University of Engineering and Technology

INT3407 21

BIOINFORMATICS

A REPORT

ON

MACHINE LEARNING IN BRAIN MRI

GROUP 8

SUBMITTED BY

Le Bang Giang - 18020428
Pham Truong Giang - 18020424
Nguyen Xuan Hieu - 18020507

(Academic Year: 2020-2021)

Contents

1 INTRODUCTION

1.1 Introduction

Machine learning has witnessed a tremendous amount of attention over the last few years. Deep neural networks are now the state-of-the-art machine learning models across a variety of areas, from image analysis to natural language processing, and widely deployed in academia and industry. These developments have a huge potential for medical imaging technology, medical data analysis, medical diagnostics and healthcare in general, slowly being realized. As this has become a very broad and fast expanding field , we will put particular focus on deep learning in MRI. MRI, which stands for magnetic resonance imaging, is a common procedure around the world. MRI uses a strong magnetic field and radio waves to create detailed images of the organs and tissues within the body. Since its invention, doctors and researchers continue to refine MRI techniques to assist in medical procedures and research. The development of MRI revolutionized medicine. In this project, we focus on applying machine learning methods to detect some of the common types of tumors from brain MRI scans.

1.2 Objectives

Apply deep learning techniques to accurately identify and segment tumor regions in MRI brain images.

2 MRI OVERVIEW

2.1 What is MRI?

Magnetic Resonance Imaging (MRI) is a non-invasive imaging technology that produces three dimensional detailed anatomical images. It is often used for disease detection, diagnosis, and treatment monitoring. It is based on sophisticated technology that excites and detects the change in the direction of the rotational axis of protons found in the water that makes up living tissues. MRI image and resolution is quite detailed, and it can detect tiny changes of structures within the body. For some procedures, contrast agents, such as gadolinium, are used to increase the accuracy of the images. MRIs are a relatively new technology to hit the medical world, and have completely revolutionized medical imaging and the diagnosing process as we know it. In-vivo images can be taken of the human body, meaning that internal images can be seen without making any incisions. Completely non-intrusive procedures are used, which makes MRI's very effective, but somewhat expensive, for doctors to use.

MRI does not involve X-rays or the use of ionizing radiation, which distinguishes it from CT and PET scans. MRI is a medical application of nuclear magnetic resonance (NMR) which can also be used for imaging in other NMR applications, such as NMR spectroscopy.

The first patent of MRI technology was filed by Raymond Damadian in 1972. The doctor and researcher discovered that images created using MRI technology could be helpful during the process of diagnosing a patient. Damadian was the first person to use MRI to make a medical diagnosis. Research completed by Damadian revealed that

MRI could be used to detect cancerous tissue in animals. His patent was approved in 1974, and Damadian developed MRI for medical use by designing and building a full-body MRI machine in 1977.

2.2 How MRI works?

MRIs employ powerful magnets which produce a strong magnetic field that forces protons in the body to align with that field. A powerful, uniform, external magnetic field is employed to align the protons that are normally randomly oriented within the water nuclei of the tissue being examined. This alignment (or magnetization) is next perturbed or disrupted by introduction of an external Radio Frequency (RF) energy. When this radiofrequency current is pulsed through the patient, the protons are stimulated, and spin out of equilibrium, straining against the pull of the magnetic field. The nuclei return to their resting alignment through various relaxation processes and in so doing emit RF energy. After a certain period following the initial RF, the emitted signals are measured. Fourier transformation is used to convert the frequency information contained in the signal from each location in the imaged plane to corresponding intensity levels, which are then displayed as shades of gray in a matrix arrangement of pixels. By varying the sequence of RF pulses applied and collected, different types of images are created. Repetition Time (TR) is the amount of time between successive pulse sequences applied to the same slice. Time to Echo (TE) is the time between the delivery of the RF pulse and the receipt of the echo signal. When the radiofrequency field is turned off, the MRI sensors are able to detect the energy released as the protons realign with the magnetic field. The time it takes for the protons to realign with the magnetic field, as well as the amount of energy released, changes depending on the environment and the chemical nature of the molecules. Thus physicians are then able to tell the difference between various types of tissues based on these magnetic.

To obtain an MRI image, a patient is placed inside a large magnet and must remain very still during the imaging process in order not to blur the image. Contrast agents (often containing the element Gadolinium) may be given to a patient intravenously before or during the MRI to increase the speed at which protons realign with the magnetic field. The faster the protons realign, the brighter the image.

T1 and T2

Tissue can be characterized by two different relaxation times – T1 and T2. T1 (longitudinal relaxation time) is the time constant which determines the rate at which excited protons return to equilibrium. It is a measure of the time taken for spinning protons to realign with the external magnetic field. T2 (transverse relaxation time) is the time constant which determines the rate at which excited protons reach equilibrium or go out of phase with each other. It is a measure of the time taken for spinning protons to lose phase coherence among the nuclei spinning perpendicular to the main field.

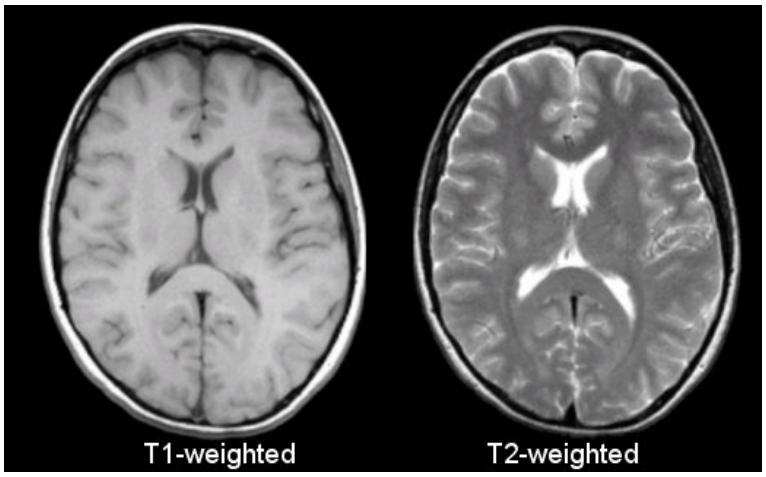

Figure 1: Comparison T1 vs. T2.

