IMPROVING PREDICTION ACCURACY OF DEEP LEARNING FOR BRAIN CANCER DIAGNOSIS USING POLYAK-RUPPERT OPTIMIZATION

M. Muthulekshmi
Department of Biomedical Engineering,
Saveetha School of Engineering, SIMATS, Saveetha University,
Chennai, Tamil Nadu, India.
muthulekshmisrinivasan@gmail.com

Azath Mubarakali
College of Computer Science,
King Khalid University,
Saudi Arabia.
mailmeazath@gmail.com

Blessy Y M
Department of Electronics and Communication Engineering,
R.M.K. Engineering College,
Chennai, Tamil Nadu, India.
ymb.ece@rmkec.ac.in

Submitted: *Apr, 28, 2024* ***Revised:*** *Jul, 12, 2024* ***Accepted:*** *Jul, 22, 2024*

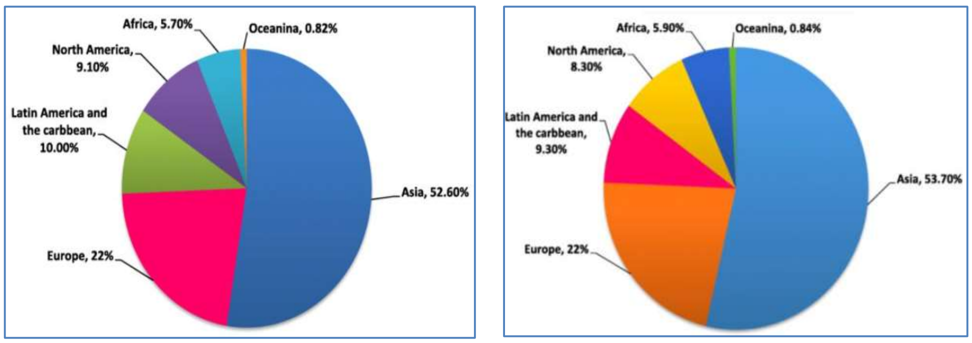
Abstract: Accurate and reliable diagnosis is critical for effective treatment planning for brain cancer. Recent advancements in deep learning have significantly enhanced diagnostic capabilities, but challenges persist in optimizing model performance for diverse and complex datasets. This study investigates the application of Polyak-Ruppert Optimization (PRO) to improve the prediction accuracy of conventional deep learning models for brain cancer diagnosis. Utilizing the REpository of Molecular BRAin Neoplasia DaTa (REMBRANDT) database, the proposed framework incorporates the advanced PRO technique to stabilize training and enhance generalization. The PRO's impacts on convergence rates, model robustness, and predictive accuracy across multiple cancer types are analyzed. Experimental results demonstrate that VGG and ResNet models employing the PRO technique outperform the conventional architectures such as VGG and ResNet in classification metrics such as accuracy, sensitivity, and specificity. The potential of advanced optimization strategies such as PRO to refine deep learning applications in oncology paves the way for more accurate, efficient, and interpretable diagnostic systems.

Keywords: Computer-aided diagnosis, brain cancer, deep learning, convolutional neural network, Polyak Ruppert optimization.

I. INTRODUCTION

Brain cancer is an abnormal proliferation of tumors inside brain tissue. These tumors can severely disrupt brain functioning, resulting in symptoms like headaches, seizures, cognitive deterioration, and movement impairment. Brain cancer is classified as primary brain tumors, which originate in the brain, and secondary or metastatic cancers, which disseminate from other regions of the body. Gliomas, meningiomas, and medulloblastomas are prevalent categories of primary brain neoplasms [1].

The incidence of brain cancer in India has been progressively rising over the years. Recent statistics indicate that brain tumors constitute roughly 2% of all cancers in India, with an estimated incidence rate of 5–10 cases per 100,000 individuals each year. The survival rate is very low, especially for high-grade Gliomas, highlighting the critical necessity for prompt and precise diagnosis. The restricted availability of advanced diagnostic facilities in rural and underprivileged areas intensifies this issue. According to the GLOBOCAN 2020 survey [2-3], over fifty percent of the population in Asia was impacted by brain tumors compared to other continents. Figure 1 illustrates brain tumour statistics from the GLOBOCAN 2020 study.



