

ECE 470

Artificial Intelligence

Summer 2021



University of Victoria

Detecting Brain Tumours with Machine Learning and MRI Scans

July 29th, 2021

David Bishop	V00884250
Derrick Ushko	V00883868
Ben Austin	V00892013

Table of Contents

List of Figures	2
List of Tables	3
Abstract	3
Introduction	3
Background	3
Motivation	4
Related Work	4
Problem Formulation	4
Problem	4
Data	5
Scope	5
Methodology	5
Results and Discussions	9
Overall Results	9
Comparisons	9
Justification	10
Challenges	11
Conclusion	12
Future Work	12
References	13

List of Figures

Figure 1: Example convolutional network	6
Figure 2: Max pooling example	6
Figure 3: The team's convolutional network implementation	7
Figure 4: Loss and accuracy graph over each epoch	9

List of Tables

Table 1: Variations of the model and results	8
--	---

Abstract

Brain tumors are a collection of abnormal cells within or around the brain which cause severe damage to the nervous system. They contribute to a significant burden of disease and are among a leading cause of solid state cancer death in Canada. Brain tumors are typically diagnosed through MRI scans, which are able to depict lesion location, size, and impact. Within the medical industry, artificial intelligence is becoming increasingly prominent in several fields. Machine learning is commonly used in image recognition applications, including MRI. The objective of this project is to develop and implement a machine learning model that can detect brain tumors in MRI scans. A dataset from Kaggle with over 1000 MRI images was used to train and test a model. The images were categorized as either “yes” or “no” to indicate the presence of tumors. A convolutional neural network was selected as the method for developing a binary classifier. The network was composed of three 2D convolutional layers, three max pooling layers, two dense layers, as well as batch normalization and dropout layers. Input samples were augmented by rescaling, flipping, and rotating the images. Several strategies were used to avoid overfitting the model. The final version of the model achieved an accuracy of 65.9%. The model was compared to similar models that have been published online. The project provided an excellent opportunity to study applications of artificial intelligence through the development of a machine learning model to detect brain tumors in MRI scans.

Introduction

To provide overall context and reasoning for the project, background information and a project motivation are detailed below.

Background

Brain tumours contribute to a significant burden of disease. In 2021, it is estimated that 27 new primary brain tumours will be diagnosed every day in Canada [1]. Over 55,000 Canadians are currently surviving with a brain tumour and they are the leading cause of solid cancer death in youth [2]. A brain tumour is the growth of abnormal cells that are either within or around the brain. Because of their location, they often cause a negative impact on quality of life, including physical and cognitive impairment. Magnetic Resonance Imaging (MRI) is the most common test used to diagnose brain tumours [3]. MRI provides detailed images of the brain that allow accurate determination of lesion location, extent and its effect on surrounding tissues. The medical team depends on neuroimaging to characterize the tumour in order to determine treatment options.

Motivation

The motivation for this project is to develop and design a machine learning model for brain tumor detection. The model will process images of MRI scans and attempt to flag any discrepancies that may align with the known patterns that indicate brain tumors. This is an important problem to investigate, as the use of artificial intelligence and machine learning is becoming increasingly predominant in the medical field, assisting medical professionals to make informed decisions about their patients and recognize discrepancies that may not be noticed with the human eye [4].

Related Work

The advancements of machine learning in the past 10 years has prompted the medical community to integrate this technology into their field. Many scholarly articles have been written illustrating the use of machine learning to detect a wide range of abnormalities in the human body. Specifically, papers that explore machine learning applications to detect brain tumors are written weekly. For example, the articles titled “Design and Implementing Brain Tumor Detection Using Machine Learning Approach” [5], “Brain tumor detection using statistical and machine learning method” [6], and “A deep learning model integrating FCNNs and CRFs for brain tumour segmentation” [7] are all papers that dive into the application of machine learning to detecting brain tumors and will contribute to our project. Convolutional neural networks are commonly used to obtain segmentation results from 2D images. Numerous methods of image segmentation have been developed to identify brain tumours in MRI scans. Proceedings from the BraTS 2016 Challenge outline various novel strategies used in brain tumour segmentation [8].

