cancer diagnoses are categorized as low-risk with less possibility of spread, and few cases may never require advanced forms of treatment. Such cases do not immediately need to be aggressively treated in the absence of worsened disease while the specialists keep records of the tumor's growth over time. Surveillance allows the specialists to monitor the disease right from the onset, thus leveraging

them the liberty to analyze the effects and progress of disease and determine the next course of action [23]. A study by Harvard researchers found that the aggressiveness of prostate cancer at diagnosis appears to remain stable over time for most men. If patients had chosen active surveillance, then this could make them feel more confident in their decision about treatment [24]. Early detection of the tumor through surveillance could assist both the patients and the specialists in taking more considered decisions about treatment.

2.4 Deep Learning Segmentation Task

Computer Vision (CV) is the field of computer science that aims to replicate (to some extent) the complex nature of the human vision system into modern-day computers and machines. It endeavors to enable machines to visually gaining a high-level understanding of objects in the imagery in its quest to mimic the humans. A chronological insight in some of the most active topics of research in computer vision can be found in [25]. Most attractive topics in today's CV tasks include object classification (i.e., categorizing objects in an image), localization (i.e., spatially locating objects in an image), detection/ recognition and segmentation (i.e., identifying the category of each pixel in an image) as shown in Fig. 3.

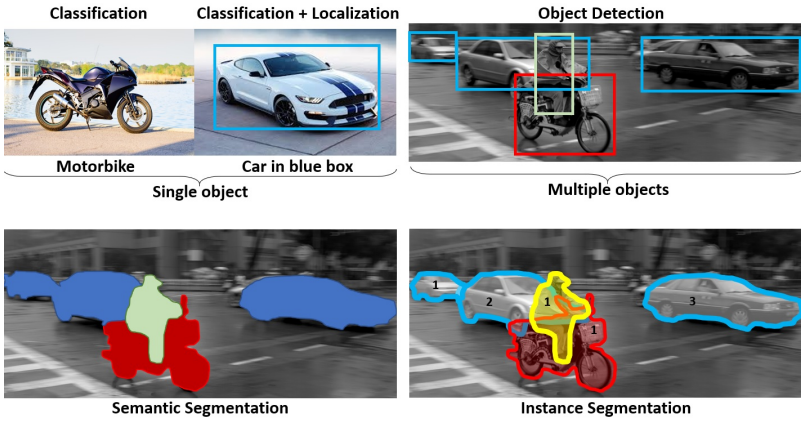

Fig. 3 : Important Computer Vision tasks.

The commonly used classical image segmentation techniques, described in section 1.1 above have been replaced by their more efficient counterparts in ML because of the former's inherently rigid algorithms and the need for human intervention. Image segmentation is the fundamental component

of DL, which is part of a broader family of ML. Compared with other DL algorithms, CNNs have proven to be the more efficient selection for segmentation tasks from imagery. Image segmentation using CNN involves feeding the CNN with the desired image as an input and getting the labels of each pixel, i.e., labeled image as an output. Instead of processing the complete image at once, CNN deals with a fraction of image conforming to its *filter*, convolving and ultimately mapping over the entire image. To learn more on CNNs, a concise explanation supported by a visualization can be referred to at [26, 27].

2.5 Motivation

A large population suffers from fatalities caused by cancer, and brain tumors are one of the leading causes of death for cancer patients, especially children and adolescents. Brain tumors account for one in every 100 cancers diagnosed annually in the United States [28]. In 2019, the *American Cancer Society* reported that 23,820 new brain cancer cases were discovered in the United States [29]. One of the most frequent primary brain tumors is glioma [30], which affects the glial cells of the brain as well as the surrounding tissues. The HGG or *GlioBlastoMa* (GBM) is the most common and aggressive type with a median survival rate of one to two years [31]. Although neurosurgery may be the only therapy for many brain tumors [32], other treatment methods such as radiotherapy and chemotherapy are also used to destroy the tumor cells that cannot be physically resected or to slow their growth. Before the treatment through chemotherapy, radiotherapy, or brain surgeries, there is a need for medical practitioners to confirm the boundaries and regions of the brain tumor and determine where exactly it is located and the exact affected area. Moreover, all of these invasive treatments face challenging practice conditions because of the structure and nature of the brain. These conditions make it very difficult to distinguish the tumor tissue from normal brain parenchyma for neurosurgeons based on visual inspection alone [33].

Moreover, such manual-visual practices usually involve a group of clinical experts to define the location and the type of the tumor accurately. This lesion localization process is laborious, and its quality depends on the physicians' experience, skills, slice-by-slice decisions, and the results may still not be universally accepted among the clinicians. Treatment protocols for high-grade pediatric brain tumors and general low-grade tumors recommend regular follow-up imaging for up to 10 years. For these longitudinal studies, a comparison of the current MRI with all prior imaging takes a very long time, which is practically infeasible. Automated computer-based segmentation methods present a excellent solution to the challenges mentioned above by saving physician's time and providing reliable and accurate

results while reducing the diagnosis efforts of surgeons on a single patient [34]. Brain tumor segmentation is motivated by assessing tumor growth, treatment responses, computer-based surgery, treatment of radiation therapy, and developing tumor growth models. Thus, the computer-assisted diagnostic system is meaningful in medical treatments to reduce the workload of doctors and to get accurate results.

2.6 Challenges

Segmentation of gliomas in pre-operative MRI scans — conventionally done by expert board-certified neuro-radiologists and other physicians — provides quantitative morphological characterization and measurement of glioma sub-regions. The quantitative analysis task is challenging due to the high variance in appearance and shape, ambiguous boundaries and imaging artifacts. Although computer-aided techniques have the advantage of fast speed, consistency in accuracy and immunity to fatigue [35], automatic segmentation of brain tumors in multi-modal MRI scans is still one of the most difficult tasks in medical image analysis and applications. Automatic segmentation involves dealing with a complicated and massive amount of data, artifacts due to patient's motion, limited acquisition time, and soft tissue boundaries that are usually not well defined. Moreover, many classes of tumors have a variety of irregular shapes, sizes, and image intensities, especially the surrounding structures of tumors. Numerous attempts have been made in developing ML algorithms for segmenting normal and abnormal brain tissues using MRI images, which will be covered in detail in section 2. However, feature selection to enable automation is challenging and requires a combination of computer engineering and medical expertise. Thus, developing fully-automated brain tumor segmentation remains a challenging task, and a large part of the research community is currently involved in overcoming these challenges in bringing state-of-the-art ideas in this field into reality.

3 Brain Tumor Segmentation Using Deep Artificial Neural Networks

The task of segmenting brain tumor in MRI images has been adopted in DNNs from the image segmentation task in CV. This section focuses on these techniques imported in DL from CV and also gives an overview of the various DL architectures employed on brain MRI datasets.

3.1 Image Segmentation in Computer Vision Realm

Image segmentation is the task of finding groups/ clusters of pixels that belong to the same category. It divides an input image into *segments* to simplify image analysis. These segments represent objects or parts of objects and comprise sets of pixels belonging to each part. Practically, the segmentation sorts pixels into larger components, eliminating the need to consider individual pixels as units of observation. In statistics, this problem is known as *cluster analysis* and is a widely studied area with many different algorithms [36, 37, 38, 39]. In CV, image segmentation is one of the oldest and most extensively used problems dates back to the 1970s [40, 41, 42, 43, 44, 45]. Some of the most extensively known techniques developed for image segmentation are: (a) active contours [46]; (b) level sets [47]; (c) region splitting and graph-based merging [48]; (d) mean shift (mode finding) [49]; (e) normalized cuts (splitting based on pixel similarity metrics, as depicted in Fig. 4. The segmentation process itself has two forms, namely; *semantic*, and *instance* segmentation. The former classifies all the pixels of an image into meaningful or *semantically interpretable* classes of objects and is usually referred to as *dense prediction*. The latter identifies each instance of each object in an image and differs from semantic segmentation in that it does not categorize every pixel. For example, in Fig. 3, semantic segmentation classified all cars, while instance segmentation identifies each one individually. Various metrics are used for performance evaluation of image segmentation including pixel accuracy P_{acc}, mean accuracy M_{acc}, *Intersection-over-Union* (IoU) M_{IU}, frequency weighted IoU F_{IU} and *Dice coefficient* [50]. Let n_{ij} indicate the number of pixels of class i predicted to belong to class j, where there are n_{cl} different classes, and let $\mathbf{t}_i = \sum_j n_{ij}$ indicates the number of pixels of class i, then the performance evaluation terms mentioned above are defined by:

$$P_{acc} = \frac{\sum_i n_{ii}}{\sum_i t_i}, \tag{1}$$

$$M_{acc} = \frac{1}{n_{cl}} \sum_i \frac{n_{ii}}{t_i}, \tag{2}$$

$$M_{IU} = \frac{1}{n_{cl}} \sum_i \frac{n_{ii}}{t_i + \sum_j n_{ji} - n_{ii}}, \tag{3}$$

$$F_{IU} = \frac{1}{\sum_k t_k} \sum_i \frac{t_i n_{ii}}{t_i + \sum_j n_{ji} - n_{ii}}, \tag{4}$$

Dice Similarity Coefficient (DSC) has also been extensively used for evaluating segmentation algorithms in medical imaging applications [51]. DSC between a predicted binary image P and ground truth binary image G, both of size N x M is given by:

$$DSC(P,G) = 2\frac{\sum_{i=0}^{N-1}\sum_{j=0}^{M-1} P_{ij}G_{ij}}{\sum_{i=0}^{N-1}\sum_{j=0}^{M-1} P_{ij} + \sum_{i=0}^{N-1}\sum_{j=0}^{M-1} G_{ij}}, \tag{5}$$

where i and j represent pixel indices for the height N and width M. The range of DSC is [0, 1], and a higher value of DSC corresponds to a better match between the predicted image P and the ground truth image G.

The application of image segmentation techniques in the medical imaging field opened a new frontier of knowledge with advances in the areas of diabetic retinopathy detection, skin cancer classification, brain tumor segmentation and many more. In this chapter, we will restrict ourselves to brain tumor segmentation only and look at the various techniques employed in *Artificial Neural Networks* (ANNs) for brain tumor segmentation.

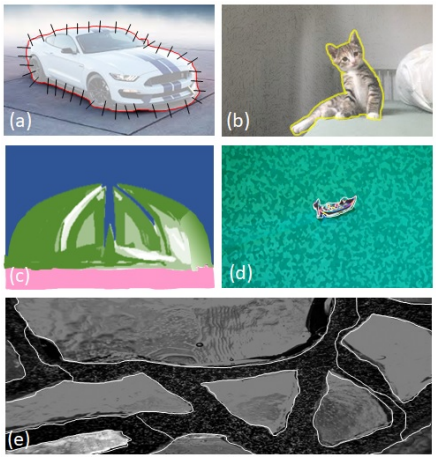

Fig. 4 : Few image segmentation techniques in computer vision: (a) *active contours* [46]; (b) *level sets* [47]; (c) *region splitting* and *graph-based merging* [48]; (d) *mean shift* (mode finding) [49]; (e) *normalized cuts* (splitting based on pixel similarity metrics) [52].

3.2 Deep Artificial Neural Networks and Image Segmentation

DNNs have achieved significant milestones in the CV field. DNNs have multiple layers between the input and output layers. The basic element of ANN, i.e., *artificial neuron*, has multiple inputs that are weighted and summed up, followed by a *transfer function* or *activation function*. Then the neuron outputs a scalar value. An example of ANN is illustrated in Fig. 5 [53]. Inspired by biological processes, ANNs use shared-weight architecture where the connectivity pattern between neurons mimics the organization of

the brain visual cortex [54, 55]. ANNs imitate the concept of receptive fields where individual cortical neurons respond to stimuli only in a restricted field of view. Because of their shared-weight architecture and translation invariance characteristics, ANNs are shift or space-invariant. Due to the linear operations followed by the non-linear activations, ANNs are capable of extracting higher-level representative features [56] and can compute any function [57].

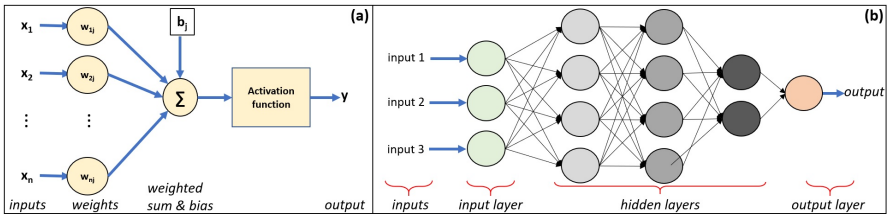

Fig. 5 : (a) artificial neuron model, (b) ANN model.

