Babylonian Journal of Machine Learning
Vol.2025, **pp**. 116–125
DOI: https://doi.org/10.58496/BJML/2025/010; ISSN: 3006–5429
https://mesopotamian.press/journals/index.php/BJML

Research Article

Automated Parkinson's disease Detection from Images Using Deep Transfer Learning and Optimization

Thanaa Alsalem [1, *], Mohammed Amin [2]

[1] *College of Computer Sciences and Information Technology, King Faisal University, Saudi Arabia*.
[2] *Fellowship Researcher, INTI International University, Nilai 71800, Malaysia*.

ARTICLEINFO

Article History
Received 04 May 2025
Revised: 05 Jun 2025
Accepted 02 Jul 2025
Published 23 Jul 2025

Keywords
Parkinson's Disease (PD)
Deep Transfer Learning
Convolutional Neural Networks (CNNs)
Genetic Algorithm Optimization
Automated Diagnosis

ABSTRACT

The diagnosis and treatment of Parkinson's disease (PD) is critical to effectively managing this progressive neurological disorder, which significantly affects motor and non-motor functions. This study presents a deep transfer learning-based algorithm for PD detection. The features are extracted from handwritten image datasets using pre-trained convolutional neural networks such as ResNet50, VGG19, and Inception-V3. To achieve precise classification, a hybrid classification framework that combines a genetic algorithm-optimized k-nearest neighbour (KNN) classifier with a support vector machine (SVM) is implemented. The proposed model offers a reliable, scalable, and efficient solution for diagnosing Parkinson's disease. The experimental results demonstrate the model's state-of-the-art accuracy. AI-driven methodologies are being used in this research to advance automated medical diagnostics, reduce diagnostic delays, and improve patient outcomes.

1. INTRODUCTION

Due to its effect on both motor and non-motor functions, Parkinson's disease (PD) is one of the most challenging conditions in the global healthcare system. To improve patient outcomes, early and accurate detection of PD is essential. Automated diagnostic systems are gaining prominence as a result of artificial intelligence and advances in medical imaging. Using pre-trained deep-learning models, this research aims to provide a reliable, efficient, and scalable method for detecting Parkinson's disease early. As a result of this advancement, diagnostic delays will be shortened, quality of life will be improved, and diagnostics for Parkinson's disease will be revolutionized. In Parkinson's disease (PD), there is insufficient dopamine in the brain, resulting in an incurable neurological disorder. Movement and coordination are controlled by the basal ganglia in the brain, which receives dopamine messages. Dopamine levels decrease when the basal ganglia lack dopamine-producing cells. Parkinson's disease is characterized by several symptoms, including tremors, slow movements (Bradykinesia), impaired balance, poor posture, involuntary movements (dyskinesia), stiff muscles, and changes to speech and language [1], [2]. Due to the lack of clinical tests such as blood tests, Parkinson's disease can be difficult to diagnose. Most people with Parkinson's disease are over 60, but the disease can begin early and go undiagnosed for a long time. In the early stages of the disease, symptoms can be managed more easily, and deterioration can be delayed [3]. At the beginning of the disease, Parkinson's disease can cause finger tremors and halts in speech and movement. As a result of finger tremors, Parkinson's patients tend to have cramped and small handwriting. Micrographia is one of the most important ways to detect Parkinson's early. The presence of micrographic or other deformations in a patient's handwriting can indicate PD. As part of transfer learning, CNNs are retrained to work on new use cases, followed by one or more additional layers [4]. Res Nets, Efficient Nets, and Mobile Nets are examples of deep learning architectures. Several years have passed since deep learning approaches were used in

*Corresponding author. Email: mohammad7ups@yahoo.com

medicine. Several studies have shown that deep learning algorithms are significantly more effective than other high-performance algorithms [5], [6].

Neurodegenerative diseases damage dopaminergic neurons in large parts of the brain early [7]. Dopaminergic dysfunction, as well as non-motor symptoms, causes different kinds of symptoms. It is also possible to experience sweating, pain, slow movement, and difficulty walking as additional motor signs. Aside from the symptoms of genitourinary problems, anxiety, hysteria, fatigue, and sleep disturbances are some of the other non-motor symptoms associated with it [8]. A wide range of ageing-related symptoms affect people's quality of life, both physically and mentally [9]. Early detection, slowing, and stopping Parkinson's disease progression are essential during the early stages of the disease. PD cannot be diagnosed with a blood test or laboratory test despite efforts to develop economical and readily available biomarkers [10]. Patients with Parkinson's disease are diagnosed based on their history and a cognitive evaluation, which includes the mini-mental state examination (MMSE) [11]. With transfer learning, deep CNNs can be trained on a limited set of data [12]. Deep CNNs can be trained using pre-trained CNN models, and their classification layers can be adjusted to fine-tune the weight parameters of the target dataset. The method can enhance efficiency and speed up training despite some shortcomings. The use of data augmentation can also enhance the effectiveness of training [13], [14]. MR imaging methods are showing promise as a means of detecting Parkinson's disease at an early stage, and they are expected to be more sensitive than usual clinical tests. MRI can track brain structural changes to detect significant black iron build-up's [15]. Structural modifications should enable the assessment of disease progression. As a result of the inherent complexity of brain development and slight variations in severity, it is difficult to identify neurodegeneration in the brain visually. At present, clinicians interpret these images based on human interpretation, and there is a possibility of human error. A study [16], found that medical diagnostics of Parkinson's disease had an 80.6% pooled accuracy.

