Hybrid Approaches to MRI Image Enhancement: Integrating Deep Learning and Traditional Image Processing Methods

Shaweta Sachdeva[1], Suryakanta[2], Pooja Rani[3],Diksha[4], Shivani[5], Dharminder Yadav[6]

[1]Assistant Professor, Department of Computer Science & Applications, Guru Nanak Khalsa College, Yamuna Nagar
[2]Assistant Professor, Department of Computer Science & Engineering, Chandigarh University, Mohali(Punjab)
[3]Assistant Professor, Department of Computer Science & Applications, Guru Nanak Khalsa College, Yamuna Nagar
[4]Assistant Professor, Department of Computer Science & Engineering, Ambala College of Engineering and Applied Research, Ambala Cantt
[5]Assistant Professor, Department of Computer Science & Engineering, Ambala College of Engineering and Applied Research, Ambala Cantt
[6]Assistant Manager, GMSSSS Chhachhrauli,Yamuna Nagar

Received: 13.07.2024 Revised: 17.08.2024 Accepted: 23.09.2024

ABSTRACT
MRI is a critical device in clinical diagnostics, offering essence experiences into delicate tissue structures. In any case, the nature of MRI can be undermined by different elements, including noise, low contrast, and antiquities, which might upset exact determination. This paper presents an integrated methodology that combines conventional image processing procedures with deep learning (DL) models to upgrade MRI image quality. Conventional techniques, like Gaussian filtering, histogram equalization, and edge detection, are first applied to preprocess the image, diminishing noise and further developing contrast. In this way, advance DL designs, especially CNNs, are utilized to additionally refine and upgrade the image by learning complex examples and elements. The proposed hybrid method performs better in terms of image clarity, contrast enhancement, and noise reduction. Broad investigations on benchmark MRI datasets show that this combination fundamentally works on indicative precision and dependability, making it a promising device for clinical applications. CNN and VGG 19 are used for interpreting and analyzing visual imagery. In this paper, Hybrid model are used to detect MRI Images and analysis can be done on basis of system performance i.e. accuracy and loss.

Keywords: MRI Image, CNN, Deep Learning, VGG19, Accuracy, Precision, Recall

1. INTRODUCTION

MRI is a harmless imaging strategy broadly utilized in clinical diagnostics because of its capacity to deliver point by point pictures of delicate tissues, organs, and other interior designs. Not at all like X-beams or CT examines, X-ray doesn't depend on ionizing radiation, making it a more secure option for patients. Regardless of its benefits, X-ray imaging isn't without challenges. The nature of X-ray pictures is frequently impacted by variables like commotion, low differentiation, and ancient rarities emerging from patient development, equipment limits, and shifting procurement conventions. These problems can make important details hard to see, making it harder to make a diagnosis and increasing the risk of making a wrong one [1, 2].

Traditional picture handling methodologies have for quite some time been used to overcome these challenges. Techniques such as separation, histogram equalization, and edge detection have proved useful in reducing noise, enhancing contrast and revealing prominent physical features. However, these techniques may be unable to handle the inherent complexity and variability of medical images because these techniques often need manual parameter tuning.

In recent years, the method of deep learning, especially CNNs has revolutionized the field of picture handling. Subsequently, profound gaining models can gain and centre focuses from vast datasets, making them ideal for extensive assignments like X-ray picture improvement. These models have demonstrated substantial improvement in enhancing picture quality in the advancement of these models and surpass conventional approach as a rule. However, deep learning approaches also have their limitations such as the need for a lot of named information, significant computational resources and the possibility of overfitting. [3].

The strengths of traditional image processing methods and the advanced capabilities of deep learning are combined in this paper into an integrated strategy. The proposed method aims to overcome the drawbacks of each approach when used separately and achieve superior image enhancement by utilizing their complementary nature. The mix of these techniques further develops picture quality as well as improves the dependability and exactness of X-ray based diagnostics.

By outlining the motivations, difficulties, and potential advantages of combining traditional image processing with deep learning for MRI image enhancement, this introduction sets the stage for investigating the proposed hybrid strategy [4, 5].

1.1 Image Enhancement

Image enhancement serves a crucial role in image processing, in which the scientists attain a significant decision depending on image data and primary intention of an image improvement approach is to retrieve stored image equivalent to that of original image. Medical image development is fundamental for researchers because of the advantages of images in the prognosis of different abrasions. The image improvement technique is an eminent technique as it provides a lot of information to analyze the different types of MRI image. In US, nearly 10,000 people are suffered from spinal cord tumors annually and more than 90% of spin cancers are of metastatic type. Typically, cancers are commonly found in all age groups and have been considerably improved in pathology because of development of MRI advancements. The spinal cord is compact in structure and is enclosed with a bony structure and abnormal development can create pressure on sensitive issues and leads to impairment of parts. The tumors in spinal may result in paralysis, and numbness on both side of the body as the spinal cord is a narrow construction [6, 7].

