iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI: https://doi.org/10.22214/ijraset.2026.79329

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

Brain Tumor Detection and Classification using Convolutional Neural Network (CNN)

Hima Rameshbhai Vankar[1], Dr. Kamal Shah[2]
[1, 2]*St. John College of Engineering and Management*

Abstract: ***Brain tumor detection plays a crucial role in medical diagnostics, carrying profound implications for patient outcomes and care. This project presents a detailed exploration into the development and deployment of an advanced brain tumor detection system. The methodology follows a multi-step process, beginning with the collection of a comprehensive and diverse dataset of brain imaging scans. Following data acquisition, preprocessing steps such as noise reduction and image enhancement are employed to improve scan quality, ensuring that the data is in an optimal form for analysis. At the core of this system lies deep learning, particularly through the use of Convolutional Neural Networks (CNNs). These networks are leveraged to extract meaningful features from the preprocessed data, enabling the model to effectively distinguish between brain scans indicative of tumors and those that are not. The model's training process is supported by a validation set, allowing for fine-tuning to achieve peak performance. To further assess its real-world application, the trained model is tested on an entirely separate, previously unseen dataset, providing valuable insights into the model's accuracy and robustness. In the implementation phase, the system is integrated into a real-time processing platform, enabling the rapid analysis of incoming brain imaging data. Predefined thresholds are established to minimize false alarms, ensuring that only the most probable tumor cases are flagged for further examination by medical professionals. Additionally, an ongoing maintenance and monitoring framework is set in place to adapt the model to evolving tumor characteristics and advances in medical research, ensuring the system remains accurate and relevant in clinical practice. Ultimately, the results of this research aim to offer a significant advancement in the field of brain tumor detection, providing a reliable, high-precision tool that can enhance patient care and outcomes in medical imaging.***
Keywords: CNN, VGG16, Deep Learning, EfficientNet-B4, ResNet50, Naïve Bayes, Logistic Regression, KNN, SVM, Stacking ensemble.

I. INTRODUCTION

Brain tumor detection is a critical component in modern medical diagnostics, as early and accurate identification of tumor growth significantly enhances a patient's chances of successful treatment and long-term survival. Traditional diagnostic methods rely heavily on manual inspection of MRI or CT scan images by medical experts, a process that is time-consuming, prone to fatigue, and may be influenced by subjective interpretation. With the rapid advancements in artificial intelligence and medical imaging technologies, automated systems have emerged as powerful tools for improving diagnostic precision. The proposed system is designed to address these challenges by integrating state-of-the-art deep learning techniques to accurately detect and classify brain tumors in medical images. The proposed brain tumor detection system adopts a structured and highly efficient pipeline, beginning with the acquisition of diverse, high-quality brain imaging datasets. These datasets incorporate variations in tumor types, sizes, shapes, and imaging conditions, ensuring that the model is capable of learning generalized patterns rather than being limited to narrow cases. To enhance data reliability, extensive preprocessing is performed on the acquired images. This includes noise reduction, contrast adjustment, skull stripping, and image normalization. Such operations are essential to improve the clarity and consistency of the images, enabling the deep learning model to extract meaningful features without interference from irrelevant artifacts. At the core of the proposed system lies the use of Convolutional Neural Networks (CNNs), which have proven tremendously successful in image classification and medical imaging tasks. CNNs are designed to identify spatial hierarchies in images, making them ideal for detecting subtle patterns in tumor-affected brain regions. The model learns these features during the training phase, where it is exposed to thousands of preprocessed MRI images labeled as tumor-present or tumor-absent. A separate validation set supports the optimization process, ensuring that the model generalizes well and avoids issues such as overfitting.

Once the model is trained, it undergoes rigorous testing on completely unseen data to evaluate its real-world applicability. This testing phase is crucial for understanding the system's accuracy, sensitivity, specificity, and robustness across a variety of cases, including rare tumor patterns. Performance metrics from this stage provide valuable insights into the reliability of the system when deployed in clinical settings.

©IJRASET: All Rights are Reserved

The proposed system's implementation focuses on real-time analysis, enabling rapid processing of incoming patient scans. Through an efficient interface, medical professionals can upload MRI images, receive instant diagnostic predictions, and review highlighted tumor regions.

Carefully defined decision thresholds are incorporated to minimize false positives and false negatives, thereby ensuring that only the most probable cases are flagged for detailed medical evaluation.

Furthermore, the system is designed with scalability and adaptability in mind. A continuous monitoring and updating framework allows the model to evolve alongside advancements in medical research and the emergence of new tumor variations. This feature ensures that the system remains clinically relevant and maintains high precision over time.

In conclusion, the proposed system aims to transform brain tumor detection by offering a fast, reliable, and high-accuracy diagnostic tool. By combining deep learning, advanced preprocessing, and robust evaluation techniques, it empowers healthcare professionals with enhanced decision-making capabilities and contributes meaningfully to improved patient outcomes.

A. Related Work

Brain tumor detection using MRI images has been an active area of research, with numerous studies exploring machine learning, deep learning, and hybrid techniques to enhance diagnostic accuracy.

Sadad et al. [1] investigated advanced deep learning models for tumor detection and multi-class classification, demonstrating that CNN-based architectures significantly outperform traditional approaches due to their ability to automatically learn complex spatial patterns. Similarly, Tiwari et al. [2] presented a multiclass brain tumor detection system using CNNs, highlighting notable improvements in classification accuracy across diverse tumor types.

Hybrid models that combine handcrafted features with deep learning methodologies have also been explored. Saba et al. [3] introduced a fusion model integrating handcrafted features with deep CNN-extracted features, resulting in enhanced diagnostic performance. In contrast, Islam et al. [4] applied a combination of superpixels, PCA, and K-means clustering for detection, showing that while classical feature engineering can be effective, it struggles with irregular tumor boundaries and complex textures.

Deep learning-based architectures continue to dominate recent research. Choudhury et al. [5] and Maqsood et al. [6] both proposed CNN and DNN hybrid frameworks, obtaining high accuracy in tumor classification tasks.

Toğaçar et al. [7] developed BrainMRNet, a novel CNN model specifically designed for MRI-based tumor identification, achieving exceptional accuracy due to enhanced convolutional layers and optimized learning strategies. Similarly, Alsubai et al. [8] introduced an ensemble deep learning model, demonstrating that combining multiple networks increases robustness and reduces prediction variance.

The development of deeper and more efficient CNN architectures has led to breakthroughs in both classification and segmentation tasks. Ayadi et al. [9] proposed a deep CNN structure optimized for tumor classification, achieving improved precision with minimal preprocessing. Noreen et al. [10] presented a concatenation-based deep learning framework, enabling effective extraction of both local and global features. Irmak [11] introduced an optimized CNN model capable of multi-class classification with minimal computational overhead, improving real-time applicability.