(a) Incidence (b) Mortality

Fig. 1 GLOBOCAN 2020 Statistics

Deep learning has arisen as an effective technique for medical image analysis, providing unparalleled precision in identifying and categorizing diseases from intricate datasets. Deep learning models, especially Convolutional Neural Networks (CNNs), have effectively analyzed Magnetic Resonance Imaging (MRI) for brain cancer diagnosis. These algorithms can aid radiologists by automating tumor identification, categorizing tumor kinds, and accurately predicting malignancy grades.

Although deep learning has many advantages, its performance mainly depends on the necessity for extensively annotated datasets, managing discrepancies in image quality, and attaining generality across varied patient groups. This study seeks to tackle these issues by creating a strong deep learning-based system for brain cancer classification. The suggested method utilizes advanced neural network topologies to improve diagnostic precision while ensuring computational efficiency.

II. RELATED WORKS

A hybrid CNN architecture for the classification of brain cancers is examined in [4]. It employs four CNN architectures: InceptionV3, ResNet, DenseNet, and VGG16 for classification purposes. The predictions from these designs are evaluated using the XAI layer for final classification. A lightweight CNN model for classifying MRI brain images is described in [5]. As the name implies, it comprises a limited quantity of convolutional and max-pooling layers, employing minimal repetitions. Batch normalization with an elevated learning rate is employed to expedite model training.

A GraphMRINet architecture is developed in [6] to classify MRI brain images. The Prewitt operator is employed to identify edges during the graph building. Pixels with an intensity exceeding 128 are designated graph nodes, and the other pixel points are regarded as graph edges. An isomorphic graph network is utilized for classification via Adam optimization. MobileNetV2 is employed for the classification of brain images in [7]. Features extracted from the pre-trained MobileNetV2 are inputted into several classifiers. Three distinct classifiers, extreme learning machine, Schmidt neural network, and random vector functional-link network, are utilized for classification purposes. The network's weights and biases are optimized with the chaotic bat algorithm.

A hybrid intelligent approach for brain tumor categorization is implemented in [8]. It employs median filtering for denoising, followed by applying a U-Net model to detect the cancer. The recovered features, including median binary pattern characteristics and local Gabor directional patterns, are classified utilizing a deep belief network and Bi-LSTM. A refined EfficientNet is discussed in [9] to classify brain images. It is a profound CNN that uses EfficientNet-B0 as its foundational model. The image quality is improved, and data augmentation is utilized to expand the dataset for efficient training. By incorporating additional layers, the base model is fine-tuned to get superior outcomes compared to VGG, ResNet, and Inception.

A Parallel Deep CNN (PDCNN) is employed to classify MRI brain images [10]. The MRI images undergo preprocessing using anisotropic diffusion filtering before feature extraction with PDCNN. An ensemble method for classification utilizes SVM, Bayes, decision trees, and KNN classifiers. Deep Neural Networks (DNN) utilize time-frequency information to classify brain images, as discussed in [11]. Initially, deep features are extracted using a deep CNN with pooling feature mapping. Subsequently, time-frequency characteristics are retrieved and integrated with deep features. Differential DNN models are ultimately constructed for classification purposes.

A hybrid CNN architecture integrated with Cubic SVM for brain tumor classification is described in [12]. It employs EfficientNetB0 and VGG-19 models for feature extraction, and the aggregated features are input into a Cubic SVM classifier for classification. A modified Visual Geometry Group (VGG) architecture is described in [13] to classify brain images. Conventional VGG architecture employs a max-pooling layer for feature reduction. The improved VGG system has a median-pooling layer instead of max-pooling. Median pooling mitigates the influence of noisy features during selection, whereas max pooling chooses noisy pixels in images impacted by salt and pepper noise.