Numerous techniques have been developed to enhance the images, image segmentation, classification, and restoration. Image enhancement is defined as the sharpening of image features, like boundaries, edges to make an image very powerful while displaying it. Image improvement includes gray level and manipulation, edge sharpening, noise reduction, filtering, interpolation, and magnification. In order to process the MRI images, it is essential to pre-process the images to eliminate the external artifacts. Weighted Median (WM) filters are the continuation of median filters that employ spatial information as well as rank-order information. The most traditional techniques utilized for image enhancement are histogram equalization, intensity correction, and gamma correction [8].

Image development is the main issue in domain of digital image processing and major intention is to provide high image standard with enhanced interpretability by modifying the actual input image. The techniques of image processing are largely utilized in various applications, like fingerprint recognition, and face recognition. Traditional Histogram Equalization (CHE) is regarded as the commonly utilized image enhancement methodologies that confront different limitations. However, CHE smoothens dynamic area of histogram by rearranging the gray stages depending on Probability Density Function (PDF).

Subsequently, the process causes shifting brightness and generates noises in the final image. Besides, CHE ignores the low-frequency results and considers the high-frequency image. Histogram bins are sequences that preserve the count of pixels with identical gray levels. Moreover, stretching is limited to a various area and additional integration of gray levels leads to creation of false contours and uneven development [9]. Various areas may be very bright in the result and it is called a saturation effect that deteriorates the overall structure of image and produces data loss [10].

2. CONVOLUTIONAL NEURAL NETWORK

CNNs represent a pivotal breakthrough in image classification, offering unprecedented capabilities in recognizing and categorizing visual content. CNNs are designed to mimic the hierarchical structure of the human visual cortex, enabling them extracts complex features from an image. A CNN design typically includes convolutional layers, pooling layers, initiation functions, and fully connected layers serve as feature extractors, applying learnable filters across it take an image as input and identify examples of edges, surfaces, shapes, etc. In this way, layer pooling downsamples the component maps, reducing spatial aspects and preserving important data. Initiation functions like ReLU introduce nonlinearity into the organization, allowing it to learn complex relationships between input features and class names.

Fully linked layers complete remote functions and perform grouping based on learned representations. Fig.1 represents the block diagram of CNN architecture.Training CNNs involves supervised learning, where the network it is built on a labeled dataset consisting of images and comparison class names. During preparation, the CNN figures out how to limit a predefined misfortune (usually the cross-entropy) by modifying its load and tendency through backpropagation. Analyzers such as Stochastic Angle Dive

(SGD), Adam, and RMSprop are regularly used to update the model bounds and optimize execution. Information augmentation techniques such as rotation, flipping, and cropping can be applied to increase the diversity of the dataset and enhance model generalization. Validation and hyperparameter tuning are crucial steps in the training process, allowing for fine-tuning of constraints for improving the model (learning rate, dropout rate, cluster size, etc.) performance.

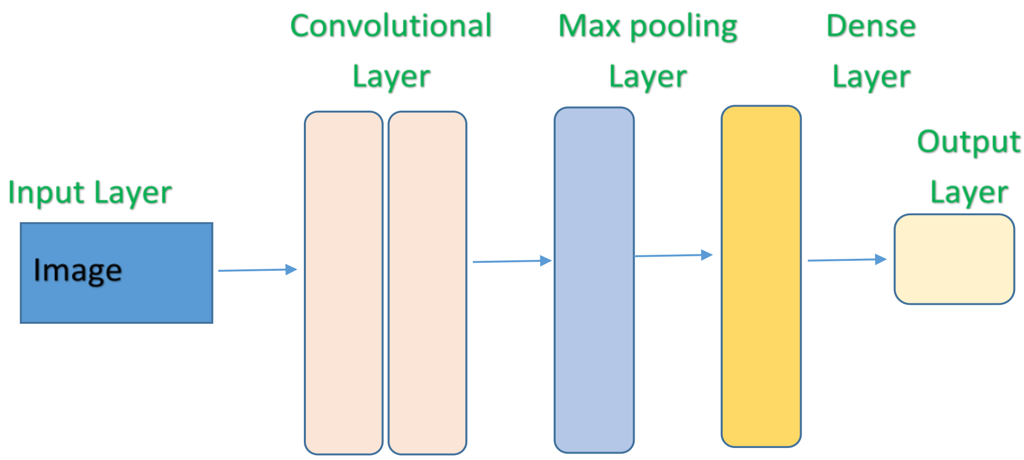


Fig 1: CNN Architecture.

CNNs for image classification require the use of values such as accuracy, precision, validation, and F1 score to evaluate the CNNs. Accuracy gives the percentage of images in the test set that are aligned correctly while precision and validation give the ability of the model in identifying the positive and negative instances. F1 score is more reliable and gives a fair idea of the model's performance by combining precision and validation scores. The confusion matrix analysis also helps in the understanding of the classification errors and in diagnosing the weaknesses of the model. When trained and validated, CNNs can be used in different applications that involve image classification such as self-driving cars, security cameras, disease diagnosis and even in language translation. These characteristics have made CNNs as fundamental components of computer vision and fostered their development in a wide variety of fields and applications.