Multiscale analysis and segmentation models have also gained attention. Díaz-Pernas et al. [12] proposed a multiscale CNN capable of tumor classification and segmentation simultaneously, showcasing improved segmentation consistency across different tumor sizes. Abiwinanda et al. [13] further validated the potential of CNNs for tumor classification, emphasizing their reliability across multiple MRI datasets. Pathak et al. [14] and Siar & Teshnehlab [15] demonstrated that CNN-based approaches consistently outperform traditional machine learning methods, particularly when large datasets are available.

Recent clinical and biological studies from NCBI and related sources [16], as well as medical organizations such as AANS and Dana-Farber [18, 20], provide valuable background on tumor characteristics, imaging challenges, and diagnostic standards, supporting the development of AI-assisted medical systems. Additionally, research on image super-resolution by Wang et al. [17] has contributed to improved medical image clarity, enhancing deep learning model performance. Contemporary scientific studies published in Nature [19] further validate the importance of AI-driven solutions for precise tumor detection.

B. Material And Methods

The methodology of the present study is illustrated in Fig. 1. Major steps in the present study comprise brain tumor dataset selection, pre-processing MRI images, feature extraction, and classification by various classifiers.

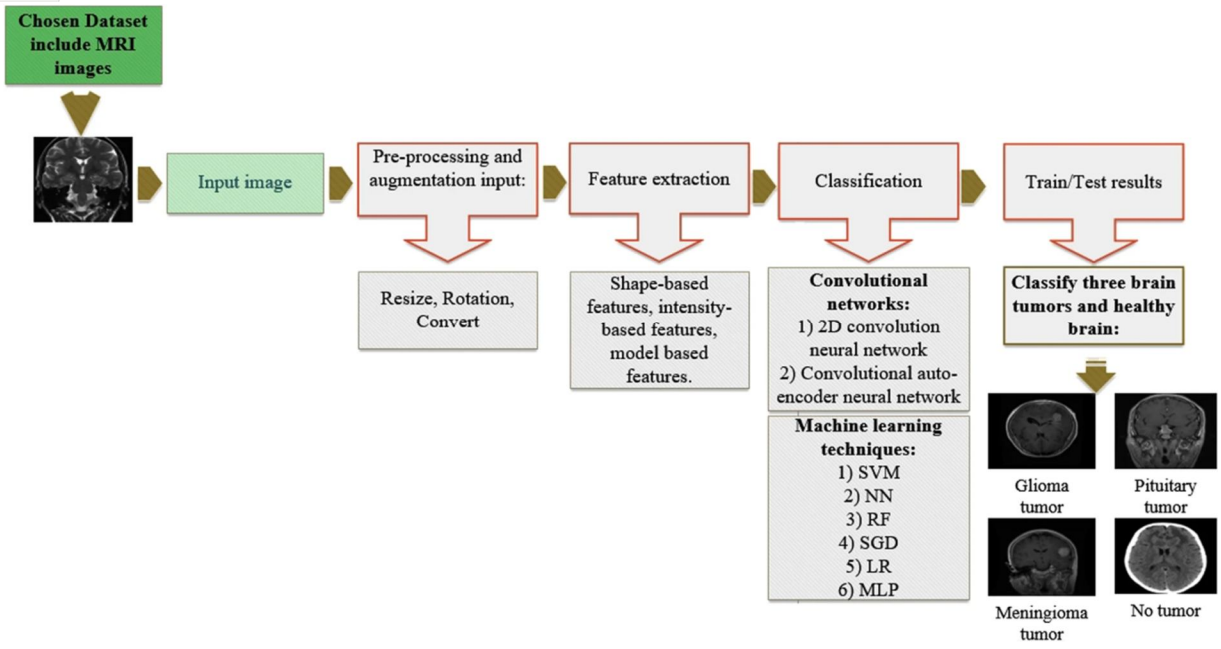


Figure no. :1 Stages of the proposed methodology

1) Data Acquisition: The process begins with the collection of a comprehensive dataset of MRI brain images. These images include multiple tumor categories such as glioma, meningioma, and pituitary tumors, as well as non-tumorous scans. Publicly available datasets such as BraTS, Figshare, and Kaggle MRI repositories are used to ensure data diversity in terms of tumor shape, location, intensity variations, and imaging conditions. A diverse dataset is essential for developing a robust model capable of generalizing across real-world clinical scenarios.

2) Image Preprocessing: Preprocessing plays a crucial role in enhancing image quality and preparing the data for analysis. Initially, noise reduction techniques such as Gaussian filtering or median filtering are applied to eliminate unwanted artifacts. Skull stripping is performed to remove non-brain tissues, ensuring the model focuses solely on the relevant region. Image normalization is then conducted to standardize pixel intensity values, improving model stability during training. Additional augmentation techniques—such as rotation, flipping, zooming, and contrast adjustment—are used to artificially expand the dataset and reduce overfitting.

3) Feature Extraction Using CNN: The core of the system utilizes Convolutional Neural Networks (CNNs), which have proven highly effective for medical image analysis. CNN layers automatically extract hierarchical features such as edges, shapes, textures, and tumor-specific patterns. Filters in the convolution layers capture local features, while deeper layers learn abstract, high-level representations of tumor structures. This automated extraction eliminates the need for manual feature engineering and significantly improves accuracy.

4) Model Training and Validation: The dataset is divided into training, validation, and testing subsets. During the training phase, MRI images are fed into the CNN model, which iteratively updates its weights using backpropagation and optimization algorithms such as Adam or SGD. A validation set is used to fine-tune hyperparameters, monitor loss and accuracy, and prevent overfitting. Techniques such as dropout, batch normalization, and early stopping are incorporated to enhance generalization.

