Brain Tumor Classification From MRI Images Using ResUNet

Eshita Pradhan
Department of Computer Science and Engineering
IIIT Naya Raipur
Raipur, India
eshita21100@iiitnr.edu.in

Riya Yadav
Department of Computer Science and Engineering
IIIT Naya Raipur
Raipur, India
riya21100@iiitnr.edu.in

Abstract—**Brain tumor diagnosis and treatment are complex endeavors heavily reliant on accurate interpretation of Magnetic Resonance Imaging (MRI) scans. However, the manual assessment process is time-consuming and susceptible to errors, necessitating innovative solutions. Enter Convolutional Neural Networks (CNNs) based on the ResUNet architecture—a cutting-edge approach proposed to streamline tumor detection and segmentation. Distinctive in its design, this methodology harnesses the power of residual units to optimize deep network training. By incorporating recurrent residual convolutional layers, it ensures robust feature accumulation, thereby enhancing the representation of tumor characteristics. Despite the inherent challenges posed by the diverse spatial and structural variations of brain tumors and their surroundings, this novel approach achieves an outstanding 95% accuracy rate on test data. This advancement not only expedites diagnosis but also alleviates the burden on neurospecialists and radiologists, who are often in short supply. By automating segmentation and classification tasks, it empowers healthcare providers to deliver more timely and accurate care, ultimately improving patient outcomes.**

Index Terms—**Brain tumor, MRI, CNNs, ResUNet, architecture, Diagnosis, Segmentation**

I. INTRODUCTION

A Brain tumour is among the most severe conditions that can strike people of any age, including adults and children. Approximately 85 to 90 percent of primary Central Nervous System (CNS) tumour cases are comprised of these tumours. About 11,700 people are diagnosed with brain tumours each year. Malignant brain or central nervous system tumour patients still have a poor prognosis, with a 34% male 5-year survival rate and a 36% female 5-year survival rate. Different types of brain tumours can be distinguished, such as pituitary tumours, benign tumours, and malignant tumours. Precise diagnosis and careful treatment planning are essential for saving lives and providing patients with better care. With its ability to produce detailed images that are essential for diagnosis, Magnetic Resonance Imaging (MRI) is the gold standard for brain tumour detection. However, because of the large volume of data generated, interpreting these images is a challenging task. The complexities surrounding brain tumours, which can differ widely in size, shape, and location, present a challenge for radiologists responsible for reviewing MRI scans. The probability of errors in manual examination is increased by this complexity. Additional difficulties arise from the nature of brain tumours, whose various manifestations make it challenging to comprehend their traits and behaviour in their entirety. As such, the analysis frequently necessitates the expertise of licenced neurosurgeons. The situation is made worse in many developing nations by a lack of qualified medical personnel and a lack of knowledge about brain tumours. Due to this scarcity, producing accurate MRI reports is difficult and time-consuming, which delays diagnosis and treatment. An automated cloud-based system might be very helpful in resolving these problems. A system like that would improve tumour detection accuracy, simplify the diagnostic process, and lighten the workload for medical personnel. This strategy seeks to get around existing constraints and greatly enhance patient outcomes by utilising cutting-edge technology.

II. LITERATURE REVIEW

The use of machine learning techniques in medical imaging has attracted a lot of attention lately, especially when it comes to the identification and categorization of brain tumours. TensorFlow is a framework that Abadi et al. (2016) presented. It allows for large-scale machine learning on heterogeneous distributed systems and offers a stable foundation for creating and implementing sophisticated algorithms. The foundation for further studies in the topic was created by this effort. Abraham and Khan (2019) presented a unique localised Tversky loss function that showed enhanced attention in lesion segmentation utilising the U-Net architecture. Their method demonstrated improvements in segmentation accuracy, especially in defining lesion boundaries, which will lead to more accurate diagnostic instruments.

Anithadevi and Perumal (2016) introduced a hybrid segmentation method that combines machine learning and conventional techniques to detect brain tumours in MRI images. Their method attempted to overcome the difficulties caused by the complexity of brain tumours and the variation in picture properties in order to improve the precision and effectiveness of tumour segmentation. Bakas et al. (2017) provided segmentation labels and radiomic features for pre-operative scans of the TCGA-LGG collection, contributing valuable resources for research and development in brain tumor analysis. Their work facilitated the benchmarking of segmentation algorithms

and the exploration of predictive biomarkers in brain tumor imaging.