3.3 DL-based Image Segmentation Architectures

Most prominent DL architectures used by the CV community include *Convolutional Neural Networks*(CNNs), *Recurrent Neural Networks* (RNNs) and *Long Short Term Memory* (LSTM), *encoder-decoders*, and *Generative Adversarial Networks* (GANs) [58, 59, 60, 61]. With our focus on CNNs, let us discuss the three important DL-based image segmentation architectures.

3.3.1 Convolutional Neural Networks

Waibel *et al.* introduced *Convolutional Neural Networks* (CNN) that had weights shared among temporal receptive fields, and it had back-propagation training for phoneme recognition [62]. LeCun *et al.* developed a CNN architecture for document recognition as shown in Fig. 6a [58]. The three basic components/ layers of a CNN are: 1) *convolutional layer*, having a kernel (or filter) of weights convolved with the input image to extract features; 2) *nonlinear layer* having an element-wise activation function applied to feature maps; and 3) *pooling layer*, which reduces spatial resolution and replaces appropriate neighborhood of a feature map with some statistical information (mean, max, etc.) [63]. Deep CNNs have performed extremely well on a wide variety of medical imaging tasks, including diabetic retinopathy detection [64], skin cancer classification [65], and brain tumor segmentation [66, 67, 68, 69, 70]. Some of the most well-known CNN architectures include *AlexNet* [71], *VGGNet* [72], *ResNet* [73], *GoogLeNet* [74] which use

Inception modules architecture (explained in detail in section 3), *MobileNet* [75], and *DenseNet* [76].

3.3.2 Fully Convolutional Networks

Fully Convolutional Networks (FCNs), proposed by Long *et al.*, use convolutional layers to process varying input sizes [50]. It was one of the first DL models for semantic segmentation. As shown in Fig. 6b, the final output layer of FCN has a large receptive field and corresponds to the height and width of the image, while the number of channels corresponds to the number of classes. The convolutional layers classify every pixel to determine the context of the image, including the location of objects. FCNs have been applied to a variety of segmentation problems, such as brain tumor segmentation [68], instance-aware semantic segmentation [77], skin lesion segmentation [78], and iris segmentation [79].

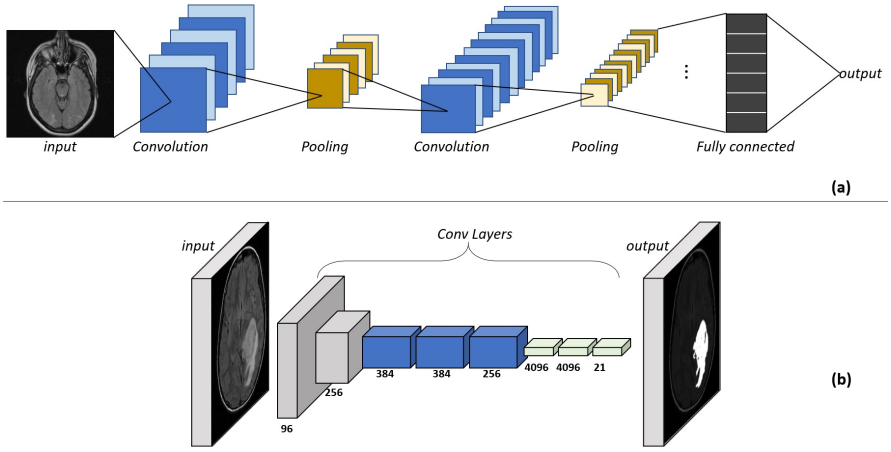

Fig. 6 : Basic architecture of (a) CNN, (b) FCN.

3.3.3 Encoder-Decoder Based Models

Encoder-decoder models are inspired by the FCNs, and the most well-known architecture of these models are *U-Net* and *V-Net* [80, 81]. *U-Net* was proposed for segmenting biological microscopy images, and it used the data augmentation technique to learn from the available annotated images more effectively. The *U-Net* architecture consists of two parts; a *contracting* or *down-sampling* path to capture the context, and a symmetric *expanding*

or *up-sampling* path for localization of the captured context. The contracting path has FCN-like architecture that extracts features with 3×3 convolutions while increasing the number of feature maps and reducing their dimensions. Contrarily, the expanding path carries out deconvolutions by reducing the number of feature maps while increasing their dimensions. The feature maps from the contracting path are concatenated to the expanding path to maintain the integrity of pattern information. Finally, a segmentation map is generated from feature maps by 1×1 convolution operation that categorizes each pixel of the input image. *U-Net* was trained on 30 transmitted light microscopy images, and it won the *International Symposium on Biomedical Imaging* (ISBI) cell tracking challenge in 2015 by a large margin. *V-Net* [81] is another well-known FCN-based model proposed for 3D medical image segmentation. It introduced a new objective function based on the Dice coefficient, which enabled the model to deal with strong class imbalance between the number of voxels in the foreground and the background. *V-Net* was trained end-to-end on MRI volumes depicting prostate, and it learned to predict segmentation for the whole volume at once.

3.3.4 Other Deep Learning Models used in Image Segmentation

In addition to the models described in previous sections, there are many families of DL architectures that are very popular for medical image segmentation. For example, convolutional graphical models (incorporating concepts of *Conditional Random Fields* (CRFs) and *Markov Random Field* (MRFs)), *Multi-scale pyramid network* models (*Feature Pyramid Network* (FPN)) [82], *Pyramid Scene Parsing Network* (PSPN) [83], *Regional CNN* (R-CNN) like *Fast* R-CNN, *Faster* R-CNN, and *Mask-RCNN*, dilated or atrous convolution (*DeepLab Family* [84]), RNN-based models (ReNet [85], *ReSeg* [86]), *Data-Associated* RNNs (DA-RNNs) [87]), and *attention-based* models (OCNet [88], *Expectation-Maximization Attention Network* (EMANet) [89], *Criss-Cross attention Network* (CCNet) [90]). Minaee *et al.* has presented an elaborate review reference of all of these models [63].

3.4 Brain Tumor Segmentation Task Challenge

In this section, we discuss the brain tumor segmentation task in the realm of DL. Internationally held challenges on medical imaging analysis have become the standard for validation of the proposed methods. *Brain Tumor Segmentation* (BraTS) Challenge [91] is one such challenge that is held in conjunction with *Medical Imaging Computing and Computer-Assisted Intervention* (MICCAI) conference [92]. The first challenge workshop was held

in 2012, followed by yearly benchmarks held with MICCAI conferences. BraTS challenge evaluates state-of-the-art segmentation methods of brain tumors in MRI scans. It has a publicly available dataset (with accompanying expert delineations), which is used for benchmarking the submitted contenders for segmenting *multi-institutional pre-operative* MRI (mpMRI) scans having intrinsically heterogeneous (in appearance, shape, and histology) brain tumors, namely *gliomas*. In addition, this challenge also encompasses the survival prediction of the patient and evaluates the algorithmic uncertainty in tumor segmentation. The challenge evaluates segmentation of tumor sub-regions of *Enhancing Tumor* (ET), *Tumor Core* (TC), and *Whole Tumor* (WT) as shown in Fig. 7. ET are regions of hyper-intensity in T1C when compared to T1, and TC is the bulk of the tumor, that is typically resected. The TC involves ET and the necrotic (fluid-filled) and the non-enhancing (solid) parts of the tumor. WT is the complete extent of the disease, as it is comprised of the TC and the peritumoral *EDema* (ED), depicted by FLAIR. The dataset for the 2018 challenge consisted of a total of 542 patients, with 285 for training, 66 for validation, and 191 for testing scans having 210 *High-Grade Glioma* (HGG) and 75 *Low-Grade Glioma* (LGG) patients with annotations approved from experienced neuro-radiologists through a hierarchical majority vote. The data consists of clinically-acquired 3T multi-contrast MR scans from around 19 institutions, with ground truth labels by expert board-certified neuro-radiologists in NIfTI files (.nii.gz). *Dice coefficient* and *Hausdorff distance* (95%) have been used as evaluation schemes. Apart from these, *Sensitivity* and *Specificity* are also used as metrics. An assessment of state-of-the-art ML methods used for brain tumor segmentation under the BraTS challenge from the period 2012-2018 has been compiled by Bakas *et al.* [93].

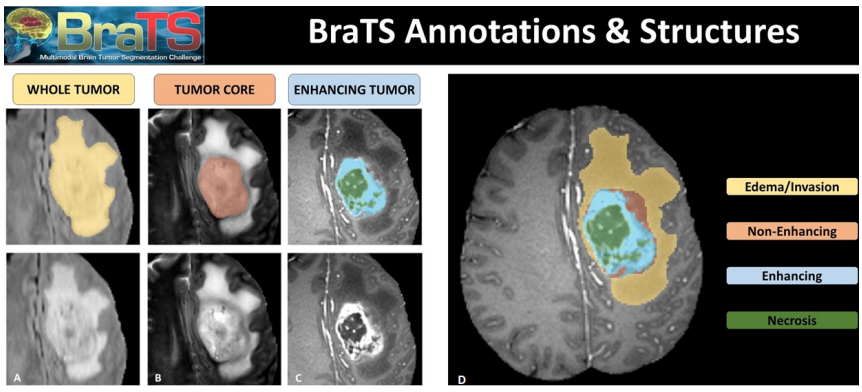

Fig. 7 : Image patches with annotated tumor (glioma) sub-regions. (A) Whole tumor (yellow) visible in T2-FLAIR, (B) Tumor core (orange) visible in T2, (C) Enhancing tumor (blue) surrounding the cystic/necrotic core (green) visible in T1c, (D) Combined segmentations [94].

4 Inception Modules in Brain Tumor Segmentation

After assimilating the brain tumor segmentation problem using DNNs, let us now look at one of the architectures that has appreciable accuracy in solving this problem. This architecture is based on *U-Net* with *Inception* modules.

4.1 Brain Tumor Segmentation Using Inception and Dilated Inception modules

Cahall *et al.* proposed an image segmentation framework for tumor delineation that benefits from two state-of-the-art ML architectures in CV: *Inception* modules and *U-Net* [70, 74, 80]. This new framework includes two learning regimes, i.e., learning to segment intra-tumoral structures (necrotic and non-enhancing tumor core, peritumoral edema, and enhancing tumor) or learning to segment glioma sub-regions (WT, TC, ET). Both learning regimes are described in section 2.4 above. These learning regimes were incorporated into a modified loss function based on the DSC described in section 2.1 equ. 5 above.

U-Net was originally developed for cell tracking. However, it has been applied recently to other medical segmentation tasks, such as, brain vessel segmentation [95], brain tumor segmentation [96], and retinal segmentation [97]. To tackle different medical imaging segmentation problems, variations and extensions of *U-Net* have also been proposed, such as 3D *U-Net* [98, 99], H-DenseUNet [100], RIC-UNet [101], and Bayesian *U-Net* [102]. Cahall *et al.* used a cascade learning approach in which three different models were used first to learn the WT, then TC, and finally, ET resulting in a proposed end-to-end implementation for all tumor sub-types [70].

4.2 BraTS Dataset and Pre-processing

We used BraTS 2018 dataset, described in section 2.4, for experiments [94, 51, 103, 93]. The dataset contains four sequences for each patient's MRI (T1, T1c, T2, and FLAIR) images. It also contains ground truth in the form of pixel-level manual segmentation markings for three intratumoral structures: necrotic and non-enhancing tumor core (labeled as 1), peritumoral edema (labeled as 2), and enhancing tumor (labeled as 4). The glioma sub-regions have been defined as WT having all three intratumoral structures (labeled as $(1 \cup 2 \cup 4)$), TC containing all except peritumoral

edema (labeled as $(1 \cup 4)$), and ET (labeled as 4). Different sequences provide complementary information for identifying the intratumoral structures:

- FLAIR highlights the peritumoral edema.
- T1c distinguishes the ET.
- T2 highlights the necrotic and non-enhancing tumor core.

BraTS images have been pre-processed for skull-stripping, re-sampled to an isotropic 1 mm^3 resolution, and co-registered all four modalities of each patient. Cahall *et al.* [70] performed additional pre-processing in the following order:

1. Discard excess background pixels from images by obtaining the bounding box of the brain and extracting the selected portion, effectively zooming on the brain.
2. Re-size the cropped image to 128×128 pixels.
3. Drop the images having no tumor regions in the ground truth segmentation.
4. Apply intensity windowing function to each image such that the lowest 1% and the highest 99% pixels were mapped to 0 and 255, respectively.
5. Normalize images by subtracting the mean and dividing by the standard deviation of the dataset.

