MRI Image Synthesis for the Diagnosis of Parkinson's Disease using Deep Learning

Name: Neeyanth Kopparapu

Abstract

Parkinsons disease (PD) is a neurodegenerative disease that affects an estimated 1% of adults over 65. While the disease itself is not fatal, complications related to PD are rated as the 14th largest cause of death in the United States by the Center of Disease Control and Prevention. In spite of technological advances, Parkinson's diagnosis methods have not changed, and the accuracy of diagnosis has remained at approximately 81% for the past 25 years. These methods include analyzing years of neurological data to determine if the patient has developed the symptoms of Parkinson's, including limb rigidity and tremors, both of which are common side effects of a number of other diseases.

With the rise of automated prediction algorithms paired with the generation of massive amounts of data, the automatic diagnosis of Parkinson's Disease has not caught up to traditional means. This is commonly attributed to the lack of useful data, as most computational systems require a tremendous amount of medical data that isn't readily available as gathering the data can be expensive.

This study presents PDGAN, a tool to aid pathologists and neurologists in the diagnosis of Parkinson's Disease. PDGAN uses a series of neural networks to classify Magnetic Resonance Images (MRI Images). PDGAN employs Generative Adversarial Networks (GANs) to synthetically generate medical images which is used to augment the classification efforts. The pair of Convolutional Neural Networks exhibited an testing accuracy of 91.4% without the augment of new data, and combined the total accuracy was 96.6%, a 16% increase compared to traditional methods at a fraction of the cost and time. PDGAN demonstrates the feasibility of utilizing GANs to generate unseen data for the improvement of classification accuracy in the medical setting.

Introduction

Parkinson's Disease (PD) is the second most prevalent neurodegernative disease, affecting approximately 1% of the population above the age of 65 [31]. Although PD is not a fatal disease, it decreases life expectancy by 16 years [36]. PD affects movement, originally with small tremors, slowness of movement, and shortness of breath, but symptoms worsen as the condition progresses while untreated [5]. Although the exact cause of PD has not been determined, PD is known to kill cells in the substanitia nigra, the area of the brain responsible for the production of dopamine, a neurotransmitter responsible for the control of movement. The loss of dopamine, an early sign of Parkinson's, leads to symptoms including bradykinesia, the slowness of movement [35].

Although there is no specific test for the diagnosis of PD, the current procedure is as follows: (1) a physical exam is conducted, (2) further lab tests are conducted (Blood Work, MRI, Brain Ultrasound), (3) manual analysis of past neurological history, (4) determine if symptoms are present [4]. These non-deterministic tests often take weeks and produce unreliable results. It is also very difficult to predict symptoms until they are already present, and the longer PD goes undiagnosed in a patient, the harder it is for treatment to be effective [36].

Therefore, it is important for patient's future for them to have a quick and accuracy diagnosis. However, due to the lack of a specific diagnosis pipeline for PD, the accuracy of diagnosis has not increased despite the advancements in the fields of medicine, technology, and bioinformatics. The accuracy of diagnosis reported in 1992 from 100 various studies was 82% [15], and the accuracy in 2014 was 83.9% [26]. Although numerous computational vision attempts at diagnosing PD have emerged, including automatically regressing neurological data and classifying patient micrography, efforts haven't reached classification metrics that compare to manual efforts, with accuracies near 67% [25].

Many previous studies attribute their difficulties to two main problems. First, there is a lack of useable data. Complex regression and deep learning models that include millions of parameters require a hundreds or thousands of data points to learn from, and many available datasets contain anywhere from 50 to 200 images, not nearly enough for capable models to

converge. Second, studies determined the hardest diagnosis class for PD is the patients right after showing signs of symptoms. At this stage, symptoms have not fully present, so it is nearly impossible for neurologists or automatic tools to detect the presence of PD [25].

This study presents an approach of the automatic approach to aid the diagnosis of PD through patient Magnetic Resonance Imaging (MRI) scans. MRI scans have historically been used to determine alterations in the Central Nervous System (CNS), some of which might be indicative of PD. MRI scans have been shown to be more effective than other imaging processes in displaying these changes [12], and most diagnosis efforts from scientists include an MRI scan of the brain. The approach uses a system of Fully Convolutional Networks (FCNs) and Convolutional Neural Networks (CNNs) to perform image analysis tasks. It also utilizes Generative Adversarial Networks (GANs) to augment the current dataset with more artificial scans that were used in the classification training.