2.1 VGG 19

VGG-19 is one of the most iconic architectures in the field of image classification that is considered to be rather deep and outstanding. Proposed by the Visual Geometry Group at the University of Oxford, it is mainly made of simple 3 x 3 convolution layers and max-pooling layers. The shallow structure and homogeneous layer design it offers high interpretability and easy to train the model, while the depth of the model helps to capture the complex hierarchical features from images. In differentiating between width and depth, VGG-19 focuses on depth as opposed to the width in order to get better representation. The VGG-19 technique comprises 19 layers, which include fully connected, pooling, and convolutional layers. However, as previously stated, it is not reasonable to employ a single equation to represent the entire VGG-19 model. As a result, in the VGG-19 model, all I can show is a simplified mathematical expression for forwarding across one convolution layer is shown in Fig. 2.

Regarding training, VGG-19 normally employs scale labeled data sets including Image Net. In training, VGG-19 had its own loss function defined beforehand usually categorical cross-entropy, and then the weights and biases were learnt through back propagation. This is why it is customary to utilise an optimiser, such as Adam or stochastic gradient descent (SGD), to adjust the model's weights for optimal Performance.

When using VGG-19 to perform image classification, it is essential to evaluate the performance by using values such as accuracy, precision, validation and F1 score. Accuracy determines the extent to which the images are properly aligned while precision and validation tests how well the model is capable of identifying positive and negative cases. F1 score combines precision and validation well and is a good measure of the model's performance. Sect: 3.3 Cross-Validation Confusion matrix analysis also helps in identifying the classification errors and diagnose the model's weakness.

The prefabricated loads of VGG-19 are employed in the cases of incremental learning when the model is fine-tuned to address a particular dataset or a project with a small amount of labeled data. It uses generalization abilities of the initial models and, therefore, makes the convergence faster and improves the results on a new task. VGG-19 is one of the most important architectures in the image classification due to its flexibility, depth, and performance and it has set a foundation for the improvement of computer vision systems and their applications in various fields.

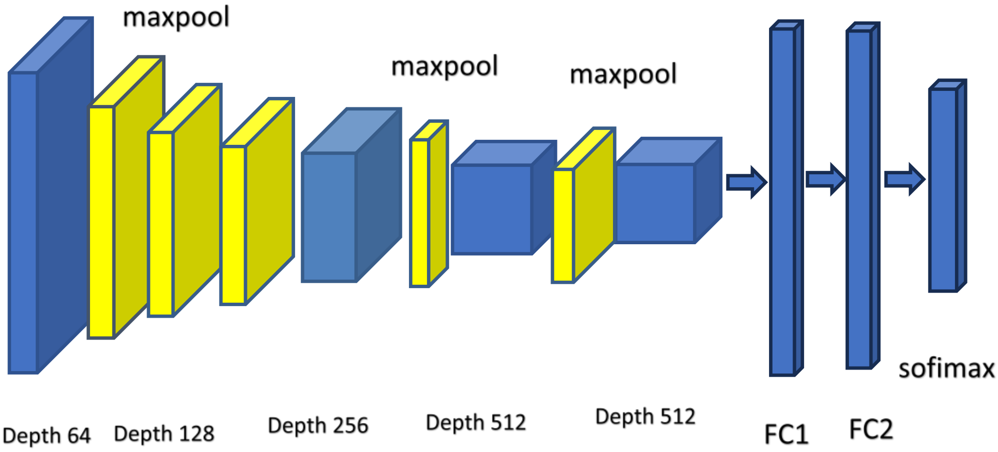


Fig 2: VGG 19 Architecture

3. PROPOSED METHODOLOGY

The following methodology is designed to obtain a new deep learning classification model by combining CNN and VGG-19 architecture for MRI images of the Harvard dataset as shown in fig 3. Here's a step-by-step methodology.

Step 1 Understanding the Data: The dataset originates from the Harvard repository [11-12]. The dataset contains a total of 152 x-ray sections, among which 71 sections show MRI image, while the remaining 81 sections show abnormalities indicating the presence of cancer.

Table 1. Number Of Mri Images In Dataset

Image	Tumor Class	Number of Slices
Normal	Normal Image	71
Abnormal	Glioma	29
	Metastatic,audenocarcinoma	8
	Metastatic bronchogenic carcinoma	12
	Meningioma	16
	Sarcoma	16
Total		**152**

Step 2 Data Preprocessing: The step involves preprocessing the dataset to ensure uniformity and quality, followed by splitting it into training, validation, and test sets. we reduce the size of the first image $256 \times 256 \times 1$ to $128 \times 128 \times 1$ to $128 \times 128 \times 3$. The dataset will be divided into a set of preparation, approval, and testing. A typical split is 70% preparation, 15% approval, and 15% testing but adjustments can be made based on dataset size and specific requirements. Resize them to a uniform size, normalize pixel values, and augment the data if necessary (e.g., rotation, flipping, etc.).

