An improved approach for early diagnosis of Parkinson's disease using advanced DL models and image alignment

S. Kanagaraj, M.S. Hema & M. Nageswara Guptha

Published online: 11 Mar 2024.

Submit your article to this journal ↗

Article views: 785

View related articles ↗

View Crossmark data ↗

Citing articles: 3 View citing articles ↗

An improved approach for early diagnosis of Parkinson's disease using advanced DL models and image alignment

S. Kanagaraj[a], M.S. Hema[b] and M. Nageswara Guptha[c]

[a]Department of Information Technology, Kumaraguru College of Technology, Coimbatore, India; [b]Department of Information Technology, Anurag University Medchal, Hyderabad, India; [c]Computer Science and Engineering, Sri Venkateshwara College of Engineering, Bengaluru, India

ABSTRACT

An innovative approach to enhance image alignment through affine transformation, allowing images to be rotated from 0 to 135 degrees. This transformation is a crucial step in improving the diagnostic process, as image misalignment can lead to inaccurate results. The accurate alignment sets the stage for a robust U-Net model, which excels in image segmentation. Precise segmentation is vital for isolating affected brain regions, aiding in the identification of PD-related anomalies. Finally, we introduce the DenseNet architecture model for disease classification, distinguishing between PD and non-PD cases. The combination of these DL models outperforms existing diagnostic approaches in terms of acceptance precision (99.45%), accuracy (99.95%), sensitivity (99.67%), and F1-score (99.84%). In addition, we have developed user-friendly graphical interface software that enables efficient and reasonably accurate class detection via Magnetic Resonance Imaging (MRI). This software exhibits superior efficiency contrasted to current cutting-edges technique, presenting an encouraging opportunity for early disease detection. In summary, our research tackles the problem of low accuracy in existing PD diagnostic models and addresses the critical need for more precise and timely PD diagnoses. By enhancing image alignment and employing advanced DL models, we have achieved substantial improvements in diagnostic accuracy and provided a valuable tool for early PD detection.

I. Introduction

Parkinson's disease (PD), a neurological condition affecting movement with symptoms like tremors and rigidity, necessitates early detection to enhance life expectancy. In this context, brain images acquired through magnetic resonance imaging (MRI) [1] have been pivotal in understanding brain functionality and neurological disorders [2]. PD's hallmark is neuronal cell loss causing synaptic dysfunction and cognitive impairments, making early prediction crucial for effective management despite existing treatment challenges [3]. This neurodegenerative disorder results from dopamine neurotransmitter deficiency in the substantia nigra, posing challenges in early diagnosis due to obscure onset symptoms [4]. Establishing a comprehensive understanding of PD's manifestations, causative factors, and treatment becomes imperative to streamline management and reduce time-consuming diagnostic procedures.

Advanced imaging techniques play a pivotal role in diagnosing and distinguishing PD from related conditions. Magnetic resonance imaging (MRI) and its sophisticated variants, including magnetic resonance spectroscopy imaging, diffusion-weighted imaging, and functional MRI, offers valuable insights into early PD detection by differentiating it from a typical parkinsonian disorder [5]. Functional MRI and diffusion tensor imaging techniques specifically reveal aberrations within the olfactory system during the prodromal phase of PD, aiding in its identification [6]. These imaging modalities enable clinicians to visualize and analyze intricate brain patterns, contributing significantly to earlier and more accurate diagnoses of PD, crucial for timely intervention and management.

In recent studies, innovative methodologies have emerged to address challenges in image analysis and cancer classification. One approach involves a self-supervised machine learning algorithm designed for live cell image segmentation across diverse imaging modalities, using optical flow for pixel self-labelling and generating objective cell/background classifications [7]. Another study introduces a comprehensive methodology combining binary particle swarm optimization with decision trees (BPSO-DT) and convolutional neural networks (CNN) for classifying various cancer types based on RNA sequence data. This method involves

CONTACT S. Kanagaraj ✉ kanagarajsphd@gmail.com 📧 Department of Computer Science and Engineering, Sri Venkateshwara College of Engineering, Visvesvaraya Technological University (VTU), Karnataka, India; M.S. Hema ✉ hema@gmail.com 📧 Department of Computer Science and Engineering, Sri Venkateshwara College of Engineering, Vidya nagar, Bengaluru 562 157, Karnataka, India; M. Nageswara Guptha ✉ nageswaraguptha@yahoo.com 📧 Sri Venkateshwara College of Engineering, Vidya nagar, Bengaluru 562 157, Karnataka, India
This article has been corrected with minor changes. These changes do not impact the academic content of the article.

preprocessing for optimal feature selection, dataset augmentation, and a robust deep CNN architecture for precise cancer classification [8]. Deep learning (DL) models, such as Dense-U-Net, have shown promise in segmenting high-resolution medical images [9]. However, challenges persist in accurate automated vertebra segmentation within CT images due to the diversity in spinal architecture among individuals [10]. In medical imaging, dealing with unbalanced datasets in supervised learning poses a challenge [11]. Estimating six-dimensional (6D) posture in computer vision remains a complex task, hindered by the scarcity of annotated data [12]. In dentistry, automatic dental image analysis for caries identification using DL methods faces obstacles due to the limited availability of labelled clinical images [13]. Moreover, research in transmission systems has explored learning-based solutions like online coherency-based controlled islanding, offering potential benefits for power grid management [14]. This encompasses various approaches investigating the impact of input data and measurement errors on methods like feed-forward neural networks (FFNN) and support vector machines (SVM) [15]. These advances in imaging techniques, DL models, and innovative approaches in different domains underline ongoing efforts to address challenges in medical imaging analysis, posture estimation, dentistry, and power grid management.

This study emphasizes the superiority of Deep Learning (DL) models over Machine Learning (ML) in feature extraction through convolution and pooling procedures. DL models, being nonlinear networks, offer scalability and enhanced performance with increased training data, although their stochastic training approach can lead to sensitivity issues to training information and variations in predictions based on weights discovered during training. To address this, we've adopted an ensemble approach, employing a U-Net segmentation model followed by DenseNet architecture to sort the segmentation outputs from raw DaTscan images for Parkinson's disease diagnosis, departing from simpler DL algorithms used in prior PD detection studies. The system workflow is visually depicted in Figure 1, illustrating the sequential application of the U-Net and DenseNet models for efficient disease diagnosis.

Figure 1 presents the proposed system's flowchart, delineating the sequential steps involved in the utilized approach for Parkinson's disease diagnosis from DaTscan images. Our objective in this research is to enhance disease diagnosis by employing a novel Deep Learning-based methodology integrating affine transformation, U-Net segmentation, and DenseNet classification. Alongside diagnostic advancements, our aim includes developing a user-friendly Graphical User Interface (GUI)-based software tool. This tool aims to facilitate rapid and accurate disease detection, potentially aiding in early diagnosis and improving patient outcomes. The integration of cutting-edge DL techniques with a user-friendly interface is poised to enhance diagnostic precision and accessibility for Parkinson's disease.

The following is the framework for this study: In Section 2, we'll talk about a few scholarly works that might help with PD classification. The experimental design, including the models and ensemble method used, is described in Section 3. We compare the suggested model's performance to that of existing approaches and provide our findings in Section 4. Here, we take a look at what the suggested process would yield: a sample application. The paper is concluded and summarized in Section 6.

II. Literature review

The literature reviewed herein encompasses diverse applications of machine learning techniques across various domains, highlighting their efficacy in data-driven enhancements. the study [17] addresses spectral identification improvement through robust Pareto analysis, consistently outperforming existing methods [16]. Following this, semi-supervised machine learning is utilized in the ANN-SoLo tool to bolster the identification of post-translationally changed peptides, significantly enhancing spectral IDs [17].

A novel strategy is used for supervised voice augmentation using deep auto encoder and neural network models, surpassing traditional approaches in non-stationary conditions and achieving superior outcomes in coding feature estimation [18]. Moreover, a self-supervised learning approach in a neural network is employed to denoised microscope-integrated 4D-OCT images, enhancing image quality and preserving anatomical details in real-time [19].

A distinctive encoding method using CGR representations and supervised autoencoders is introduced for peptide/protein sequences, significantly impacting drug sensitivity studies, notably outperforming existing approaches in HIV protease mutants and hemolytic/non-hemolytic peptide analysis [20].