The most common MRI sequences are T1-weighted and T2-weighted scans. T1-weighted images are produced by using short TE and TR times. The contrast and brightness of the image are predominately determined by T1 properties of tissue. Conversely, T2-weighted images are produced by using longer TE and TR times. In these images, the contrast and brightness are predominately determined by the T2 properties of tissue. In general, T1- and T2-weighted images can be easily differentiated by looking the CSF. CSF is dark on T1-weighted imaging and bright on T2-weighted imaging.

2.3 What is MRI used for?

MRI scanners are particularly well suited to image the non-bony parts or soft tissues of the body. They differ from computed tomography (CT), in that they do not use the damaging ionizing radiation of x-rays. The brain, spinal cord and nerves, as well as muscles, ligaments, and tendons are seen much more clearly with MRI than with regular x-rays and CT; for this reason MRI is often used to image knee and shoulder injuries. In the brain, MRI can differentiate between white matter and grey matter and can also be used to diagnose aneurysms and tumors. Because MRI does not use x-rays or other radiation, it is the imaging modality of choice when frequent imaging is required for diagnosis or therapy, especially in the brain. However, MRI is more expensive than x-ray imaging or CT scanning.

One kind of specialized MRI is functional Magnetic Resonance Imaging (fMRI.) This is used to observe brain structures and determine which areas of the brain "activate" (consume more oxygen) during various cognitive tasks. It is used to advance the understanding of brain organization and offers a potential new standard for assessing neurological status and neurosurgical risk.

2.4 MRI safety

Unlike other imaging forms like X-rays or CT scans, MRI doesn't use ionizing radiation. MRI is increasingly being used to image fetuses during pregnancy, and no adverse effects on the fetus have been demonstrated. Still, the procedure can have risks, and

medical societies don't recommend using MRI as the first stage of diagnosis. Because MRI uses strong magnets, any kind of metal implant, such as a pacemaker, artificial joints, artificial heart valves, cochlear implants or metal plates, screws or rods, pose a hazard. The implant can move or heat up in the magnetic field. Patients should notify their physicians of any form of medical or implant prior to an MR scan.

During the MRI scan, patient lies in a closed area inside the magnetic tube. Some patients can experience a claustrophobic sensation during the procedure. Claustrophobia is the fear of confined spaces. Some people with claustrophobia experience mild anxiety when in a confined space, while others have severe anxiety or a panic attack. Therefore, patients with any history of claustrophobia should relate this to the practitioner who is requesting the test, as well as the radiology staff. A mild sedative can be given prior to the MRI scan to help alleviate this feeling. It is customary that the MRI staff will be nearby during MRI scan. Furthermore, there is usually a means of communication with the staff (such as a buzzer held by the patient) which can be used for contact if the patient cannot tolerate the scan.

3 PROBLEM

3.1 Image data

In clinical settings, usually only a certain number of slices of brain CE-MRI with a large slice gap, not 3D volume, are acquired and available. A 3D model is difficult to construct with such sparse data. Hence, in this project we use the data set based on 2D slices. The brain T1-weighed CE-MRI dataset was acquired from Nanfang Hospital, Guangzhou, China, and General Hospital, Tianjing Medical University, China, from 2005 to 2010. The dataset consists of 3064 slices from 233 patients, containing 708 meningiomas, 1426 gliomas, and 930 pituitary tumors. The images have an inplane resolution of 512×512 with pixel size 0.49×0.49 mm^2. The slice thickness is 6 mm and the slice gap is 1 mm. The tumor border was manually delineated by three experienced radiologists. Some of the examples are illustrated in Figure 2.

3.2 Three types of tumor

All the images in the dataset belong to one of the three types of tumor,

1. pituitary tumor
2. glioma
3. meningioma

Pituitary tumor

Pituitary tumors are abnormal growths that develop in pituitary gland. The pituitary is a small gland found inside the skull just below the brain and above the nasal passages, which are above the fleshy back part of the roof of the mouth (known as the soft palate). The pituitary sits in a tiny bony space called the sella turcica. The nerves that connect the eyes to the brain, called the optic nerves, pass close by it.

