Journal of Applied Informatics and Computing (JAIC)
Vol.10, No.2, April 2026, pp. 1182~1190
e-ISSN: 2548-6861

Parkinson's Disease Detection from Handwriting Using VGG-16 and Random Forest

Billy Franko [1]*, Derry Alamsyah [2]*
* Informatics Study Program, Multi Data Palembang University
billy.franko@mhs.mdp.ac.id [1], derry@mdp.ac.id [2]

Article Info

Article history:

Received 2026-02-04
Revised 2026-02-21
Accepted 2026-02-27

Keyword:

Early Detection,
Parkinson's Disease,
Random Forest,
VGG-16.

ABSTRACT

Parkinson's disease is one of the most common neurodegenerative diseases in Indonesia and remains incurable. However, early recognition of Parkinson's disease makes it possible for timely intervention to be conducted. Unfortunately, Parkinson's disease detection process can sometimes take a long time. This study proposes a system for early Parkinson's disease using offline handwriting images and machine learning techniques. The system employs VGG-16 for feature extraction and Random Forest classifier for prediction to recognize early signs of Parkinson's disease through three handwriting tasks, namely spirals, meander and circle using publicly available NewHandPD dataset containing 594 samples across all tasks. The model will be trained using original data as well as images processed with three preprocessing techniques, namely grayscale, grayscale with CLAHE and grayscale with CLAHE and Otsu thresholding. In the final testing phase using a held-out test set, the model trained on the original data achieved the best performance, achieving average accuracy of 94%. The best performing model will be hosted in a cloud-based environment and accessed by a developed software application through an API. A questionnaire was conducted using the USE Questionnaire, resulting in average score of 90,97%. Indicating a high level of user satisfaction for the application developed in this study.

I. Introduction

Neurodegenerative diseases consist of a diverse group of neurological disorders that affect the lives of millions of individuals worldwide resulting in progressive loss of neurons in central nervous systems (CNS) and/or peripheral nervous system (PNS). The collapse of neural network structure and function, together with the loss of neurons that cannot be renewed efficiently because of their differentiated nature, leads to the breakdown of the core communicative circuitry. This breakdown ultimately results in impairments on memory, cognition, behavior, and both sensory and motoric function [1].

Huntington's disease, Alzheimer's disease and Parkinson's disease are among the most common neurodegenerative diseases in Indonesia [2]. Parkinson's disease is a degenerative condition of the brain, primarily associated with motor symptoms including bradykinesia, tremor, rigidity, walking and balance trouble followed with a variety of non-motor complications including neuropsychiatric and autonomic symptoms, sleep-wake disturbances, pain and other sensory disturbances [3]. As of 2023, it was estimated that there are at least 880000 individuals in Indonesia living with this conidtion, and this number is expected to increase in the following years [4]. Although no definitive cure is currently available, early-stage identification offers the greatest opportunity for timely intervention. Neuroprotective treatment initiated during the prodromal stages has the potential to prevent clinical Parkinson's disease from emerging. Therefore, early detection is critically important to maximize the effectiveness of the neuroprotective treatment [5] [6].

Parkinson's disease can be diagnosed using both invasive and non-invasive methods. Invasive methods such as cerebrospinal fluid (CSF) and skin biopsy may reduce patient acceptance and limit their routine application in clinical practice [7]. Non-invasive methods include blood sampling and neuroimaging such as single-photon emission computed tomography (SPECT) and magnetic resonance imaging (MRI). However, blood sampling and neuroimaging may exhibit signs that overlap to other neurological diseases. In addition, SPECT and MRI require substantial maintenance and incur high operational costs [8][9]. Changes to the brain associated with Parkinson's disease such as neuronal loss, synaptic dysfunction, and brain atrophy, might cause malfunction in the motor system and its components. The alterations were reflected in the deterioration of previously acquired motor skills. Therefore, handwriting assessment can be considered to play a unique role in Parkinson's disease detection [10].

A need for speedy and reliable tools for the detection of Parkinson's disease arises from the lack of sufficient resources and the increasing number of suspected patients, which hinder the ability of practitioners to treat and diagnose each case effectively. As a result, Parkinson's disease is often only diagnosed only after it has progressed to later stages [8]. To address this challenge, numerous studies has proposed diagnosed tools to detect Parkinson's disease utilizing handwriting as its indicator. Examples of such approaches include ones developed in [10] and [11]. Handwriting based detection methods can be categorized into two approaches such as online and offline handwriting. Offline detection relies on images captured from handwriting samples produced on paper after patients has completed a predefined task. On the other hand, Online detection employ digital means such as tablet with a specialized pen that enable this approach to provide more information than offline detection. In summary, offline approach provides more accesibility, where as online approach provides more information.

One of the methods widely adopted by researchers to detect Parkinson's disease is machine learning. Machine learning techniques enable the identification of relevant features that are not traditionally considered in clinical diagnosis, thereby facilitating early stage detection of Parkinson's disease [12]. Research conducted by Ranjan et al. [13] which employs a Random Forest classifier with Histogram of Oriented Gradients (HOG) feature extraction, successfully developed a Parkinson's disease detection tool with accuracy reaching 86.67% and 83.3%.

