Modified VGG-19 Deep Learning Strategies for Parkinson's Disease Diagnosis - A Comprehensive Review and Novel Approach

Aruna Kokkula [1*], P. Chandra Sekhar [1], T. M. Praneeth Naidu [1]

Abstract

Background: The incidence of neurodegenerative diseases like Alzheimer's and Parkinson's is increasing among ageing populations in industrialized nations, highlighting the need for better and faster diagnostic tools for early detection. Advanced machine learning algorithms have emerged to assist in the categorization and preliminary risk evaluation of Parkinson's disease (PD) patients, leveraging publicly available data. This study aimed to enhance PD diagnosis using Magnetic Resonance Imaging (MRI) by examining deep learning architectures and evaluating their performance using various criteria. Methods: This study developed an accurate PD detection technique using MRI scans. A novel deep learning system was implemented, incorporating a Fully Connected (FC) layer and dropout, with a proposed VGG19 model that evolves through normalization and incrementing layers. These enhancements improve model normalization, ability, and task-specific flexibility. The training process utilized MRI scans from both healthy individuals and PD patients. Results: The network was tested on extensive MRI datasets for tasks including dementia grading, brain tumor classification, and disease classification, using 10-fold cross-validation. The Attention Feature Fusion VGG19 (AFF-VGG19) network achieved an accuracy of 0.984 in differentiating between three classes of brain tumors, 0.976 in distinguishing between Alzheimer's and Parkinson's diseases, and 0.977 in grading dementia cases. Conclusion: This research suggests that integrating attention modules, feature-fusion blocks, and baseline convolutional neural networks significantly enhances the performance of deep learning models in diagnosing neurodegenerative diseases using MRI scans.

Keywords: Deep learning, Parkinson's disease, VGG19, MRI images.

Significance | This study showed VGG-19 model with superior training speed, accuracy, and iteration efficiency, significantly enhancing Parkinson's disease diagnosis from MRI images.

*Correspondence. Aruna Kokkula, Department of ECE, University College of Engineering, Osmania University, Hyderabad, India, 500007.
E-mail: kokkulaaruna@gmail.com

Editor Deepak Kumar Sahu, And accepted by the Editorial Board Mar 04, 2024 (received for review Jan 10, 2024)

Introduction

The PD becoming a risk for the clinical experts, as it progresses quickly, thus affecting PD persons badly (Kaplan et al., 2022). It is a progressive neurological state, that can be categorized through the motor and non-motor signs. Motor indications may influence the daily routine activities, like stiffness and bradykinesia. Non-motor indicators (Solana-Lavalle and Rosas-Romero 2021) includes mental illness, problems with movement, and illnesses such as stress, as it will be difficult for the detection and treatment.

Initial PD may be uncritical, but if it is not properly treated on time (Noor et al., 2020). It is a serious issue as well as complicated process depends on the medical test, and doctor's experience (Shu et al., 2021). Millions of people got suffered by this PD in both motor and

Author Affiliation.
[1] Department of ECE, University College of Engineering, Osmania University, Hyderabad, India, 500007.

Please cite this article.
Aruna Kokkula, P. Chandra Sekhar (2024). Modified VGG-19 Deep Learning Strategies for Parkinson's Disease Diagnosis - A Comprehensive Review and Novel Approach, Journal of Angiotherapy, 8(3), 1-11, 9559

non-motor signs (Pourzinal et al., 2022). Evidence-strategy is applied for the diagnosis.

One feature of ML is that sets it apart from the others is its capacity to extract complex patterns from massive datasets. In the area of medical testing, DL has emerged as a potentially useful technique, especially for the diagnosis of PD (Wingate et al., 2020). The most minor PD signs can be detected by this type of apparatus (Camacho et al., 2023), (Wang et al., 2023). It also has the ability to sift through enormous volumes of patient data, including data from electronic devices, imaging results, and clinical records. DNN (Deep NN) power is applied to do this. An early and accurate diagnosis is made possible by DL algorithms, and this is essential for selecting the most effective course of treatment and enhancing patient outcomes. To do this, they do thorough data analysis.

When Parkinson's disease is in its early stages, it might be mistaken for other diseases. Atypical Parkinsonian disorders (APDs) are common, leading many to mistake PD for one of these other diseases. For example, disorders like corticobasal degeneration (CBD), multiple system atrophy (MSA), and progressive supranuclear palsy (PSP) are examples. Further complicating matters is the difficulty of classifying PD clinical entities. Moreover, APD and PD categorization are crucial for precise prodromal diagnosis and therapeutic selection. Magnetic resonance imaging (MRI), a game-changing and extensively utilized diagnostic tool, has shown great promise in the detection of Parkinson's disease (PD). Typically, MRI pictures might vary in resolution, contrast, and signal-to-noise ratio as a result of scanning procedures, hardware acquisitions, or both. These elements affect how these pictures are put together, which in turn affects how well training models work. Deep Learning allows for a better quality learning representation of MRI image data (Basnin et al., 2021). It allows for various degrees of abstraction, accomplished by multiple processing of layers, which is why we are applying it in this study. This results in the best model performance rate when compared to other learning methods. The MRI data samples are trained using an LSTM-integrated DenseNet Model. Because DenseNet uses the input from each layer as the output from the next layer, the use of learning parameters is drastically minimized. Making it possible to reuse and ensure that significant characteristics may flow freely across the layer without loss. Additionally, there is a system within each layer that chooses characteristics based on how well they are suited to class prediction in terms of their temporal relatedness. In order to find out how these properties rely on one another over time, they are fed into the LSTM layer. Exposing it to more intricate temporal dynamics of images improves the model's Learning and classification capabilities (Sailaja and VenuGopal , 2023). Along with this, LSTM is able to train datasets well since it addresses back prop and vanishing gradient difficulties. For further assessment, the proposed model is compared to various CNN state-of-the-art models, namely DenseNet, VGG19, InceptionV3, ResNet, and MobileNet.