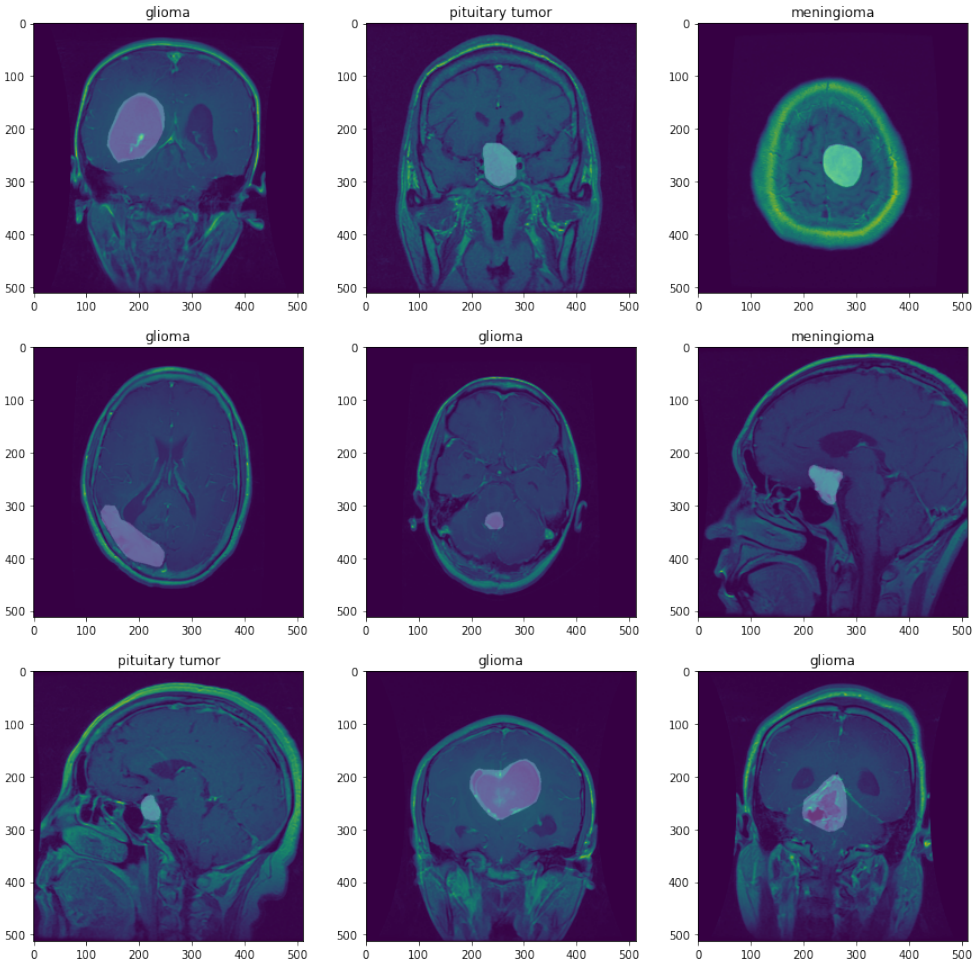

Figure 2: Sample images with the corresponding labels and masks, the shaded regions in the images indicate the areas in which the tumors are.

Almost all pituitary tumors are benign (not cancer) glandular tumors called pituitary adenomas. These tumors are called benign because they don't spread to other parts of the body, like cancers can. Still, even benign pituitary tumors can cause major health problems because they are close to the brain, may invade nearby tissues (like the skull or the sinuses), and because many of them make excess hormones the body doesn't need (functioning tumors). These benign tumors do not spread outside the skull. They usually stay in the sella turcica (the tiny space in the skull that the pituitary gland sits in). Sometimes they grow into the boney walls of the sella turcica and nearby tissues, like blood vessels, nerves, and sinuses. They don't grow very large, but they can have a big impact on a person's health. There is very little room for tumors to grow in this part of the skull. So, if the tumor gets larger than about a centimeter (about half an inch) across, it may grow upward, where it can press on and damage nearby parts of the brain and the nerves that arise from it. This can lead to problems like vision changes or headaches. They're typically treated with surgery, medicine, or radiation.

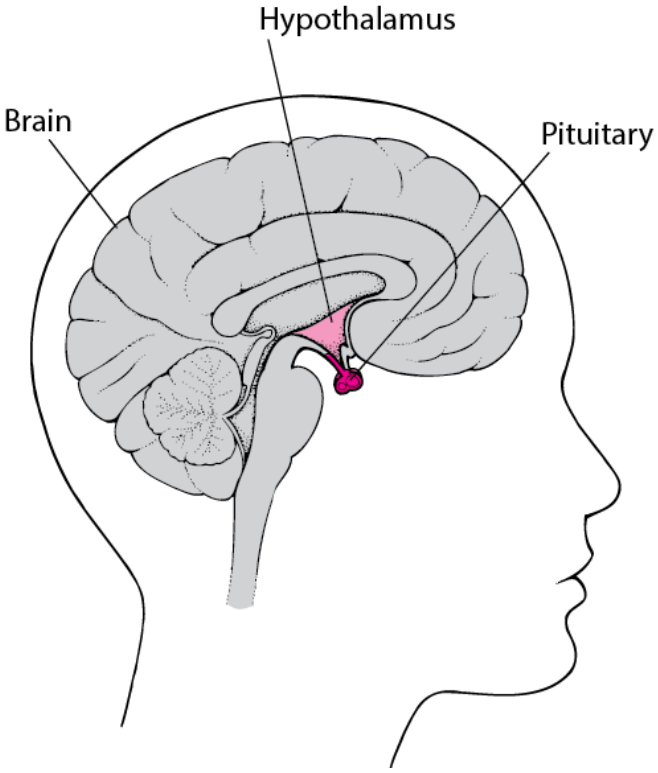

Figure 3: Locating the Pituitary Gland.

Scientists don't know exactly what causes most pituitary tumors. The genes of some pituitary cells change at the start, but the change seems to happen at random. But a condition called multiple endocrine neoplasia type I (MEN 1) may raise the risk of getting the tumor. This condition is passed down through families.

About 10,000 pituitary tumors are diagnosed each year in the United States (we cited this number since we could not find a worldwide statistics). The actual number of pituitary tumors may be much higher than the number of tumors that are found each year. When examining people who have died or who have had imaging tests (like MRI scans) of their brain for other health problems, doctors have found that as many as 1 out of 4 people may have a pituitary adenoma without knowing it. These tumors are often small and never cause any symptoms or health problems, so very few of them would normally be diagnosed at all.

Glioma

A glioma is a type of tumor that starts in the glial cells of the brain or the spine. Gliomas are called intra-axial brain tumors because they grow within the substance of the brain and often mix with normal brain tissue. A brain glioma can cause headaches, vomiting, seizures, and cranial nerve disorders as a result of increased intracranial pressure. A glioma of the optic nerve can cause visual loss. Spinal cord gliomas can cause pain, weakness, or numbness in the extremities.

Glioma is a type of brain cancer that is often – but not always – malignant. In some cases, the tumor cells do not actively reproduce and invade nearby tissues, which makes them noncancerous. However, in most cases, gliomas are cancerous and likely to spread.

The exact causes of gliomas are not known. Diet, radiation, virses and genes are

believed to be the main contributions to the development of gliomas. Gliomas can occur in people of all ages but are more common in adults. Gliomas are slightly more likely to affect men than women.

Gliomas are one of the most common types of primary brain tumors, gliomas comprise about 30 percent of all brain tumors and central nervous system tumours, and 80 percent of all malignant brain tumours.

