Journal of Innovative Image Processing (ISSN: 2582-4252)
www.irojournals.com/iroiip/

Brain Tumor Identification using YOLO Network

T. R. Ganesh Babu[1], R. Praveena[2], M. Manoharan[3], A. Rajadurai[4], M. Sridharan[5]

Department of Electronics and Communication Engineering, Muthayammal Engineering College, Rasipuram, India.

E-mail: [1]ganeshbabutr@gmail.com, [2]praveenajuhi@gmail.com, [3]mahimano2003@gmail.com, [4]rajadurai3112003@gmail.com, [5]sridharm575@gmail.com.

Abstract

CAD systems for brain MRI analysis employ various AI techniques to assist radiologists in interpreting images and detecting abnormalities. These systems must be trained on large datasets encompassing diverse brain pathologies to ensure accurate detection and classification of different diseases. In this research, the use of YOLOv4 and YOLOv5 architectures for brain tumour detection in MRI images is an interesting application of deep learning technology. The performances metrices such as Precision, Recall, F1 Score and mAP are analysed. The coding for this work was developed using Python, utilizing TensorFlow as the platform. Simulations were carried out on Google Colab.

Keywords: YOLO Network, Glioma Tumor, Meningioma Tumor, Pituitary Tumor, MRI Images

1. Introduction

Tumours in brain are clumps of abnormal cells that grow in your head. These can be really serious, causing death or long-term problems for people around the world. Brain tumours come in two categories, namely: cancerous and non-cancerous. Cancerous ones, also called malignant brain tumours or gliomas, are particularly dangerous because the cancer cells spread quickly inside your brain. The good news is that not all brain tumours are cancerous.

DOI: https://doi.org/10.36548/jiip.2024.2.009
Received: 16.04.2024, received in revised form: 08.05.2024, accepted: 21.05.2024, published: 03.06.2024

Brain tumours can be life-threatening, turning lives upside down for both patients and their families. Because they grow aggressively inside the brain, they can be very dangerous. But the good news is, catching them early makes a big difference in how well someone can be treated. Doctors use special machines to look for brain tumours. These machines take pictures of the brain, and some of the most common ones are CT scans, MRIs, and X-rays.

The old ways of taking pictures of the brain inside someone's head have some downsides. For one, the machines sometimes use radiation, which can be bad for your health in the long run. They also might not always catch the problem, and the radiation can mess with brain cells. These tests can be expensive and not always suitable for everyone, especially older folks and pregnant women. There's a new way to do things, though scientists are working on computer programs that can automatically find and outline brain tumors in MRI scans. These programs are like super-powered assistants, trained with a special technique called deep learning to be really good at their job. They can even use something called transfer learning to identify scary brain tumors like glioblastoma. This is exciting because it could be a safer, more accurate way to find brain tumors.

This method uses a powerful object detection tool called YOLO and deep learning to find and classify brain tumours in MRI scans. It basically trains a computer program to recognize tumors by using a technique called transfer learning. For the task of brain tumour identification, YOLOv4 and YOLOv5 architectures were selected due to their superior performance in object detection tasks. The primary criteria for selecting these models included their high accuracy, speed, and ability to detect small objects within an image.

1.1 Classification of Tumor in Brain

Four types of Tumor in brain. They are,

- Glioma Tumor
- Meningioma Tumor
- Pituitary Tumor
- No Tumor

Glioma Tumor

Gliomas start in the brain or spinal cord. These tumors grow from cells that look like healthy helper cells in the brain, called glial cells. As a glioma gets bigger, it can press on the brain or spinal cord, causing problems depending on the location. Different areas control different functions, so symptoms will vary.

Meningioma Tumor

Meningiomas are the most common brain tumor, originating from the protective layers surrounding your brain and spinal cord (the meninges). Even though they aren't technically brain tumors themselves, they can grow and press on the brain, nerves, and blood vessels nearby. This is why they're often classified as brain tumors.

Pituitary Tumor

A tiny pea-sized gland at the base of your brain, can sometimes develop abnormal growths called pituitary tumors. Thankfully, most of these tumors are benign, meaning they're not cancerous and won't spread. However, they can disrupt hormone production in your pituitary gland, leading to various health issues. The pituitary gland is like the control center for many hormones in your body, so problems there can cause a variety of symptoms.

No Tumor

Soft tissue, like muscles and fat, can sometimes develop lumps called benign tumors. These are not cancer and won't spread throughout your body. They usually don't pose a threat to your life and can often be removed with surgery. There are many different types of these benign tumors, each categorized by the specific soft tissue they originate from.

