Detection Brain Tumor Using CNN and YOLO

Bommisetty Naveen[1], Mandi Prathyusha[1], Damai Lokesh[1], Kunche Sathwika[1], Himabindu Sathyaveti[2]

[1]B TECH Student, [2]Guide and M. Tech (Ph. D)

Department of Electronics and Communication Engineering, Audisankara College of Engineering and Technology, Gudur, India

---

ARTICLEINFO

ABSTRACT

Different multimodalities, such as magnetic resonance imaging (MRI) and computed tomography (CT), are mashed together in the medical field to create a fused image. Image fusion (IF) is a method for preserving crucial information by combining all pertinent details from numerous photographs into a single fused image. Brain CT and MRI scans are merged in this study and labelled as normal or abnormal. If the connection network is discovered to be abnormal, the portion of the cancer region is determined. These demonstrations make use of YOLOV2 architecture, convolutional neural networks, and image processing methods. The experimental results are evaluated based on model accuracy.

Keywords: Convolutional Neural Network, YOLO, Magnetic Resonance Imaging (MRI), and Computed Tomography (CT).

I. INTRODUCTION

A brain tumour is a tumour or swelling that originates from the building up of a surplus of abnormal neural cells. There are several distinct types of brain tumours. Benign (not cancerous) or malignant (cancerous) brain tumours are also conceivable. The most perennial breed of brain tumours are primary tumours, however tertiary (metastatic) brain tumours can also arise from cancer that first emerged in another part of your body. How quickly a brain tumor grows can vary greatly. The growth rate as well as the location of a brain tumor determines how it will affect the function of your nervous system. The type of brain tumour you have, as well as its size and location, all affect the possibilities for treatment.

Brain-covering membranes (meninges), cranial nerves, the pituitary gland, or the pineal gland are only a few examples of nearby tissues where primary brain tumours might develop. Primary brain tumours start when DNA alterations (mutations) occur in normal cells. The instructions directing a cell's functionality are encoded in its DNA. When healthy cells would die, the mutations instruct the cells to rapidly divide and expand in order to maintain life. The end consequence is a tumour, which is a mass of aberrant cells. Adults with primary brain tumours are much less likely to develop secondary brain tumours,

---

which develop after cancer has spread from the primary site to the brain.

Medical imaging technologies such as magnetic resonance imaging (MRI) and computed tomography (CT) offer precise illustrations of inside body components. To help with medical evaluation and treatment, computer vision techniques can be implemented to analyse and understand these images. In MATLAB, there are several toolboxes and functions available for processing CT and MRI images. For example, the Image Processing Toolbox provides functions for image filtering, segmentation, and feature extraction, which can be useful for analyzing medical images. Additionally, the Medical Image Processing Toolbox is specifically designed for processing medical images, and includes functions for volume rendering, registration, and segmentation.

Convolutional Neural Networks (CNNs) are a ration of deep learning algorithms that have manifest superior efficacy for plenty of image recognition tasks, including medical picture analysis. For developing, training, and deploying CNNs, MATLAB's Deep Learning Toolbox provides functions. Pretrained CNN models are also included in this toolbox for jobs involving medical image analysis. The real-time landmark recognition system YOLO (You Only Look Once) leverages an instrumental neural network capable of recognising many different objects in an image. YOLO is frequently used in computer vision applications, such as picture analysis for medicine. A YOLOv3 object detector is provided by the Computer Vision Toolbox in MATLAB and can be used to find issues in medical images.

Using this technique introduce the YOLO method. A very accurate object identification algorithm called YOLO (You Only Look Once) can identify everything in real-time. Recent deep learning algorithm YOLO can recognise and find objects in videos and pictures.

II. RELATED WORKS

Image Fusion is a process of merging of two images or more than two images or it is a process of blending the contrasting in addition to the shared characteristics of a collection of related but distorted or biassed images to create a final image. For enhanced vision, picture fusion is necessary to obtain high resolution on multispectral and panchromatic pictures as well as realistic images. There are numerous ways for fusing images, ranging from IHS, PCA, DWT, Laplacian pyramids, Gradient Pyramids, DCT, and SF. In the current investigation, every approach is assessed together with its intended inputs and results. In entirety, image fusion gets utilised or works in all digital fields, encompassing medical, military, remote sensing, and broad uses where input images or comparable images are difficult to discern; Fused images assist in providing high-resolution vision [1].

