

Generative Models for Synthesizing PET Images for Alzheimer Disease

MANISH PANDA

*Master of Science in COMPUTER VISION, ROBOTICS & MACHINE
LEARNING*
from the
University of Surrey



School of Computer Science and Electrical and Electronic Engineering
Faculty of Engineering and Physical Sciences
University of Surrey
Guildford, Surrey, GU2 7XH, UK

September 2024

Supervised by: Prof. Kevin Wells

©MANISH PANDA 2024

DECLARATION OF ORIGINALITY

I confirm that the project dissertation I am submitting is entirely my own work and that any material used from other sources has been clearly identified and properly acknowledged and referenced. In submitting this final version of my report to the JISC anti-plagiarism software resource, I confirm that my work does not contravene the university regulations on plagiarism as described in the Student Handbook. In so doing I also acknowledge that I may be held to account for any particular instances of uncited work detected by the JISC anti-plagiarism software, or as may be found by the project examiner or project organiser. I also understand that if an allegation of plagiarism is upheld via an Academic Misconduct Hearing, then I may forfeit any credit for this module or a more severe penalty may be agreed.

MSc Dissertation Title: Generative Models for Synthesizing PET Images for Alzheimer Disease

Author Name: Manish Panda

Author Signature 

Date: 02/09/2024

Supervisor's name: Prof. Kevin Wells

WORD COUNT

Number of Pages: 69

Number of Words: 13900

ABSTRACT

Alzheimer's disease, one of the most common causes of dementia, is difficult to diagnose at an early stage because of its complicated course and the limits of available imaging methods. Alzheimer's disease is mostly detected by Positron Emission Tomography (PET) imaging, which provides important insights into brain metabolism. However, the high cost, restricted availability, and requirement for radioactive tracers limit the broad usage of PET.

This study starts with a thorough examination of a large PET dataset that has been divided into different phases of cognitive impairment. The dataset was meticulously preprocessed, using dimensionality reduction using Principal Component Analysis (PCA) to simplify the data structure, augmentation through adaptive histogram equalization to boost contrast, and standardization to assure consistency. These actions were crucial in establishing the foundation for the subsequent stage of the undertaking.

The preparation of the dataset enabled the application of Generative Adversarial Networks (GANs), namely Deep Convolutional GANs (DCGAN) and Super-Resolution GANs (SRGAN). Realistic PET images with great resolution were created and trained on these architectures. Thorough hyperparameter fine-tuning was necessary during the training process in order to achieve the highest levels of image fidelity and quality. Using GANs' power for medical imaging has advanced significantly during this stage.

Making use of measures like Mean Squared Error (MSE) and Structural Similarity Index (SSIM), the investigation shed light on the synthetic pictures' potential uses and therapeutic relevance. The findings showed that GANs might provide a scalable, affordable substitute for conventional PET imaging, potentially transforming the way Alzheimer's disease is diagnosed.

CONTENTS

Declaration of Originality	ii
Word Count	iii
Abstract	iv
List of figures	viii
List of tables	xvi
1 Introduction	1
1.1 Background and Context	1
1.2 Objectives	2
1.3 Achievements	2
1.4 Overview of Dissertation	3
2 Background Theory and Literature Review	5
2.1 World of Dementia	5
2.2 Alzheimer Disease	6
2.2.1 Stages of Alzheimer disease	6
2.2.1.1 Cognitively Normal	7
2.2.1.2 Early Moderate Cognitive Impairment	7
2.2.1.3 Mild Cognitive Impairment	7
2.2.1.4 Late Mild Cognitive Impairment	8
2.3 Diverse Techniques in Medical Imaging	8
2.3.1 Computed Tomography(CT) Scanning	8
2.3.2 Magnetic Resonance Imaging(MRI)	9
2.3.3 Positron Emission Tomography(PET)	10
2.4 Unveiling PET	10