2. RELATED WORK

There are numerous datasets that have been used by researchers to diagnose Parkinson's disease. Other symptoms of Parkinson's disease include loss of olfactory sensation, problems with speech, handwriting, and walking patterns. A new approach to detecting Parkinson's disease based on entropy has been proposed in [17]. Based on the UCI dataset, this improved algorithm was compared to existing approaches to estimate its efficiency. The feasibility of improved algorithms was verified by comparing KNN (k-nearest neighbours), Random Forest, and Nave Bayes algorithms. Our cross-validation scheme consisted of fivefolds [18]. In comparison with traditional algorithms, entropy weights improved the accuracy of the improved KNN algorithm [19]. To categorize the symptoms of Parkinson's disease, an MRI scan was performed. This study sought to improve Parkinson's disease detection by using artificial intelligence. It was necessary to perform three classification tasks, each focusing on a different stage of Parkinson's disease and a different symptom type. In addition to motor skills, clinical stage, and dementia status, symptoms were also assessed. Combining handcrafted textural feature engineering, multiple feature selectors, patch-based learning, and IMV techniques, a novel model was created tailored to each patient's characteristics. A good performance was achieved by the model for all classification tasks[20].

Author [21] proposes an ensemble of deep learning architectures to detect PD from offline handwriting. A total of two datasets were used for this study: the PaHaW data and the New Hand PD data. Transfer learning was considered to improve the model's generality. Five CNN models were used to create the ensemble classifier. There are eight separate handwritten tasks in the PaHaW dataset, so the prediction accuracy was calculated for each one. Each handwritten task in the PaHaW dataset was analyzed to determine the accuracy of the prediction. The authors have developed a number of CNNs to diagnose Parkinson's disease. Author [22] spiral and meander drawings can be used to detect Parkinson's disease using a vision transformer (ViT). A pytorch layer (base model), a dropout layer (dropout value = 0.1), and finally, a linear layer as a classifier were used in the model. It was based on DeiT, pretrained on ImageNet, and self-supervised with DINO. The author of this study [23] included 22 healthy individuals and several Parkinson's disease patients. In an effort to collect data on all subjects except one who is right-handed, we collaborated with the Casa di Cura Le Terrazze Institute. This project was in collaboration with Casa di Cura Le Terrazze, which included all but one right-handed subjects. Researchers used handwritten data collected from subjects to identify patterns for the diagnosis of Parkinson's disease by using a tablet application. On the basis of three different samples of handwritten data, three classification models have been proposed. The Mann-Whitney test was used to identify highly discriminatory features [24].

The author used electroencephalogram (EEG) signals to detect Parkinson's disease [25]. In comparison with other approaches that typically analyze PD using handwritten data, this approach was considerably different. A spectrogram was obtained by splitting EEG signals in half and transforming them using Gabor transforms. Each EEG signal was analyzed two times, resulting in two spectrograms. 2D-CNNs were employed to classify the spectrogram signals of control,

medication-treated, and non-treated Parkinson's disease patients[26]. PD patients and healthy individuals were evenly divided into two groups in the study presented in [27]. Seventeen subjects in each group participated in the study. Various data from patients were collected using an iPad Pro equipped with a stylus for this study. AlexNet architecture has been proposed based on data from writing and drawing tests. Using deep learning networks, their research found 93% accuracy in diagnosing Parkinson's disease. As a result of the author's work [28], a low-cost classifier for diagnosing Parkinson's disease was developed. Signals recorded for the subject's response to vertical ground forces are included in the VGFR dataset. A dataset containing responses to voice impairments can be found in the voice impairment dataset. The Max Little University of Oxford has donated a dataset of 91 voice measurements for people with voice impairments. Two Parkinson's symptoms are detected using deep learning, namely impairment of gait and impairment of speech. This paper presents a two-module model; one identifies distorted walking patterns using VGFR spectrograms, and the other identifies patients' speech distortions as a classifier for voice impairment. A sensor reading for each patient is converted into a pattern as part of the first module. Spectrograms are made using these signal values to build patterns.

3. METHODOLOGY

PD detection model materials and methods are described in this section. The proposed model analyses three models of learning: transfer learning, K-nearest neighbour classification, and genetic algorithm optimization.

3.1. Transfer Learning

Pre-trained networks may be able to transfer knowledge to new use cases, reducing resource usage and improving efficiency. As a general rule, transfer learning refers to using previously learned knowledge to solve a new, more complicated problem. In addition to saving valuable training time, transfer learning can also allow a new model to be trained without having to start from scratch. Additionally, this approach is beneficial when there is not enough data. It enables a user to solve a completely different problem by using a completely new dataset. Using transfer learning, users can adjust other hyperparameters and weights in the pretrained model as well, including the last layer's dimensions. Users can also fine-tune other hyperparameters by adjusting the output layer dimensions. Typically, the starting layer in transfer learning is fixed or locked and cannot be modified, while the last layer can be modified.