Problem Formulation

The problem formulation of the project is outlined in this section to provide context to the specific problem and an overview of the proposed problem solving method.

Problem

Provided MRI scans of a patient's brain sections, medical professionals can perform close analysis of these images to recognize and confirm the presence of brain tumors, and perform a subsequent surgical procedure to remove these tumours. However, due to the tendency for tumor cells to infiltrate healthy tissues, surgeons may find it difficult to identify the border between tumor and healthy tissue with the naked eye during surgery. By applying artificial intelligence image processing technology, surgeons can quickly and clearly identify areas containing tumor or healthy tissue, where the tumor cells would otherwise be virtually invisible during surgery.

With this technology, a patient is more likely to have better treatment outcomes with fewer surgery-related complications, due to more healthy brain tissue being preserved during the procedure [9].

Data

The data for the project will consist of a collection of MRI brain scan images collected from a public Kaggle dataset [10]. This dataset contains over 1000 categorized jpg images of various brain MRI scans with and without tumor presence. While there are more comprehensive and complicated datasets, such as the BraTS challenge data [11], we believe that for the purpose and scope of our project, the size and simplicity of the mentioned dataset is perfectly appropriate.

Scope

For the purpose of the project, the images would be used to train our model over several iterations, then subsequently the accuracy of the model would be analyzed by determining whether or not it correctly classified the scans into one of two categories: has tumour present, or no tumour present.

Methodology

In this section, the methodology of our project and artificial intelligence model are discussed to help provide understanding of our model and the thought process behind our design.

In order to properly process our images, an approach of using a convolutional neural network was taken. A convolutional neural network is composed of multiple layers of artificial neurons. These artificial neurons are mathematical functions that calculate the weighted sum of multiple inputs and outputs and activation value. The behaviour of each neuron is defined by its weights. When a convolutional neural network is fed a variety of pixel values, the neurons of the network pick out various visual features.

When an image is provided to a convolutional neural network, each of the network's layers generate several activation maps. These activation maps highlight the relevant features of the image. Each neuron accepts a patch of pixels as input, multiplies their color values by its weights, sums them, then runs them through the activation function.

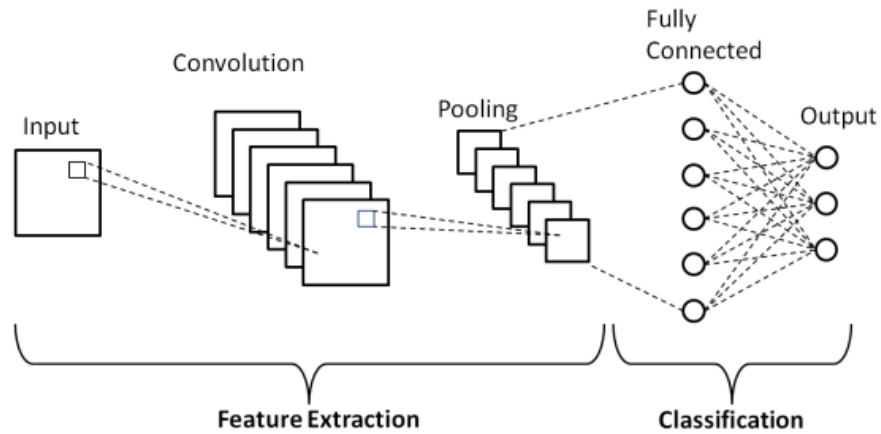


Figure 1: Example convolutional network

The first layer of the network typically detects basic features such as vertical and horizontal edges on the image. The output of the first layer is provided as the input to the next layer, which can extract more complex features and details. The details and level of features detected increase the deeper we move into the network [12].

To reduce the dimensions of the feature maps and reduce computations performed in the network, max pooling layers were used in between each convolutional network layer. Max pooling is a pooling operation that selects the maximum element from the region of the feature map encompassed by the filter. As a result, the output after the max pooling layer is a feature map containing the most prominent features of the previous feature map [13].