2.4.1	Understanding Positron Emission Tomography	10
2.4.2	The Science Behind Positron Emission Tomography	11
2.4.3	Vital Role of Positron Emission Tomography in Medical Imaging	12
2.4.4	Fluorodeoxyglucose Positron Emission Tomography	12
2.5	Manifold Learning	13
2.5.1	Linear	13
2.5.1.1	Principal Component Analysis(PCA)	13
2.5.2	Non-Linear	15
2.5.2.1	Uniform Manifold Approximation and Projection(UMAP) . . .	15
2.5.2.2	t-Distributed Stochastic Neighbor Embedding(t-SNE)	16
2.5.2.3	Isometric Mapping(Isomap)	18
2.6	Generative Adversarial Networks	19
2.6.1	Deep Convolutional Generative Adversarial Network(DCGAN)	21
2.6.2	Super Resolution Generative Adversarial Network(SRGAN)	22
2.7	Summary	24
3	Extensive Analysis of PET Image Dataset Using Manifold Learning	25
3.1	Overview of the Dataset	25
3.2	Data Preprocessing	25
3.2.1	Standardization	25
3.2.2	Enhancement	27
3.3	Principal Component Analysis	28
3.3.1	K-means Clustering	32
3.4	Manifold Learning	33
3.4.1	t-Distributed Stochastic Neighbor Embedding(t-SNE)	33
3.4.2	Uniform Manifold Approximation and Projection(Umap)	34
3.4.3	Isometric Mapping(Isomap)	36
4	Implementation of GAN	38

4.1	Preprocessing and Fixing Hyperparameters	38
4.2	Deep Convolution Generative Adversarial Network (DCGAN)	39
4.3	Super Resolution Generative Adversarial Network (SRGAN)	46
5	Comparative Analysis of DCGAN and SRGAN Using Evaluation Metrics	54
6	Conclusions	58
6.1	Evaluation	59
6.2	Future Work	60
	Bibliography	61
	Appendix	66

LIST OF FIGURES

2.1 An example of how dementia progresses, showing a steady fall from cognitive function to severe impairment. [19]	5
2.2 A comparative analysis of the major anatomical variations between an Alzheimer's disease and a normal brain. In the Alzheimer's brain, notable alterations include considerable shrinkage of the hippocampus and white matter, enlarged ventricles, widening sulci, and cortical thinning. These changes show how Alzheimer's disease profoundly affects brain structure, resulting in memory loss and cognitive decline.[47]	6
2.3 An illustration of a CT scanner's main parts, including the motorized platform, detector array, X-ray source, X-ray beam, and data gathering system. The patient is exposed to beams from the X-ray source, and the detector array gathers the rays that are transmitted. With the use of this data, the data gathering system generates intricate cross-sectional photographs of the body, which are necessary for precise medical diagnosis.[2]	9
2.4 An illustration of an MRI scanner highlighting the main parts and how they work. The X, Y, and Z magnetic coils change the magnetic field in different directions, but the main magnetic coil produces a powerful, homogeneous magnetic field. The body's internal architecture can be seen in great detail because to the processing of signals that are sent and detected by the radio frequency transmitter and receiver.[31]	9
2.5 Figure illustrating the main parts of a PET scan machine: the motorized exam table moves the patient for scanning, and the PET machine, with its spinning scanner, detects gamma rays released by a radiotracer in the body. The resultant brain scan, which shows metabolic activity in various brain regions and is essential for identifying diseases including cancer, neurological problems, and heart disease, is displayed on the monitor.[8]	10

2.6 An illustration showing the intravenous injection of a radioactive tracer to identify the body's areas of metabolic activity during a PET-CT scan. Positioned on a table, the patient passes through the scanner, which combines CT and PET imaging to provide detailed images that are displayed on a monitor. For the detection and assessment of diseases including cancer, heart disease, and neurological disorders, this method is essential.[36]	11
2.7 Fluorodeoxyglucose (FDG) molecular diagram: Utilized in positron emission tomography (PET) imaging, FDG is a glucose analog that is radiolabeled. Fluorine-18, a radioactive isotope, takes the place of a hydroxyl group in FDG, making it comparable to glucose. PET scans that show metabolic activity can show the accumulation of FDG after injection in regions like the brain that have significant glucose absorption.[12]	12
2.8 Data transformation from the original data space to the principle component space is illustrated in this principle Component Analysis (PCA) illustration. In order to project the data points along the directions of largest variation in the data, the principal components (PC1 and PC2) are used. By lowering the number of dimensions in the data, this method preserves the most important characteristics while simplifying the data and making it simpler to see and analyze.[5]	14
2.9 An example of dimensionality reduction with UMAP is seen in the left image, which displays a point cloud in three dimensions that represents various data point clusters. The data is reduced to a 2-dimensional space while maintaining the structure and linkages between clusters through the use of UMAP (Uniform Manifold Approximation and Projection), as shown in the image on the right. By improving data presentation and interpretation, this method increases accessibility to complicated, high-dimensional data.[29]	15
2.10 An example of the t-Distributed Stochastic Neighbor Embedding (t-SNE) method is displayed on the left, where distances are measured from a location of interest to every other point in the original high-dimensional space. To determine similarity scores, these distances are then displayed on a normal distribution curve (right side). t-SNE preserves the local structure and facilitates the visualization of the relationships between data points by projecting these similarities onto a lower-dimensional space.[43]	17