For studies involving a small amount of dataset, researchers have found a way to mitigate data scarcity through Transfer Learning (TL) in Deep Learning (DL) models. By using TL, knowledge received by a domain can be transferred to the other domains, thereby increasing the accuracy of the other domain [14]. One example of using Convolutional Neural Network (CNN) with TL is presented in [14] which utilize VGG-16 and SqueezeNet architecture to NewHandPD and PaHaw datasets resulting in highest accuracy of 94.7%.

In addition, some researchers have adopted hybrid modelling approaches by combining multiple models at once. One such study was conducted by Das et al. [15], which utilize HOG and VGG-16 as feature extractor, with the extraction features being provided as input various classifiers such as KNN (K-Nearest Neighbor), SVM (Support Vector Machine), MLP (Multi Layer Perceptron), Random Forest and Naïve Bayes Model across two different datasets. The result shows the peak accuracy from dataset 1 of 96% by VGG-16 and Random Forest, while dataset 2 shows peak accuracy of 98% achieved by VGG-16 paired with Random Forest and SVM.

Based on these findings, VGG-16 and Random Forest were chosen as the primary models used to develop early Parkinson's disease detection tool in this study. These models were chosen in order to achieve high detection accuracy. Compared to previous studies, this study will emphasize both classification accuracy and practical implementation, resulting in a mobile-based software for early Parkinson's disesase screening. Additionally, this study will also make use of different preprocessing methods to evaluate its impact on the model effectiveness at recognizing Parkinson's disease through handwriting images. The proposed system is expected to result in a non-invasive and accessible Parkinson's disease early detection tool that may assist clinicians in identifying early signs of Parkinson's disease more effectively using handwriting images.

II. METHOD

This study employed an experimental approach by applying different preprocessing techniques using NewHandPD dataset to create a Parkinson's disease detection tool. The overall procedure is presented on Figure 1.

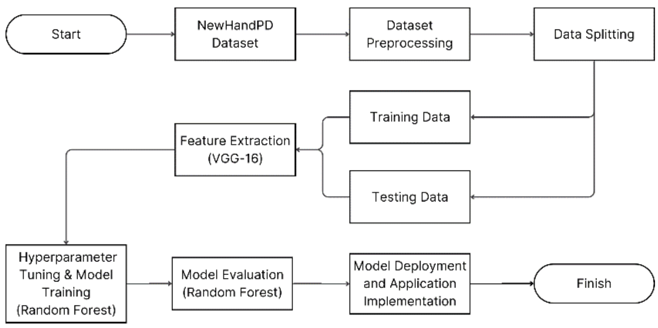


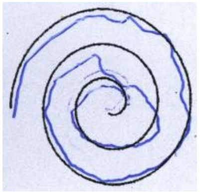
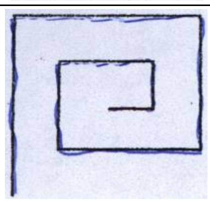
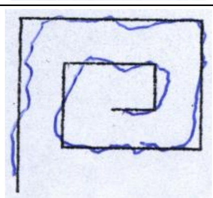
Figure 1. Study Procedure Workflow

A. Dataset Description

The study utilizes the publicly available NewHandPD dataset, which was collected in Botucatu Medical School, São Paulo State University, Brazil [16]. The dataset used in this study does not contain any personally identifiable information. It comprises handwritings images from 35 healthy individuals and 31 Parkinson's disease patients. The healthy group consists of 18 males and 17 females, with ages ranging from 14 to 79 years old. Among these individuals, 5 are left handed and 30 are right handed. Patient group consists of 21 males and 10 females with ages ranging from 38 to 78 years old. Among these individuals, 2 are left handed and 29 are right handed. Each participant completed twelve tasks, namely four spiral drawings, four meander drawings, two circle drawings (one performed in the air and one performed on paper) followed with diadochokinesis using both the left and right hands.

In this study, diadochokinesis signals and circle drawings that were performed in the air be excluded. Consequently, only spiral, meander and on-paper circle drawings were utilized, resulting in 9 handwriting images per individual and 594 data samples. Examples of the handwriting images included in the dataset are provided in Table 1.

TABLE I
HANDWRITING IMAGE EXAMPLE

Task	Image	
	Healthy	**Patient**
Spiral		
Meander		
Circle		

B. Dataset Preprocessing

Prior to the model training phase, all input data undergo through preprocessing to ensure maximal compatibility and effectiveness with the VGG-16 model. Each image is resized into 224x224 pixels while preserving its original size ratio with padding. The size was chosen to correspond with the standard input dimension for VGG-16 model.

In this study, the model will be trained on 3 different preprocessing techniques along with the original image, namely grayscale conversion, grayscale with Contrast Limited Adaptive Histogram Equalization (CLAHE), grayscale with CLAHE followed by Otsu thresholding and dilation. These pre-processing techniques were chosen for its respective advantages. Grayscale conversion simplifies colored image, while retaining its important structural information [17]. CLAHE enhances quality of image with better contrast without significant degradation on its quality [18]. Meanwhile, Otsu Thresholding is able to maximize threshold compatibility value, improving segmentation and enhancing feature discrimination [19].

C. Data Splitting

After preprocessing, the dataset was divided into training and testing data with the ratio of 80:20, where 80% of the data will be allocated for training and 20% for model testing. Testing data will be preserved until final evaluation to evaluate model generalization.