Materials and Methods

Dataset Description The data used for the system came from the University of Southern California's Labratory of Neruological Imaging (LONI). The Parkinson's Progression Markers Initiative under LONI's Image Data Archive contains MRI scans, genetic data, medical history, and motor assessments from over 1400 Parkinson's and Control patients. The Image Archive contains 921 Control MRIs and 2633 PD MRIs viewed in the Axial form[29].

Magnetic Resonance Imaging (MRI) is an imaging procedure that uses magnets and radio waves to capture detailed images of the brain. Although there are many types of MRI scans, including functional MRIs (fMRI), cardiac MRIs, Magnetic resonance angiography (MRA), and Magnetic resonance venography (MRV), the most common MRI is a anatomical scan of the brain. This anatomical scan displays the shape, volume, and developmental changes in the brain as a three-dimensional image.

The anatomical MRIs obtained from the data archive came in 43 different sizes, ranging from $1024 \times 1024 \times 256$ to $32 \times 24 \times 32$. Due to the large variance in sizes, the data actually used was one that had a high volume of images, but also had a representative split compared to the total dataset. The chosen size was $256 \times 240 \times 176$, as it had a total of 612 images, 146 from the Control group, and 466 from the PD group, representative of the approximate $1:3$

split in data throughout the study. The 86 generated images had 23 MRIs of the control group and 63 MRIs of the PD group, keeping consistent with the same ratio as the previous data as shown in Figure 1.

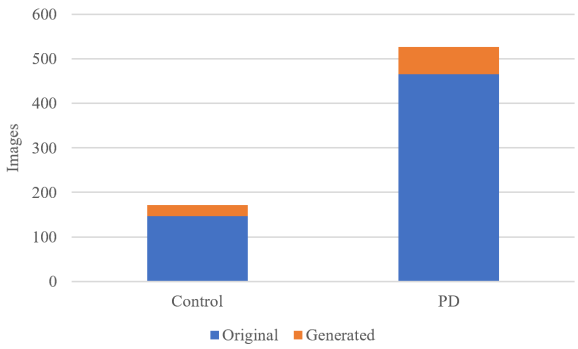

Figure 1: Chart displaying sizes of generated and original dataset belonging to each class.

Software and Hardware The software platform used for the creation of this project was Python. Python was used due to its modularity and compatibility with numerous standard machine learning, plotting, and artificial intelligence frameworks. To perform standard image preprocessing and analysis tasks, the OpenCV and PIL (Python Imaging Library) libraries were used [20].

Both Tensorflow and Keras were used as deep learning frameworks. Keras was used due to its code readability and modularity, and Tensorflow was used for its extensive online documentation, as well as its use for machine learning, mathematical computation, and deep learning [32, 21]. Artificial Neural Networks (ANNs) were used to perform the classification task. ANNs have grown in popularity for outperforming humans in various tasks, including computer analysis tasks [23].

Throughout the study, Convolutional Neural Networks (CNNs) and Generative Adversarial Networks (GANs) were utilized. CNNs are a flavor of ANNs that specialize in image processing. Figure 2 shows the basis of what a Convolutional Network does.

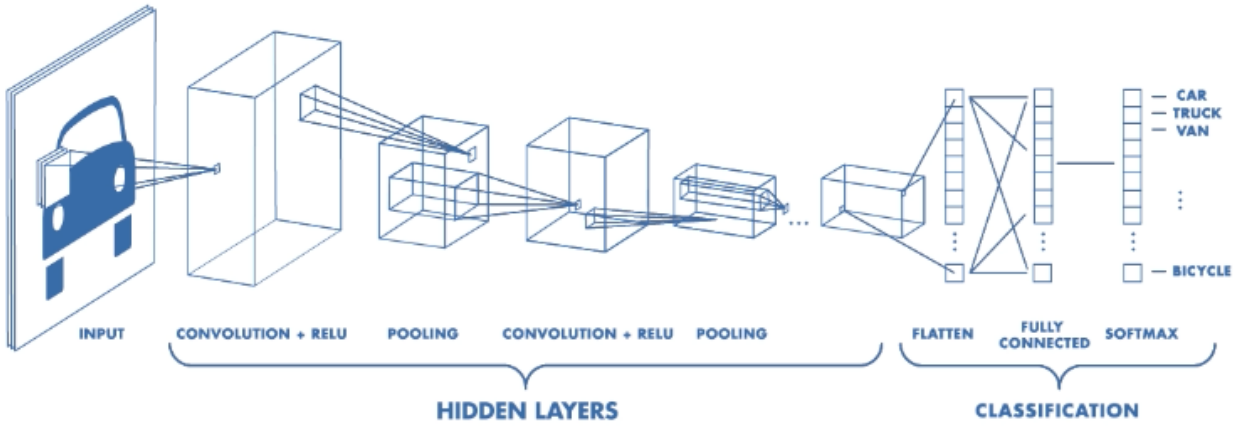

Figure 2: Example Fully Convolutional Network that has Convolution layers [22].