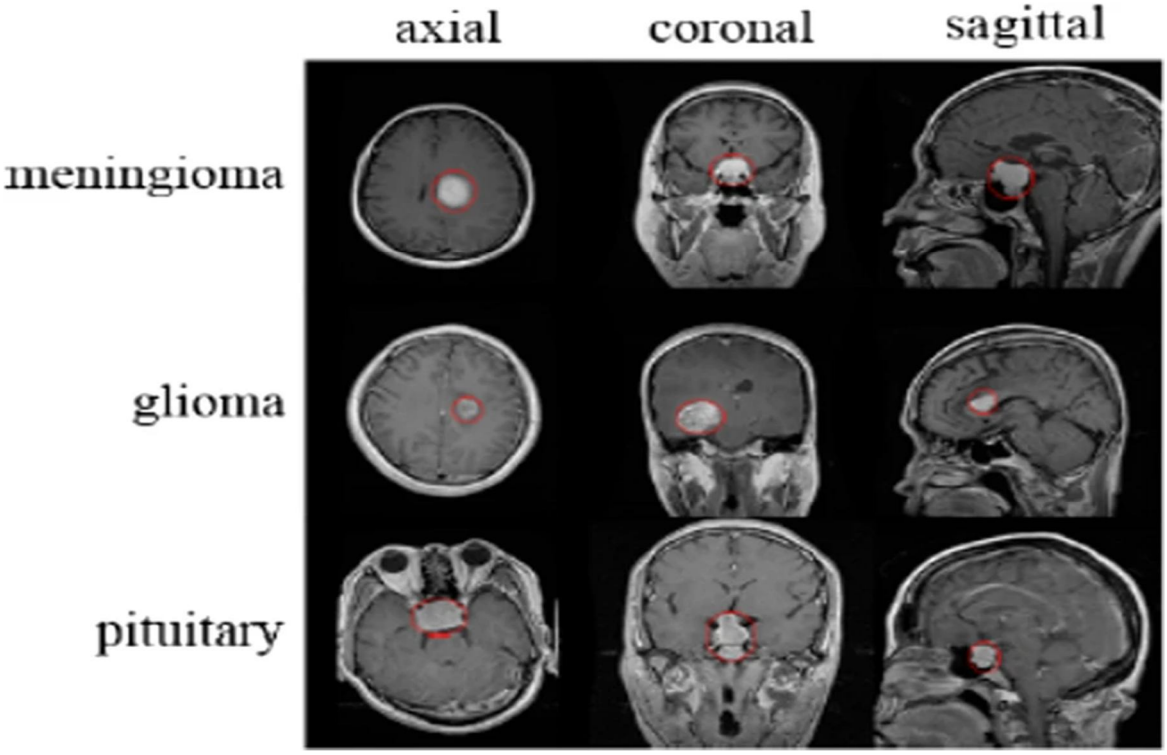
5) Model Testing and Evaluation: After training, the model is evaluated using an unseen test set to measure metrics such as accuracy, precision, recall, F1-score, specificity, sensitivity, and confusion matrices. This evaluation phase ensures the model's robustness and ability to handle real clinical cases. Performance comparison with existing methods from literature highlights the superiority and reliability of the proposed approach.

6) System Deployment: The final stage involves integrating the trained model into a real-time application or web-based interface. The system accepts MRI scans as input, processes them instantly, and produces predictions along with highlighted tumor regions. Decision thresholds are set to minimize false positives and false negatives. A maintenance framework ensures periodic retraining based on new medical data.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

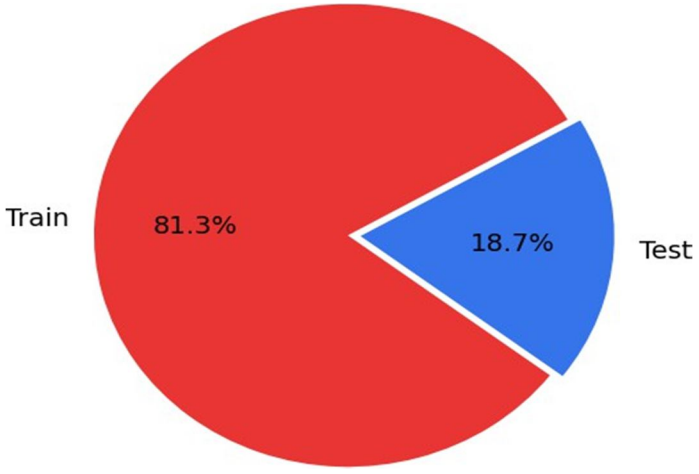
II. DATASET

The dataset used in this work was created by combining three datasets: figshare, SARTAJ, and Br35H. The dataset includes 7023 human brain MRI scans that have been classified into four groups: pituitary, glioma,meningioma, and no tumor. [8] In particular, the images from the "no tumor" class were obtained using theBr35H dataset. The comprehensive coverage of different types of brain tumors and non-tumor instances in the diversified dataset facilitates the development and testing of accurate and dependable brain tumor detection algorithms. With multiple classes and a large number of images, the dataset is more valuable for creating and testing deep learning models for brain tumor detection.



FigureNo:2. Description of normalized MRI images presenting diverse varieties of tumor in a different plane

The training and testing subsets of the dataset are separated, and the ratio of the data is plotted in the figure below, where all four types of tumors are present in the respective sets of the training and testing datasets. Simple random sampling ensured that each image had an equal likelihood of being included in the training or the testing subset. With over 80% of the images in the training subset of the data and less than 20% going toward actual testing, it is evident from the pie chart that the focus is on the model's training. However, in these situations, the most crucial aspect is on accurately training the model.



FigureNo3 Dataset Ratio Representation

©IJRASET: All Rights are Reserved

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

The dataset includes the four major groups discussed previously, and Figure 3 shows their uniform distribution. For the purposes of testing and training, each group includes the following numbers of images: There are 1595 images in the training dataset for "notumor," 405 in the testing dataset, 1457 images in the training dataset and 300 images in the testing dataset for "pituitary," 1339 images in the training dataset and 306 images in the testing dataset for "meningioma," and 1321 images in the training dataset and 300 images in the testing dataset for "glioma."

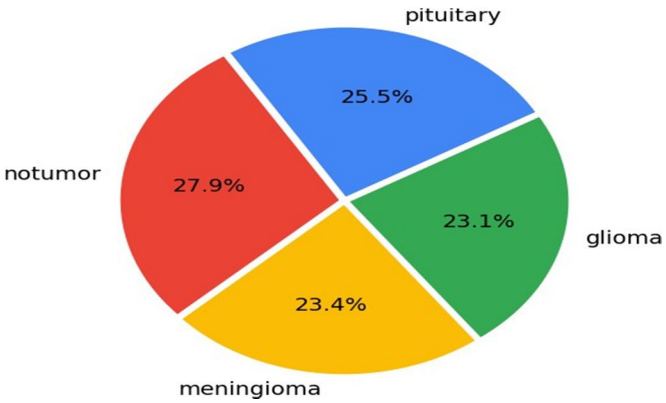


Figure No 4 Dataset Distribution

A. Dataset Source

In this study, the brain tumor classification dataset used is available at https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumorclassification-mri. It comprises 3264 T1-weighted contrastenhanced MRI images. The brain images used from the dataset of the categories 'no tumor' and 'glioma_tumor'. A sample of data that has been used in this paper prior to preprocessing is