4.3 Deep Artificial Neural Network Architectures

In medical imaging, semantic segmentation's accuracy depends on the ability to extract the local structural as well as global contextual information from MRI scans while training the model. For this reason, many multi-path architectures in the context of medical imaging have been proposed, and all of them extract the structural and contextual information from input data at multiple scales [98, 104, 105]. This features extraction-aggregation concept at various scales was also done in *Inception* modules [74]. However, the feature extraction mechanism in the *Inception* module is different from the multi-path architectures. The *Inception* module applies filters of various sizes at each layer and concatenates resulting feature maps [74]. Cahall *et al.* [70] proposed a modified version of *Dilated Residual Inception* (DRI) [106] based on *U-Net* and factorized convolution *Inception* module [80, 74]. DRI's special blocks were inspired from *Inception* module [107] and dilated convolution [108]. DRI has fewer parameters than the original *Inception* module and employs residual connections to alleviate the vanishing gradients problem at a faster convergence rate [73]. *MultiResUNet* combined a *U-Net* with residual *Inception* modules for multi-scale feature extraction, applying the architecture to several multimodal medical imaging datasets [109]. Integration of *Inception* modules with *U-Net* has been evaluated for

left atrial segmentation [110], liver and tumor segmentation [111], and brain tumor segmentation [112].

4.3.1 *Inception* Module

The convolutional layer in the proposed *Inception* module [70] in the original *U-Net* was replaced with an *Inception* module having multiple sets of 3×3 convolutions, 1×1 convolutions, 3×3 max pooling, and cascaded 3×3 convolutions as depicted in Fig. 8(B). At each layer on the contracting path, the height and width of the feature maps are halved, and the depth is doubled until reaching the bottleneck, i.e., the center of the U. On the corresponding expanding path at each layer, the height and width of feature maps are doubled, and the depth is halved until having the segmentation mask as the output. As with *U-Net*, feature maps generated on the contracting path are concatenated to the corresponding expanding path. The authors employed a *Rectified Linear Unit* (ReLU) as the activation function, with batch normalization [113] in each *Inception* module. The architecture setting receives an input image of size $N \times M \times D$ and outputs an $N \times M \times K$ tensor where $N = M = 128$ pixels, $D = 4$ represents the four MRI modalities (T1, T1c, T2, FLAIR), and $K = 3$ represents the segmentation classes (intra-tumoral structures or glioma sub-regions). The output image of K slices is a binary image representing the predicted segments for the i^{th} class ($0 \leq i \leq K - 1$). Pixel-wise activation functions (sigmoid [114] for *glioma* and *softmax* [114] for intra-tumoral structures) are used to generate the output binary images.

4.3.2 Dilated Inception *U-Net*

Another useful architecture, called *Dilated Inception U-Net* (DIU-Net), integrates dilated or astrous convolutions [84] and *Inception* modules in the *U-Net* architecture [115] as shown in Fig. 8(A). Here, each *dilated Inception* module consists of three 1×1 convolutional operations, followed by one l-dilated convolutional filter (with $l = 1, 2, 3$), as illustrated in Fig. 8(C). The 1×1 convolutional filters perform dimensionality reduction, while three l-dilated convolutional filters each of size 3×3 implement atrous convolutions. In dilated convolutions, an image I of size $m \times n$ and a discrete convolutional filter w of size $k \times k$ are convolved by:

$$(I * w)(p) = \sum_{s} I\left[p + s\right] w\left[s\right]. \tag{6}$$

Simple convolution operation of equ. 6 can be generalized to l-dilated convolution ($*_l$) as [108]:

$$(I *_l w)(p) = \sum_s I\left[p + l_s\right] w\left[s\right]. \tag{7}$$

For $l = 1$, we get the simple convolutional operation of equ. 6. For $l > 1$, l-1 zeroes are inserted between each filter element, creating a scaled and sparse filter of size $k_\mathrm{s} \times k_\mathrm{s}$, where k_s is defined by:

$$k_s = k + (k-1)(l-1), \tag{8}$$
$$= l(k-1) + 1. \tag{9}$$

The scaling s increases the receptive field of the filter by a factor $\frac{k_\mathrm{s}}{k}$.

$$\frac{k_s}{k} = \frac{k + (k-1)(l-1)}{k}, \tag{10}$$
$$= l + \left(\frac{-l+1}{k}\right). \tag{11}$$

The receptive field of the filter increases linearly with l, while the number of elements $(k \times k)$ remains fixed.

4.3.3 Modified DSC as Objective/Loss Function

Cahall *et al.* used a modified version of DSC (equ. 5) as an objective/ loss function, after incorporating three changes: (1) the sign of DSC was changed to convert it into a minimization problem, (2) a log function was introduced, and (3) a new parameter γ was used to cater for extremely large values of the loss function [70]. From initial experiments, it was empirically observed that $\gamma = 100$ provided the best segmentation performance. Modified DSC as a loss function for a binary class (tumor or not tumor) and multi-class (for K classes) are given in the following two equations:

$$\mathcal{L}_{DSC}(P,G) = -log\left[2\frac{\sum_{i=0}^{N-1}\sum_{j=0}^{M-1} P_{ij}G_{ij} + \gamma}{\sum_{i=0}^{N-1}\sum_{j=0}^{M-1} P_{ij} + \sum_{i=0}^{N-1}\sum_{j=0}^{M-1} G_{ij} + \gamma}\right], \tag{12}$$

$$\mathcal{L}_{DSC}(P,G) = -log\left[\frac{1}{K}\sum_{i=0}^{K-1} DSC(P_i, G_i)\right]. \tag{13}$$

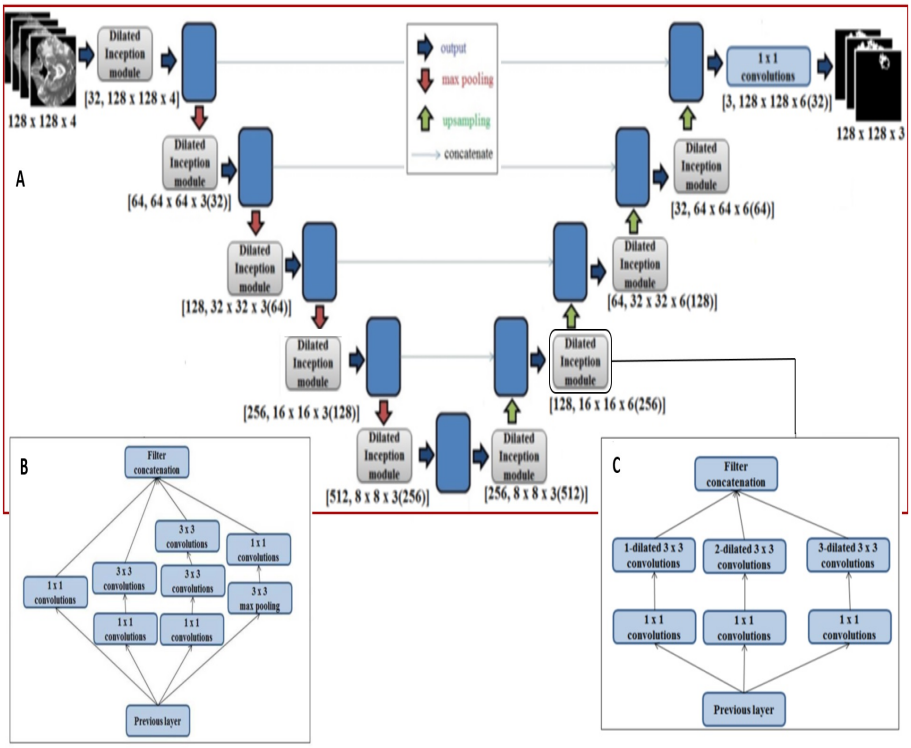

Fig. 8 : (A) DIU-Net architecture with contracting and expanding path and a bottleneck in the middle. On the contracting path, the multiplication by 3 indicates three l-dilated convolutional filters. On expanding path, concatenation of feature maps from contracting path doubles the depth of output feature map, hence the multiplication by 6. (B) *Inception* module architecture, (C) *Dilated Inception* module with three l-dilated convolutional filters and 1x1 dimensional reduction convolution filters.

4.4 Experimental Setup and Results

Four different models were trained by Cahall *et al.* [70], two for the *U-Net* architecture (intra-tumoral structures and glioma sub-regions), and two for the *U-Net* with *Inception* module (intra-tumoral structures and glioma sub-regions). All four models were trained using k-fold *cross-validation* on the dataset that was randomly split into k mutually exclusive subsets of equal or near-equal size. Each algorithm was run k times subsequently, and each time one of the k splits was taken as a validation subset and the rest as the training subset. *Stochastic gradient descent* [116] with an adaptive moment estimator (*Adam*[116]) was used for training all models and their variations [117]. With a batch size of 64 and 100 epochs, the learning rate was initially set to 10^{-4}, which was exponentially decayed every 10 epochs. All

learnable parameters (weights and biases) were initialized based on the He initialization method [118]. *Keras* [119] *Application Programming Interface* (API) with *TensorFlow* [120] backend was used for implementation, and all models were trained on a Google Cloud Compute [121] instance with 4 NVIDIA TESLA P100 *Graphical Processing Units* (GPUs).

4.4.1 Results from *Inception* Modules

For intra-tumoral structures, the addition of *Inception* modules to *U-Net* resulted in statistically significant improvements in WT (DSC improved from 0.903 to 0.925, $p < 0.001$), TC (0.938 to 0.952, $p < 0.001$), and ET (0.937 to 0.948, $p <$0.001). For glioma sub-regions, significant improvements were also noticed in WT (0.898 to 0.918, $p < 0.001$), TC (0.942 to 0.951, $p = 0.001$), and ET (0.942 to 0.948, $p = 0.002$). Changing the objective from intra-tumoral structures to glioma sub-regions learning in the *U-Net* resulted in no difference in performance for WT (0.903 to 0.898, $p =$ 0.307), TC (0.938 to 0.942, $p = 0.284$), and ET (0.937 to 0.942, $p = 0.098$). However, *U-Net* with *Inception* modules, which learned the intra-tumoral structures outperformed those which learned glioma sub-regions in WT (0.918 to 0.925, $p = 0.007$), but there was no difference in the performance for TC (0.952 to 0.951, $p = 0.597$) and ET (0.948 to 0.948, $p = 0.402$). This implies that integrating *Inception* modules in the *U-Net* architecture resulted in statistically significant improvement in tumor segmentation performance that was quantified using k-fold cross-validation ($p < 0.05$ for all three glioma sub-regions). The improvement in the validation accuracy can be attributed to the multiple convolutional filters of different sizes employed in each *Inception* module. These filters are able to capture and retain contextual information at multiple scales during the learning process, both in the contracting as well as expanding paths. We also consider that the improvement in the tumor segmentation accuracy is linked to the new loss function based on the modified DSC (equ. 13). DSC scores for *Inception* modules are comparable or exceed the results of No New-Net [122], which achieved second place in the BraTS 2018 competition, and the ensemble approach proposed in [122, 123, 124].

4.4.2 Results from DIU-Net

DIU-Net showed significant improvement in the WT sub-region with an increase in the Dice score from 0.925 to 0.931 with $p < 0.05$. For the TC sub-region, the Dice score improved from 0.952 to 0.957 with $p < 0.05$. However, for the ET, the change was not statistically significant, $p = 0.114$. Interestingly, DIU-Net is computationally more efficient. DIU-Net has 2.5 million fewer parameters than *U-Net* with *Inception* modules. In contrast,

DIU-Net achieves significantly better results at a lesser computational cost (15% fewer parameters). The *Dice* scores for each glioma sub-region are comparable or exceed the results of other recently published architectures, including *No New-Net*, SDResU-Net and the ensemble approach proposed in [122, 123, 124].

5 Uncertainty Estimation in Brain Tumor Segmentation

As mentioned before, accurate segmentation of brain tumors is crucial for treatment planning and follow-up evaluations. Furthermore, the robustness and trustworthiness of the segmentation results are of particular interest in medical imaging and in the clinic for diagnosis and prognosis due to their link to human health. In this section, we propose a new DL framework, named *extended Variational Density Propagation* (exVDP), that can quantify uncertainty in the output decision [125]. In exVDP, we adopt the *Variational Inference* (VI) [126] framework and propagate the first two moments of the variational distribution through all ANN's layers (convolution, max-pooling and fully-connected) and non-linearities. We use the first-order Taylor series linearization [127] to propagate the mean and covariance of the variational distribution through the non-linear activation functions in the DNNs.

