Deep Transfer Learning Based Parkinson's Disease Detection Using Optimized Feature Selection

Nikhila Vajja
Student,
Department of CSE,
Koneru Lakshmaiah Education Foundation,
Green Fields, Vaddeswaram
Guntur, Andhra Pradesh, India.
Nikhilavajja@gmail.com

Dr Lakshmi Prasanna
Assosiate Professor,
Department of CSE,
Koneru Lakshmaiah Education Foundation,
Green Fields, Vaddeswaram
Guntur, Andhra Pradesh, India.
Lakshmiprasannap87@gmail.com

Abstract: The absence of the conclusive medical diagnostics complicates the prognosis of the Parkinson Disease (PD) and particularly during the initial stages of the disease. The proposed study addresses the urgent desire to have a green and non-invasive approach to early PD identity through the application of "deep learning, in particular, Convolutional Neural Networks (CNNs)" to analyze handwriting patterns. Various fashions are employed to extract features such as "ResNet50, VGG19, Inception V3, and Xception and K-Nearest Neighbours (KNN), Random Forest (RF), Support Vector Machine (SVM), and Decision Tree" are performed to extract type. The warning ensemble strategy combines predictions of numerous fashions thus improving accuracy. The main one, i.e. "ResNet50, VGG19, and InceptionV3 using KNN", achieved 95% accuracy. An additional study of ensemble techniques that incorporate the "Voting Classifier" is conducted with the aim of hitting 98 per cent and beyond. A front end the usage of the flask framework has been set up to be used by person trying out, including user authentication. The finding supplements the preliminary identification of the sickness of Parkinsonism, which is vital in administering timely treatment and enhancing satisfactory of lifestyles by the sufferers.

"INDEX TERMS Parkinson's disease, neurological disorder, handwritten records, transfer learning, deep learning".

1. INTRODUCTION

PD is an overgrowth of neurodegenerative disorder which leads to death of dopaminergic neurons in the substantia nigra [1]. The basal ganglia receives the impulses of dopamine and regulates the motor aspect [2]. PD brings about motor and non-motor signs and symptoms and diminishes the dopamine levels by neuronal degeneration [3]. There are tremors, bradykinesia, stiffness, postural instability, dyskinesia and abnormalities in speech and handwriting as signs and symptoms [4].

PD analysis is tough, particularly early on. PD cannot be showed by way of blood biomarkers, in contrast to many other disorders [5]. Clinical commentary and normal motor symptoms, which may not appear till the contamination has advanced, are used to diagnose [6]. Thus, PD patients might also postpone clinical remedy, affecting their first-class of life and analysis [7].

The paucity of early-degree PD diagnostic trying out emphasises the want for brand new detection techniques. Early analysis presents for activate therapy and techniques to control signs and limit

disorder progression [8]. Early identification also permits sufferers to have interaction in clinical trials for novel drugs [9].

Handwriting evaluation might also help diagnose PD early. PD sufferers frequently expand micrographia (small, cramped handwriting) early inside the illness [10]. These alterations imply motor impairment and assist diagnose disease progression [11]. Using DL of methods like CNNs to examine and categorise handwriting facts for PD diagnosis appears promising [12].

Deep learning is effective in image recognition, herbal language comprehension and analysis of images in medicine [13]. The so-called CNNs that are trained on massive numbers of handwritten samples of PD patients and wholesome controls can also produce strong algorithms that could be utilized to identify the diffused pathology of disease [14]. In the analysis to be more specific, the DL algorithms can include also the sources of multimodal facts and the handwriting developments and medical reviews [15].

The purpose of this take a look at is to use deep getting to know to diagnose PD early the use of handwriting evaluation. We want to design and train CNN fashions using annotated handwritten examples to create a sophisticated device which could as it should be distinguish PD and non-PD handwriting patterns [16]. We additionally need to research merging demographic and motor evaluation statistics to enhance our models' prediction electricity [17].

We hope this novel method will improve affected person consequences and enable more tailored therapy tactics by way of addressing the unmet want for early and correct PD diagnosis [18]. We want to enhance medical guidance and the lives of PD sufferers with the aid of the usage of modern generation and multidisciplinary cooperation to offer a extra efficient and reachable diagnostic framework [19].