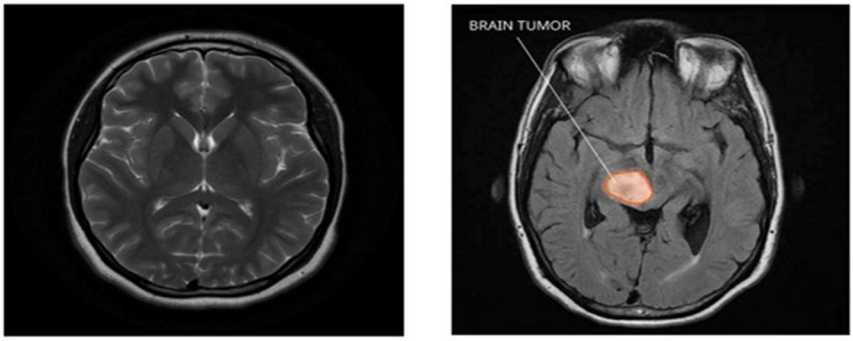


Figure 5 a) Normal Brain and b)with tumour

B. Pre-Processing

The size of data set images and that of network's input must match in order to train the network and generate predictions on new data. Therefore, the size of the images is adjusted to fit the network and then data is scaled or cropped. 0.1 is the zoom range, horizontality is the rotation, and 0.5 and 1.0 are the brightness ranges respectively, and 0.1 is the preprocessing range for rescaling, width and height shift [1,5,6,11]. Then, 80% dataset is used for training and 20% for validation [1,6]

C. CNN

The CNN model has multiple convolutional layers followed by max pooling, which allows the network to capture and abstract complex features from the brain scans. They work to detect various aspects of the images, such as edges and textures that are essential for distinguishing between healthy tissue and tumor affected tissue. Batch normalization that the model's representation remains stable across different images with varying position of the tumor was also done.

During the training process, certain neurons are randomly deactivated by dropout layers in order to reduce the risk of overfitting. Thus, enhancing the ability for the model to generalize its learning to new, unseen data. During the training phase, which spanned 10 epochs, it was observed that there was a continual improvement in accuracy and a gradual decrease in loss on the training set, which shows efficient learning by the model. The model used was Sequential and the batch size was 32.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

The Adam optimizer was chosen for the model having a learning rate of 0.0001[1-8]. The learning rate is kept constant in all the model implementations. Adam, which stands for Adaptive Moment Estimation, uses the combined advantages of two popular optimizers: RMSProp and AdaGrad. Adam calculates adaptive learning rates for each parameter. An activation function of ReLu(Rectified Linear Unit) has been used. For the loss function calculation, the categorical_crossentropy has been used. Meanwhile, the validation accuracy and loss provided an insight into the model's performance on data it had not previously encountered.

D. VGG-16

The VGG-16-based model utilizes a pre-trained VGG16 network, a renowned architecture known for its depth and effectiveness in image classification tasks. Unlike our initial CNN model, VGG16 comes with a series of convolutional layers pre-configured and trained on a vast dataset (ImageNet).In adapting VGG16, modification of the base network to suit the binary classification task – identifying the absence or presence of a brain tumor was done. This was achieved this by appending additional layers to the pretrained VGG16 model, which was loaded with its original ImageNet weights. The model's top layers, originally designed for 1000- class classification, were replaced with a Flatten layer to reshape the output for dense layers, and a Dense layer with 64 nodes (using ReLU activation) for further feature interpretation. A Dropout layer (0.5) randomly disables International Journal of Computer Applications (0975 – 8887) Volume 186 – No.50, November 2024 33 neurons to mitigate overfitting risk and thus enhances the model's generalization capabilities. The model's final layer is dense having the softmax activation function and is updated to output two classes, corresponding to 'tumor present' or 'no tumor'. Adam optimizer is used to compile the model and categorical cross entropy is used as the loss function to maintain consistency with our initial CNN model

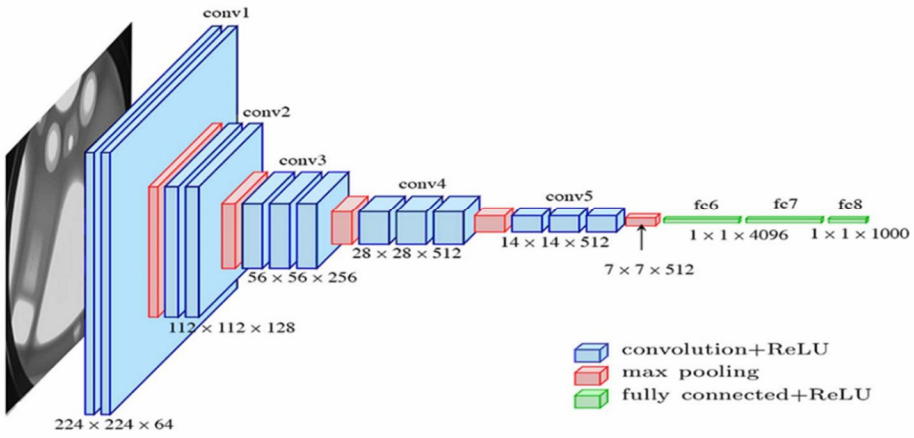


Figure 6 vgg16 architecture

E. RESNET-50

On the ResNet-50 base, a GlobalAveragePooling2D layer was added. This facilitates the dimensionality reduction of the feature maps, condensing the spatial information. This reduces the computational load and the number of parameters in the network. Then a fully connected Dense layer with 256 neurons and ReLU activation is added. This layer serves as a classifier that processes the features extracted by the ResNet-50 model. The final layer is another Dense layer having softmax activation function. Adam optimizer is used to compile this model and the categorical_crossentropy loss function is employed to maintain homogeneity for comparison.

F. MobileNet

The MobileNet model precisely, MobileNetV1 used in our implementation is pre-trained on the ImageNet dataset. This pre-training provides an advantage, as the model has already learned a wide range of features from a diverse set of images. Using include_top=False to discard the top classification layer, allowing us to customize the network for our task. The input shape is adjusted to match our MRI images (100x100 pixels with 3 color channels). Following the MobileNet base, a GlobalAveragePooling2D layer was added. Thus, the spatial dimensionality of the feature maps is reduced by summarizing the most critical features in every map. Next, 1024 neurons in a fully connected Dense layer with ReLU activation function[3,6].

The final layer is also Dense with a softmax activation and two neurons, designed for our task. Adam optimizer is used to compile the model and the categorical_crossentropy loss function is used for comparison.