3.2. KNN Classifier

An uncomplicated and straightforward machine learning algorithm, k-NN, is known as k-nearest neighbour. As a result of its non-parametric nature, the algorithm achieves its simplicity since no parameters are considered. During the final stages of this algorithm, the algorithm acts on the data, which is known as lazy learning. In addition to classification, the algorithm is widely used for regression problems as well. Since k represents the neighbouring data points surrounding the new data point, understanding and applying the algorithm is straightforward. Data points are compared with their neighbours (k) and then grouped based on their similarity [29]. Typically, K is between three and five at the beginning of the algorithm as part of the randomness process. Comparing data points based on distances, especially Euclidean distances, will determine similarity. Data points that are closest to each other have the least Euclidean distance and are assigned to the least Euclidean distance group.

Currently, no method has been developed to determine the optimal number, so there will be some trial and error involved. Generally, k is set at 5, but there may be times when this value must be changed due to certain problems. It is possible to get misleading interpretations if the values of k are very small, for example, 1 or 2. Once the feature vector has been optimized, KNNs are also used for classification after being used as the objective function during feature optimization.

3.3. Genetic Algorithm

Once the feature vector has been optimized, KNNs are also used for classification after being used as the objective function during feature optimization. A genetic algorithm shares a close relationship with evolution theory, much like natural selection. Natural selection defines survival as adjusting to new environments and habitats by mutating. It is essential that the fittest survive in natural selection. Every generation contains unique individuals, known as populations. Search space consists of the points of a population [29]. Additive models consist of five phases: initial population, fitness function, crossover, mutation, and selection. GAs follow a basic process, as shown in Figure 1.

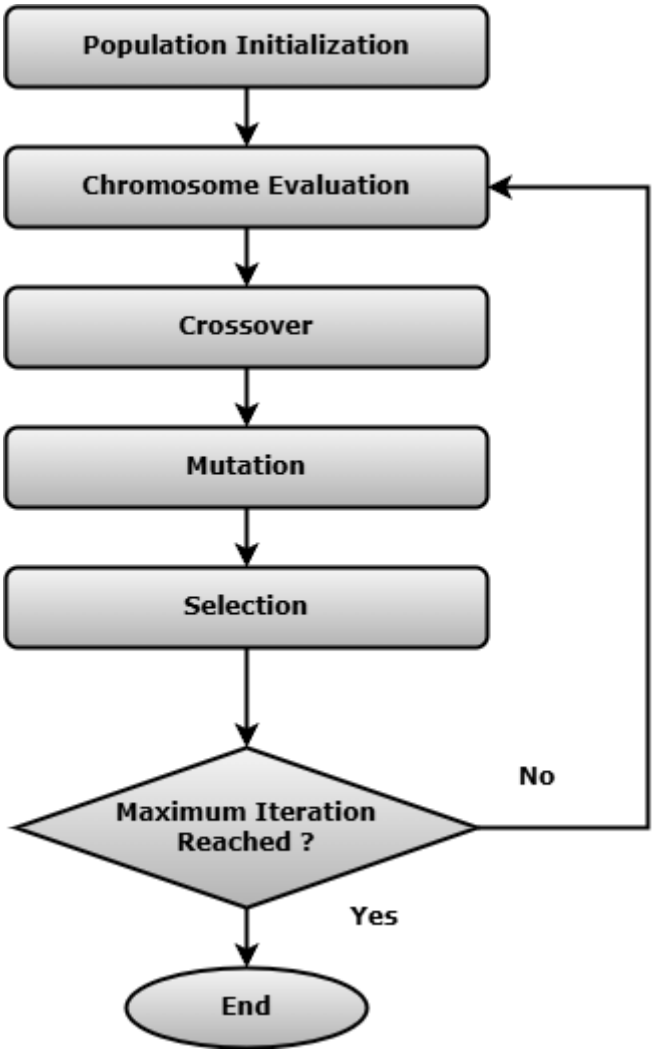

Fig. 1. Process involving genetic algorithms.

Let's examine the population's initial phase, which includes all of them. Individuals are represented by unique strings, meaning each individual can solve any problem. Following this phase is the fitness function. Individual fitness levels are displayed based on the fitness levels of the population. The ability to compete with other individuals is essential for survival in one's habitat. Using the fitness function, you can calculate a fitness level score, and based on that score, you can select the individuals. The third phase consists of selecting the most fit individuals to pass on their genes. The fitness scores of a few individuals are used to select them in this phase. A high fitness score is also associated with a higher chance of reproducing offspring for parents. As part of the crossover process, a very significant phase is taking place after the selection phase. Genetic material is randomly selected from the chromosomes of both parents and swapped. This offspring then becomes a part of the existing population since it has half the genes of both parents. During the mutation phase of the genetic algorithm, each newly created offspring's genes are altered slightly. As the environment changes, these changes will occur. Because of mutations, the new population can cope with the changes and thrive instead of going extinct as a result of not adapting. When there are no mutations in the offspring, the algorithm terminates.

3.4. Dataset Description

Experimental aspects of the study were conducted based on the NewHandPD [30] dataset. Our dataset is entirely composed of handwritten specimens that provide useful information for our research. We have published this document. For the NewHandPD, images were collected using a smart pen and a tablet. The HandPD dataset has been extended [31]. This data set contains 594 images of 160 males and 104 females. There are two types of groups in these data sets: Healthy Groups and Patient Groups. In the healthy group, 315 samples were collected, while in the patient group, 279 samples were collected. Men and women in each category are sampled based on the drawing they receive; the data are divided into three

groups: Circle, Meader, and Spiral. It is possible to represent a group using a circle, in which a member in good health and a patient draw a circle around the given area. A comparison of healthy and patient images was made in our study by combining images from all groups. In the healthy group, images are included from the three healthy groups; in the patient group, images are included from all three patient groups, and examples are included in each group.