2.11 An illustration of the generator architecture of a DCGAN (Deep Convolutional Generative Adversarial Network): An input vector that is 100 dimensions in size is first projected and reshaped to begin the process. The data is then incrementally upsampled by the network as it goes through a sequence of deconvolutional (deconv) layers with increasing spatial dimensions (4x4x1024 to 64x64x3). After applying a 5x5 filter to every deconvolution layer, a 64x64 image with three color channels is produced. The creation of realistic images from random noise in GAN models relies heavily on this architecture.[9]	21
2.12 An example of a discriminator architecture using a DCGAN (Deep Convolutional Generative Adversarial Network): A 64x64x3 image is entered at the start of the process and is routed via several convolutional (conv) layers. From 32x32x64 to 4x4x512, each layer decreases the spatial dimensions while increasing the depth. Subsequently, the network reduces the final feature map to a single output value, which indicates the likelihood that the input image is real instead of contrived. In order to differentiate between authentic and fraudulent images in GAN models, this architecture is necessary.[9]	22
2.13 Architecture of the SRGAN generator: A low-resolution input is fed into the first convolutional layer of the process, which is activated by PReLU. After that, it processes a number of residual blocks (B blocks) that have PReLU, batch normalization, and convolutional layers in them. The network employs pixel shufflers for upsampling and a final convolutional layer to produce the high-quality output image, improving image details and resolution, following residual learning.[11] .	22
2.14 SRGAN discriminator architecture: An input image is first processed by batch normalization, multiple convolutional layers with leaky ReLU activations, and subsequent downsampling of the picture. After that, the features are flattened and sent through a series of thick layers, culminating in a final sigmoid activation that yields a probability that either the input image is created (super-resolved) or real (high-resolution). To distinguish between created and genuine images, this output is then compared to the comparable super-resolved (SR) image from the generator.[11]	23

3.1 Each categorization in the dataset—Cognitively Normal (CN), Early Mild Cognitive Impairment (EMCI), Late Mild Cognitive Impairment (LMCI), Mild Cognitive Impairment (MCI), and Alzheimer’s Disease (AD)—is represented by one example brain image in this figure.	25
3.2 A clearer depiction of the structure of the brain has been achieved by the standardizing process. Differentiating between distinct brain areas is made simpler by the standardized image’s more consistent distribution of intensity values. Various shades of gray represent intermediate levels in the intensity values, which range from 0 (black) to 1 (white). A proper study and interpretation of the brain depend on the ability to visually discriminate between its functional and anatomical parts, which is made possible by this improved contrast.	26
3.3 With a notable peak at the lowest intensities, the histogram shows the normal distribution of the intensity values of pixels in the standardized PET brain image. The frequency of pixels gradually drops as intensity values rise, suggesting that higher intensity values are less frequent. The large range of pixel frequencies is effectively visualized by the logarithmic scale on the y-axis.	26
3.4 This image depicts the PET brain scan following adaptive histogram equalization, which results in increased contrast. The improved image exposes more distinct anatomical and functional brain regions, allowing for greater visualization and analysis.	27
3.5 The histogram has a prominent peak at low intensities and a wider spread throughout the range, showing better contrast and visibility of brain characteristics following adaptive histogram equalization.	28
3.6 Correlation matrix illustrating the connections between kurtosis, skewness, standard deviation, and mean intensity. Negative correlations are shown in blue, and strong positive correlations are shown in red. While mean intensity and skewness have a strong negative correlation (-0.92), skewness and kurtosis have a substantial positive correlation (0.97).	30