GANs [11], in the class of unsupervised machine learning, are a version of an adversarial network which works to generate images similar to the samples given using a noise vector. As shown in Figure 3, the Generator takes a noise vector which it uses, in addition to its trained weights, to create images. The Discriminator is tasked to determine if the sample given to it is real. The adversarial networks's goals are to make it harder for each other, and in turn improve. GANs are trained over thousands of epochs to converge, but these epochs take less time compared to traditional training epochs [11, 2].

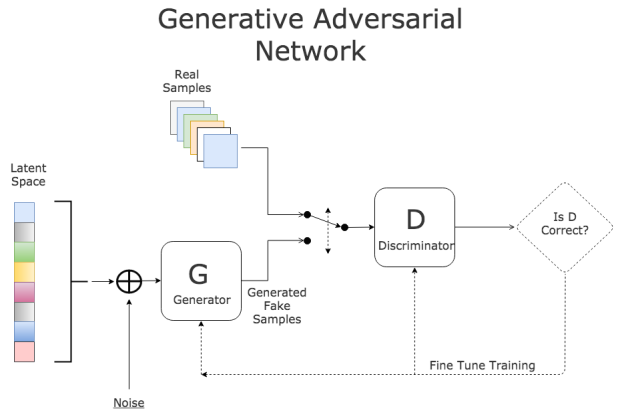

Figure 3: Example GAN data flow (Stack Exchange).

To complete the task, a Tesla K80 GPU was used, which included 11.2 GB of GPU memory, and 4992 CUDA cores. A p2.xlarge AWS EC2 instance was used with an Ubuntu backend and 100 GB extra SWAP space to store data during program calls.

Approach - Preprocessing Minimal preprocessing steps were taken. First, the images were normalized. The images were not subjected to standard augmentation procedures including cyclic pooling and tilting because hemispherical asymmetries in MRI images can reveal important features of the brain anatomy [34]. The dimensions of the image array after preprocessing was $256 \times 240 \times 176$.

Approach - First Iteration The first iteration of the system consisted of the preprocessing steps and the image classification neural network, trained only on the images in LONI's PPMI database. Figure 4 shows the flow of data in the system.

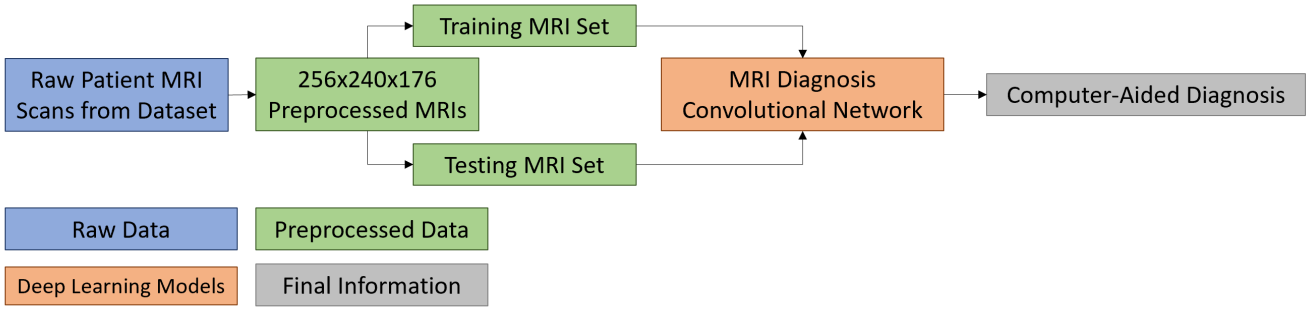

Figure 4: Graphical Description of the flow of data and processing through the system.

There were various models used, in an attempt to maximize accuracy. First, the traditional CNNs, including the VGG-19, GoogLeNet, and Resnet-50, were used due to their high accuracy in standard benchmarks like ImageNet [27, 30, 14]. The VGG model is described in Figure 5, where "E" is the VGG-19 model. Because MRIs are three dimensional, the CNNs were repurposed for 3D classification in Keras.

To further improve the accuracy, the Spearmint [28] package was used. Spearmint, a Bayseian optimization package which changes hyperparameters to minimize loss, was used on the VGG-19 model which had the highest accuracy before using the package. With 5 variables to change the different size of the filter layers, the model's accuracy was increased

slightly to 91.4%. Spearmint was not allowed to fully complete its optimization due to time constraints, as each trial would take upwards of 6 hours with early stopping. Thus, only 18 spearmint programs were run.