D. Model Architecture

This study employs a hybrid Convolutional Neural Network (CNN) architecture named VGG-16 and Random Forest classifier. In this study, VGG-16 was employed solely as feature extractor while Random Forest classifier served as the primary prediction model. Random Forest is composed of a collection of independent hierarchical classifiers in form of decision trees and employs a majority voting mechanism to predict whether an instance belong to a positive, negative or a neutral class [20] [21] [22]. VGG-16 was introduced by Karen Simonyan and Andrew Zisserman in 2014 in their paper entitled "*Very Deep Convolutional Networks for Large-Scale Image Recognition*". VGG-16 architecture accepts input images of size 224x224x3 and consist of 16 learnable layers including convolutional layers with 3 x 3 kernels designed for architectural simplicity, five max-pooling layers with a pooling size of 2 x 2, and two fully connected layers followed by a softmax classifier [23] [24]. The overall architecture of Random Forest and VGG-16 together with the image of the proposed hybrid model used in this study are presented in Figure 2, 3 and 4.

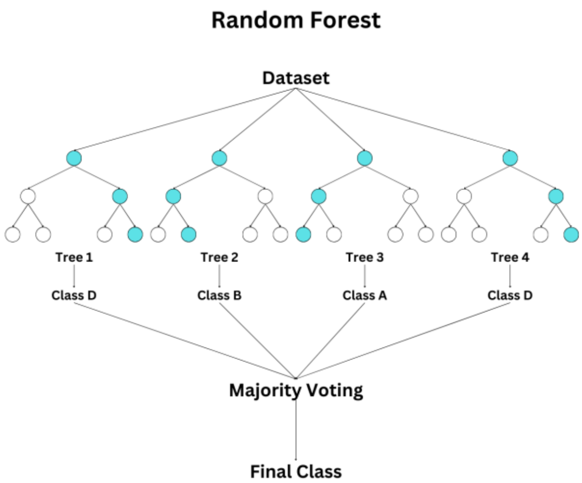


Figure 2. Random Forest Classifier (adapted from [25])

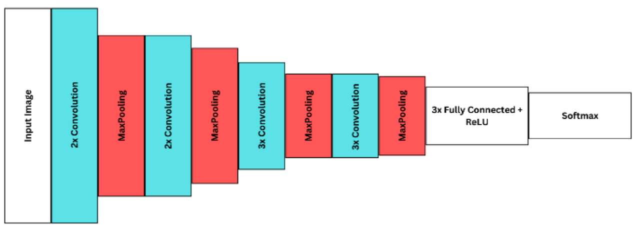


Figure 3. VGG-16 Architecture (adapted from [26])

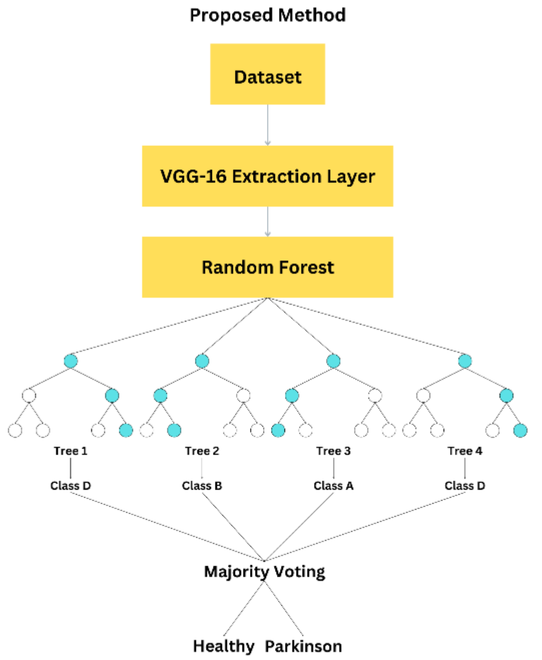


Figure 4. Proposed Method

E. Hyperparameter Tuning and Model Training

In this phase, feature vectors extracted by VGG-16 model were used as inputs for the Random Forest classifier. Before training the final model, hyperparameter tuning is performed to identify the combination that yielded the optimal performance. For the spiral and meander datasets, Stratified Group K-Fold Cross Validation will be applied while Stratified K-Fold Cross Validation will be used for the circle model with K = 5. Cross validations were conducted to all three extracted data using two different approach, (1) the original extracted feature and (2) features normalized with StandardScaler and reduced using Principal Component Analysis (PCA) to retain 85% of the variance (n_components =0.85). The hyperparameter combination that will be explored in this study are summarized in Table 2.

TABLE II
HYPERPARAMETER TUNING COMBINATION

Hyperparameter Combination (Stratified K-Fold Cross Validation)	
n_estimators	100, 200, 300
Max_depth	2, 4, 6, 8, 10
Min_samples_split	2, 5, 10
Min_samples_leaf	1, 2, 4

The combination of hyperparameters in Table 2 results in 135 distinct experimental scenarios, yielded a total of 675 during 5 splits of cross validation. After obtaining the best performing hyperparameters, the final model will be retrained using the parameter and the complete training dataset, after which the model can proceed to the testing phase for evaluation.