A brain image classification for cancer detection utilizing DNN is discussed in [14]. A deep wavelet autoencoder is employed for feature extraction, and an uncomplicated autoencoder model performs classification using three hidden layers. A DNN called BrainCDNet was developed in [15] to classify MRI brain images. A nimble filter is utilized to enhance the borders of the brain. It comprises three CNN blocks and employs concatenated pooling layers. To prevent overfitting, the weights are initialized using 'He Normal' initialization, supplemented by batch normalization and Global Average Pooling (GAP) through the initialization of weights in layers.

III. PROPOSED SYSTEM

Deep learning involves training models by adjusting their parameters (weights and biases) to minimize a predefined loss function, quantifying the error between predicted and true outputs. Optimization algorithms such as Stochastic Gradient Descent (SGD) and Adam are used to iteratively update these parameters to reduce the loss and improve the model's performance. Neural networks have millions or billions of parameters; thus, finding the best combination of these parameters to make accurate predictions is a non-trivial task. Optimization helps navigate the high-dimensional parameter space to converge toward the optimal solution. The loss function of a deep learning model often has a complex landscape with local minima, saddle points, and flat regions. Effective optimization strategies help escape poor local minima and avoid stagnation in saddle points, steering the model toward global minima or acceptable solutions. A proper optimization technique can also significantly reduce the time it takes for a model to converge to a solution and influence how well it generalizes to unseen data. Hence, by applying the right optimization techniques, deep learning models become more accurate, computationally feasible, and practically useful in real-world applications. This paper employs the PRO to train conventional deep learning architectures such as VGG and ResNet.

A. VGG Architecture

VGG models [16] set new benchmarks by achieving high accuracy on image recognition tasks during their introduction. VGG networks use a straightforward architecture consisting of sequential convolutional layers with small (3x3) filters, followed by pooling layers. This consistent design is easy to understand and implement. VGG demonstrated that increasing network depth (up to 19 layers in VGG-19) could significantly improve performance, providing a baseline for deeper architectures. Small receptive fields (3x3 convolutions) enhance feature extraction, capturing spatial hierarchies of image patterns. Pretrained VGG models on ImageNet are widely used for transfer learning in tasks like image classification and object detection.

Figure 2(a) shows the VGG16 architecture. VGG16 comprises 16 layers, including 13 convolutional layers and 3 fully connected layers. All convolutional layers in VGG16 utilize exclusively 3x3 kernels with a stride of 1 for feature map extraction. Consequently, the computational complexity of VGG designs is inferior to that of other deep learning models. In the initial two convolutional layers, 64 filters are employed to extract the preliminary feature map, followed by a reduction in feature map size utilizing a 2x2 max pooling layer. Following the initial block, a pair of convolutional filters comprising 128 filters is employed in conjunction with a max pooling layer in the subsequent block. The subsequent three blocks employ three convolutional layers with 256, 512, and 512 filters to extract feature maps. VGG16 contains three fully connected layers where classification occurs via the back-propagation technique. The softmax function in the output layer categorizes the specified input test sample, which is defined by

$$SoftMax(O_t) = \frac{e^{O_t}}{\sum_n e^{O_n}} \tag{1}$$

where O_t is the output of *the* t^{th} layer and n is the number of classes.

B. ResNet Architecture

ResNet [17] introduced skip connections, which help gradients flow backward through the network, addressing issues like vanishing gradients in deep networks. ResNet allows scale to hundreds or even thousands of layers, improving accuracy without increasing the risk of overfitting or degraded performance. ResNet architectures generalize well across different datasets, making them versatile for image classification and medical imaging tasks. Pretrained ResNet models are among the most widely used backbones in deep learning pipelines, especially for transfer learning tasks.

Figure 2(b) shows the ResNet architecture. The main differences between VGG and ResNet architectures are (1) the absence of a max pooling layer between convolutional layer blocks, (2) the presence of a shortcut connection in each residual block, and (3) the implementation of average pooling before the fully connected layer. In each residual block, the shortcut connection extends from input to output. In ResNet18, the shortcut link occurs between two convolutional layers. All convolutional layers employ the rectified linear activation function, while the output layer utilizes the softmax activation function. Conventional architectures such as VGG and ResNet have 1000 neurons in the output layer to predict 1000 objects. However, this layer is modified to diagnose brain cancer with only two output neurons (normal/abnormal).