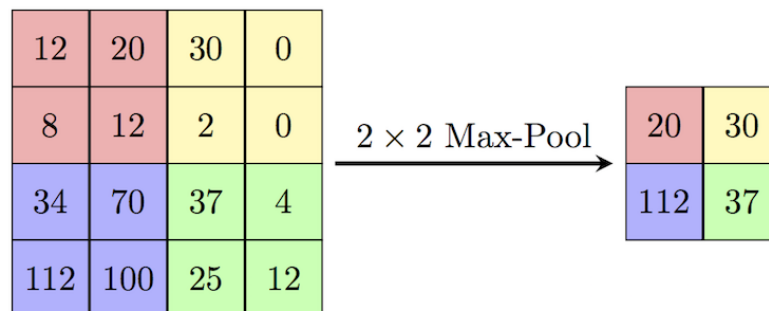


Figure 2: Max pooling example

While training the model, batch normalization and dropout layers were implemented to prevent overfitting, as shown in Figure 3 below. Batch normalization assists the model to train on more consistent values between batches. This allows the model to train on regularized data without constantly adjusting the weights due to abnormal values [14]. Dropout layers prevent overfitting by randomly ignoring select neurons when training the model on an input. This allows the model

to be less prone to overfitting because nearby neurons need to adjust to the sudden change in information [15].

Before being input to the model, each of the samples were pre-processed to crop the image, removing any unnecessary border outside of the primary brain shape. Next, the model underwent multiple layers of pre-processing. The first two layers were to resize the images from their original height and width to a standard 200x200 pixels and scale the pixel values to a range between 0 and 1. The next two layers were to introduce randomness of how the image would be presented to the model. This helped the model stay away from finding features in known areas of the images, but instead to curate its activation maps to potential features all around the images allowing it to be more flexible. After pre-processing, there were three sets of convolution, normalization, max pooling, and drop out layers. Lastly, a 128 neuron dense layer is used to curate the data from the previous layers, and a final layer with a single neuron, zero for negative and one for positive tumor is used to determine the final output.

Layer (type)	Output Shape	Param #
rescaling (Rescaling)	(None, 200, 200, 3)	0
random_rotation (RandomRotat	(None, 200, 200, 3)	0
random_flip (RandomFlip)	(None, 200, 200, 3)	0
conv2d (Conv2D)	(None, 198, 198, 16)	448
batch_normalization (BatchNo	(None, 198, 198, 16)	64
max_pooling2d (MaxPooling2D)	(None, 99, 99, 16)	0
dropout (Dropout)	(None, 99, 99, 16)	0
conv2d_1 (Conv2D)	(None, 97, 97, 32)	4640
batch_normalization_1 (Batch	(None, 97, 97, 32)	128
max_pooling2d_1 (MaxPooling2	(None, 48, 48, 32)	0
dropout_1 (Dropout)	(None, 48, 48, 32)	0
conv2d_2 (Conv2D)	(None, 46, 46, 64)	18496
batch_normalization_2 (Batch	(None, 46, 46, 64)	256
max_pooling2d_2 (MaxPooling2	(None, 23, 23, 64)	0
dropout_2 (Dropout)	(None, 23, 23, 64)	0
flatten (Flatten)	(None, 33856)	0
dense (Dense)	(None, 128)	4333696
batch_normalization_3 (Batch	(None, 128)	512
dropout_3 (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 1)	129

Figure 3: The team's convolutional network implementation

After initial efforts of trial and error while building and testing the model, as well as addressing overfitting issues, multiple tests were recorded in an effort to identify critical areas of improvement. The effects of pre-processing the data and varying relevant parameters were observed in order to obtain optimal values for the model. These test results are shown in Table 1 below.

EPOCHS	BATCH SIZE	CONV2D LAYERS	MAX POOLING LAYERS	DATA AUGMENTATION	TEST ACCURACY
10	10	1	1	no	0.491
15	10	1	1	no	0.490
20	10	1	1	no	0.493
10	20	1	1	no	0.495
15	20	1	1	no	0.501
20	20	1	1	no	0.504
10	32	1	1	no	0.506
15	32	1	1	no	0.505
20	32	1	1	no	0.515
10	32	2	2	no	0.525
10	32	3	3	no	0.526
10	32	2	2	yes	0.583
15	32	2	2	yes	0.584
15	32	3	3	yes	0.595
15	8	3	3	yes	0.598
20	8	3	3	yes	0.618
25	8	3	3	yes	0.659

Table 1: Variations of the model and results.