Various studies have explored the utility of deep learning architectures in brain tumor classification and segmentation. Khan et al. (2020) proposed a convolutional neural network (CNN) approach for brain tumor classification from MRI images, demonstrating promising results in automated diagnosis. Similarly, Rai and Chatterjee (2020) introduced the Lu-Net model for detecting brain abnormalities, showcasing the potential of deep learning in medical image analysis. With its introduction by Ronneberger et al. (2015), the U-Net architecture has become a popular option for biomedical picture segmentation tasks, such as brain tumour segmentation. Studies have shown that its architecture, which includes symmetric contracting and expanding routes, allows for accurate tumour region delineation from MRI scans.

Salçin et al. (2019) demonstrated the usefulness of object detection approaches in medical imaging analysis by investigating the use of quicker R-CNN in the detection and classification of brain tumours from MRI images. Their research serves as an excellent example of the multidisciplinary use of computer vision techniques in medical diagnosis. With its introduction by Ronneberger et al. (2015), the U-Net architecture has become a popular option for biomedical picture segmentation tasks, such as brain tumour segmentation. Studies have shown that its architecture, which includes symmetric contracting and expanding routes, allows for accurate tumour region delineation from MRI scans.

Salçin et al. (2019) demonstrated the usefulness of object detection approaches in medical imaging analysis by investigating the use of quicker R-CNN in the detection and classification of brain tumours from MRI images. Their research serves as an excellent example of the multidisciplinary use of computer vision techniques in medical diagnosis. When taken as a whole, these studies demonstrate the continuous attempts to progress the area of medical image analysis, with an emphasis on enhancing accessibility, effectiveness, and accuracy of diagnosis. Through the utilisation of cutting-edge algorithms and technology, scientists endeavour to tackle the intricate problems presented by brain tumours, with the ultimate goal of improving patient outcomes and treatment.

III. PROPOSED METHODOLOGY

The Cancer Imaging Archive (TCIA), notably the TCGA-LGG dataset, provided the MRI images used in this project. These details are available on the TCIA website. Hand-drawn segmentation masks for fluid-attenuated inversion recovery (FLAIR) are included with these images, which are essential for defining the borders of the tumour. One noteworthy feature of FLAIR sequences is their well-known capacity to generate excellent digital tumour representations (Rucco, Viticchi, Falsetti, 2020).

This dataset, which includes FLAIR sequences and genomic cluster data, is made up of images from 109 patients in The Cancer Genome Atlas (TCGA) LGG collection. The dataset, which includes more than 20 image pairs and matching
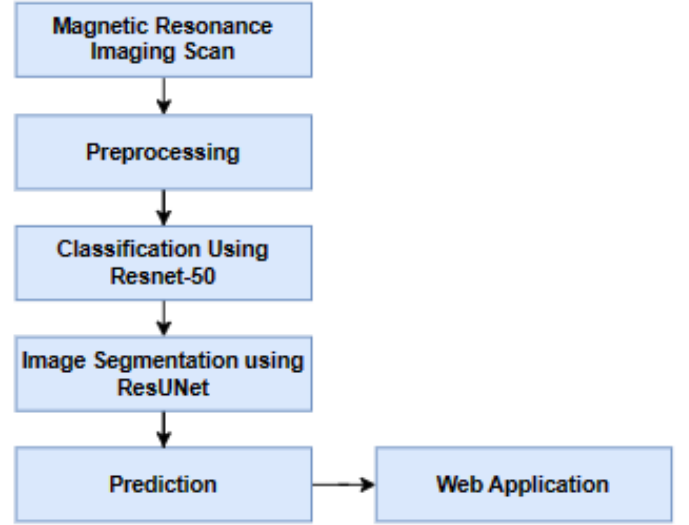

Fig. 1. Proposed Model

masks for every patient, is extensive and publicly available, highlighting its importance in furthering this field of study. The MR images are displayed as three-channel RGB images with 256 by 256 pixels in each channel. Mask images of the same size but with a single channel accompany these pictures, which help with accurate tumour segmentation (Yang et al., 2021).