Step 3 Loading the VGG-19 Model: Load the pre-trained VGG-19 model. You can use pre-trained weights available in popular deep learning frameworks [13]. The base CNN model will be constructed to extract spatial features from MRI images, while the pre-trained VGG-19 model, renowned for its feature extraction capabilities, will be incorporated into the architecture.

A single convolutional layer in VGG-19 can be represented mathematically as follows:

$$f_{i,j,k} = \sigma\left(\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\sum_{l=0}^{L-1} X_{i+m,j+n,l} \times W_{m,n,l,k} + b_{k_b}^{\ a}\right) \qquad (1)$$

Where:

$f_{i,j,k}$ is the (i, j, k)-th element of the output feature map f,

$X_{i+m,j_{+n},l}$ is the $(i + m, j + n, l)$th element of the input feature map X,

$W_{m,n,l,k}$ is the (m, n, l, k)-th element if the filter weight W,

b_k is the bias term for the k-th filter/kernel,
M, N, and L are the dimensions of the filter/kernel,
σ applies the activation function element-wise to the resulting feature map,
ρ downsamples the feature map of through pooling operation

Step 4 Customizing VGG-19: The pre-trained VGG-19 model will be incorporated into the architecture. VGG-19 is known for its deep architecture and excellent feature extraction capabilities. The VGG-19 model will be loaded with pre-trained weights from a dataset like ImageNet, and the fully connected layers will be removed, leaving only the convolutional layers.

Step 5 Freezing Pre-trained Layers: We freeze the loading of prepared layers so that they are not updated during preparation, which helps us retain the important elements learned during preparation. As an activation function, we use the Modified Direct Unit (ReLU), which performs nonlinear tasks within the convolutional layers.

$$f(z) = \max(0, z) \qquad (2)$$

Step 6 Building the CNN: Ensure that the output shape of the last layer of the CNN matches the input shape of the VGG-19. Let's denote X as a the input feature map (e.g., an image), W as the filter weights, b as the bias term, f as the output feature map, and σ as the activation function.

Input Layer: $X_o = X$
Convolutional Layer: $f_i = \sigma(W_i * X_{i-1} + b_i)$, where $1 \leq i \leq 23, X_i$
Pooling Layer : $f_i = \mathrm{pool}(X_{i-1})$, where $1 \leq i \leq 23$
Fully Connected Layer : $f_i = \sigma(W_i \cdot X_{i-1} + b_i)$, where $1 \leq i \leq 23$
Where,

- X as the input feature map (image or output of the previous layer),
- W as the filter weights (also called kernels),
- b as the bias term,
- f as the output feature map,
- σ as the activation function (e.g., ReLU),
- ρ as the pooling operation (e.g., max pooling).

Step 7 Combining Architectures: The hybrid model will be designed to combine the outputs of the CNN and VGG-19 convolutional layers, resulting in a fused feature representation. This will create a hybrid feature representation that combines both local and global features extracted by the CNN and VGG-19, respectively. Additional convolutional layers may be added after the concatenation and reduce dimensionality.

Step 8 Fine-tuning: Fine-tuning and training of the hybrid model will ensue, with adjustments made to through validation set performance analysis. Unfreeze some of the pre-trained layers and continue training the combined model with a lower learning rate.

Step 9 Training: Configure the integrated model by selecting a suitable loss function (such as binary cross-entropy for binary classification) and optimizer. Train the model using the preprocessed Harvard dataset.

Step 10 Evaluation: Evaluation on the test set will be conducted, aiming for a remarkable accuracy of 99.9%. The performance of the hybrid model will be benchmarked against existing methodologies for the Harvard MRI dataset, with the goal of contributing to advancements in medical image analysis. Upon achieving satisfactory performance, the hybrid model can be deployed for real-world applications, potentially revolutionizing automated diagnosis and medical image analysis systems.