This research investigates the use of DL strategies for PD treatment, offering an overview of the approaches and developments that might remake this industry's diagnostic area. The probe specifically focuses on PD. This initiative looks at previous research in great detail and develops novel methods and strategies in an attempt to solve the difficulties and complications involved with PD detection. The study's ultimate objective is to improve patient care by advancing our understanding of this debilitating neurological condition.

Literature Review

Using DL techniques, Tarjni Vyas et al. (Vyas et al., 2022) exhibited two cutting-edge approaches.

Convolution neural networks (CNN) in two dimensions and three dimensions are employed, and they are trained using MRI images performed in the axial plane. Images obtained via the PPMI (Parkinson's Progression Markers Initiative) were utilized in the construction of the dataset. This study makes use of four different pre-processing approaches: histogram matching, Z-score normalization, bias field correction, and picture scaling. Each of these methods is described below. Inaccurate training models were almost often produced by pre-processing approaches. DL is a strong notion because it does not rely on a single or a small number of essential characteristics to make all class detection analyzed by the method. Instead the model probably considered multiple traits from different brain levels.

A DL NN (Neural Network) was used by (Sivaranjini and Sujatha 2020) to classify MR (Magnetic Resonance) images of individuals with PDs and healthy controls. The use of the CNN (Convolutional Neural Network) architecture known as AlexNet has been employed to enhance the diagnostic accuracy of Parkinson's disease. The MR pictures undergo training using a transfer learning network, and afterwards undergo testing to get accuracy values.

(Loh et al., 2021) examined a total of 63 scholarly articles, published from January 2011 to July 2021. The motion symptoms such as gait, handwriting, speech, and Electromyography (EMG) were also included in the development of these models. The present research allows us to ascertain the most effective DL model documented for each modality, while also emphasizing the existing constraints that impede the widespread use of Computer-Aided Diagnosis (CAD) technologies in the healthcare sector. In conclusion, we put up novel avenues for further research on DL in the automated identification of PD, with the aim of enhancing the effectiveness, versatility, and significance of these apparatuses in facilitating early diagnosis of PD on a worldwide scale.

(Sahu et al., 2022) presented two hybridized deep learning techniques, namely RA and ANN, is performed to boost the precision of illness identification via the calculation of probabilities.

The collective benefits of individual methodologies within the current techniques are recognized in this particular context to achieve precise assessment of probabilities. The process of data preprocessing and subsequent chances estimate of the preprocessed information is conducted within the realm of RA (Research Analysis). The second established methodology involves the identification of persons with PD by a comparison with a predetermined threshold value of a neuron.

(Zhao et al.,2022) proposed a system that utilized a greedy algorithm to merge models from various sites into a single composite model. The study involved 305 patients with PD, aged 59.9±9.7 years on average, and 227 healthy control subjects, aged 61.0±7.4 years on average. The subjects were selected from three distinct retrospective studies.

(Pahuja and Prasad 2022) introduced two distinct frameworks, namely the feature-level and modal-level frameworks, both of which rely on DL techniques. These frameworks were used to categorize the provided participants into two groups: individuals with PD and individual health. The dataset utilized in this study consisted of neuroimaging information, namely T1 weighted MRI scans and SPECT scans, as well as biological characteristics derived from cerebrospinal fluid (CSF) samples.

(Noor et al., 2019) analyzed the methodological research articles that proposed to diagnose neurodegenerative illnesses with deep machine learning approaches based only on MRI data.

(Madan et al., 2021) developed a novel algorithm and used to determine the optimal range of MRI slices at the individual level for the purpose of distinguishing between patients in the NC (Normal Cohorts) and those diagnosed with PD. Furthermore, the suggested study has focused on efficiently handling data leaks and utilising Stratified k-fold CV (Cross-Validation) to evaluate the model's generalizability.

The technique outperforms the current state of the art in terms of training speed, accuracy, and the number of training samples each iteration. In order to make model training easier, the raw data gathered from various sensors is first standardized and then integrated into a graph.VGG-19 is a 19-layer convolutional neural network. A pre-trained network trained on over a million photos is available for loading from the ImageNet database. Images may be classified into one thousand item categories by the pre-trained network. Among these categories are several animals, keyboards, mice, and pencils. Extremely Deep Convolutional Networks for Large-Scale Image Recognition VGG 19-layer model. A minimum input size of 32x32 is necessary for the model. Sixteen convolution layers are organized into five blocks in VGG-19. The Maxpool layer, which appears after each block, doubles the size reduction of the input picture and doubles the number of filters in the convolution layer.

VGG-19 Image Classification Model

The VGG-19 model has been known for being simple and architecturally constant which makes it useful for a wide range of computer vision applications, including object detection and classification of pictures. A detailed explanation of the VGG-19 algorithm's functioning is listed below:

The Input Layer: The model requires an input image that can be modified to accommodate images of different sizes. An image's pixel count is typically 224 × 224. The RGB (Red, Green, and Blue) channels are the 3 parts of the input image.

1. Convolutional layer: CNN's (Convolutional NN) convolutional layer is an important component. It uses a convolution method, where a small window called a filter or kernel runs across the input data while the local input sector is multiplied element by element. This method creates an FM (Feature Map) that highlights specific features in the input.