Researchers can investigate and develop in the field of medical imaging analysis with the help of this freely accessible open-source dataset, which will help advance the identification and characterization of tumours.

Fig. 2. Distribution of brain with and without tumour

A. Data Preprocessing

According to Bakas et al. (2017), preprocessing is the cornerstone of data improvement and analysis preparedness. First, co-registration and skull-stripping were applied to the provided

images. An experienced neuroradiologist then carefully corrected the tumour segmentation labels that had been produced by an automated hybrid generative method. The resulting images are rich in features, including volumetric, morphologic, histogram-based, textural, and intensity parameters.

Three RGB channel colour maps are used to represent brain data, which is measured in voxels, which are similar to 3D pixels. In order to enable pixel-level segmentation (i.e., distinguishing between the presence and absence of tumours), all images and their masks were cropped to remove any noise. The CLAHE algorithm (Zuiderveld, 1994) was used before cropping to improve image contrast, which was especially useful for images with low contrast. After identifying image edges with Canny edge detection (Canny, 2002), a dilation operation was performed to remove small noise regions.

Finally, extreme points and contours were found and used to crop the images successfully during the cropping phase (Rai and Chatterjee, 2020). The resulting MRI image dataset, known as post-cropping, demonstrates how well these preprocessing procedures worked to improve the quality of the data in preparation for further analysis.
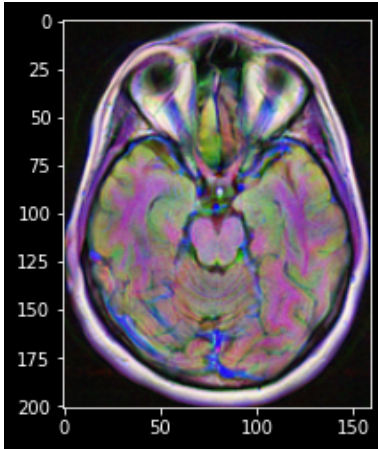

Fig. 3. Cropped Image

B. Creating Train, Test and Validation set

Within this project, the dataset undergoes a meticulous division into three crucial segments: the train set, validation set, and test set. As articulated by Shah (2017), the train set lays the groundwork for model training, allowing the algorithm to grasp intricate patterns and features within the data. Subsequently, the validation set, as underscored by Shah (2017), assumes a pivotal role in fine-tuning model parameters and evaluating performance. Its insights serve as a compass, guiding refinements in subsequent iterations.

In alignment with Brownlee's principles (2017), the test set serves as an unbiased yardstick for gauging the final model's efficacy, providing an impartial assessment of its performance. With a total of 3929 images, the dataset allocation designates 2828 images for training, 708 for validation, and the remaining 393 for testing. Each subset contributes uniquely to the model development journey, ensuring robustness and accuracy.

Featured in section 2.2 of the project, a snapshot of the validation set offers a glimpse into the dataset's composition, enriching understanding of its distribution. This meticulous partitioning underscores the project's methodological rigor, facilitating comprehensive model evaluation and refinement.

C. Data Augmentation

In computer vision, image data augmentation is a commonly used strategy that aims to generate additional samples with random mutations while maintaining the original class labels, hence augmenting the training dataset (Gu, Pednekar, & Slater, 2019). By exposing machine learning models to a wider variety of training cases, data augmentation aims to improve the generalizability of these models. The model performs better on unseen data because of this diversity, which aids in the model's learning of robust features and patterns.

Frequently used techniques in data augmentation include random cropping, rotating images at different angles, flipping images horizontally or vertically, and rearranging them while applying changes. By introducing variances in the training data, these strategies improve the model's ability to handle various input variations and emulate real-world circumstances. It is crucial to remember that data augmentation is usually limited to the training dataset because adding more data to the test dataset may create bias and affect how well the model performs. The test dataset's augmentation is withheld, allowing for a more accurate assessment of the model's accuracy on unknown data.