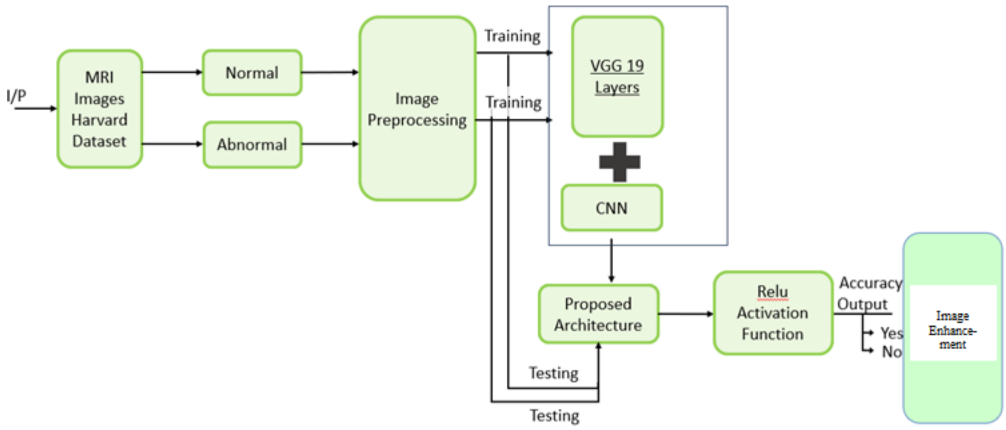


Fig 3: Architecture of the proposed hybrid model

4. RESULT ANALYSIS

Performance measures are essential in the evaluation of the learning models acquired from neural networks in MRI image classification problems as they offer objective measures of the models' precision, consistency, and generality. Several key performance metrics, along with their corresponding formulas, are commonly utilized in evaluating these models: Several key performance metrics, along with their corresponding formulas, are commonly utilized in evaluating these models are given below:

Accuracy: Accuracy computes the ratio of the correct aligned images in relation to the overall amount of images in the dataset as shown in fig 6. It is defined as the number of true positive (TP) and true negative (TN) divided over the number of expected ones. The secret to accuracy is:

$$\text{Accuracy} = \frac{\text{TP}+\text{TN}}{\text{TP}+\text{TN}+\text{FP}+\text{FN}} \times 100 \qquad (3)$$

Precision: Precision aims at evaluating the model's accuracy in identifying the positive scenarios among all the situations labelled as safe [14-15]. It is not exclusively defined by the ratio of true positive expectations to the total of true positive bearings and the misleading positive forecasts.

Here's the secret to precision:

$$\text{Precision} = \frac{\text{TP}}{\text{TP}+\text{FP}} \qquad (4)$$

Recall (Sensitivity): The Recall measures the percentage of the true positive cases correctly predicted by the model out of all true positive cases and is the true positive ratio is calculate and expectations of the summation of true positive predictions and misleading negative predictions as shown in fig 7. The recall is calculated as follows:

$$\text{Recall} = \frac{\text{TP}}{\text{TP}+\text{FN}} \qquad (5)$$

Specificity: Recall/Accuracy asymmetric measures specificity assesses the ability of the model in correctly identifying negative cases among all the cases that have been classified as negative. It is calculated based on the relation of true negatives with the total of true negatives and false positives. The formula for specificity is:

$$\text{Specificity} = \frac{\text{TN}}{\text{TN}+\text{FP}} \qquad (6)$$

F1-score: The F1-score takes the harmonic average of precision and recall, making it give a balanced measure of the model. It is calculated as follows: It is calculated as follows:

$$\text{F1} - \text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \qquad (7)$$

Area under the Receiver Operating Mark Curve (AUC-ROC): The AUC-ROC values have a range of 0 to 1 where the higher the value the better the differentiation of the classes.

Loss: It is calculated as how much the predicted output of the model differs from the actual output value

Confusion Matrix Analysis: Chaotic systems provide a detailed overview of the model's predictions, showing the number of TP, TN, FP, and FN. Clutter grids can be used to determine various performance metrics such as precision, accuracy, summary, and explicitness.

Our experimental setup made use of the fully cloud-based Google Colab Pro+ platform. Leveraging this highly customized environment, the training of machine learning models was not only expedited but also significantly more efficient compared to conventional setups as shown in Fig 5.