As seen in Table 1, the model's final testing accuracy was evaluated after it was compiled and trained on various combinations of different epochs, batch sizes, convolutional layers, max pooling layers, and pre-processing of data. The testing accuracy was calculated by running each combination of values over four separate training and testing phases, and subsequently averaging the testing accuracy. The model was determined to perform the best with 25 epochs, 3 convolutional layers paired with 3 max pooling layers, a batch size of 8, and when only the training data was pre-processed. This configuration reliably produced a testing error of 0.351, or about 65.9% accuracy when the model was run against testing data.

Results and Discussions

In this section, the results of our model and testing are discussed and explained with the help of graphs and images, with comparisons to other solutions. Justifications for our results are presented, and challenges are outlined.

Overall Results

With the model outlined in Figure 3 above, our team found that the average accuracy of the model on a set of unseen testing data was about 65% after training for 25 epochs with a final training loss of 0.21, a final training accuracy of 0.92, a final validation loss of 0.29, and a final validation accuracy of 0.89, as shown in Figure 4 and 5 below. The model was trained with a set of over 1000 images from a public data set. These images contained images with and without tumours, as well as images taken from various positions and angles around the head.

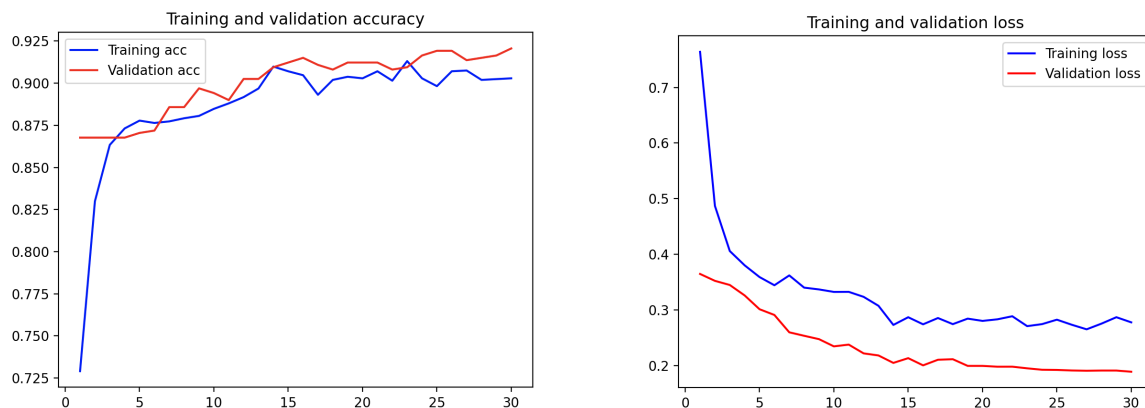


Figure 4: Loss and accuracy graph over each epoch.

As we can see from the results of our testing, we successfully trained and tested the model without much overfitting or other prominent issues. The model was reasonably successful in detecting brain tumours on unseen data considering the lack of image thresholding during the preprocessing phase.

Comparisons

CNN is a common method used for image classification and many models have been developed for brain tumor detection. Published models have reported testing accuracy as high as 94-97% [16][17][18]. Although the model in this report did not achieve the same level of accuracy, it shares many similarities with the better performing models. The related models utilize a similar layout containing various combinations of CNN layers. Data augmentation was also a common technique used to increase the size of datasets as well as avoid overfitting. One article outlines a