In actuality, techniques such as Keras' ImageDataGenerator can be used to effectively implement data augmentation (Gulli & Pal, 2017). This tool applies a sequence of random augmentations to batches of input images to make the transformation process easier. After randomly transforming each image in an input batch of photos, the ImageDataGenerator returns the augmented batch for training. Crucially, the ImageDataGenerator ensures that the original dataset is preserved and that the model gains knowledge from a variety of augmented samples by only offering the randomly modified data for training. This method improves the model's overall performance in real-world applications by enriching the training dataset and enhancing the model's capacity to generalise well to new data.

D. Feature Extraction

In the realm of image analysis, feature extraction stands as a pivotal process, crucial for enhancing system accuracy. A landmark contribution by Haralick, Shanmugam, and Dinstein (1973) introduced the Gray Level Co-occurrence Matrix (GLCM), a cornerstone method for extracting features from images. This technique meticulously tabulates the frequency of different combinations of pixel brightness values, or grey levels, within an image, as highlighted by Hall-Beyer (2017). GLCM texture analysis delves into the spatial relationships between pairs of pixels, known as reference and neighbor pixels, providing valuable insights into textural properties.

The computation of GLCM involves constructing a matrix framework that captures the spatial correlation between

reference and neighbor pixels, thereby facilitating texture analysis essential for tumor detection. Notably, Hall-Beyer (2017) delineates several key statistical features pivotal in texture analysis: Correlation, elucidating spatial dependencies between pixels, offers crucial insights into pixel relationships. Contrast serves as a measure of intensity disparity between pixels and their neighbors, aiding in the delineation of tissue boundaries within the image. Energy, quantifying the extent of pixel pair repetitions, serves as a metric for assessing image similarity. Homogeneity reflects the scaled local changes in image texture, providing insights into the uniformity of pixel distribution within the GLCM. Dissimilarity facilitates the comparison of segmentations generated by different algorithms, offering valuable insights into segmentation quality and differences.

Leveraging these statistical features extracted from GLCM analysis holds immense promise for enhancing tumor detection and characterization, paving the way for improved diagnostic accuracy and treatment efficacy in medical imaging applications.

E. Classification

1) Convolution Neural Network (CNN): Within this study, the utilization of a Convolutional Neural Network (CNN) model to categorize tumors in MRI Brain scans takes center stage. Drawing from its widespread application in object recognition, as outlined by O'Shea and Nash (2015), CNNs are structured with key layers—convolution, pooling, and fully-connected—forming the architectural backbone depicted in Figure 2.3.

The CNN's operational essence unfolds across four pivotal domains. Initially, the input layer encapsulates the pixel values from MRI brain scans, setting the stage for subsequent analysis. Next, the Convolution layer undertakes the critical task of discerning tumor presence within the brain. This layer scrutinizes localized input regions, determining correlations via scalar products with their respective weights. Following this, the introduction of the rectified linear unit (ReLu) activation function infuses non-linearity, enhancing model complexity. The pooling layer then executes spatial down-sampling, optimizing computational efficiency while retaining pertinent features. Finally, the fully connected layers culminate in class score generation, propelling the subsequent classification endeavor.

This iterative transformational journey epitomizes the CNN's prowess, systematically processing input layers to distill crucial class scores vital for MRI brain scan classification. This approach underscores CNNs' efficacy in medical imaging analysis, promising strides in precise tumor detection and classification, thereby advancing diagnostic accuracy and patient care.

F. Segmentation

Pixel-level segmentation is required, and the U-Net architecture proves to be an extremely effective solution. U-Net, so named because of its unique architectural resemblance
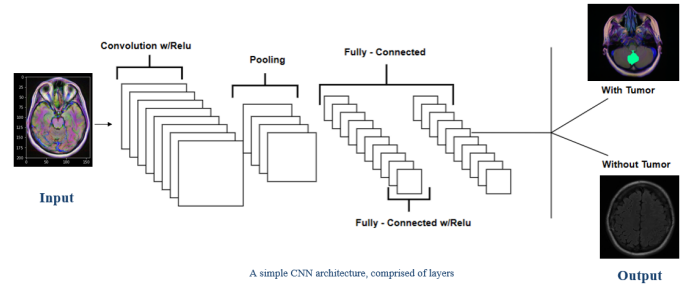

