International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
IJISAE
ISSN:2147-6799 www.ijisae.org Original Research Paper

A Novel Mask R-CNN based Approach to Brain Tumour Detection

Avishek Mondal[1], Arnab Sardar[2], Rohini Basak*[3], Sourav Mandal [*4]

Submitted: 22/**07**/**2022** **Accepted:** 25/**09**/**2022**

Abstract: When abnormal cells form in the brain, it is called a brain tumour. The unique blocking and proliferation of abnormal tissues in the brain can be found using magnetic resonance imaging (MRI) imaging of the brain. Due to advancements in medical imaging technology for auto detecting equipment, it is no longer necessary for a doctor to look at MRI scans to determine whether a patient has a tumour. As a result, it has proven handy for patients who do not wish to see a doctor immediately. Our study describes a method for segmenting abnormal brain tissues and determining whether the patient has a tumour. This approach detects a unique area of the brain and forecasts the likelihood of a tumour developing there. Mask regional-based convolutional neural network (Mask R-CNN) is a pre-trained deep neural network model that is used to distinguish objects from an image such as cars, animals, trees, and other objects. In comparison to many other similar methods based on MLP, VGG-16 model, and U-net model, we discovered that Mask R-CNN method performs the best. The clarity of the MRI scans has a big impact on the accuracy. The proposed system was able to outperform similar systems on the same dataset, achieving 74 percent *Intersection over Union* (IoU) score on the reference dataset, *Brain MRI Images for Brain Tumour Detection.*

Keywords: Brain tumour detection, Medical imaging, Image segmentation, Convolutional neural network, Mask R-CNN.

1. Introduction

According to [1], there are two types of brain tumours: malignant (cancerous) and benign (non-cancerous) (non-cancerous). Symptoms include migraines, convulsions, low vision, vomiting, and so on. According to the World Health Organization (WHO) [2] brain tumour categorization is mostly based on histogenesis notions, which state that tissues are commonly classified based on their little resemblance to cells of various origins and their levels of differentiation. The light-reflecting properties of hematoxylin and eosin-colored phases, as well as the immune-histochemical expression of lineage proteins linked with structure creation, were all important in the formation of such histological commonalities. It uses Medical Imaging [3] techniques and data gathered from several biological engineering technologies to get the diagnosis. Imaging procedures include MRI scans, CT scans, and X-rays, among others.

Proper medical image processing is essential to determine whether someone has a brain tumour. The best technique to get a proper brain scan is with an MRI. A large magnet and radio waves are used in magnetic resonance imaging (MRI) [4] to examine at structures and organs within a living body. Healthcare experts utilize MRI images to identify a wide range of ailments, from tissue to injured muscles. In spinal cord and brain testing, MRIs are quite useful. As a result, a good MRI scan is essential for brain tumour classification. It is necessary to detect and study brain tumours utilizing a machine learning-based approach for awareness and analysis.

The following are some of the identified obstacles and limits [5] of deep machine learning-based techniques.

- Noise reduction is a difficult task [6-8].
- Accurate segmentation is a challenging task [9].
- Tentacles and dispersed structures are found in brain tumours [10-12].
- For improved classification, selecting and extracting suitable features and enough training/testing samples is also critical [13].

The process of identifying a brain tumour entail detecting a brain and looking for strange things on it, such as aberrant tissue growth. It is frequently utilized in the realm of medicine. This work uses Mask R-CNN to provide a novel deep learning-based solution to serve the above objective. An MR imaging of the brain is used to detect tumour tissue in the case of brain tumour detection. The task is to determine whether or not there is a tumour in an MRI image of the brain. If there is one, the shape and size of the tumour tissue must be identified as well. To classify the brain tumour and determine the findings, we employed a Convolutional Neural Network model whose fundamental

[1]*Dept. of Information Technology, Jadavpur University, Kolkata-700032, West Bengal, India. ORCID ID: 0000-0002-6288-0360*
[2]*Dept. of Information Technology, Jadavpur University, Kolkata-700032, West Bengal, India. ORCID ID: 0000-0001-9881-3496*
[3]*Dept. of Information Technology, Jadavpur University, Kolkata-700032, West Bengal, India. ORCID ID: 0000-0001-9662-3074*
[4]*School of Computer Science and Engineering (SCSE), XIM University, Harirajpur-752050, Odisha, India. ORCID ID: 0000-0002-6066-8008*
** Corresponding Author Email: visitrohinihere@gmail.com, sourav_officials@yahoo.co.in*

structure is depicted in Fig. 1, specifically a Mask R-CNN [14] model. From images of MRI scans of the brain, the Mask R-CNN model can determine if a tissue is a tumour or not.

This is an excellent piece of work with the potential to influence current research in this field. Some of our specific contributions to this project are listed below.

- The Mask R-CNN technology was applied for the first time to detect brain tumours.
- Annotated and modified "Brain MRI Images for Brain Tumor Detection" to create a new dataset.
- The model can detect partially occluded brain tumours and pinpoint the exact location of the tumour item, as well as provide a visual depiction of the mask position in relation to the tumour.

