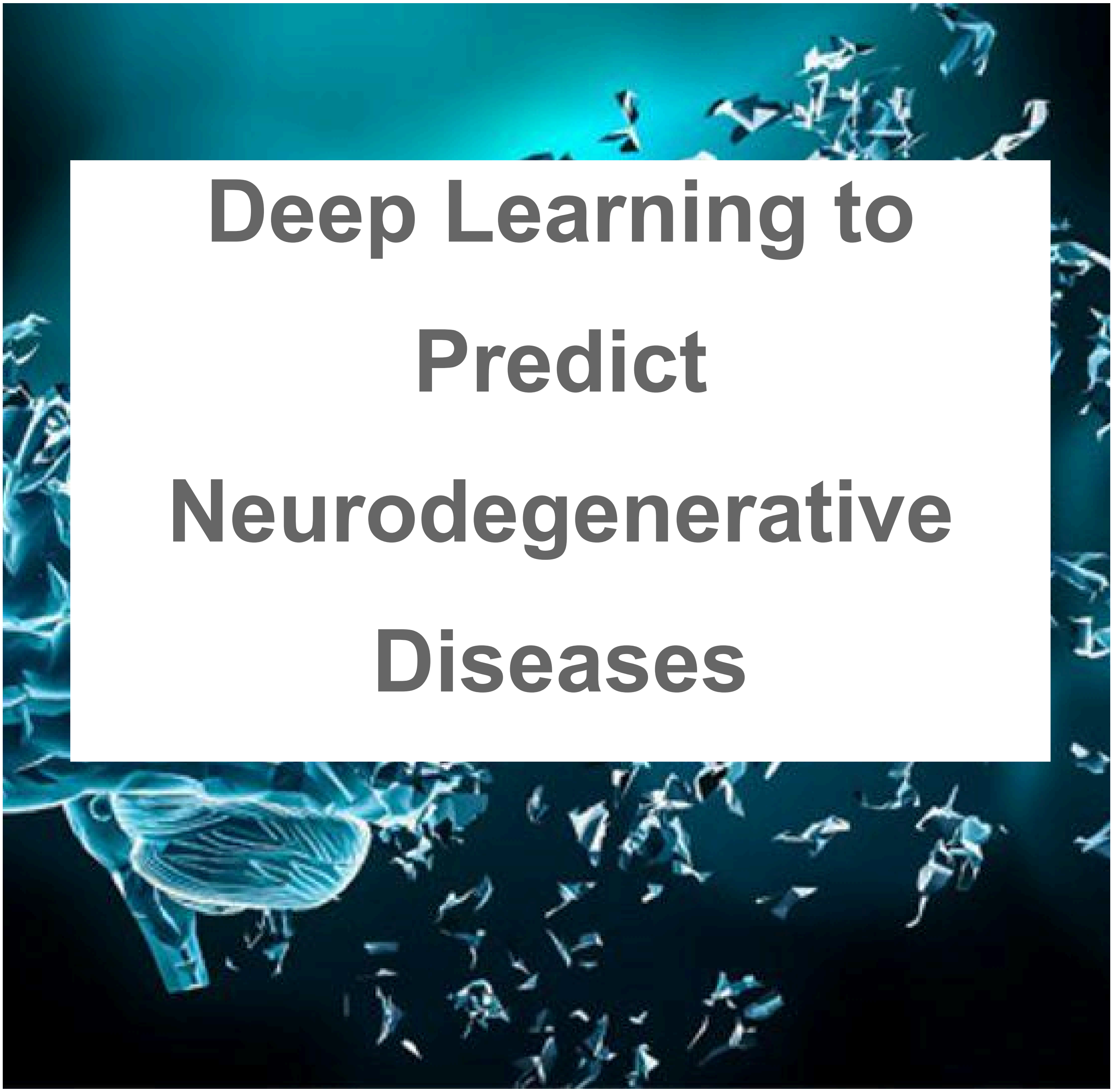
Deep Learning to Predict Neurodegenerative Diseases