Image fusion is a technique that involves integrating different types of source images, such as multi-modal and multi-focus ones, to create a new, more detailed image. A number of images with different gadgets in focus are assembled using the multi-focus image fusion technique. Digital cameras and remote sensing are two specific instances of image fusion applications. Spatial frequency, image Laplacian energy, morphological bringing it and closure, and other focus metrics are used by the assortment of multi-focus picture fusion techniques covered in the current publication. The capability of the aforementioned algorithms is gauged based on how intent areas in images are chosen to discharge a fused image. The multi-focus image fusion tackle uncovers the focused zones and combines them to yield a more compelling Image [2].

In medical imaging using different modalities such as MRI and CT, complementary information of a targeted organ will be captured. Every appropriate gathered of information gleaned from just about the aforementioned processes has to be joined forces into

one cohesive picture in order for the person being treated to benefit from more accurate diagnosis and therapy. The technique of fusing data contained in multiple images that is pertinent or beneficial into a single image is professionally referred to as image fusion [1]. In this dissertation, we lay out a novel blended average fusion approach relying on a steered image filter and image data gathered to bring together MRI and CT scans of the cerebral cortex. This is the demonstrated algorithm: Each raw picture's detail layers get returned using a steered image filter. Weights corresponding to each source image are calculated from the detail layers with help of image statistics. Then a weighted average fusion strategy is implemented to integrate source image information into one continuous image. Both qualitative and quantitative metrics are used to evaluate fusion performance [3].

Convolutional neural networks have multiple exciting applications one of which is image classification. Other than straightforward image classification, there are many exciting issues in computer vision. One of the most intriguing of these is object detection. YOLO ("You Only Look Once") is a formidable real-time object identification method that was initially presented by Joseph Redmon et al. in their historic 2015 circulation. The YOLO method and a notion of object detection become privy in this article [4].

Convolutional neural networks are constructed from up of two incredibly fundamental building blocks: convolutional layers and pooling layers. For a noteworthy computer vision issue, there are almost limitless ways to assemble these layers, despite their simplicity. Nonetheless, you can construct quite deep convolutional neural networks using both uniformed methods for mounting these layers and novel architectural techniques. The rationale and an intuitive understanding concerning how to use these architectural design choices tailored to forward-thinking image categorization issues may be attained

by studying them whilst fabricating your own deep convolutional neural network theories [5].

Another manageable strategy approaches any sort of machine learning expert desire to be able grasp is the support vector machine. Quite a few individuals prefer the support vector machine because it lends obvious accuracy while using fewer processors. The SVM, or Support Vector Machine, may be utilised for problems with classification as well as regression, according to a tool. It nevertheless can be used frequently in classification ambitions [6].