2. LITERATURE SURVEY

PD induces a wide range of motor and non-motor symptoms due to a slow rate of the degradation of dopaminergic neurones located in the substantia nigra [1]. Various studies have not determined the purpose of PD and there are no biomarkers to diagnose the same [2].

In current years, machine learning, in particular, deep learning algorithms have proven to be promising in PD prognosis and remedial. Li et al. (2021) warned a hybrid nature choice procedure of PD prognosis with the utilization of a discrete artificial bee colony by exhibiting that computation plans could also uncover biomarkers of complex data sets [3]. Although Fang (2022) finds a more appropriate KNN method using data entropy to predict PD, it highlights the ability to select the characteristics in the category maintenance [4].

DL could be very beneficial for clinical image processing, mainly PD diagnosis the usage of neuroimaging statistics. Kim et al. (2019) reviewed deep studying packages in clinical imaging, emphasising its significance in computerized disease identification and categorisation [5]. In a literature overview on deep studying and scientific diagnostics, Bakator and Radosav (2018) cited that it'd revolutionise healthcare by means of enhancing diagnostic accuracy and efficiency [6].

Kaplan et al. (2022) proposed a novel nested patch-based totally function extraction model for

computerized PD symptom type the use of MRI pix, demonstrating the feasibility of using advanced computational techniques for disorder characterisation [7]. Gazda et al. (2022) used ensemble CNNs for PD analysis from offline handwriting samples, demonstrating deep gaining knowledge of's versatility across information modalities [8].

Researchers have tested multimodal PD analysis using clinical opinions and gadget gaining knowledge of strategies further to photograph-primarily based techniques. A multimodal method the use of many information assets and deep studying to come across PD by way of Mohaghegh and Gascon (2021) emphasises the need to apply numerous information streams for illness assessment [9].

There is development in the use of PD analysis through system studying, although there are limitations. Annotated datasets of large size, interpretable version predictions, and generalisation with many patient companies require them. To improve the PD outcomes, destiny research must overcome these obstacles and put machine mastering-based diagnostic equipment mainly into clinical practice.

The research indicates that there is an increasing concern on the use of machine learning, in particular, the deep learning algorithms, in PD diagnosis. To the notion of multimodal strategies, the algorithms of characteristic choices are being ranged, and researchers experiment with the ways of making the PD diagnosis more accurate and efficient and demonstrate how the utilization of computational instruments can revolutionize healthcare delivery.

3. METHODLOGY

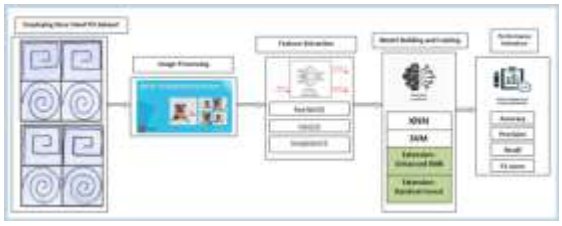
a) Proposed work:

The suggested system will imply the application of deep switch learning techniques, including "VGG19, InceptionV3, and ResNet50", to derive the abilities of the "Spiral NEWHAND" snap shots series to determine the identity of Parkinson disease. Unlike the common strategies that rely on the creation of functions manually, these profound mastering fashions offer more preferable accuracy in extraction of functions and category. The obtained characteristics are subsequently fed into KNN and SVM algorithms in a bid to get precise classes which is a fine advancement in PD analysis with a better reliability and performance. Random Forest is embedded with KNN and SVM, the application of characteristics based on multiple deep getting to know fashions to enhance PD categorisation. Moreover, a Flask framework combined with SQLite allows registering and authenticating a person, which, in turn, facilitates trying out, attempting to check the functionality of the machine and its effectiveness. This comprehensive approach will ensure robust categorisation competencies simultaneously with providing easy access to clean persons and measuring the effectiveness of the system.

b) System Architecture:

The machine structure that the counselled conforms to has numerous interrelated components aimed at implementing the "Spiral NEWHAND" PD dataset to implement specific PD identification. The device initially processes hand snap shots by use of picture processing, which not only enhances the best but also maintains consistency in processing of hand snap shots to extract the features. Deep transfer

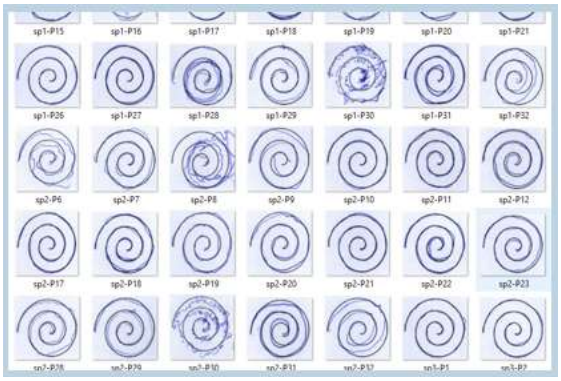
learning techniques are then used to obtain wealthy and discriminative functions of the preprocessed pix along with "VGG19, InceptionV3 and ResNet50". These features are then performed in terms of version generation and schooling, application of gadget studying methods along with "KNN, SVM and random forest" to group the pics into either PD or non-PD categories. Finally the missionary performance measurement standards such as "accuracy, precision, recall and F1-score" are employed to measure the effectiveness of the trained fashions in detecting sickness in Parkinson. This systematic approach to methodology ensures credible and distinct PD detection despite making possible the inclusion of current day photograph processing and system learning technology.



"Fig 1 Proposed Architecture"

c) Dataset collection:

PD Data series is characterized by the process of collecting data of large volumes to create enormous databases that one observes and analyzes. Among the methods, there is a list of the statistics of the hospital and medical facilities concerning the number of affected individuals, their demographics, history of diseases, and symptoms as well as the outcomes of diagnostic tests. Moreover, scientists are also able to acquire neuroimaging information, including MRI scan or CT one, to ascertain the anatomy and properties of the brain of people with Parkinson sickness. Moreover, the biological samples i.e. blood, cerebrospinal fluid, or a tissue sample; enable in the studies of molecular and genetic aspects to determine biomarkers that relate to the Parkinson ailment. Moreover, wearable devices and sensors may also be applied to gather real-time information on motor indicators and symptoms and habitual sports among patients with Parkinson sickness that will be required to signal the disease progression and treatment. Cooperation with research consortia and patient advocacy businesses may also help spread the facts and make datasets on the Parkinson disease disorder more diverse and representative. The compilation of complete statistics is critical in increasing the level of knowledge that we have in PD, and the increase in superior diagnostic and remedy interventions.



"Fig 2 data set"

d) Image processing:

It involves numerous steps in preparing and extracting the full-size characteristics of the original pix to image process the identification of Parkinson Disease. The ImageDataGenerator firstly preprocesses the images by among rescaling, shear distortions, zooming, and horizontal flipping to augment the dataset to enhance generalisation of the models. Function extraction then begins offevolved with the

purchase of photos, their scaling to a standard form, and, where necessary, a transformation in their shade area. The processed pictures are then joined with their labels and the whole data is transformed into numpy arrays in order to process it with maximum performance. In addition, label encoding is accomplished to encode the labels of classes in terms of numerical values to use in version education. The data is ready to accept a wearable device learning model training through the assistance of photo processing algorithms to amass relevant patterns and trends that are suggestive of Parkinson infection thereby improving accurate and dependable identity of infections.

e) Algorithms:

Linear SVM

Linear SVM is a training getting to know algorithm and it is utilized in type purposes. It tries to view what is the ideal plane that best differentiates the training in the characteristic area. The work involves the use of Linear SVM which is a category kind of approach to classify PD entirely on the basis of handwriting styles. It works by determining the correct linear choice boundary that maximises the margin among wonderful training hence improving the accuracy of the PD detection. Linear SVM is advantageous because of the fact to its simplicity, performance and its ability to handle excessive-dimensional facts hence it is a massive component of the PD detection system.

