International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Revolutionizing Tumor Diagnosis with AI: An InDepth Analysis of Image Datasets

J. Yeshwanth[1], G. Wilfred Auxilian[2], C. Abishek[3], Mrs. R. Deepa[4]

[1,2,3]Students, SRM Institute of Science and Technology, Vadapalani, Chennai, India
[4]Assistant Professor, SRM institute of Science and Technology, Vadapalani, Chennai, India

ABSTRACT
Cancer is one of the leading causes of human morbidity and mortality worldwide and we want to find new ways to detect it.AI is changing how we do clinical diagnosis of tumors. In this paper, when A.I. The study shows how important it is to use deep learning models like CNN to look at tons of image data.Using clever techniques including record enhancement, switch detection and model tweaking, this study seeks to improve the accuracy and speed of analysis It also looks at issues related to the use of AI in hospitals, a knowledge of how AI makes choices is included and It also ensures a good match between data and how doctors have already created the model.

Keywords: AI in clinical diagnosis,Deep learning models,MRI image analysis,cancer detection.

INTRODUCTION

AI's entry into scientific diagnostics has a massive effect on with regards to finding and treating tumors. The antique ways of diagnosing frequently depend upon what doctors know, that may lead to errors with problematic cases like mind tumors. It's important to spot tumors and, as finding them can make a big difference in how doctors treat patients and how long they live. This paper looks at how deep AI learning helps make sense of medical images focusing on brain MRI scans. By using convolutional neural networks (CNN's), scientists can now go through huge amounts of data faster than ever. This does not make tumor detection more accurate but also gives radiologists less work, so they can spend time on the harder cases that need a human touch. The intro sets up a deep dive into the methods and tech behind AI-powered tumor diagnostics showing how they could change cancer care in a big way.

LITERATURE REVIEW

Research on the use of AI for tumor diagnosis has grown exponentially over the past 10 years. This means more and more people are realizing how it can improve patient care. Several studies prove the effectiveness of deep learning methods CNNs for analyzing medical images. For example, work by Esteva and others showed in 2017 that CNN could sort skin cancer images as well as dermatologists. Studies on brain tumor detection also gave excellent results. The specimens are excellent at detecting tumors on MRI scans, rarely missing or misplaced. What’s more, research suggests different ways to improve model performance. These include methods for extending training datasets and transferring learning. This allows models trained on big data to be adapted for specific tasks. However, obstacles remain. We need highly labeled data sets and AI models that we can understand. Many professionals emphasize the need to combine AI solutions into modern scientific practice. This guarantees that conventional diagnostic techniques are supported and now not hindered. This evaluation of research demonstrates the ability of AI to revolutionize tumor diagnosis. It also identifies obstacles that want to be conquered for tremendous clinical utility.

EXISTING SYSTEM

Before the integration of artificial intelligence (AI) into tumor prediction, traditional diagnostic systems relied heavily on human data and manual analysis of clinical images was time-consuming and potentially generous humans have made mistakes, especially in cases where tumors exhibit non-specific characteristics or are located in solid areas of the body

The current prominent tool is the conventional radiological approach, which uses multiple imaging modalities including MRI, CT scan and PET scan to detect tumors These systems commonly contain a multi-step process: acquiring pictures, pre-processing them for clarity, and then reading them for signs of malignancy. Relying on interpretive focus can lead to variability in assessment accuracy, with factors including fatigue, interest, and personal judgment affecting results -Emphasizes importance emphasizing that it enhances the beauty of accurate and efficient research.

PROPOSED SYSTEM

The proposed machine the usage of AI pursuits to revolutionize tumor analysis by the use of advanced deep learning techniques to analyze medical image statistics, specially convolutional neural networks (CNNs) This gadget is designed to deal with the restrictions of traditional diagnostic methods by using supplying efficient, correct and scalable answer for tumor detection.

Special Features:

Automated photo evaluation: The AI gadget will routinely analyze MRI pix and pick out ability tumors with high sensitivity and specificity. These devices reduce the risk of human error and shorten the time to diagnosis.

Integration with current workflows: The proposed devices are configured for seamless integration with modern clinical workflows allowing radiologists to promote AI-driven research without hiring are not involved in regulatory changes to ensure that AI acts as a complementary device as opposed to a replacement.