The remainder of the paper is written as follows. We discuss several related papers on brain tumour detection using various approaches in Section 2 with critical analysis. We presented the dataset we used for this project, as well as several DL-based tools for implementation. The proposed Mask R-CNN based methodology is demonstrated in Section 4 with detailed descriptions of each component. We examine our system's performance in Section 5 with results and a comparative analysis, and then we conclude in Section 6.

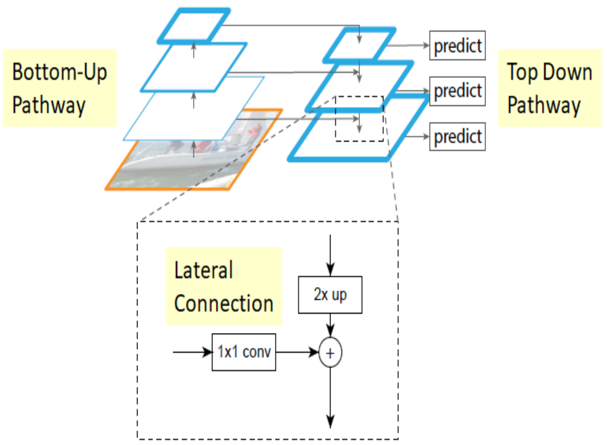

Fig. 1. Structure of a Mask R-CNN model [5, 42]

2. Related Work

In this section, we'll go over some of the related works that have similar goals. In the instance disjunction of the model, the mask R-CNN has become the new contemporary. Machine learning uses Mask R-CNN technology to extract objects from images or videos. It is a more accurate variation of Faster R-CNN [15], in which Faster R-CNN provides bounding box-based segmentation whereas Mask R-CNN performs pixel level segmentation, resulting in more accurate and efficient segmentation. We begin by providing a masked input image, with the masked object to be detected as a separate entity.

Fast R-CNN: Let's start with a quick overview of the Faster R-CNN identifier [16]. There are two sections to Fast R-CNN. The Regional Proposal Network (RPN) is the initial part, and it proposes a set of element binding boxes. The platform, which is essentially Fast R-CNN [17], dispatches feature that employ RoIPool in each child box and executes binding box separation and reversal. Prompt acquisition could be used to distribute features that employ both parts of it. The youngsters are then directed to [18] for up to date, completed comparisons of various methods and Faster R-CNN.

Mask R-CNN: Mask R-CNN uses two similar types of processes, with the initial phase differing slightly (RPN). Mask R-CNN and gives every binary mask RoI in the second component, which is identical to box prediction offset and object. It will then be compared to more recent systems, and the mask prediction will be used to determine the distinguishing properties (e.g. [19, 20, 21]). This approach follows the Fast R-CNN [22] morals, which does binding box division as well as reverse lateral division (this is easy to do in several steps in actual R-CNN workflow [23]).

Then it goes through two steps of processing. First, it suggests where an object might be based on the supplied image. Then, based on the first phase of the proposal, it forecasts the class of that object, refines the enclosing object, and creates a mask at the pixel level of the object recognized as a different business.

Based on Feature Pyramid Networks [24] of Object Detection, these two categories or layers are related to the core. A bottom-up line, a top-down method, and a lateral link are all included. Convolutional neural networks are used in the sub-top route [25]. This bottom-up approach retrieves characteristics from raw photos directly. The top-to-bottom pathway is made up of feature pyramids, whereas the lateral linkages are made up of convolution and addition operations between two related paths. Feature Pyramid Networks, on the other hand, outperform ConvNets because they keep significant semantic features at different resolution scales.

Previous research studies [26] reveal that Convolutional Neural Networks [27] can detect brain tumours and that they can be quite accurate. However, it can only detect whether there is a tumour or not; it cannot predict the location or coordinates where the tumour is present. However, we may find the tumour in a bounded box using the Faster RCNN [28] approach, and it will also supply the coordinates of those bounding boxes to be drawn in the final image. However, in this study, we will discuss how Mask R-CNN detects the tumour pixel by pixel, which will aid medical specialists in determining the tumor's intensity. Because different shapes of tumours indicate different types of cancers [29], this method can be used to anticipate the exact location of the resident tumour and to draw or highlight the area. Mask R-CNN will return the object mask in addition to the class label and the enclosing box object on each item using the image provided. As a result, Mask R-CNN is the most accurate method for detecting malignancies in the brain, because each pixel in a brain MRI picture matters when detecting a tumour. As a result, this problem can be solved using Mask R-CNN. In the following sections, we'll go through the dataset we used and the software we needed to carry out our study approach and goals.

3. Dataset and Packages

For this procedure, only one dataset has been used. The dataset [30] has 253 images, with 155 MRI images containing brain cancers (sample shown in Fig. 2) and the remaining 98 photos (sample shown in Fig. 3) containing MRI images without brain tumours. These are the MRI pictures that have been processed to determine if there are any cancers present.