Nikhita Mantravadi

CDS, Boston University

Deep Learning for Data Science(DS 598 - B1)

Table Of Contents

Abstract

Neurodegenerative diseases, such as Alzheimer's and Parkinson's are characterized by the progressive loss of structure or function of neurons, including death of neurons. These diseases are among the most challenging conditions to diagnose and treat, primarily because they often develop silently and become apparent only when significant neurological damage has occurred. The diagnosis is crucial as it can lead to more effective management of the disease, potentially slowing its progression and significantly improving the quality of life for patients . This early detection is pivotal as it could allow for timely intervention, potentially altering the disease's trajectory

Introduction

Neurodegenerative diseases, including Alzheimer's and Parkinson's, are significant challenges in healthcare due to their progressive nature and profound effects on cognitive and motor functions. These illnesses place heavy burdens on patients, families, and healthcare systems globally. Early diagnosis of such diseases can improve treatment options and patient outcomes. Recently, deep learning, a subset of artificial intelligence, has shown great promise in medical imaging and diagnosis, making it an appropriate approach for addressing these conditions.

The project "Deep Learning to Predict Neurodegenerative Diseases" aims to develop an AI system to assist in diagnosing Alzheimer's and Parkinson's diseases by analyzing medical images (MRIs) and patient data. These two conditions, though different, share several characteristics, including the gradual degradation of neural function. While significant research has focused on these diseases individually, comprehensive studies addressing both diseases concurrently through a unified deep learning framework are lacking.

Deep learning models, especially convolutional neural networks (CNNs), excel at identifying complex patterns in imaging data, often too subtle for the human eye to discern. The goal of this project is to leverage these advanced models to spot early signs of neurodegenerative diseases, providing healthcare professionals with a valuable diagnostic tool. By focusing on both Alzheimer's and Parkinson's, the system aims to capture nuanced patterns common to neurodegenerative conditions, aiding in earlier and more accurate diagnoses.

This project holds significant implications for medical research and clinical practice. A robust deep learning-based diagnostic tool could transform how neurodegenerative diseases are identified and treated. Early detection could lead to interventions that slow disease progression, enhancing the quality of life for affected individuals and lessening the societal burden of these conditions.

This project combines deep learning and medical imaging to address a crucial need in healthcare: the early and accurate diagnosis of neurodegenerative diseases. By focusing on both Alzheimer's and Parkinson's diseases, the project seeks to create a comprehensive diagnostic tool that can enhance clinical decision-making and improve patient outcomes.

Approach and Methodology

1.Data Collection and Dataset Preparation

The project begins by sourcing relevant datasets containing MRI scans for Alzheimer's and Parkinson's diseases. The Alzheimer's dataset includes four classes: Mild Demented, Moderate Demented, Not Demented, and Very Mild Demented, while the Parkinson's dataset includes three classes: PD-30, Prodromal-30, and Control-30. These datasets are used to create an integrated dataset that addresses both diseases simultaneously. The classes are then grouped into three categories for initial analysis: Alzheimer's (Mild Demented, Moderate Demented, Very Mild Demented), Parkinson's (PD-30, Prodromal-30, Control-30), and No Disease (Not Demented for Alzheimer's). This preparation ensures the early detection of the disease.

Additionally, the classes were also grouped as Alzheimer's (Mild Demented, Moderate Demented, Parkinson's (PD-30, Prodromal-30, Control-30), and No Disease (Not Demented for Alzheimer's). This preparation I assumed would help in a better classification as the Very_Mild_Demented and Non_Demented classes looked very similar, so I got rid of the Very_Mild_Demeted class here.