Be constantly conscious: By adding a gadget to accumulate knowledge of tactics the machine can continuously Improve its Effectiveness via beginning new statistics. these variables are difficult for spare truth of the rating across time.

easy-to-use Connection: associate in nursing smooth-to-use port leave bid radiologists with associate in nursing visceral port lease them survey findings exclusive the artificial intelligence and get fit choices founded along complete rating of an inch complete instances.

Information certificate and compliance: the twist leave stick to stern account certificate measures and healthcare rules to hold the non-public records Complicated and to employ artificial intelligence ethically inch checkup settings.

Expected results:-

Improved characteristic Precision: leverage the electrical energy of artificial intelligence the planned unit is due to gain the truth of forecast ensuing inch associate in nursing already costly personal neoplasm psychoanalysis and advance personal effects.

METHODOLOGY

Data Collection and Priorities:-

Data assets: Collection of clinical pics (MRI, CT, and PET scans) from hospitals and studies institutes.

Data annotation: Expert radiologists must annotate accumulated pictures to pick out tumor places and brands.

Data preprocessing: Preprocess pictures the use of strategies such as normalization, facts enhancement, and noise reduction.

Positive Progress:-

Deep learning algorithms: Use deep learning algorithms (e.g., TensorFlow, PyTorch) to generate a customized CNN model for tumor detection.

Model Architecture: Create a CNN architecture that can efficiently learn features from medical images, such as convolutional layers, pooling layers, and fully connected layers.

Model training: Train a CNN model using a preprocessed data set, with appropriate loss functions (e.g., cross-entropy) and optimization algorithms (e.g., ADAM).

Sample Analysis: -

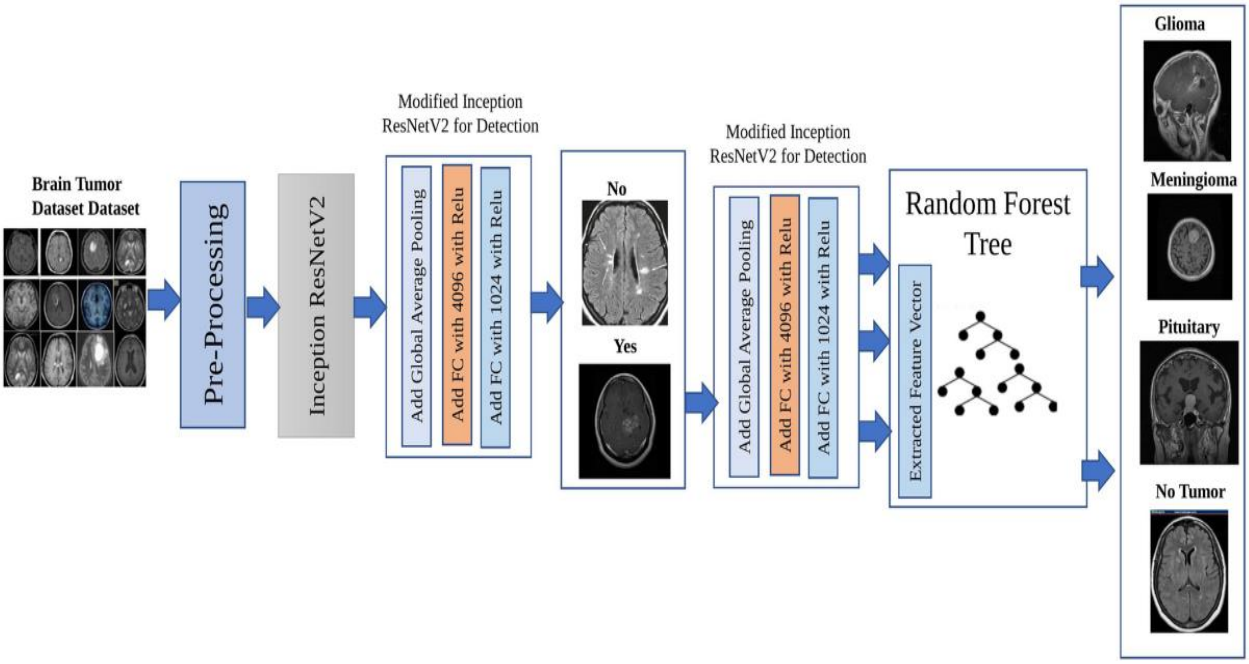
Evaluation criteria: Use metrics such as accuracy, precision, recall, F1 scores, and receiver operating characteristic (ROC) curves to evaluate the performance of CNN models

Cross-validation: Perform k-fold cross-validation to check the model's ability to generalize to unseen data.