G. INCEPTIONV3

InceptionV3 has a 'network within a network' approach and efficiency in handling various scales of features. It was developed by Google and pre-trained on ImageNet, and this model is adapted for binary classification of brain tumor in an MRI image. It makes use of the Inception modules which allows incorporating multiple filters within the same layer. Factorisation is used to decompose the large convolutional layers to smaller manageable operations to achieve a more robust model. This also uses a batch normalization approach so that different images can be effectively classified. Modification of InceptionV3 was done to accept 100x100 pixel images with three color channels, replacing the top classification layers to suit our binary classification task of detecting brain tumors. The GlobalAveragePooling2D a layer was also used, followed by the layer that is fully connected with a softmax activation function. Parameters were all kept the same for the sake of fair comparison.

Below is the Image graph for the comparison for all the five models used and their respective performance measures.

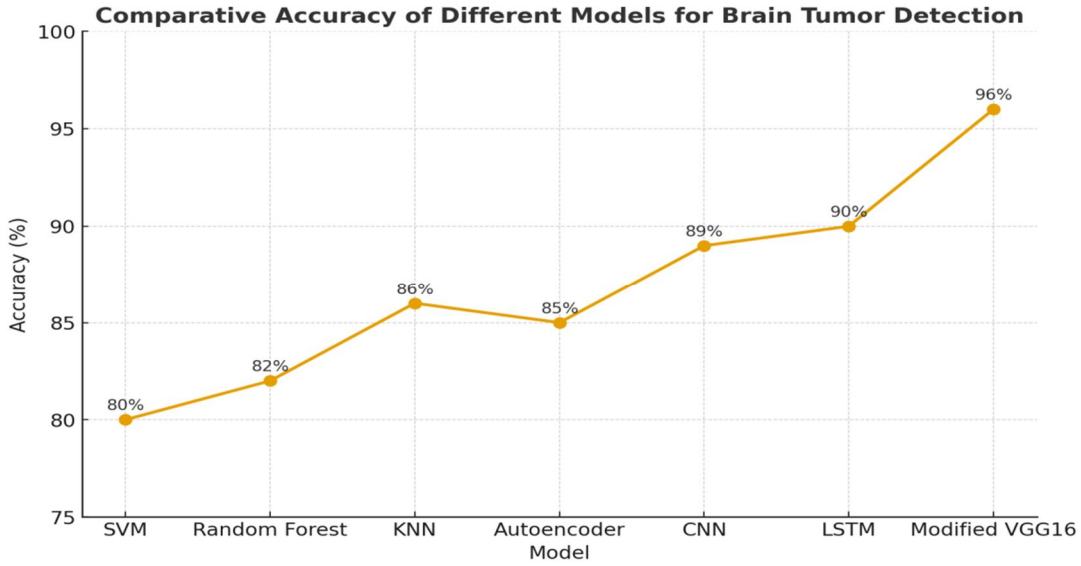


Figure 7 Comparision of accuracy with other Model

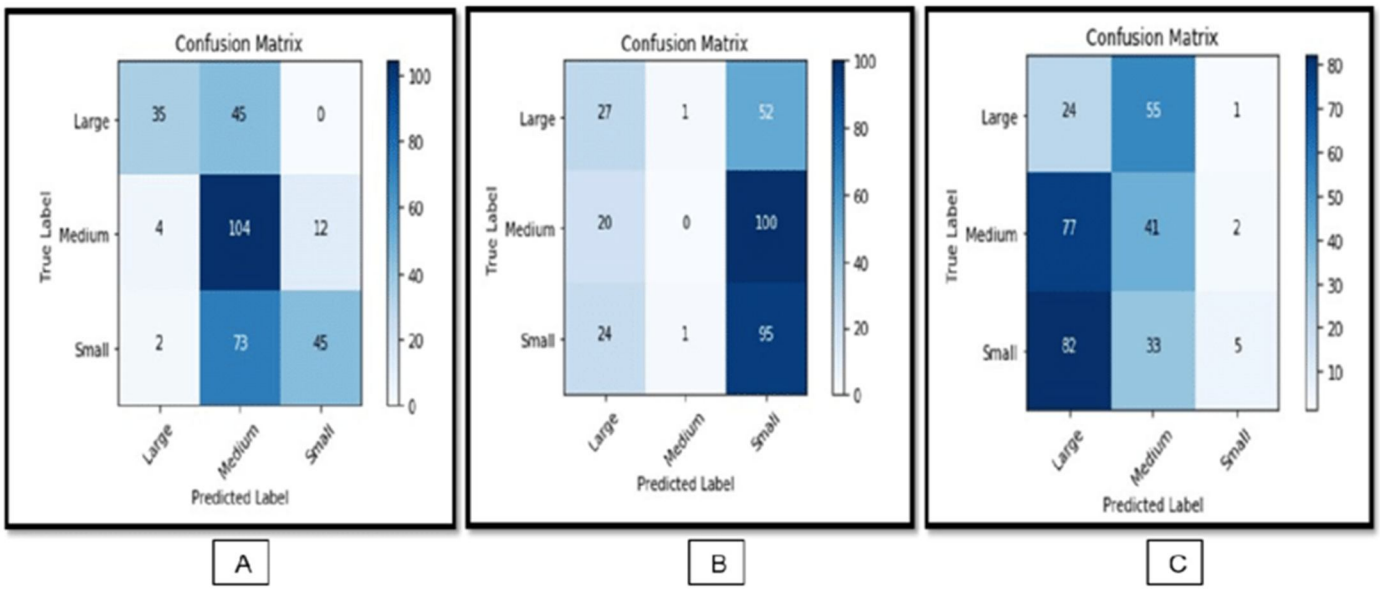


Figure 8 Confusion matrix for deep learning models (A) VGG16, (B) InceptionV3, (C) ResNet50

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

III. RESULTS

Image processing and model definition are the two primary steps we discussed in our proposed work. We will review and evaluate the outcomes we were able to attain in this section. Initially, we processed the images using PIL, and the images in Figure 7 illustrate the outcomes of the uniform image processing.