2. Related Work

Yang et al. [4] combined U-Net and residual networks to develop the Deeper ResU-Net, effectively utilizing the strengths of both architectures for enhanced feature extraction. Zhang et al. developed AResUNet [5] by integrating a series of attention units into the downsampling and upsampling processes. This model adaptively resizes features to enhance local responses of the downsampled residual features, which are then used for feature recovery

in the upsampling process. Gan et al. [6] introduced a global attention mechanism to address the limitation of convolution operations, which can only extract local information, by capturing long-distance dependencies. Aboelenein et al. [7] introduced a hybrid two-track U-Net (HTTU-Net), where one track focuses on the tumor's shape and size, while the other captures contextual information. Zhou et al. [8] designed an efficient encoder-decoder architecture for brain tumor segmentation, using the lightweight ShuffleNetV2 neural network as the encoder to minimize parameters and achieve a large receptive field. Wang et al. [9] introduced a spatial dilated feature pyramid (DFP) module. Many models fail to fully exploit global context information. Chen et al. [10] developed a two-stage automated brain lesion segmentation framework by combining cascaded RF and dense CRF. Huang et al. [11] introduced a group cross-channel attention residual U-Net, designed to fully utilize low-level fine details in tumor regions.

Abd-Ellah et al. [12] conducted an in-depth research study on several diagnostic methodologies for brain MRI images. In their study [13], the authors presented various strategies for detecting brain cancers from MR images. They focused on using three-dimensional CNNs, SVMs, and multi-class SVMs for deeper segmentation. Ganesh et al. [14] Proposed a semantic classification techniques applied to medical imaging data, especially for identifying and outlining brain cancer zones, hold significant promise for improving diagnosis and treatment planning.

3. Materials and Method

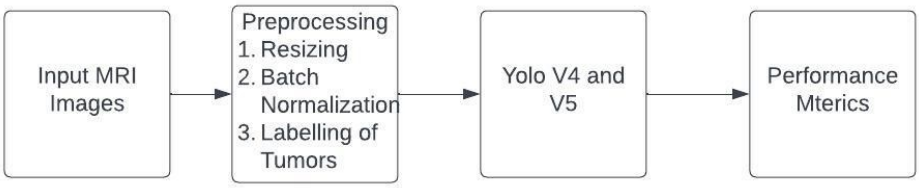

Figure 1. Proposed Block Diagram

Figure 1 shows the block diagram of proposed method. The MRI images, resized to 224x224 pixels, were obtained from the Kaggle [17].

3.1 Preprocessing

The input MRI images were resized and normalized using batch normalization. Subsequently, 70% of the brain MRI images with different tumour conditions were used to train the YOLO network. Batch normalization is a technique used to improve the training of deep neural networks by normalizing the inputs to each layer. This helps in speeding up the training process and achieving higher overall accuracy. Images were resized to a consistent dimension suitable for the YOLO models, typically 224x224 pixels, using OpenCV (cv2). This resizing step is crucial for maintaining uniformity and improving computational efficiency. Batch normalization is performed by using deep learning frameworks that is cv2. For labelling tumors, a detailed annotation process was undertaken using tools like LabelImg to create bounding boxes around the tumors in MRI images.

3.2 Description of Dataset

The dataset consists of a total of 2,402 images, divided into four classes: Glioma, Pituitary, Normal, and Meningioma. The distribution of images across the classes is as follows:

Glioma: 826 training images, 100 testing images, and 8 validation images.

Pituitary: 827 training images, 105 testing images, and 8 validation images.

Normal: 493 training images, 74 testing images, and 8 validation images.

Meningioma: 822 training images, 115 testing images, and 8 validation images.

3.3 YOLO V4 and V5 Architecture

In this research, detection of brain tumour is carried out by using YOLOv4 and YOLOv5 architecture. In YOLOv4 indeed utilizes the smaller version of the CSPDarknet53 feature extractor as its backbone. CSPDarknet53 consists of three CSP (Cross-Stage Partial) blocks, each incorporating "leaky" activation functions. These blocks help in enhancing feature representation and facilitating better object detection performance within the YOLOv4 architecture. The Backbone: YOLOv4 uses CSPDarknet53 as its backbone network. CSP stands for Cross Stage Partial network. This backbone is a modified version of Darknet, a deep neural network architecture designed for object recognition, particularly in real-time scenarios.