Fig. 4. A simple CNN architecture, comprised of layers

to the letter U, has received a lot of praise for its efficient pixel-based picture segmentation. Because of its creative architecture, which combines encoding and decoding routes, the network can capture fine details and context at various image scales. While extracting hierarchical features, the encoding path which resembles the left half of the letter U—gradually decreases spatial resolution. On the other hand, the decoding path, which mirrors the U's right side, gradually upsamples the features and combines them with high-resolution data to produce accurate segmentation maps. Because of this special architecture, U-Net performs segmentation jobs better than typical CNNs, especially in the biomedical arena where precise pixel-by-pixel delineation of structures is critical. Its efficacy and adaptability make it a top option for a wide range of applications, from semantic segmentation in computer vision to medical picture analysis. Therefore, U-Net is a major breakthrough in the field that will enable academics and practitioners to improve the interpretation of biomedical pictures for study and diagnosis as well as obtain superior segmentation performance.

The contraction path and the expansion path make up the two primary structural elements of the ResUNet architecture. Because of its architecture, the network is especially well-suited for tasks like image segmentation, as it can learn complicated data and rebuild images with accuracy. The contraction path is made up of various contraction blocks, each of which has twox2 max pooling operations after a number of residual blocks. These basic building components, known as residual blocks, enable the network to pick up residual information, which makes training deep architectures easier. The number of feature maps doubles as the input moves through each contraction block, improving the model's ability to identify minute details and patterns in the input data.

Nonetheless, the ResUNet architecture's expansion or decoder part is its most important component. Here, the upsampled input from the previous layer is received by each block in the expansion path, and it is concatenated with the equivalent output features from the remainder blocks in the contraction path. This method makes sure that the features that were acquired during the contraction phase are used to their fullest potential when the image is rebuilt. The network may better maintain geographical information and semantic context by combining data from both the contraction and expansion paths,

which produces segmentation results that are more accurate.

The ResUNet architecture is a strong and adaptable solution for a variety of biological imaging tasks and beyond, marking a significant leap in the field of image segmentation. It is a useful tool for both researchers and practitioners because of its capacity to integrate residual connections and use both local and global information to ensure higher performance and generalizability.
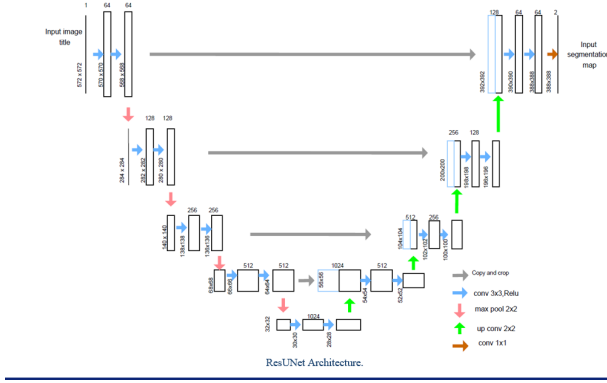

Fig. 5. ResUNet Architecture

IV. RESULT

The ImageDataGenerator module in the Keras library is used to supplement the brain pictures. Using the train test split function from the Scikit-Learn library, the dataset is split into three sets after successful preprocessing: the training set, the test set, and the validation set. A transfer learning model, namely ResNet50, is built for classification purposes in order to identify the presence or absence of tumours. The input data format that ResNet50 wants is (256 x 256 x 3). Furthermore, a tumour localization model, ResUNet, is created for the segmentation of the pictures. Rectified Linear Unit (ReLU) is the activation function, and each ResUNet model is trained using normal backpropagation procedures over a period of generally 40 epochs. The prediction function makes two different kinds of predictions on the image after receiving as input a DataFrame with an ImageID. To start, the image is classified by the network to ascertain if a tumour is there or not. The image is classified as no-tumor if the classification model accurately predicts that there isn't a tumour. On the other hand, the segmentation network receives the image if the classification model is deemed questionable. The segmentation network is used to look for tumours in the image more closely. The segmentation network identifies the kind and position of the defect in the image if a tumour is found. As a basic statistic for evaluating classification models, accuracy may be obtained with ease using the confusion matrix. Conversely, precision indicates the degree of purity of positive detections with respect to the ground truth and is intimately associated with random mistakes. An algorithm with a high precision yields more relevant results than irrelevant ones. The confusion matrix can also be used to determine precision.
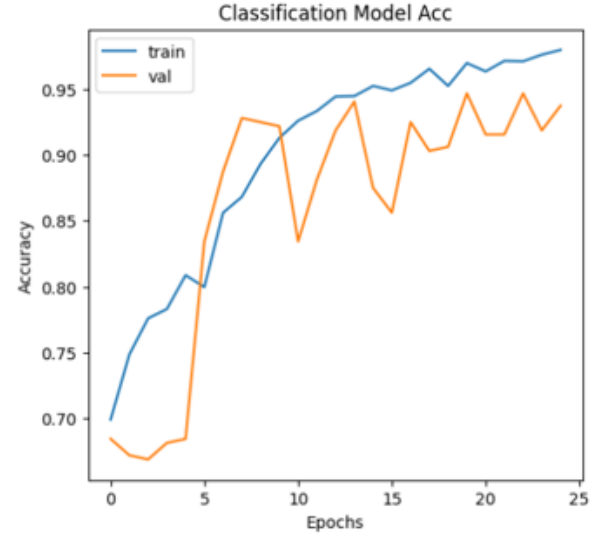