This research delves into diverse applications of machine learning in medical imaging and data augmentation methodologies. Primarily, a representation learning strategy is demonstrated to autonomously enhance data without requiring prior information, thereby refining labels for better applicability to new cases in various machine learning settings [21]. Subsequently, a notable contribution lies in a proposed architecture for image segmentation tasks using U-Net networks, showcasing exceptional accuracy in segmenting CT thorax and 3D MRI brain scans without preprocessing, surpassing human specialists in accuracy
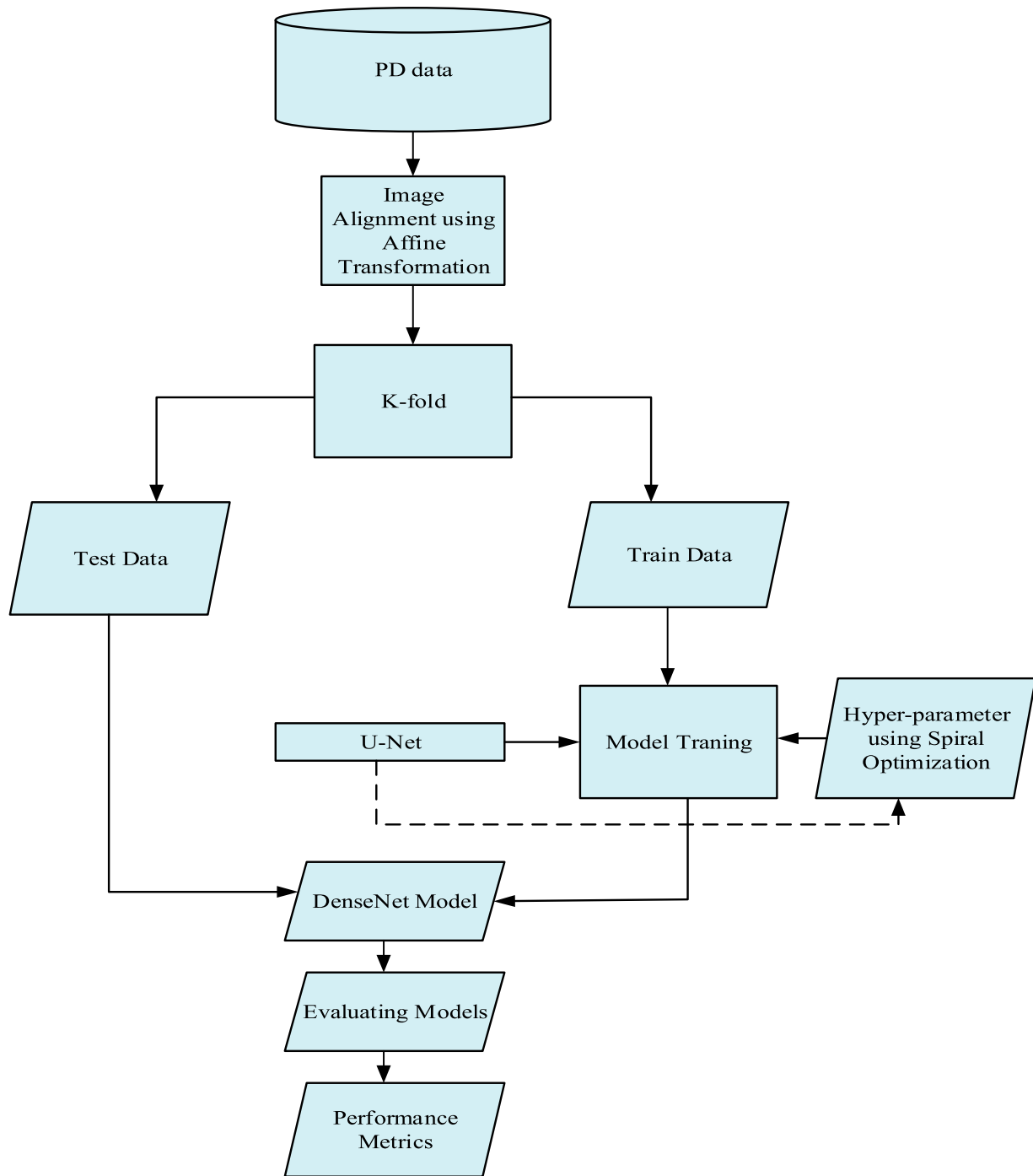
Figure 1. Proposed system flowchart.

[22]. This architecture, available under an open-source license, exhibits potential applicability in improving segmentation outcomes across different tasks.

Furthermore, the study introduces a novel deep learning approach, PaDBNs, for automating CT vertebra segmentation. Leveraging patch-based techniques, this model efficiently performs feature selection, class differentiation, and fine-tuning, markedly reducing processing costs while significantly improving segmentation results compared to established methods [23]. Additionally, an innovative Implicit Semantic Data

Augmentation (ISDA) method is presented, diversifying augmentation samples through deep characteristic translations along semantically relevant orientations, especially beneficial for underrepresented classes. Utilizing meta-learning to automate altered semantic directions, the ISDA approach proves successful across various image datasets [24].

In the field of machine learning, focusing on diverse strategies that advance unsupervised image representation learning, semi-supervised training, and enhancing model performance. One such advancement is the

Bootstrap Your Own Latent (BYOL) approach, introducing a cooperative learning paradigm between online and offline networks to enhance image representations without negative pairings. BYOL demonstrates superior performance on transfer and semi-supervised models, achieving state-of-the-art accuracy on ImageNet classification tasks [25].

In a study, the Wild6D dataset takes centre stage, addressing category-level 6D object position estimation challenges. The Render for Pose network (RePoNet) emerges as a novel model, leveraging silhouette-matching objectives on both synthetic and real-world data. This model attains state-of-the-art performance without relying on 3D annotations, signifying its efficacy in object pose estimation scenarios [26].

Another significant contribution, MixMatch, merges MixUp and generative model-based label uncertainty propagation approaches to improve semi-supervised learning outcomes. MixMatch showcases substantial accuracy improvements on various datasets with limited labelled data, achieving significant error rate reductions on STL-10 and CIFAR-10 datasets. Moreover, MixMatch demonstrates its utility in enhancing accuracy-privacy trade-offs in differential privacy scenarios [27].

In the dental radiography a self-training-based approach is proposed caries for identification and segmentation. By training a student model on a large collection of unlabelled images and leveraging centre-cropped decayed area photos with enhancement techniques, this method outperforms traditional self-supervised learning, exhibiting notable gains in mean intersection over union and average pixel accuracy [28].

In the realm of power distribution networks, a rapid load-shedding methodology is introduced, streamlining prediction and optimization following network disturbances. Unlike previous optimization-based techniques, this method employs a classification model to quickly forecast splitting schemes, simplifying load-shedding optimization problems and demonstrating its effectiveness through simulations on a 16-machine, 68-bus system [29].

Another facet of interest involves estimating distribution network conditions, particularly challenging in scenarios lacking regular monitoring. Addressing this, studies have explored state estimation techniques to manage the complexity of distribution networks, considering uncertainty factors in load profiles for improved accuracy [30]. Moreover, the research investigates the potential of deep learning models in analyzing and forecasting biomedical signals. This exploration holds promise for medical diagnosis and therapy by utilizing deep learning's capabilities to analyze biomedical data [30–34].

This study is focused on early (PD) detection by utilizing non-motor manifestations, which significantly impact patients' lives. It employs nine machine learning algorithms to differentiate PD patients from controls, using diverse non-motor clinical features from the PPMI datasets. This research highlights the promise of non-motor variables for accurate PD screening and the role of interpretable rules in this context [35].

III. Proposed model

A. Dataset description

Progression Markers Initiative (PPMI) dataset consists of 449 early-stage PD individuals and 210 healthy controls (HCs). Baseline scans and subsequent scans at 1, 2, 3, and 5 years are available for PD individuals. HCs mostly have one scan each. It available at https://paperswithcode.com/dataset/ppmi.