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 **LRN**	conv3-64 **conv3-64**	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 **conv3-128**	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 **conv1-256**	conv3-256 conv3-256 **conv3-256**	conv3-256 conv3-256 conv3-256 **conv3-256**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

Figure 5: VGG System Architecture [27]

To improve the accuracy, some problems were highlighted.

1. The size of the dataset is small - with over 400 million parameters, only a few hundred images is not enough to make the model converge.

2. With deep networks like the vanilla GoogLeNet, the *vanishing gradient problem* only allowed small changes to be made progressively, with large changes not having much of an impact on the early layers.

Many steps were taken to combat these problems, all of which were reflected in the second interaction of the model.

Approach - Motivation Wang et al. described one of the first methods of augmenting a dataset during low-shot learning [33]. The hallucinator, G, described is able to take an input (x, y), with a noise vector z and output a "hallucinated" input (x', y) that would

then be used to augment the original training data. The hallucinator used was a multilayer perceptron. Similar to [33], the goal of this project was to similarly augment the training set with new MRI scans. However, the study used a three layer MLP, which would not be able to make the complex transformations necessary to generate the 3-dimensional scans.

Kingma and Welling described the first methods of generation using encoders [18], but like Wang et al., the networks described were used for simple image tasks with small data sets and small dimensional images. Goodfellow et al.'s work with GANs and the image generation process created not only a generator (similar to the hallucinator previously), but also a discriminator - serving as the adversarial network in the pair [11]. The GAN would fix the problem of a small dataset.

Meanwhile, like in [30], auxiliary classifiers were used to solve the problem of the vanishing gradient.

Approach - Second Iteration The second iteration of the system consists of 2 network models—one for image classification and the final diagnosis, and one for artificial image synthesis for data augmentation—shown in Figure 6. The images are from a shared dataset, but the images used to train the GAN as well as the ones generated by the GAN were only used in the training set for the classifier. This meant that only real images not seen by either the GAN or classifier in training made up the evaluation (test) set.

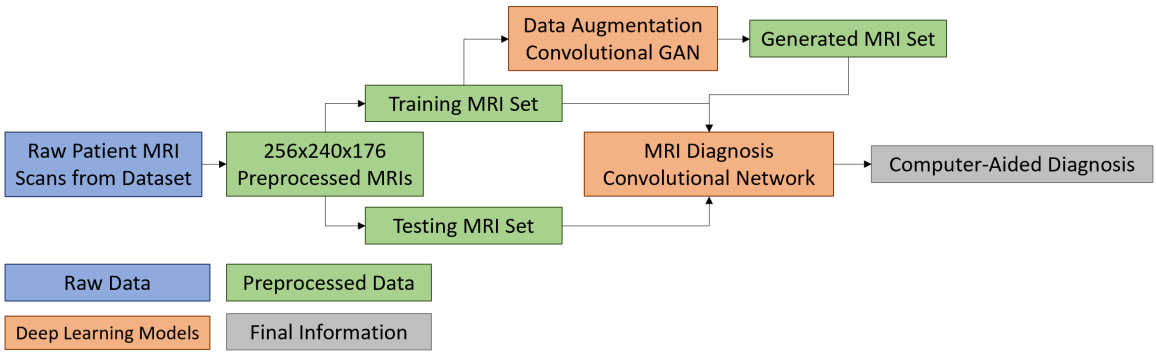

Figure 6: Graphical Description of the flow of data and processing through the system.

Evidently the change made is the GAN and its ability to add to the training data for the diagnosis network. The same architectures from the first iteration (VGG, GoogLeNet, and Resnet) were trained again to determine the effects of the addition to the dataset.

Additionally, one other set of networks was used, including a feature extractor appended to a classification network. Auxilliary classifiers were added throughout the VGG, GoogLeNet, Resnet, and new models. Similarly to [19], the generator network's weights were initialized as block diagonal identity matrices. The GAN model was inspired by [2]. The architecture of the new model is described in Figure 7.

Convolutional Network Configuration
Convolutional 3D (512) x 3 Layers
Max-Pooling 3D (4x4x4)
Convolutional 3D (128) x 4 Layers
Max-Pooling 3D (4x4x4)
Convolutional 3D (16) x 4 Layers
Max-Pooling 3D (2x2x2)
Convolutional 3D (8) x 2 Layers
Global Averaging 3D (2x2x2)
FC – 512
FC – 16
Softmax

Figure 7: Graphical Description of the flow of data and processing through the system.