Fig. 6. Classification model accuracy.

A high recall rate means that the model is able to retrieve the majority of the relevant findings. Recall measures the completeness of positive predictions in relation to the ground truth. There is a formula used to compute it. The accuracy

	Precision	Recall	F1 - score
0	0.96	0.96	0.96
1	0.93	0.93	0.93
accuracy			0.95
macro avg	0.94	0.95	0.94
weighted avg	0.95	0.95	0.95

Fig. 7. Precision, recall and accuracy

of a test is measured in statistical analysis by the F-score, especially the F1, which is created by averaging the weights of precision and recall. Stochastic gradient descent is a technique used by deep learning algorithms to optimise and learn model objectives. It is vital to make sure that the mathematical representations of objectives take edge situations into account in order to accelerate accurate learning. Class imbalance is a prevalent problem in datasets that arises during segmentation operations. In order to solve this issue, the weight vector must be updated using loss functions in accordance with both the labelled output and the model's computed output. The Dice score coefficient (DSC) is a commonly used metric for segmentation task evaluation in the medical domain. The focal Tversky loss function (FTL), a novel method for model evaluation presented by Abraham and N. M. Khan in 2019, is used,

even though Dice coefficient largely takes the segmentation class into account rather than the background class.

In actuality, the FTL is unaffected even if a pixel is incorrectly labelled with a high Tversky index. On the other hand, the FTL significantly decreases if the pixel is misclassified and the Tversky index is tiny.With an impressive 95% test accuracy, CNN was able to classify photos into tumour and normal patients with effectiveness.Furthermore, this CNN's accuracy with ResUNet achieved 95%. Interestingly, since the feature values are taken straight from the CNN, the suggested CNN-based classification method does away with the requirement for further feature extraction stages.
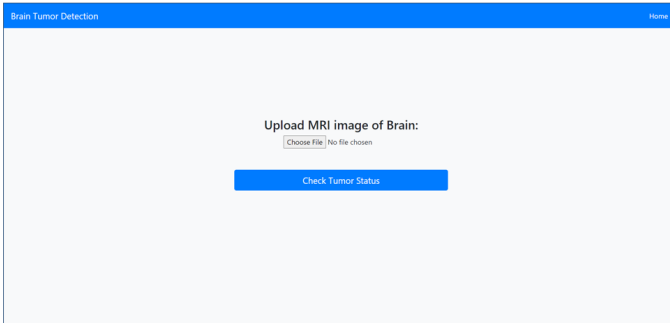