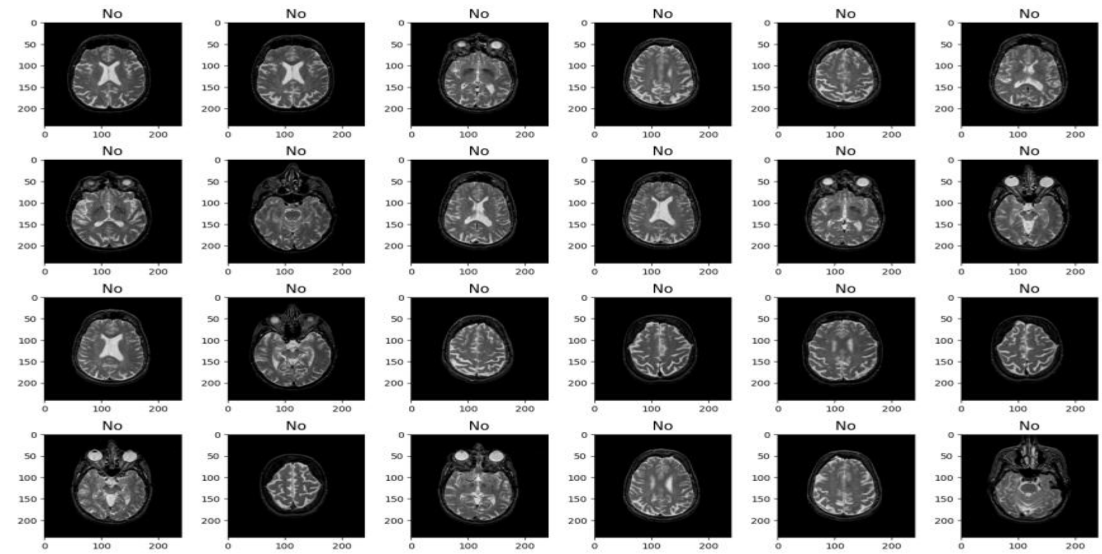


Fig 4: Input Image

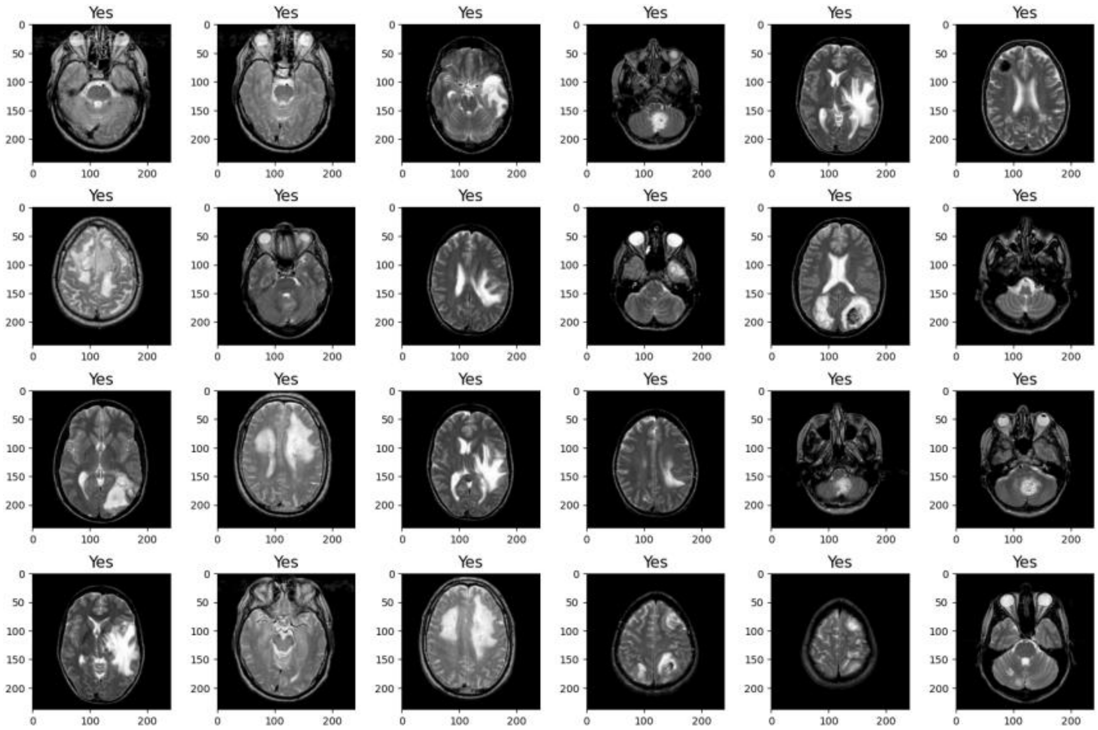
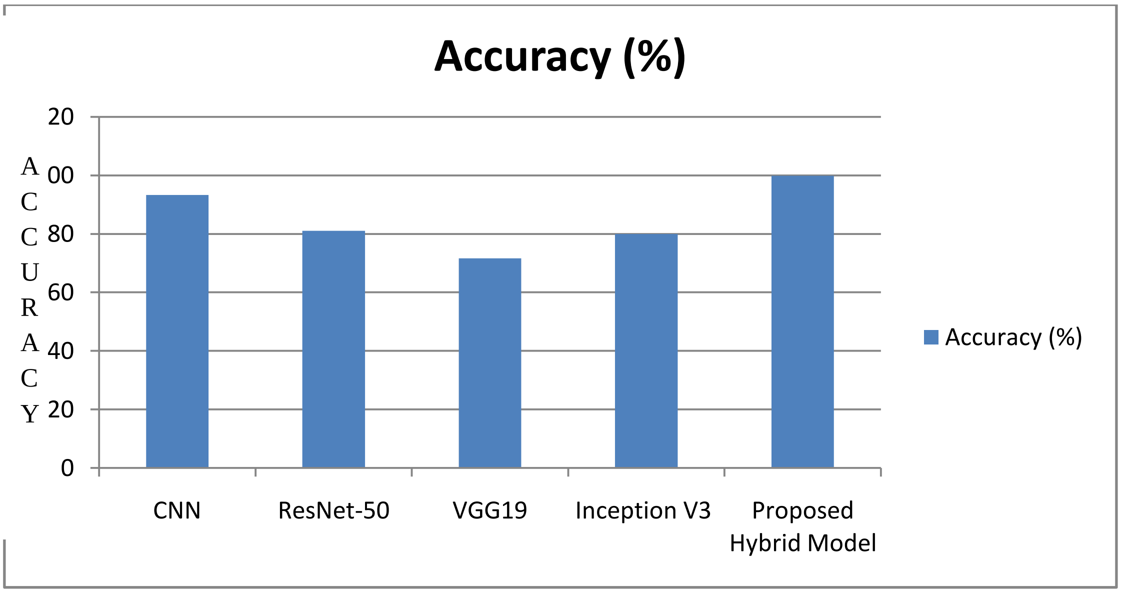