We consider a CNN with a total of C convolutional layers and L fully-connected layers, where the convolutional kernels and the weights of the fully-connected layers are random tensors. A non-linear activation function follows every convolutional and fully-connected layer. Moreover, the ANN contains max-pooling layers. The ANN's weights (and biases) are represented by $\mathbf{\Omega} = \{\{\{\mathbf{\mathcal{W}}^{(k_c)}\}_{k_c=1}^{K_c}\}_{c=1}^{C}, \{\mathbf{W}^{(l)}\}_{l=1}^{L}\}$, where $\{\{\mathbf{\mathcal{W}}^{(k_c)}\}_{k_c=1}^{K_c}\}_{c=1}^{C}$ is the set of K_c kernels in the c^{th} convolutional layer, and $\{\mathbf{W}^{(l)}\}_{l=1}^{L}$ is the set of weights in L fully-connected layers. We consider input tensor $\mathbf{\mathcal{X}} \in \mathbb{R}^{I_1 \times I_2 \times K}$, where I_1, I_2, and K represent image height, width, and number of channels, respectively.

5.1 Variational Learning

We introduce a prior distribution over ANN weights, $\mathbf{\Omega} \sim p(\mathbf{\Omega})$. We assume that convolutional kernels are independent of each other within a layer as well as across different layers. This independence assumption is desirable as it promotes convolutional kernels to extract uncorrelated features within and across layers. Given the training data $\mathcal{D} = \{\mathbf{\mathcal{X}}^{(i)}, \mathbf{y}^{(i)}\}_{i=1}^{N}$ and the prior $p(\mathbf{\Omega})$, the posterior $p(\mathbf{\Omega}|\mathcal{D})$ is given through the Bayes' rule. However, $p(\mathbf{\Omega}|\mathcal{D})$ is typically intractable. VI methods approximate the true posterior

$p(\mathbf{\Omega}|\mathcal{D})$ with a simpler parametrized variational distribution $q_\phi(\mathbf{\Omega})$. The optimal parameters of the variational posterior ϕ^* are estimated by minimizing the *Kullback-Leibler* (KL) divergence between the approximate and the true posterior [128, 126].

$$\begin{aligned}\phi^* &= \operatorname{argmin} \mathrm{KL}\left[q_\phi(\mathbf{\Omega}) \| p(\mathbf{\Omega}|\mathcal{D})\right] \\ &= \operatorname{argmin} \int q_\phi(\mathbf{\Omega}) \log \frac{q_\phi(\mathbf{\Omega})}{p(\mathbf{\Omega}) p(\mathcal{D}|\mathbf{\Omega})} d\mathbf{\Omega} \\ &= \operatorname{argmin} \mathrm{KL}\left[q_\phi(\mathbf{\Omega}) \| p(\mathbf{\Omega}\right] - E_{q_\phi(\mathbf{\Omega})}\left\{\log p(\mathcal{D}|\mathbf{\Omega})\right\}.\end{aligned} \tag{14}$$

The optimization objective is given by the *Evidence Lower BOund* (ELBO) $\mathfrak{L}(\phi; \mathbf{y}|\mathbf{\mathcal{X}})$:

$$\mathfrak{L}(\phi; \mathbf{y}|\mathbf{\mathcal{X}}) = E_{q_\phi(\mathbf{\Omega})}(\log p(\mathbf{y}|\mathbf{\mathcal{X}}, \mathbf{\Omega})) - \mathrm{KL}(q_\phi(\mathbf{\Omega}\|p(\mathbf{\Omega})). \tag{15}$$

ELBO consists of two parts, the expected log-likelihood of the training data given the weights and a regularization term, which can be re-written as:

$$\mathrm{KL}(q_\phi(\mathbf{\Omega}\|p(\mathbf{\Omega}) = \sum_{c=1}^{C} \sum_{k_c=1}^{K_c} \mathrm{KL}(q_\phi(\mathbf{\mathcal{W}}^{(k_c)})\|p(\mathbf{\mathcal{W}}^{(k_c)})) - \sum_{l=1}^{L} \mathrm{KL}(q_\phi(\mathbf{W}^{(l)})\|p(\mathbf{W}^{(l)})). \tag{16}$$

5.2 Variational Density Propagation

We propose to approximate the true unknown posterior $p(\mathbf{\Omega}|\mathcal{D})$ by a variational distribution $q_\phi(\mathbf{\Omega})$. We have defined *Gaussian* distribution as a prior over convolutional kernels and weights of the fully-connected layers [125]. The task is now to propagate the moments of the variational distribution $q_\phi(\mathbf{\Omega})$ through various layers, i.e., convolution, activation, max-pooling, fully-connected, and *softmax*. It is important to note that in our settings, the convolutional kernels, resulting activations, extracted features, logits, and output of the *softmax* function are all random variables. Therefore, instead of performing algebraic operations on real numbers, we are confronted with operations on random variables, including (1) multiplication of a random variable with a constant, (2) multiplication of two random variables, and (3) non-linear transformations [125] operating over random variables. As a result of the multiplication of two Gaussian random variables [127] or non-linear transformation, the resulting random variables may not have *Gaussian distribution* [127]. Our goal is to propagate the mean and covariance of the variational distribution and later obtain the mean and covariance of the predictive distribution, $p(\mathbf{y}|\mathbf{\mathcal{X}}, \mathcal{D})$. The mean of $p(\mathbf{y}|\mathbf{\mathcal{X}}, \mathcal{D})$ represents the ANN's prediction, while the covariance matrix reflects the

uncertainty associated with the output decision. An illustration of the proposed variational density propagation CNN with one convolutional layer, one max-pooling and one fully-connected layer is shown in Fig. (9).

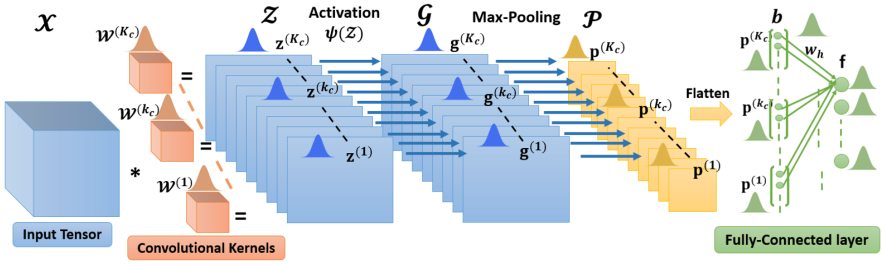

Fig. 9 A schematic layout of the proposed variational density propagation CNN is presented. We show the propagation of the mean and covariance of the variational distribution $q_\phi(\mathbf{\Omega})$ through multiple layers of a CNN.

5.3 Extended Variational Density Propagation

We start with our mathematical results for the propagation of the mean and covariance of the variational distribution $q_\phi(\mathbf{\Omega})$ through convolutional layers, activation functions, max-pooling, fully-connected layers, and the *softmax* function. We use first-order Taylor series for the approximation of the first two moments (mean and covariance) after a non-linear activation function and refer to this method as the exVDP [125].

5.3.1 First convolutional layer

The convolution operation between a set of kernels and the input tensor is formulated as a *matrix-vector multiplication*. We first form sub-tensors $\mathbf{\mathcal{X}}_{i:i+r_1-1,j:j+r_2-1}$ from the input tensor $\mathbf{\mathcal{X}}$, having the same size as the kernels $\mathbf{\mathcal{W}}^{(k_c)} \in \mathbb{R}^{r_1 \times r_2 \times K}$. These sub-tensors are subsequently vectorized and arranged as the rows of a matrix $\tilde{\mathbf{X}}$. Thus, convolving $\mathbf{\mathcal{X}}$ with the k_c^{th} kernel $\mathbf{\mathcal{W}}^{(k_c)}$ is equivalent to the multiplication of $\tilde{\mathbf{X}}$ with $\text{vec}(\mathbf{\mathcal{W}}^{(k_c)})$. Let

$$\mathbf{z}^{(k_c)} = \mathbf{\mathcal{X}} * \mathbf{\mathcal{W}}^{(k_c)} = \tilde{\mathbf{X}} \times \text{vec}(\mathbf{\mathcal{W}}^{(k_c)}), \tag{17}$$

where $*$ denotes the convolution operation and $\times$ is a regular matrix-vector multiplication.

We have defined a multivariate Gaussian distribution over the vectorized convolutional kernels, i.e. $\text{vec}(\mathbf{\mathcal{W}}^{(k_c)}) \sim \mathcal{N}\left(\mathbf{m}^{(k_c)}, \mathbf{\Sigma}^{(k_c)}\right)$. It follows that:

$$\mathbf{z}^{(k_s)} \sim \mathcal{N}\left(\boldsymbol{\mu}_{\mathbf{z}^{(k_c)}} = \tilde{\mathbf{X}}\mathbf{m}^{(k_c)},\ \boldsymbol{\Sigma}_{\mathbf{z}^{(k_c)}} = \tilde{\mathbf{X}}\boldsymbol{\Sigma}^{(k_c)}\tilde{\mathbf{X}}^T\right). \tag{18}$$

5.3.2 Non-linear activation function

We approximate the mean and covariance after the non-linear activation function ψ using the first-order Taylor series approximation [127]). Let $\mathbf{g}_i^{(k_c)} = \psi[\mathbf{z}_i^{(k_c)}]$ be the element-wise i^{th} output of ψ. We have $\boldsymbol{\mu}_{\mathbf{g}^{(k_c)}}$ and $\boldsymbol{\Sigma}_{\mathbf{g}^{(k_c)}}$:

$$\begin{aligned}
&\mu_{\mathbf{g}_i^{(k_c)}} \approx \psi(\mu_{\mathbf{z}_i^{(k_c)}}),\\
&\boldsymbol{\Sigma}_{\mathbf{g}^{(k_c)}} \approx \begin{cases} \sigma^2_{\mathbf{z}_i^{(k_c)}} \left(\frac{d\psi(\mu_{\mathbf{z}_i^{(k_c)}})}{d\mathbf{z}_i^{(k_c)}}\right)^2, & \text{if } i = j.\\ \sigma_{\mathbf{z}_i^{(k_c)}\mathbf{z}_j^{(k_c)}} \left(\frac{d\psi(\mu_{\mathbf{z}_i^{(k_c)}})}{d\mathbf{z}_i^{(k_c)}}\right)\left(\frac{d\psi(\mu_{\mathbf{z}_j^{(k_c)}})}{d\mathbf{z}_j^{(k_c)}}\right), & \text{if } i \neq j. \end{cases}
\end{aligned} \tag{19}$$

5.3.3 Max-pooling layer

For the max-pooling, $\boldsymbol{\mu}_{\mathbf{p}^{(k_c)}} = \text{pool}(\boldsymbol{\mu}_{\mathbf{g}^{(k_c)}})$ and $\boldsymbol{\Sigma}_{\mathbf{p}^{(k_c)}} = \text{co-pool}(\boldsymbol{\Sigma}_{\mathbf{g}^{(k_c)}})$, where pool represents the max-pooling operation on the mean and co-pool represents down-sampling the covariance, i.e., we keep only the rows and columns of $\boldsymbol{\Sigma}_{\mathbf{g}^{(k_c)}}$ corresponding to the pooled mean elements.

5.3.4 Flattening operation

The output tensor $\boldsymbol{\mathcal{P}}$ of the max-pooling layer is vectorized to form the input vector $\mathbf{b}$ of the fully-connected layer such that, $\mathbf{b} = \left[\mathbf{p}^{(1)T}, \cdots, \mathbf{p}^{(K_c)T}\right]^T$. The mean and covariance matrix of $\mathbf{b}$ are given by:

$$\boldsymbol{\mu}_{\mathbf{b}} = \begin{bmatrix} \boldsymbol{\mu}_{\mathbf{p}^{(1)}} \\ \vdots \\ \boldsymbol{\mu}_{\mathbf{p}^{(K_c)}} \end{bmatrix}, \boldsymbol{\Sigma}_{\mathbf{b}} = \begin{bmatrix} \boldsymbol{\Sigma}_{\mathbf{p}^{(1)}} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \boldsymbol{\Sigma}_{\mathbf{p}^{(K_c)}} \end{bmatrix}. \tag{20}$$

5.3.5 Fully-connected layer

Let $\mathbf{w}_h \sim \mathcal{N}(\mathbf{m}_h, \boldsymbol{\Sigma}_h)$ be h^{th} weight vector of the fully-connected layer, where $h = 1, \cdots, H$, and H is the number of output neurons. We note that f_h is the product of two independent random vectors $\mathbf{b}$ and $\mathbf{w}_h$. Let $\mathbf{f}$ be

the output vector of the fully-connected layer, then the elements of $\boldsymbol{\mu}_{\mathbf{f}}$ and $\boldsymbol{\Sigma}_{\mathbf{f}}$ are derived by the following proposition:

Proposition 1.

$$\begin{aligned} \mu_{f_h} &= \mathbf{m}_h^T \boldsymbol{\mu}_{\mathbf{b}}, \\ \boldsymbol{\Sigma}_{\mathbf{f}} &= \begin{cases} \operatorname{tr}\left(\boldsymbol{\Sigma}_h \boldsymbol{\Sigma}_{\mathbf{b}}\right) + \mathbf{m}_h^T \boldsymbol{\Sigma}_{\mathbf{b}} \mathbf{m}_h + \boldsymbol{\mu}_{\mathbf{b}}^T \boldsymbol{\Sigma}_h \boldsymbol{\mu}_{\mathbf{b}}, & \\ \mathbf{m}_{h_1}^T \boldsymbol{\Sigma}_{\mathbf{b}} \mathbf{m}_{h_2}, & h_1 \neq h_2, \end{cases} \end{aligned} \tag{21}$$

where $h_1, h_2 = 1, \cdots, H$ represent any two weight vectors in the fully-connected layer.