F. Model Evaluation

After all the preceding phases were completed, a final evaluation was carried out using a confusion matrix on the held-out test set to measure its accuracy, precision, recall and f1-score, to assess model generalization. Based on the evaluation results, the best performing model will be selected for deployment. Other than confusion matrix, additional metrics will also be presented, such as sensitivity, specificity and ROC-AUC. Sensitivity is a percentage of true positives correctly identified among all patientswith a condition while specificity is percentage of actual negatives correctly identified among all healthy individuals [27]. AUC is are under curve while ROC-AUC is a single number that summarized predictive performance, allowing comparison of curves [28]. The score of AUC are better the closer it is to 1 with indicators presented in Table 3 [29].

TABLE III
HYPERPARAMETER TUNING COMBINATION

Score	Indicator
$0.5 \leq AUC < 0.6$	Very Weak
$0.6 \leq AUC < 0.7$	Weak
$0.7 \leq AUC < 0.8$	Fair
$0.8 \leq AUC < 0.9$	Good
$0.9 \leq AUC < 1$	Excellent

G. Deployment and Implementation

The selected model will be selected to be deployed to a cloud-based environment on Microsoft Azure. Following that, a RESTful web API will be developed using Laravel framework then hosted through Vercel to serve as an intermediary layer between the Model and Android application. The android application itself, named Parkinvent will be developed in Android Studio using Kotlin programming language. In addition, Supabase was utilized as an online database to store user information, and patient test results. The application communicates with the API by sending input images and receiving prediction result from the deployed model. This android application represents the final output of this study.

H. Study Result Evaluation

An evalution will be conducted using a survey to receive user feedback on the developed android application. The questions on the survey are designed based on aspects included in USE questionnaire. USE questionnaire covers four main aspects such as Usefulness, Satisfaction, Ease of Learning and Ease of Use [30]. The collected responses were analyzed to asssess user satisfaction with the proposed system.

III. RESULT AND DISCUSSION

This section will present the results and describe more in depth about the implementation method and results achieved from this study, which are divided into two parts: the machine learning model and the software system which includes the API and Android app.

A. Model Training Result

This subsection presents the result achieved from machine learning model training as described in the previous section. The training performance is illustrated using bar charts in Figure 5-8, while a more detailed results derived from confusion matrix are provided in Table 4.