- *Image Characteristics*: Images have a resolution of $2\,mm^3$ per voxel and measure $109 \times 91 \times 91$ voxels in size. They are mapped to the Montreal Neurological Institute's (MNI) atlas by PPMI.
- *Data Cleaning:* Misregistered photos were discarded, resulting in 208 HCs and 365 PD individuals for analysis.
- *Participant Information:* Mean age of participants: 60.6 ± 11.2 years. The male/female ratio was not specified.
- *Imaging Details:* The DaTscan images were rotated to ensure the more affected hemisphere was on the right due to PD's asymmetric effects.
- *Normalization and Classification:* Occipital lobe served as the normalization area (N), and the striatum was used for classification (C) in normalized classification.
- *Mask Extraction:* Otsu's thresholding was applied to the mean HC image to remove background, then again to eliminate nonspecific binding voxels, resulting in the extraction of the striatum mask.
- *Mask Range:* The range of masks used was from slice 29 to slice 55 for analysis.

B. Data pre-processing

(a) Image alignment using affine transformation

The shape of the image's contents may be altered through transformations. Any alteration of an image's geometry will alter the spatial connections between the image's pixels. A pixel in the input picture located at coordinates (x,y) is mathematically transformed to a new location (x', y') in the resulting image. The spatial connection between pixels may be altered in a linear (Affine transform) or irregular (Projective transform) fashion.

$$G(x, y) \rightarrow H(x', y') \qquad (1)$$

The new coordinates for every pixel in the input picture are determined by a mapping function that results from

an affine transformation, which is a linear transformation. It is possible to provide the mapping function as a pair of functions, such as,

$$(x', y') = M(x, y)$$
$$x' = M_x(x, y) \tag{2}$$
$$y' = M_y(x, y) \tag{3}$$

It may be written in polynomial form as,

$$x' = a_0 x + a_1 y + a_2$$
$$y' = b_0 + b_1 y + b_2 \tag{4}$$

In matrix form,

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} a_0 & a_1 & a_2 \\ b_0 & b_1 & b_2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{5}$$

Matrix notation is often used in image processing. Using the Affine Transform, you can keep the same distance ratio between points and maintain the image's parallelism. That example, the transformation may turn squares or rectangles into a parallelogram, but it wouldn't produce a trapezium.

Commonly used Affine transformations include the following:

- Translation
- Rotation
- Scaling

A linear translation of pixels in the X and Y axes is the result of an affine transformation. Here is the translation matrix,

$$\begin{bmatrix} 1 & 0 & \Delta x \\ 0 & 1 & \Delta y \\ 0 & 0 & 1 \end{bmatrix} \tag{6}$$

If you apply a rotation matrix to a picture, you may flip it by a certain amount. The matrix for rotation is written as,

$$\begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{7}$$

The picture is resized when the scale is applied. An example of a scaling matrix is (Figure 2),

$$\begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{8}$$

C. Classification

(a) U-Net for segmentation

The U-net architecture works effectively for a range of biomedical segmentation applications. Figure 3 shows the general architecture of the network. One path widens as it moves to the right, and another narrows as it moves to the left. Similar in design to all other convolutional neural networks is the decreasing path.

The total amount of feature channels is increased by the factor of two for each downsampling step. The process of ups the feature maps, performing 22 convolutions ("up-convolution") to divide the feature channels, combining with the resulting cropped include diagram from the path of contraction, and performing two 33 convolutions, everyone which is tracked by a ReLU operation, are all examples of operations on the expansive path. Since convolution eliminates pixels at the image's edges, cropping is always necessary. The last layer uses a 1×1 convolutional operation to translate each feature vector with 64 components to a class. Twenty-three convolutional layers make up the network as a whole. Selecting an input tile number that guarantees all 22 greatest pooling operations are performed to an element having equal x and y sizes results in an even tiling of the segment map's output. (see Figure 3). The total amount of feature channels is increased by a factor of two for each down sampling step.

Using segmentation maps and input images, Caffe's stochastic gradient descent (SGD) implementation trains the network. The finished picture is proportionately smaller than the original due to the unpadded convolutions. In order to minimize overhead and make the greatest use of the GPU RAM, we select a large input tile over a high batch size and compress the entire batch into a single image. Therefore, we choose a high velocity (0.99) to ensure that many previously observed training samples are used to inform an update in the present-day optimization phase.

To get the energy function, we use the function of cross-entropy loss and perform a soft-max over the final characterization map, pixel by pixel. In this article, we define the softmax as

$$p_k(x) = \exp(a_k(x)) / \left(\sum_{k'}^{K} \exp(a_{k'}(x)) \right)$$

where Z2 is the intensity of the activations and $a\_k(x)$ is the amount of data being processed by channel k of the set of feature channels at position x in the image. If K is the overall amount of categories, then $p\_k(x)$ is the approximate maximum function. In other words, for the k with the greatest activation $a\_k(x)$, $p\_k(x)1$ is calculated, whereas for the remaining k values, p_k (x)0 is calculated. The distance of $p\_(l(x))(x)$ from 1 is then penalized at each position by the cross entropy.

$$E = \sum_{x \in \Omega} w(x) \log(p_{l(x)}(x)) \tag{9}$$

Where, l:1, … , K is the actual label for all pixels and w: R is the amount maps that we implemented to prioritize certain pixels during training.
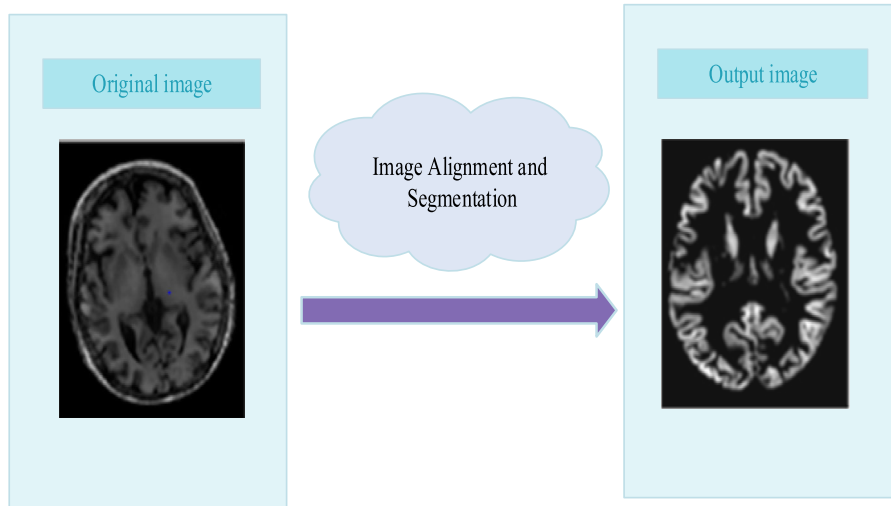
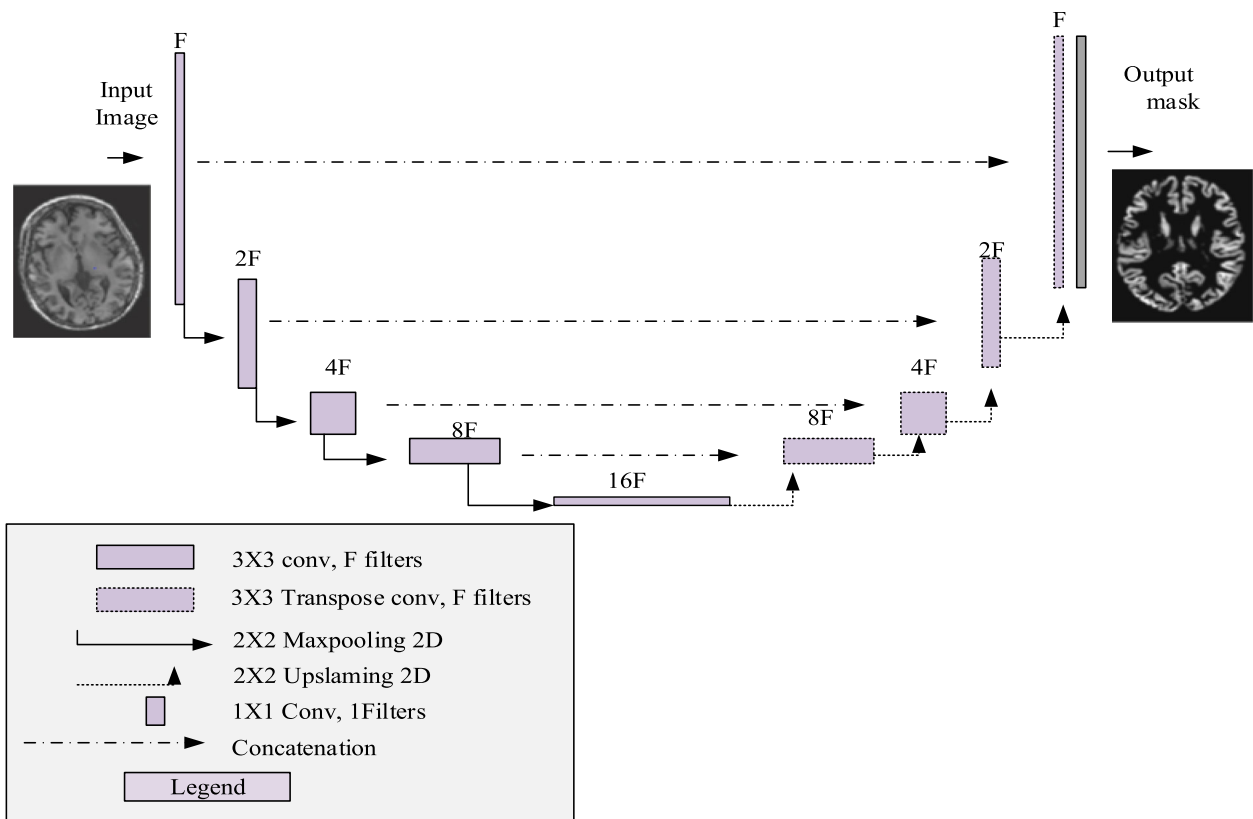**Figure 2.** Image alignment using affine transformation.