Meningioma

Meningioma, also known as meningeal tumor, is typically a slow-growing tumor that forms from the meninges, the membranous layers surrounding the brain and spinal cord. Symptoms depend on the location and occur as a result of the tumor pressing on nearby tissue. Many cases never produce symptoms. Occasionally seizures, dementia, trouble talking, vision problems, one sided weakness, or loss of bladder control may occur. Often, meningiomas cause no symptoms and require no immediate treatment. But the growth of benign meningiomas can cause serious problems. In some cases, such growth can be fatal.

The causes of meningiomas are not well understood. Most cases are sporadic, appearing randomly, while some are familial. Persons who have undergone radiation, especially to the scalp, are more at risk for developing meningiomas, as are those who have had a brain injury. Atomic bomb survivors from Hiroshima had a higher than typical frequency of developing meningiomas, with the incidence increasing the closer that they were to the site of the explosion. Dental x-rays are correlated with an increased risk of meningioma, in particular for people who had frequent dental x-rays in the past, when the x-ray dose of a dental x-ray was higher than in the present. Having excess body fat increases the risk. A 2012 review found that mobile telephone use was unrelated to meningioma. People with neurofibromatosis type 2 (NF-2) have a 50% chance of developing one or more meningiomas. Previous injury may also be a risk factor, but a recent study failed to confirm this. Meningiomas have been found in places where skull fractures have occurred. They've also been found in places where the surrounding membrane has been scarred. Some researches suggest a link between meningiomas and the hormone progesterone.

Meningiomas are the most common type of tumor that originates in the central nervous system. Onset is usually in adults. Women are affected about twice as often as men. They are very rare in children. Ninety-two percent of meningiomas are benign. Eight percent are either atypical or malignant. A small number of meningiomas are cancerous. They tend to grow quickly. They also can spread to other parts of the brain and beyond, often to the lungs.

3.3 Data preprocessing

Standardize images

One important constraint that exists in some machine learning algorithms, such as CNN, is the need to resize the images in the dataset to a unified dimension. This implies that the images must be preprocessed and scaled to have identical widths and heights before fed to the learning algorithm. Additionally, resizing images provides a way to control the size of input, because some machine learning tasks do not require high resolution images to achieve good performance, and large-size images make the

models run much slower and consume a lot of resources. In our setting, we resized all the images in the data set to a fixed size of 128 pixels width and height.

Other important standardized preprocessing step before feeding data to neural net is to normalize the input. Normalization consists of subtracting a quantity related to a measure of localization or distance and dividing by a measure of the scale. We applied the normalization procedure for each image data as $x' = \frac{x-\mu}{\sigma}$ where μ is the mean and σ is the standard deviation of the images, x' is the new value of each pixel in the images. After the normalization, each image has the mean of 0 and standard deviation of 1.

Data augmentation

Data augmentation is essential to teach the network the desired invariance and robustness properties, when only few training samples are available. Common preprocessing technique involves augmenting the existing dataset with perturbed versions of the existing images. Scaling, rotations and other affine transformations are typical. This is done to enlarge the dataset and expose the neural network to a wide variety of variations of images. This makes it more likely that the model recognizes objects when they appear in any form and shape, from different perspectives. In our experiment, we applied all of these techniques with some randomness, such that at every epoch the data set will appear slightly different from the previous epoch, making the models more robust to variance and thus achieve better generalization. More specifically, we applied random rotation between 30 and -30 degree, center cropping, horizontal flipping and scaling with probabilities of 0.8, 0.8, 0.8 and 0.7 respectively for each individual image in the dataset. Figure 4 shows some examples after the augmentation is applied.

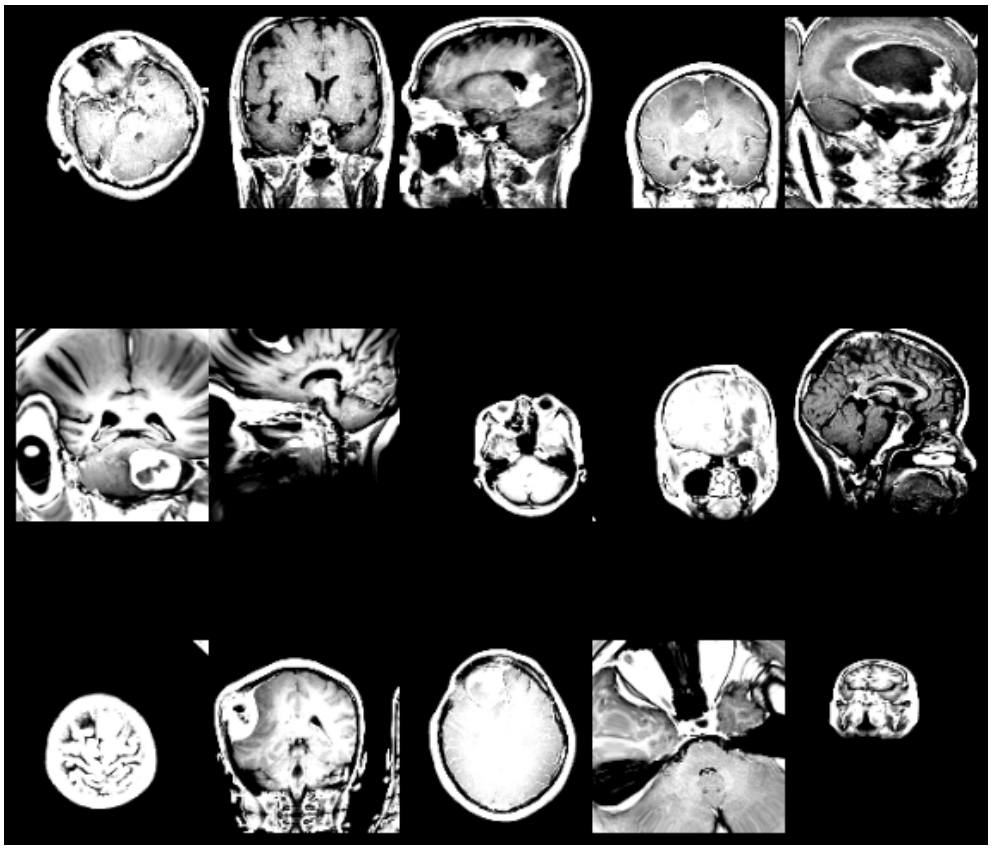

Figure 4: Examples after applying data augmentation.