C. Optimization Technique

PRO is an advanced statistical technique that enhances the convergence and stability of iterative algorithms, particularly in machine learning and deep learning. Originating from stochastic approximation methods, this approach involves averaging the sequence of iterates generated during training, which helps to smooth out fluctuations and reduce variance. The Key features of PRO are as follows:

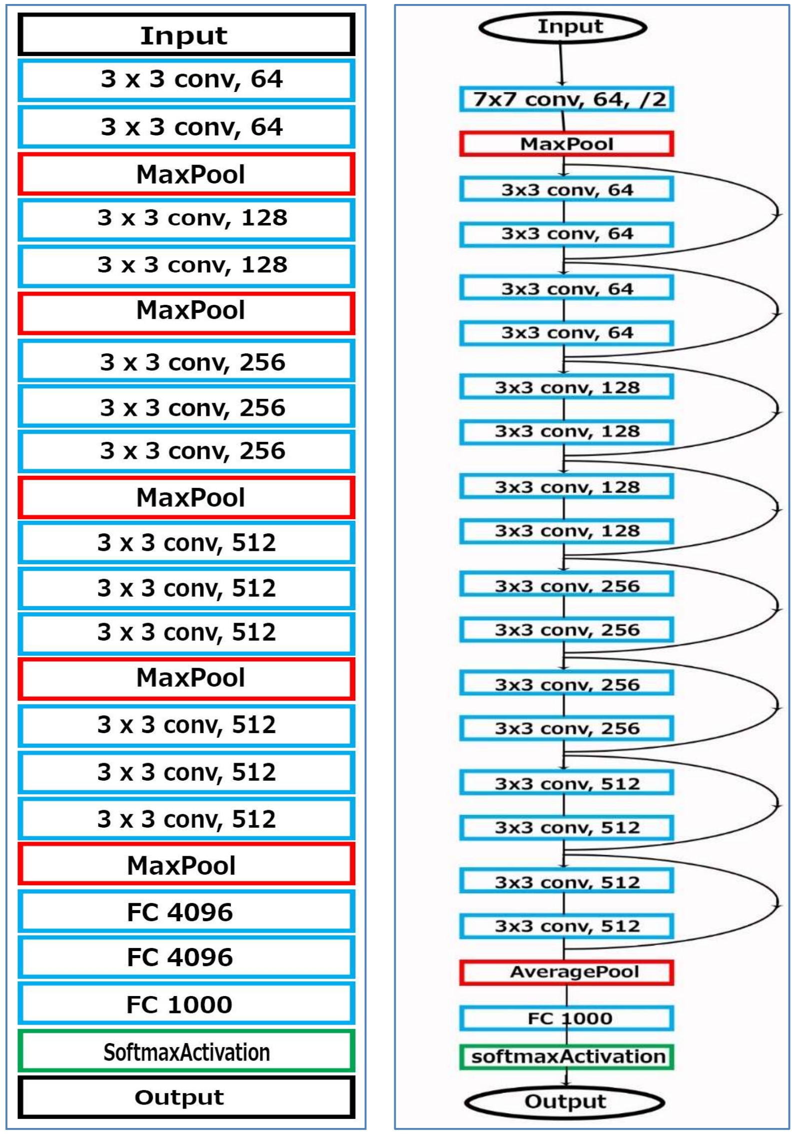
- **Iterate Averaging:** Instead of using the final iteration as the output, the algorithm computes the average of all previous iterations. This often results in better convergence properties and enhances the stability of training, particularly on noisy or complex datasets.
- **Variance Reduction:** Averaging helps mitigate noise's impact in SGD updates, leading to more stable optimization. It improves the model's ability to generalize to unseen data.
- **Faster Convergence:** It improves the convergence rate for models trained on non-convex loss functions, making it particularly effective in deep learning applications.