CSPDarknet53 helps in extracting features from input images efficiently. The Neck: YOLOv4 has a spatial pyramid pooling (SPP) module and a PANet (Path Aggregation Network) as its neck. These modules help in capturing features at different scales and integrating them for better detection performance. The SPP module captures context information at multiple scales, while the PANet module aggregates features from different network layer.

The head detection of YOLOv4 predicts bounding boxes, objectless scores, and class probab6ilities for detected objects. It predicts bounding boxes directly, along with confidence scores indicating the presence of an object within each bounding box and class probabilities for each detected object. YOLOv4 uses a modified version of the YOLOv3 head that includes various improvements such as Mish activation function, which helps in better gradient flow and convergence during training. Feature Fusion: YOLOv4 employs feature fusion techniques to combine features from different layers of the network effectively. This helps in enhancing the representation power of the model and improving its detection performance.

Training of YOLOv4 is typically trained using the COCO (Common Objects in Context) dataset, which contains various objects belonging to 80 different classes. During training, YOLOv4 optimizes a loss function that penalizes errors in predicting bounding boxes, object ness scores, and class probabilities. YOLOv4 network balance between speed and accuracy. In YOLOv5 training pipeline is simpler compared to YOLOv4. It involves single-stage training with default configurations, making it more straightforward and easier to implement. YOLOv5 implements more advanced data augmentation techniques like CutMix and MixUp, which have been shown to improve generalization and robustness of the model.

YOLOv5 marks a significant departure from its predecessors by adopting PyTorch instead of Darknet as its framework. PyTorch also provides flexibility and ease of experimentation. Additionally, YOLOv5 utilizes CSPDarknet53 as its backbone, which is a variant of the Darknet architecture that incorporates Cross-Stage Partial connections to improve performance and efficiency. These changes have contributed to YOLOv5's popularity and effectiveness in real-time object detection tasks. Figure 2 and 3 illustrates architecture of YOLOv4 and YOLOv5 respectively. [15, 16]

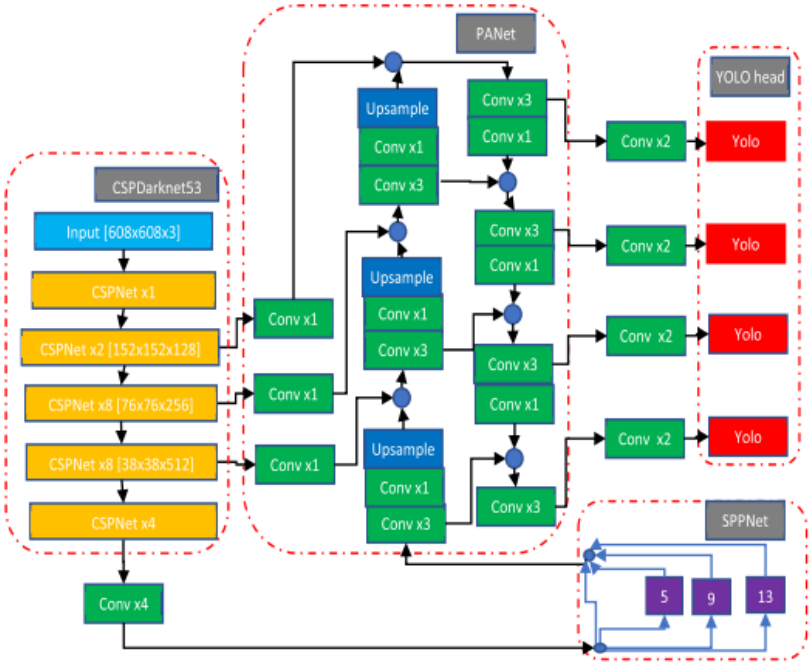

Figure 2. Architecture of YOLOv4 [16]

3.4 Hyper Tuning Parameters of YOLO V5 Architecture

Based on the information provided, here is a summary of the hyperparameter tuning plan for the YOLO object detection model:

Learning Rate: Adjusted to 0.0001

Dataset Split: Training and testing data was split in the ratio of 70:30 respectively.

Number of Epochs: 50

Backbone Network: ResNet, was used to achieve better accuracy by leveraging its deep architecture and pretrained weights.

Confidence Threshold: Was set to 90% to filter out predictions with lower confidence scores, potentially improving the precision of the model.

In the context of deep neural networks, Convolutional Neural Networks (CNNs) can suffer from the vanishing gradient problem, especially in very deep networks. This limitation arises when gradients become increasingly small as they propagate backward through the

network during training, making it challenging to update the weights of early layers effectively. To address this issue, the YOLO architecture utilizes ResNet as a backbone.