Figure 3. U-Net system for segmentation process.

As seen in Figure 3, we pre-compute the weight map for each ground truth segment in order to account for the varying frequency of the outcomes of a particular group in the first training set. This forces the system to determine the modest separation boundaries that we create between touching cells.

The dividing line is determined by a series of morphological calculations. After that, we get the weight distribution by

$$w(x) = w_c(x) + w_0 \cdot \exp\left(-\frac{(d_1(x) + d_2(x))^2}{2\sigma^2}\right) \tag{10}$$

where w_c: R is the normalization weight map, d_1: R is the distance that runs between the cell centre to the first border, and d_2:R is the length from the cell centre to the second border. We tested with $w\_0 = 10$ and a focal distance of 5 pixels.

An effective weight initialization is crucial for the success of deep neural networks, which have several convolutional layers and multiple potential routing options. If not, certain network nodes might be always active while others would never accomplish anything. If the beginning weights are adjusted correctly, each feature map in the network should have a variance of around one. We may accomplish this by taking a sample

of the initial weights for a network with our architecture (alternating convolution & ReLU layers) from a Gaussian distribution with an average deviation of (2/N), where N is the total number of input nodes for a whole neuron. For instance, if the layer before it was a 33 convolution with 64 feature channels $N = 9 \cdot 64 = 576$.

(b) Densenet architecture for PD classification

Recent years have seen a surge of interest in the study of dense convolutional networks (DenseNet), which have found useful applications in the processing of medical pictures. In this article, we will briefly discuss the following elements of DenseNet.

the dense net architecture. (a) DenseNet's foundational layers include a transition layer, a convolutional layer, a fully connected layer, and a dense block. a transition layer, the convolutional layers, a layer that is completely connected, and a dense block make up DenseNet's basic layers. (b) To provide faster network training and better generalization performance, dense blocks are composed of densely linked units that use nonlinear mapping functions like BN, ReLU, and Conv. These units are constructed using a preactivation technique. To prevent the algorithm from overfitting the input data, we may minimize the number of dense blocks inputs and the size of the map of features. By integrating category parameters into network features and performing a weighted categorization on the feature data, the fully connected layer, often referred to as the categorization prediction layer, reduces the impact of feature placement on classification.

DenseNet is distinguished by its feature-sharing and freely interconnected layers. DenseNet's strengths lie in its ability to effectively reduce the difficult-to-optimize gradient disappearance problem in deep networks, the provision of compact as well as distinguished input attributes by shortcut connections of varying lengths, and the reuse of feature maps from multiple layers. The conclusion is that using features from every layer will get the best results in terms of performance and model resilience on a benchmark dataset while requiring less computing effort and a smaller model. DenseNet's feature maps for each layer are combined with those of preceding layers, as well as the data is repeated many times over. The exponential development in compute and memory cost during training occurs because the number of parameters for the model grows linearly with the number of networks input layers.

To begin, the vanishing-gradient issue is mitigated via DenseNet. The second is that they minimize the number of input parameter and boost feature propagation and feature regeneration. Each layer's outputs are sent into the next dense layer and added together along the depth axis. DenseNet is a technique that combines Dense Blocks and transition Layers to classify

a given input data (see Figure 4). Picture of the Day. DenseNet takes an input picture and processes it via numerous layers of filters, each of which has the same feature maps across layers but a different number of filters. The transition layer follows the dense block. The transition layer's duties include both convolution and pooling. The down-sampling procedures outside a thick block are executed by the transition layer. The feature maps in the dense block must all have the same size before feature concatenation can be performed. A bottleneck convolution layer may be included before the convolutions to minimize the amount of feature maps and boost performance. DenseNet's transition layers include the batch normalization (BN) input layers, the convolution input layers, and the average pooling input layers.

Figure 4 depicts a comprehensive operational model for the dense block's operations. DenseNet architecture's dense block consists of a batch normalization (BN) input layer, a rectified linear unit (ReLU) stimulation, and convolution (convs). After the last dense block, a pooling input layer with global averages is used to provide input for a Softmax classifier. Since DenseNet contains L levels, these connections will be made directly between L layers: L(L + 1)/2. See also: Figure Inputs to layer x_l for non-linear H_l transformations are as follows:

$$x_l = H_l([x_0, x_1, \ldots, x_{l-1}]) \tag{11}$$

where $[x_0, x_1, \ldots, x_{l-1}]$ represents to combining input feature maps from layer zero to layer l-1. Layers' input features maps are utilized as inputs by the layer above it, and layers' input feature maps are used as inputs by the layer below it. Consequently, as shown in Figure 5, the DenseNet displays the following results for level l, and the collections of levels on top of depths dimension H:

$$x[l] = f\left(w^* H\left(x[l-1], x[l-2], \right.\right.$$
$$\left.\left. x[l-3], \ldots, x[1]\right)\right) \tag{12}$$

At runtime, an ideal tensor is constructed by concatenating all of H's inputs from Equation (2). By use of convolution and pooling, feature-map sizes may be adjusted in the DenseNet architecture.