method for optimizing brain tumor classification using a multi-layered CNN [4]. The main techniques listed to enhance the models include increasing the size of datasets by creating synthetic images, concentrating multiple models into one, and utilizing strategies that avoid overfitting. The model in this report attempted to implement most of these features as the dataset went through a pre-processing phase that included image augmentation, and overfitting was avoided using batch normalization and dropout. Given higher quality training data and greater computational power, it is possible that variations of this model could achieve much higher accuracy. The BraTS 2016 proceedings outline multiple models that attempt to solve a similar problem [8]. The challenge varies slightly as the task involves segmentation and classification of 3D MRI brain samples. 3D volumetric CNN, decision forest, and random field are some of the techniques used to complete the task. A 3D volumetric CNN utilized a very similar approach to the one outlined in this report. Four 3D convolution layers, two dense layers, and two dropout layers were implemented to classify samples into five categories. The model achieved an average score of 0.725 for whole tumor segmentation. Typically, segmentation is achieved through analysis of 2D slices, similar to the samples in this report. This particular study showed that a similar technique can be expanded to a 3D volumetric CNN, given sufficient computing power. Considering all the studies mentioned, the developed model lacked the high accuracy of more complete CNNs, however, included many of the key features needed to build a successful version. The final variation provided reasonable accuracy with signs of promise for improvement given the proper resources.

Justification

The final model utilized CNN with three 2D convolution layers, three max pooling layers, batch size of 8, and 25 epochs. A CNN model was selected for this task because it is a relatively fast and efficient method of using machine learning for image classification [19]. Multiple convolutional layers were incorporated to extract key features from the input MRI images and a ReLU layer was used as an activation function on the feature map. A max pooling layer was applied after each convolution layer in order to reduce the size of the input while maintaining key information. Between these sets of layers, additional functions were incorporated in the CNN to avoid overfitting. Batch normalization and dropout layers were included after convolution and max pooling stages. By normalizing abnormal values and randomly reducing the number of active neurons, these layers were meant to help stabilize the training process and reduce the total number of epochs required. The final layers included two dense layers, with a sigmoid activation function used to determine the final output. Pre-processing was also implemented on the dataset prior to training. Multiple augmentation techniques were carried out on the dataset, such as rescaling, flipping, and rotating, in order to make the model flexible for various orientations of brain scans. Input samples were also shuffled to balance training between “yes” and “no” image classifications. Combining all these layers would provide a flexible framework for which a relatively reliable model could be trained to detect brain tumors in MRI scans.

Challenges

A few challenges were faced during the process of designing and testing our model.

Firstly, we found that there was an amount of data inconsistency in the data sets that we explored and used. Notably, we found that there were MRI scans that were taken from the side of the head, while the majority of the images were taken from an overhead view, providing a considerably different outline and overall shape of the head. In another dataset that was explored [20], we found that this was even more prominent, and we suspect it may have led to the observed decline in testing accuracy when running the model. In addition to this point, we had noticed that several of the testing and training images appear to have been categorized incorrectly, with images seeming to contain visible tumours appearing in the “no” category, and images that had no tumours appearing in the “yes” category. This may have also confused our model during the training process, as it learns based on this categorization.

Secondly, we struggled with pre-processing our images in a meaningful way. Our original goal was to compute a thresholding image, where the original grayscale image is segmented into a binary image of pixels given a 1/0 value based on their brightness in the original image. Thresholding an image provides the algorithm with an image that is easier to analyze, by creating clear, stark boundaries around features in images. Unfortunately, when we attempted to implement thresholding as part of our image pre-processing, we quickly discovered that it was difficult for the image libraries to threshold the MRI scans as the majority of the tumours did not have a significantly different shade or brightness when compared with the rest of the brain. Even worse, some tumours were significantly darker than their surrounding brain tissue, which resulted in the thresholding algorithm ignoring the tumour area entirely as it perceived it as a background. Often, the images were too sensitive or not sensitive enough to the thresholding process, which we can attribute to the different brightnesses of images and different variations of tumours in the dataset. Even when thresholded with what appeared to be reasonable results, the accuracy of the model when ran against testing data was not as accurate compared to testing without thresholded images.

Finally, one of the main challenges we faced was overfitting of the model. Overfitting is a common phenomenon in artificial intelligence and machine learning when the model learns the detail and noise in the training data to the extent that it negatively impacts the performance of the model on new data. Since our model was suffering from overfitting, we were able to employ the Keras data augmentation library to create additional modified and synthetic data from our pre-existing dataset, which helped us find a correct fit for our model, eliminating the issue of overfitting.