2. The total number of element-wise products between the local region of the input f and the filter g can be used to numerically express the convolution operation (f*g)(s,t) at location (s,t):

$$(f * g)(s,t) = \sum_{i=1}^{m}\sum_{j=1}^{n} f(s+i-1, t+j-1)g(i,j) \quad (1)$$

Here:

- Input data f, i.e., image.
- g is the filter (kernel).
- The position in the output FM can be symbolized as (s, t)
- The dimensions of the filter are m and n.

Stride: It establishes how much the filter moves across the input data. The prolonged phase results in a smaller final product.

Padding: Adding extra pixels across the input data to prevent data loss at the edges is known as padding. One frequent padding technique is zero-padding.

Activation Function - ReLU (Rectified Linear Unit) introduces non-linearity by setting all negative pixel values in the feature map to zero. The mathematical formula for the ReLU activation function is:

$$ReLU(x) = max(0, x) \quad (2)$$

Here:

- The input to the activation function is x.
- In simple terms, max sets all negative values to zero by returning the largest value among 0 to x.

The convolution process generates an output where each element is carried out with a different ReLU activation function. This characteristic allows the network to be non-linear, which makes it easier to discover complex relationships in the data. The following is a representation of the Convolutional Layer with ReLU activation:

ReLU(Conv(f,g))

Here:

- f is the input data.
- g is the filter (kernel).

1. Batch Normalization layer: This layer normalises each layer's inputs in order to deal with the internal covariate shift problem. It works with a mini-batch of training samples, subtracting the mean from each feature inside the mini-batch and dividing the result by the feature's standard deviation. The network is then given the opportunity to determine the ideal scale and shift for every feature through a scaling and shifting operation.

Considering the feature x in a mini-batch, the Batch Normalisation operation can be expressed as follows:

$$\text{Batch Norm}(x) = \gamma \frac{x-\mu}{\sigma} + \beta \quad (3)$$

Here:

- x - input.
- μ mean of x.
- σ standard deviation.
- γ is a scale parameter.
- β is a shift parameter.

2. Max-Pooling Layers: Following a number of the convolutional layers, the max-pooling layers are included in VGG-19. When using max-pooling, FMs spatial dimensions may be reduced whereas the essential information is kept intact. It contributes to the network's overall reduction in the computing complexity of its operations.
3. Dense Layers: Following the convolutional layers is a set of three fully linked layers in VGG-19. These levels are the standard feed forward neural network layers, which means that every neuron in these layers is coupled to every neuron in the layers below and above it. The last layer of the fully linked network contains the same number of neurons as there are categories in the issue being solved.
4. Dropout layer: It is a regularization method frequently employed in neural networks to combat overfitting and enhance the model's generalization capacity. Dropout layers are typically inserted between the FC (Fully Connected) layers of a NN. The dropout probability (p) is a hyperparameter that can be tuned based on the specific characteristics of the dataset and the network architecture.

During training, the VGG-19 model is taught to recognize objects by making use of labeled data (input photos with associated class labels) and a loss function, such as category cross-entropy. Backpropagation and gradient descent are used by the model to make any necessary adjustments to its weights and biases. It makes the process of classifying images into the relevant categories possible.

VGG-19 is a deep CNN architecture-based classification system. It is composed of FC (Fully Connected) layers for classification, max-pooling layers for minimizing the spatial dimensions, and convolutional layers for extracting features. The model is trained on labeled information that can be applied as a feature extractor in TL (Transfer Learning) scenarios or tailored for specific purposes. Its growing prominence in the field of computer vision can be attributed to its standard architecture, which makes it easy to understand and apply.

Hypothetical thinking encompasses a wide range of mental processes, including speculation, imagination, logical and counterfactual reasoning, and generalizations about potential future states of the world. Over 8.5 million people were diagnosed with PD in 2019, according to research undertaken by the World Health Organisation. The fact that only 4% of those afflicted are less than 50 years old suggests that the prevalence of this disease increases with age. On a global scale, PD affects millions of people and ranks second among neurodegenerative disorders, behind only Alzheimer's disease. Due to the early stages of treatment, doctors are unable to do much to alleviate the symptoms of this condition at this time. Most of the time, the patient's medical history plays a deciding role in making a diagnosis of this illness. It is critical to find a simple and reliable way to diagnose this disease since invasive diagnosis and treatment are time-consuming and expensive.

A wide range of symptoms, including mood disorders and depressed states, may be noted in association with nonmotor manifestations displayed by patients with PD. In addition to language and other pertinent aspects, a patient's facial expressions may reflect these symptoms. This study aims to fill a gap in the literature by investigating the effects of PD on motor and nonmotor functions via the use of handwriting modelling approaches, spirals, and waves. Specifically, it will examine the effects of PD on drawing spirals and waves. The field of biomedical and medical image analytics has been greatly transformed by the application of deep learning (DL) models. Many domains have made use of DL models, including detection, segmentation, and illness categorization. Disease classification accuracy is enhanced by DL models' exceptional capacity to extract high-level features. One possible explanation is their remarkable ability to generalize. Another area that has benefited greatly from convolutional neural networks (CNNs) is medical image processing. In several medical imaging applications, CNNs have shown remarkable success.

Proposed Modification in VGG19 model

In comparison to the conventional VGG19 model, the suggested model, a variant of that model includes more layers and modifies the architecture. Below is a summary of the suggested model and a clarification of the ways it differs from the conventional VGG19.

An Additional Global Max Pooling Layer:

The suggested model introduces the global max-pooling layer after the fifth and final block of convolutional layers. To reduce the spatial dimensions to a single number for every FM, this layer performs global max-pooling for all FMs.