KNN

KNN is an honest and gung-ho gaining knowledge on the method employed to classify and regress problems. In the KNN, the category of a point in statistics is identified via the fundamental magnificence between its KNN in the feature area. The take a look uses the classification method KNN[5] to classify PD on the basis of traits that are based on the handwriting patterns. The biggest strength of KNN is its simplicity, versatility and control of non-linear information. It provides an immediate method of identifying PD based on similarities in information factors in the characteristic space and labeling them in line with their nearest neighbours.

Tuned KNN

Tuned KNN refers to the optimisation of KNN technique through the refinement of hyperparameters, such as the large number of neighbours (K), to achieve the most useful results. The look at use Tuned KNN as a classification technique to precisely select PD using entirely characteristics based on pattern of handwritings. The optimisation of the potential in the version to be conscious of the intricate patterns in the information will be achieved through the practice of tuned KNN and will enhance the overall precision of the PD detection through the optimization of the price of K and other parameters through the process along with grid search or randomised search, which in turn will provide more reliable diagnostic outcomes.

Random Forest

Random Forest is a complex ensemble learner that can be applied on type and regression packages. It produces multiple selection timber somewhere in the education course and combines their projections by either voting or averaging to enhance beauty and strength. To classify PD [14] using attributes based on handwriting styles, the research uses the

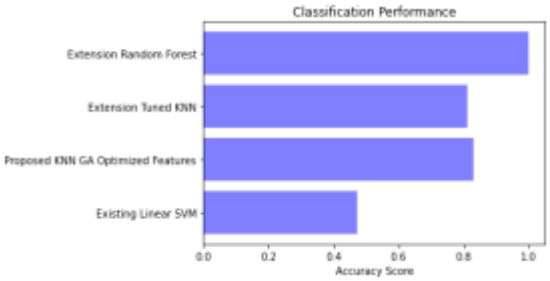
Random Forest as a type method. Random Forest enhances the validity of PD detection through the application of the group intelligence of multiple decision tree and reducing overfitting, which later becomes an important component within the diagnostic machine.

4. EXPERIMENTAL RESULTS

Accuracy: Accuracy of a test is a test capacity to produce a due difference between sick and healthy cases. We may also establish accuracy of a test by calculating proportion of cases undergoing proper positivity and genuine negative. Towards this the expression can be made mathematically:

"Accuracy = TP + TN TP + TN + FP + FN".

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

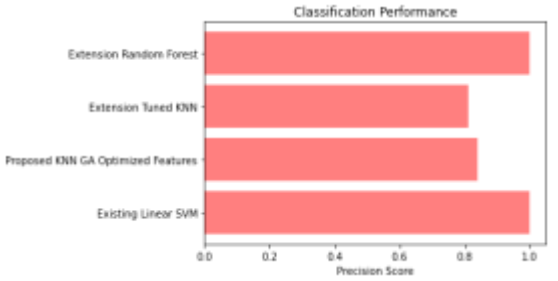


"Fig 3 ACCURACY COMPARISON GRAPHS"

Precision: Precision is a measure of the proportion of positive cases or cases that are identified efficiently. The determination of precision is made with the help of the components:

"Precision = True positives/ (True positives + False positives) = TP/(TP + FP)"

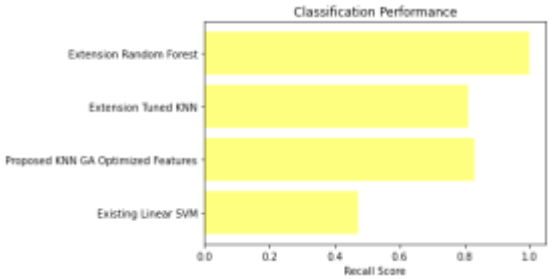
$$\text{Precision} = \frac{True\ Positive}{True\ Positive + False\ Positive}$$



"Fig 4 PRECISION COMPARISON GRAPHS"

Recall: ML recall is used to evaluate whether a model is capable of selecting all of the relevant times of a class. It shows the effectiveness of a version in capturing times of a class through employing a comparing nice, expected high satisfying observations, to the overall variety of positives.