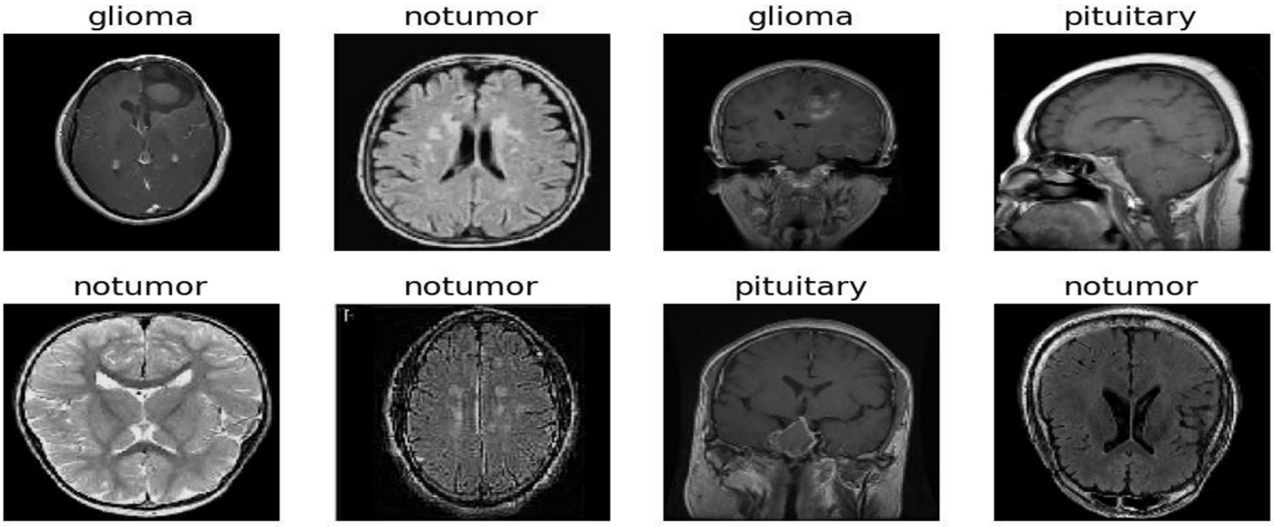


Figure 9 Images after PIL processing

The next stage after processing the images was to add data to our model and train it. The mode will train on 20 images at a time since we fitted the model with the data and set the batch size to 20. We also set the epochs to 5, which means the model will go through the training dataset five times, continuously obtaining random images from the dataset. Figure 8. is a graphic representation of the accuracy and loss results from our five epochs, as we can see from Figures 8 and 9, each epoch produces results with higher accuracy and lower loss. As we can see from the exact values displayed, the accuracy increased with each epoch, going from 83% to 96,9%. These results, given the size of the dataset, are satisfactory.



Figure 10 Model Training History

Figure 11 Accuracy and loss values after each epoch

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

The classification report for our model is shown in Figure 10. As we can see, the model performed extremely well in every class, with the lowest results being 89% for meningioma and the highest being 99% for pituitary, 96% for no tumor, and 98% for glioma. All of these findings are highly consistent with one another, and the data indicates that the model has performed extremely well in its task of classifying the type of tumor.

Additionally, the model's overall accuracy of 96,9% is quite promising for this particular problem and its potential applications in the future. Another figure 11, belove, shows these values in a chart to provide a more readable view of the information obtained from the model. The model's performance in various classes is visually represented by the confusion matrix in Figure 12. It is clear that the model misclassified 32 images in the "meningioma" class, where it performed lowest. This misprediction is consistent with the accuracy metric found in the classification report, which shows that the "meningioma" class has a lower accuracy than other classes. Furthermore, there are tiny percentages of inaccurately predicted images for other classes, which are also represented in the accuracy values given in the classification report. For those particular classes, the differences between the expected and actual classes lead to somewhat lower accuracy percentages. The confusion matrix provides a thorough summary of the model's advantages and disadvantages in every class, effectively demonstrating these misclassifications.

Class	Precision	Recall	F1-Score	Support	Overall Accuracy
notumor	0.96	0.89	0.93	300.00	
meningioma	0.89	0.97	0.93	306.00	
pituitary	0.99	0.98	0.99	405.00	
glioma	0.98	0.98	0.98	300.00	
Overall					0.96