The same optimizations made after model creation for the first iteration were applied to the second, including using Spearmint.

Results

First Iteration The following table indicates accuracy, sensitivity, and specificity of the 3 classification algorithms after the 15 training epochs.

Table 1: Table detailing the accuracy, sensitivity, and specificity of the 3 models.

Model	Accuracy	Sensitivity	Specificity
VGG-19	90.20%	93.10%	81.08%
GoogLeNet	84.97%	86.21%	81.08%
Resnet-50	88.89%	92.24%	78.38%

$$Sensitivity = \frac{TP}{TP + FN} \qquad Specificity = \frac{TN}{TN + FP}$$

.

The number of epochs, 15, was chosen as a part of the EarlyStopping callback in Keras [32], which continues the training of the algorithm until the validation loss experiences multiple epochs of no decrease. The graphs of loss are depicted in Figure 8.

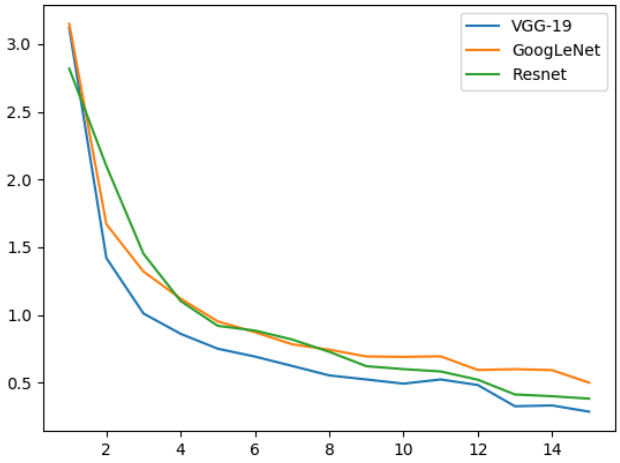

Figure 8: Validation Loss Graphs of the 3 models through the 15 epochs.

The model's weights were fine-tuned to attain improved performance on the validation set, and had minimal overfitting due to many empirically tested methods, including adding dropout, regularization, and patience.

After using the spearmint package, the accuracy of the VGG model was increased to 91.4%. The ROC curve for the finalized model is given in Figure 9.

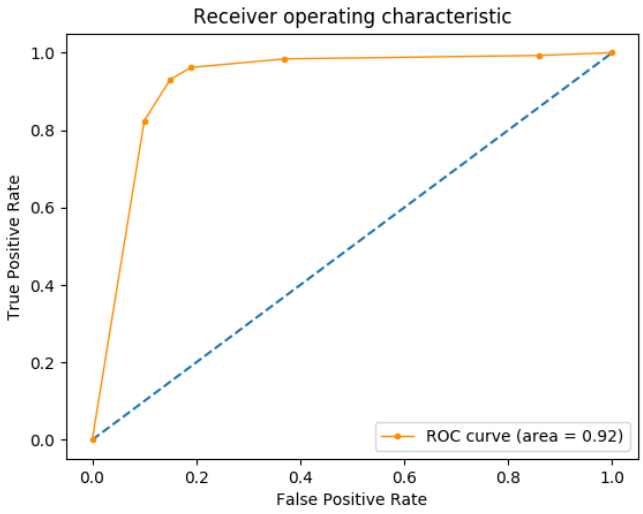

Figure 9: Receiving Operating Characteristic curve for the VGG Model.

Second Iteration With the addition of the GAN, the 3 models performed significantly better. In addition, the new model created surpassed the accuracy of the previous and new models. Table 2 gives the accuracy, sensitivity, and specificity of the new models.

Table 2: Table detailing the accuracy, sensitivity, and specificity of the 4 new models.

Model	Accuracy	Sensitivity	Specificity
VGG-19	94.12%	94.83%	91.89%
GoogLeNet	91.50%	92.24%	89.19%
Resnet-50	89.54%	87.93%	94.59%
PDGAN	96.62%	97.41%	94.59%

To determine the effects of only the addition of the new data, the same number of epochs was used for each system, namely 15. However, it was evident the model was still able to learn more, as the loss formed a monotonic decreasing sequence. Looking at the ROC curve in Figure 10, as it is a common evaluation metric among medical devices [7], it is evident that PDGAN's performance is better than all the first iteration models due to the increase in dataset.