Normalization Layer:

In the next step, which comes after the global max-pooling layer, a batch normalization layer is implemented. By standardizing the activations in each mini-batch, batch normalization makes training more stable and increases the generalization of the network.

An Additional Dense Layer:

Following the batch normalization layer is an additional fully connected layer called dense_4 that has 1024 neurons and an activation function called ReLU. This layer is part of the proposed model.

Dropout Layer:

Following the insertion of the extra completely linked layer, the next layer to be added is a dropout layer with a dropout rate of half. Dropout is a strategy for regularization that deactivates neurons at random during training in order to assist avoid overfitting. Dropout works to prevent overfitting.

Dense layer:

The model is completed with the final dense layer with three output neurons that have a softmax activation function. This layer is completely coupled. This is used in classification jobs involving three different categories. Figure 1 shows the proposed model architecture.

The following are some benefits of applying the modified VGG model:

- **Improved Regularization**: The addition of a dropout layer makes the model more resilient to data that has not yet been seen. This is accomplished by enhancing the model's capacity to generalize and lowering the danger of overfitting.
- **Increased Capacity**: The addition of one more fully connected layer to the model (dense_4) results in an increased capacity for the model to recognize intricate patterns in the input, which may result in an improvement in the model's overall performance on difficult tasks.
- **Normalization**: The use of batch normalization contributes to the preservation of stable training dynamics and has the potential to hasten the process of convergent learning.
- **Fine-tuning**: The improved model may be more flexible in terms of transfer learning situations. The extra layers make it possible to hone down on certain jobs while still maintaining the information gained from the VGG19 levels that were pre-trained.
- **Potential for Improved Accuracy**: Depending on the Particular job and Dataset, the Modified VGG Model May Achieve Higher Accuracy Than the Standard VGG19 Due to the Incorporation of These Additional Layers and Regularization Techniques The potential for improved accuracy of the modified VGG model depends on the particular job and dataset.

Our feature vector, φ (l1), was created by activating the FC layer in VGG19 for deep feature extraction, and we used the first convolutional layer as an input layer. The size of the feature vector is 1×4096. Multiple sets of pooling layers and stacks of convolutional layers make up the neural network used for feature extraction. Exactly what it sounds like the convolution layer applies a convolutional transformation to the picture. To create the final picture, the neural style transfer paper employs feature maps that are produced by the intermediate layers of the VGG-19 network. Using the characteristics derived from the convolution layers of a VGG network, this architecture saves style and content pictures as input. VGG-19 is a 19-layer convolutional neural network. A pretrained network trained on over a million photos is available for loading from the ImageNet database. Images may be classified into one thousand item categories by the pretrained network. Among these categories are several animals, keyboards, mice, and pencils. Among the 19 connection layers that make up the VGG-19 deep learning neural network are 3 completely connected layers and 16 convolution layers. Fully linked layers will use the information extracted from the input photos by the convolution layers to categorize the leaf images. This technique outperforms the current state of the art in terms of training speed, accuracy, and the number of training samples in each iteration. In order to make model training easier, the raw data gathered from various sensors is first standardized and then integrated into a graph.

Hyperparameters

Convolutional Neural Networks (CNNs) have several hyperparameters that need to be tuned to achieve optimal performance. Here are some of the key hyperparameters in a typical CNN:

- **Learning Rate:** It is an essential hyperparameter in the training of DL models. It specifies the magnitude of the increment by which the model adjusts its parameters throughout the optimization process. Excessive learning rates might cause the model to surpass the optimum solution and result in failure to converge. On the other hand, if the value is very low, the model may experience prolonged convergence or get trapped in a poor solution.
- **Categorical Cross Entropy:** It is a popular loss function used in classification issues, particularly for tasks involving several classes. It is possible to quantify the difference between expected and actual distributions for probabilities.

Adam Optimizer: Frequently, the Adam optimization technique is used to repeatedly adjust network weights using information from training to minimize the loss function. It incorporates ideas from both adjustable LR (Learning Rate) and momentum-based approaches. The algorithm determines a unique adaptive learning rate for every parameter and modifies it based on the gradient magnitudes during training. More flexibility than traditional GD (Gradient Descent) optimizers allows for faster convergence and improved efficiency on a wide range of problems. To train and optimise DL models, three key elements are required: the learning