The PRO can address challenges in training deep learning models on diverse and high-dimensional datasets like REMBRANDT, improving prediction accuracy and reliability. The parameter update in SGD is defined by

$$\theta_{t+1} = \theta_t - \eta \nabla_\theta L(\theta_t; x_t, y_t) \tag{2}$$

where η is the learning rate, θ_t are the model parameters at iteration *t*, $\nabla_\theta L(\theta_t; x_t, y_t)$ is the gradient of the loss function (*L*) with respect to θ_t and the data samples are represented by (x_t, y_t). In PRO, the averaged iterate (*T*) is defined in Eqn. (2) instead of simply using the final iteration θ_t. The average iterate in PRO is defined as

$$\theta_T = \frac{1}{T}\sum_{t=1}^{T}\theta_t \tag{3}$$



(a) VGG-16 **(b) ResNet-18**

Fig. 2 Conventional CNN architectures

III. REpSULTS AND DISCUSSIONS

The REMBRANDT dataset is an extensive compilation of brain cancer data, encompassing imaging, clinical, and genomic information. It was established to enhance research by amalgamating clinical data with molecular characterizations of brain tumors. The imaging component includes pre-surgical MRI scans from

130 patients, amounting to roughly 10.59 GB of data. The images are accessible in DICOM format via the cancer imaging archive [18-20]. In addition to imaging, REMBRANDT encompasses comprehensive clinical data and genetic profiles from glioma tissues. This includes gene expression arrays, copy number arrays, and clinical phenotypic data. The datasets are stored in Georgetown University's G-DOC System, a platform that consolidates diverse biological data types for thorough analysis. Access to G-DOC necessitates registration. Figure 3 shows the brain MRI images in the REMBRANDT database.

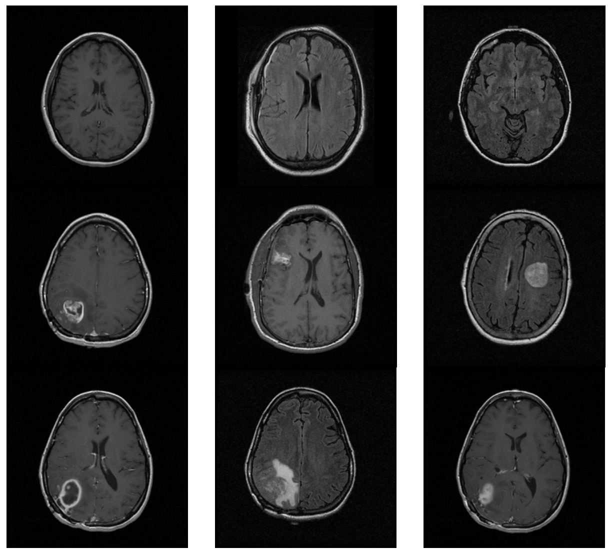
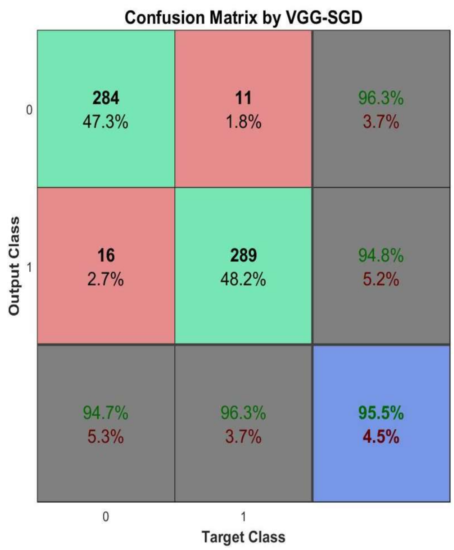
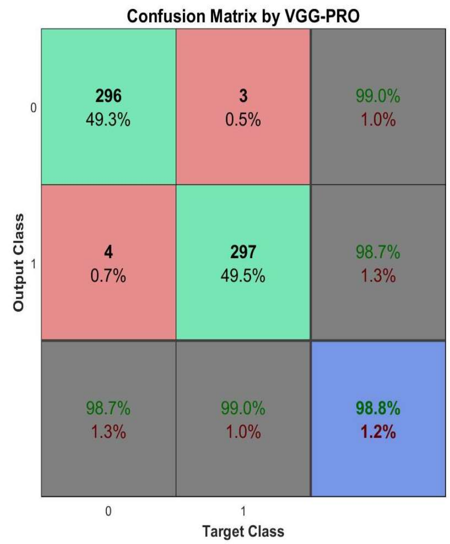