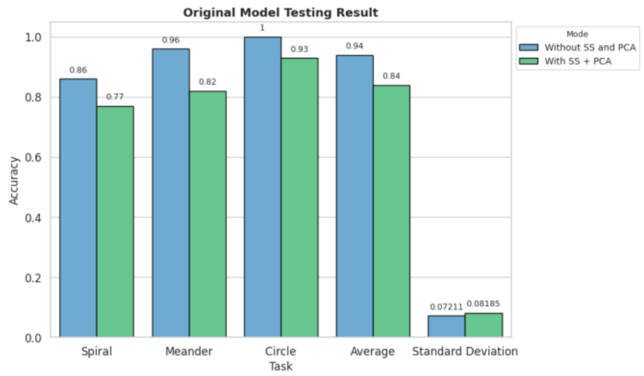


Figure 5. Original Image Model Testing Result

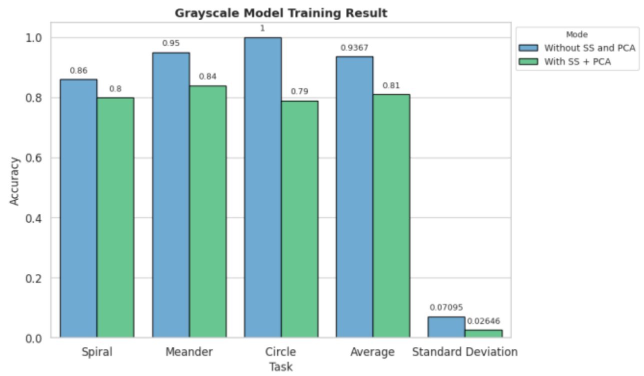


Figure 6. Grayscale Image Model Testing Result

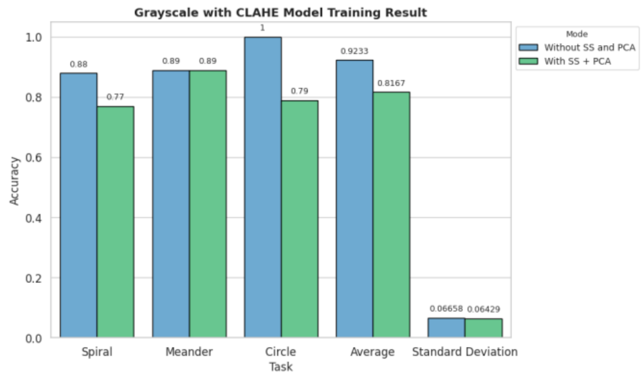


Figure 7. Grayscale and CLAHE Image Model Testing Result

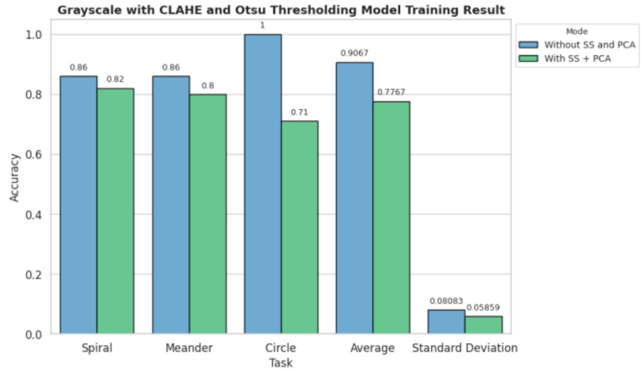


Figure 8. Grayscale, CLAHE and Otsu Image Model Testing Result

TABLE IV
CONFUSION MATRIX FOR EACH IMAGE TYPE

Image Type	Task	*Accuracy*	*Precision*	*Recall*	*F1-Score*
Original	Spiral	86%	87%	86%	86%
	Circle	100%	100%	100%	100%
	Meander	96%	97%	96%	96%
Original (SS + PCA)	Spiral	77%	82%	77%	76%
	Circle	93%	94%	93%	93%
	Meander	82%	87%	82%	82%
G	Spiral	86%	86%	86%	86%
	Circle	100%	100%	100%	100%
	Meander	95%	95%	95%	95%
G (SS + PCA)	Spiral	80%	81%	80%	80%
	Circle	79%	79%	78%	78%
	Meander	84%	88%	84%	84%
GC	Spiral	88%	89%	88%	87%
	Circle	100%	100%	100%	100%
	Meander	89%	91%	89%	89%
GC (SS + PCA)	Spiral	77%	79%	77%	76%
	Circle	79%	85%	79%	78%
	Meander	89%	91%	89%	89%
GCO	Spiral	86%	89%	86%	85%
	Circle	100%	100%	100%	100%
	Meander	86%	89%	86%	85%
GCO (SS + PCA)	Spiral	82%	87%	82%	82%
	Circle	71%	82%	71%	69%
	Meander	80%	86%	80%	80%
G = Grayscale; C = CLAHE; O = Otsu; SS = Standard Scaler					

TABLE V
ADDITIONAL METRICS

Image Type	Task	*Specificity*	*Sensitivity*	*ROC-AUC*
Original	Spiral	96%	75%	0.98
	Circle	100%	100%	1
	Meander	100%	93%	1
Original (SS + PCA)	Spiral	96%	57%	0.96
	Circle	100%	86%	0.96
	Meander	100%	64%	0.98
G	Spiral	93%	79%	0.98
	Circle	100%	100%	1
	Meander	100%	89%	1
G	Spiral	86%	75%	0.91

(SS + PCA)	Circle	86%	71%	0.92
	Meander	100%	68%	0.99
GC	Spiral	96%	79%	0.98
	Circle	100%	100%	1
	Meander	100%	79%	1
GC (SS + PCA)	Spiral	89%	64%	0.94
	Circle	100%	57%	1
	Meander	100%	79%	0.96
GCO	Spiral	100%	71%	0.96
	Circle	100%	100%	1
	Meander	100%	71%	1
GCO (SS + PCA)	Spiral	100%	64%	0.91
	Circle	100%	43%	1
	Meander	100%	61%	0.93
G = Grayscale; C = CLAHE; O = Otsu; SS = Standard Scaler				

The model testing results indicate that the model trained using original images demonstrated the highest ability to recognize unseen data compared to the other models, with an average accuracy of 94%. This was followed by model trained on grayscale images (93.67%), grayscale and CLAHE (92.33%) and grayscale with CLAHE and Otsu (90.67%).

The superior performance of the original images may be attributed to the preservation of complex RGB information, which provides more intensity and texture variations for feature extraction. Grayscale conversion simplifies the image by reducing the RGB channels, potentially removing subtle color cues that may be informative for CNN, resulting in a slight decrease in accuracy. Although CLAHE enhances contrast and improve stroke visibility, it may also amplify regions that might introduce additional noise that affect feature consistency, which explains the drop in accuracy. Furthermore, while Otsu thresholding simplifies segmentation, it may remove textural and gradient information that could be important for deep learning, leading to a decline in accuracy.

The results also suggest that, in the proposed system, the inclusion of StandardScaler and PCA did not improve the performance, as both consistently led to lower accuracy. This might be because the CNN-extracted features are already compact and discriminative, and further dimensionality reduction through PCA could remove useful information. Therefore, both normalization and feature reduction may be unnecessary.

Additionally, Table 5 displays complimentary metrics such as specificity, sensitivity and ROC-AUC obtained from each models trained across three handwriting tasks. In general, each model achieved consistently high specificity, indicating a strong ability to correctly identify healthy subjects. However, greater variations were observed in sensitity, suggesting that the detection of Parkinson’s disease case were more challenging. Similar to results observed in Table 4, models trained without using StandardScaler and PCA consistently produced higher sensitivity and ROC-AUC values. Overall, model trained using grayscale and original images provided the most stable and balanced performance across tasks.

The comparison of the result of proposed system and previous studies is summarized in Table 6.

TABLE VI
COMPARISON TO OTHER STUDIES

Study	Dataset	Method	Accuracy
Proposed System	NewHandPD	VGG-16 + Random Forest	100% (Best) / 94%*
Gazda et al. [14]	PaHaw + NewHandPD	VGG-16, SqueezeNet	94.7% (Best)
Ranjan et al. [13]	Unspecified	HOG + Random Forest	86.67% (Best)
Huang et al. [31]	NIATS Handwriting Dataset	Various methods, VGG-19 (Best)	96,67% (Best)
*Average over all tasks			

Compared with previous studies, the proposed hybrid CNN-Random Forest approach demonstrates competitive performance. Using the NewHandPD dataset, the proposed system achieved an average accuracy of 94%. This performance is comparable with results reported by Gazda et al. using end-to-end Deep Learning with VGG-16 and SqueezeNet (94.7%), and received higher accuracy than Ranjan et al. using HOG and Random Forest (86.67%). The proposed system has also achieved a result close to the one reported by Huang et al. using VGG-19 (96.67%). However, since these studies employed different datasets and experimental settings, the comparison should never be indicated as direct benchmark. Nevertheless, the results suggest that the proposed hybrid system is capable of achieving performance on par with the other approaches.