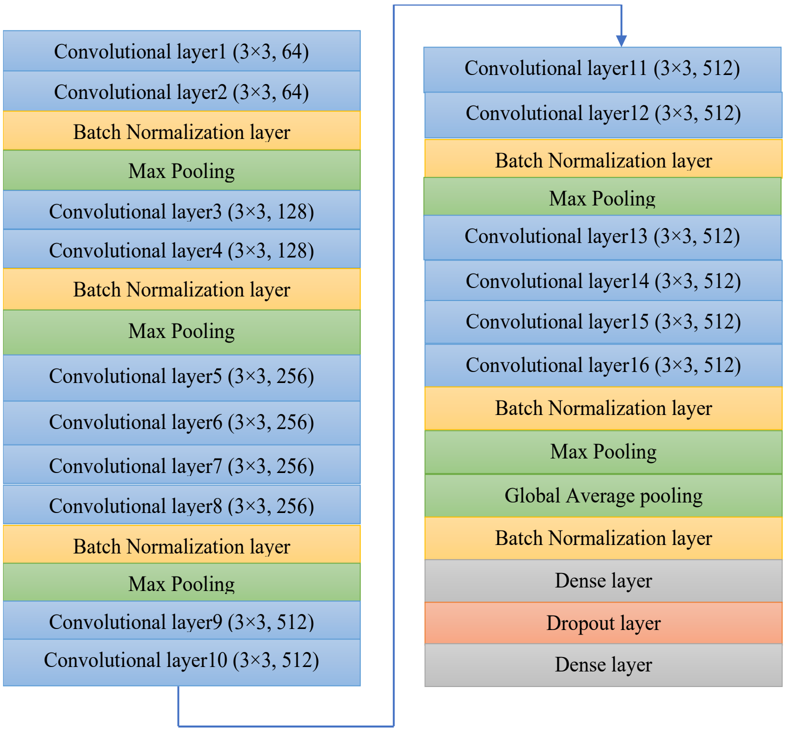

Figure 1. Proposed model architecture

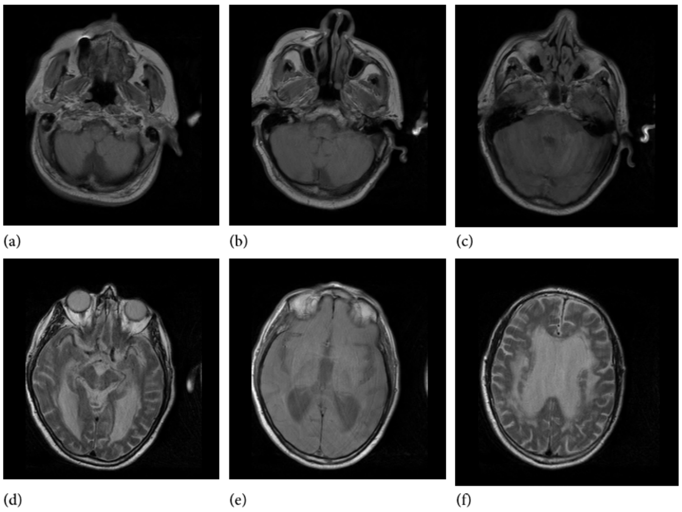

Figure 2. MRI of patients with PD

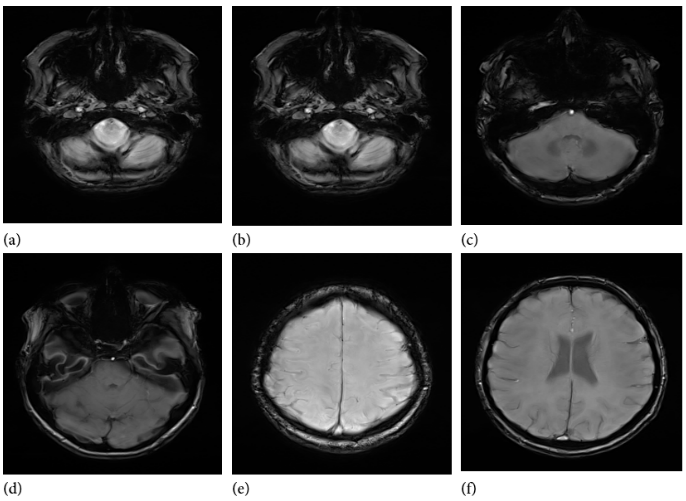

Figure 3. MRI of normal people

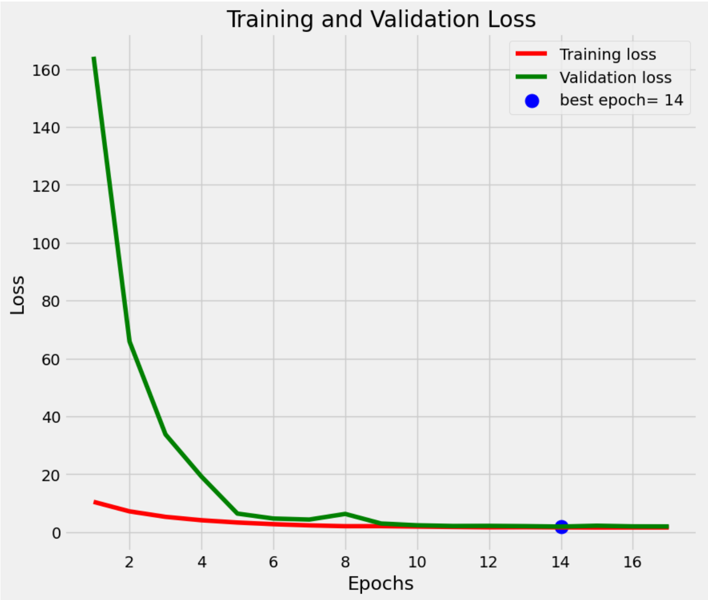

Figure 4. Training and validation loss

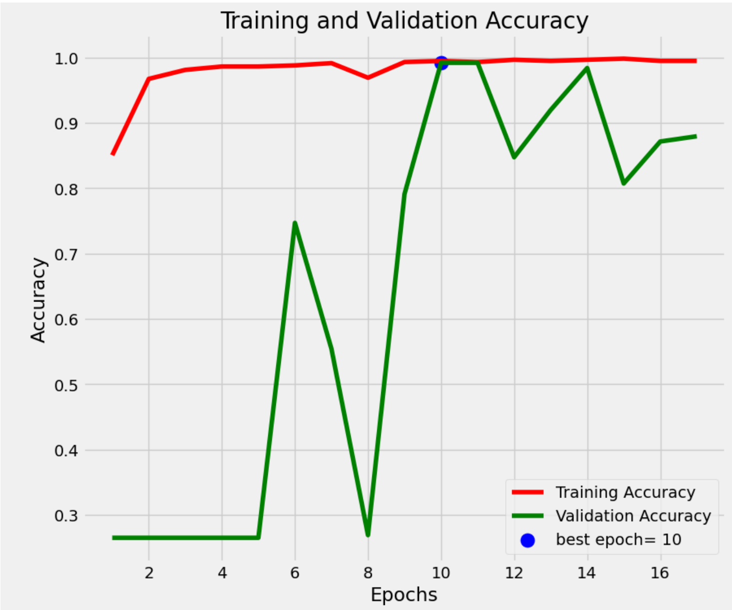

Figure 5. Training and validation accuracy

Table 1. Comparative performance metrics

Model	Precision	Recall	F1-Score	Accuracy
DenseNet201 [17]	0.7	0.71	0.71	0.71
ResNet50V2 [18]	0.701	0.720	0.710	0.71
InceptionV3 [19]	0.863	0.842	0.849	0.851
Proposed VGG19	0.992	0.976	0.977	**0.984**