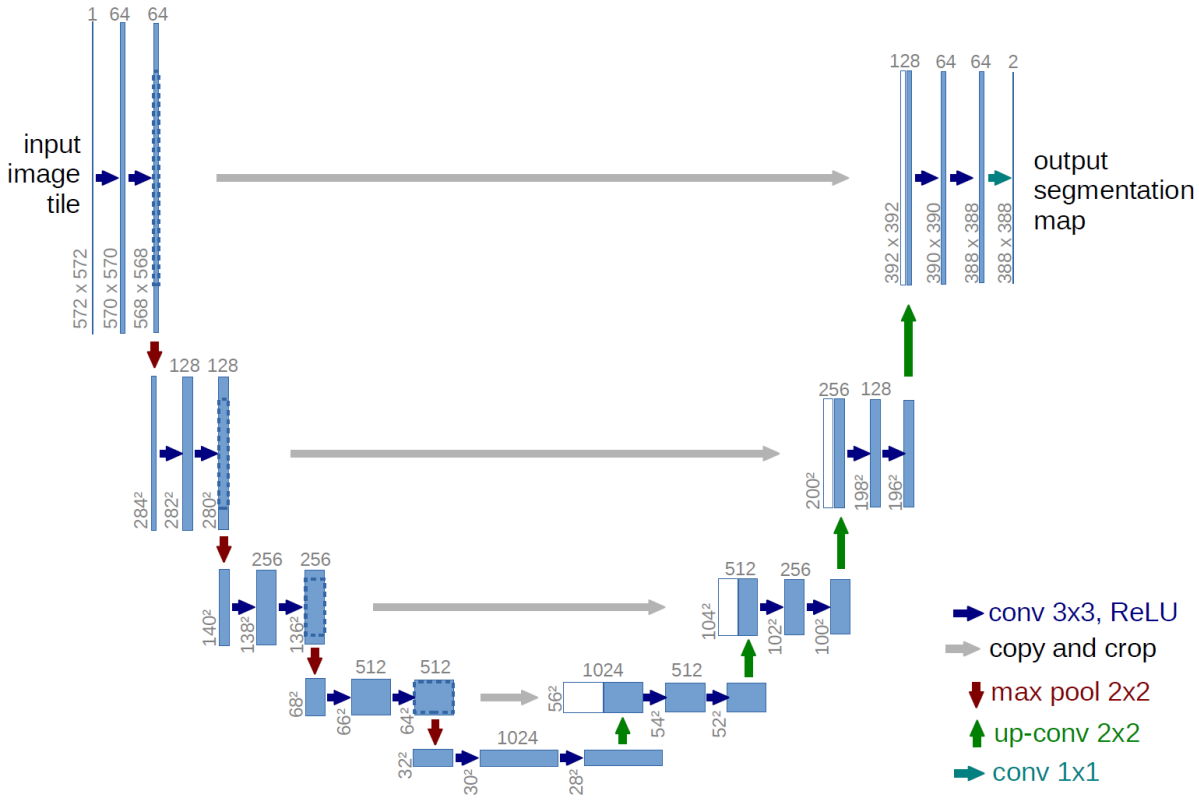

Figure 5: U-net architecture from the author's paper

4 PAPER AND METHOD

4.1 Paper

In this project, we used U-net, a convolutional neural network that was developed for biomedical image segmentation at the Computer Science Department of the University of Freiburg, Germany. U-Net was created by Olaf Ronneberger, Philipp Fischer, Thomas Brox in 2015 at the paper "U-Net: Convolutional Networks for Biomedical Image Segmentation".

4.2 Network architecture

The network architecture is illustrated in Figure 5. It consists of a contracting path (left side) and an expansive path (right side), which gives it the u-shaped architecture. The contracting path is a typical convolutional network that consists of repeated application of two 3x3 convolutions, each followed by a batchnorm layer and a rectified linear unit (ReLU) with a 2x2 max pooling operation with stride 2 at the end for downsampling. At each downsampling step we double the number of feature channels. Thus, the spatial information is reduced while feature information is increased. Every step in the expansive path consists of an upsampling of the feature map followed by a 2x2 convolution ("up-convolution") that halves the number of feature channels. It then combines the feature and spatial information through a concatenation with high-resolution features from the corresponding contracting path, each followed by two 3x3 convolutions with a batchnorm and ReLU layer in between as in the contracting path.

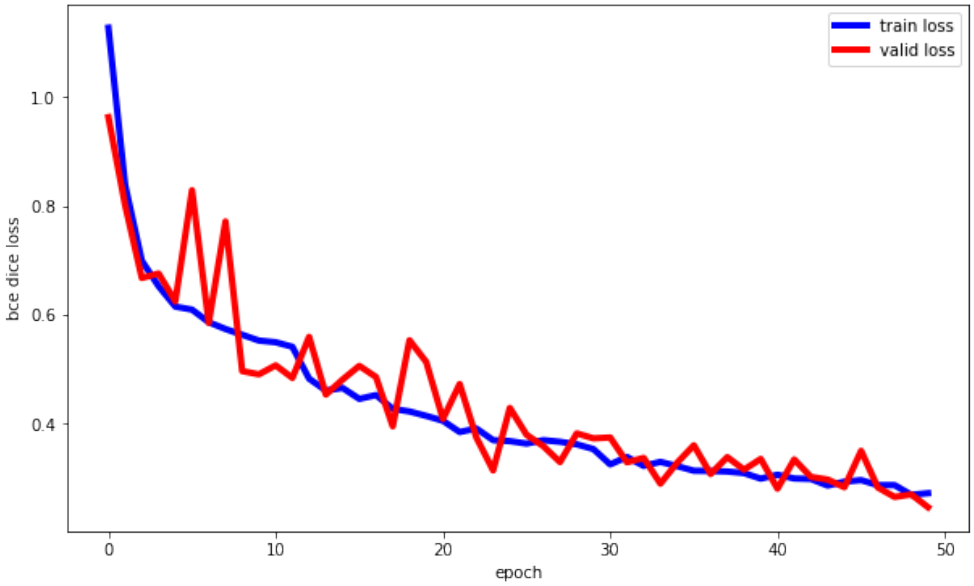

Figure 6: Train and Validation Learning Curves

The outputs of the model have the same size as the input images with three channels correspond to the masks of three different tumor classes.