Network inputs may be standardized using the batch normalization method, which can be applied to either the reactions of a preceding layer or the data inputs themselves. Rectified linear action is used as the default activation for building transfer and learning convolutional neural networks. DenseNet splits up into Density blocks, each with its filters but identical overall dimensionality. The Transition Layer implements batch normalization through down-sampling. One step of average pooling is averaging across the various feature map segments. This indicates that the average value of each
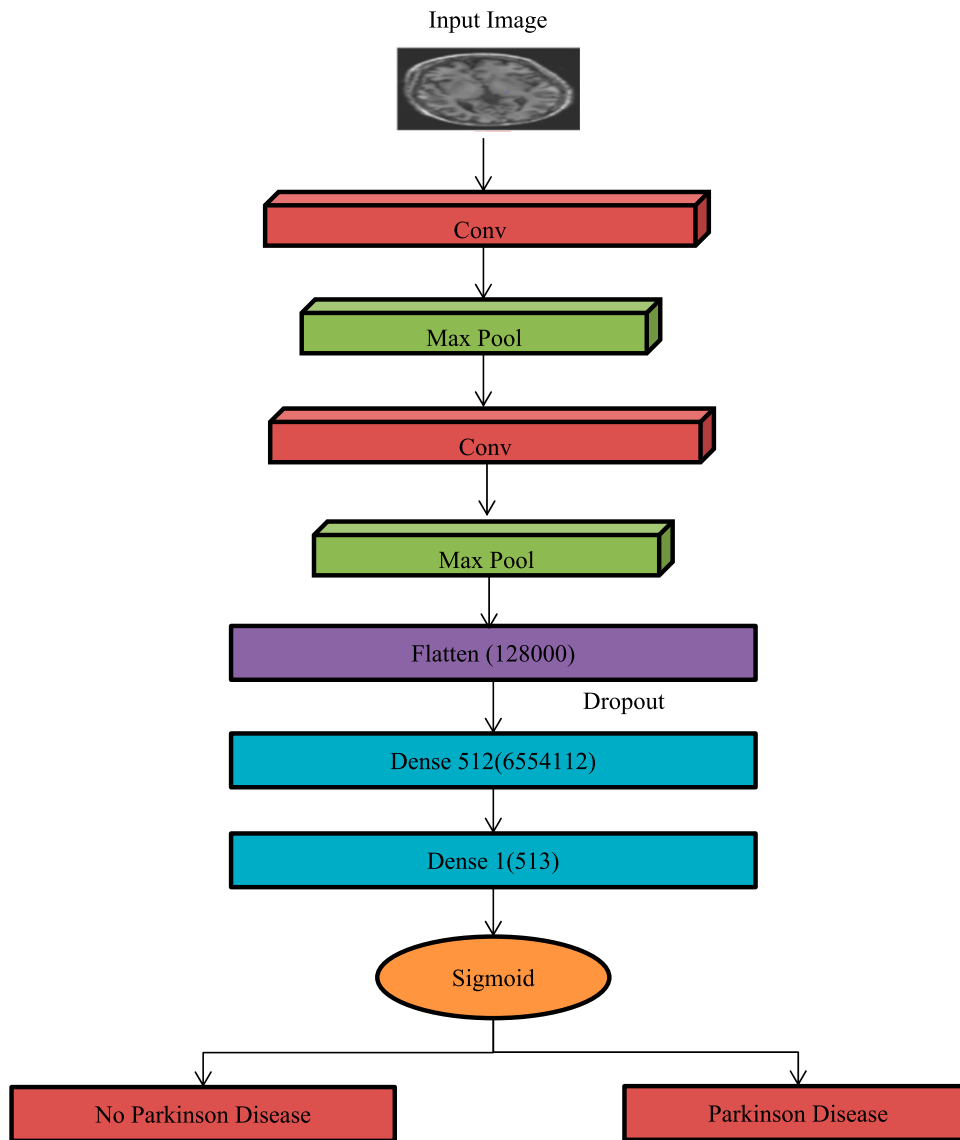
Input Image

Figure 4. DenseNet architecture.

square in the function map is used to down-sample the squares.

The training of the model took 250 iterations. Additionally, the Cross-Entropy measure of loss was utilized, a batch size of 8 was utilized, a 1e3 learning rate was employed, and an Adamax optimization was used to optimize the weights. In all, there are 6,955,906 parameters in the deep learning model that was built. The proposed model was trained on the Windows computers with the 64-core, 128-thread AMD Threadripper CPU and 256MB of L3 cache. RAM was about 128 GB, and the graphics card utilised was an RTX3080.

To get the best accuracy of our suggested framework, we have examined several optimizers based on deep instruction and the DenseNet framework. The Spiral Optimisation Algorithm as shown in Figures 6 and 7 is one example of a divergence-creating optimizer. Each parameter's learning rate is calculated independently by the SOA. The first two minutes of SOA are used to derive the gradient. SOA is superior to other optimizers because few models contain embeddings.

$$\text{Adam} : m_n = E[A^n] \tag{13}$$

Where m is the current value of the variable under consideration, A is any other variable, and E is the anticipated value of n variables.

$$\text{Adamax} : v_t = \beta_2 v_t + (1 - \beta_2)|g_t|^2 \tag{14}$$

$$u_t = \beta_2^\infty v_{t-1} + (1 - \beta_2^\infty)|g_t|^\infty \tag{15}$$

$$u_t = \max(\beta_2 \cdot v_{t-1}, |g_t|) \tag{16}$$

If v_t is an updated norm and u_t is a survey of the scope of that norm

D. Experimental setup

The dataset was divided into three distinct subsets: training, validation, and testing. Specifically,
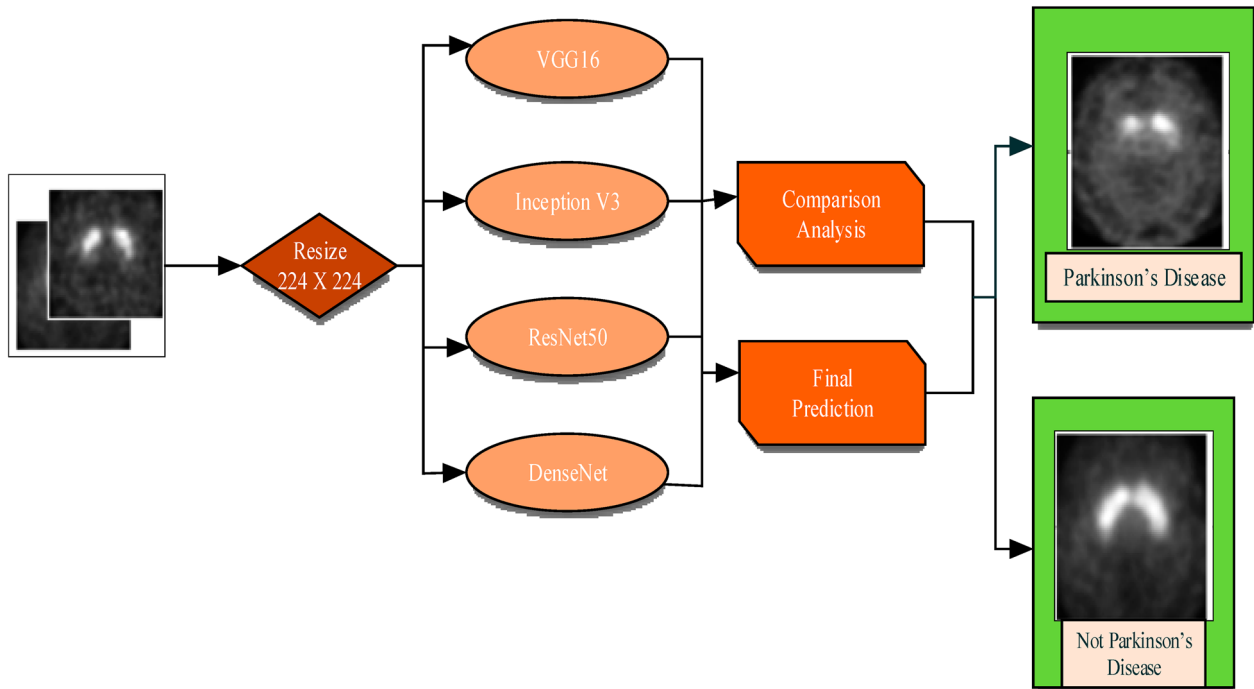
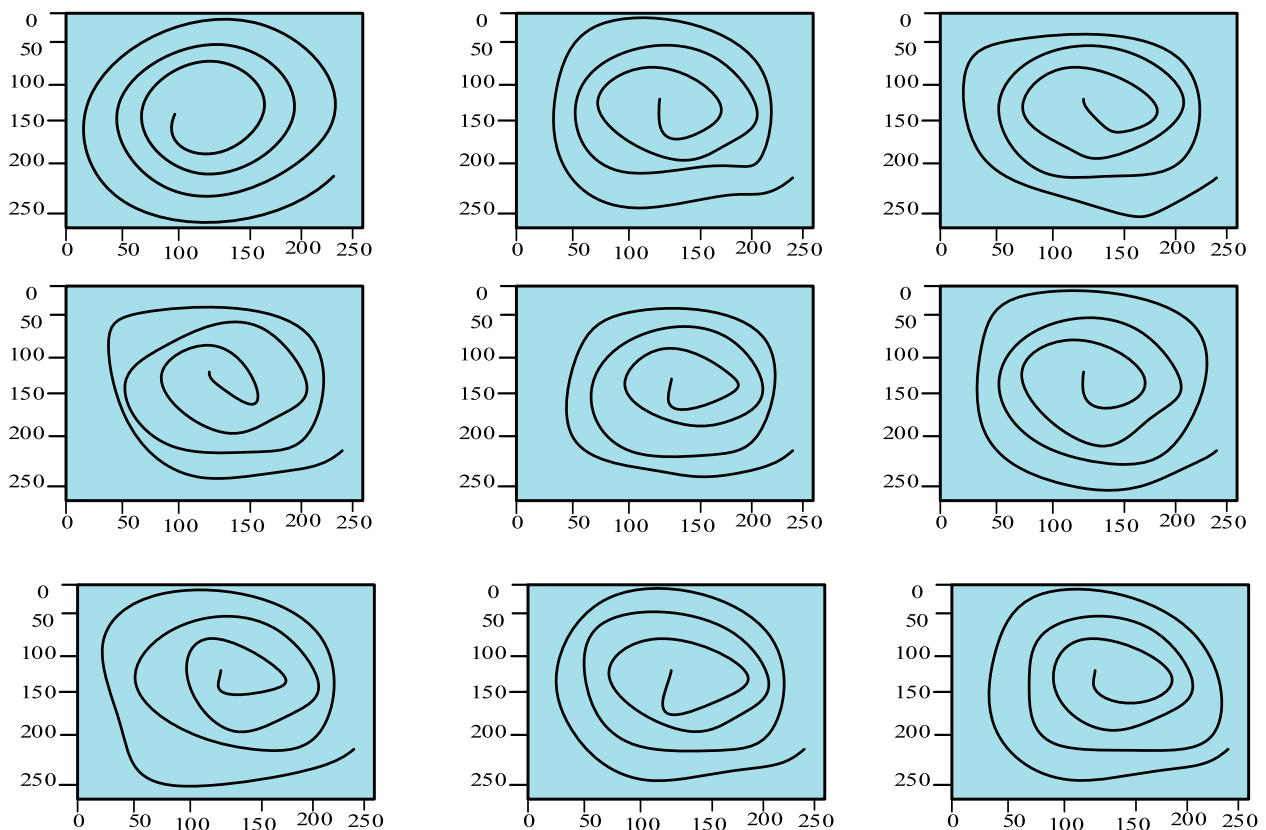**Figure 5.** Comparison of DL architectures.