B. Deployment

The best performing model selected from the training process will be deployed to a cloud based environment and later be accessed through an API which makes it accessible from the Parkinvent mobile application to support Parkinson’s disease screening process. Representative images of the Parkinvent app are presented in Figure 9 and 10.

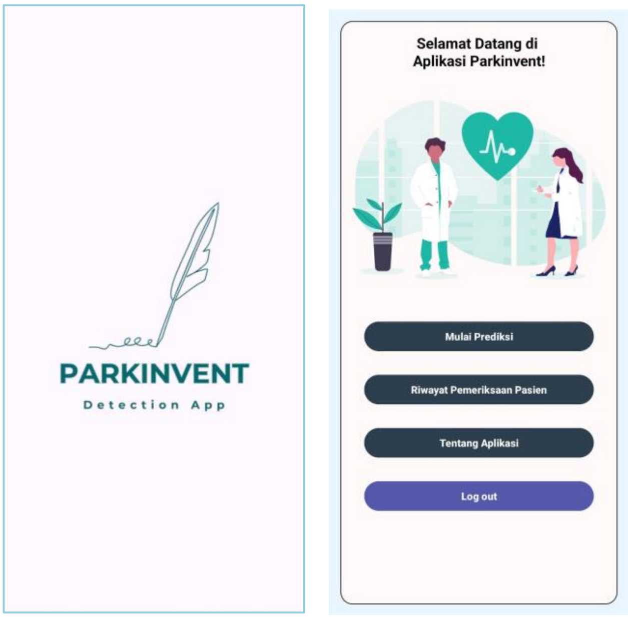


Figure 9. *Parkinvent App's* Starting Screen

Figure 10. *Parkinvent App's* Prediction Result Interface

Figure 9 present splash screen and main menu of the Parkinvent application, while Figure 10 illustrates an example of a prediction result received after a patient submits all the handwriting images. In this example, one of the three handwriting tasks is classified as Parkinson's disease by the model. Consequently, the application displays a recommendation prompting the user to immediately consult a medical professional for further evalation, in order to determine whether its an early indication of Parkinson's disease or a false positive.

C. Study Result Evaluation

The Parkinvent application will go through usability testing through USE questionnaire consisting of 12 questions within 4 different criterias. The usability testing was conducted at a government-owned hospital in Indonesia. Each participant was allowed to test the application and complete the questionnaire after finishing the app testing process. Each question inside the questionnaire will be answered by Likert Scale as shown in Table 7.

TABLE VII
LIKERT SCALE

Description	Score	Percentage (%)
Strongly Agree	5	80-100
Agree	4	60-79
Neutral	3	40-59
Disagree	2	20-39
Strongly Disagree	1	0-19

After the questionnaire responses were collected, the scores will be calculated using equation (1) and (2) [32].

$$Total\ Score = \sum (Tn * Pn) \quad (1)$$

Where *Tn* denotes the frequency of respondents selecting response category n and *Pn* represents the likert score assigned to that category.

$$Score\ Index = \frac{Total\ Score}{Highest\ Score * Total\ Respondent} * 100 \quad (2)$$

The questions and the result of the questionnaire are presented in Table 8.

TABLE VIII
QUESTIONNAIRE RESULT

No	Question	Percentage (%)	Description
1	The software can effectively be used to predict early Parkinson's disease	89.09	Strongly Agree
2	The software makes it easier for medical personel to perform early Parkinson's prediction as supporting information for specialists	90.90	Strongly Agree
3	The software help saves time for early Parkinson's detection	90.9	Strongly Agree
4	I can use this app without any special training	90.9	Strongly Agree
5	This app is user friendly and easy to use	90.9	Strongly Agree
6	If an error occurs, the app shows messages that are easy to understand	91.81	Strongly Agree
7	I can understand the flow of the app easily	92.72	Strongly Agree
8	Steps of using the app can be remembered easily	90	Strongly Agree
9	I can become proficient in using this app quickly	88.18	Strongly Agree

10	I am satisfied on features and performance given by this app	92.72	Strongly Agree
11	The selection of color, icon and image are appropriate and able to give a clear image to users	91.81	Strongly Agree
12	As a whole, I am satisifed with this application	91.81	Strongly Agree

Based on the questionnaire results, the study output achieves an average satisfaction percentage of 90.97%, which falls within the strongly agree category.

Overall, these results suggest that the proposed system has successfully met its intended objective by providing an application capable to assist medical personnels detect early signs of Parkinson's disease. Nevertheless, this study has its own limitations and further research may be conducted in the future to develop a more robust and efficient system for Parkinson's disease detection.