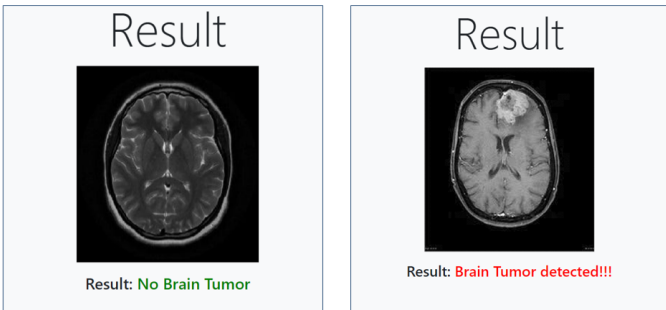Fig. 8. Focal Tversky Loss

Fig. 9. Home Page of Application

Fig. 10. Result

V. CONCLUSION

The primary objective of this project is to develop an efficient automatic brain tumor classification and localization system characterized by high accuracy, performance, and low complexity. Initially, conventional brain tumor classification is conducted using a Convolutional Neural Network (CNN) based on ResNet50. This stage focuses on determining whether the tumor exists or not, achieving a balance between low complexity and high accuracy, albeit with increased computational time. Subsequently, to precisely localize the tumor within the given MRI images and delineate its boundaries, another CNN-based approach known as ResUNet-based segmentation is introduced. This segmentation method predicts the position of the tumor by generating a mask around its edges, providing valuable insights for medical diagnosis.

To further enhance accuracy, the focal Tversky loss function is applied during training, resulting in a remarkable training accuracy of 96%. This heightened accuracy not only bolsters the reliability of the automated diagnosis process but also complements the expertise of physicians in diagnosing and treating patients. By leveraging the strengths of automated classification and segmentation techniques, this proposed method aims to significantly improve medical accuracy, ultimately enhancing patient outcomes and treatment efficacy.

VI. FUTURE SCOPE

A meticulous focus must be directed towards both unclassified and misclassified samples in the context of brain tumor classification. The presence of unclassified samples typically arises from detections with low scores, potentially leading to instances where tumors are erroneously categorized as "no tumors." To address this challenge in future iterations, integrating healthy images into the dataset could help mitigate such occurrences. Additionally, refining the pre-processing step with algorithms designed to accentuate subtle features in the images may enhance classification accuracy.

Looking ahead, the incorporation of optimization algorithms such as genetic algorithms, particle swarm optimization, or simulated annealing holds promise for identifying the optimal parameter set that maximizes classification accuracy, as proposed by Salçin et al. (2019). These algorithms offer avenues for fine-tuning model parameters and optimizing performance, thereby contributing to more precise and reliable brain tumor classification outcomes. By addressing these considerations and leveraging advanced optimization techniques, future iterations of the classification system can strive to achieve even greater accuracy and robustness in tumor detection and localization.

ACKNOWLEDGMENT

REFERENCES

[1]. Abadi, Martın et al. (2016). 'Tensorflow: Large-scale machine learning on heterogeneous distributed systems'. In: arXiv preprint arXiv:1603.04467 (cit. on p. 17).

[2].Abraham, Nabila and Naimul Mefraz Khan (2019). 'A novel focal tversky loss function with improved attention u-net for lesion segmentation'. In: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019). IEEE, pp. 683–687(cit. on p. 16).

[3].Anithadevi, D and K Perumal (2016). 'A hybrid approach based segmentation technique for brain tumor in MRI Images'. In: arXiv preprint arXiv:1603.02447 (cit. on pp. 14, 15).

[4].Bakas, Spyridon et al. (2017). 'Segmentation labels and radiomic features for the pre-operative scans of the TCGA-LGG collection'. In: The cancer imaging archive286 (cit. on p. 6).

[5].Brownlee, Jason (2017). 'What is the difference between test and validation datasets'. In: Machine Learning Mastery 14 (cit. on p. 7).

[6].Gu, Shanqing, Manisha Pednekar and Robert Slater (2019). 'Improve Image Classification Using Data Augmentation and Neural Networks'. In: SMU Data Science Review 2.2, p. 1 (cit. on p. 8)

[7].Gulli, Antonio and Sujit Pal (2017). Deep learning with Keras. Packt Publishing Ltd (cit. on p. 8).

[8].Hall-Beyer, Mryka (2017). 'GLCM texture: a tutorial v. 3.0 March 2017'. In: (cit. on p. 9).

[9].Haralick, Robert M, Karthikeyan Shanmugam and Its' Hak Dinstein (1973). 'Textural features for image classification'. In: IEEE Transactions on systems, man, and cybernetics 6, pp. 610–621 (cit. on p. 9).