Figure 6. Spiral optimization algorithm tuning.

approximately 70% of both PD individuals and healthy controls were allocated to the training set for model training purposes. The remaining 15–25% of the data was distributed between the testing set, utilized to objectively evaluate the final model's performance, and the validation set, employed for hyperparameter tuning and model selection. This allocation strategy aimed for a deliberate 70% split for training, ensuring a significant portion of the dataset was dedicated to training the model effectively.
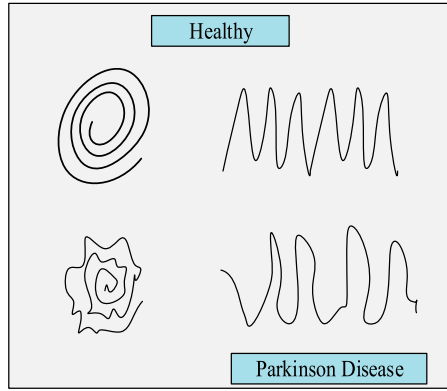
Figure 7. Healthy vs. PD disease prediction.

E. Evaluation metrics

In this section, demonstrates the equations and the metrics utilized to obtain their scores for evaluation of the gathered findings,

Accuracy is a measure of overall effectiveness in classification tasks.

$$A_c = \frac{TP + TN}{TP + TN + FP + FN} \qquad (17)$$

Precision is a metric that focuses on the accuracy of positive predictions.

$$P_r = \frac{TP}{TP + FP} \qquad (18)$$

The sensitivity of a model also called the True Positive Rate or Recall, measures how well it can recognize every positive case.

$$S_n = \frac{TP}{TP + FN} \qquad (19)$$

Specificity evaluates how well the model can recognize every negative case.

$$S_p = \frac{TN}{FP + TN} \qquad (20)$$

The harmonic mean of sensitivity and precision is known as the F1-Score.

$$F_m = \frac{2 \times P_r \times S_n}{P_r + S_n} \qquad (21)$$

Where TP, TN, FP, and FN have the meanings given below:

- When both the actual lesson of the data piece and the expected class are True (1), we have a true positive.
- A True Negative (TN) occurs when both the real and forecasted classes of an information points are False (0).
- When a data item's expected class is True (1) but its actual class is False (1), this is known as a false positive (FP).
- When a data point's projected class is False (0) but its actual class is True (1), this is known as a false negative (FN).

Table 1. Predictions of PD using DL models have been shown.

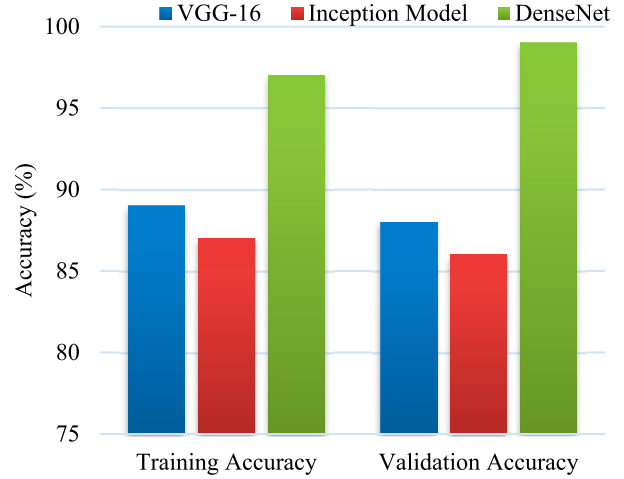| Prototypical	Accuracy	Precision	Sensitivity	Specificity	F1-value
DenseNet	99.34%	99.51%	99.51%	99.02%	99.51%
VGG16	93.02%	95.29%	94.19%	90.69%	94.74%
Inception-V3	94.03%	93.32%	98.68%	84.73%	95.82%
Exception	96.35%	95.45%	99.85%	89.38%	97.58%

Figure 8. Performance of accuracy.

IV. Experimental results

Here, we provide our thoughts on the performance of the four different base learners – DenseNet, VGG-16, Inception-V3, and Xception – and the outcomes they produced. The Grad-Cam explainability of the foundational learners and the results obtained utilizing the suggested approach is reported further in this section. Table 1 represents the comparative analysis of existing and proposed system.

VGG16, ResNet50, Inception-V3, & Xception were all tested on the test set after they had been trained using the training data. Table 1 displays the outcomes achieved by the four foundational students. According to Table 1 and the Figure 8, VGG16 & Xception models achieved 95.34% accuracy, which was the highest of any of the base learners. The investigational analysis shown in Figures 9 and 10 also demonstrates that the Inception-V3 and Xception models are capable of making accurate predictions about PD patients, with the lowest probability of misclassification. DenseNet, on the other hand, is superior at predicting which patients would not get PD. The proposed method performance shown in Figures 11 and 12.