4.3 Training

The input images and their corresponding segmentation maps are used to train the network with the Adam optimizer implementation of Pytorch. The loss function is computed by a Dice loss over the final feature map combined with the cross entropy loss function. Dice loss is based on the Sørensen–Dice coefficient or Tversky index, which can be defined as $D = 2\sum_i^N p_i g_i / (\sum_i^N p_i^2 + \sum_i^N g_i^2)$, where p_i and g_i represent pairs of corresponding pixel values of prediction and ground truth, respectively. In segmentation detection scenario, the values of p_i and g_i is either 0 or 1, representing whether the pixel is mask (value of 1) or not (value of 0). Therefore, the denominator is the sum of total mask pixels of both prediction and ground truth, and the numerator is the sum of correctly predicted mask pixels because the sum increments only when p_i and g_i match (both of value 1). The Dice loss then can be defined as $L = 1 - D$ as to maximize the overlap between two sets.

To make the training process more efficient, we combined both Dice loss with cross-entropy loss as an aggregate loss function. Cross-entropy is the concept borrowed from information theory, the cross-entropy between two probability distributions p and q over the same underlying set of events measures the average number of bits needed to identify an event drawn from the set if a coding scheme used for the set is optimized for an estimated probability distribution q, rather than its true distribution p. However, cross-entropy loss is not significant in the context of image segmentation. In our experiment, after several update steps, cross-entropy loss became negligible and was dominated by the Dice loss.

From 3064 images in the dataset, we split them into training set, test set and validation set with the sizes of 2619, 307 and 138 images for each set respectively. We trained the model with NVIDIA TESLA P100 GPUs provided by Kaggle platform.

After 50 epochs (about 40 minutes), we achieved an accuracy of 74% DSC on the test set. The train and validation curve is illustrated as in Figure 6.

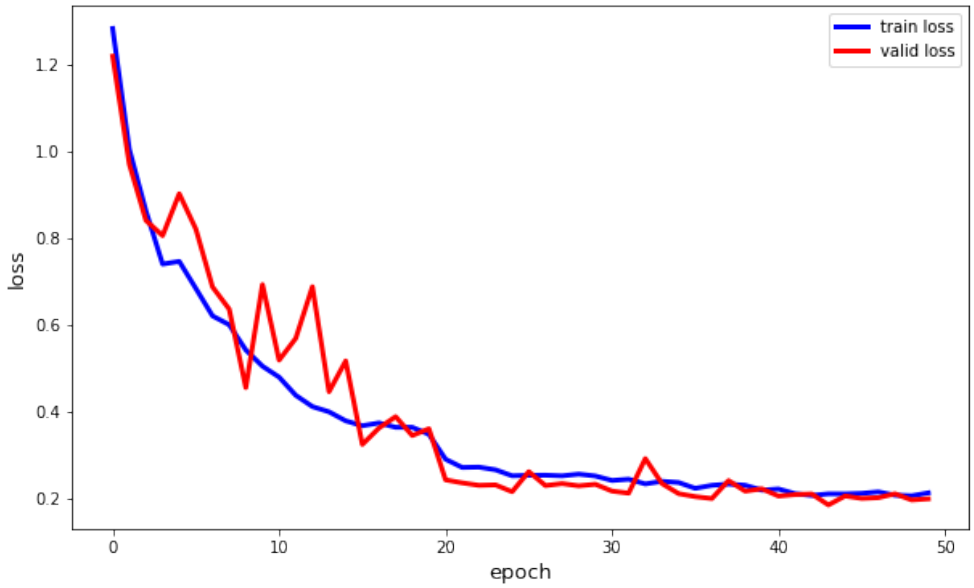

Figure 7: Train and Validation Learning Curves after modifications

4.4 Improving model

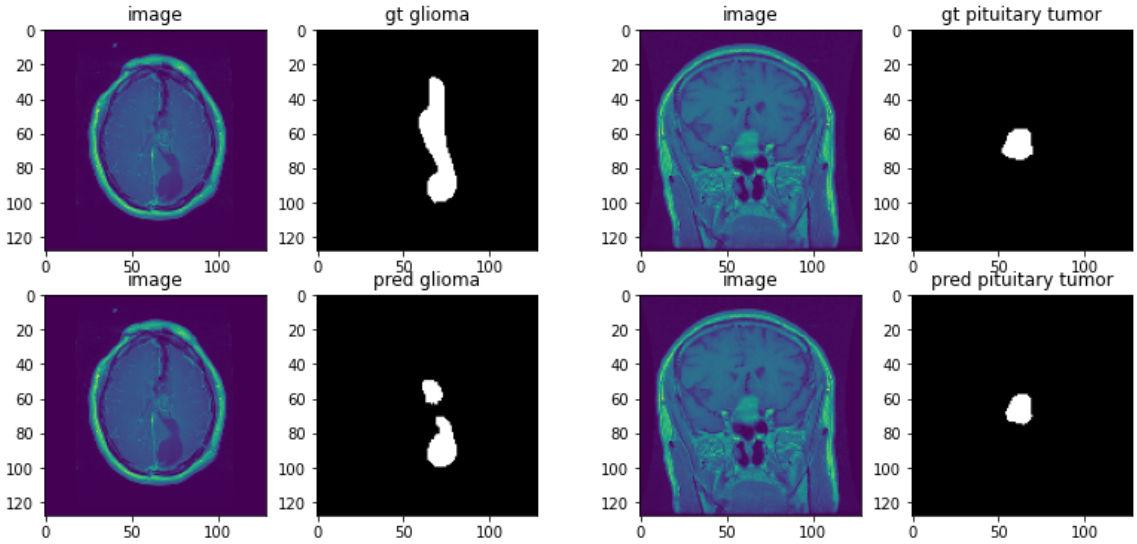

Figure 8: Two sample images with their ground truth and predicted masks.