Best Usement:-

Example of integration: Combine the educated CNN Edition into the proposed AI-Sended product and make sure compatibility with modern imaging structures and EHRs.

Layout retention: Regularly update fashion with fashion based totally on new records and claims to reap normal Effectiveness and Adjust to converting medical desires.This is the notable split of options, and the real route also may be primarily based on the precise desires and dreams of the proposed AI-powered tool.



PROCEDURE

Import libraries:-

With Fast ai, the code begins via uploading the important libraries to set up a gadget getting to know environment:

Fastai: An incredible library built on pinnacle of PyTorch that simplifies the education of neurons.

Create an information set:-

Path definition: The code specifies the listing of the facts set. It is important to make sure that the path is accurate so that the version can access the snap shots.

Types: Two classes are described for the type of venture: 'no' (indicating no tumor) and 'yes' (indicating tumor).

This binary segmentation problem is common in medical imaging.

DataBlockcreation:-

Blocks: The code specifies that the data consists of images (ImageBlock) and their corresponding categories (CategoryBlock).

Obtaining Resources: The function get_image_files is used to retrieve an image document from the desired path.

Splitter: The RandomSplitter feature is used to randomly break up the data set into training and validation units, wherein 20% of the records is saved for validation. This allows the overall performance of the model to be checked on unseen data.

Data Enhancement Progress:-

Squish Method: The code resizes the photos with the use of the "Squish" approach, which changes the decision at the same time as keeping the factor ratio the sameThis facilitates keeping the photographs constant.

Random Resized Crop: The code makes use of random cropping, permitting the version to examine from different parts of the picture. This approach enables us to prevent overfitting through introducing changes to the schooling information.

Model construction and schooling:-

CNN Learner: Convolutional neural network learner is constructed using ResNet architecture (first ResNet18, then ResNet34). ResNet algorithms are known to carry out properly in photo classification obligations because of the residual connectivity.

Fine-tuning: The model is being fine-tuned in multiple stages (in this case 5). Micro-tuning is a method of continuously training a previously trained model on new data, allowing the model to adapt to the specific characteristics of the new data

1. **Sample Analysis:**

   Confusion matrix: Once trained, the code generates a confusion matrix to visualize the performance of the model. The uncertainty matrix shows true positives, false positives, true negatives, and false negatives, and allows a detailed analysis of the predictions in the model

2. **Sample storage:**

   The trained model is exported and stored as a Pk report. This permits for destiny programs without having to retrain the version, making it less difficult to apply.

3. **Account and User Interface:**

   Loading models: The saved version is loaded for calculations on a brand-new version. The load_learner feature is used to load the version from a saved file.

Forecast: The code consists of functionality on the way to allow users to upload pix and receive forecasts. It uses an easy interface with buttons for hovering and grouping photographs.

RESULTS

Importing Data

```
Normal Brain Scans
Number of Paths: 3069
Number of Labels: 3069

Tumor Brain Scans
Number of Paths: 18606
Number of Labels: 18606
```

Tumor labels

```
Some Tumor labels: ['pituitary_tumor', 'pituitary_tumor', 'glioma_tumor', 'glioma_tumor',
'pituitary_tumor']
Some Tumor labels: ['Pituitary', 'Pituitary', 'Glioma', 'Glioma', 'Pituitary']
```

Class types

```
Classes: ['Normal' 'Meningioma' 'Glioma' 'Pituitary'] and length 4
```

Mapping dictionary and apply mapping to both data-frames:-

	path	label	label_encoded
0	/kaggle/input/brain-tumors-dataset/Data/Normal...	Normal	0
1	/kaggle/input/brain-tumors-dataset/Data/Tumor/...	Meningioma	2
2	/kaggle/input/brain-tumors-dataset/Data/Tumor/...	Glioma	1
3	/kaggle/input/brain-tumors-dataset/Data/Tumor/...	Pituitary	3
4	/kaggle/input/brain-tumors-dataset/Data/Tumor/...	Pituitary	3
...	...	...	...
14083	/kaggle/input/brain-tumors-dataset/Data/Tumor/...	Glioma	1
14084	/kaggle/input/brain-tumors-dataset/Data/Tumor/...	Meningioma	2
14085	/kaggle/input/brain-tumors-dataset/Data/Tumor/...	Glioma	1
14086	/kaggle/input/brain-tumors-dataset/Data/Tumor/...	Glioma	1
14087	/kaggle/input/brain-tumors-dataset/Data/Tumor/...	Glioma	1