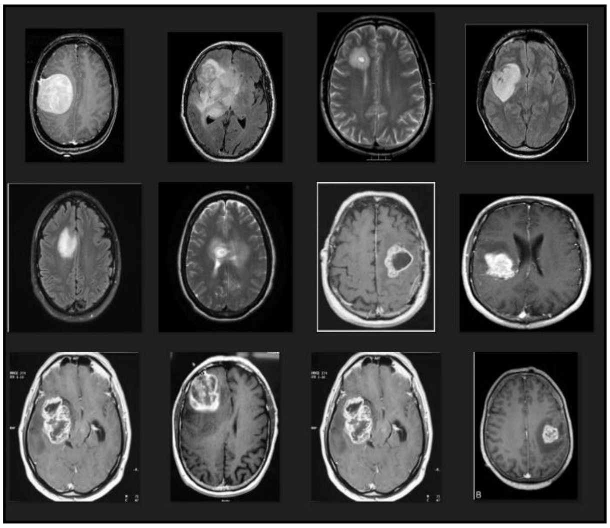

Fig. 2. Images with Tumors [30]

The irregular white segments observed in Fig. 2 suggest the presence of a tumour in the brain. Our objective now is to use Mask R-CNN to study the irregular white section and identify whether it has a tumour or not.

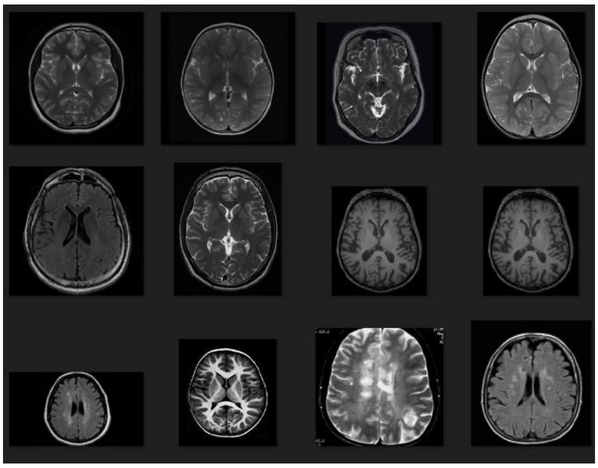

Fig. 3. Images without Tumors [30]

TensorFlow [31] is an open-source deep learning framework that can be used for fault detection, speech recognition, computer vision, sentiment analysis, geographic information extraction, text summarization, computational drug development, and information retrieval, among other things. This was utilized to apply the proposed Mask-RCNN model in a variety of our tasks, with a focus on deep neural network training and inference [32]. Keras is an open-source library that operates on top of TesnsorFlow and is used for deep artificial neural networks. It takes full advantage of Tensor-Flow cross-thickness platforms and strength. Layers and models are the foundational data structures in Keras [33]. Keras is used inside the Mask-RCNN model as a TensorFlow library interface in the proposed method. OpenCV is a real-time computer vision library that contains a variety of capabilities such as 2D and 3D feature extraction toolkits, facial recognition systems, gesture recognition systems, and much more [34]. The suggested solution makes advantage of OpenCV's image processing features, which will be explored in more detail in the data processing section.

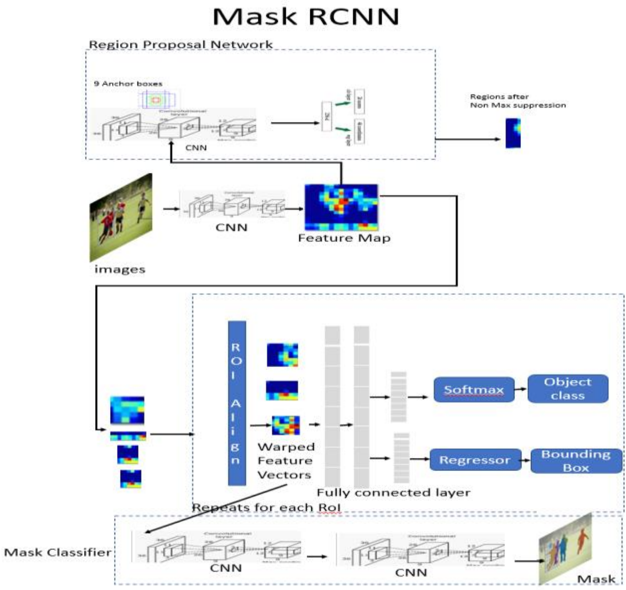

Fig. 4. Working of Mask R-CNN [40, 41]

4. Proposed Method

This section delves into the specifics of our proposed approaches. The proposed system includes a mask classifier that was trained on our proprietary dataset of brain MRI images, as well as a masked coordinate of polygons that defines where the tumour is on the x-y axis. The Mask R-CNN model is first configured using the config instance we established during training, and then we load the pre-trained weights for the Mask R-CNN from the COCO dataset [35], excluding the last few layers. COCO is a dataset that includes object detection, segmentation, and captioning on a huge scale. COCO has several characteristics. We don't need to train for a long period because we're not using a large dataset and started with COCO trained weights. It is also not necessary to train all the layers; simply training the head layer accomplished our purpose [35]. Fig. 4 depicts the proposed architecture of a Mask R-CNN based system. The following sub-sections go through the specifics of each component and how it operates.