3.5. Data Preprocessing

If the images of each group belong to the same category, they are categorized as Healthy and Patient. Transfer Learning (TL) models use 256x256 images. For some TL models, image size requirements are set in advance to prevent future problems. Since TL models are trained only on coloured images, RGB images must be read for them to work. A 256-bit reduction is also performed on the grey values of the image pixels so that they can be reduced from 0 to 1. As a result, deep learning models perform best when the value is between 0 and 1.

3.6. Feature Extraction

Three transfer learning models are employed in this study for feature extraction [29].

3.6.1. ResNet50

Using ImageNet, ResNet50 was trained as a 50-layer deep neural network. The network has been trained using more than a million images. This network's input shape is 2242243. It was originally designed for computer vision, but it has proved useful in other applications as well.

3.6.2. VGG19

In addition to ImageNet, VGG19 was also educated on ImageNet. There are 19 layers in the network, and several different characteristics were used to train it. Additionally, this network can be used for other image-related tasks, including computer vision.

3.6.3. Inception-V3

The Inception model uses deep neural networks. The model is comprised of 48 convolutional, pooling, dropout, and asymmetric layers. It is more accurate and computationally efficient than its predecessors (v1 and v2).
Using these transfer learning models, we freeze their training and remove their tops. We have stopped training these models and are relying on weights from previous training (ImageNet) to obtain features. For our study, we do not need up to 20,000 classes in the top set of these models. Our approach was to layer our neural network (top) over each TL model in order to employ TL models effectively. A thick layer is layered atop a flattened layer as our top layer. Each model's features are extracted from the dense layer by adding nodes. Diagram showing the structure of a model for extracting features from input images using individual transfer learning. A diagram showing all three TL models together illustrates the entire feature extraction process.
To determine the final input image format, each TL model extracts 100 features from each input image and layers them horizontally on top of each other. The final input image will be formatted as 300 (100 + 100 + 100). We extracted 300 features from each image in our dataset, which resulted in 594 images after we extracted all their features.

3.7. Support Vector machine

As a discriminative classifier designed by Vapnik for healthcare, it is effective at detecting anomalies in biomedical signals because it is capable of analyzing nonlinear and high-dimensional data [32]. Using the training set of some well-known data, this classifier differentiates the unknown validation set of data into appropriate classes.
An optimal hyperplane is constructed for binary classification to differentiate samples. The input sample is represented by $y \in \{+1, -1\}$, and the class label is represented by $y \in \{+1, -1\}$. Hence, the hyperplane is defined as follows:

$$w^T.x + b = 0 \tag{1}$$

An orthogonal coefficient vector w is defined in Eq. (1) above. A distance between its points describes a dataset's origin. An SVM is primarily used to determine b and w parameters. It is necessary to minimize the hyperplane $||w^2||$ Under the conditions of $y_i(w^T.x + b) \geq 1$ in order to obtain the ideal hyperplane. This leads to the following formulation of the optimization problem:

$$\begin{gathered} minimize \frac{1}{2}||w^2|| \\ Subject\ to\ y_i(w^T.x + b) \geq 1, \qquad i = 1,2, \ldots . n \end{gathered} \tag{2}$$

A linear problem is solved by applying Lagrange multipliers. An example of a support vector is a data point that lies on the decision boundary. It is, therefore, possible to calculate the value as follows:

$$w = \sum_{i=1}^{n} \alpha_i y_i x_i \tag{3}$$

Equation (3) expresses n support vectors and α_i as Lagrange multipliers. Using w as a starting point, we can calculate b as follows:

$$y_i(w^T . x_i + b) - 1 = 0 \tag{4}$$

Based on this equation, we can express the linear discriminant function as follows:

$$\hat{y} = sgn\left(\sum_{i=1}^{n} \alpha_i y_i x^T x_i + b\right) \tag{5}$$

With SVM, a kernel trick is employed to deal with nonlinear problems. Based on this, we can give the decision function as follows:

$$\hat{y} = sgn\left(\sum_{i=1}^{n} \alpha_i y_i K(x_i, x) + b\right) \tag{6}$$

An unlabeled input x is given the kernel zed label $\hat{y}$ and the sgn function defines a positive or negative classification. A positive definite kernel function, such as $K(x_i, x) = \exp(-\gamma||x - x_i||^2$ or $K(x_i, x) = (\mathrm{x}^{\mathrm{T}} x_i + I)^d$, which meets Mercer's constraint, is usually acceptable as a kernel function [33]. The purpose of this subsection is to give a brief overview of SVM. A comprehensive description of SVM concepts can be found in [34], which provides further reading.

3.8. Binary GWO

Grey wolf optimization is a bio-inspired metaheuristic that was introduced by [35]. Grey wolves are imitated in their leadership and hunting behaviours by GWO. The grey wolf belongs to the Canidae family and is a socially dominant species. It is well known that they hunt in swarms based on swarm intelligence. Unlike other predators, the grey wolf hunts its prey in packs of five to twelve individuals (i.e., a pack). Alphas, betas, deltas, and omegas are the four levels of leaders. α(dominant) wolves are the leaders of the cluster, occupying the highest hierarchy. Alpha makes all hunting decisions, sleeping locations, wake-up times, and discipline. [35]. All members of the group are required to follow its verdicts.