Figure 12 Classification report

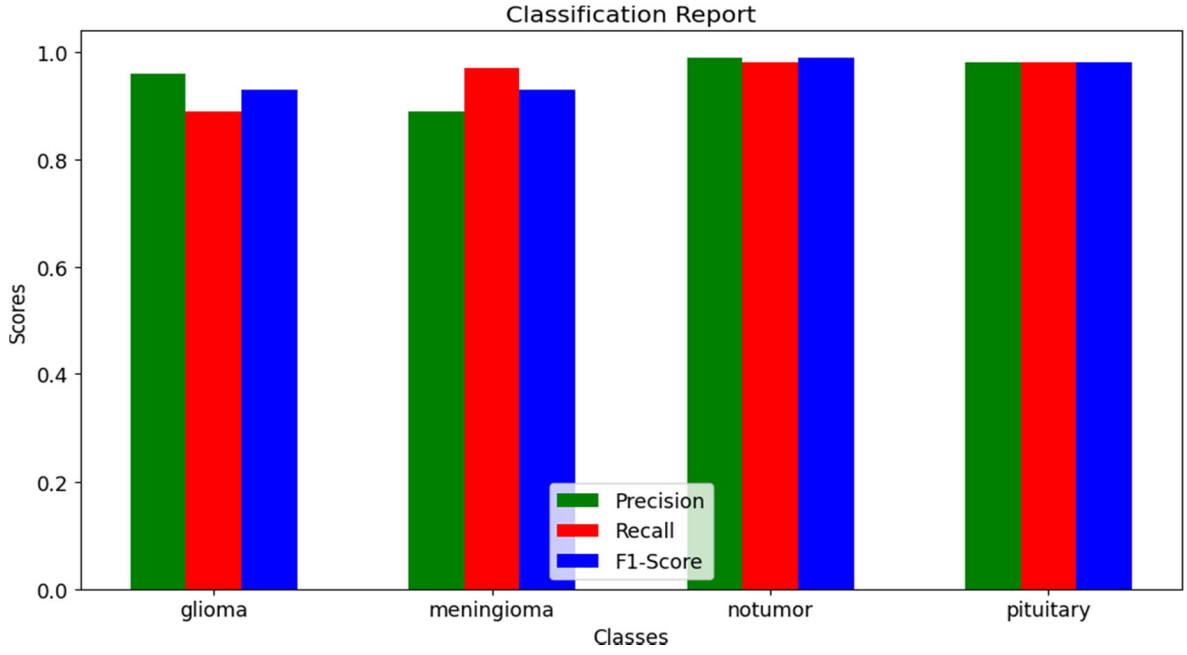


Figure 13 Plot representation of classification report

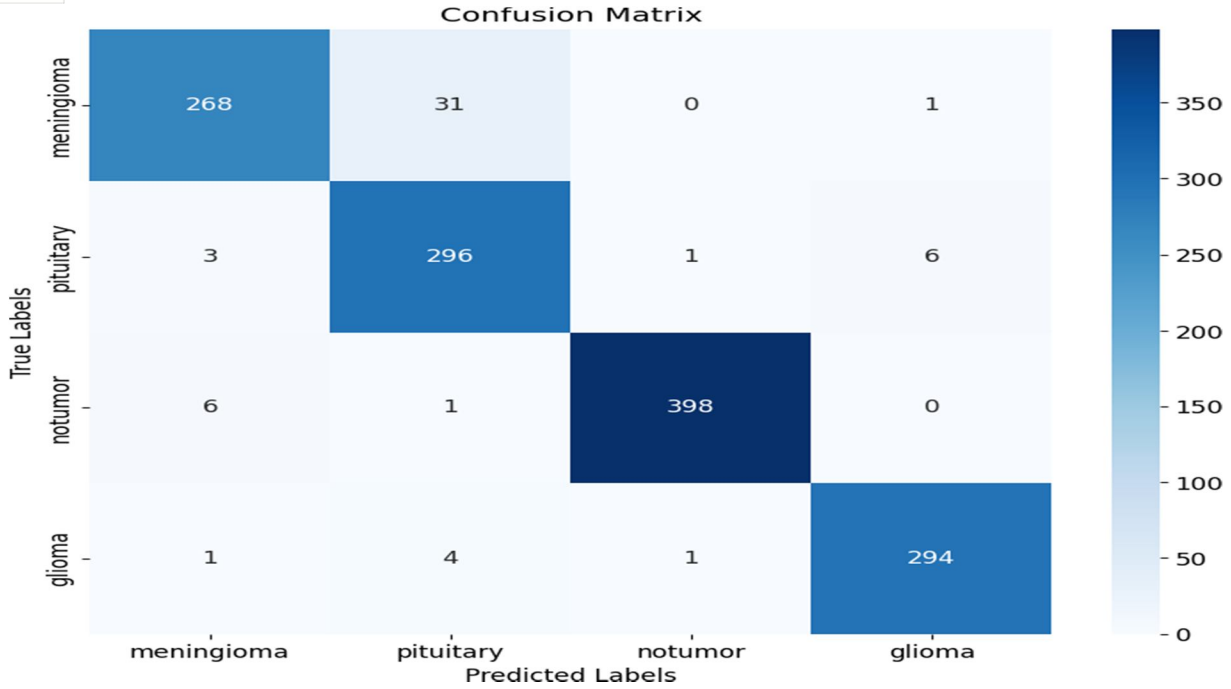


Figure 14. Balanced accuracy confusion matrix

The prediction made from a single image was one aspect of the work that we wanted to be unique. We developed a method that takes in a single image and, after the model has been trained and assessed, returns a percentage indicating the likelihood that a tumor is present in both the processed image and the MRI. And figure 13 shows the outcome that was obtained for a random image.

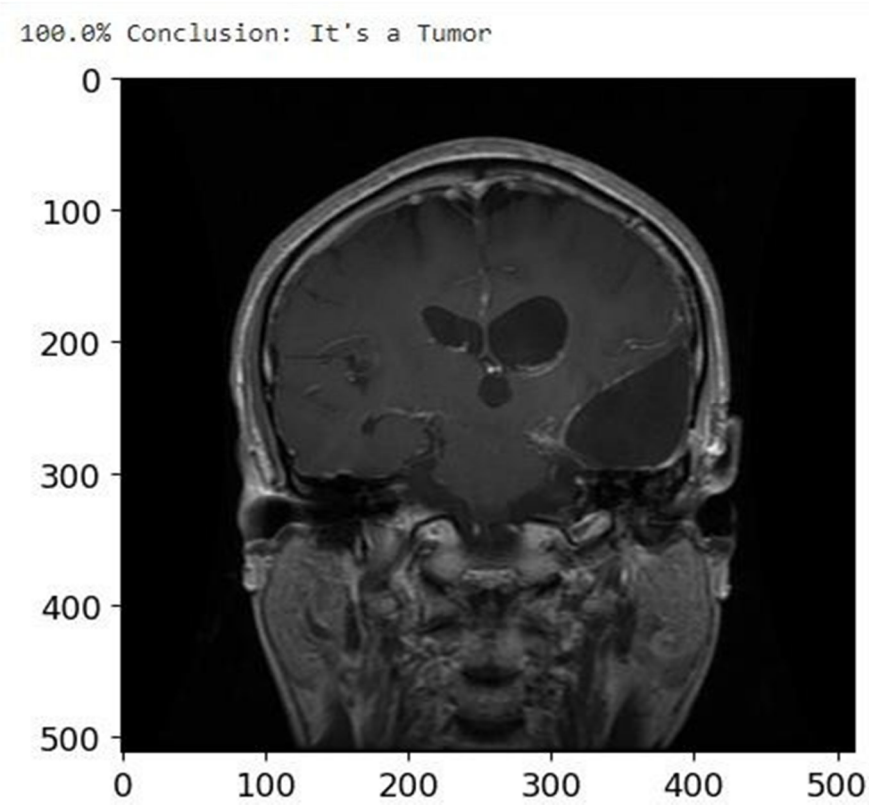


Figure 15 Specific image prediction

IV. DISCUSSION

The findings of this study show how the PIL-VGG16 technique may successfully identify and categorize brain cancers from MRI data. High classification accuracy has been achieved because to the model's ability to learn and extract significant features thanks to the combination of PIL's image preprocessing capabilities and VGG16's deep learning architecture. The outcomes obtained are encouraging and show this method's potential for use in clinical settings. This study's use of a sizable dataset of 7023 MRI scans, which enhances the robustness and generalizability of the proposed model, is one of its major strengths. The addition of several tumor forms, such as glioma, meningioma, pituitary, and non-tumor instances, improves the dataset's diversity and offers a thorough illustration of the range of different types of brain tumors. But there are some restrictions that must be understood. First off, even if the achieved classification accuracy is great, further analysis and validation on a different dataset are required to guarantee the model's dependability. Additionally, the current study only considers image-based tumor detection and ignores additional clinical data or patient-specific factors that can improve the model's accuracy and clinical applicability.

But when comparing the results with similar work from this are we can see that our model has achieved above average results in this problems solution which can make a statement to further keep evolving this model results. When comparing with work of Chandra [3] and hers colleagues we can see that they used the VGG16 algorithm as well, but their dataset only consists of around 260 MRI images and they achieved accuracy of 84% which can be seen from the table below, as our dataset consisted of 7000 MRI images. In Dheiver's [4] work we can see the utilization of the VGG16 algorithm as well, but he didn't disclose the layer structure of the model nor the size of the dataset, but his results yielded results of 80% and 88% which are quality results but most likely the models layer structure can be improved as well as the datasets size. The work of Younis [5] and her colleagues achieved very high results 89% using CNN, 97% using VGG16 and 91% using an ensemble model, from this we can see the results for the VGG16 are in similar values with our result even though they used a smaller dataset of 253 images with a good model they achieved very high accuracy results.