rate, the Adam optimizer (AO), and the category CE (Cross-Entropy). They significantly affect the models' convergent, accurate, and generalizable to new data. The model's effectiveness may be strongly impacted by properly optimizing these hyperparameters.

The pre-processed photos may have their characteristics automatically extracted using automated diagnostic deep model techniques. Quantitative measurement of brain tissue volume and large-scale research on intracranial volume need accurate and automated brain tissue segmentation in imaging, including GM, WM, and CSF. Both the pattern recognition technique and the Atlas-based method are significant examples of classical methods for brain tissue segmentation. Intensity information in target pictures is matched using atlas-based approaches. Techniques such as the pattern recognition technique and atlas-based approaches are often used for brain segmentation due to limitations in recognition and differences in reliable real data. The hippocampus, on the other hand, is both small and very dynamic; thus, their results are not dependable. Using a set of image intensity settings, pattern recognition algorithms classify brain tissues. A biomarker for Alzheimer's disease (AD) has recently been proposed, namely hippocampal atrophy. A variety of brain tissues encase the hippocampus, which is part of the limbic system. This initiative aims to identify motor problems connected to PD via noninvasive tracking of the vital signs of patients using reusable and handheld gadgets combined with accelerometers and other inertial measurements. When applied to certain motions, these methods reliably identify traits that indicate an abnormal tremor during particular tasks [6]. The proliferation of connected devices has provided researchers with an enormous data set from sensors, which they can now analyze and mine for insights. When multiple devices are used to track and transmit different types of data at the same time, the resulting data is often very dimensional and very large in volume. When faced with such data, one of the first options for deciphering it and finding hidden patterns for a more thorough and reliable study is automated Learning. Examples of algorithms that benefit from large-scale data include neural networks, which are able to handle the constraints of this kind of data while simultaneously boosting their performance. Machine learning methods for PD data collected from sensors have made tremendous strides in this field. Increasingly, sensor-based systems using predictive methods are necessary due to the increasing complexity of both accessible data and technology.

Experimental results

This section displays the results of the tests that were run to confirm the suggested technique. The input for the model consists of brain MRI data from both PD patients and healthy individuals. These pictures are necessary for an ML algorithm to be trained. 582 images from MRI that were specially chosen for the algorithm's training make up the dataset. persons with PD and persons without it are represented in the two categories of the photographs. Moreover, an additional set of 249 MRI pictures is intended to assess how well the learnt method is implemented. This section estimates the degree to which the model can apply its learned patterns to fresh, unknown data, hence offering a trustworthy evaluation of its predictive power.

The paper provides a detailed description of how the DL models were trained and tested. We set out to build a model that could reliably diagnose PD by analyzing spiral and wave drawings. This is important since elderly patients often become tired quickly, making it unrealistic to expect them to complete a large number of drawing assignments while they are in a primary care environment. We also showed that DL models have diagnostic potential for PD, which is important from a clinical perspective. Using analysis of spiral and wave drawing pictures, the DenseNet201 and VGG19 DL models have been very effective in reliably discriminating between healthy people and patients with PD. Here, we compared the results of enhanced DL models to those of other existing systems and an open-source program that made use of the same dataset. When applied to two distinct picture drawings, our findings show that the suggested approach detects PD with a high degree of accuracy (Figure 2, Figure 3).

To build and evaluate the model, the dataset must be divided into training and testing sets. The training set's reasonably equal photo distribution between the two groups helps the algorithm comprehend the special traits connected to Parkinson's disease. The approach looks for features that differentiate people with PD from people without it using a large variety of samples. Principal objective: create a DL methodology capable of categorizing brain MRI pictures and distinguishing between PD patients and normal people.

Diagnosing Parkinson's disease is currently not possible with a single test. A neurologist is a medical specialist who specializes in diagnosing disorders of the neurological system. Medical history, symptoms, and neurological and physical examinations are the three main components of a Parkinson's disease diagnosis (Rajanbabu et al., 2022). The diagnosis of PD cannot be confirmed by any conclusive testing at this time. One of the most popular ways to measure PD patients' health is using the UPDRS. Possible causes of PD symptoms include structural lesions, infections, metabolic problems, hydrocephalus, toxins, and other neurodegenerative diseases, as well as genetic disorders. Here are some possible treatments for Parkinson's disease: Prescribed drugs, Surgery, Diet, exercise, PT, OT, and speech therapy are all examples of supplementary and supporting treatments.

The partitioning of this dataset into separate training and testing sets provides the basis for training an algorithm to identify specific characteristics in MRI images of the brain.

The allocation of 582 training images into two distinct categories, together with an extra 249 images for the purpose of testing, guarantees a thorough assessment of the methods capability in creating exact predictions. The successful creation of a precise model has the ability to greatly impact an initial detection and PD diagnosis. This may be achieved by using non-invasive brain imaging methods to enhance patient care and improve medical results. Once data is trained by using proposed modified VGG19 model the validation and loss and accuracy are obtained and depicted in Figure 4 and 5.