Fig. 3 Brain images in REMBRANDT database - Normal (Top row), Low-grade (Middle row) and High-grade (Bottom row)

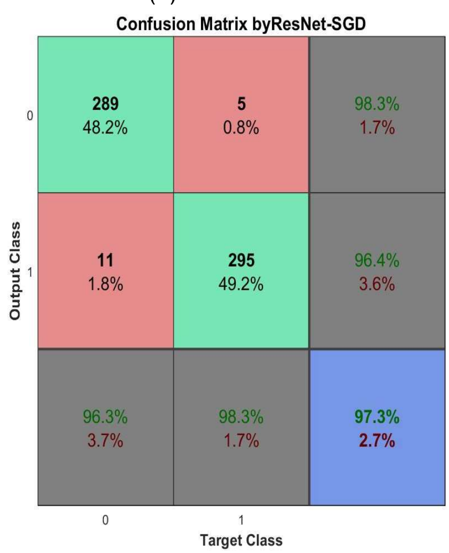
The PRO-based deep learning system's capability to classify all brain MRI images is assessed by classification accuracy. The capacity to differentiate between abnormal and normal brain MRI images is assessed by sensitivity and specificity. It is widely recognized that an increase in training samples enhances classification accuracy. Consequently, data augmentation is implemented via rotation and flipping procedures, resulting in 1000 images per class. The PRO-based deep learning system is trained using 70% of images per class, and the remaining 30% are employed for testing the system. The performance of the proposed system is compared with that of conventional architectures such as VGG, ResNet, and AlexNet. The obtained confusion matrices are shown in Figure 4.



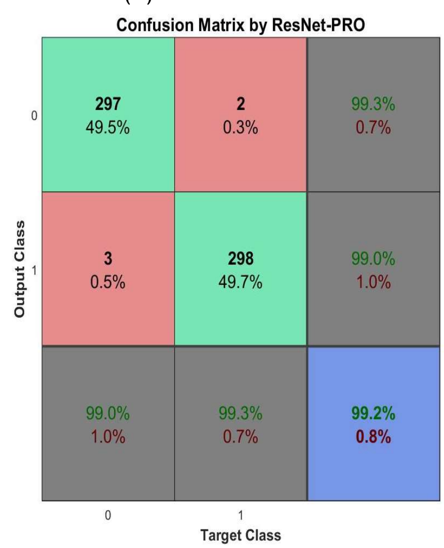
(a) VGG-SGD



(b) VGG-PRO



(c) ResNet-SGD



(d) ResNet-PRO

Fig. 4 Confusion matrices by VGG and ResNet with SGD and PRO techniques

The confusion matrices in Figure 4 indicate that the proposed PRO-based system accurately identifies more images than standard architectures such as VGG and ResNet. The proposed system accurately classifies 20 additional images compared to VGG-SGD and 11 more than ResNet-SGD. Based on the confusion matrices, the performance measures such as sensitivity and specificity are computed and are shown in Figure 5.

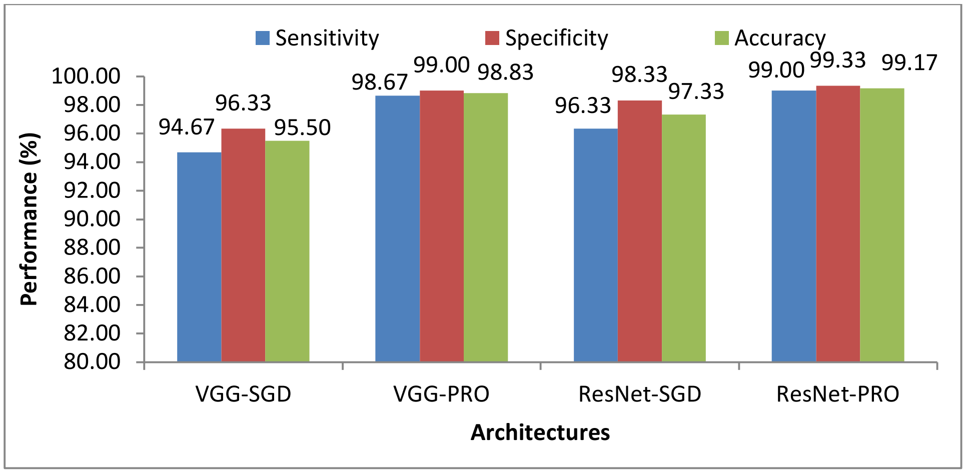


Fig. 5 Comparative analysis of conventional architectures with SGD and PRO techniques