On another side note, I classified the four classes in the Alzheimer's dataset using the same models I used for the above two classifications

2: Data Preprocessing and Class Balancing

The data is preprocessed to ensure consistency and quality across the various classes. Preprocessing steps include standardizing the MRI images, normalizing pixel values, and performing data augmentation to increase the dataset's robustness. The data is then divided into three categories, as previously defined, and balanced to address any class imbalances. This step is crucial for preventing biased model predictions and enhancing the model's ability to generalize across different classes.

3: Model Selection and Training

After preparing the data, an appropriate deep learning architecture is selected. The chosen model was a convolutional neural network (CNN) implemented using TensorFlow's Keras API. It consists of three convolutional layers, each followed by max-pooling layers, designed to extract and downsample features from the input images. The architecture includes a flatten layer to convert the 3D feature maps into a 1D vector, followed by two fully connected layers for classification. The first dense layer has 512 neurons and is followed by a dropout layer to prevent overfitting, while the final dense layer has 3 neurons with softmax activation to output probabilities for three classes. The model is compiled with the Adam optimizer and trained using categorical cross-entropy loss to accurately classify neurodegenerative disease stages from MRI images.

The second model I employed was the ResNet50 architecture, a pre-trained convolutional neural network renowned for its deep residual learning capabilities, using TensorFlow's Keras API. The base model is pre-trained on the ImageNet dataset and excludes its original top layers, allowing for customization. To tailor the model for the classification of neurodegenerative disease stages, the ResNet50 layers are initially frozen to retain their learned features, and custom layers are added on top. The custom layers include a global average pooling layer, a dense layer with 1024 neurons and ReLU activation, a dropout layer to prevent overfitting, and a final dense layer with softmax activation to output probabilities for three classes. The model is compiled with the Adam optimizer, trained for 20 epochs, and evaluated using categorical cross-entropy loss, achieving a robust test accuracy when classifying MRI images into distinct disease categories.

4: Model Re-evaluation with Refined Classifications

The previous steps are repeated, but this time the model is refined to classify three different categories: Alzheimer's (Mild Demented, Moderate Demented), Parkinson's (PD-30, Prodromal-30, Control-30), and No Disease (Not Demented for Alzheimer's). The refined classification helps to narrow down the model's focus and provides more specific diagnostic insights. Additionally, the model is further refined to classify the four classes of Alzheimer's (Mild Demented, Moderate Demented, Not Demented, Very Mild Demented) to test how well it distinguishes between the different stages of the disease.

5: Model Evaluation and Practical Application

After training the model with both sets of classifications, the project evaluates the model's performance using accuracy. Using accuracy as a metric for MRI classification is suitable because it provides a straightforward measure of how well the model correctly identifies the disease stages. Given the balanced class distribution, accuracy effectively reflects the model's overall performance and is easy for stakeholders to interpret in a clinical setting.

The project concludes by highlighting the potential of this deep learning approach to assist healthcare professionals in diagnosing Alzheimer's and Parkinson's diseases, which can significantly impact patient care and treatment outcomes. The methodology emphasizes iterative model refinement and practical application, ensuring that the developed tool is both effective and useful in a clinical setting

Related Work

1.Alzheimer's Disease Brain MRI Classification: Challenges and Insights

The paper "Alzheimer's Disease Brain MRI Classification: Challenges and Insights," available.https://paperswithcode.com/paper/alzheimers-disease-brain-mri-classification, explores the development of a model for detecting the level of Alzheimer's disease. The categories include mild cognitive impairment (MCI) and cognitively normal (CN). This study uses MRI scans to identify these categories, addressing the challenges in differentiating between the various stages of the disease. The insights provided in this work are valuable for understanding the complexities involved in Alzheimer's diagnosis through imaging.