3.7 Dataset's 3D PCA plot: The data points mapped onto the first three major components are visualized in this scatter plot. Plotting the variance that each component captures aids in illuminating the linkages and underlying structure of the data in a lower dimensional space.	30
3.8 Every primary component loading is displayed in the bar plots. For PC1, skewness and kurtosis have the greatest positive loading, and kurtosis is the most influential. Skewness (negative loading) and kurtosis (positive loading) have an inverse connection that PC2 depicts. Other variables have negligible impact on PC3, which is mostly determined by mean intensity.	31
3.9 The 3D scatter plot shows the results of K-means clustering on the PCA-reduced dataset. Data points are colored according to their assigned clusters, revealing distinct groupings in the principal component space. The plot illustrates the separation of clusters based on the principal components, highlighting the inherent structure and variability within the dataset.	32
3.10 The PCA-reduced dataset's two-dimensional embedding is displayed in the t-SNE plot, where data points are colored based on their K-means cluster allocations. The display highlights the underlying linkages and groupings in the data by revealing distinct clusters and maintaining local structures.	34
3.11 The UMAP map displays the PCA-reduced dataset's two-dimensional embedding, with data points colored based on their K-means cluster assignments. The display highlights the innate linkages and groupings within the data by revealing unique clusters while maintaining local and global structures.	35
3.12 The PCA-reduced dataset's two-dimensional embedding is displayed in an Isomap plot, where data points are colored based on their K-means cluster assignments. By emphasizing discrete clusters and retaining the separations between data points, the display maintains the overall geometric structure of the data. This representation sheds light on the relationships and underlying groups within the collection, offering an understanding of its general topological structure.	36

4.1 Visualization of the label distribution: The dataset's counts for each label category (AD, CN, MCI, LMCI, and EMCI) are shown in the left bar chart, which exhibits an equal distribution with 50 cases for each label. The same distribution is shown in the pie chart on the right, which shows that each category makes up 20% of the entire dataset, indicating that the label distribution is balanced.	38
4.2 Grid of 20 preprocessed brain scan pictures: A collection of brain slices that have undergone preprocessing, like normalization, are displayed in this grid. The consistent appearance of the images makes analysis and comparison easier, which helps with jobs like medical imaging research machine learning model training and feature extraction.	39
4.3 Comparison of images at different epochs: (a) The brain images generated at epoch 100 show the early phases of the model's training, when the fundamental structural patterns of the brain are starting to take shape. However, the high levels of noise and blurring suggest that the model has not yet fully caught the intricate details of the architecture of the brain. (b) The brain images generated at epoch 300 show improved structural correctness and decreased noise, suggesting that as training goes on, the model has significantly improved its capacity to recreate the minute intricacies of brain architecture.	43
4.4 Comparison of images at different epochs: (a) Images of the brain generated at epoch 400 show additional structural and lucid refinement along with notable enhancements in brain feature representation, suggesting that the model is becoming more and more capable of producing realistic images. (b) Brain images generated at epoch 500 show reduced artifacts and great structural fidelity, indicating that the model has successfully learned to capture and duplicate the fine aspects of brain anatomy as training achieves maturity.	44
4.5 Plot showing discriminator and generator loss during training: The generator (G) and discriminator (D_real and D_fake) loss values are plotted on the graph as they increase with each repetition. As a result of the adversarial nature of GAN training, there is a notable initial dip in both losses. With time, the discriminator retains its capacity to discern between actual and created data, and the generator becomes more adept at producing realistic images, suggesting that the model is gradually getting closer together.	45