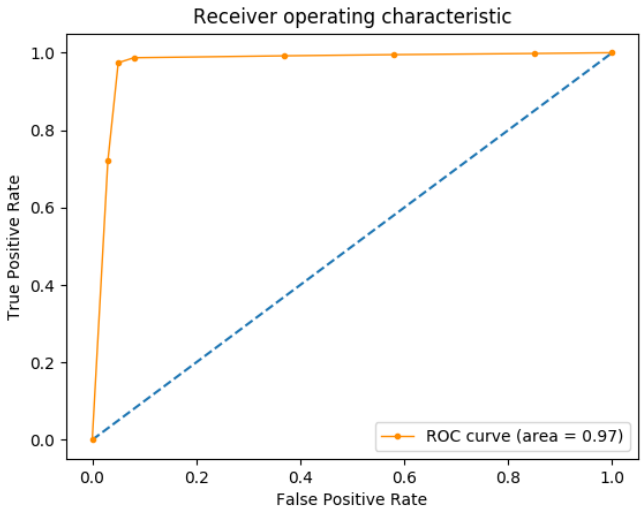

Figure 10: Receiving Operating Characteristic curve for the new PDGAN Model.

Discussion

The primary objective of the work was to improve on the current diagnosis system for PD and increase the chance for early diagnosis among patients. As many patients get MRI scans done as a preliminary test during PD screening, analyzing MRI scans for early signs of Parkinson's Disease was ideal because of the MRI's ability to analyze the anatomy of the brain with high precision [12]. Computationally generating an alternative diagnosis gives neurologists the power to access second opinions in a time and money efficient manner. This is crucial for the well-being of patients, as treatment plans become less accessible for patients with Parkinson's in later stages [6].

With around 12 thousand MRI machines in the US, the applicability of the system is widespread. Still, in over 20 years, a firm technological solution has not emerged for the diagnosis of PD in spite of the MRI process being standard procedure during the screening of MRIs. Multiple reports have cited an technological solution as a next step in the diagnostic efforts[15, 26, 16].

A search of scientific literature reveals that the PDGAN is unique in its ability to diagnose Parkinson's Disease with an augmented, generated dataset. Table 3 illustrates several related

studies, and their difference to the PDGAN's methodology.

Table 3: Table detailing information about several related studies and the difference between the study and the current study [3, 24, 9, 10, 1].

Study	Description	Input Data	Methodology	Accuracy	Difference between study
Chen, 2013	FKNN – based Diagnosis	Voice Measurements	Fuzzy K-Nearest Neighbors	96.07%	Had thousands of sample data
Frid-Adar et al.	Liver Lesion Classification	Liver Lesion Images	GANs, CNNs	88.4%	Used GANs but with a different classifier – low accuracy
Gil et al.	MLP – based Diagnosis	Voice Measurements	MLP and SVM	92.31%	Had thousands of sample data
Adams, 2017	Typing based Diagnosis	Typing Movements	Various Machine Learning Models	96.1%	Had thousands of sample data
Pereisa et al.	Writing and Medical Exam Diagnosis	Handwriting, Medical Exam Information	Computer Vision Processing, CNNs, MLP	67%	Low accuracy, used a combination of tests.

Although this method of improving accuracy hasn't been used in the context of Parkinson's Disease, there is still room for improvement. Evidently, the accuracy could be increased with a larger starting dataset. Additionally, due to hardware constraints, the spearmint code was stopped prematurely, only making 2 optimizations. With more time and better hardware, the optimization would have improved.

Conclusion and Future Work

This study presented the first data augmented approach at classifying Parkinson's Disease. It is able to classify MRI scans with a high accuracy by augmenting the originally small dataset with more images to learn upon. To obtain this information, the study used a combination of state of the art CNNs and GAN models, as well as a self-created CNN able to outperform the defined ones using the spearmint package.

Several performance metrics for all 7 systems are summarized in Table 4, and demonstrate the positive effect of adding generated images to the dataset.

Table 4: Table detailing the accuracy, sensitivity, and specificity of the 7 models. The bolded models indicate ones that had access to the augmented data.

Model	Accuracy	Sensitivity	Specificity
PDGAN	**96.62%**	**97.41%**	**94.59%**
VGG-19	94.12%	94.83%	91.89%
VGG-19	**94.12%**	**94.83%**	**91.89%**
GoogLeNet	84.97%	86.21%	81.08%
GoogLeNet	**91.50%**	**92.24%**	**89.19%**
Resnet-50	88.89%	92.24%	78.38%
Resnet-50	**89.54%**	**87.93%**	**94.59%**

A future step for this study would be to broaden the scope of analysis of patients. Another method of analysis apart from looking at MRI scans would be to use genetic markers. Cited as another method of early detection [6, 37], looking at methylation profiles as a way to detect early signs of Parkinson's Disease would be a clear next step [8]. Already being used as a diagnosis tool for cancer [13] and muscular dystrophy [17], the PPMI database already has genetic markers and methylation profiles for patients with Parkinson's Disease.