4.1 Data Pre-processing

Data pre-processing is the process of preparing raw data for use in a machine learning model. To develop a machine learning model, this is a critical stage. Tables, photos, videos, graphs, and other formats of data are all acceptable. The quality of the data, as well as the information that can be obtained from the data, affects the ability of a machine learning model. As a result, preprocessing data before using it to train a machine learning model is critical.

The steps of data preparation differ depending on the type of problem we're attempting to address using the data. Although some common methods of data pre-processing include collecting data, cleaning data, checking for NULL values, checking for categorical data, standardizing data, and splitting data into different sets for testing, training, and validation. We solved a classification challenge in this example. As a result, the processes to pre-process the photographs should be tailored to this type. The pre-processing includes the stages listed below because this proposed method deals with MRI images as well as the tumour coordinates on the x-y plane.

4.1.1. Conversion of Image Colour

The conversion of the image's colour from RGB to grayscale is a crucial step. When it comes to picture classification for machine learning, grayscale photos are always easier to deal with than RGB images. RGB photos may contain additional information that is important for solving other issues. However, RGB information is not necessary for this problem because the Mask R-CNN model and the bulk of image classification methods assume images to be grayscale. As a result, transforming the photos to grayscale not only reduces code complexity, but it also improves the performance of the machine learning model by eliminating the need to calculate the extra RGB bytes. Although some MRI images appear to be grayscale, this is not the case for all MRI images. As a result, before proceeding to the next stage, the colours of the MRI pictures must be converted to grayscale. We utilized the OpenCV method "cv2.cvtColor(img, cv2.COLOR RGB2GRAY)" for this case. Fig. 5 shows an example.

Fig. 5. An example of a colour to black and white conversion

Only changing the image's colour isn't enough; some blur needs also be added to the MRI image to ensure that it has extremely little noise. To blur an image, Gaussian blur is utilized. We used the OpenCV method "cv2.GaussianBlur(gray, (5, 5), 0)" to solve this problem. Fig. 6 depicts one example.

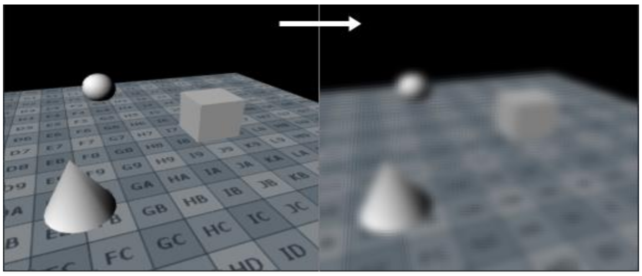

Fig. 6. An image that has been blurred as an example

Thresholding is also essential for the pre-processing of MRI pictures, as it ensures that the tumours in the MRI images are brighter than the background. For this challenge, we utilized OpenCV's "cv2.threshold(gray, 45, 255, cv2.THRESH BINARY)" technique. Fig. 7 shows one example.

Fig. 7. An image thresholding example

4.1.2. Removing Noise

On the threshold image, a succession of erosion and dilation must also be applied. Any minor noise patches are removed by erosion and dilatation. For this task, the OpenCV library functions cv2.erode() and cv2.dilate() are utilized with two rounds of each for erosion and dilation.

4.1.3. Finding Contours

Finding the contours of the brain's parameters and cropping the image based on the maximum values of the brain's contours are required for fine tuning the MRI images. Finding the contours of the brain's total parameter reduces the image's size while also assisting the machine learning model in obtaining a more accurate image for training. As a result, it is no longer required to analyze components that are outside the boundaries of the brain. An example can be shown in Fig. 8.

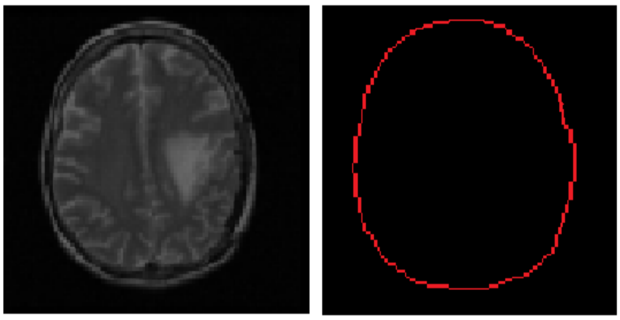

Fig. 8. An image of a brain's contours

4.1.4. Resizing the image

One of the most critical processes in pre-processing is resizing images. The length of time it takes to train a machine learning model is proportional to the size of the training image. The training step takes substantially longer to complete as the image size grows greater. This resizing may alter the quality of the results depending on the solution's requirements, but in this model, a larger image size merely increases the training time because we don't need to worry about how much more information and details we can pull out of a larger image. As a result, each image is reduced to **256 pixels by 256 pixels**.