The alpha wolf is followed by the beta wolf, who acts as the alpha's adviser and enforces the alpha's orders. The act of dominating lower-level wolves is the result of β aiding α in making decisions and assisting A with other tasks. In the event that one of the subleaders in the group becomes very old or dies, betas have the highest likelihood of becoming true wolves. As a result, δ follows. During the chorus, they bow to the wolves of both levels, and they instruct their subordinates to do the same (i.e., below-level wolves). A hunter, sentinel, scout, caretaker, or elder falls into this level. The hunters provide food for the group and aid in hunting targets. A sentinel protects and ensures the group's safety. Any hazards within the search space are reported to the group by the scouts. Wolves are cared for by caregivers when they are injured, sick, or weak. In addition, older wolves are experts who were previously α or β.

At the very bottom of the dominance structure, there are omega wolves. Scapegoats are used. The wolves are always subordinated to other wolves at a higher level. Among all the omega wolves, only they are allowed to consume food. Despite being a not so important w wolf, dropping the wolf causes some complications and internal fighting among the cluster. As a result of the fury of all wolves, all frustration and violence has been extinguished. In this way, the hierarchy is preserved, and everyone is satisfied. Some groups also include babysitters.

Additionally, grey wolves hunt in packs, which is a distinctive characteristic. Specifically, [36] specifies a number of steps for group hunting: (i) tracking the prey, chasing it, grasping it, enclosing it and harassing it until stability is achieved, and (ii) attacking and killing it. Wolf encirclement can be defined by the following equations [35].

$$\vec{N} = \left|\overrightarrow{K}.\vec{P}_{prey}(i) - \vec{P}_{SA}(i)\right| \tag{7}$$

$$\vec{P}_{wolf}(i+1) = \vec{P}_{prey}(i) - \vec{M}.\vec{N} \tag{8}$$

The current version is specified here. $\vec{P}_{prey}$ vectors represent the position vectors of $\vec{P}_{SA}$ whereas search agent vectors represent the position vectors of search agents. A coefficient vector $\vec{M}$ and a coefficient vector $\vec{N}$ are estimated as follows:

$$\vec{M} = 2.\vec{m}.\overrightarrow{a_1} - \vec{m} \tag{9}$$

$$\vec{K} = 2.\vec{a}_2 \tag{10}$$

An arbitrary value for A is in the [0,1] range, and a value for B is in the [0,1] range. Where $\vec{a}_1$ and $\vec{a}_2$ has arbitrary values in [0,1]. $\vec{m}$ slowly decreases from 2 to 0 during iterations, acting as a controlling element. As the agents will be able to locate the target, they will be able to surround it easily. The alpha wolf directs all tasks. As each wolf hunts, it updates its location based on the optimal location of its α, β, and δ. Based on the following model, searching agents encircle each other.

$$\overrightarrow{N_\alpha} = \left|\overrightarrow{K_1}.\vec{P}_\alpha - \vec{P}\right| \quad (11)$$
$$\overrightarrow{N_\beta} = \left|\overrightarrow{K_2}.\vec{P}_\beta - \vec{P}\right| \quad (12)$$
$$\overrightarrow{N_\delta} = \left|\overrightarrow{K_3}.\vec{P}_\delta - \vec{P}\right| \quad (13)$$
$$\overrightarrow{P_1} = \left|\overrightarrow{P_\alpha} - \overrightarrow{M_1}.\vec{N}_\alpha\right| \quad (14)$$
$$\overrightarrow{P_\beta} = \left|\overrightarrow{P_\beta} - \overrightarrow{M_2}.\vec{N}_\beta\right| \quad (15)$$
$$\overrightarrow{P_\delta} = \left|\overrightarrow{P_\gamma} - \overrightarrow{M_1}.\vec{N}_\gamma\right| \quad (16)$$
$$\vec{P}_{SA}(i+1) = \frac{\vec{P}_1 + \vec{P}_2 + \vec{P}_3}{3} \quad (17)$$

Agents searching for targets attack only when they have found a stable target. According to $\vec{m}$, the following calculation is made for this process.

$$\vec{m} = 2 - \frac{2_i}{I_{max}} \quad (18)$$

The searching agent moves in the direction of the target when moving towards it, while $\left|\vec{M}\right| > 1$ deviates from the target and looks for a different prey when it deviates from the target. Grey wolves determine what is a good position by combining α, β, and δ. In the same way, $\vec{M}$ and $\vec{K}$ determine when the optimization process will go into the exploitation phase and the exploration phase. Searching agents can be made to diverge or converge by specifying an arbitrary value for $\vec{M}$. The arbitrarily small values of $\vec{K}$ in the range [0,2] are important in preventing local solutions from stagnating.

It is a challenge for search agents to determine distances between themselves and targets by weighing them on their own. A target with $\left|\vec{K}\right| > 1$ maximizes its impact, whereas one with $\left|\vec{K}\right| < 1$ minimizes it randomly. This course emphasizes or de-emphasizes exploitation or exploration processes while $\vec{M}$ and $\vec{K}$ Are tuned cautiously. The algorithm will ultimately terminate when a certain condition is met, and the best value for α will be returned.