The results of this approach will not only aid doctors in making accurate diagnoses but will also allow them to intervene before their patients' conditions worsen. Our technique is successful on the PPMI data set, with an accuracy of 98.45%, as indicated in the Figure 13. A wrong diagnosis may have serious physical, emotional, and psychological consequences for the patient and their loved ones, making it extremely important to
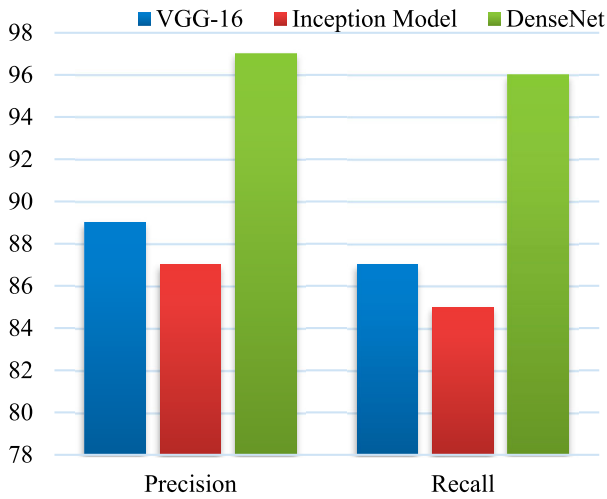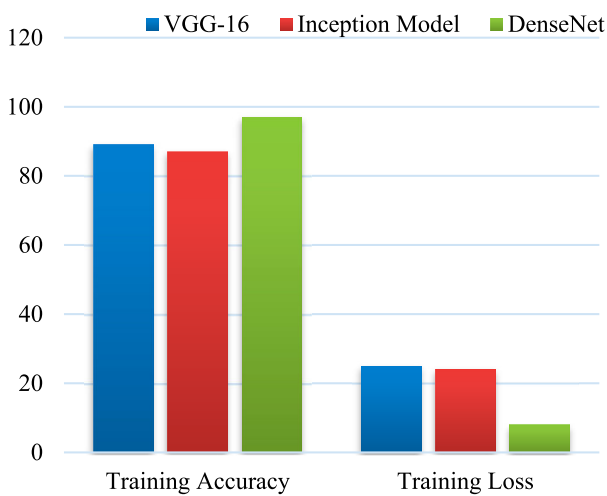
Figure 9. Performance of precision and recall.
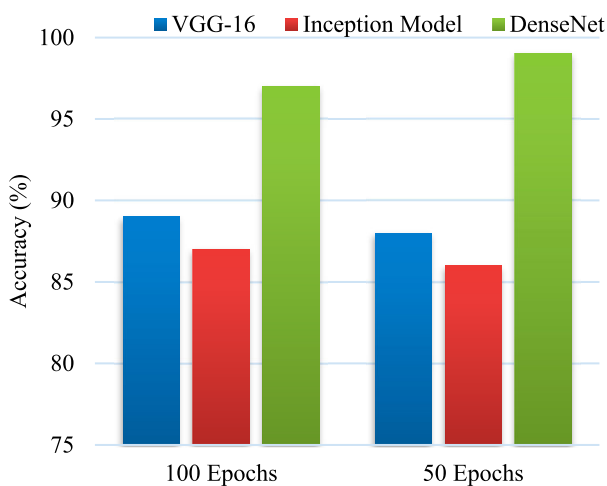

Figure 10. Performance of training accuracy and loss.

Figure 11. Performance of accuracy.

Figure 12. Performance of precision and recall.
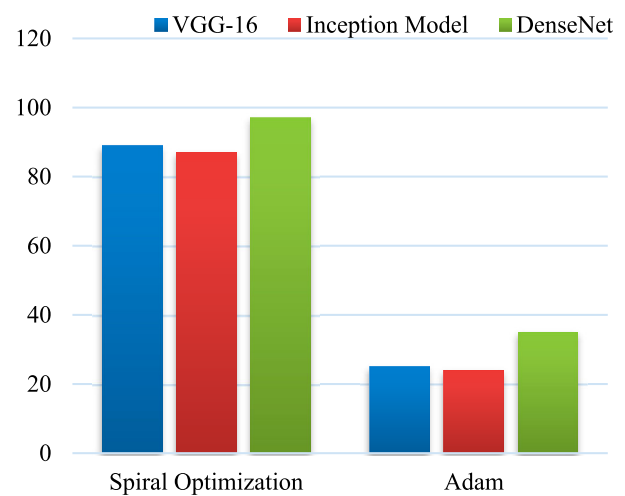
Figure 13. Performance of optimization algorithm.

A. Software application

The developed application, detailed at https://gitlab.com/digiratory/biomedimaging/parkinson-detector. It is a versatile tool intended to assist medical professionals in their rapid preliminary diagnosis tasks using MRI images. It is a Python-based application compatible with both windows and linux operating systems. The users-friendly interface is built utilizing the Qt library and supports the seamless handling of various image formats, including Dicom files from MRI machines and common image formats like jpg and png. Notably, the application offers drag-and-drop functionality, enhancing user convenience.

This application employs a variety of ensemble techniques, such as the Product Rule, Majority Voting, Sum Rule, and the novel FRLF method. These ensemble techniques leverage the capabilities of various neural networks. The minimal system requirements of the application call for an operating system without GPU capability, including Windows 7 or later, Ubuntu 16.04 or later, or Mac OS 10.12.6 (Sierra) or later (64-bit). For

minimize misclassifications in the medical sector analysis domain. When comparing the deep learning method to the absolute structure of the U-net (segmentation) & DenseNet (PD classification), we found that the deep learning method considerably decreased the amount of false positives as well as false negatives.
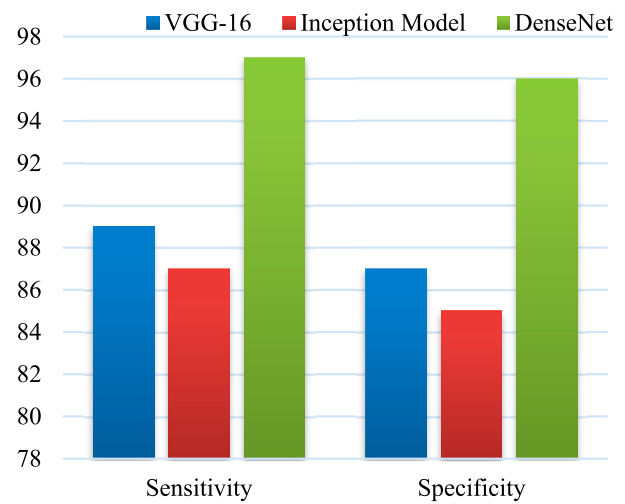
the initial neural network model downloads, users must have Python 3.6 or later, at least 4GB of free disk space, an Intel Core i3 processor, and a broadband internet connection. Admin privileges are not a prerequisite for running this application.

B. Discussion

The central focus of this discussion revolves around the profound impact of image segmentation on the predictive results achieved in this study. By incorporating an image segmentation approach, specifically the U-Net model, the study significantly enhances the accuracy and effectiveness of Parkinson's disease prediction utilizing MRI images from the PPMI datasets. The segmentation process facilitates the isolation and extraction of pertinent features from the raw MRI images, ensuring that the deep learning models work with more refined and relevant data.

The results clearly demonstrate that the utilization of image segmentation has transformative effects on the model's efficiency. The incorporation of segmentation notably elevates precision, sensitivity, and specificity, showing substantial improvements. Segmentation enhances precision to 99.51%, sensitivity to 99.51%, and specificity to 99.02%. This technique refines input data, notably boosting predictive accuracy to 99.34%. The results underscore segmentation's pivotal role in advancing model efficiency, leading to robust early-stage disease detection. It allows for a more precise analysis of the regions of interest, such as the striatum, crucial for Parkinson's disease diagnosis. This segmentation step refines the input data, improving the quality of information fed into the deep learning models. As a result, the predictive accuracy, precision, sensitivity, and specificity are significantly enhanced, ultimately providing a robust tool for early disease detection. The integration of U-Net-based segmentation into the proposed system showcases the pivotal role of image preprocessing in medical image analysis. This discussion underscores how the careful handling of medical images through segmentation can yield more accurate and reliable diagnostic outcomes. Furthermore, it highlights the potential for image segmentation techniques to be applied in various medical imaging contexts, contributing to improved disease detection and patient care.

V. Conclusion

In this paper, we offer a set of DL models that can efficiently predict Parkinson's disease using PPMI DaTscan pictures. To improve the overall outcomes of the model, We built a deep learning model that uses the trustworthiness assessments of VGG16, DenseNet, Inception-V3, and Xception. Based on these findings,

it is secure to say that the suggested model outperforms the other current method alternatives. The suggested model achieves an impressive 98.45% accuracy in recognition, 98.84% precision, 98.84% sensitivity, 97.67% specificity, and 98.84% F1 score. Our approach has also been implemented in a publicly available, software application with a graphical user interface (GUI) for the fast parkinson's disease diagnosis by DaTscan scans. Potentially helpful for Parkinson's disease screening and early detection. We rely heavily on DaTscan images as our starting point. While these results are promising, we recognize the need for further research to address limitations and enhance the models' capabilities. Our future work will focus on refining the hybrid deep learning architecture and expanding the scope of diagnostic images to include MRI and CT scans, thereby contributing to more comprehensive and accurate Parkinson's disease diagnosis and early detection.