The model must be trained with a given dataset and tested on a different dataset once the MRI pictures have been resized. Appropriate model and classification of prepared train tests aid in the production of reliable predictions. A total of 80% of the images in the dataset are trained, 10% are validated, and the remaining 10% are tested. As a result, we split photos into 8:1:1 ratio and saved them in their corresponding test, train, and validation folders.

4.1.5. Masking

To mask all the cleaned photos, **LabelMe** [36] is used. It creates a file with data, x-y coordinates of a polygon that represents the custom masked tumours, and other information. LabelMe is a computer program that allows us to label and categorize any image. It employs the Python-based Image **Polygonal Annotation** technology. Using LabelMe, we manually masked the tumour area of the brain and saved the data for model training. Fig. 9 and 10 show how to label a tumour with LabelMe software and give it the name tumour.

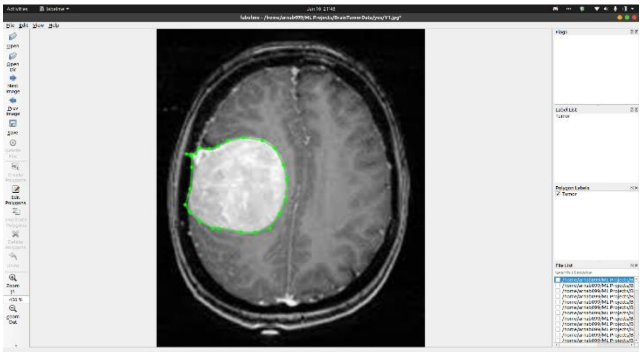

Fig. 9. Labeling the tumor

Fig. 10. Naming the label as Tumor

As we can see we plot the tumor in the image, LabelMe saves it by coordinates as a polygon. Then it creates a **JSON** (JavaScript Object Notation) file where all the coordinates of the polygon which contains the tumor is saved which is used to train our model. Sample data is shown in Fig. 11.


```
{
  "version": "3.16.7",
  "flags": {},
  "shapes": [
    {
      "label": "Tumor",
      "line_color": null,
      "fill_color": null,
      "points": [[28.77, 81.55], [35.66, 82.0], ... ,[26.33, 85.33]]],
      "shape_type": "polygon",
      "flags": {}
    }
  ],
  "lineColor": [0, 255, 0, 128],
  "fillColor": [255, 0, 0, 128],
  "imagePath": "Y1.jpg",
  "imageData": <Binary Data of the Image>,
  "imageHeight": 218,
  "imageWidth": 180
}
```


Fig. 11. Data of the coordinates of the polygon

4.2. Training of Model

4.2.1. Mask R-CNN and brain MRI dataset

Mask RCNN is available online[1]. All the repositories in the model and dataset directories have been cloned at this point. Then we enter the data we retrieved with LabelMe, as well as the coordinates of the tumour polygons in the training and validation sets. This information was saved in the annotation's directory.

4.2.2. Configuring the brain tumour dataset's training settings

We now set up the configurations, which include properties such as the number of Graphics processing unit(s) or GPUs to use, as well as the number of images per GPU (we only used one GPU), and the number of classes, which is set to two, one for the backdrop and the other for the tumour. We've set the number of training steps per epoch to 100, the learning rate to 0.001, and the confidence level for skip detection to be less than 85 percent. As a result, these are the model-training setups for the brain tumour dataset.

4.2.3. Creating the custom data set

We go through all the picture files in the datasets to add new annotations or labels, resulting in a dataset that includes each image's output label as a binary class tumour or not. From the annotation file, we now extract each of the bounding boxes. It holds the box's information, such as its height and breadth. The masks for each object in the image are then generated from the log data. Returns a single mask, as well as category ids and a 1D list of class ids for that specific instance of the mask. We've now created our own

[1] https://github.com/matterport/Mask RCNN.git

dataset, and our model is ready to be trained.

4.2.4. Initializing the Mask R-CNN model for training

We must utilize an instance of the Config that was created before to initialize the Mask R-CNN. The weights of the previously trained Mask R-CNN model are then downloaded from the COCO Dataset [27], except the final couple of layers. We're working with a tiny dataset here, and the model doesn't need to be trained for as long thanks to COCO trained instruments. The head layers of Mask-RCNN should also enough for this training data; no more than the head layer is required, therefore **ResNet101's** last few layers are omitted for this model's training. Now we load the dataset and train the model with 15 epochs at 0.001 learning rate.