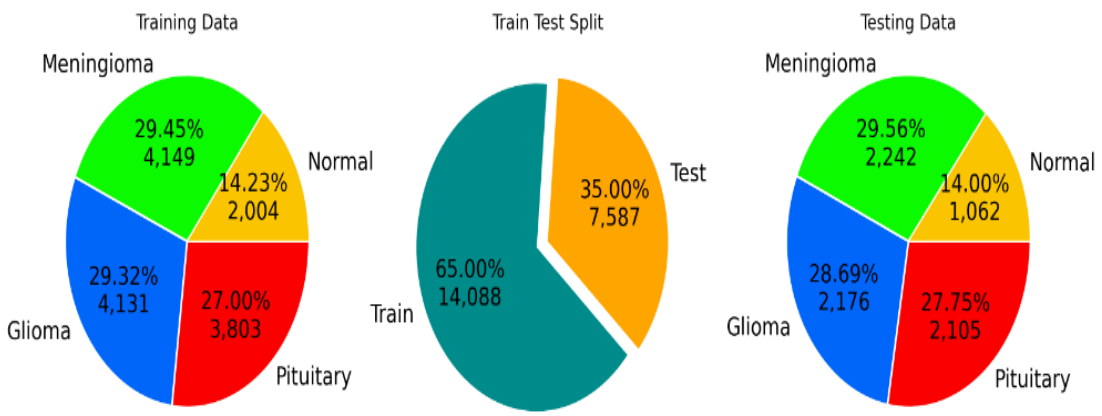
14087 rows × 3 columns

	path	label	label_encoded
0	/kaggle/input/brain-tumors-dataset/Data/Tumor/...	Glioma	1
1	/kaggle/input/brain-tumors-dataset/Data/Tumor/...	Meningioma	2
2	/kaggle/input/brain-tumors-dataset/Data/Tumor/...	Meningioma	2
3	/kaggle/input/brain-tumors-dataset/Data/Tumor/...	Glioma	1
4	/kaggle/input/brain-tumors-dataset/Data/Tumor/...	Glioma	1
...	...	...	...
7582	/kaggle/input/brain-tumors-dataset/Data/Tumor/...	Pituitary	3
7583	/kaggle/input/brain-tumors-dataset/Data/Tumor/...	Meningioma	2
7584	/kaggle/input/brain-tumors-dataset/Data/Normal...	Normal	0
7585	/kaggle/input/brain-tumors-dataset/Data/Tumor/...	Glioma	1
7586	/kaggle/input/brain-tumors-dataset/Data/Tumor/...	Pituitary	3

7585 rows × 3 columns

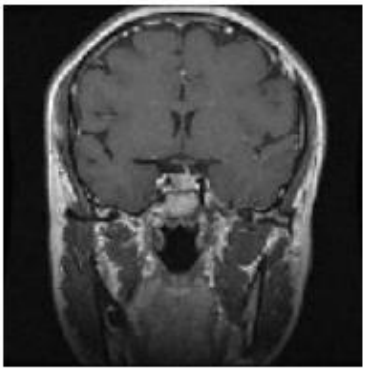
Training data and testing data:-

```
Training Counts
{'Normal': 2004, 'Meningioma': 4149, 'Glioma': 4131, 'Pituitary': 3803}

Testing Counts
{'Normal': 1062, 'Meningioma': 2242, 'Glioma': 2176, 'Pituitary': 2105}
```