It is the original grey wolf optimizer that is used to solve continuous optimization problems. Discrete optimization snags require a binary optimizer. This work implements GWO in binary form. According to the proposed BGWO, each searching agent is composed of a fag vector whose size is equal to the attribute's size. Using Eq. (17), we can express discrete positions of searching agents in the following manner:

$$F = \begin{cases} 1 & P_{ij} > 0.5 \\ 0 & Otherwise \end{cases} \quad (19)$$

In this case, P_{ij} represents the location of the wolf in the j^{th}

4. RESULT AND DISCUSSION

Two approaches were used for Parkinson's Disease (PD) classification: fine-tuning pretrained CNN models and freezing. AlexNet, VGG16, and ResNet-50 networks were used to process the enhanced handwriting images. Based on the BGWO algorithm, the features of the best-trained network were selected and classified using SVMs. On a 16GB RAM and 512GB SSD Intel i7 system, the experiment used holdout (70:30, 80:20, 90:10) and cross-validation methods for performance evaluation.

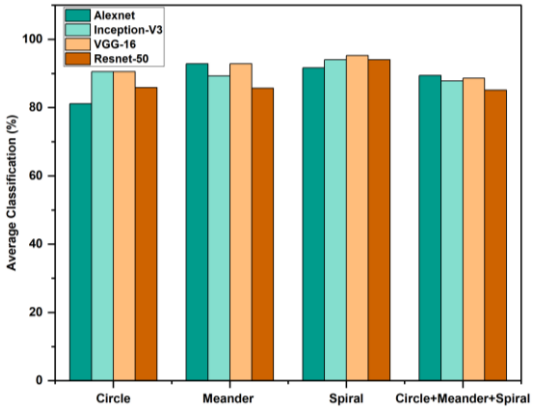

Fig. 2. Classification accuracy was obtained using fine-tuning with different deep-learning models.

VGG16 was used to extract features, BGWO was used to optimize, and SVM was used to classify the results. Metrics such as accuracy, sensitivity, and specificity are used to evaluate a classifier's performance. TPs (True Positives) and TNs (True

Negatives) represent correctly classified positives and negatives, whereas FPs and FNs (False Positives) are misclassified positives and negatives, respectively.

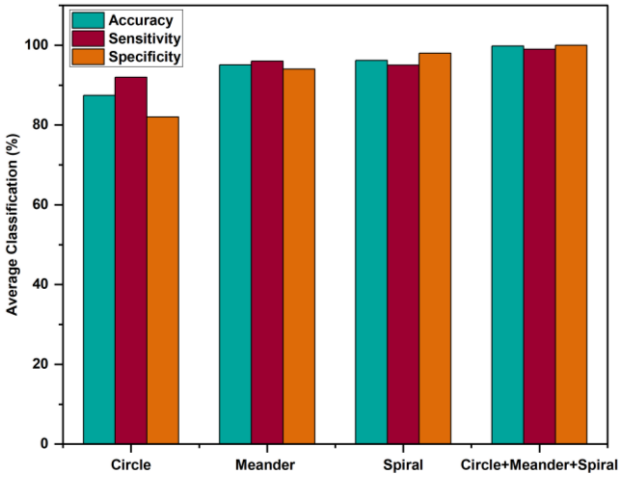

Fig. 3. Results were obtained after extracting features using VGG16.

A comparison between the proposed approach and previously published methods is provided in this study. In Figure 4, a comparative study of the suggested approach and relevant studies is presented. The proposed approach is benchmarked against previously published methods using the same dataset. In comparison with existing approaches, the proposed method achieves a slight improvement, possibly paving the way for new advances. Furthermore, the proposed model achieved 94.7% accuracy, compared to 95.5%, 95.5%, and 83.6% accuracy for other models.

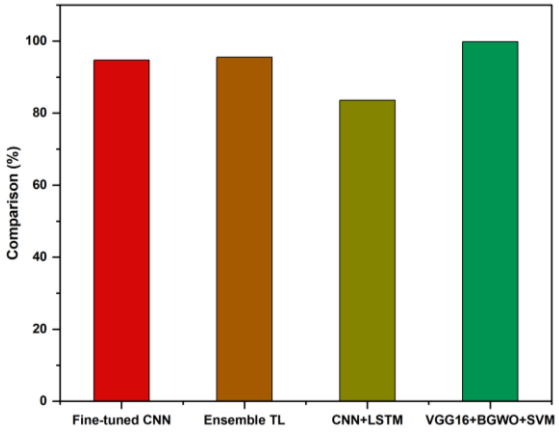

Fig. 4. Comparison to the related work.

5. CONCLUSION

By combining deep transfer learning with optimization, this study presents a robust, AI-driven method for diagnosing Parkinson's disease. By combining CNN-based feature extraction with SVM classification and genetic algorithm-based feature selection, the classification accuracy of Parkinson's disease is exceptionally high. Compared to traditional methods, this approach is more accurate and reliable. Compared to traditional methods, this approach is more accurate and reliable. A number of further improvements can be made to the framework, such as incorporating larger and more diverse datasets, optimizing compute efficiency, and extending its application to other neurological disorders.

Conflicts Of Interest

The paper's disclosure section confirms the author's lack of any conflicts of interest.

Funding

The author's paper does not provide any information on grants, sponsorships, or funding applications related to the research.