5. Result And Analysis

5.1. System performance and evaluation

The loss function can be used to validate the results of fine-tuning the model. For optimization, the model employs the Binary-Cross entropy loss function. The log mask is adjusted for each erroneous prediction. Mask R-loss CNN's function includes classification, localization, and segmentation losses. Average Precision is a popular metric for evaluating acquisition items such Mask R-CNN and Faster R-CNN. To calculate the recall and precision for an object detector, another factor known as the Intersection over Union (IoU) score [37] and the ratio of overlap and union of the ground truth and predicted mask is taken into consideration. As a result, an IoU score larger than a preset threshold is required to identify a prediction as right or true positive, and an IoU score of 50% is regarded good. On the test dataset, this model's average IoU score is 0.74, which is close to 75 percent, indicating that it performs admirably. The model can detect partially blocked brain tumours and pinpointing the tumour object's location.

5.2. Visualizing the results

The function "display differences" in the Mask R-CNN implementation plots the Mask, Polygon result on the provided picture. The results of the Mask R-CNN model are then fed into the function, and a visual representation of the mask position, where the tumour is located, can be seen on top of the brain MRI image, as shown in Fig. 12 and Fig. 13, where Fig. 12 is fed into the Mask R-CNN model as input image and Fig. 13 is generated as result image. The Training and Validation loss curves can be seen in Fig. 14's Epoch vs. Loss graph, while the Training and Validation accuracy can be seen in Fig. 15's Epoch vs. Accuracy graph.

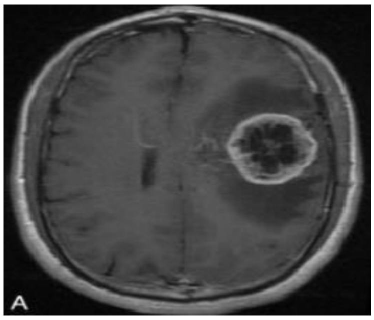

Fig. 12. Input Image

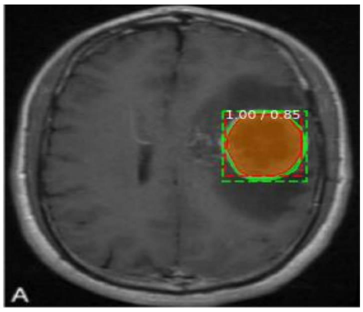

Fig. 13. Resultant Image

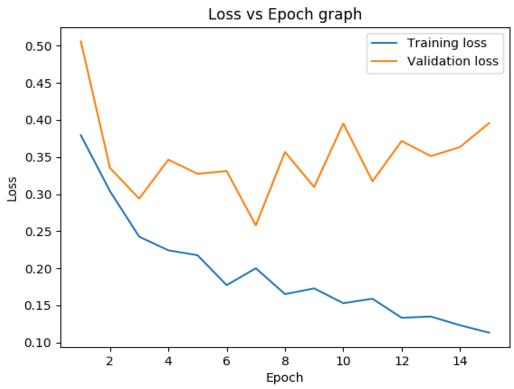

Fig. 14. Epoch vs. Loss

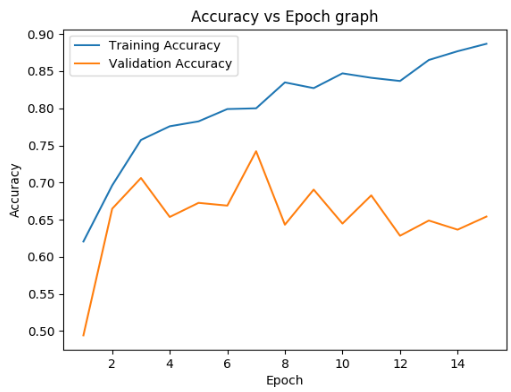

Fig. 15. Epoch vs. Accuracy

Other similar techniques used deep neural networks to detect tumours [38] and image-based breast cancer diagnosis [39]. Other than [20], we haven't found any other system that has worked on a same dataset to compare accuracy and system performance with. However, in terms of the final IoU-based score, our system exceeded [20].

6. Conclusion and Future Work

In this paper, we will briefly discuss brain tumours, their classifications, characteristics, and how to detect them. We discovered that using MRI scanned images and then applying Mask R-CNN to detect the tumour in them is the best technique to predict brain malignancies. To implement our proposed solution, we used the Mask R-CNN model, as well as different deep learning algorithms and certain necessary pre-processing processes. The approach has a 74 percent accuracy rate in predicting the tumour and its location from a brain MRI image. This method is highly reliant on the quality of the brain's MRI image; therefore, if the image is incorrect, blurred, blackened, overexposed to light, low brightness, or incomplete, this method will not produce accurate findings. Completing all these stages, as well as having a good machine that can conduct all these activities efficiently, takes a long time. Even if all those prerequisites are met, it will still take a long time. This method can only be used for report generation and automation. In comparison to other medical professions, this method demonstrates and opens a large chance for deep learning to have a significant impact on radiology and medical imaging. The proposed approach might be expanded to include a vast collection of MRI images of brains with and without tumours, potentially improving accuracy. The proposed systems can be tested with a higher number of epochs to train the model; however, one of the primary challenges with such research is the lack of system resources. Additional hyperparameter tunings and the application of the attention mechanism with certain fresh features may help the system perform better.