Overall, it is clear with this research and the PDGAN model that the addition of a Generative Adversarial Network for the purpose of data augmentation is beneficial to the classification accuracy where there is not an large, accessible dataset.

References

[1] Warwick R. Adams. "High-accuracy detection of early Parkinson's Disease using multiple characteristics of finger movement while typing". In: *PLOS ONE* 12.11 (Nov. 2017), pp. 1–20. DOI: 10.1371/journal.pone.0188226. URL: https://doi.org/10.1371/journal.pone.0188226.

[2] Antreas Antoniou, Amos Storkey, and Harrison Edwards. "Data Augmentation Generative Adversarial Networks". In: abs/1803.01229 (2018). URL: https://arxiv.org/abs/1803.01229.

[3] Huiling Chen. "An efficient diagnosis system for detection of Parkinson's disease using fuzzy k-nearest neighbor approach". In: *Expert Systems with Applications* (Jan. 2013).

[4] Mayo Clinic. *Parkinson's Disease Diagnosis and Treatment*. URL: https://www.mayoclinic.org/diseases-conditions/parkinsons-disease/diagnosis-treatment/drc-20376062.

[5] Mayo Clinic. *Parkinson's Disease Symptoms and Causes*. URL: https://www.mayoclinic.org/diseases-conditions/parkinsons-disease/symptoms-causes/syc-20376055.

[6] Emily Downward. *Diagnosis & Early Symptoms & Early Diagnosis*. 2017. URL: https://parkinsonsdisease.net/diagnosis/early-symptoms-signs/.

[7] Tom Fawcett. "An introduction to ROC analysis". In: *Pattern Recognition Letters* (2006).

[8] Suhua Feng et al. "Determining DNA methylation profiles using sequencing." In: *Methods Molecular Biology* 733 (2011), pp. 223–238. DOI: 10.1007/978-1-61779-089-8_16. URL: https://link.springer.com/protocol/10.1007/978-1-61779-089-8_16.

[9] Maayan Frid-Adar et al. "GAN-based Synthetic Medical Image Augmentation for increased CNN Performance in Liver Lesion Classification". In: *CoRR* abs/1803.01229 (2018). arXiv: 1803.01229. URL: http://arxiv.org/abs/1803.01229.

[10] David Gil and Magnus Johnsson. "Diagnosing Parkinson by using artificial neural networks and support vector machines". eng. In: Global Journal of Computer Science and Technology 9.4 (2009), pp. 63–71. ISSN: 0975-4172.

[11] Ian J. Goodfellow et al. "Generative Adversarial Networks". In: abs/1406.2661 (2014). URL: https://arxiv.org/abs/1406.2661.

[12] B Griewing, H Hielscher, and A Lutcke. "The importance of MRI (magnetic resonance imaging) for the diagnosis of brainstem infarction." In: *Bildgebung* 59.2 (1992), pp. 94, 97.

[13] Xiaoke Hao et al. "DNA methylation markers for diagnosis and prognosis of common cancers". In: *Proceedings of the National Academy of Sciences* 114.28 (2017), pp. 7414–7419. ISSN: 0027-8424. DOI: 10.1073/pnas.1703577114. eprint: http://www.pnas.org/content/114/28/7414.full.pdf. URL: http://www.pnas.org/content/114/28/7414.

[14] Kaiming He et al. "Deep Residual Learning for Image Recognition". In: *CoRR* (2015). eprint: 1512.03385. URL: http://arxiv.org/abs/1512.03385.

[15] A J Hughes et al. "Accuracy of clinical diagnosis of idiopathic Parkinson′s disease: a clinico-pathological study of 100 cases." In: *Journal of Neurology, Neurosurgery & Psychiatry* 55.3 (1992), pp. 181–184. DOI: http://doi.org/10.1136/jnnp.55.3.181.

[16] Andrew J. Hughes et al. "What features improve the accuracy of clinical diagnosis in Parkinson′s disease". In: *Neurology* 42.6 (1992), pp. 1142–1142. ISSN: 0028-3878. DOI: 10.1212/WNL.42.6.1142. eprint: http://n.neurology.org/content/42/6/1142.full.pdf. URL: http://n.neurology.org/content/42/6/1142.

[17] Takako Jones et al. "Identifying diagnostic DNA methylation profiles for facioscapulohumeral muscular dystrophy in blood and saliva using bisulfite sequencing". In: *Clin Epigenetics* 6.1 (2014), p. 23. DOI: 10.1186/1868-7083-6-23. URL: https://clinicalepigeneticsjournal.biomedcentral.com/articles/10.1186/1868-7083-6-23.