5.3.6 *Softmax* function

Let the output of the ANN be $\mathbf{y} = \varphi(\mathbf{f})$, where φ is the *softmax* function. Using the first-order Taylor series approximation, the mean and covariance of the output vector, i.e., $\boldsymbol{\mu}_{\mathbf{y}}$ and $\boldsymbol{\Sigma}_{\mathbf{y}}$, are derived as follows [129]:

$$\boldsymbol{\mu}_{\mathbf{y}} \approx \varphi(\boldsymbol{\mu}_{\mathbf{f}}); \;\; \boldsymbol{\Sigma}_{\mathbf{y}} \approx \mathbf{J}_{\varphi} \boldsymbol{\Sigma}_{\mathbf{f}} \mathbf{J}_{\varphi}^T, \tag{22}$$

where $\mathbf{J}_{\varphi}$ is the *Jacobian* matrix of φ with respect to $\mathbf{f}$ evaluated at $\boldsymbol{\mu}_{\mathbf{f}}$ [129].

5.3.7 Objective function

Assuming a diagonal covariance matrix for the variational posterior distribution, N independently and identically distributed (iid) data points and using M Monte Carlo samples to approximate the expectation by a summation, the expected log-likelihood in the ELBO objective function is given as follows:

$$\begin{aligned} &E_{q_{\phi}(\boldsymbol{\Omega})}(\log p(\mathbf{y}|\boldsymbol{\mathcal{X}}, \boldsymbol{\Omega})) \approx \\ &-\frac{NH}{2}\log(2\pi) - \frac{1}{M}\sum_{m=1}^{M}\Big[\frac{N}{2}\log(|\boldsymbol{\Sigma}_{\mathbf{y}}|) + \frac{1}{2}\sum_{i=1}^{N}(\mathbf{y}^{(i)} - \boldsymbol{\mu}_{\mathbf{y}}^{(m)})^T(\boldsymbol{\Sigma}_{\mathbf{y}}^{(m)})^{-1}(\mathbf{y}^{(i)} - \boldsymbol{\mu}_{\mathbf{y}}^{(m)})\Big] \end{aligned} \tag{23}$$

The regularization term in (16) is the KL-divergence between two multivariate Gaussian distributions [127]. If we have a CNN with one convolutional layer followed by the activation function, one max-pooling and one fully-connected layer, thus the regularization term in the ELBO objective function is derived as follows:

$$\begin{aligned}
&\mathrm{KL}(q_\phi(\mathbf{\Omega}\|p(\mathbf{\Omega}) = \\
&\frac{1}{2}\sum_{k=1}^{K_1}\left(r_1 r_2 K\ \sigma^2_{r_1,k}\ \sigma^2_{r_2,k}\ \sigma^2_{K,k} + \|\mathbf{\mathcal{M}}^{(k)}\|_F^2 - r_1 r_2 K - r_1 r_2 K\ \left(\log(\sigma^2_{r_1,k}\ \sigma^2_{r_2,k}\ \sigma^2_{K,k})\right)\right) \\
&+ \frac{1}{2}\sum_{h=1}^{H}\Big(n_f\ \sigma_h^2 + \|\mathbf{m}_h\|_F^2 - n_f - n_f\ \log\sigma_h^2\Big),
\end{aligned} \tag{24}$$

where $(r_1 \times r_2 \times K)$ is the size of the kernels, K_1 is the number of kernels in the convolutional layer, H is the number of output neurons and n_f is the length of the weight vector $\mathbf{w}_h$ in the fully-connected layer.

5.3.8 Back-propagation

During back-propagation, we compute the gradient of the objective function $\nabla_\phi \mathfrak{L}(\phi; \mathcal{D})$ with respect to the variational parameters:

$$\phi = \left\{\left\{\{\mathbf{\mathcal{M}}^{(k_c)}, \sigma^2_{r_1,k_c}, \sigma^2_{r_2,k_c}, \sigma^2_{K_{c-1},k_c}\}_{k_c=1}^{K_c}\right\}_{c=1}^{C}, \{\mathbf{m}_h, \sigma_h^2\}_{h=1}^{H}\right\}, \tag{25}$$

where $(r_1 \times r_2 \times K_{c-1})$ is the size of the k_c^{th} kernel, K_c is the number of kernels in the c^{th} convolutional layer and H is the number of output neurons. We use $\nabla_\phi \mathfrak{L}(\phi; \mathcal{D})$ to update our parameters ϕ using the gradient descent update rule.

5.4 Application to Brain Tumor Segmentation in MRI Images

The performance of the proposed exVDP model on the HGG brain tumor segmentation task using the BraTS 2015 dataset has been evaluated. The dataset consists of 5 classes, i.e. class 0 - normal tissue, class 1 - necrosis, class 2 - edema, class 3 - non-enhancing, and class 4 - enhancing tumor [94]. The evaluation of segmentation is based on three regions, (1) complete tumor (1, 2, 3 and 4), (2) tumor core (1, 3 and 4), and (3) enhancing tumor (class 4) [94].

Brain tumor segmentation has been formulated as a multi-class classification problem by randomly sampling patches from four MRI modalities, i.e., FLAIR, T1, T2 and T1c [130]. The label of each patch has been manually set to the label of the center pixel. The sampled patches are balanced over all classes, and a total of $100,000$ patches of size 33×33 are extracted from the BraTS data of 20 patients. These patches are divided into

training and validation bins (95% for training and 5% for validation). The test set included randomly sampled 372 images, i.e., 43, 264 patches, from each of the four modalities. The proposed exVDP model is compared with a deterministic CNN, presented in [130]. The following CNN architecture has been used: six convolution layers (all kernels were 3×3, and we had 32, 32, 64, 64, 128, 128 kernels in layers one to six, respectively, followed by ReLU activation), two max-pooling layers, and a fully-connected layer. The architecture is shown in Table (1).

DSC has been used to evaluate the segmentation results before and after adding Gaussian noise or targeted adversarial attack (targeted class is class 3, i.e., *non-enhancing tumor*). The evaluation of the proposed model on the BraTS dataset is done without doing any pre-processing or data augmentation techniques.

In Table (2), DSC values for three test cases have been presented, i.e., noise-free, Gaussian, and adversarial noise. We note that the DSC values of the proposed model are significantly higher than that of the deterministic CNN for all cases in general and adversarial noise in particular. Fig. (10) shows segmentation results for exVDP and a deterministic CNN for a representative HGG image (with and without adversarial noise). The uncertainty map associated with each segmentation is also presented for the exVDP model. The uncertainty map allows physicians to review the segmentation results quickly and, if needed, make corrections of tumor boundaries in the regions where the uncertainty is high.

Table 1 *Architecture of the two models, i.e., exVDP and deterministic CNN*

Layer	Type	Filter size	HGG stride	No. kernels	FC units	Input
1	Conv.	3×3	1×1	32	-	33×33×4
2	Conv.	3×3	1×1	32	-	33×33×32
3	Conv.	3×3	1×1	64	-	33×33×32
4	Max-pool.	3×3	2×2	-	-	33×33×64
5	Conv.	3×3	1×1	64	-	16×16×64
6	Conv.	3×3	1×1	128	-	16×16×64
7	Conv.	3×3	1×1	128	-	16×16×128
8	Max-pool.	3×3	2×2	-	-	16×16×128
9	FC	-	-	-	5	6272

Table 2 *Segmentation results measured using the DSC for the BraTS test dataset*

Method	Tumor Regions	No noise	Adversarial noise	Gaussian noise
exVDP	Complete	**80.8**%	**77.4**%	**80.6**%
	Core	74.6%	72.6%	74.5%
	Enhancing	74.0%	69.8%	73.9%
Deterministic CNN	Complete	78.0%	43.4%	66.9%
	Core	65.0%	47.1%	51.9%
	Enhancing	75.0%	43.9%	55.7%

6 Tumor Surveillance

As defined in section 2.3, tumor surveillance is the process of monitoring patient's tumor in longitudinal studies to establish severity of the disease and planning treatment accordingly. It helps identifying early signs of tumor occurrence which is critical especially in case of the cancerous tumors.

6.1 Rationale for Tumor Surveillance

Temporal medical imaging data is widely used in oncology as well as radiology for visual comparison of disease over an extended period of time. The 2D medical images (CT or MRI scans) are examined by the physicians to diagnose the disease in 4D (3D tumor volume over time) usually referred to as *change in volume over time*. A detection at an earlier stage of disease is more responsive to treatment, resulting in improved outcomes for the patient. Biological characteristics of various tumor types such as growth, location, and patterns of local as well as metastatic disease are the basis for surveillance scheduling, protocols, and selection of imaging techniques. Usually low-risk tumor is subjected to active surveillance as part of a patient's treatment plan and an ideal candidate for active surveillance is the one which has one or more of these conditions: (1) disease has not spread, (2) tumor is small and growing slowly, or (3) patient exhibits no symptoms of specific cancer. The data from active surveillance is also used to look for trends and patterns over time in certain regions/ groups of people, and to see if preventive measures are making a difference among the sample population. Apart from brain cancer, active surveillance has been widely employed in other forms of cancer diagnosis. It has been shown that use of *Prostate-Specific Antigen* (PSA) [131] testing as part of active surveillance of prostate cancer helps in understanding tumor progression and prognosis, enabling the patients diagnosed with lower grade disease feel more comfortable [24].

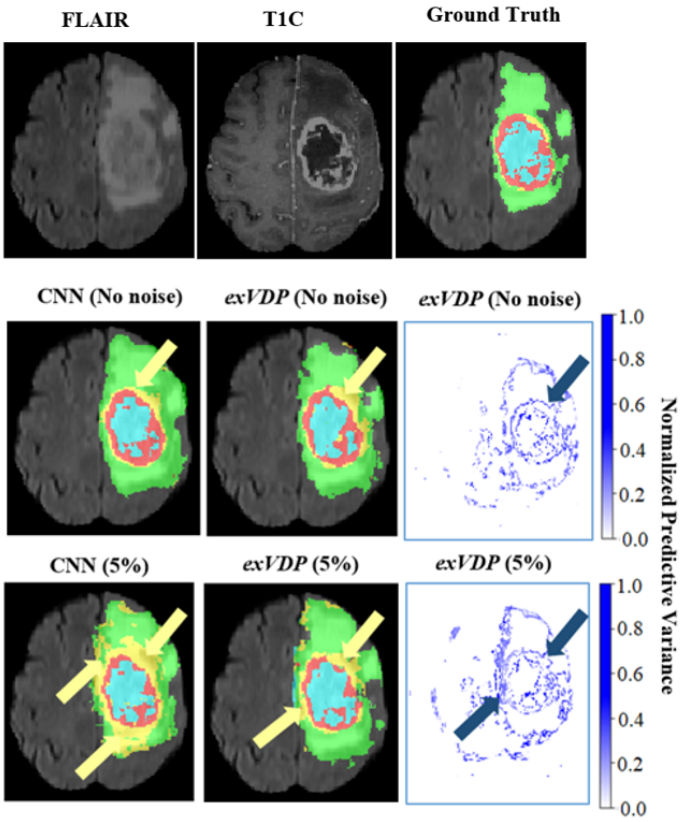

Fig. 10 Example 1: Segmentation results of the proposed exVDP and a deterministic CNN on the BraTS 2015 dataset with and without adding adversarial noise. The uncertainty map associated with each segmentation is also shown for each of the two models. The class label *non-enhancing tumor*, which is the target of the attack, is represented in yellow color in the ground truth image. The green color refers to the edema class, the red color refers to the enhancing tumor, and the blue color refers to the necrosis.