7. References and Footnotes

Author contributions

Avishek Mondal: Implementation and coding, Field study **Arnab Sardar:** Implementation and coding, Methodology, Field study. **Rohini Basak:** Conceptualization, Methodology, Data curation, Writing-Reviewing and Editing, Validation. **Sourav Mandal:** Conceptualization, Methodology, Investigation, Writing-Reviewing and Editing.

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] The Wikipedia page about “Brain tumor”, available: https://en.wikipedia.org/wiki/Brain_tumor

[2] “The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary”, Springer-Verlag Berlin Heidelberg, 9 February 2016, available: http://braintumor.org/wp-content/assets/WHO-Central-Nervous-System-Tumor-Classification.pdf

[3] “Medical imaging”, from wikipedia, available: https://en.wikipedia.org/wiki/Medical_imaging

[4] The Wikipedia page about “Magnetic resonance imaging”, available: https://en.wikipedia.org/wiki/Magnetic_resonance_imaging

[5] Amin, Javaria, Muhammad Sharif, Anandakumar Haldorai, Mussarat Yasmin, and Ramesh Sundar Nayak. Brain tumor detection and classification using machine learning: a comprehensive survey. Complex Intell. Syst. (2021). https://doi.org/10.1007/s40747-021-00563-y

[6] Khan MA, Lali IU, Rehman A, Ishaq M, Sharif M, Saba T et al (2019) Brain tumor detection and classification: A framework of marker-based watershed algorithm and multilevel priority features selection. Microsc Res Tech 82:909–922

[7] Agrawal U, Brown EN, Lewis LD (2020) Model-based physiological noise removal in fast fMRI. Neuroimage 205:1–18

[8] Malla, S., M. J. . Meena, O. . Reddy. R, V. . Mahalakshmi, and A. . Balobaid. “A Study on Fish Classification Techniques Using Convolutional Neural Networks on Highly Challenged Underwater Images”. International Journal on Recent and Innovation Trends in Computing and Communication, vol. 10, no. 4, Apr. 2022, pp. 01-09, doi:10.17762/ijritcc.v10i4.5524.

[9] Dubey YK, Mushrif MM (2016) FCM clustering algorithms for segmentation of brain MR images. Adv Fuzzy Syst 2016:8

[10] Irshad M, Muhammad N, Sharif M, Yasmeen M (2018) Automatic segmentation of the left ventricle in a cardiac MR short axis image using blind morphological operation. Eur Phys J Plus 133:1–14

[11] Rewari R (2021) Automatic tumor segmentation from MRI scans. Stanford University, Stanford

[12] Havaei M, Davy A,Warde-Farley D, Biard A, Courville A, Bengio Y, Pal C, Jodoin PM, Larochelle H (2017) Brain tumor segmentation with deep neural networks. Med Image Anal 35:18–31

[13] Jiang J, Wu Y, Huang M, Yang W, Chen W, Feng Q (2013) Brain tumor segmentation in multimodal MR images based on learning population-and patient-specific feature sets. Comput Med Imaging Graph 37:512–521

[14] Ortiz A, Gorriz JM, Ramírez J, Salas-Gonzalez D, Alzheimer’s Disease Neuroimaging Initiative (2013) Improving MRI segmentation with probabilistic GHSOM and multiobjective optimization 114:118–131

[15] He, Kaiming, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. "Mask r-cnn." In Proceedings of the IEEE international conference on computer vision, pp. 2961-2969. 2017.available: https://arxiv.org/abs/1703.06870

[16] Girshick, Ross. "Fast r-cnn." In Proceedings of the IEEE international conference on computer vision, pp. 1440-1448. 2015., available: https://arxiv.org/abs/1504.08083

[17] Ren, Shaoqing, Kaiming He, Ross Girshick, and Jian Sun. "Faster r-cnn: Towards real-time object detection with region proposal networks." Advances in neural information processing systems 28 (2015)., available: https://arxiv.org/pdf/1506.01497.pdf

[18] Zhang, L., Lin, L., Liang, X., He, K. (2016). Is Faster R-CNN Doing Well for Pedestrian Detection?. In: Leibe, B., Matas, J., Sebe, N., Welling, M. (eds)

Computer Vision – ECCV 2016. ECCV 2016. Lecture Notes in Computer Science(), vol 9906. Springer, Cham. https://doi.org/10.1007/978-3-319-46475-6_28

[19] Huang, Jonathan, Vivek Rathod, Chen Sun, Menglong Zhu, Anoop Korattikara, Alireza Fathi, Ian Fischer et al. "Speed/accuracy trade-offs for modern convolutional object detectors." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 7310-7311. 2017.

[20] O Pinheiro, Pedro O., Ronan Collobert, and Piotr Dollár. "Learning to segment object candidates." Advances in neural information processing systems 28 (2015).

[21] Dai, Jifeng, Kaiming He, and Jian Sun. "Instance-aware semantic segmentation via multi-task network cascades." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 3150-3158. 2016.