IV. CONCLUSION

Based on the developed software and prediction model combined with the results of the testings done in this study, a few conclusions can be concluded as such: the application developed in this study demonstrates the potential to assist medical personals in early detection of Parkinson's disease. The prediction generated by the application may serve as additional information for specialists during diagnostic process. Among all kinds of preprocessing performed and evaluated in this study, original image without preprocessing and normalization performed the best with an average accuracy of 94%, followed by models trained with grayscale (93,67%), grayscale with CLAHE (92.33%) and grayscale with CLAHE and Otsu thresholding (90.67%). The application of PCA and StandardScaler did not contribute positively, as their inclusion consistently resulted in reduced accuracy across models. The result of questionnaire conducted through USE questionnaire yielded an average result of 90.97% which falls into a Strongly Agree category. Indicating a high level of user satisfaction with the developed application. This study was implemented using the proposed system on NewHandPD dataset. Future improvements aimed at addresing greater variations in handwriting styles and accounting comorbid conditions that may alter writing patterns are expected to enhance the system's robustness and generability even further.

REFERENCES

[1] D. M. Wilson, M. R. Cookson, L. Van Den Bosch, H. Zetterberg, D. M. Holtzman, and I. Dewachter, "Hallmarks of neurodegenerative diseases," Cell, vol. 186, no. 4, pp. 693–714, Feb. 2023, doi: 10.1016/j.cell.2022.12.032.

[2] I. Karlina et al., "Gambaran penyakit neurodegeneratif: Huntington, Alzheimer, dan Parkinson: Sebuah tinjauan," Jurnal Biomedika dan Kesehatan, vol. 7, no. 1, pp. 113–123, Mar. 2024, doi: 10.18051/JBiomedKes.2024.v7.113-123.

[3] World Health Organization, Parkinson disease: A public health approach. Technical brief. World Health Organization, 2022. [Online]. Available: https://books.google.co.id/books?id=CngOEQAAQBAJ

[4] N. Raisa, A. F. Insanitaqwa, and M. Rahayu, "The depiction of general physician's knowledge level of Parkinson's disease in Indonesia," MNJ (Malang Neurology Journal), vol. 9, no. 2, pp. 129–133, Jul. 2023, doi: 10.21776/ub.mnj.2023.009.02.10.

[5] R. B. Postuma and D. Berg, "Advances in markers of prodromal Parkinson disease," Nat. Rev. Neurol., vol. 12, no. 11, pp. 622–634, Nov. 2016, doi: 10.1038/nrneurol.2016.152.

[6] J. Zhang, "Mining imaging and clinical data with machine learning approaches for the diagnosis and early detection of Parkinson's disease," NPJ Parkinsons Dis., vol. 8, no. 1, p. 13, Jan. 2022, doi: 10.1038/s41531-021-00266-8.

[7] Z. Wang et al., "A minimally invasive biomarker for sensitive and accurate diagnosis of Parkinson's disease," Acta Neuropathol. Commun., vol. 12, no. 1, p. 167, Oct. 2024, doi: 10.1186/s40478-024-01873-1.

[8] S. Ali, A. Hashmi, A. Hamza, U. Hayat, and H. Younis, "Dynamic and static handwriting assessment in Parkinson's disease: A synergistic approach with C-Bi-GRU and VGG19," Journal of Computing Theories and Applications, vol. 1, no. 2, pp. 151–162, Dec. 2023, doi: 10.33633/jcta. v1i2.9469.

[9] E. V. Altay and B. Alatas, "Association analysis of Parkinson disease with vocal change characteristics using multi-objective metaheuristic optimization," Med. Hypotheses, vol. 141, p. 109722, Aug. 2020, doi: 10.1016/j.mehy.2020.109722.

[10] M. Diaz, M. Moetesum, I. Siddiqi, and G. Vessio, "Sequence-based dynamic handwriting analysis for Parkinson's disease detection with one-dimensional convolutions and BiGRUs," Expert Syst. Appl., vol. 168, p. 114405, Apr. 2021, doi: 10.1016/j.eswa.2020.114405.

[11] A. Ammour, I. Aouraghe, G. Khaissidi, M. Mrabti, G. Aboulem, and F. Belahsen, "A new semi-supervised approach for characterizing the Arabic on-line handwriting of Parkinson's disease patients," Comput. Methods Programs Biomed., vol. 183, p. 104979, Jan. 2020, doi: 10.1016/j.cmpb.2019.07.007.

[12] J. Mei, C. Desrosiers, and J. Frasnelli, "Machine learning for the diagnosis of Parkinson's disease: A review of literature," Front. Aging Neurosci., vol. 13, May 2021, doi: 10.3389/fnagi.2021.633752.

[13] N. M. Ranjan, G. Mate, and M. Bembde, "Detection of Parkinson's disease using machine learning algorithms and handwriting analysis," Journal of Data Mining and Management, vol. 8, no. 1, pp. 21–29, Mar. 2023, doi: 10.46610/JoDMM.2023.v08i01.004.

[14] M. Gazda, M. Hires, and P. Drotar, "Multiple-fine-tuned convolutional neural networks for Parkinson's disease diagnosis from offline handwriting," IEEE Trans. Syst. Man Cybern. Syst., vol. 52, no. 1, pp. 78–89, Jan. 2022, doi: 10.1109/TSMC.2020.3048892.