Table 1. Comparison of results using similar models

Authors	Model	Accuracy
Dheiver Francisco Santos [4]	VGG16	88%
Chandra, K. Sarath; Priya, A. Sai; et al. [3]	VGG16	84%
Younis, A; Qiang, L; Nyatega, et al [5]	CNN VGG16 Ensemble	89% 97% 91%

As per the works that have used the same dataset as in our research the following work of Md. Monirul Islam [6] and his colleagues have conducted research on the same dataset and have obtained quite substantial results they have implemented four models: InceptionV3, VGG19, DenseNet121 and MobileNet with their respective accuracy's 98.76%, 98.97%, 99.12% and 99.60% on the training datasets, while on the testing dataset the accuracy's were the following InceptionV3 (96.8%), VGG19 (95.5%), DenseNet121 (97.41%) and MobileNet (98.4%), from the following data we can see that two of the models outperformed ours model. DenseNet121 is a deep neural network architecture comprising 121 layers, characterized by dense connectivity between layers, enabling feature maps concatenation and fostering effective gradient flow and feature reuse. In contrast, MobileNet optimizes parameter efficiency by utilizing depth-wise separable convolutions, dividing conventional convolutions into depth-wise and point wise convolutions, resulting in a highly efficient network design with fewer parameters. As per the image processing they mention using image augmentation to create altered duplicates of the images, which allows the model to be trained on more image variants, but the most interesting part of this work comes from results investigation as per the author the best performing model with minimal loss was InceptionV3 which architecture combines different convolutions, along with max pooling, average pooling, batch normalization, dropout, factorized convolution, and auxiliary classifiers, achieving state-of-the-art performance on various image recognition benchmark datasets. Another interesting research on the same dataset was performed by Gomez-Guzman [7] and his colleagues they have as well tested multiple models with the following having the best results in their research: InceptionV3 (97.1%), ResNet50 (96.97%) and InceptionResNetV2 (96.78%). In their work they have used InceptionV3 as one of the models they mention that the model consists of 48 layers. ResNet-50 is a 50-layer deep convolutional neural network architecture. Gradients can flow through the network more directly thanks to these connections, which makes it possible to train very deep networks and increases accuracy across a range of computer vision tasks.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

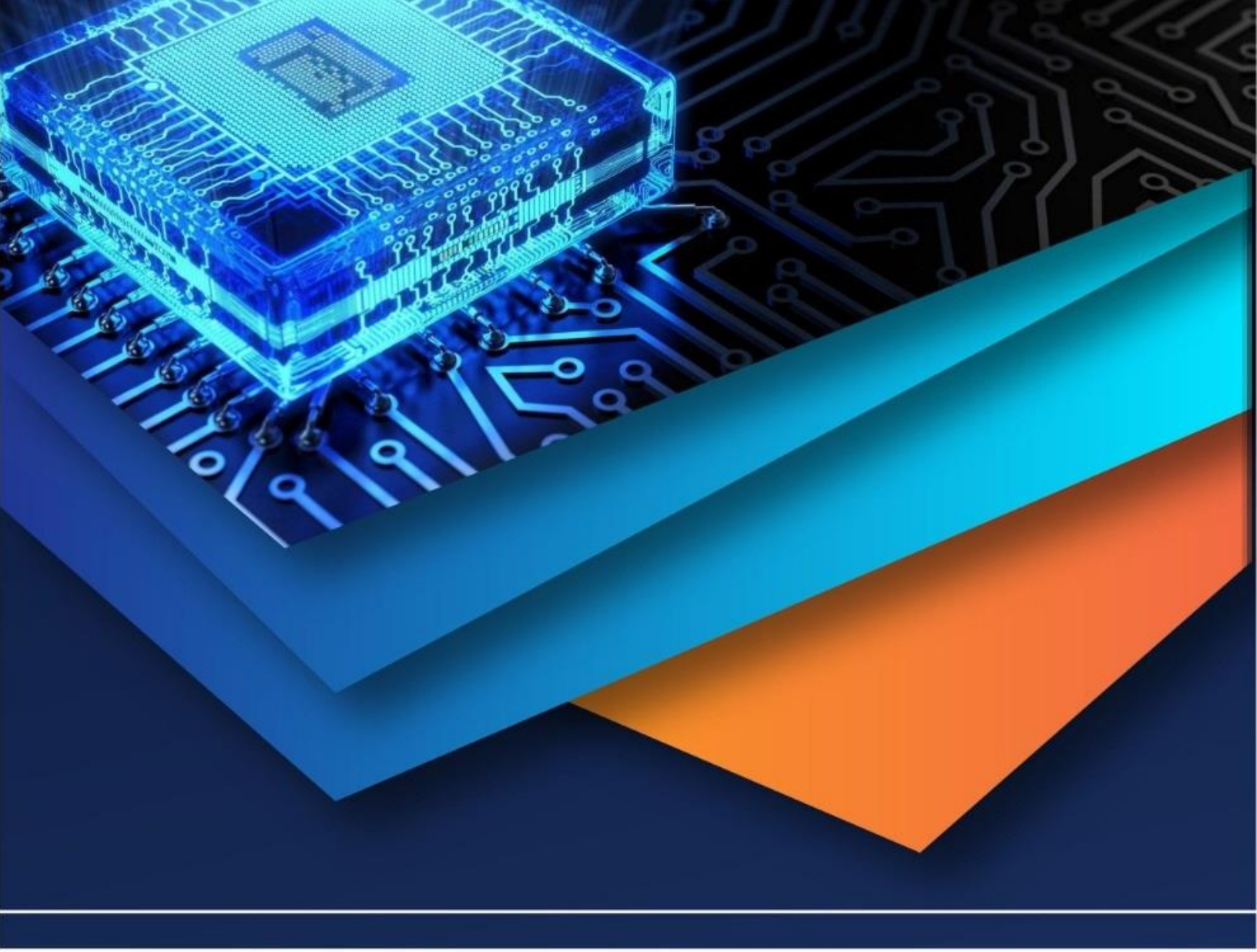
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

On the other hand, InceptionResNetV2 is a deep convolutional neural network that combines modified Inception blocks with residual connections and filter expansion layers, utilizing batch normalization to enhance its performance. Before the models can be trained the images are preprocessed with adjustment of the brightness, rotation and zoom to prepare the images for the mentioned models. The results of the mentioned works are displayed in Table 2.

Table 2. Comparison of results using the same dataset

Authors	Model	Accuracy
P. B. M. Md. Monirul Islam [6]	InceptionV3 VGG19 DenseNet121 MobileNet	96.8% 95.5% 97.41% 98.4%
M. Gómez-Guzmán, L. Jiménez-Beristaín, E. García-Guerrero, et al. [7]	InceptionV3 ResNet50 InceptionResNetV2	 97.1% 96.97% 96.78%

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)