6.2 Surveillance Techniques

Change-point detection is the classical technique of detecting abrupt changes in sequential data, which focuses predominantly on datasets with a single observable. It has been a long-standing research area in statistics [132, 133], with applications in fields ranging from economics [134], bioinformatics [135, 136], and climatology [137], wherein it is dealt as the problem of detecting abrupt changes in temporal data. The objective is to determine if the observed time series is statistically homogeneous or otherwise to find the point in time when the change happens. There are two variations to

the change-point detection technique: posterior and sequential. Posterior tests are done offline after entire data is collected, and a decision of homogeneity or change-point is made based on the analysis of all the collected data. Whereas, sequential tests are done on-the-fly as the data is presented sequentially, and the decisions are made online. *Gleason grade* is another technique used for pathological scoring of the differentiation of prostate cancer, and it has been the most widely used grading system for prostate tumor differentiation and prognostic indicator for prostate cancer progression [138].

6.3 Community-level Active Surveillance

Apart from the individual tumor surveillance, the term of active surveillance is also often used for collective cancer surveillance data and programs in the United States through *Cancer Registries*. A cancer registry is an information system designed for the collection, storage, and management of data on persons with cancer [139]. Data on cancer in the United States is collected through two types of registries: hospital registries, which are the part of a facility's cancer program, and population-based registries, usually tied to state health departments. Hospital registries provide patient's data on care within the hospital for evaluation. Population-based registries, under state health departments, collect information on all cases diagnosed within a certain geographic area from multiple reporting facilities, including hospitals, doctors' offices, nursing homes, pathology laboratories, radiation and chemotherapy treatment centers, etc. The collected data is used to build statistics like new cancer cases (incidence), death rates (mortality), cancer types related to types of jobs, cancer trends over time to keep an eye on age and racial groups that are most affected by different types of cancer. Registries are staffed under the *Certified Tumor Registrar* (CTR), having pre-defined standards of training, testing, and continuing education, and they compile timely, accurate, and complete cancer information to report to the registry. The major cancer surveillance programs in the United States are the *National Cancer Data Base* (NCDB) [140], *National Cancer Institute*'s (NCI) *Surveillance, Epidemiology and End Results* (SEER) program [139], and *National Program of Cancer Registries* (NPCR) of the *Center for Disease Control and Prevention* (CDC) [141]. *Central Brain Tumor Registry of the United States* (CBTRUS) is a registry dedicated to collecting and disseminating statistical data on all primary benign and malignant brain tumors [142]. A recent study [143] tries to estimate excess mortality in people with cancer and multi-morbidity in the COVID-19 affected patients through analysis of surveillance data DATA-CAN, the UK National Health Data Research Hub for cancer emergency [144].

6.4 Surveillance of Brain Tumor

Tumors of the *Central Nervous System* (CNS) are the second most common tumors among children after leukemia. Treatment protocols for high-grade pediatric brain tumors recommend regular follow-up imaging for up to 10 years. Based on the surveillance data of high-grade childhood brain tumor patients, a review of maximal time to recurrence and minimal time to radiologically detectable long-term sequel such as secondary malignancies, vascular complications, and white matter disease found that there was no recurrence of the primary brain tumor, either local or distant, 10 years or more after the end of treatment in the reviewed literature and so the results do not justify routine screening to detect tumor recurrence more than 10 years after the end of treatment [145].

Tumor surveillance is being used for building statistical figures in a broad spectrum of ways, including adult glioma incidence and survival by race or ethnicity in the United States [146], county-level glioma incidence and survival variations [147], and accurate population-based statistics on the brain and other central nervous system tumors [148]. A CBTRUS statistical report on the primary brain and *Other Central Nervous System* (OCNS) tumors data diagnosed in the United States in the period 2011-2015 states that brain and OCNS tumors (both malignant and non-malignant) were the most common cancer types in persons age 0–14 years for both males and females. For age 15–39 years, these tumors were the second most common cancer in males and the third most common among females in this age group. For age 40+, these were the eighth most common cancer type, with males having eighth and females having the fifth most common brain cancer. These results were based on the NPCR data of 388,786 brain and OCNS tumors, and 16,633 tumor cases from SEER [149].

Tumor Surveillance among patients also enables the authorities to predict cancer cases and death in advance and respond in time to offset the predicted scores. A recent report by the *American Cancer Society* on cancer statistics in 2020 projected the number of new cancer cases and deaths that will occur in the United States. Incidence data from 2002 to 2017 were collected, and it was estimated that in 2020, 23,890 new cases and 18,020 deaths related to brain and ONS tumors were projected to occur in the United States [150]. Similarly, tumor surveillance is equally important to avoid the side effects of the aggressive forms of treatment. Patients treated for glioma, meningioma, and brain metastases may develop side effects of treatment, including neuropathy (with visual loss), cataracts, hypopituitarism, cognitive decline, increased risk of stroke, and risk of secondary tumor occurring months or even years later. The surgical treatment causes immediate side effects, chemotherapy-caused side effects occur early after treatment (but infertility may not manifest itself until later), and radiotherapy's side effects occur months or even years after treatment. The risks vary depending on the technique used and the area of the brain treated.

Surveillance enables the physicians to identify these potential late side-effects earlier which increases the length and quality of life for patients [151].

6.4.1 An example of Surveillance Study

In this section, we will study an example of tumor surveillance, specifically of patients having low-grade gliomas [152]. Low-grade gliomas, constituting around 15% of all adult brain tumors, significantly affect neurological morbidity by brain invasion. Generally, there is no universally-accepted technique available for the detection of growth of low-grade gliomas in the clinical setting. Clinicians usually consider visual comparisons of two or more longitudinal radiological scans through subjective evaluation for detecting the growth of low-grade gliomas. The paper [152] suggests a *Computer-Assisted Diagnosis* (CAD) method to help physicians detect earlier growth of low-grade gliomas. This method consists of tumor segmentation, computing volumes, and pointing to growth by the online abrupt change-of-point method considering only past measurements. The study suggests that early growth detection of tumor sets the stage for future clinical studies to decide upon the type of treatment-path to be undertaken and whether early therapeutic interventions prolong survival and improve quality of life. Longitudinal (temporal) radiological studies of 63 patients were carried out with a median follow-up period of 33.6 months. These patients were diagnosed with grade 2 gliomas by expert physicians through manual (visual) procedures as well as detection of growth with that of the CAD method, and both detection methods were compared by 7 expert physicians [152] . Each patient had at least 4 MRI scans available for review either after the initial diagnosis or after the completion of chemotherapy with temozolomide (if applicable). The researchers calculated the time to growth detection from the impressions of the radiological reports of these patients from 627 MRI scans. Unexpectedly, the study found large differences in growth detection by visual comparison and by physicians aided by the CAD method. The reasons for missing growth by the visual inspection can be attributed to one or more of these reasons: (1) interpreting a large number of prior studies by physicians takes a very long time, (2) the practice in vogue of comparing the current MRI to a couple of MRI scans immediately preceding it, (3) the lack of determination of baseline MRI, (4) small changes from one scan to the next, and (5) comparing single 2-D images overlooks the growth in the third dimension.

The study [152] showed that the CAD method helped physicians detect growth at earlier times and significantly smaller tumor volumes than the manual standard method. Moreover, physicians aided by the CAD method diagnosed tumor growth in 13 of 22 glioma patients labeled as clinically stable by the standard method. Fig. 11 shows the volume growth curves of

grade 2 gliomas of two patients diagnosed with oligos, seen at the University of Alabama at Birmingham clinics between 1 July 2017 and 14 May 2018 [152]. The x-axis corresponds to the time interval from the baseline MRI, and the y-axis corresponds to the change in the volume of tumors from the baseline. The volume at each time step until the growth detected by CAD is colored in yellow, and the manual (visual) detection of change-point time is colored in red. The CAD for patient 1 detected a change-point in 20 months from the baseline, whereas visual detection by a physician was done in 80 months.

Similarly, CAD detection time for patient 2 was also around 20 months, where visual detection was in 150 months, primarily because this tumor did not grow at a faster pace. The detection of tumor volume growth in time enabled the researchers to identify tumors with nonlinear and non-homogeneous growth. Early growth detection holds the potential of lowering the morbidity, and perhaps mortality of patients with low-grade gliomas. The decision to treat a patient would be determined by the rate of growth and proximity to critical areas of the brain, once they have been measured. The study also suggested early interventions for cases where (1) the new growth is in the proximity of key nonsurgical structures like the corpus callosum, (2) the rate of growth is elevated, or (3) the tumor is sensitive to chemotherapy.

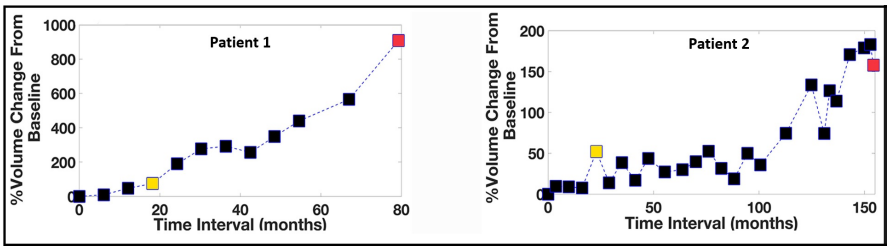

Fig. 11 Volume growth curves of grade 2 gliomas of two patients diagnosed with oligos. x-axis corresponds to the time interval from the baseline MRI, y-axis corresponds to change in volume of tumor from baseline. The volume at the time-to-growth detected by computer-assisted diagnosis is colored in yellow and manual (visual) detection time is colored in red.

7 Conclusion

In this chapter, we have thoroughly reviewed the image segmentation task in the classical CV field and examined various techniques of CV employed in DL frameworks for brain tumor segmentation. We have also assessed multiple DL architectures having varying attributes that make them suit-to-task employment. We have looked into a case study for in-depth analysis of *U-Net* with *Inception* and *dilated Inception* modules in the context of brain tumor segmentation. A new DL framework, called exVDP, that can quantify uncertainty in the output decision of an ANN, has also been discussed. In the last section, we have discussed the concept of tumor surveillance, its rationale, techniques, and an example study on low-grade gliomas surveillance.

The brain tumor segmentation community has achieved substantial progress in the last decade because of the advances in DL. Although efforts have been made in commercializing the technology for clinicians, there is still a long way to make brain tumor segmentation a reliable and routine tool broadly applied to practical clinical decisions with minimal human interventions. This is due to the lack of existing methods in the face of adversarial examples and research-oriented frameworks that are not suited to production environments. Breakthrough is likely to come with the advent of effective and scalable platforms by the ML community, and direction of research towards adversarial learning.

Acknowledgement

This work was supported by the National Science Foundation Awards NSF ECCS-1903466 and NSF DUE-1610911.

References

1. Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," *Nature*, vol. 521, no. 7553, p. 436, 2015.
2. A. A. Aly, S. B. Deris, and N. Zaki, "Research review for digital image segmentation techniques," *International Journal of Computer Science & Information Technology*, vol. 3, no. 5, p. 99, 2011.
3. Stanford, "Tutorial 3: Image segmentation," https://ai.stanford.edu/~syyeung/cvweb/tutorial3.html.
4. Marvel, "Marvel movies," https://www.marvel.com/movies.
5. S. Gould, T. Gao, and D. Koller, "Region-based segmentation and object detection," in *Advances in neural information processing systems*, 2009, pp. 655–663.
6. S. Yuheng and Y. Hao, "Image segmentation algorithms overview," *arXiv preprint arXiv:1707.02051*, 2017.