III. METHODOLOGY

Blending information from many images of the same occurrence that were taken swiftly or at various points in time in order to generate photos with a greater amount of information than the corresponding shots quietly is more commonly referred to as combining images. Feature-based methods, decision-based methods, and pixel-based methods constitute a number of the manner exploit in image fusion. A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm the fact that can take in an input image, give significance (learnable weights and biases) to varying key elements/objects in the image, and be taught to differentiate one regarding the other. A ConvNet needs far less pre-processing than distinct algorithms for classification need. Contrary to basic techniques, where filters must be hand-engineered, ConvNets are capable of learning these filters and the attributes they possess.
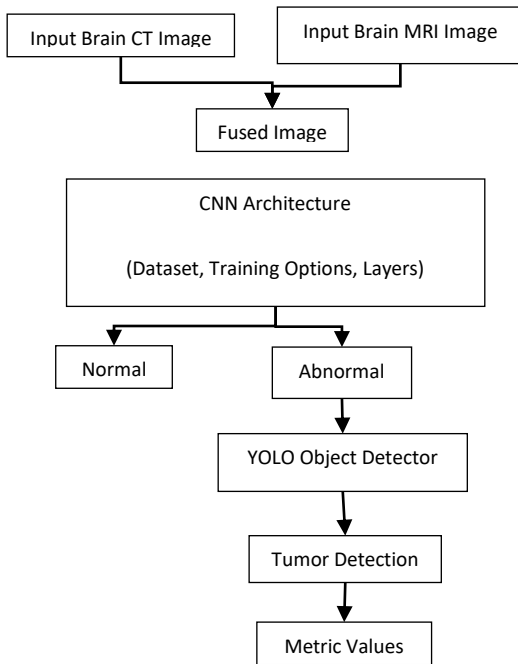
```
┌─────────────────────┐   ┌─────────────────────┐
│ Input Brain CT Image│   │ Input Brain MRI Image│
└─────────────────────┘   └─────────────────────┘
              │                   │
              └─────────┬─────────┘
                        ▼
                ┌──────────────┐
                │  Fused Image │
                └──────────────┘
                        │
        ┌───────────────────────────────────┐
        │         CNN Architecture          │
        │                                   │
        │ (Dataset, Training Options, Layers)│
        └───────────────────────────────────┘
              │                   │
              ▼                   ▼
        ┌──────────┐        ┌──────────┐
        │  Normal  │        │ Abnormal │
        └──────────┘        └──────────┘
                                 │
                                 ▼
                    ┌──────────────────────┐
                    │  YOLO Object Detector │
                    └──────────────────────┘
                                 │
                                 ▼
                    ┌──────────────────┐
                    │  Tumor Detection │
                    └──────────────────┘
                                 │
                                 ▼
                    ┌──────────────────┐
                    │   Metric Values  │
                    └──────────────────┘
```

Fig: Proposed Method Block Diagram

Multi-modal In a technique known as image fusion, data from multiple imagery of various modalities, such as visible and infrared, is merged to generate a single image whose appearance improves the information richness of the input images. The starting point of the strategy underneath MATLAB's multi-modal image fusion algorithms is as outlined in the following

1. Browse through the input images: Using MATLAB's "imread" function, acquire the input images from the various modalities (such as CT and MRI scans) that you wish to combine.

2. Preprocessing: To guarantee that the donated images have a uniform size, resolution, and zestful range, preprocess them. homogenize, noise reduction, and scaling are frequently used preprocessing strategy.

3. Fused Image: The imfuse function can be used to overlay one image over another or to see the juxtaposes between two images. The optional blending ratio, with a value of 0.5 signifying equal blending of both images, defines how much of each image is displayed.

4. CNN Architecture: emblematic, the layers in a

CNN are structured as follows: convolutional layers, activation layers (such as ReLU), pooling layers, and fully connected layers. The convolutional layers build a handful emphasise representations of the input depict after applying multiple layers of filters to it. These feature portrayals capture various facets of the input image. The network is made more non-linear by the activation layers, supplying a non-linear function that matches the convolutional layers' output. due to their assistance of downsampling, the pooling layers diminish the size of the feature maps, while the fully pertaining layers furnish the final output during classification.

5. YOLO object Detector: Users might use the incorporated "yolov2ObjectDetector" and "yolov3ObjectDetector" functions in MATLAB to leverage the YOLO object detector. These operations return the position and kind of items that were identified from an input picture or video. Additionally, users can alter the YOLO object detector's network architecture, train their own model, or use transfer learning to improve an already-trained model.

For multi-modal image fusion in MATLAB, the Image Fusion Toolbox comprises a number of tools and functions, including several fusion methods, registration algorithms, and evaluation metrics. A graphical user interface for interactive fusion and assessment is also provided by the toolkit.

Convolutional Neural Networks (CNNs), that incorporate a multitude of layers, are used to learn and extract traits from images. underlying layer types that are utilised by CNNs include convolutional layers, pooling layers, fully connected layers, and activation layers.

➤ Convolutional layers: Convolutional layers are used to extract information from the image being used by applying some sort of filters to it. Every filter produces an assortment of feature maps that capture various facets of the input image as it slides across the motion picture industry. The

"convolution2dLayer" function in MATLAB may be used to add convolutional layers and enables users to define the number of filters, filter size, and padding.