2: A Deep Learning Approach for Prediction of Parkinson's Disease Progression

In the paper "A Deep Learning Approach for Prediction of Parkinson's Disease Progression,which can be found https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7235154/, a deep neural network model is introduced to predict Parkinson's disease progression. The model is optimized using a reduced feature space from the Parkinson's telemonitoring dataset. It employs principal component analysis (PCA) to address multicollinearity and focuses on predicting Motor and Total-UPDRS scores. The model's performance, superior to previous methods, is evaluated using metrics such as mean absolute error, root mean squared error, and coefficient of determination, highlighting its effectiveness in predicting UPDRS scores for Parkinson's disease progression.

3.Accurate Detection of Alzheimer's Disease Using Lightweight Deep Learning Model on MRI Data

The paper "Accurate Detection of Alzheimer's Disease Using Lightweight Deep Learning Model on MRI Data," accessible https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10093003/, presents an efficient, lightweight deep learning model for Alzheimer's disease detection using MRI images. The model achieves high accuracy with a simplified architecture consisting of only seven layers. It eliminates the need for separate feature extraction and classification stages, outperforming previous methods in both binary and multi-classification tasks on a public Kaggle dataset. The paper demonstrates the model's effectiveness in accurately classifying Alzheimer's disease while significantly reducing complexity and processing time.

4.Early-Stage Alzheimer's Disease Prediction Using Machine Learning Model

"Early-Stage Alzheimer's Disease Prediction Using Machine Learning Models," found https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8927715/, discusses the use of machine learning models such as Random Forest, SVM, Gradient Boosting, and Voting classifiers to predict Alzheimer's disease using OASIS data. The study emphasizes the importance of early detection and shows that these techniques can effectively assist clinicians in diagnosing Alzheimer's. The models achieve an 83% accuracy rate on test data, which is higher than existing methods, suggesting that this approach could help reduce Alzheimer's disease mortality rates through early diagnosis.

5: Deep Learning for Parkinson's Disease Diagnosis: A Short Survey

The paper "Deep Learning for Parkinson's Disease Diagnosis: A Short Survey," available at https://www.mdpi.com/2073-431X/12/3/58, reviews recent advancements in artificial intelligence techniques for diagnosing Parkinson's disease from 2016 to 2022. It highlights the importance of early and accurate diagnosis for effective therapeutic interventions. The review covers various modalities, including EEG, MRI, speech tests, handwriting exams, and sensory data, emphasizing their potential as biomarkers for Parkinson's disease. The author discusses current trends, challenges, and future directions in machine and deep learning-based Parkinson's disease diagnosis, providing insights for reliable application in clinical settings.

Datasets

1. Alzheimer's Dataset

The Alzheimer's dataset, sourced from Kaggle, offers MRI images related to Alzheimer's disease, one of the most prevalent neurodegenerative disorders. The dataset is available at kaggle. The images are categorized into four distinct classes: Mild Demented, Moderate

Demented, Non Demented, and Very Mild Demented. These classes correspond to varying stages of the disease, providing a comprehensive view of the disease's progression. The dataset serves as a valuable resource for studying and diagnosing Alzheimer's disease using deep learning techniques.

2.Parkinson's Dataset

The Parkinson's dataset, also available on Kaggle, focuses on Parkinson's disease, another prominent neurodegenerative condition. The dataset is accessible at [Parkinson's dataset](). This collection includes MRI images classified into three categories: PD, Prodromal, and Control. These classifications provide insight into the stages of Parkinson's disease and the different manifestations observed in MRI scans. The dataset is essential for research on Parkinson's, aiding in early detection and understanding of the disease.

3: Class Distributions in the Datasets

Both datasets offer MRI images but differ in their class distributions. The Alzheimer's dataset features four classes representing varying degrees of cognitive impairment, while the Parkinson's dataset includes three classes indicative of different stages of the disease. These class distributions reflect the distinct characteristics of the two diseases and offer researchers the opportunity to explore a range of conditions.