From the result of the previous section, it was suggested that we could improve the model by modifying its output structure. Concretely, we know that each image in the data set belongs to exactly one class, so instead of predicting three classes of tumors separately, which is an error-prone process, we can adjust the model to predict just the label and the mask of that label. In this manner, we reduce the amount of information that the model need to learn, hence increase the accuracy by a significant margin. For

labeling classes, we attached an additional neural network at the bottom layer of U-net, this subnetwork works as a classification network and consists of several CNN layers followed by BatchNorm + ReLU with a Softmax layer at the end to normalize the output of the network to a probability distribution over predicted output classes. The overall loss is computed as the sum of the loss from the output mask of the original net and the loss from the additional net. This new setting achieved a higher accuracy of 79% DSC on the test set. Some of the results are shown in the Figure 8.

5 APPENDIX

5.1 Convolutional neural networks

A convolutional neural network (CNN, or ConvNet) is a class of deep neural networks, most commonly applied to analyzing visual imagery. The name "convolutional neural network" indicates that the network employs a mathematical operation called convolution. Convolutional networks are a specialized type of neural networks that use convolution in place of general matrix multiplication in at least one of their layers.

Convolutional layers are the major building blocks used in convolutional neural networks. Convolutional layers convolve the input and pass its result to the next layer. This is similar to the response of a neuron in the visual cortex to a specific stimulus. Each convolutional neuron processes data only for its receptive field. Although fully connected feedforward neural networks can be used to learn features as well as classify data, it is not practical to apply this architecture to images.

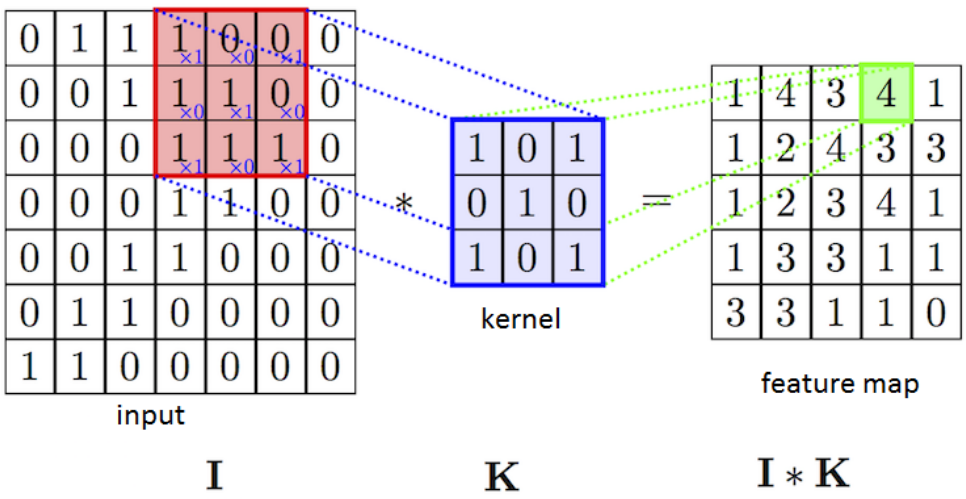

Figure 9: Dot-product between Filter and Input, resulting as a feature map.

Repeated application of the same filter (aka kernel) to an input results in a map of activations called a feature map, indicating the locations and strength of a detected feature in an input, images in our case.

A simplified illustration of how a convolutional layer works is shown in Figure 9. Each pixel in the feature map is calculated by taking the dot product between the kernel and some region in the original image. The kernel will then move step by step

from left to right, top to bottom on the input. At each step, it will move by number of strides that was specified (1 in this case).

5.2 Activation function

An activation function then is applied to the feature map. Typically, the activation function is commonly a ReLU layer, which can be defined as $f(x) = x^+ = max(0, x)$. They introduce a non-linearity at zero that can be used for decision making. The output of the activation layers and is subsequently followed by additional convolutions such as pooling layers, fully connected layers and normalization layers.

Other functions are also used to increase nonlinearity, for example the saturating hyperbolic tangent $f(x) = tanh(x)$, $f(x) = |tanh(x)|$, and the sigmoid function $\sigma(x) = (1 + e^{-x})^{-1}$. ReLU is often preferred to other functions because it trains the neural network several times faster without a significant penalty to generalization accuracy.

Finally, in order for the output to have a nice probability interpretation, a softmax function is often applied as a final activation function. The softmax function is a function that turns a vector of K real values into a vector of K real values that sum to 1. The input values can be positive, negative, zero, or greater than one, but the softmax transforms them into values between 0 and 1, so that they can be interpreted as probabilities. If one of the inputs is small or negative, the softmax turns it into a small probability, and if an input is large, then it turns it into a large probability, but it will always remain between 0 and 1. The softmax formula is as follows:

$$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}$$

Where all the z_i values are the elements of the input vector $\vec{z}$ and can take any real value, K is the number of classes in the multi-class classifier. In case of binary-class classification, the softmax function is similar to the sigmoid function.

5.3 Pooling layers

The Pooling layer is responsible for reducing the spatial size of the Convolved Feature. This is to decrease the computational power required to process the data through dimensionality reduction. Furthermore, it is useful for extracting dominant features which are rotational and positional invariant, thus maintaining the process of effectively training of the model.

There are two types of Pooling: Max Pooling and Average Pooling. Max Pooling returns the maximum value from the portion of the image covered by the Kernel. On the other hand, Average Pooling returns the average of all the values from the portion of the image covered by the Kernel. The most commonly used pooling layer is Max Pooling with filters of size 2×2 applied with a stride of 2 downsamples at every depth slice in the input by 2 along both width and height, discarding 75% of the activations.