7. X. Xu, G. Li, G. Xie, J. Ren, and X. Xie, "Weakly supervised deep semantic segmentation using cnn and elm with semantic candidate regions," *Complexity*, vol. 2019, 2019.
8. mayoclinic, "Braintumor," https://www.mayoclinic.org/diseases-conditions/brain-tumor/symptoms-causes/syc-20350084.
9. movies.effects, *Instagram*, 2019 (accessed August 27, 2020). [Online]. Available: https://www.instagram.com/p/BuOlFbHhjr7/
10. P. Medicine, "Mipg," https://www.pennmedicine.org/departments-and-centers/department-of-radiology/radiology-research/labs-and-centers/biomedical-imaging-informatics/medical-image-processing-group.
11. A. A. of Neurological Surgeons, "Braintumortypes," https://www.aans.org/en/Patients/Neurosurgical-Conditions-and-Treatments/Brain-Tumors.
12. PennMedicine, "Common types of brain tumors," https://www.pennmedicine.org/updates/blogs/neuroscience-blog/2018/november/what-are-the-most-common-types-of-brain-tumors.
13. S. Bauer, R. Wiest, L.-P. Nolte, and M. Reyes, "A survey of MRI-based medical image analysis for brain tumor studies," *Physics in Medicine & Biology*, vol. 58, no. 13, pp. R97–R129, 2013.
14. A. Bousselham, O. Bouattane, M. Youssfi, and A. Raihani, "Towards reinforced brain tumor segmentation on mri images based on temperature changes on pathologic area," *International journal of biomedical imaging*, vol. 2019, 2019.
15. MayoFoundation, "Biomedical imaging resource," https://analyzedirect.com/.
16. NIH, "Neuroimaging informatics technology initiative," https://nifti.nimh.nih.gov/ and https://brainder.org/2012/09/23/the-nifti-file-format/.
17. R. D. Vincent, P. Neelin, N. Khalili-Mahani, A. L. Janke, V. S. Fonov, S. M. Robbins, L. Baghdadi, J. Lerch, J. G. Sled, R. Adalat, D. MacDonald, A. P. Zijdenbos, D. L. Collins, and A. C. Evans, "Minc 2.0: A flexible format for multi-modal images," *Frontiers in Neuroinformatics*, vol. 10, p. 35, 2016. [Online]. Available: https://www.frontiersin.org/article/10.3389/fninf.2016.00035
18. P. Scanners, "parrec," https://nipy.org/nibabel/reference/nibabel.parrec.html.
19. NRRD, "Nearly raw raster data," http://teem.sourceforge.net/nrrd/.
20. DICOM, "Digital imaging and communications in medicine," https://www.dicomstandard.org/about-home.
21. M. Larobina and L. Murino, "Medical image file formats," *Journal of digital imaging*, vol. 27, no. 2, pp. 200–206, 2014.
22. N. C. Institute, "surveillance," https://www.cancer.gov/publications/dictionaries/cancer-terms/def/surveillance.
23. RCCA, "Active surveillance: Its role in low-risk cancer," https://www.regionalcancercare.org/services/active-surveillance/.
24. K. L. Penney, M. J. Stampfer, J. L. Jahn, J. A. Sinnott, R. Flavin, J. R. Rider, S. Finn, E. Giovannucci, H. D. Sesso, M. Loda *et al.*, "Gleason grade progression is uncommon," *Cancer research*, vol. 73, no. 16, pp. 5163–5168, 2013.
25. R. Szeliski, *Computer vision: algorithms and applications*. Springer Science & Business Media, 2010.
26. V. Dumoulin and F. Visin, "A guide to convolution arithmetic for deep learning," *arXiv preprint arXiv:1603.07285*, 2016.
27. ——, "Convolution arithmetic," https://www.github.com/vdumoulin/conv_arithmetic.
28. C. Sinai, "Brain tumors and brain cancer," https://www.cedars-sinai.org/health-library/diseases-and-conditions/b/brain-tumors-and-brain-cancer.html.
29. R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2019," *CA: a cancer journal for clinicians*, vol. 69, no. 1, pp. 7–34, 2019.

30. E. C. Holland, "Progenitor cells and glioma formation," *Current opinion in neurology*, vol. 14, no. 6, pp. 683–688, 2001.
31. J. C. Buckner, "Factors influencing survival in high-grade gliomas," in *Seminars in oncology*, vol. 30. Elsevier, 2003, pp. 10–14.
32. J. Lemke, J. Scheele, T. Kapapa, S. Von Karstedt, C. R. Wirtz, D. Henne-Bruns, and M. Kornmann, "Brain metastases in gastrointestinal cancers: is there a role for surgery?" *International journal of molecular sciences*, vol. 15, no. 9, pp. 16 816–16 830, 2014.
33. R. C. Miner, "Image-guided neurosurgery," *Journal of medical imaging and radiation sciences*, vol. 48, no. 4, pp. 328–335, 2017.
34. M. E. Karar, D. R. Merk, V. Falk, and O. Burgert, "A simple and accurate method for computer-aided transapical aortic valve replacement," *Computerized Medical Imaging and Graphics*, vol. 50, pp. 31–41, 2016.
35. N. Sharma and L. M. Aggarwal, "Automated medical image segmentation techniques," *Journal of medical physics/Association of Medical Physicists of India*, vol. 35, no. 1, p. 3, 2010.
36. L. Kaufman and P. J. Rousseeuw, *Finding groups in data: an introduction to cluster analysis*. John Wiley & Sons, 2009, vol. 344.
37. A. K. Jain, A. Topchy, M. H. Law, and J. M. Buhmann, "Landscape of clustering algorithms," in *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, vol. 1. IEEE, 2004, pp. 260–263.
38. A. K. Jain, R. P. Duin, and J. Mao, "Statistical pattern recognition: A review, ieee transactions on pattern analysis and machine intelligence," *vol*, vol. 22, 2000.
39. A. K. Jain and R. C. Dubes, *Algorithms for clustering data*. Prentice-Hall, Inc., 1988.
40. C. R. Brice and C. L. Fennema, "Scene analysis using regions," *Artificial intelligence*, vol. 1, no. 3-4, pp. 205–226, 1970.
41. E. M. Riseman and M. A. Arbib, "Computational techniques in the visual segmentation of static scenes," *Computer Graphics and Image Processing*, vol. 6, no. 3, pp. 221–276, 1977.
42. T. Pavlidis, *Structural pattern recognition*. Springer, 1977.
43. R. Ohlander, K. Price, and D. R. Reddy, "Picture segmentation using a recursive region splitting method," *Computer graphics and image processing*, vol. 8, no. 3, pp. 313–333, 1978.
44. A. Rosenfeld and L. S. Davis, "Image segmentation and image models," *Proceedings of the IEEE*, vol. 67, no. 5, pp. 764–772, 1979.
45. R. M. Haralick and L. G. Shapiro, "Image segmentation techniques," *Computer vision, graphics, and image processing*, vol. 29, no. 1, pp. 100–132, 1985.
46. A. Blake and M. Isard, *Active contours: the application of techniques from graphics, vision, control theory and statistics to visual tracking of shapes in motion*. Springer Science & Business Media, 2012.
47. D. Cremers, M. Rousson, and R. Deriche, "A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape," *International journal of computer vision*, vol. 72, no. 2, pp. 195–215, 2007.
48. P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *International journal of computer vision*, vol. 59, no. 2, pp. 167–181, 2004.
49. D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 5, pp. 603–619, 2002.
50. J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.

51. S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. S. Kirby, J. B. Freymann, K. Farahani, and C. Davatzikos, "Advancing the cancer genome atlas glioma mri collections with expert segmentation labels and radiomic features," *Scientific data*, vol. 4, p. 170117, 2017.
52. J. Malik, S. Belongie, T. Leung, and J. Shi, "Contour and texture analysis for image segmentation," *International journal of computer vision*, vol. 43, no. 1, pp. 7–27, 2001.
53. X. Liu, Z. Deng, and Y. Yang, "Recent progress in semantic image segmentation," *Artificial Intelligence Review*, vol. 52, no. 2, pp. 1089–1106, 2019.
54. K. Fukushima and S. Miyake, "Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition," in *Competition and cooperation in neural nets*. Springer, 1982, pp. 267–285.
55. Y. LeCun, Y. Bengio *et al.*, "Convolutional networks for images, speech, and time series," *The handbook of brain theory and neural networks*, vol. 3361, no. 10, p. 1995, 1995.
56. J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, G. Wang, J. Cai *et al.*, "Recent advances in convolutional neural networks," *Pattern Recognition*, vol. 77, pp. 354–377, 2018.
57. M. Nielsen, "Ch-4: A visual proof that neural nets can compute any function," http://neuralnetworksanddeeplearning.com/chap4.html.
58. Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
59. S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
60. I. Goodfellow, Y. Bengio, and A. Courville, "Deep learning book," *MIT Press*, vol. 521, no. 7553, p. 800, 2016.
61. I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in neural information processing systems*, 2014, pp. 2672–2680.
62. A. Waibel, T. Hanazawa, G. Hinton, K. Shikano, and K. J. Lang, "Phoneme recognition using time-delay neural networks," *IEEE transactions on acoustics, speech, and signal processing*, vol. 37, no. 3, pp. 328–339, 1989.
63. S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image segmentation using deep learning: A survey," *arXiv preprint arXiv:2001.05566*, 2020.
64. V. Gulshan, L. Peng, M. Coram, M. C. Stumpe, D. Wu, A. Narayanaswamy, S. Venugopalan, K. Widner, T. Madams, J. Cuadros *et al.*, "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *Jama*, vol. 316, no. 22, pp. 2402–2410, 2016.
65. A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *nature*, vol. 542, no. 7639, pp. 115–118, 2017.
66. Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3d u-net: learning dense volumetric segmentation from sparse annotation," in *International conference on medical image computing and computer-assisted intervention*. Springer, 2016, pp. 424–432.
67. F. Isensee, P. Kickingereder, W. Wick, M. Bendszus, and K. H. Maier-Hein, "Brain tumor segmentation and radiomics survival prediction: Contribution to the brats 2017 challenge," in *International MICCAI Brainlesion Workshop*. Springer, 2017, pp. 287–297.
68. G. Wang, W. Li, S. Ourselin, and T. Vercauteren, "Automatic brain tumor segmentation using cascaded anisotropic convolutional neural networks," in *International MICCAI brainlesion workshop*. Springer, 2017, pp. 178–190.

69. L. Sun, S. Zhang, and L. Luo, "Tumor segmentation and survival prediction in glioma with deep learning," in *International MICCAI Brainlesion Workshop*. Springer, 2018, pp. 83–93.
70. D. E. Cahall, G. Rasool, N. C. Bouaynaya, and H. M. Fathallah-Shaykh, "Inception modules enhance brain tumor segmentation," *Frontiers in computational neuroscience*, vol. 13, p. 44, 2019.
71. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.
72. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
73. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
74. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1–9.
75. A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017.
76. G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.
77. Y. Li, H. Qi, J. Dai, X. Ji, and Y. Wei, "Fully convolutional instance-aware semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 2359–2367.
78. Y. Yuan, M. Chao, and Y.-C. Lo, "Automatic skin lesion segmentation using deep fully convolutional networks with jaccard distance," *IEEE transactions on medical imaging*, vol. 36, no. 9, pp. 1876–1886, 2017.
79. N. Liu, H. Li, M. Zhang, J. Liu, Z. Sun, and T. Tan, "Accurate iris segmentation in non-cooperative environments using fully convolutional networks," in *2016 International Conference on Biometrics (ICB)*. IEEE, 2016, pp. 1–8.
80. O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.
81. F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," in *2016 fourth international conference on 3D vision (3DV)*. IEEE, 2016, pp. 565–571.
82. T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2117–2125.
83. H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2881–2890.
84. L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs," *IEEE transactions on pattern analysis and machine intelligence*, vol. 40, no. 4, pp. 834–848, 2017.
85. F. Visin, K. Kastner, K. Cho, M. Matteucci, A. Courville, and Y. Bengio, "Renet: A recurrent neural network based alternative to convolutional networks," *arXiv preprint arXiv:1505.00393*, 2015.
86. F. Visin, M. Ciccone, A. Romero, K. Kastner, K. Cho, Y. Bengio, M. Matteucci, and A. Courville, "Reseg: A recurrent neural network-based model for semantic

segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2016, pp. 41–48.
87. Y. Xiang and D. Fox, "Da-rnn: Semantic mapping with data associated recurrent neural networks," *arXiv preprint arXiv:1703.03098*, 2017.
88. Y. Yuan and J. Wang, "Ocnet: Object context network for scene parsing," *arXiv preprint arXiv:1809.00916*, 2018.
89. X. Li, Z. Zhong, J. Wu, Y. Yang, Z. Lin, and H. Liu, "Expectation-maximization attention networks for semantic segmentation," in *Proceedings of the IEEE International Conference on Computer Vision*, 2019, pp. 9167–9176.
90. Z. Huang, X. Wang, L. Huang, C. Huang, Y. Wei, and W. Liu, "Ccnet: Criss-cross attention for semantic segmentation," in *Proceedings of the IEEE International Conference on Computer Vision*, 2019, pp. 603–612.
91. P. S. of Medicine University of Pennsylvania, "Brain tumor segmentation," http://braintumorsegmentation.org/.
92. D. Group, "Medical imaging computing and computerassisted intervention," https://miccai2020.org/en/.
93. S. Bakas *et al.*, "Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge," *arXiv [Preprint]*, 2018.
94. J. Menze, Bjoern H *et al.*, "The multimodal brain tumor image segmentation benchmark (brats)," *IEEE transactions on medical imaging*, vol. 34, no. 10, pp. 1993–2024, 2014.
95. M. Livne, J. Rieger, O. U. Aydin, A. A. Taha, E. M. Akay, T. Kossen, J. Sobesky, J. D. Kelleher, K. Hildebrand, D. Frey *et al.*, "A u-net deep learning framework for high performance vessel segmentation in patients with cerebrovascular disease," *Frontiers in neuroscience*, vol. 13, p. 97, 2019.
96. H. Dong, G. Yang, F. Liu, Y. Mo, and Y. Guo, "Automatic brain tumor detection and segmentation using u-net based fully convolutional networks," in *annual conference on medical image understanding and analysis*. Springer, 2017, pp. 506–517.
97. F. Girard, C. Kavalec, and F. Cheriet, "Joint segmentation and classification of retinal arteries/veins from fundus images," *Artificial intelligence in medicine*, vol. 94, pp. 96–109, 2019.
98. K. Kamnitsas, C. Ledig, V. F. Newcombe, J. P. Simpson, A. D. Kane, D. K. Menon, D. Rueckert, and B. Glocker, "Efficient multi-scale 3d cnn with fully connected crf for accurate brain lesion segmentation," *Medical image analysis*, vol. 36, pp. 61–78, 2017.
99. P. Sandur, C. Naveena, V. M. Aradhya, and K. Nagasundara, "Segmentation of brain tumor tissues in hgg and lgg mr images using 3d u-net convolutional neural network," *International Journal of Natural Computing Research (IJNCR)*, vol. 7, no. 2, pp. 18–30, 2018.
100. X. Li, H. Chen, X. Qi, Q. Dou, C.-W. Fu, and P.-A. Heng, "H-denseunet: hybrid densely connected unet for liver and tumor segmentation from ct volumes," *IEEE transactions on medical imaging*, vol. 37, no. 12, pp. 2663–2674, 2018.
101. Z. Zeng, W. Xie, Y. Zhang, and Y. Lu, "Ric-unet: An improved neural network based on unet for nuclei segmentation in histology images," *Ieee Access*, vol. 7, pp. 21 420–21 428, 2019.
102. J. I. Orlando, P. Seeböck, H. Bogunović, S. Klimscha, C. Grechenig, S. Waldstein, B. S. Gerendas, and U. Schmidt-Erfurth, "U2-net: A bayesian u-net model with epistemic uncertainty feedback for photoreceptor layer segmentation in pathological oct scans," in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*. IEEE, 2019, pp. 1441–1445.
103. S. Bakas, H. Akbari, A. Sotiras, M. Bilello, M. Rozycki, J. Kirby, J. Freymann, K. Farahani, and C. Davatzikos, "Segmentation labels and radiomic features for

the pre-operative scans of the tcga-lgg collection," *The cancer imaging archive*, vol. 286, 2017.
104. M. Havaei *et al.*, "Brain tumor segmentation with deep neural networks," *Medical image analysis*, vol. 35, pp. 18–31, 2017.
105. S. S. M. Salehi, D. Erdogmus, and A. Gholipour, "Auto-Context Convolutional Neural Network (Auto-Net) for Brain Extraction in Magnetic Resonance Imaging," *IEEE Transactions on Medical Imaging*, vol. 36, no. 11, p. 2319–2330, 2017.
106. S. M. Shankaranarayana, K. Ram, K. Mitra, and M. Sivaprakasam, "Fully Convolutional Networks for Monocular Retinal Depth Estimation and Optic Disc-Cup Segmentation," *IEEE Journal of Biomedical and Health Informatics*, 2019.
107. C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning," *arXiv preprint arXiv:1602.07261*, 2016.
108. F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," *arXiv preprint arXiv:1511.07122*, 2015.
109. N. Ibtehaz and M. S. Rahman, "MultiResUNet: Rethinking the U-Net Architecture for Multimodal Biomedical Image Segmentation," *arXiv [Preprint]*, 2019.
110. C. Wang, M. Rajchl, A. Chan, and E. Ukwatta, "An ensemble of U-Net architecture variants for left atrial segmentation," in *Medical Imaging 2019: Computer-Aided Diagnosis*, vol. 10950. International Society for Optics and Photonics, 2019, p. 109500M, (San Diego, CA, USA).
111. S. Li and G. K. F. Tso, "Bottleneck supervised u-net for pixel-wise liver and tumor segmentation," *arXiv preprint arXiv:1810.10331*, 2018.
112. H. Li, A. Li, and M. Wang, "A novel end-to-end brain tumor segmentation method using improved fully convolutional networks," *Computers in biology and medicine*, vol. 108, pp. 150–160, 2019.
113. S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167*, 2015.
114. C. Nwankpa, W. Ijomah, A. Gachagan, and S. Marshall, "Activation functions: Comparison of trends in practice and research for deep learning," *arXiv preprint arXiv:1811.03378*, 2018.
115. W. Shi, F. Jiang, and D. Zhao, "Single image super-resolution with dilated convolution based multi-scale information learning inception module," in *2017 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2017, pp. 977–981.
116. S. Ruder, "An overview of gradient descent optimization algorithms," *arXiv preprint arXiv:1609.04747*, 2016.
117. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.
118. K. He, X. Zhang, S. Ren, and J. Sun, "Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification," *2015 IEEE International Conference on Computer Vision (ICCV)*, 2015, (Santiago, Chile).
119. F. Chollet *et al.*, "Keras," https://keras.io, 2015.
120. M. Abadi *et al.*, "TensorFlow: A system for large-scale machine learning," in *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, 2016, pp. 265–283, (Savannah, GA, USA). [Online]. Available: https://www.usenix.org/system/files/conference/osdi16/osdi16-abadi.pdf
121. Google, "Gpus on compute engine," https://cloud.google.com/compute/docs/gpus.
122. F. Isensee, P. Kickingereder, W. Wick, M. Bendszus, and K. H. Maier-Hein, "No New-Net," in *International MICCAI Brainlesion Workshop*. Springer, 2018, pp. 234–244, (Granada, Spain).

123. P.-Y. Kao, T. Ngo, A. Zhang, J. W. Chen, and B. Manjunath, "Brain Tumor Segmentation and Tractographic Feature Extraction from Structural MR Images for Overall Survival Prediction," in *International MICCAI Brainlesion Workshop*. Springer, 2018, pp. 128–141, (Granada, Spain).
124. J. Zhang, X. Lv, Q. Sun, Q. Zhang, X. Wei, and B. Liu, "Sdresu-net: Separable and dilated residual u-net for mri brain tumor segmentation," *Current Medical Imaging Formerly Current Medical Imaging Reviews*, vol. 15, 08 2019.
125. D. Dera, G. Rasool, and N. Bouaynaya, "Extended variational inference for propagating uncertainty in convolutional neural networks," in *IEEE 29th International Workshop on Machine Learning for Signal Processing (MLSP)*, Oct 2019, pp. 1–6.
126. D. M. Blei, A. Kucukelbir, and J. D. McAuliffe, "Variational inference: A review for statisticians," *Journal of the American Statistical Association*, vol. 112, no. 518, pp. 859–877, 2017.
127. A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*, 4th ed. McGraw-Hill Higher Education, 2002.
128. M. Bishop, *Pattern Recognition & Machine Learning*. New York: Springer, Aug. 2006.
129. D. Simon, *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. Wiley-Interscience, 2006.
130. S. Pereira, A. Pinto, V. Alves, and C. A. Silva, "Brain tumor segmentation using convolutional neural networks in MRI images," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1240–1251, 2016.
131. N. C. Institute, "Prostate-specific antigen," https://www.cancer.gov/types/prostate/psa-fact-sheet/.
132. M. Basseville, I. V. Nikiforov *et al.*, *Detection of abrupt changes: theory and application*. prentice Hall Englewood Cliffs, 1993, vol. 104.
133. E. Brodsky and B. S. Darkhovsky, *Nonparametric methods in change point problems*. Springer Science & Business Media, 2013, vol. 243.
134. C.-J. Kim, J. C. Morley, and C. R. Nelson, "The structural break in the equity premium," *Journal of Business & Economic Statistics*, vol. 23, no. 2, pp. 181–191, 2005.
135. A. B. Olshen, E. Venkatraman, R. Lucito, and M. Wigler, "Circular binary segmentation for the analysis of array-based dna copy number data," *Biostatistics*, vol. 5, no. 4, pp. 557–572, 2004.
136. A. Futschik, T. Hotz, A. Munk, and H. Sieling, "Multiscale dna partitioning: statistical evidence for segments," *Bioinformatics*, vol. 30, no. 16, pp. 2255–2262, 2014.
137. J. Reeves, J. Chen, X. L. Wang, R. Lund, and Q. Q. Lu, "A review and comparison of changepoint detection techniques for climate data," *Journal of applied meteorology and climatology*, vol. 46, no. 6, pp. 900–915, 2007.
138. O. Andrén, K. Fall, L. Franzén, S.-O. Andersson, J.-E. Johansson, and M. A. Rubin, "How well does the gleason score predict prostate cancer death? a 20-year followup of a population based cohort in sweden," *The Journal of urology*, vol. 175, no. 4, pp. 1337–1340, 2006.
139. N. C. Institute, "Seer program," https://seer.cancer.gov/registries/cancer_registry/index.html, 2020.
140. K. Y. Bilimoria, A. K. Stewart, D. P. Winchester, and C. Y. Ko, "The national cancer data base: a powerful initiative to improve cancer care in the united states," *Annals of surgical oncology*, vol. 15, no. 3, pp. 683–690, 2008.
141. CDC, "Ncpr," https://www.cdc.gov/cancer/npcr/index.htm, 2020.
142. CBTR-US, "Central brain tumor registry of the us," https://cbtrus.org/, 2020.
143. A. G. Lai, L. Pasea, A. Banerjee, S. Denaxas, M. Katsoulis, W. H. Chang, B. Williams, D. Pillay, M. Noursadeghi, D. Linch *et al.*, "Estimating excess

mortality in people with cancer and multimorbidity in the covid-19 emergency," *medRxiv*, 2020.
144. HDR-UK, "The health data research hub for cancer," https://www.hdruk.ac.uk/help-with-your-data/our-hubs-across-the-uk/data-can/, 2020.
145. M. Otth and K. Scheinemann, "Surveillance imaging for high-grade childhood brain tumors: What to do 10 years after completion of treatment?" *Pediatric blood & cancer*, vol. 65, no. 11, p. e27311, 2018.
146. Q. T. Ostrom, D. J. Cote, M. Ascha, C. Kruchko, and J. S. Barnholtz-Sloan, "Adult glioma incidence and survival by race or ethnicity in the united states from 2000 to 2014," *JAMA oncology*, vol. 4, no. 9, pp. 1254–1262, 2018.
147. D. J. Cote, Q. T. Ostrom, H. Gittleman, K. R. Duncan, T. S. CreveCoeur, C. Kruchko, T. R. Smith, M. J. Stampfer, and J. S. Barnholtz-Sloan, "Glioma incidence and survival variations by county-level socioeconomic measures," *Cancer*, vol. 125, no. 19, pp. 3390–3400, 2019.
148. C. Kruchko, Q. T. Ostrom, H. Gittleman, and J. S. Barnholtz-Sloan, "The cbtrus story: providing accurate population-based statistics on brain and other central nervous system tumors for everyone," 2018.
149. Q. T. Ostrom, H. Gittleman *et al.*, "CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2011–2015," *Neuro-Oncology*, vol. 20, 10 2018. [Online]. Available: https://doi.org/10.1093/neuonc/noy131
150. R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2020," *CA: a cancer journal for clinicians*, vol. 70, no. 1, pp. 7–30, 2020.
151. N. I. for Health and C. E. G. Britain), *Brain tumours (primary) and brain metastases in adults*. National Institute for Health and Care Excellence, 2018.
152. H. M. Fathallah-Shaykh, A. DeAtkine, E. Coffee, E. Khayat, A. K. Bag, X. Han, P. P. Warren, M. Bredel, J. Fiveash, J. Markert, N. Bouaynaya, and L. B. Nabors, "Diagnosing growth in low-grade gliomas with and without longitudinal volume measurements: A retrospective observational study," *PLOS Medicine*, vol. 16, no. 5, pp. 1–16, 05 2019. [Online]. Available: https://doi.org/10.1371/journal.pmed.1002810