4.6 Depiction of generated images across training phases, accompanied by relevant measurements: Along with generated brain pictures at various training epochs (0 to 50), the grid also shows mean squared error (MSE), structural similarity index (SSIM), discriminator loss (D Loss), and generator loss (G Loss). The model's rising picture quality is demonstrated by the images and metrics, which show how training improves structural correctness and similarity to the target images, with dropping MSE and increasing SSIM.	49
4.7 Gradually, the clarity and realism of early visuals (Epochs 120–140) improve. But, as seen by the rise in G Loss and fall in SSIM, the generated images start to look less realistic about Epochs 320–330. Even if the image quality has somewhat improved by Epoch 340, more tweaks are need to guarantee consistent progress. . .	50
4.8 The grid displays the generated brain images at training epochs 460 to 490, together with the accompanying Mean Squared Error (MSE), Generator Loss (G Loss), Discriminator Loss (D Loss), and Structural Similarity Index (SSIM). With consistent SSIM and MSE values, these images and metrics demonstrate the stabilization of the model and show that, as the training process approaches its conclusion, the model has successfully captured the structural properties of the brain images.	51
4.9 Over 500 epochs, the training metrics plot: The Structural Similarity Index (SSIM), Mean Squared Error (MSE), Generator Loss, and Discriminator Loss progressions during the training process are displayed on the graph. An adversarial dynamic between the discriminator and generator is seen in the large oscillations in the Generator Loss, particularly around the halfway point of the training. A comparatively low discriminator loss indicates that actual and produced images may be distinguished from one another effectively. The model appears to be getting better at producing realistic and structurally accurate images as evidenced by the stabilization of MSE and SSIM measures over time.	52

5.1 Comparing the brain pictures produced by DCGAN and SRGAN When it comes to capturing the finer elements of brain anatomy, DCGAN outperforms SRGAN in terms of structural details and clarity with the images (a) it generates. DCGAN generates images that are higher quality and more realistic than SRGAN, while the SRGAN-generated images (b) are more blurry and less detailed, emphasizing an unexpected scenario.	56
6.1 A few real images from the dataset.	66
6.2 Generated DCGAN images.	67
6.3 Generated SRGAN images.	68

LIST OF TABLES

3.1	Table of intensity statistics.	29
3.2	Table of principal component analysis results.	29
3.3	Table of principal component analysis results for intensity statistics.	31
4.1	Hyperparameters for the DCGAN model.	40
4.2	Layer summary for the DCGAN generator.	41
4.3	Layer summary for the DCGAN discriminator.	42
4.4	Layer summary for the DCGAN.	42
4.5	Hyperparameters for the SRGAN model.	46
4.6	Layer summary for the SRGAN generator.	47
4.7	Layer summary for SRGAN discriminator.	48
4.8	Layer summary for the SRGAN.	48
5.1	Comparison of DCGAN and SRGAN Performance Metrics	54

1 INTRODUCTION

Generative models for synthesizing Positron Emission Tomography (PET) images are a breakthrough method at the convergence of medical imaging, computer vision, and artificial intelligence. PET imaging is an important technique in detecting and monitoring Alzheimer's disease because it offers extensive information about the metabolic processes in the brain. This dissertation focuses on the creation and implementation of generative models, specifically those designed to generate PET images, which can help with the early identification along with comprehension of Alzheimer's disease.

Generative adversarial networks (GANs), a type of generative model, is employed to generate realistic synthetic PET images. GANs, which are known for their capability to learn and reconstruct complex data distributions, are particularly effective in creating highly detailed and realistic synthetic images. The ultimate objective of this research is to develop a robust and scalable method for creating synthetic medical images that may be used to train diagnostic algorithms, perform research, and, ultimately, improve clinical outcomes in Alzheimer's disease.

1.1 Background and Context

Dementia refers to a group of neurological conditions that cause a progressive and long-term deterioration in cognitive functions and memory, significantly interfering with daily activities. Alzheimer's disease, the most common type of dementia, is characterized by the accumulation of amyloid plaques and tau tangles in the brain. Early and accurate Alzheimer's diagnosis is critical for optimal management and treatment. Medical imaging is essential in this situation, particularly Positron Emission Tomography (PET). Since the introduction of early imaging technologies like X-rays and MRI, the area of medical imaging has experienced considerable improvements. The 1970s saw the development of PET, which allowed for the viewing of bodily metabolic processes, most notably those occurring in the brain thus revolutionizing medical diagnosis. This capacity has proven to be extremely helpful in the investigation and identification of neurological disorders, such as Alzheimer's disease.