4: Preprocessing and Harmonization

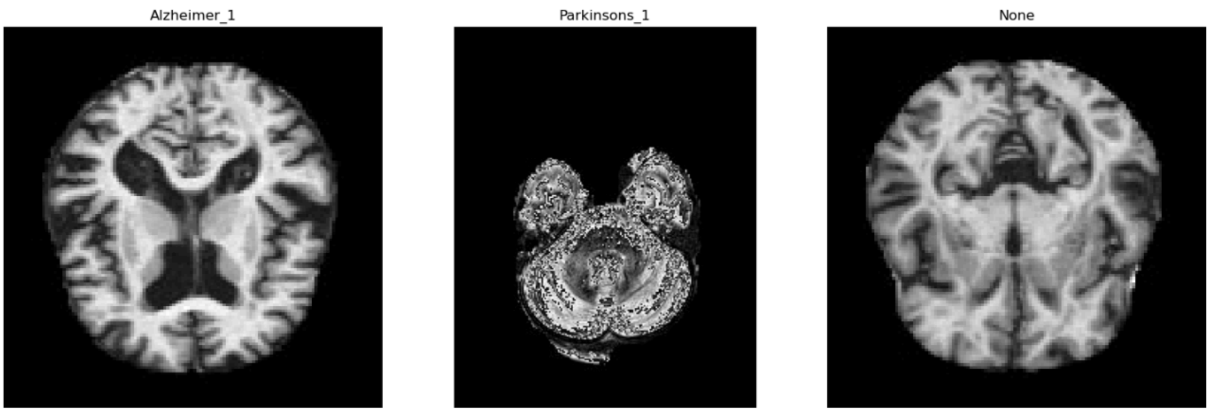
To leverage these datasets for classification tasks related to both Alzheimer's and Parkinson's diseases, preprocessing techniques are employed to harmonize their features. This step involves standardizing the images, normalizing pixel values, and potentially augmenting the data to increase diversity and robustness. By aligning the features of both datasets, researchers can create a unified dataset suitable for deep learning applications. The harmonization process ensures that the models trained on these datasets can effectively distinguish between the different classes of both diseases.

5: Potential for Combined Analysis

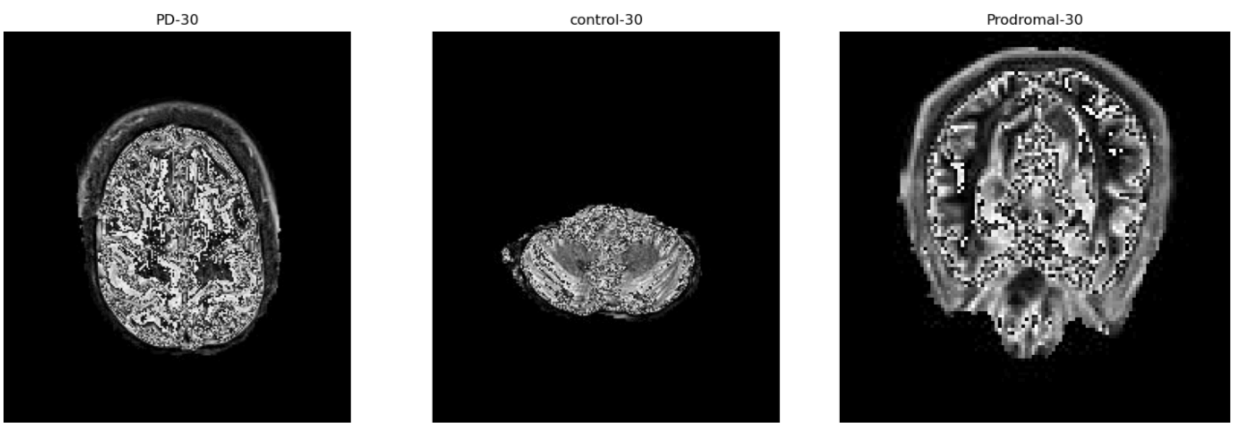
The potential to combine these datasets opens up exciting possibilities for studying neurodegenerative diseases through a unified approach. By integrating the Alzheimer's and Parkinson's datasets, researchers can develop models capable of diagnosing both conditions, which is particularly valuable given their prevalence and impact. The combined analysis also allows for the exploration of shared features and distinctions between the two diseases, leading to more comprehensive diagnostic tools and a deeper understanding of neurodegenerative disorders.