[22] Ananthakrishnan, B., V. . Padmaja, S. . Nayagi, and V. . M. "Deep Neural Network Based Anomaly Detection for Real Time Video Surveillance". International Journal on Recent and Innovation Trends in Computing and Communication, vol. 10, no. 4, Apr. 2022, pp. 54-64, doi:10.17762/ijritcc.v10i4.5534.

[23] Li, Yi, Haozhi Qi, Jifeng Dai, Xiangyang Ji, and Yichen Wei. "Fully convolutional instance-aware semantic segmentation." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2359-2367. 2017.

[24] Andrie Dazlee, N. M. A., Abdul Khalil, S., Abdul-Rahman, S., & Mutalib, S. (2022). Object Detection for Autonomous Vehicles with Sensor-based Technology Using YOLO. International Journal of Intelligent Systems and Applications in Engineering, 10(1), 129–134. https://doi.org/10.18201/ijisae.2022.276

[25] Ren, Shaoqing, Kaiming He, Ross Girshick, and Jian Sun. "Faster r-cnn: Towards real-time object detection with region proposal networks." Advances in neural information processing systems 28 (2015).

[26] Girshick, R., Donahue, J., Darrell, T. and Malik, J., 2014. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 580-587).

[27] "Deep Residual Learning for Image Recognition", Computer Vision and Pattern Recognition, by KAIMING HE, XIANGYU ZHANG, SHAOQING REN, JIAN SUN, 10 DEC 2015, AVAILABLE: HTTPS://ARXIV.ORG/ABS/1512.03385

[28] "Convolutional neural network", From Wikipedia, available: https://en.wikipedia.org/wiki/Convolutional_neural_network

[29] Pawan Kumar Tiwari, Mukesh Kumar Yadav, R. K. G. A. . (2022). Design Simulation and Review of Solar PV Power Forecasting Using Computing Techniques. International Journal on Recent Technologies in Mechanical and Electrical Engineering, 9(5), 18–27. https://doi.org/10.17762/ijrmee.v9i5.370

[30] T. Hossain, F. S. Shishir, M. Ashraf, M. A. Al Nasim and F. Muhammad Shah, "Brain Tumor Detection Using Convolutional Neural Network," *2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT)*, 2019, pp. 1-6, doi: 10.1109/ICASERT.2019.8934561.

[31] S. Albawi, T. A. Mohammed and S. Al-Zawi, "Understanding of a convolutional neural network," *2017 International Conference on Engineering and Technology (ICET)*, 2017, pp. 1-6, doi: 10.1109/ICEngTechnol.2017.8308186.

[32] R. Ezhilarasi and P. Varalakshmi, "Tumor Detection in the Brain using Faster R-CNN," *2018 2nd International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2018 2nd International Conference on*, 2018, pp. 388-392, doi: 10.1109/I-SMAC.2018.8653705.

[33] "Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme", 2010 Dec 1, available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2863141

[34] "Brain Tumor Dataset", By Navoneel Chakrabarty, Nov 2018, available: https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection

[35] "TensorFlow", by Google, 9 November 2015, available: https://www.tensorflow.org/

[36] "Deep Neural Networks", by Wikipedia, available: https://en.wikipedia.org/wiki/Deep_learning#Deep_neural_networks

[37] "Keras", by François Chollet, 27 March 2015, available: https://keras.io/

[38] "OpenCV", by Intel, June 2000, available: https://opencv.org/about/

[39] "COCO dataset", COCO is a large-scale object detection, segmentation, and captioning dataset, available: https://cocodataset.org/

[40] "Labelme Software", Image Polygonal Annotation with Python, by Kentaro Wada, available: https://github.com/wkentaro/labelme

[41] "Intersection over Union", from Wikipedia, available: https://en.wikipedia.org/wiki/Jaccard_index

[42] Shah, P.J. and Shah, T., 2021. Identification of Breast Tumor Using Hybrid Approach of Independent Component Analysis and Deep Neural Network. International Journal of Intelligent Systems and Applications in Engineering, 9(4), pp.209-219.

[43] Altan, G., 2020. Deep Learning-based Mammogram Classification for Breast Cancer. International Journal of Intelligent Systems and Applications in Engineering, 8(4), pp.171-176.

[44] He, Kaiming, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. "Mask r-cnn." In *Proceedings of the IEEE international conference on computer vision*, pp. 2961-2969. 2017.

[45] "Computer Vision: Instance Segmentation with Mask R-CNN", by Renu Khandelwal, available: https://towardsdatascience.com/computer-vision-instance-segmentation-with-mask-r-cnn-7983502fcad1

[46] "Mask R-CNN", available: https://developpaper.com/mask-r-cnn/

[47] Ghazaly, N. M. . (2022). Data Catalogue Approaches,

Implementation and Adoption: A Study of Purpose of Data Catalogue. International Journal on Future Revolution in Computer Science &Amp; Communication Engineering, 8(1), 01–04. https://doi.org/10.17762/ijfrcsce.v8i1.2063