ResNet is employed in YOLO to leverage its residual connections, which are also known as skip connections. These connections allow gradients to flow more easily through the network by providing shortcuts that bypass several layers. By enabling the direct flow of gradients to earlier layers, ResNet helps mitigate the vanishing gradient problem, allowing for the training of deeper networks more effectively. This plan outlines a systematic approach to optimizing the YOLO model's performance by tuning key hyperparameters and selecting appropriate components based on the desired accuracy and performance metrics. Table 1 represents the comparison between Pretrained Yolo V5 with Yolo V5 with Backbone (ResNet).

Table 1. Comparison with Yolo V5 with Backbone (ResNet)

Parameter	Pretrained Yolo V5	Yolo V5 with Backbone (ResNet)
maP@90% Confident Level	0.86	0.921

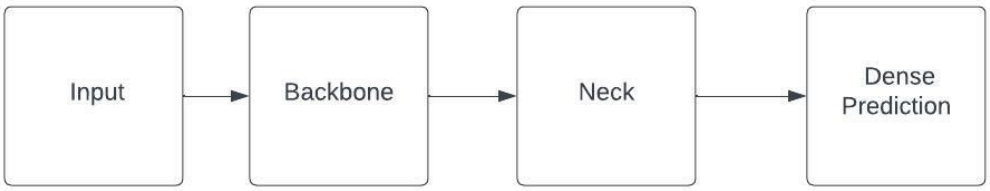

Figure 3. YOLOv5 Architecture [16]

4. Result and Discussion

The performance of the YOLOv4 and YOLOv5 architectures is evaluated based on several parameters. The mean Average Precision (mAP) is used to measure the average precision across different classes. The F1 score, which is the harmonic mean of precision and recall, provides a balanced measure of the model's performance, particularly useful when dealing with imbalanced class distributions. Precision represents the proportion of true positive predictions among all positive predictions made by the model, while recall measures the

proportion of true positive predictions among all actual positive instances. From the Table 2 infer that performance of YOLOv4 and YOLOv5 in terms of mAP, F1 score, precision, and recall. Based on absorbed values it appears that YOLOv5 outperforms YOLOv4 across all these metrics, indicating that YOLOv5 achieves higher accuracy, better balance between precision and recall, and overall better performance in object detection tasks. The equations 1-4 indicates mAP, F1 score, precision, and recall.

$$mAP = \sum_{q=1}^{Q} \frac{AveP(q)}{Q}, \tag{1}$$

$$F1score = 2 * \frac{(precision * recall)}{precision + recall}, \tag{2}$$

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive}, \tag{3}$$

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative}, \tag{4}$$

Table 2. Performance Analysis of Proposed Networks

Measure	YOLOv4	YOLov5
Precision	0.84	0.921
Recall	0.52	0.702
F1 Score	0.64	0.786
mAP	0.57	0.758

To diagnose the Brain Tumour with the help of MRI Images using YOLO Architecture and identifies the Glioma Tumour, Meningioma Tumour, Pituitary Tumour and No Tumour shown in Figure 4.