[15] A. Das, H. S. Das, A. Neog, R. B. Bharat, A. Choudhury, and M. Swargiary, "Detection of Parkinson's disease from hand-drawn images using machine learning algorithms," in Soft Computing and Signal Processing, V. K. and W. J. and R. K. T. V Reddy V. Sivakumar and Prasad, Ed., Singapore: Springer Singapore, 2021, pp. 241–252.

[16] C. R. Pereira, S. A. T. 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, Oct. 2016, pp. 340–346. doi: 10.1109/SIBGRAPI.2016.054.

[17] K. Kusnadi and D. A. S. P. Atmaja, "Implementation of grayscale image transformation and histogram equalization methods in digital image processing," Krisnadana Journal, vol. 4, no. 2, pp. 111–121, 2025.

[18] M. Hayati et al., "Impact of CLAHE-based image enhancement for diabetic retinopathy classification through deep learning," Procedia Comput. Sci., vol. 216, pp. 57–66, 2023, doi: 10.1016/j.procs.2022.12.111.

[19] A. Riadi and R. Sulaehani, "Analisis implementasi preprocessing dengan Otsu-Gaussian pada pengenalan wajah," ILKOM Jurnal Ilmiah, vol. 11, no. 3, pp. 200–205, Dec. 2019, doi: 10.33096/ilkom.v11i3.457.200-205.

[20] V. Jackins, S. Vimal, M. Kaliappan, and M. Y. Lee, “AI-based smart prediction of clinical disease using random forest classifier and Naive Bayes,” J. Supercomput., vol. 77, no. 5, pp. 5198–5219, May 2021, doi: 10.1007/s11227-020-03481-x.

[21] A. Parmar, R. Katariya, and V. Patel, “A review on random forest: An ensemble classifier,” International Conference on Intelligent Data Communication Technologies and Internet of Things, pp. 758–763, 2019, doi: 10.1007/978-3-030-03146-6_86.

[22] T. F. Basar, D. E. Ratnawati, and I. Arwani, “Analisis sentimen pengguna Twitter terhadap pembayaran cashless menggunakan ShopeePay dengan algoritma random forest,” Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer, vol. 6, no. 3, pp. 1426–1433, Feb. 2022, [Online]. Available: https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/10830

[23] K. Kamal and H. Ez-Zahraouy, “A comparison between the VGG16, VGG19 and ResNet50 architecture frameworks for classification of normal and CLAHE processed medical images,” Apr. 28, 2023. doi: 10.21203/rs.3.rs-2863523/v1.

[24] S. Tammina, “Transfer learning using VGG-16 with deep convolutional neural network for classifying images,” International Journal of Scientific and Research Publications (IJSRP), vol. 9, no. 10, p. p9420, Oct. 2019, doi: 10.29322/IJSRP.9.10.2019.p9420.

[25] A. P. Siregar, D. P. Purba, J. P. Pasaribu, and K. R. Bakara, “Implementasi algoritma random forest dalam klasifikasi diagnosis penyakit stroke,” Jurnal Penelitian Rumpun Ilmu Teknik, vol. 2, no. 4, pp. 155–164, Nov. 2023, doi: 10.55606/juprit.v2i4.3039.

[26] D. Theckedath and R. R. Sedamkar, “Detecting affect states using VGG16, ResNet50 and SE-ResNet50 networks,” SN Comput. Sci., vol. 1, no. 2, p. 79, Mar. 2020, doi: 10.1007/s42979-020-0114-9.

[27] E. Bolin and W. Lam, “A review of sensitivity, specificity, and likelihood ratios: Evaluating the utility of the electrocardiogram as a screening tool in hypertrophic cardiomyopathy,” *Congenit. Heart Dis.*, vol. 8, no. 5, pp. 406–410, Sep. 2013, doi: 10.1111/chd.12083.

[28] E. Richardson, R. Trevizani, J. A. Greenbaum, H. Carter, M. Nielsen, and B. Peters, “The receiver operating characteristic curve accurately assesses imbalanced datasets,” *Patterns*, vol. 5, no. 6, Jun. 2024, doi: 10.1016/j.patter.2024.100994.

[29] K. Kristiawan and A. Widjaja, “Perbandingan algoritma machine learning dalam menilai sebuah lokasi toko ritel,” *Jurnal Teknik Infor matika dan Sistem Informasi*, vol. 7, no. 1, pp. 35–46, Apr. 2021, doi: 10.28932/jutisi.v7i1.3182.

[30] W. A. Kusuma, V. Noviasari, and G. I. Marthasari, “Analisis usability dalam user experience pada sistem KRS online UMM menggunakan USE questionnaire,” Jurnal Nasional Teknik Elektro dan Teknologi Informasi, vol. 5, no. 4, pp. 294–301, Nov. 2016.

[31] Y. Huang, K. Chaturvedi, A.-A. Nayan, M. H. Hesamian, A. Braytee, and M. Prasad, “Early Parkinson’s disease diagnosis through hand-drawn spiral and wave analysis using deep learning techniques,” *Information*, vol. 15, no. 4, p. 220, Apr. 2024, doi: 10.3390/info15040220.

[32] S. Sumartini, K. S. Harahap, and S. Sthevany, “Kajian Pengendalian Mutu Produk Tuna Loin Precooked Frozen Menggunakan Metode Skala Likert di Perusahaan Pembekuan Tuna,” Aurelia Journal, vol. 2, no. 1, p. 29, Nov. 2020, doi: 10.15578/aj.v2i1.9392.