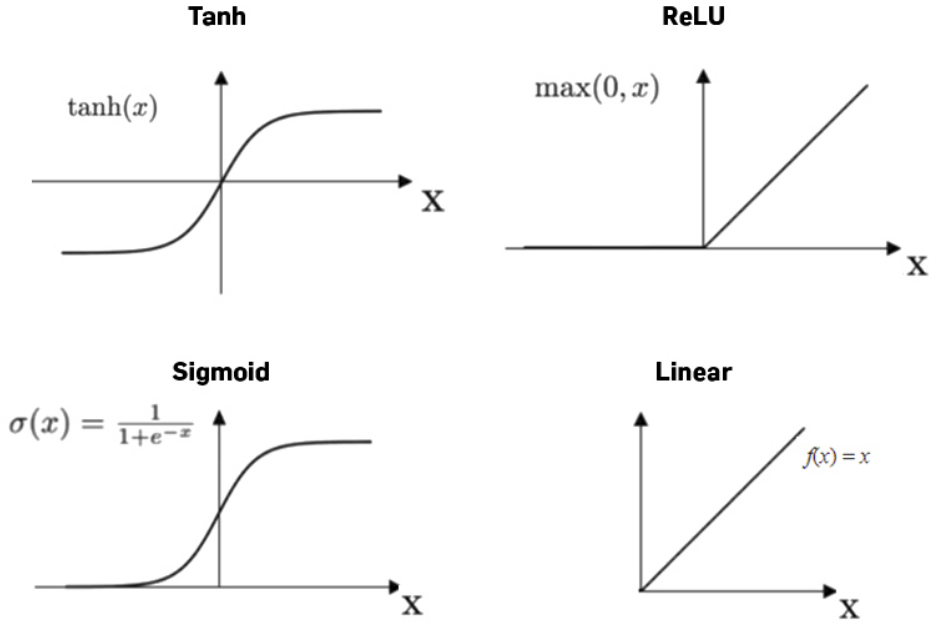

Figure 10: Some common activation functions.

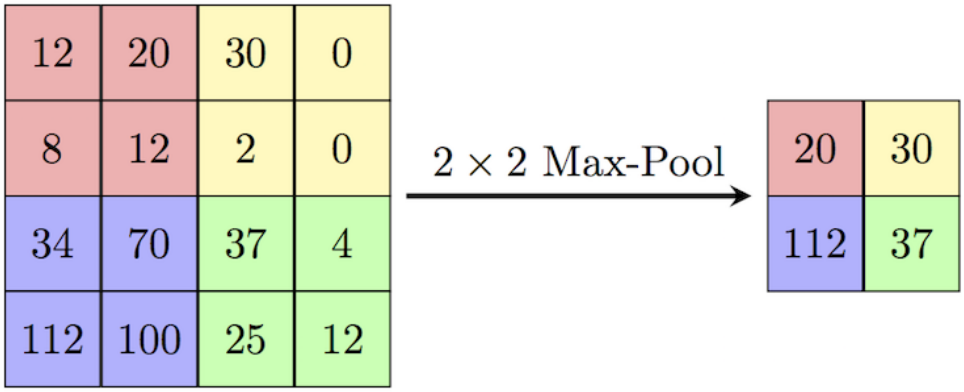

Figure 11: Max Pooling Layer with the stride of 2 and filter size 2x2 reduces size of input volume.

5.4 Batch normalization

Batch normalization (also known as batch norm) is a technique for training very deep neural networks that standardizes the inputs to a layer for each mini-batch. This has the effect of stabilizing the learning process and dramatically reducing the number of training epochs required to train deep networks.

Training deep neural networks is challenging because of a number of problems. One possible reason for this difficulty is the distribution of the inputs to layers deep in the network may change after each mini-batch when the weights are updated. When the parameters of a layer change, so does the distribution of inputs to subsequent layers. These shifts in input distributions can be problematic for neural networks, especially deep neural networks that could have a large number of layers. Batch normalization is a method intended to mitigate internal covariate shift for neural networks.

Use B to denote a mini-batch of size m of the entire training set. The empirical mean

and variance of B could thus be denoted as $\mu_B = \frac{1}{m}\sum_{i=1}^{m} x_i$ and $\sigma_B = \frac{1}{m}\sum_{i=1}^{m}(x_i - \mu_B)^2$. For a layer of the network with d-dimensional input, $x = (x^{(1)}, x^{(2)}, ...x^{(d)})$, each dimension of its input is then normalized separately.

$$\hat{x}_i^{(k)} = \frac{x_i^{(k)} - \mu_B^{(k)}}{\sqrt{{\sigma_B^{(k)}}^2 + \epsilon}}$$

In the equation above, ϵ is added in the denominator for numerical stability and is an arbitrarily small constant to avoid zero division error. The final normalized activation $\hat{y}_i^{(k)}$ can then be described as $\hat{y}_i^{(k)} = \gamma^{(k)}\hat{x}_i^{(k)} + \beta^{(k)}$, where the parameters $\gamma^{(k)}$ and $\beta^{(k)}$ are learnable parameters.

6 REFERENCE

Convolutional neural networks: https://en.wikipedia.org/wiki/Convolutional_neural_network
MRI: https://en.wikipedia.org/wiki/Magnetic_resonance_imaging
MRI: https://www.nhs.uk/conditions/mri-scan/
Meningioma: https://en.wikipedia.org/wiki/Meningioma
Glioma: https://en.wikipedia.org/wiki/Glioma
Pituitary tumor: https://www.cancer.org/cancer/pituitary-tumors/about/what-is-pituitary-tumor.html
Activation function: https://en.wikipedia.org/wiki/Activation_function
Batch normalization: https://en.wikipedia.org/wiki/Batch_normalization
Pooling layer: https://computersciencewiki.org/index.php/Max-pooling_/_Pooling
Softmax layer: https://deepai.org/machine-learning-glossary-and-terms/softmax-layer

Additional resources
U-Net: Convolutional Networks for Biomedical Image Segmentation,
https://arxiv.org/pdf/1505.04597.pdf.
Link code: https://www.kaggle.com/banggiangle/multi-class-unet
Link dataset: https://www.kaggle.com/awsaf49/brain-tumor