➤ Max Pooling Layers: The max pooling layer is an instance of a kind of layer the fact that is used in convolutional neural networks (CNNs) to diminish the geographic extent (height and breadth) of the input feature maps while maintaining their essential properties. After the input feature maps have been sliced into rectangular, non-overlapping sections, the maximum value beneath each sector gets picked as the output in the operation of the max pooling layer. The network's parameter count is reduced as a result of avoiding the phenomenon of overfitting.

➤ Fully Connected Layers: The culminating lumping output of the network is provided by fully linked layers. MATLAB's "fullyConnectedLayer" function, which lets users define the number of output neurons, may be used to build completely linked layers.

➤ RELU Layers: Convolutional neural networks (CNNs) necessitate non-linearity, hence the ReLU (Rectified Linear Unit) layer represents a particular type of layer that has been developed. After element-wise applying it to the layer's input feature maps, the ReLU activation function simply outputs the most significant value between 0 and the input value.

Various additional kinds of layers, including dropout layers, normalisation layers, and concatenation layers, can also be used in CNNs in addition to these primary layer types. Overall, the order and mix of these layers in a CNN design can have a considerable influence on how well the network performs in a variety of tasks, such as segmentation, object recognition, and image categorization.

Every neighbourhood in YOLOv2 can have its information shown in the visualisation by moving the cursor over it. With the exception of the last Convolution block, each Convolution block starts with BatchNorm normalisation and ends with Leaky Relu activation. The supplied image is used by YOLO for help creating an SS grid. Each grid cell predicts a single item.

For each grid cell,

• It forecasts B boundary boxes, with a box confidence score for each box,

• Regardless of the quantity of boxes B, it only detects one object,

• It makes predictions about C conditional probabilities, one for each type of objects to be probable.

The box trustworthiness score is displayed in the margins of boxes. The confidence score indicates the precision of the bounding box and the possibility that it is empty (or contains nothing). We subtract the image's width and height from the bounding box's dimensions. The offsets to the crucial cell are X and Y. Thus, x, y, w, and h are all integers between 0 and 1. There are 20 probabilistic class probabilities in each cell. The conditional class probability (one probability per category for each cell) represents the likelihood that the regarded object falls under a certain class. Each prediction box's class confidence score is calculated as follows:

class confidence score = box confidence score * conditional class probability

It monitored the accuracy of the localization (the specific position of an object) as well as the classification. These score and probability phrases are straightforward to translate. The mathematical definitions are provided here for your future use.

The loss will be determined by YOLO using the sum-squared error between the expected and actual data. Regardless of what creates the loss function,

• The classification loss.

• The differences between the actual boundary box and the projected boundary box, or localization loss.

• The confidence loss (the objectness of the box)

Classification loss

If an item arises, the squared error of the class conditional probability for each class represents the classification loss for each cell. The sizes and locations of the projected border boxes are determined using the localization loss. Only boxes that can detect the item are counted. Absolute errors shouldn't be weighed equally in large and small boxes. Comparable to one with an enormous box is a little box with a 2-pixel inaccuracy.

IV. RESULTS AND DISCUSSIONS

Medical imaging methods like computed tomography (CT) and magnetic resonance imaging (MRI) use a variety of ideas to generate complex pictures of the interior of the human body. While MRI creates pictures using powerful magnets and radiation, CT creates cross-sectional images using X-rays. You may analyse and analyse pictures from MRI and CT scans using built-in MATLAB tools and functions, such as the Image Processing Toolbox and the Medical Image Processing Toolbox. With the aid of these tools, you may, among other things, edit, segment, and extract characteristics from the images. Convolutional neural networks (CNNs), a family of deep learning methods, perform at the cutting edge in a variety of computer vision tasks like segmentation, object identification, and image categorization. The Deep Learning Toolbox may be used in MATLAB to train and use CNNs for a variety of tasks. The popular object identification technique You Only Look Once (YOLO) uses one neural network to forecast bounding boxes and class probabilities for numerous items in a picture. You may use MATLAB's Deep Learning Toolbox and Computer Vision Toolbox to implement YOLO. These programmes offer features for conducting object recognition on fresh photos and training YOLO models on unique datasets.
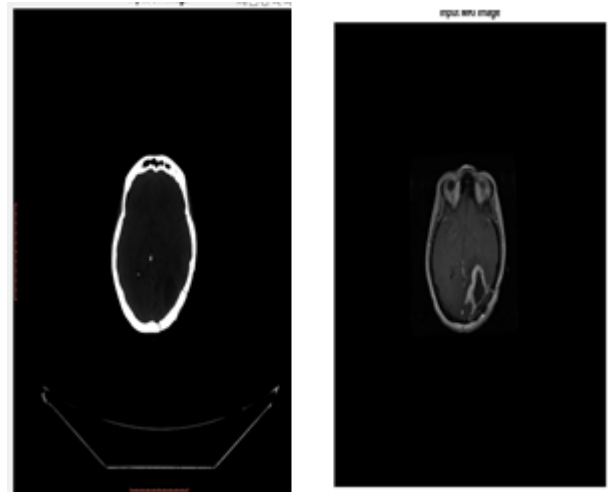