Acknowledgment

The author acknowledges the assistance and guidance received from the institution in various aspects of this study.

References

[1] R. Katzenschlager, C. Sampaio, J. Costa, A. Lees, and C. M. D. Group, "Anticholinergics for symptomatic management of Parkinson s disease," Cochrane Database of Systematic Reviews, vol. 2010, no. 1, 1996, Accessed: Mar. 16, 2025. [Online]. Available: https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD003735/abstract

[2] S. Verma et al., "An automated face mask detection system using transfer learning based neural network to preventing viral infection," Expert Systems, vol. 41, no. 3, p. e13507, Mar. 2024, doi: 10.1111/exsy.13507.

[3] H. Li et al., "A hybrid feature selection algorithm based on a discrete artificial bee colony for Parkinson's diagnosis," ACM Transactions on Internet Technology, vol. 21, no. 3, pp. 1–22, 2021.

[4] K. Suzuki, "Overview of deep learning in medical imaging," Radiol Phys Technol, vol. 10, no. 3, pp. 257–273, Sep. 2017, doi: 10.1007/s12194-017-0406-5.

[5] M. Bakator and D. Radosav, "Deep Learning and Medical Diagnosis: A Review of Literature," MTI, vol. 2, no. 3, p. 47, Aug. 2018, doi: 10.3390/mti2030047.

[6] P. Rani, S. P. Yadav, P. N. Singh, and M. Almusawi, "Real-World Case Studies: Transforming Mental Healthcare With Natural Language Processing," in Demystifying the Role of Natural Language Processing (NLP) in Mental Health, A. Mishra, S. P. Yadav, M. Kumar, S. M. Biju, and G. C. Deka, Eds., IGI Global, 2025, pp. 303–324. doi: 10.4018/979-8-3693-4203-9.ch016.

[7] W. Poewe et al., "Parkinson disease," Nat Rev Dis Primers, vol. 3, no. 1, p. 17013, Mar. 2017, doi: 10.1038/nrdp.2017.13.

[8] S. H. Fox et al., "The Movement Disorder Society Evidence-Based Medicine Review Update: Treatments for the motor symptoms of Parkinson's disease," Movement Disorders, vol. 26, no. S3, Oct. 2011, doi: 10.1002/mds.23829.

[9] K. R. Chaudhuri and A. H. Schapira, "Non-motor symptoms of Parkinson's disease: dopaminergic pathophysiology and treatment," The Lancet Neurology, vol. 8, no. 5, pp. 464–474, May 2009, doi: 10.1016/S1474-4422(09)70068-7.

[10] C. R. Pereira, D. R. Pereira, J. P. Papa, G. H. Rosa, and X.-S. Yang, "Convolutional Neural Networks Applied for Parkinson's Disease Identification," in Machine Learning for Health Informatics, vol. 9605, A. Holzinger, Ed., in Lecture Notes in Computer Science, vol. 9605. , Cham: Springer International Publishing, 2016, pp. 377–390. doi: 10.1007/978-3-319-50478-0_19.

[11] J. Xu and M. Zhang, "Use of magnetic resonance imaging and artificial intelligence in studies of diagnosis of Parkinson's disease," ACS chemical neuroscience, vol. 10, no. 6, pp. 2658–2667, 2019.

[12] S. H. Kassani, P. H. Kassani, M. J. Wesolowski, K. A. Schneider, and R. Deters, "Breast Cancer Diagnosis with Transfer Learning and Global Pooling," in 2019 International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, Korea (South): IEEE, Oct. 2019, pp. 519–524. doi: 10.1109/ICTC46691.2019.8939878.

[13] S. Liu, G. Tian, and Y. Xu, "A novel scene classification model combining ResNet based transfer learning and data augmentation with a filter," Neurocomputing, vol. 338, pp. 191–206, Apr. 2019, doi: 10.1016/j.neucom.2019.01.090.

[14] P. Rani, S. Verma, S. P. Yadav, B. K. Rai, M. S. Naruka, and D. Kumar, "Simulation of the lightweight blockchain technique based on privacy and security for healthcare data for the cloud system," International Journal of E-Health and Medical Communications (IJEHMC), vol. 13, no. 4, pp. 1–15, 2022.

[15] K. Wu, D. Zhang, G. Lu, and Z. Guo, "Learning acoustic features to detect Parkinson's disease," Neurocomputing, vol. 318, pp. 102–108, Nov. 2018, doi: 10.1016/j.neucom.2018.08.036.

[16] J.-S. Provost, A. Hanganu, and O. Monchi, "Neuroimaging studies of the striatum in cognition Part I: healthy individuals," Frontiers in systems neuroscience, vol. 9, p. 140, 2015.

[17] Z. Fang, "Improved KNN algorithm with information entropy for the diagnosis of Parkinson's disease," in 2022 International Conference on Machine Learning and Knowledge Engineering (MLKE), IEEE, 2022, pp. 98–101. Accessed: Mar. 16, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9763618/

[18] P. Rani and R. Sharma, "Intelligent transportation system for internet of vehicles based vehicular networks for smart cities," Computers and Electrical Engineering, vol. 105, p. 108543, 2023.