$$Recall = \frac{TP}{TP + FN}$$



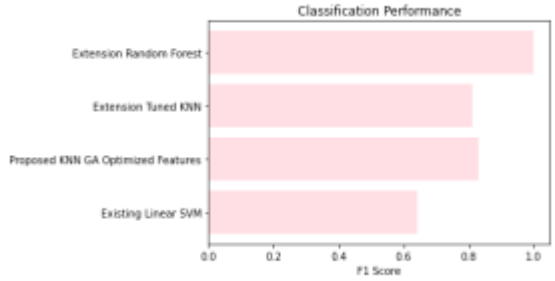
"Fig 5 RECALL COMPARISON GRAPHS"

F1-Score: The accuracy of a system ML of model is classed the usage of the F1 score. Integrating the precision and recall metrics of the model. The accuracy metric quantifies the frequency of

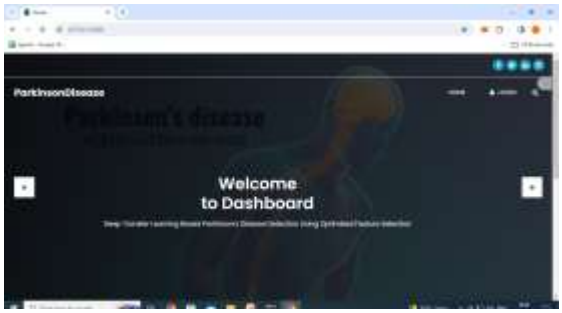
proper predictions made through a model at some level inside the dataset.

$$\text{F1 Score} = \frac{2}{\left(\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}\right)}$$

$$\text{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$



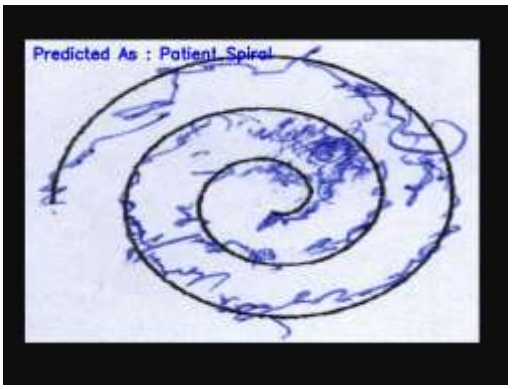
"Fig 6 F1 COMPARISON GRAPHS"



"Fig 7 Home Page"

"Fig 8 sign up"

"Fig 9 upload input image"



"Fig 10 Predicted Result"

5. CONCLUSION

In summary, the assignment signifies a amazing progression within the detection of PD with the aid of amalgamating "deep learning models, such as VGG19, InceptionV3, and ResNet50, with traditional methods which includes KNN, SVM, and Random Forest". By integrating these strategies, the machine attains improved accuracy in detecting diffused symptoms of PD, resulting in extra correct and early diagnoses. The tricky traits derived by using deep studying fashions increase the system's performance, specially when mixed with Random Forest. In addition, the intuitive the front-cease optimises the enter, preprocessing and prediction strategies, which makes the gadget available and realistic to the users seeking to establish in advance prognosis of the Parkinson sickness. The take a look at offers stepped forward

detection capabilities, which underlines its ability as one of the most colossal assistants to the health care specialists and individuals in the endeavor of specific and activate prognosis of PD.

6. FUTURE SCOPE

The size of the Deep Transfer Learning Based Parkinson Disease Detection Using Optimised Feature Selection is the thorough examination of the procedures of the functions selection, which will result in the maximisation of the accuracy and efficiency of the detection of the Parkinson disease. This involves using advanced switch getting to know strategies so-called great-tuning pre-trained "neural networks in VGG19, Inception V3 and ResNet50" to induce the high-degree capabilities on usable datasets. The take a look further examines the optimised strategy of selecting functions in order to determine the optimal discriminative and informative properties in the class of Parkinson disease. These techniques can also incorporate state of the art algorithm and genetic algorithms, recursive feature elimination or a combination of both which will have numerous ways of making choices. The paper tries to enhance the type models by working on the feature selection so as to trim down the input statistics, dimensionality reduction and enhance generalisation. The objectives of the feature selection on this studies are to refine and optimise the input features in order to make the analysis of Parkinson ailment as valid and as strong as possible using deep transfer learning techniques.