Conclusion

This project aimed to apply artificial intelligence towards a biomedical application in the form of detecting brain tumors in MRI scans. A dataset containing over 1000 MRI images was used to train and test a machine learning model. Brain scans were labelled as either “yes” or “no” depending on whether or not a tumor was present. A CNN was selected as the framework for developing a binary classifier, due to its effectiveness and efficiency with image recognition. The network was composed of three 2D convolutional layers, three max pooling layers, two dense layers, as well as batch normalization and dropout layers. Pre-processing was also applied to input samples which included rescaling, flipping, and rotating images, in order to increase variability. The final version of the model achieved an accuracy of 65.9% and was characterized by a batch size of 8, 25 epochs, and data augmentation. Significant improvements were obtained between initial and final model variations by adjusting the CNN layers and implementing strategies to avoid overfitting. Although there are better performing models that exist for this particular application, the project was successful in developing and applying a binary classification model which demonstrated the benefits of AI and displayed multiple machine learning techniques.

Future Work

Due to the rather low accuracy of the model when it is tasked with sorting new data, we identified several areas for improvement if the model were to be developed and tested further.

Firstly, we believe access to a more comprehensive and detailed data set would help with both the training and testing process. The standard for MRI brain scan images seems to be the Brain Tumour Segmentation (BraTS) challenge [8]. The yearly BraTS challenge is presented by the Center for Biomedical Image Computing & Analytics by the University of Pennsylvania. The BraTS dataset is a very comprehensive set of multi-parametric MRI (mpMRI) scans of glioma with pathologically confirmed diagnosis routinely clinically-acquired from a range of different institutes [21]. While more comprehensive, the data is not immediately publicly accessible, and is presented in a more complicated fashion that was beyond the scope of this project.

Additionally, as we struggled with properly creating thresholded images, we believe an investigation into other methods of image pre-processing and thresholding would be beneficial to the overall accuracy of the model. Investing into research of thresholding algorithms and approaches would allow for a better understanding of different thresholding methods such as clustering methods, entropy methods, spatial methods, and local methods. With more testing and fine tuning, the addition of a proper image thresholding step would likely lead to a higher testing accuracy.

References

- [1] “Fact Sheet,” Brain Tumour Registry, 10-Mar-2020. [Online]. Available: <https://braintumourregistry.ca/fact-sheet/>. [Accessed: 17-May-2021].
- [2] “Facts About Brain Tumours,” Brain Tumour Foundation of Canada, 06-Dec-2019. [Online]. Available: <https://www.braintumour.ca/facing-a-brain-tumour/facts-about-brain-tumours/>. [Accessed: 17-May-2021].
- [3] L. D. H. X. M. T. N. LB; “Role of MRI in primary brain tumor evaluation,” Journal of the National Comprehensive Cancer Network : JNCCN. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/25361803/>. [Accessed: 17-May-2021].
- [4] S. Kulkarni, N. Seneviratne, M. S. Baig, and A. H. A. Khan, “Artificial Intelligence in Medicine: Where Are We Now?,” Academic Radiology, 19-Oct-2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1076633219304581?casa_token=Gar58-POnpoAAAAA%3A0cxjNezys8kYETSHBiALbbv2sqX0u5gLS95FTB7zE0CTU5wcu_I_0Of-tgG8009yfXzP_Y7WInvH. [Accessed: 17-May-2021].
- [5] “Design and Implementing Brain Tumor Detection Using Machine Learning Approach,” IEEE Xplore. [Online]. Available: <https://ieeexplore.ieee.org/document/8862553>. [Accessed: 17-May-2021].
- [6] J. Amin, M. Sharif, M. Raza, T. Saba, and M. A. Anjum, “Brain tumor detection using statistical and machine learning method,” Computer Methods and Programs in Biomedicine, 17-May-2019. [Online]. Available: <https://www.sciencedirect.com/science/article/abs/pii/S0169260718313786>. [Accessed: 17-May-2021].
- [7] X. Zhao, Y. Wu, G. Song, Z. Li, Y. Zhang, and Y. Fan, “A deep learning model integrating FCNNs and CRFs for brain tumor segmentation,” Medical Image Analysis, 05-Oct-2017. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S136184151730141X>. [Accessed: 17-May-2021].
- [8] “BraTS 2016 - MICCAI BRATS 2017,” Google Sites. [Online]. Available: https://sites.google.com/site/braintumorsegmentation/home/brats_2016. [Accessed: 17-May-2021].