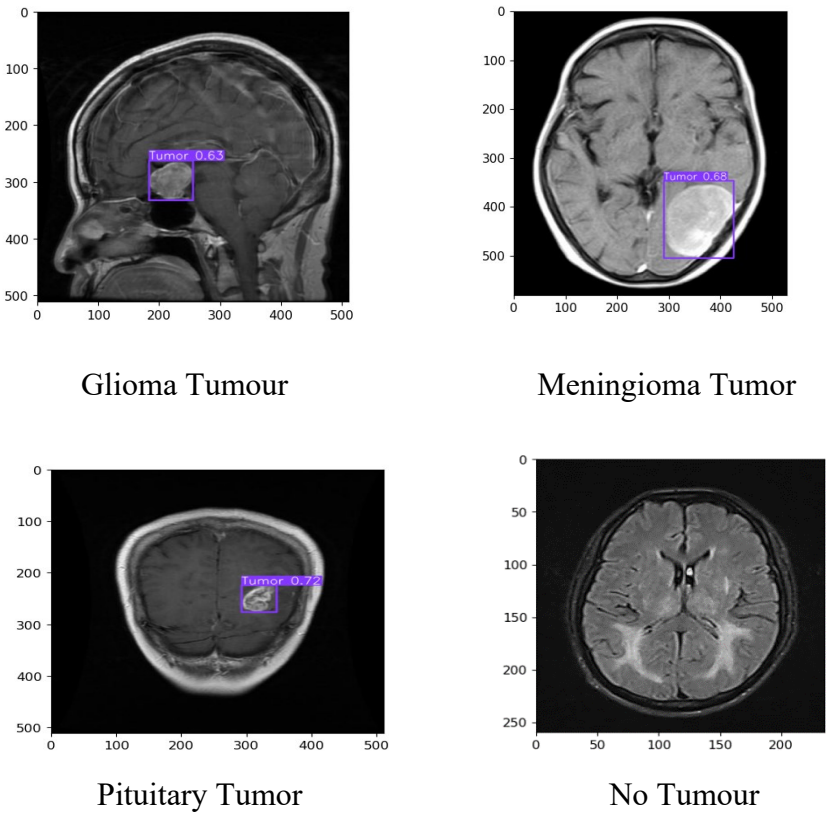

Figure 4. Tumour Classification

Figure 5 and 6 shows the training and validation loss and accuracy over 50 epochs for a deep learning model. The x-axis represents the number of epochs, while the y-axis represents the loss and accuracy .

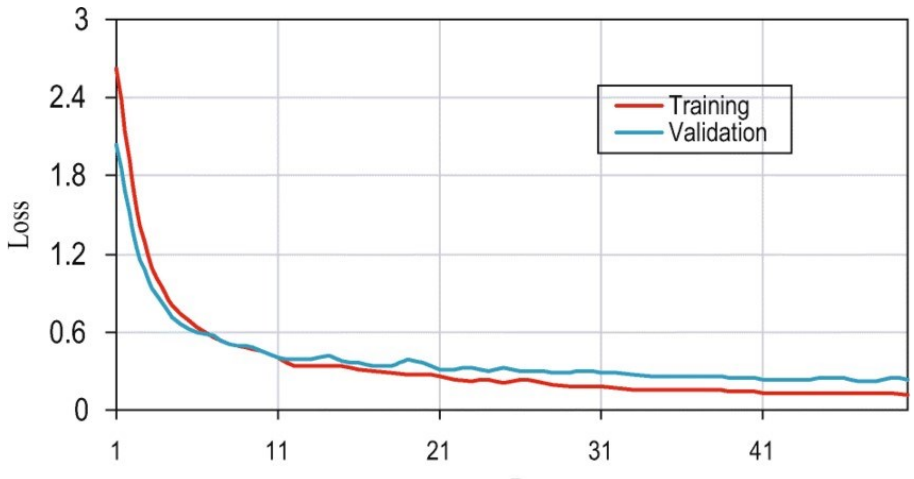

Figure 5. Epochs Vs Loss

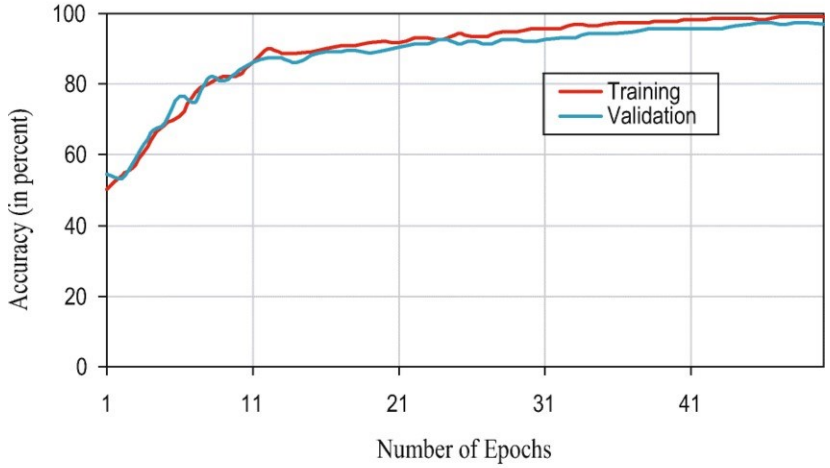

Figure 6. Epochs Vs Accuracy

5. Conclusion

The research work describing focuses on using both YOLOv4 and YOLOv5 architectures for the detection of brain tumour types. The performance metrics including mAP, F1 score, precision, and recall have been analysed, and it's found that YOLOv5 performs better compared to YOLOv4 in terms of these metrics. Furthermore, in future work, hybrid deep learning classifiers will be developed to classify specific tumor types. This approach suggests combining the strengths of different AI techniques to improve the accuracy and reliability of tumor classification.