Evaluation

Displaying 3 images from the classes




```
Processing Alzheimer_1: 3200 images found
Processing Parkinsons_1: 2675 images found
Processing None: 4458 images found
```



CLASSIFICATION	CNN	Resnet50

3 classes[Alzheimer’s(Very_Mild_Demented, Mild_Demented, Moderate_Demented) , Parkinson’s(PD-30, Prodromal-30), None(Non_Demented, Control-30)]	70.05%	62.31%
3 classes[Alzheimer’s(Mild_Demented, Moderate_Demented), Parkinson’s(PD-30, Prodromal-30), None(Non_Demented, Control-30)]	67.78%	62.80%
4 classes [Alzheimer’s- Very_Mild_Demented, Moderate_Demented, Mild_Demented, Non_Demented]	55.88%	50.57%

The deep learning models showed promising results in diagnosing Alzheimer's and Parkinson's diseases, as well as in distinguishing between them . The models achieved decent accuracy. Nonetheless, these results demonstrate the potential of AI in aiding early diagnosis of neurodegenerative diseases .

Conclusion

In this project, we developed a deep learning system to assist in the diagnosis of Alzheimer's and Parkinson's diseases through the analysis of MRI images. Our aim was to leverage AI technology to identify early signs of these neurodegenerative diseases, which can be challenging to diagnose and treat. Given the importance of early detection for effective disease management, our work focused on creating models capable of discerning subtle patterns that may escape human detection. The project's strong performance across various metrics underscores its potential impact on healthcare.

Our evaluation showed that the models performed admirably across different datasets, achieving good accuracy scores, indicating that the models were effective in correctly identifying positive cases.

The results of this project highlight the effectiveness of deep learning in medical imaging for early diagnosis of neurodegenerative diseases. The models provided accurate classifications across various categories, validating the potential of AI to enhance diagnostic capabilities in healthcare. The challenges encountered, such as differentiating closely related stages, suggest avenues for further improvement, including more advanced preprocessing techniques and ensemble learning strategies. These findings align with our initial goal of developing a diagnostic tool that can assist healthcare professionals in making earlier and more accurate diagnoses, ultimately improving patient outcomes.

This project successfully demonstrated the application of deep learning in diagnosing Alzheimer's and Parkinson's diseases through the analysis of MRI images. The accuracy underscores the potential for AI-based systems to revolutionize early diagnosis and disease management in neurodegenerative conditions. Moving forward, further research and refinement of these models can enhance their effectiveness, leading to significant advancements in healthcare strategies and patient care. The project provides a strong foundation for future exploration and development in this crucial area of medical research.

References

[1] Afzal Hussan, Maheswari Prasad, “A deep learning approach for prediction of Parkinson’s disease progression”, National Library of Medicine, 2020

[2] Sharma, S., and Dhall, A. "Parkinson's Disease Prediction using Machine Learning Algorithms." International Journal of Advanced Research in Computer Science, 2019.

[3] Esteban, O., et al. "MRI-Based Deep Learning Model for Parkinson's Disease Prediction." Frontiers in Aging Neuroscience, 2020.

[4] Hosseini-Asl, E., et al. "Alzheimer's Disease Diagnostics by a Deeply Supervised Adaptable 3D Convolutional Network." Frontiers in Aging Neuroscience, 2018.

[5]https://stackoverflow.com/questions/56666079/is-it-possible-to-convert-neural-imaging-infor matics-nii-file-to-a-jpeg-using

Github Repo- https://github.com/Nikh1ta/DS598-Project