PET imaging is useful, but it has drawbacks such as high costs, scarcity, and the use for radioactive tracers. Large-scale research and the creation of reliable diagnostic instruments are

hampered by these restrictions. These problems have seen encouraging resolutions in recent years with the introduction of artificial intelligence (AI) and machine learning into medical imaging. Since their introduction by Ian Goodfellow and others in 2014, Generative Adversarial Networks (GANs) have demonstrated extraordinary potential in producing synthetic data that closely resembles real-world datasets.

In order to improve the visual appeal of the generated images over time, GANs work by competing two neural networks: a discriminator that assesses the authenticity of the created data and a generator that generates synthetic data. Applications for this novel method have been found in several fields, such as data augmentation, image enhancement, and medical image synthesis. GANs can be used to produce high-quality synthetic PET pictures in the setting of Alzheimer's disease, offering a scalable and affordable substitute for conventional imaging techniques.

The pressing need to overcome the limits of conventional PET imaging drives fresh research in this field. The ability to make synthetic PET images can greatly improve diagnostic algorithm training, allow for broader Alzheimer's disease research, and, ultimately, improve clinical outcomes. This dissertation aims to help enhance diagnostic techniques and therapeutic solutions for Alzheimer's disease by addressing the limited availability of PET imaging data and lowering reliance on expensive procedures. The particular goals of this work, together with the techniques used and anticipated contributions to the area, are described in the sections that follow.

1.2 Objectives

- Design and implement GAN architecture suitable for synthesizing PET images.
- Adjust model hyperparameters for better PET picture synthesis and performance.
- Analyse synthetic picture quality qualitatively using metrics such as Mean Squared Error (MSE) and Structural Similarity Index (SSIM).
- Compare the produced PET images to actual PET scans to confirm their clinical applicability in depicting the metabolic pathways linked to AD pathogenesis.

1.3 Achievements

In this dissertation, a great deal of progress has been made in the synthesis of PET images for Alzheimer's disease study using Generative Adversarial Networks (GANs). Super-Resolution

GAN (SRGAN) and Deep Convolutional GAN (DCGAN), two GAN designs, were successfully implemented and optimized, which was one of the major accomplishments. GANs have the potential to be used in medical imaging; via much experimentation, DCGAN was refined to generate high-quality synthetic PET pictures that closely mimic genuine scans. Furthermore, the research offered an extensive comparative evaluation of various models, emphasizing the difficulties and achievements associated with each architecture. By incorporating manifold learning techniques such as PCA, t-SNE, and UMAP, the data analysis was further refined, facilitating a more profound comprehension of the dataset and enhancing the generative models' performance. Future developments in the utilization of synthetic data to assist in the diagnosis and study of Alzheimer's disease are made possible by these accomplishments.

1.4 Overview of Dissertation

- **Chapter 1** explains the study's background, context, and goals, laying the groundwork for the investigation by talking about the significance of PET imaging in Alzheimer's disease and the possibility of GANs to get around the drawbacks of conventional imaging techniques.
- **Chapter 2** gives an in-depth review of the literature, including a summary of medical imaging techniques and the theoretical underpinnings of dementia and Alzheimer's disease. The ideas of Generative Adversarial Networks (GANs) and manifold learning, which are essential to the research, are also introduced in this chapter.
- **Chapter 3** focuses on the study's dataset and describes the manifold learning and data augmentation strategies utilized to evaluate and prepare the PET pictures. The foundation for comprehending the data preparation process for the GAN models is laid out in this chapter.
- **Chapter 4** explains how the GAN designs are put into practice, covering DCGAN and SRGAN design, training, and optimization. The technical difficulties faced and the solutions created to enhance model performance are highlighted in this chapter.
- **Chapter 5** delivers a comparison of the effectiveness of SRGAN and DCGAN, evaluating both models' efficacy and image quality using a range of evaluation parameters. Future research directions are informed by this analysis, which offers important insights into the advantages and disadvantages of each model.

- **Chapter 6** summarises the main conclusions, assesses the accomplishments in relation to the original goals, and suggests possible directions for further research to close the dissertation. The research's ramifications for the larger area of medical imaging are also discussed in this chapter.