Nomenclature

Acronym	Description
DL	Deep Learning
DDL	DaTscan and Deep Learning
ATIA	Affine Transformation for Image Alignment
PPMI	Parkinson's Progression Markers Initiative
MRI	Magnetic Resonance Imaging
SSL	Self-supervised learning
FSL	Few-Shot Learning
NCF	Noise Classification And Fusion Strategy
OCT	Optical Coherence Tomography
CGR	Chaos Game Representation
CT	Computed Tomography images
6D	Six-dimensional
DAE	Deep Autoencoder
CNN	Convolutional Neural Networks
SAE	Supervised Autoencoders
PaDBNs	Patch-Based Deep Neural Networks
ISDA	Implicit Semantic Data Augmentation
BYOL	Bootstrap Your Own Latent
RePoNet	Render for estimate of posture network
HCs	Healthy Controls
MNI	Montreal Neurological Institute
C	Classification Region
N	Normalization Area
SGD	Stochastic Gradient Descent
DenseNet	Dense Convolutional Networks
BN	Batch Normalization
ReLU	Rectified Linear Unit
conv	Convolution

Acknowledgement

Authorship contributions

All authors are contributed equally to this work.

Data availability statement

Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Ethics approval and consent to participate

No participation of humans takes place in this implementation process.

Human and animal rights

No violation of Human and Animal Rights is involved.

Funding

No funding is involved in this work.

References

[1] Vyas T, Yadav R, Solanki C, et al. Deep learning-based scheme to diagnose Parkinson's disease. Expert Syst. 2022;39(3):e12739. doi:10.1111/exsy.12739

[2] Noor MBT, Zenia NZ, Kaiser MS, et al. 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 Inform. 2020;7:1–21.

[3] Gupta R, Kumari S, Senapati A, et al. New era of artificial intelligence and machine learning-based detection, diagnosis, and therapeutics in Parkinson's disease. Ageing Res Rev. 2023;102013.

[4] Aggarwal N, Saini BS, Gupta S. Role of artificial intelligence techniques and neuroimaging modalities in detection of Parkinson's disease: a systematic review. Cognit Comput. 2023: 1–38.

[5] Zhang J. Mining imaging and clinical data with machine learning approaches for the diagnosis and early detection of Parkinson's disease. NPJ Parkinson's Dis. 2022;8(1):13. doi:10.1038/s41531-021-00266-8

[6] Marino S, Ciurleo R, Di Lorenzo G, et al. Magnetic resonance imaging markers for early diagnosis of Parkinson's disease. Neural Regen Res. 2012;7(8):611.

[7] Robitaille MC, Byers JM, Christodoulides JA, et al. A self-supervised machine learning approach for objective live cell segmentation and analysis. bioRxiv. 2021.

[8] Khalifa NM, Taha MH, Ezzat Ali D, et al. Artificial intelligence technique for gene expression by tumor RNA-seq data: a novel optimized deep learning approach. IEEE Access. 2020;8:22874–22883. doi:10.1109/ACCESS.2020.2970210

[9] Cardone D, Perpetuini D, Filippini C, et al. Driver stress state evaluation by means of thermal imaging: a supervised machine learning approach based on ECG signal. Appl Sci. 2020;10(16):5673.

[10] Chang C, Hung J, Hu Y, et al. Prediction of preoperative blood preparation for orthopedic surgery patients: a supervised learning approach. Appl Sci. 2018.

[11] Zhang B, Wang Y, Hou W, et al. Flexmatch: Boosting semi-supervised learning with curriculum pseudo labeling. Adv Neural Inf Process. 2021;34:18408–18419.

[12] Han D, Jung J, Kwon S. Comparative study on supervised learning models for productivity forecasting of shale reservoirs based on a data-driven approach. Appl Sci. 2020.

[13] Afham M, Dissanayake I, Dissanayake D, et al. Cross-Point: self-supervised cross-modal contrastive learning for 3D point cloud understanding. 2022 IEEE/CVF conference on computer vision and pattern recognition (CVPR); 2022. p. 9892–9902.

[14] Huang L, You S, Zheng M, et al. Learning where to learn in cross-view self-supervised learning. 2022 IEEE/CVF conference on computer vision and pattern recognition (CVPR), 2022. p. 14431–14440.

[15] Sajun AR, Zualkernan IA. Survey on implementations of generative adversarial networks for semi-supervised learning. Appl Sci. 2022;12(3):1718.

[16] Chen Z, Ge J, Zhan H, et al. Pareto self-supervised training for few-shot learning. 2021 IEEE/CVF conference on computer vision and pattern recognition (CVPR); 2021. p. 13658–13667.

[17] Arab I, Fondrie WE, Laukens K, et al. Semi-supervised machine learning for sensitive open modification spectral library searching. bioRxiv. 2022.

[18] Pashaian M, Seyedin S, Ahadi SM. A novel jointly optimized cooperative DAE-DNN approach based on a new multi-target step-wise learning for speech enhancement. IEEE Access. 2023;11:21669–21685. doi:10.1109/ACCESS.2023.3250820

[19] Nienhaus J, Matten P, Britten A, et al. Live 4D-OCT denoising with self-supervised deep learning. Sci Rep. 2023;13(1):5760.

[20] Huang B, Zhang E, Chaudhari R, et al. Sequence-based optimized chaos game representation and deep learning for peptide/protein classification. bioRxiv. 2022.

[21] Yang K, Sun Y, Su J, et al. Adversarial auto-augment with label preservation: a representation learning principle guided approach. ArXiv, abs/2211.00824. 2022.

[22] Kolařík M, Burget R, Uher V, et al. Optimized high resolution 3D dense-U-Net network for brain and spine segmentation. Appl Sci. 2019.

[23] Qadri SF, Ai D, Hu G, et al. Automatic deep feature learning via patch-based deep belief network for vertebrae segmentation in CT images. Appl Sci. 2018.

[24] Li S, Gong K, Liu CH, et al. MetaSAug: meta semantic augmentation for long-tailed visual recognition. 2021 IEEE/CVF conference on computer vision and pattern recognition (CVPR); 2021. p. 5208–5217.

[25] Grill J, Strub F, Altch'e F, et al. Bootstrap your own latent: a new approach to self-supervised learning. ArXiv, abs/2006.07733. 2020.

[26] Fu Y, Wang X. Category-level 6D object pose estimation in the wild: a semi-supervised learning approach and a new dataset. ArXiv, abs/2206.15436. 2022.

[27] Berthelot D, Carlini N, Goodfellow IJ, et al. MixMatch: a holistic approach to semi-supervised learning. ArXiv, abs/1905.02249. 2019.

[28] Qayyum A, Tahir A, Butt MA, et al. Dental caries detection using a semi-supervised learning approach. Sci Rep. 2023;13.

[29] Sadeghi M, Akbari H, Mousavi S, et al. Coherency-based supervised learning approach for determining optimal post-disturbance system separation strategies. IEEE Access. 2023;11:5894–5907. doi:10.1109/ACCESS.2022.3227512

[30] Hong G, Kim Y. Supervised learning approach for state estimation of unmeasured points of distributionnetwork. IEEE Access. 2020;8:113918–113931. doi:10.1109/ACCESS.2020.3003049

[31] Indira DNVSLS, Ganiya RK, Ashok Babu P, et al. Improved artificial neural network with state order dataset estimation for brain cancer cell diagnosis. BioMed Res Int. 2022;2022.

[32] Zhai X, Rajaram A, Ramesh K. Cognitive model for human behavior analysis. J Interconnect Netw. 2022;22(Supp04):2146013. doi:10.1142/S0219265921460130

[33] Poloju N, Rajaram A. Data mining techniques for patients healthcare analysis during covid-19 pandemic conditions. J Environ Prot Ecol. 2022;23(5):2105–2112.

[34] Kalaivani K, Kshirsagarr PR, Sirisha Devi J, et al. Prediction of biomedical signals using deep learning techniques. J Intell Fuzzy Syst. 2023;Preprint:1–14.

[35] Martinez-Eguiluz M, Arbelaitz O, Gurrutxaga I, et al. Diagnostic classification of Parkinson's disease based on non-motor manifestations and machine learning strategies. Neural Comput Appl. 2023;35(8):5603–5617. doi:10.1007/s00521-022-07256-8