Fig 5: Final processed image

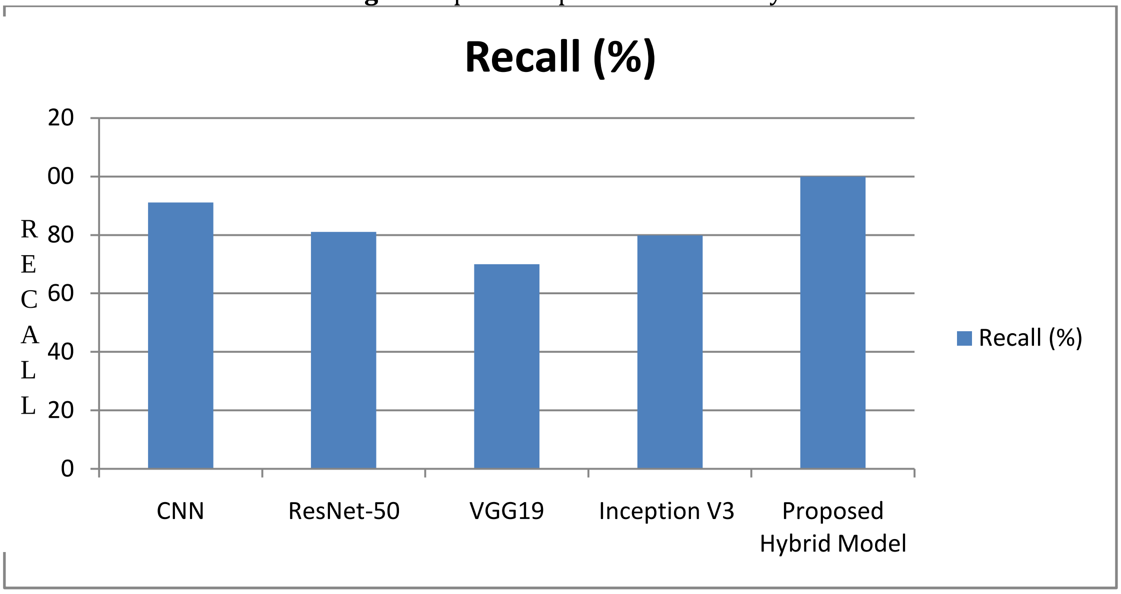
Table 2: Deep Learning Model's Performance On Mri Identification

Models	Accuracy (%)	AUC (%)	Recall (%)	Loss
CNN	93.30	98.43	91.13	0.25
ResNet-50	81.10	94.20	81.04	0.85
VGG19	71.60	89.60	70.03	1.18
Inception V3	80.00	89.14	79.81	3.67
Proposed Hybrid Model	**99.9**	100	100	0.02