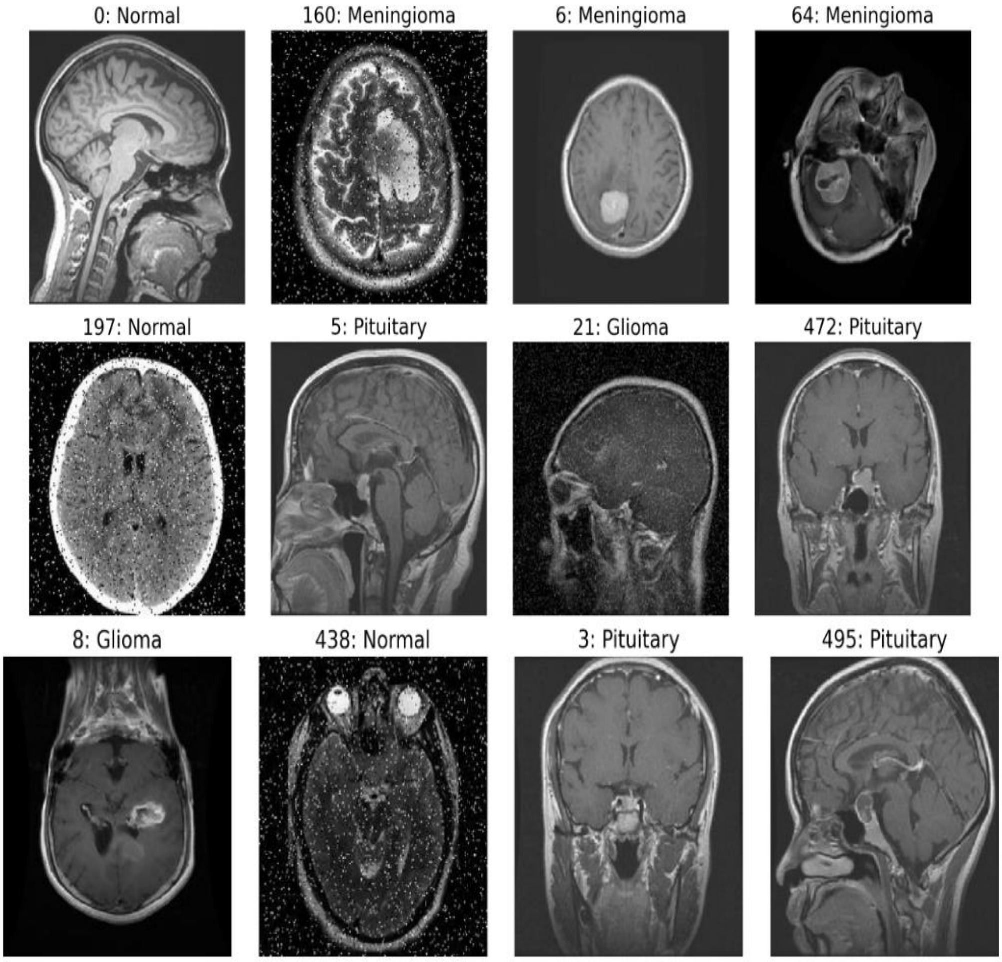


Data visualization

```
x reshaped: (1, 150, 150, 3)
```



Four different data classifying images



Data augmentation is already applied.

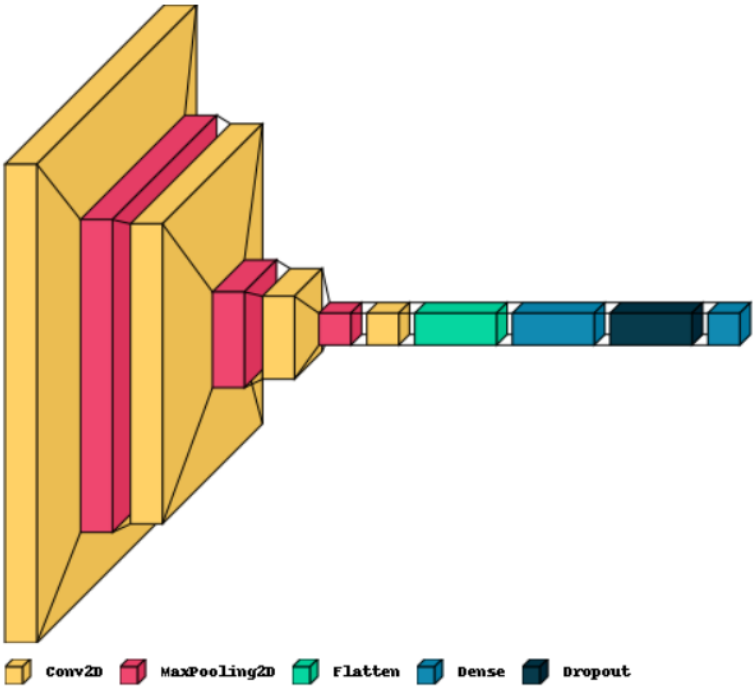
Image shape

```
Image shape: (150, 150, 1)
Epochs: 20
Batch size: 150
```