This work also includes the Receiver Operator Characteristic (ROC) curve to graphically depict the proposed system's performance. The graphic illustrates the relationship between two critical parameters: the true positive rate (sensitivity) and the false positive rate (1-specificity). Figure 6 shows the ROCs for the proposed PRO-based deep learning system and other conventional architecture.

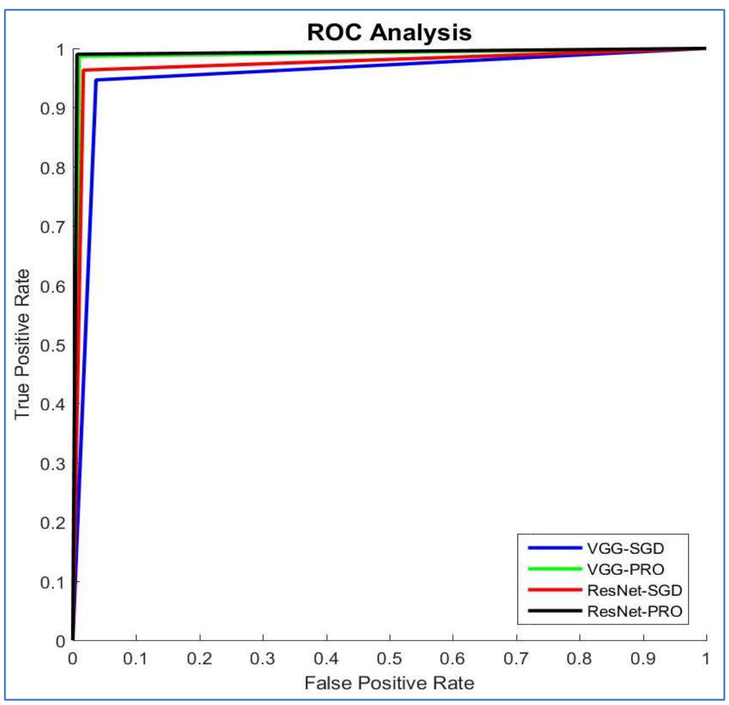


Fig. 6 Receiver Operator Characteristic Curves for Brain Image Diagnosis

It can be seen from Figure 6 that the ROC curve generated by the proposed system (VGG-PRO and ResNet-PRO) has a greater area under the curve. The proposed system provides superior classification performance by ResNet-PRO, with an AUC of 0.992.

IV. CONCLUSIONS

The proposed study examined the integration of PRO into deep learning to enhance accuracy in brain cancer diagnosis. This work analyses brain MRI images to identify samples based on their clinical conditions, contributing to exploring and improving deep learning systems. The main objective of this study is to increase the performance of conventional deep learning architectures such as VGG and ResNet for brain cancer diagnosis. An efficient optimization technique, PRO, is employed to train these architectures instead of SGD. Experimental results prove that the PRO technique outperforms SGD and performs better on the REMBRANDT database for classifying brain images. The future work involves developing and conducting research on CNNs by utilizing other activation functions, such as the exponential linear unit and the swish function.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

REFERENCES

[1]. N. Salari, H. Ghasemi, R. Fatahian, K. Mansouri, S. Dokaneheifard, M.H. Shiri and M. Mohammadi, "The global prevalence of primary central nervous system tumors: a systematic review and meta-analysis," European journal of medical research, vol. 28, no. 1, 2023, pp. 1-16.

[2]. S.V.S. Deo, J. Sharma and S. Kumar, "GLOBOCAN 2020 Report on Global Cancer Burden: Challenges and Opportunities for Surgical Oncologists," Annals of Surgical Oncology, vol. 29, no.11, 2022, pp.6497-6500.