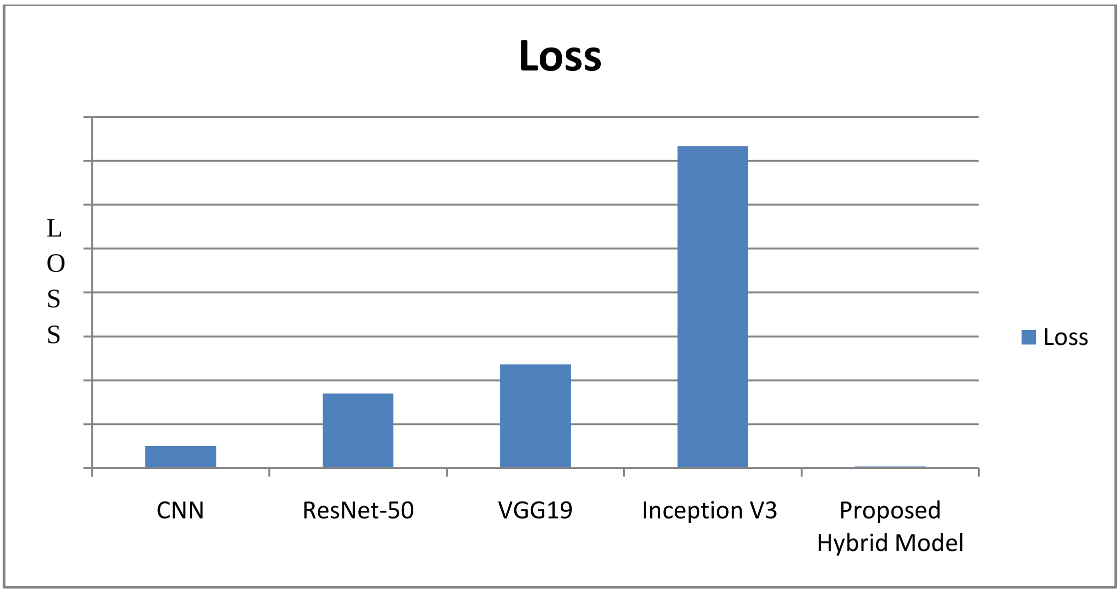
Different Deep Learning Model

Fig 6: Graphical Represent of Accuracy



Different Deep Learning Model

Fig. 7: Graphical Represent of Recall



Different Deep Learning Model

Fig. 8: Graphical Represent of Loss

5. CONCLUSION

The combination of traditional and deep learning image processing techniques when used for enhancement of MRI images provides a strong foundation for improving the images' quality and diagnostic worth. The integration of deep learning capabilities that include automatic feature extraction and high adaptability with the stability of conventional techniques makes these compound models eliminate the weaknesses of the singular approach. Deep learning models are effective for training models on large data sets while traditional methods offer interpretability and repeatability. Altogether, they improve the signal-to-noise ratio, contrast, and resolution of MRI images to help physicians make more accurate diagnoses. This integration does not only enhance the MRI imaging but also create the basis for other higher level automated medical imaging. The future studies should concentrate on the improvement of these combined approaches, increasing the speed of computations, and extending the usage of these approaches to different types of images and diseases. The upgrade of MRI is a basic part of exact clinical diagnostics, empowering better perception of physical designs and working on the recognition and investigation of obsessive circumstances. This paper investigated an incorporated methodology that consolidates customary image handling strategies with DL procedures to improve MRI image quality. Customary strategies, for example, sifting, histogram adjustment, and edge discovery, have for quite some time been viable in diminishing commotion, further developing differentiation, and featuring fundamental highlights in X-ray pictures. In any case, their limits, especially in taking care of mind boggling and fluctuated picture information, require the joining of additional modern methodologies. These models can perform errands, for example, super-goal, and antiquity decrease with momentous exactness and effectiveness.

REFERENCES

[1] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In Medical Image Computing and Computer-Assisted Intervention (pp. 234-241). Springer, Cham.

[2] Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. IEEE Transactions on Image Processing, 13(4), 600-612.

[3] Chen, Y., Shi, F., Christodoulou, A. G., Xie, Y., & Zhou, Z. (2019). Efficient and accurate MRI super-resolution using a generative adversarial network and 3D multi-level densely connected network. Medical Image Analysis, 55, 93-104.

[4] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[5] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 25, 1097-1105.

[6] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708).

[7] Dong, C., Loy, C. C., He, K., & Tang, X. (2016). Image super-resolution using deep convolutional networks. IEEE Transactions on Pattern Analysis and Machine Intelligence, 38(2), 295-307.

[8] Wang, H., Xia, Y., Zhang, L., & Liang, J. (2019). Denoising and enhancement of MRI based on modified deep learning framework. Journal of Ambient Intelligence and Humanized Computing, 10, 1635-1646.

[9] Deshmukh, N. R., & Patil, V. (2019). A review on deep learning and traditional methods for image processing. In 2019 International Conference on Computing, Power and Communication Technologies (GUCON) (pp. 876-881). IEEE.

[10] Lai, W. S., Huang, J. B., Ahuja, N., & Yang, M. H. (2017). Deep laplacian pyramid networks for fast and accurate super-resolution. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 624-632).

[11] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A., Ciompi, F., Ghafoorian, M., ... & van Ginneken, B. (2017). A survey on deep learning in medical image analysis. Medical Image Analysis, 42, 60-88.

[12] Deshmukh, N. R., & Patil, V. (2019). A review on deep learning and traditional methods for image processing. In 2019 International Conference on Computing, Power and Communication Technologies (GUCON) (pp. 876-881). IEEE.

[13] Tustison, N. J., Avants, B. B., Cook, P. A., Zheng, Y., Egan, A., Yushkevich, P. A., & Gee, J. C. (2010). N4ITK: improved N3 bias correction. IEEE Transactions on Medical Imaging, 29(6), 1310-1320.

[14] Sachdeva S, Ali A (2021)," A Hybrid approach using digital Forensics for attack detection in a cloud network environment", International Journal of Future Generation Communication and Networking 14(1):1536–1546.

[15] S. Sachdeva, Aleem Ali," Machine learning with digital forensics for attack classification in cloud network environment", International Journal of System Assurance Engineering Management, Springer, 2021. https://doi.org/10.1007/s13198-021-01323-4.