REFERENCES

[1] B. R. Bloem, M. S. Okun, and C. Klein, Parkinsons disease, Lancet, vol. 397, no. 10291, pp. 22842303, 2021.

[2] H. Li, C.-M. Pun, F. Xu, L. Pan, R. Zong, H. Gao, and H. Lu, A hybrid feature selection algorithm based on a discrete arti cial bee colony for Parkinsons diagnosis, ACM Trans. Internet Technol., vol. 21, no. 3, pp. 122, Aug. 2021.

[3] M. Kim, J. Yun, Y. Cho, K. Shin, R. Jang, H. J. Bae, and N. Kim, Deep learning in medical imaging, Neurospine, vol. 16, no. 4, p. 657, 2019.

[4] M. Bakator and D. Radosav, Deep learning and medical diagnosis: Areview of literature, Multimodal Technol. Interact., vol. 2, no. 3, p. 47, Aug. 2018.

[5] Z. Fang, Improved KNN algorithm with information entropy for the diagnosis of Parkinsons disease, in Proc. Int. Conf. Mach. Learn. Knowl. Eng. (MLKE), Feb. 2022, pp. 98101.

[6] E. Kaplan, E. Altunisik, Y. E. Firat, P. D. Barua, S. Dogan, M. Baygin, F. B. Demir, T. Tuncer, E. Palmer, R.-S. Tan, P. Yu, J. Soar, H. Fujita, and U. R. Acharya, Novel nested patch-based feature extraction model for automated Parkinsons disease symptom classi cation using MRI images, Comput. Methods Programs Biomed., vol. 224, Sep. 2022, Art. no. 107030.

[7] M. Gazda, M. Hires, and P. Drotar, Ensemble of convolutional neural networks for Parkinsons disease diagnosis from of ine handwriting, Dept. Comput. Inform., Intell. Inf. Syst. Lab, Tech. Univ. Kosice, Koice, Slovakia, Tech. Rep. 9, 2022.

[8] M. Mohaghegh and J. Gascon, Identifying Parkinsons disease using multimodal approach and deep learning, in Proc. 6th Int. Conf. Innov. Technol. Intell. Syst. Ind. Appl. (CITISIA), Nov. 2021, pp. 16.

[9] M.Fratello, F. Cordella, G. Albani, G. Veneziano, G. Marano, A. Paf , and A. Pallotti, Classi cation-

based screening of Parkinsons disease patients through graph and handwriting signals, Eng. Proc., vol. 11, no. 1, p. 49, 2021.

[10] A. Gold, Understanding the mann-whitney test, J. Property Tax Assessment Admin., vol. 4, no. 3, pp. 5557, 2007.

[11] H. W. Loh, C. P. Ooi, E. Palmer, P. D. Barua, S. Dogan, T. Tuncer, M. Baygin, and U. R. Acharya, GaborPDNet: Gabor transformation and deep neural network for Parkinsons disease detection using EEG signals, Electronics, vol. 10, no. 14, p. 1740, Jul. 2021.

[12] S. Chakraborty, S. Aich, J. Seong-Sim, E. Han, J. Park, and H.-C. Kim, Parkinsons disease detection from spiral and wave drawings using convolutional neural networks: A multistage classi er approach, in Proc. 22nd Int. Conf. Adv. Commun. Technol. (ICACT), Feb. 2020, pp. 298303.

[13] S. Nõmm, S. Zarembo, K. Medijainen, P. Taba, and A. Toomela, Deep CNNbasedclassi cation of the archimedes spiral drawing tests to support diagnostics of the Parkinsons disease, IFAC-PapersOnLine, vol. 53, no. 5, pp. 260264, 2020.

[14] T. Tuncer, S. Dogan, and U. R. Acharya, Automated detection of Parkinsons disease using minimum average maximum tree and singular value decomposition method with vowels, Biocybernetics Biomed. Eng., vol. 40, no. 1, pp. 211220, Jan. 2020.