[3]. H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal and F. Bray, "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," CA: A Cancer Journal for Clinicians, vol. 71, no. 3, 2021, pp. 209–249.

[4]. A. B. Ramakrishnan, M. Sridevi, S. K. Vasudevan, R. Manikandan and A. H. Gandomi, "Optimizing brain tumor classification with hybrid CNN architecture: Balancing accuracy and efficiency through oneAPI optimization," Informatics in Medicine Unlocked, vol. 44, 2024, pp. 1-8.

[5]. A.S. Musallam, A.S. Sherif and M.K. Hussein, "A New Convolutional Neural Network Architecture for Automatic Detection of Brain Tumors in Magnetic Resonance Imaging Images," IEEE Access, vol. 10, 2022, pp. 2775–2782.

[6]. B. Liao, H. Zuo, Y. Yu and Y. Li, "GraphMriNet: a few-shot brain tumor MRI image classification model based on Prewitt operator and graph isomorphic network," Complex & Intelligent Systems, vol. 10, 2024, pp. 6917-6930.

[7]. S.Y. Lu, S.H. Wang and Y.D. Zhang, "A classification method for brain MRI via MobileNet and feedforward network with random weights," Pattern Recognition Letters, vol. 140, 2020, pp. 252–260.
[8]. A.M. Simo, A.T. Kouanou, V. Monthe, M.K. Nana and B.M. Lonla, "Introducing a deep learning method for brain tumor classification using MRI data towards better performance," Informatics in Medicine Unlocked, vol. 44, 2024, pp. 1–24.
[9]. H.A. Shah, F. Saeed, S. Yun, J.H. Park, A. Paul and J.M. Kang, "A Robust Approach for Brain Tumor Detection in Magnetic Resonance Images Using Fine-Tuned EfficientNet," IEEE Access, vol. 10, 2022, pp. 65426–65438.
[10]. M. Bourennane, H. Naimi and E. Mohamed, "Deep Feature Extraction with Cubic-SVM for Classification of Brain Tumor," Studies in Engineering and Exact Sciences, vol. 5, no. 1, 2024, pp. 19–35.
[11]. T. Renukadevi, K. Saraswathi, P. Prabu and K. Venkatachalam, "Brain image classification using time-frequency extraction with histogram intensity similarity," Computer Systems Science and Engineering, vol. 41, no. 2, 2022, pp. 645–460.
[12]. T. Rahman, M.S. Islam and J. Uddin, "MRI-Based Brain Tumor Classification Using a Dilated Parallel Deep Convolutional Neural Network," Digital, vol. 4, no. 3, 2024, pp. 529–554.
[13]. N. Veni and J. Manjula, "Modified visual geometric group architecture for MRI brain image classification," Computer Systems Science and Engineering, vol. 42, no. 2, 2022, pp. 825–835.
[14]. P.K. Mallick, S.H. Ryu, S.K. Satapathy, S. Mishra and G.N. Nguyen, "Brain MRI image classification for cancer detection using deep wavelet autoencoder-based deep neural network," IEEE Access, vol. 7, 2019, pp. 46278–46287.
[15]. K.R. Reddy, K.N. Rajesh, R. Dhuli and V.R. Kumar, "BrainCDNet: a concatenated deep neural network for the detection of brain tumors from MRI images," Frontiers in Human Neuroscience, vol. 18, 2024, pp. 1–11.
[16]. K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *3rd International Conference on Learning Representations*, 2015. pp. 1–14.
[17]. K. He, X. Zhang, S. Ren and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778
[18]. Scarpace, Lisa, Flanders, E. Adam, Jain et al., "Data from REMBRANDT", The Cancer Imaging Archive. http://doi.org/10.7937/K9/TCIA.2015.588OZUZB.
[19]. K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel and F. Prior, "The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository," Journal of Digital Imaging, vol. 26, no. 6, 2013, pp. 1045–1057.
[20]. Brain MRI images: https://www.cancerimagingarchive.net/collection/rembrandt/