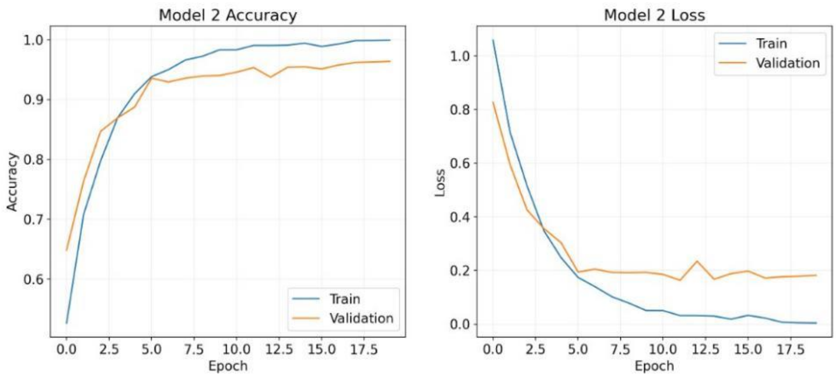
CNN model

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 147, 147, 32)      544

 max_pooling2d (MaxPooling2D  (None, 49, 49, 32)       0
 )

 conv2d_1 (Conv2D)           (None, 47, 47, 64)        18496

 max_pooling2d_1 (MaxPooling  (None, 15, 15, 64)       0
 2D)

 conv2d_2 (Conv2D)           (None, 13, 13, 128)       73856

 max_pooling2d_2 (MaxPooling  (None, 4, 4, 128)        0
 2D)