[18] Diederik Kingma and Max Welling. "Auto-Encoding Variational Bayes". In: abs/1312.6114 (2014). URL: https://arxiv.org/abs/1312.6114.

[19] Quoc V. Le, Navdeep Jaitly, and Geoffrey E. Hinton. "A Simple Way to Initialize Recurrent Networks of Rectified Linear Units". In: *CoRR* abs/1504.00941 (2015). arXiv: 1504.00941. URL: http://arxiv.org/abs/1504.00941.

[20] Fredrik Lundh. *Python Imaging Library (PIL)*. 2018. URL: http://www.pythonware.com/products/pil/.

[21] Martín Abadi et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Software available from tensorflow.org. 2015. URL: http://tensorflow.org/.

[22] MathWorks. *Introduction to Deep Learning - what are Convolutional Neural Networks*. 2017. URL: https://www.mathworks.com/videos/introduction-to-deep-learning-what-are-convolutional-neural-networks--1489512765771.html.

[23] Dimitrios Siganos Christos Stergiou. *Neural Networks*. Imperial College London, 1996. URL: https://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html.

[24] Clayton R. Pereira et al. "A Step Towards the Automated Diagnosis of Parkinson's Disease: Analyzing Handwriting Movements". In: *2015 IEEE 28th International Symposium on Computer-Based Medical Systems* (2015), pp. 171–176.

[25] CR Pereira et al. "A new computer vision-based approach to aid the diagnosis of Parkinson's disease." In: *Comput Methods Programs Biomed* 136 (2016), pp. 79–88. DOI: http://doi.org/10.1016/j.cmpb.2016.08.005..

[26] G Rizzo et al. "Accuracy of clinical diagnosis of Parkinson disease: A systematic review and meta-analysis". In: *Neurology* 86.6 (2016), pp. 566–576. DOI: http://doi.org/10.1212/WNL.0000000000002350.

[27] Karen Simonyan and Andrew Zisserman. "Very Deep Convolutional Networks for Large-Scale Image Recognition". In: *CoRR* abs/1409.1556 (2014). arXiv: 1409.1556. URL: http://arxiv.org/abs/1409.1556.

[28] Jasper Snoek, Hugo Larochelle, and Ryan Adams. "Practical Bayesian Optimization of Machine Learning Algorithms". In: *Advances in Neural Information Processing Systems* (2012). eprint: 1206.2944. URL: https://arxiv.org/abs/1206.2944.

[29] University of Southern California. *Parkinson's Progression Markers Initiative*. URL: https://ida.loni.usc.edu/login.jsp?project=PPMI.

[30] Christian Szegedy et al. "Going Deeper with Convolutions". In: *CoRR* abs/1409.4842 (2014). arXiv: 1409.4842. URL: http://arxiv.org/abs/1409.4842.

[31] C. Tanner and S. Goldman. "Epidemiology of Parkinsons disease". In: *Neurologic Clinics* 14 (1996), pp. 317–335.

[32] Keras Team. *Keras*. 2018. URL: https://github.com/keras-team/keras.

[33] Yu-Xiong Wang et al. "Low-Shot Learning from Imaginary Data". In: *CoRR* abs/1801.05401 (2018). arXiv: 1801.05401. URL: http://arxiv.org/abs/1801.05401.

[34] K.E. Watkins et al. "Structural Asymmetries in the Human Brain: a Voxel-based Statistical Analysis of 142 MRI Scans". In: *Cerebral Cortex* 11.9 (2001), pp. 868–877. DOI: 10.1093/cercor/11.9.868. eprint: /oup/backfile/content_public/journal/cercor/11/9/10.1093_cercor_11.9.868/1/1100868.pdf. URL: http://dx.doi.org/10.1093/cercor/11.9.868.

[35] WebMD. *Parkinson's Causes*. URL: https://www.webmd.com/parkinsons-disease/guide/parkinsons-causes#1.

[36] WebMD. *Parkinson's Later Diagnosis*. URL: https://www.webmd.com/parkinsons-disease/news/20101004/parkinsons-later-diagnosis-earlier-death#1.

[37] Tao Wu and Mark Hallett. "A functional MRI study of automatic movements in patients with Parkinson's disease". In: *Brain* 128.10 (2005), pp. 2250–2259. DOI: 10.1093/brain/awh569. eprint: /oup/backfile/content_public/journal/brain/128/10/10.1093/brain/awh569/2/awh569.pdf. URL: http://dx.doi.org/10.1093/brain/awh569.