Fig: Input CT & MRI Images

Input Brain CT image and input brain MRI image are medical images that provide different types of information about the brain tissues and structures.
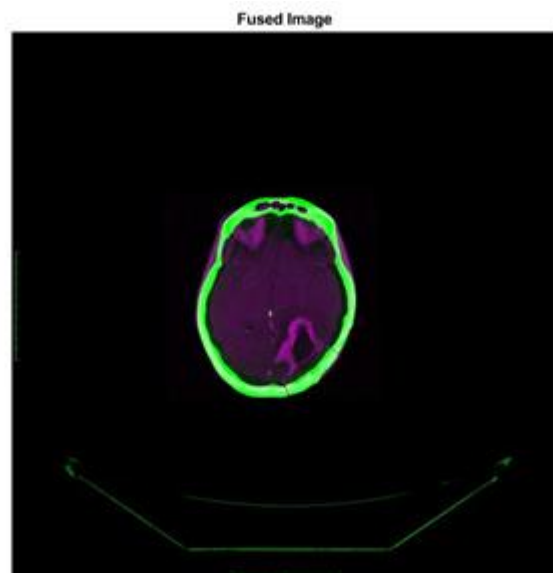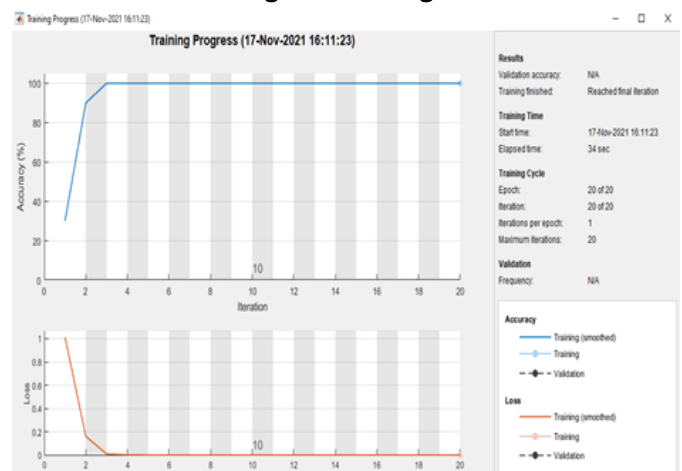

Fig: Fused Image

Fig: Training Progress

A Convolutional Neural Network (CNN) can be used to classify the brain images as normal or abnormal, based on a dataset of brain CT and MRI images with If the CNN classifies the brain images as normal, the result is a diagnosis of normal brain tissue and structures. If the condition is abnormal, we can use a YOLO object detector to detect tumors in the brain images.
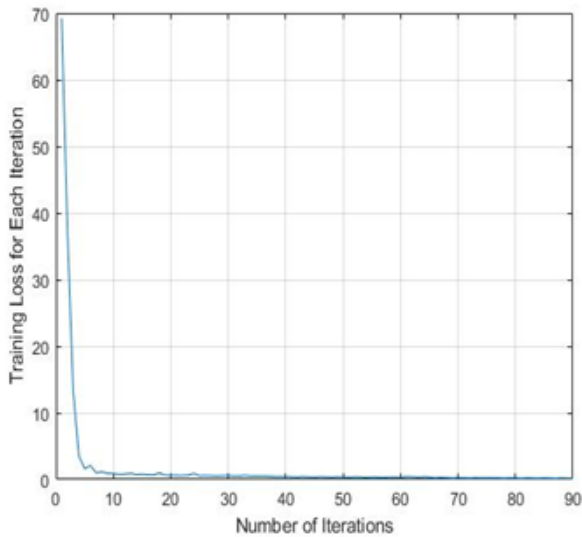