- [9] 2021 May 27, 2021 May 20, and 2021 April 22, “Artificial Intelligence Expedites Brain Tumor Diagnosis,” National Cancer Institute. [Online]. Available: <https://www.cancer.gov/news-events/cancer-currents-blog/2020/artificial-intelligence-brain-tumor-diagnosis-surgery>. [Accessed: 08-Jul-2021].
- [10] Sartaj, “Brain Tumor Classification (MRI),” *Kaggle*, 24-May-2020. [Online]. Available: <https://www.kaggle.com/sartajbhuvaji/brain-tumor-classification-mri>. [Accessed: 15-Jul-2021].
- [11] “BraTS 2016 - MICCAI BRATS 2017,” Google Sites. [Online]. Available: https://sites.google.com/site/braintumorsegmentation/home/brats_2016. [Accessed: 17-May-2021].
- [12] B. D. -, “What are convolutional neural networks (CNN)?,” TechTalks, 06-Jan-2020. [Online]. Available: <https://bdtechtalks.com/2020/01/06/convolutional-neural-networks-cnn-convnets/>. [Accessed: 09-Jul-2021].
- [13] K. Patel, “Convolution Neural Networks-A Beginner's Guide [Implementing a MNIST Hand-written Digit...,” Medium, 18-Oct-2020. [Online]. Available: <https://towardsdatascience.com/convolution-neural-networks-a-beginners-guide-implementing-a-mnist-hand-written-digit-8aa60330d022>. [Accessed: 09-Jul-2021].
- [14] J. Brownlee, “A gentle introduction to batch normalization for deep neural networks,” *Machine Learning Mastery*, 03-Dec-2019. [Online]. Available: <https://machinelearningmastery.com/batch-normalization-for-training-of-deep-neural-networks/>. [Accessed: 28-Jul-2021].
- [15] J. Brownlee, “Dropout regularization in deep learning models with keras,” *Machine Learning Mastery*, 26-Aug-2020. [Online]. Available: <https://machinelearningmastery.com/dropout-regularization-deep-learning-models-keras/>. [Accessed: 28-Jul-2021].
- [16] “Brain tumor classification using convolutional neural network,” *IEEE Xplore*. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/8934603>. [Accessed: 28-Jul-2021].
- [17] M. M. Badža and M. Č. Barjaktarović, “Directory of open access journals,” *Applied Sciences*, 01-Mar-2020. [Online]. Available: <https://doaj.org/article/5b1a4e0dfb5743e59f03717d043b38ee>. [Accessed: 28-Jul-2021].

[18] R. Anitha and D. S. S. Raja, "Segmentation of glioma tumors using convolutional neural networks," *Wiley Online Library*, 20-Nov-2017. [Online]. Available: <https://onlinelibrary.wiley.com/doi/abs/10.1002/ima.22238>. [Accessed: 28-Jul-2021].

[19] A. Bonner, "The complete beginner's guide to deep Learning: Convolutional neural networks," *Medium*, 01-Jun-2019. [Online]. Available: <https://towardsdatascience.com/wtf-is-image-classification-8e78a8235acb>. [Accessed: 28-Jul-2021].

[20] J. Brownlee, "A gentle introduction to batch normalization for deep neural networks," *Machine Learning Mastery*, 03-Dec-2019. [Online]. Available: <https://machinelearningmastery.com/batch-normalization-for-training-of-deep-neural-networks/>. [Accessed: 28-Jul-2021].

[21] "RSNA/ASNR/MICCAI Brain Tumor Segmentation (BraTS) Challenge 2021," *RSNA/ASNR/MICCAI Brain Tumor Segmentation (BraTS) Challenge 2021 | CBICA | Perelman School of Medicine at the University of Pennsylvania*. [Online]. Available: <https://www.med.upenn.edu/cbica/brats2021/>. [Accessed: 15-Jul-2021].