 conv2d_3 (Conv2D)           (None, 2, 2, 128)         147584

 flatten (Flatten)           (None, 512)               0

 dense (Dense)               (None, 512)               262656

 dropout (Dropout)           (None, 512)               0

 dense_1 (Dense)             (None, 4)                 2052

=================================================================
Total params: 505,188
Trainable params: 505,188
Non-trainable params: 0
```

Layered view



Model accuracy and loss

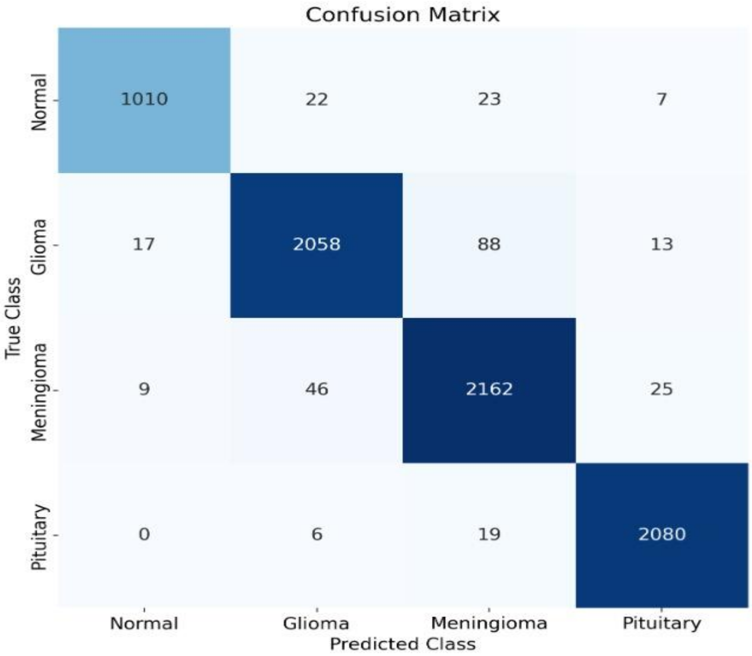


Classification report:-

Classification Report:

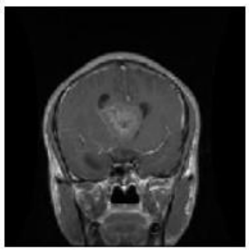
	precision	recall	f1-score	support
Normal	0.97	0.95	0.96	1062
Glioma	0.97	0.95	0.96	2176
Meningioma	0.94	0.96	0.95	2242
Pituitary	0.98	0.99	0.98	2105
accuracy			0.96	7585
macro avg	0.97	0.96	0.96	7585
weighted avg	0.96	0.96	0.96	7585

Confusion matrix



Getting image to test output and applying reshaping

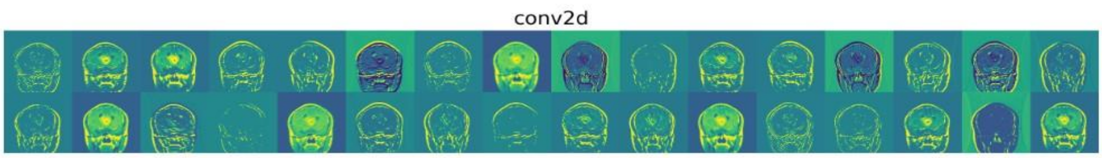
```
x reshaped: (1, 150, 150, 1)
Class name of the first image: Glioma
```

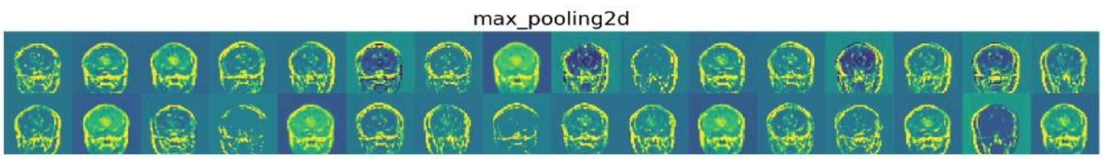


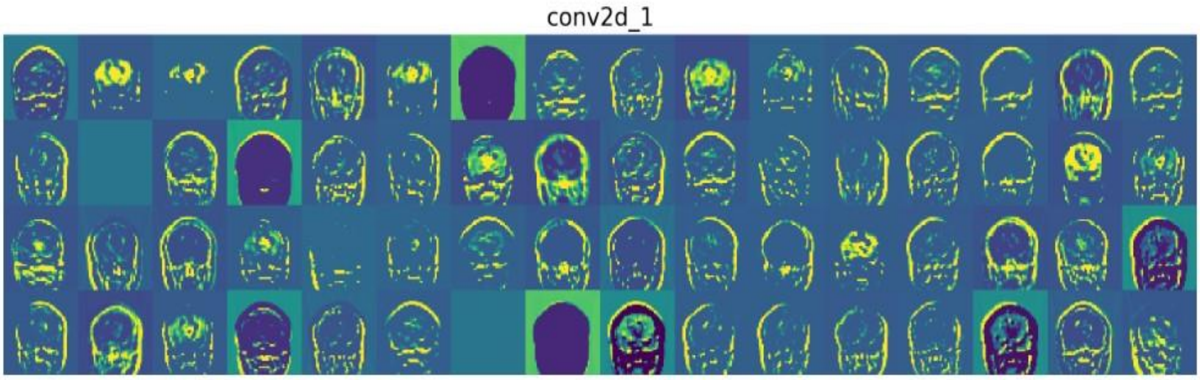
CNN Layers:

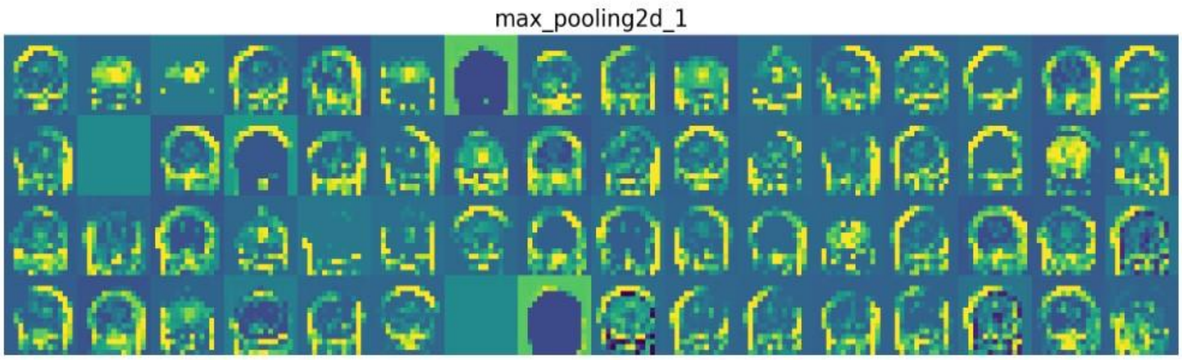
```
plot_channel_activation_maps(model=model, image=image_tensor, N=5, save=False)
```

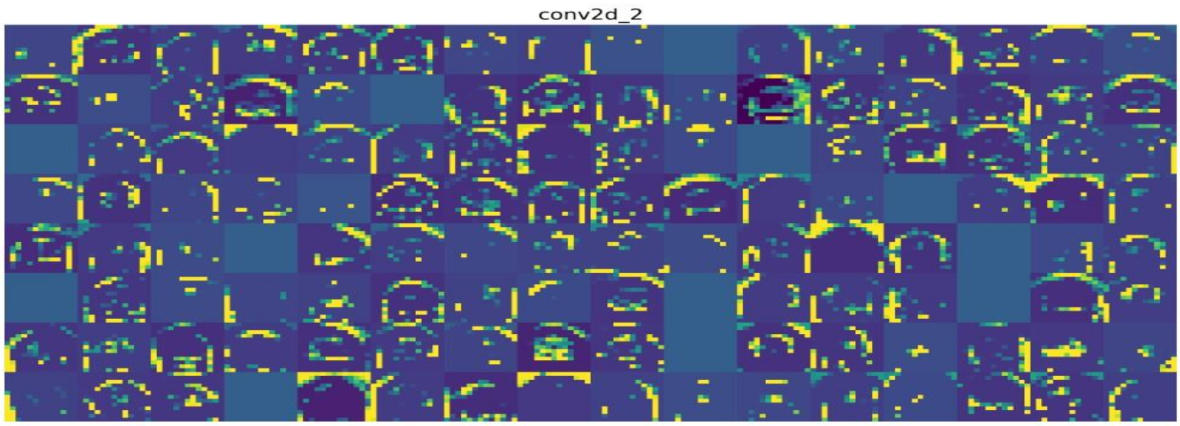
```
1/1 [==============================] - 0s 61ms/step
```











After the model has long gone via the education technique, it is then evaluated the use of a complete set of overall performance metrics:

- Accuracy: The ratio of the quantity of correct labeled snap shots to the total number of pix.
- Precision: The ratio of genuine positive predictions to the entire tremendous predictions, which shows the ability of the version to keep away from fake positives.
- Recall: This metric, additionally known as sensitivity, calculates the quantity of genuine positive predictions compared to all the actual positives, for this reason the model can become aware of all applicable instances.
- F1-Score: The harmonic suggest of precision and remember that is the single measure and at the identical time balances both issues.