Fig: Tumor Detected Image

Fig: Final Accuracy Result

Fig: Training Loss Graph

Following their discovery, tumours can be measured using a variety of metrics to determine their size, shape, and location, which can aid in the formulation of treatment options for brain tumours. Finally, the well-known object identification system YOLO (You Only Look Once) predicts bounding boxes and class probabilities from whole photos in real-time using a single neural network.
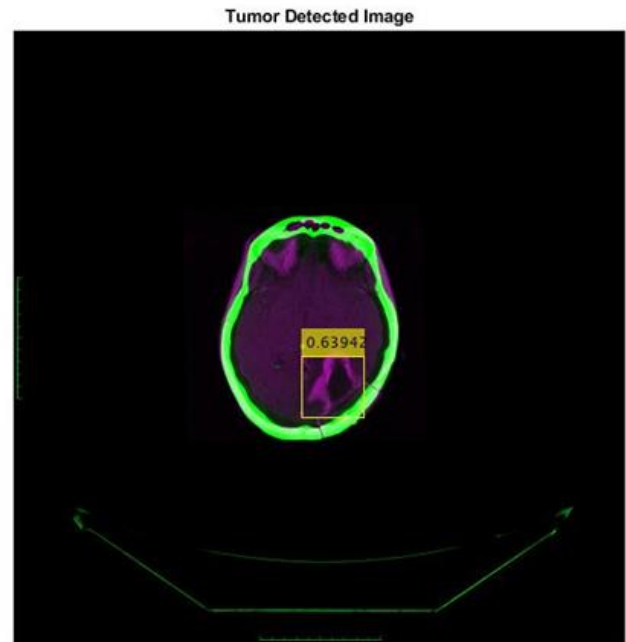
V. CONCLUSION

The foremost objective of the endeavour is to develop a machine learning approach for discovering and recognising brain tumours that is precise, rapid, and easy to use. Segmentation, feature extraction for texture and form, and an SVM classifier are all steps in the conventional approach to spot brain tumours, and they all need a substantial amount of processing time and complexity. The approved method, in addition to other enhancements in performance and computation time, incorporates a classification based on convolution neural networks and a detection based on YOLO. Brain pictures with tumours (abnormal) or without them are included as the categorization results as well. When compared to previous works, accuracy and training loss metrics are assessed and produce the best results that are satisfactory.

VI. REFERENCES

[1]. Kaur H, Rani E J, "Analytical comparison of various image fusion techniques", International and Software Engineering 5 (2015).

[2]. Garg R, Gupta P, Kaur H, "Survey on multi-focus image fusion algorithms", Recent Advances in Engineering and Computational Sciences (RAECS) (2014) pp. 1-5.

[3]. Bavirisetti DP, Kollu V, Gang X, Dhuli R, "Fusion of MRI and CT images using guided image filter and image statistics", International Journal of Imaging Systems and Technology. 2017 Sep;27(3):227-37.

[4]. Jędrzej Świeżewski, Ph.D, YOLO Algorithm & Object Detection, Appsilon, Copyrights © appsilon.com, 10 October 2020, https://appsilon.com/object-detection-yolo-algorithm/

[5]. Jason Brownlee, "Convolutional Neural Network Model Innovations for Image Classification", Machine Learning Mastery, © 2021 Machine Learning Mastery, July 5, 2019. https://machinelearningmastery.com/review-of-architectural-innovations-for-convolutional-neural-networks-for-image-classification/

[6]. Rohit Gandhi, "Support Vector Machine — Introduction to Machine Learning Algorithms", https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca4

Cite this Article