When it comes to the training of ML and DL methods, 2 key parameters are training loss and validation loss. During the process of the model being trained, they play a significant part in evaluating the method's efficiency as well as its capacity for generalization. To lessen this loss, the weights of the technique are adjusted at the start of every training iteration, or epoch. The goal of the procedure is to iteratively lower the training loss, which indicates that the model is getting better at matching the information used for training. While the technique is being trained on the validation dataset, it is frequently checked (usually after each epoch). The validation loss is determined by contrasting the model's predicted values with the actual values in the validation dataset.

Important factors in the training and assessment of ML and DL models include training and validation accuracy. They provide essential data about the effectiveness and generalization of a model. One can use the model's training accuracy to determine whether it has fit the training set. A high degree of accuracy throughout training indicates that the output of the algorithm and the training data set are well matched. During the validation phase, the algorithm's predicted accuracy is assessed with data that has never been seen before. Accuracy in variation provides an evaluation of the algorithm's predicted effectiveness when applied to information from the real world.

Performance metrics

Metrics for performance are essential for assessing the effectiveness of DL algorithms, since they provide important details regarding the algorithms' anticipated accuracy, particularly for classification tasks. Important measures include accuracy, precision, recall, and F1score. Precision indicates the rate at which the algorithm minimises FP (False Positives). Precision is defined as a percentage of TP (True Positive) predictions to all positive predictions. A high recall score indicates successful detection while reducing (False Negatives). Recall measures the model's ability to identify every positive case. The F1 score, which is determined via the harmonic mean of memory and precision, achieves its greatest value when recall and precision are at the greatest points. This gives a fair evaluation of the algorithm's correctness.

For data findings to be considered statistically significant, analysts must conclude that they cannot be explained by random chance alone. The analyst arrives at this conclusion via the use of statistical hypothesis testing. It is feasible to conclude statistical significance using the confidence interval. It is safe to presume a statistically significant result if the confidence interval does not include the value of zero effect. To quantify the degree of uncertainty associated with a sample variable, statisticians use confidence intervals. To find out how well a sample reflects the true value of a population variable, a researcher might, for example, construct a confidence interval for each sample after randomly selecting samples from the same population. When comparing two models, even when one is nested beneath the other, the likelihood ratio test may be used. For any given model, the likelihood is the same as the chance that the model would replicate our observer's findings exactly.

A comparison analysis between the recommended method and alternative SOTA (State-Of-The-Art)DL algorithms is shown in Table 1.

Table 1 compares the accuracy, recall, F1-score, precision, and other performance metrics of different DL models. Every aspect of performance for both ResNet50V2 and DenseNet201 is comparable, with DenseNet201 marginally exceeding ResNet50V2 in terms of precision. InceptionV3 exhibits improved precision and recall over previous models, indicating its increased ability to accurately identify positive examples while reducing false positives. The VGG19 algorithm performs better than all others, with precision of 0.992, recall of 0.976, F1-score of 0.977, and accuracy of 0.984. The results demonstrate that the Vgg19 model's remarkable capacity to uncover relevant cases and minimise misclassifications makes it appropriate for the task at hand.

By comparing our suggested model to other current systems, we were able to prove that the DenseNet201 and VGG19 models were accurate in their validation. In order to identify PD in pictures of spirals and waves, the research employed CNN and Resnet50. It was noted that CNN and ResNet50 obtained 90% using a comparable dataset to the one the author used. Using wave pictures, RestNet50 achieved 87% accuracy, in contrast to CNN's spiral images. The authors were able to get a 63.33% accuracy rate by using the lightning CNN model. Finding PD in spiral and wave pictures was made possible via a convolutional neural network (CNN). The findings showed that the CNN model was able to recognize PD from wave pictures with an accuracy of 89% and recognize PD from spiral images with an accuracy of 88%. Training and validation were performed on about 80% of the models.

Conclusion

A novel DL model for PD recognition utilising MRI images is presented in the study. It has been shown that the improved VGG19 architecture which includes more layers and regularisation methods performs exceptionally well in this crucial medical application. The model had a 98% accuracy rate, highlighting its potential as a very reliable tool for initial PD diagnosis. The capacity to detect PD, non-invasively and with remarkable accuracy highlights the significance of this discovery. It allows for immediate treatment and better patient outcomes. Early diagnosis is critical for the management of Parkinson's disease (PD), and our approach has the potential to revolutionise the diagnostic process by making it more accurate and efficient.

Author contributions

A.K. conducted the methodology, collected and analyzed the data, and wrote the original draft. P.C.H. conceptualized the project, supervised the work, reviewed and edited the writing, and managed the project. T. M. Praneeth Naidu developed the methodology, interpreted the data, and reviewed and edited the writing.

Acknowledgment

The authors were grateful to the Department.

Competing financial interests

The authors have no conflict of interest.

References

Basnin, N., Nahar, N., Anika, F.A., Hossain, M.S. and Andersson, K., (2021). Deep learning approach to classify Parkinson's disease from MRI samples. In International conference on brain informatics, pp. 536-547. https://doi.org/10.1007/978-3-030-86993-9_48.

Camacho, M., Wilms, M., Mouches, P., Almgren, H., Souza, R., Camicioli, R., Ismail, Z., Monchi, O. and Forkert, N.D., (2023). Explainable classification of Parkinson's disease using deep learning trained on a large multi-center database of T1-weighted MRI datasets. NeuroImage: Clinical, 38, pp. 103405. https://doi.org/10.1016/j.nicl.2023.103405.