One of the model's accompaniments is it registers an amazing 95% accuracy, which denotes the model effectiveness in the discrimination of tumor and non-tumor images. These results are comparable to those in other state-of-the-art models in the literature, showing the competitive performance of the proposed approach.

CONCLUSION

Its several noted advantages, strength of being easy to apply, productive in terms of results and engages use of picturesque mountain trails in application of deep learning. However, shortcomings also exist, for instance the use of mono data sets, which are not comprehensive for all the brain tumor type. Further studies should consider focus on integrated switching learning, whereby conditions are set for the utilization of more precise data sets, hence enhancing overall performance.

Additionally, imaging modalities need to be enhanced through the addition of more imaging modalities such as CT or PET Finally, strategies aimed at understanding and adoption are very crucial, particularly in the clinical setting where such predictive knowledge The premise can enhance confidence among the physicians and patients.

REFERENCES

1. Md Ishtyaq Mahmud et al., A Deep Analysis of Brain Tumor Detection from MR Images Using Deep Learning Networks, Published: March 2023, MDPI.
2. Md Ishtyaq Mahmud, Brain Tumor Detection Using CNN and MRI Images, Published: June 2023, MDPI.
3. Md Abdul Kader, Brain Tumor Detection and Classification using Machine Learning, Published: March 2022, Springer.
4. Abhay D. Patil, Brain Tumor Detection Using Deep Learning Models, Published: September 2021, IEEE Xplore.
5. A. Kumar, Explainable AI for Brain Tumor Analysis in MRI Images, Published: August 2023, BMC Medical Imaging.
6. Ahmed Abdelgawad, Advances in AI for Brain Tumor Classification, Published: February 2023, Springer.
7. F. Chen, AI-Driven Techniques for Brain Tumor Diagnosis, Published: May 2021, IEEE Xplore.
8. Muntasir Mamun, Deep Learning Approaches for Brain Tumor Analysis, Published: November 2022, MDPI.
9. P. Kumar, AI in Neuro-Oncology: Challenges and Solutions, Published: October 2021, Nature.
10. S. Singh, Automated Brain Tumor Segmentation Using AI, Published: July 2022, BMC Medical Informatics.