[19] E. Kaplan et al., "Novel nested patch-based feature extraction model for automated Parkinson's Disease symptom classification using MRI images," Computer Methods and Programs in Biomedicine, vol. 224, p. 107030, Sep. 2022, doi: 10.1016/j.cmpb.2022.107030.
[20] G. Ansari, P. Rani, and V. Kumar, "A novel technique of mixed gas identification based on the group method of data handling (GMDH) on time-dependent MOX gas sensor data," in Proceedings of International Conference on Recent Trends in Computing: ICRTC 2022, Springer, 2023, pp. 641–654.
[21] M. Gazda, M. Hires, and P. Drotár, "Ensemble of convolutional neural networks for Parkinson's disease diagnosis from offline handwriting," 2022, Accessed: Mar. 16, 2025. [Online]. Available: https://accedacris.ulpgc.es/handle/10553/116094
[22] M. Mohaghegh and J. Gascon, "Identifying Parkinson's Disease using Multimodal Approach and Deep Learning," in 2021 6th International Conference on Innovative Technology in Intelligent System and Industrial Applications (CITISIA), Sydney, Australia: IEEE, Nov. 2021, pp. 1–6. doi: 10.1109/CITISIA53721.2021.9719945.
[23] M. Fratello et al., "Classification-Based Screening of Parkinson's Disease Patients through Graph and Handwriting Signals," in The 2nd International Electronic Conference on Applied Sciences, MDPI, Oct. 2021, p. 49. doi: 10.3390/ASEC2021-11128.
[24] A. Gold, "Understanding the mann-whitney test," Journal of Property Tax Assessment & Administration, vol. 4, no. 3, pp. 55–57, 2007.
[25] H. W. Loh et al., "GaborPDNet: Gabor Transformation and Deep Neural Network for Parkinson's Disease Detection Using EEG Signals," Electronics, vol. 10, no. 14, p. 1740, Jul. 2021, doi: 10.3390/electronics10141740.
[26] P. Rani, P. N. Singh, S. Verma, N. Ali, P. K. Shukla, and M. Alhassan, "An implementation of modified blowfish technique with honey bee behavior optimization for load balancing in cloud system environment," Wireless Communications and Mobile Computing, vol. 2022, pp. 1–14, 2022.
[27] S. Nõmm, S. Zarembo, K. Medijainen, P. Taba, and A. Toomela, "Deep CNN Based Classification of the Archimedes Spiral Drawing Tests to Support Diagnostics of the Parkinson's Disease," IFAC-PapersOnLine, vol. 53, no. 5, pp. 260–264, 2020, doi: 10.1016/j.ifacol.2021.04.185.
[28] Shivangi, A. Johri, and A. Tripathi, "Parkinson Disease Detection Using Deep Neural Networks," in 2019 Twelfth International Conference on Contemporary Computing (IC3), Noida, India: IEEE, Aug. 2019, pp. 1–4. doi: 10.1109/IC3.2019.8844941.
[29] C.-L. Liu, C.-H. Lee, and P.-M. Lin, "A fall detection system using k-nearest neighbor classifier," Expert Systems with Applications, vol. 37, no. 10, pp. 7174–7181, Oct. 2010, doi: 10.1016/j.eswa.2010.04.014.
[30] C. R. Pereira, S. A. Weber, C. Hook, G. H. Rosa, and J. P. Papa, "Deep learning-aided Parkinson's disease diagnosis from handwritten dynamics," in 2016 29th SIBGRAPI conference on graphics, patterns and images (SIBGRAPI), Ieee, 2016, pp. 340–346. Accessed: Mar. 16, 2025. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/7813053/
[31] C. R. Pereira et al., "A new computer vision-based approach to aid the diagnosis of Parkinson's disease," Computer Methods and Programs in Biomedicine, vol. 136, pp. 79–88, Nov. 2016, doi: 10.1016/j.cmpb.2016.08.005.
[32] V. Vapnik, The nature of statistical learning theory. Springer science & business media, 2013. Accessed: Mar. 16, 2025. [Online]. Available: https://books.google.com/books?hl=en&lr=&id=EqgACAAAQBAJ&oi=fnd&pg=PR7&dq=related:WqvJVU5Hj0kJ:scholar.google.com/&ots=g5H-fC7_83&sig=KDPtj20HxdodI7WFmUSqvXspVb4
[33] B. Schölkopf, C. J. Burges, and A. J. Smola, Advances in kernel methods: support vector learning. MIT press, 1999. Accessed: Mar. 16, 2025. [Online]. Available: https://books.google.com/books?hl=en&lr=&id=_NYamXKkNM8C&oi=fnd&pg=PP13&dq=related:23zawhiI6fYJ:scholar.google.com/&ots=RBdL8H5IB8&sig=1JJOni7tswaTbvCKGmHHNrIT_Xo
[34] N. Cristianini and J. Shawe-Taylor, An introduction to support vector machines and other kernel-based learning methods. Cambridge university press, 2000. Accessed: Mar. 16, 2025. [Online]. Available: https://books.google.com/books?hl=en&lr=&id=I_0gAwAAQBAJ&oi=fnd&pg=PR13&ots=g4bp9akp5I&sig=Et-LA8DZPu4504MTmogMqC53sqY
[35] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," Advances in engineering software, vol. 69, pp. 46–61, 2014.
[36] C. Muro, R. Escobedo, L. Spector, and R. P. Coppinger, "Wolf-pack (Canis lupus) hunting strategies emerge from simple rules in computational simulations," Behavioural processes, vol. 88, no. 3, pp. 192–197, 2011.