References

[1] S. Subramanian, K. Scheufele, N. Himthani, and G. Biros, “Multiatlas calibration of biophysical brain tumor growth models with mass effect,” 2020, arXiv:2006.09932. 551–560

[2] K. Scheufele, S. Subramanian, and G. Biros, “Automatic MRI-driven model calibration for advanced brain tumor progression analysis,” 2020, arXiv:2001.09173. 1-7

[3] K. Scheufele, S. Subramanian, A. Mang, G. Biros, and M. Mehl, “Image-driven biophysical tumor growth model calibration,” 2019, arXiv:1907.07774. B549–B580

[4] T. Yang, J. Song, L. Li, and Q. Tang, ‘‘Improving brain tumor segmentation on MRI based on the deep U-Net and residual units,’’ J. X-Ray Sci. Technol., vol. 28, no. 1, pp. 95–110, Feb. 2020.

[5] J. Zhang, X. Lv, H. Zhang, and B. Liu, ‘‘AResU-Net: Attention residual U-Net for brain tumor segmentation,’’ Symmetry, vol. 12, no. 5, p. 721, May 2020.

[6] X. Gan, L. Wang, Q. Chen, Y. Ge, and S. Duan, ‘‘GAU-Net: U-Net based on global attention mechanism for brain tumor segmentation,’’ J. Phys., Conf. Ser., vol. 1861, no. 1, Mar. 2021, Art. no. 012041.

[7] N. M. Aboelenein, P. Songhao, A. Koubaa, A. Noor, and A. Afifi, ‘‘HTTUNet: Hybrid two track U-Net for automatic brain tumor segmentation,’’ IEEE Access, vol. 8, pp. 101406–101415, 2020.

[8] X. Zhou, X. Li, K. Hu, Y. Zhang, Z. Chen, and X. Gao, ‘‘ERV-Net: An efficient 3D residual neural network for brain tumor segmentation,’’ Expert Syst. Appl., vol. 170, May 2021, Art. no. 114566.

[9] J. Wang, J. Gao, J. Ren, Z. Luan, Z. Yu, Y. Zhao, and Y. Zhao, ‘‘DFPResUNet: Convolutional neural network with a dilated convolutional feature pyramid for multimodal brain tumor segmentation,’’ Comput. Methods Programs Biomed., vol. 208, Sep. 2021, Art. no. 106208.

[10] G. Chen, Q. Li, F. Shi, I. Rekik, and Z. Pan, ‘‘RFDCR: Automated brain lesion segmentation using cascaded random forests with dense conditional random fields,’’ NeuroImage, vol. 211, May 2020, Art. no. 116620.

[11] Z. Huang, Y. Zhao, Y. Liu, and G. Song, ‘‘GCAUNet: A group crosschannel attention residual UNet for slice based brain tumor segmentation,’’ Biomed. Signal Process. Control, vol. 70, Sep. 2021, Art. no. 102958.

[12] M. K. Abd-Ellah, A. I. Awad, A. A. M. Khalaf, and H. F. A. Hamed, ‘‘A review on brain tumor diagnosis from MRI images: Practical implications, key achievements, and lessons learned,’’ Magn. Reson. Imag., vol. 61, pp. 300–318, Sep. 2019.

[13] S. Somasundaram and R. Gobinath, ‘‘Current trends on deep learning models for brain tumor segmentation and detection—A review,’’ in Proc. Int. Conf. Mach. Learn., Big Data, Cloud Parallel Comput. (COMITCon), Feb. 2019, pp. 217–221.

[14] T.R. Ganesh Babu , V. Varsha , T. S Yamuna Sri , S. Priyanka Kumari. (2024) “Brain Tumor Classification using Convolutional Neural Network, Journal of Innovative Image Processing, , Vol.6, Issue 1, pp. 16-26

[15] Upesh Nepal and Hossein Eslamiat (2022) ‘ Comparing YOLOv3, YOLOv4 and YOLOv5 for Autonomous Landing Spot Detection in Faulty UAVs” Sensors, MDPI, Vol. 22, 464, pp.1-15.

[16] Anand koirala , meena jha , srinivas bodapati , animesh mishra , girija chetty, praveen kishore sahu , sanjib mohanty , timir kanta padhan , jyoti mattoo, and ajat hukkoo(2022) ‘deep learning for real-time malaria parasite detection and counting using yolo-mp” ieee engineering in medicine and biology society section, ieee access, vol.10 pp.102157-102172

[17] https://www.kaggle.com/ sartajbhuvaji/brain-tumor-classification-mri