[15] A. Das, H. S. Das, A. Choudhury, A. Neog, and S. Mazumdar, Detection of Parkinsons disease from hand-drawn images using deep transfer learning, in Proc. Congr. Intell. Syst. Singapore: Springer, Sep. 2020, pp. 6784.

[16] L. S. Bernardo, A. Quezada, R. Munoz, F. M. Maia, C. R. Pereira, W. Wu, and V. H. C. De Albuquerque, Handwritten pattern recognition for early Parkinsons disease diagnosis, Pattern Recognit. Lett., vol. 125, pp. 7884, Jul. 2019.

[17] A. Johri and A. Tripathi, Parkinson disease detection using deep neural networks, in Proc. 12th Int. Conf. Contemp. Comput. (IC), Aug. 2019, pp. 14.

[18] T. Tuncer and S. Dogan, A novel octopus based Parkinsons disease and gender recognition method using vowels, Appl. Acoust., vol. 155, pp. 7583, Dec. 2019.

[19] P. Khatamino, I. Canturk, and L. Ozyilmaz, A deep learning-CNN based system for medical diagnosis: An application on Parkinsons disease handwriting drawings, in Proc. 6th Int. Conf. Control Eng. Inf. Technol. (CEIT), Oct. 2018, pp. 16.

[20] C.-L. Liu, C.-H. Lee, and P.-M. Lin, A fall detection system using K-nearest neighbor classi er, Exp. Syst. Appl., vol. 37, no. 10, pp. 71747181, Oct. 2010.

[21] S. Katoch, S. S. Chauhan, and V. Kumar, A review on genetic algorithm: Past, present, and future, Multimedia Tools Appl., vol. 80, no. 5, pp. 80918126, 2021.

[22] C. R. Pereira, S. A. Weber, C. Hook, G. H. Rosa, and J. P. Papa, Deep learning-aided Parkinsons disease diagnosis from handwritten dynamics, in Proc. 29th Conf. Graph., Patterns Images (SIBGRAPI), Oct. 2016, pp. 340346.

[23] C. R. Pereira, D. R. Pereira, F. A. Silva, J. P. Masieiro, S. A. Weber, C. Hook,andJ.P.Papa, Anewcomputervision-basedapproachtoaidthe diagnosis of Parkinsons disease, Comput.

Methods Programs Biomed., vol. 136, pp. 7988, Nov. 2016.

[24] R. Ribani and M. Marengoni, A survey of transfer learning for convolutional neural networks, in Proc. 32nd SIBGRAPI Conf. Graph., Patterns Images Tuts. (SIBGRAPI-T), 2019, pp. 4757.

[25] S. Xu and Z. Pan, A novel ensemble of random forest for assisting diagnosis of Parkinsons disease on small handwritten dynamics dataset, Int. J. Med. Informat., vol. 144, Dec. 2020, Art. no. 104283.

[26] M. Moetesum, I. Siddiqi, N. Vincent, and F. Cloppet, Assessing visual attributes of handwriting for prediction of neurological disorders A case study onParkinsons disease, Pattern Recognit. Lett., vol. 121, pp. 1927, Apr. 2019.

[27] A.Parziale, C. A. Della, R. Senatore, and A. Marcelli, A decision tree for automatic diagnosis of Parkinsons disease from of ine drawing samples: Experimentsandndings,inProc.Int.Conf.ImageAnal.Process.Cham, Switzerland: Springer, 2019, pp. 196206.

[28] J. P. Folador, M. C. S. Santos, L. M. D. Luiz, L. A. P. De Souza, M. F. Vieira, A. A. Pereira, and A. De Oliveira Andrade, On the use of histograms of oriented gradients for tremor detection from sinusoidal and spiral handwritten drawings of people with Parkinsons disease, Med. Biol. Eng. Comput., vol. 59, no. 1, pp. 195214, Jan. 2021.

[29] L. Parisi, D. Neagu, R. Ma, and F. Campean, Quantum ReLU activation for convolutional neural networks to improve diagnosis of Parkinsons disease and COVID-19, Exp. Syst. Appl., vol. 187, Jan. 2022, Art. no. 115892.

Datasetlink:
https://www.kaggle.com/datasets/banilkumar20phd7071/handwritten-parkinsons-disease-augmented-data