Kaplan, E., Altunisik, E., Firat, Y.E., Barua, P.D., Dogan, S., Baygin, M., Demir, F.B., Tuncer, T., Palmer, E., Tan, R.S. and Yu, P., (2022). Novel nested patch-based feature extraction model for automated Parkinson's Disease symptom classification using MRI images. Computer Methods and Programs in Biome`dicine, 224, pp. 107030. https://doi.org/10.1016/j.cmpb.2022.107030.

Loh, H.W., Hong, W., Ooi, C.P., Chakraborty, S., Barua, P.D., Deo, R.C., Soar, J., Palmer, E.E. and Acharya, U.R., (2021). Application of deep learning models for automated identification of Parkinson's disease: A review. Sensors, 21(21), pp. 1-25. https://doi.org/10.3390/s21217034.

Madan, Y., Veetil, I.K., Sowmya, V., Gopalakrishnan, E.A. and Soman, K.P., (2021). Deep learning-based approach for parkinson's disease detection using region of interest. In Intelligent Sustainable Systems: Proceedings of ICISS, pp. 1-13. https://doi.org/10.1007/978-981-16-2422-3_1.

Noor, M.B.T., Zenia, N.Z., Kaiser, M.S., Mahmud, M. and Al Mamun, S., (2019). Detecting neurodegenerative disease from MRI: a brief review on a deep learning perspective. In Brain Informatics: 12th International Conference, BI, Proceedings 12, pp. 115-125. https://doi.org/10.1007/978-3-030-37078-7_12.

Noor, M.B.T., Zenia, N.Z., Kaiser, M.S., Mamun, S.A. and Mahmud, M., (2020). Application of deep learning in detecting neurological disorders from magnetic resonance images: a survey on the detection of Alzheimer's disease, Parkinson's disease and schizophrenia. Brain informatics, 7, pp. 1-21. https://doi.org/10.1186/s40708-020-00112-2.

Pahuja, G. and Prasad, B., (2022). Deep learning architectures for Parkinson's disease detection by using multi-modal features. Computers in Biology and Medicine, 146, pp. 105610. https://doi.org/10.1016/j.compbiomed.2022.105610.

Pourzinal, D., Yang, J., Lawson, R.A., McMahon, K.L., Byrne, G.J. and Dissanayaka, N.N., (2022). Systematic review of data-driven cognitive subtypes in Parkinson disease. European journal of neurology, 29(11), pp. 3395-3417. https://doi.org/10.1111/ene.15481.

Rajanbabu, K., Veetil, I.K., Sowmya, V., Gopalakrishnan, E.A. and Soman, K.P., (2022). Ensemble of deep transfer learning models for parkinson's disease classification. In Soft Computing and Signal Processing: Proceedings of 3rd ICSCSP, 2, pp. 135-143. https://doi.org/10.1007/978-981-16-1249-7_14.

Sahu, L., Sharma, R., Sahu, I., Das, M., Sahu, B. and Kumar, R., (2022). Efficient detection of Parkinson's disease using deep learning techniques over medical data. Expert Systems, 39(3), p.e12787. https://doi.org/10.1111/exsy.12787.

Sailaja, B. and VenuGopal, T., (2023). Brain MRI Image Classification and Analysis Using Modified ResNet50V2 for Parkinson's Disease Detection. SN Computer Science, 4(6), pp. 854. https://doi.org/10.1007/s42979-023-02313-y.

Shu, Z.Y., Cui, S.J., Wu, X., Xu, Y., Huang, P., Pang, P.P. and Zhang, M., (2021). Predicting the progression of Parkinson's disease using conventional MRI and machine learning: An application of radiomic biomarkers in whole-brain white matter. Magnetic resonance in medicine, 85(3), pp. 1611-1624. https://doi.org/10.1002/mrm.28522.

Sivaranjini, S. and Sujatha, C.M., 2020. Deep learning based diagnosis of Parkinson's disease using convolutional neural network. Multimedia tools and applications, 79(21), pp. 15467-15479.https://doi.org/10.1007/s11042-019-7469-8.

Solana-Lavalle, G. and Rosas-Romero, R., (2021). Classification of PPMI MRI scans with voxel-based morphometry and machine learning to assist in the diagnosis of Parkinson's disease. Computer Methods and Programs in Biomedicine, 198, pp. 105793. https://doi.org/10.1016/j.cmpb.2020.105793.

Vyas, T., Yadav, R., Solanki, C., Darji, R., Desai, S. and Tanwar, S., (2022). Deep learning-based scheme to diagnose Parkinson's disease. Expert Systems, 39(3), pp. e12739. https://doi.org/10.1111/exsy.12739.

Wang, Y., He, N., Zhang, C., Zhang, Y., Wang, C., Huang, P., Jin, Z., Li, Y., Cheng, Z., Liu, Y. and Wang, X., 2023. An automatic interpretable deep learning pipeline for accurate Parkinson's disease diagnosis using quantitative susceptibility mapping and T1-weighted images, 44(12), pp. 4426-4438. https://doi.org/10.1002/hbm.26399.

Wingate, J., Kollia, I., Bidaut, L. and Kollias, S., (2020). Unified deep learning approach for prediction of Parkinson's disease. IET Image Processing, 14(10), pp. 1980-1989. https://doi.org/10.1049/iet-ipr.2019.1526.

Zhao, H., Tsai, C.C., Zhou, M., Liu, Y., Chen, Y.L., Huang, F., Lin, Y.C. and Wang, J.J., (2022). Deep learning based diagnosis of Parkinson's Disease using diffusion magnetic resonance imaging. Brain imaging and behavior, 16(4), pp. 1749-1760. https://doi.org/10.1007/s11682-022-00631-y.