2 BACKGROUND THEORY AND LITERATURE REVIEW

2.1 World of Dementia

Dementia has a long history that predates humankind. The comprehension of the epidemiology, pathophysiology, and diagnosis of Alzheimer's disease (AD) and related illnesses has significantly improved recently, and the nosology of these disorders is being closely examined. The word dementia is used to describe a wide range of cognitive impairments that are severe enough to interfere with day-to-day functioning, including memory, language, problem-solving, and other cognitive functions. This deterioration has an effect on daily functioning, mental abilities, behavior, emotions, and interpersonal interactions. Dementia is not a single disease; rather, it is a collection of symptoms spurred on by a number of illnesses. Memory loss, cognitive, linguistic, and problem-solving impairments, as well as behavioral, mood, and perception abnormalities, are typical symptoms.[33]

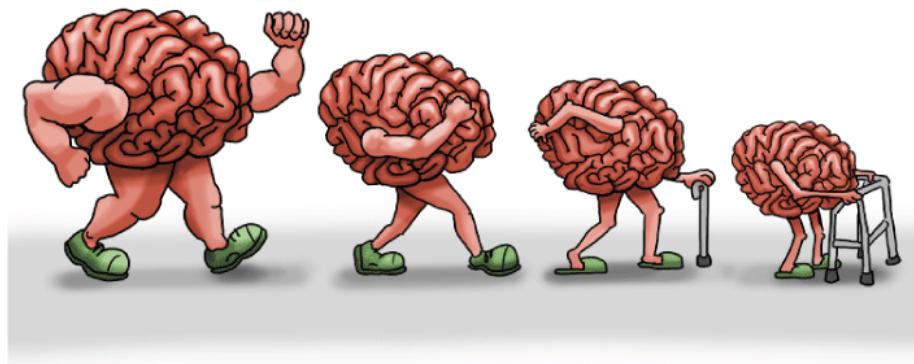


Figure 2.1: An example of how dementia progresses, showing a steady fall from cognitive function to severe impairment .[19]

Although Alzheimer's disease is the most prevalent cause of dementia, frontotemporal dementia, Lewy body dementia, and vascular dementia are all possible causes. Age is one of the main risk factors for dementia; those over 65 are most likely to get the illness, and genetics and family history can further raise one's risk. Significant roles are also played by cardiovascular variables such as excessive cholesterol and hypertension. Additional lifestyle factors that impact the risk include nutrition, physical activity, alcohol consumption, and smoking. Furthermore, there is a correlation between a higher risk of dementia and other illnesses including diabetes and depression.[34]

2.2 Alzheimer Disease

The primary cause of dementia in older persons is Alzheimer's disease (AD), a neurological illness that progresses over time. It is related to the accumulation of tau neurofibrillary tangles and amyloid-beta plaques in the brain, which impair neuronal function and cause cell death.[35] The hippocampus, which is important for memory, is often where this deterioration begins and then progresses throughout the brain, causing a large amount of atrophy. Gradually, the illness worsens to the point of significant cognitive impairments and behavioral changes from minor memory loss and disorientation. People eventually lose their capacity to carry out daily duties and need full-time assistance. The majority of Alzheimer's cases involve those over 65, although younger people can also develop early-onset Alzheimer's, frequently as a result of genetic factors.[34][47]

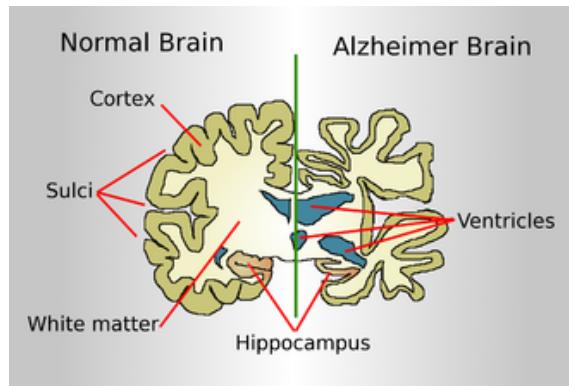


Figure 2.2: A comparative analysis of the major anatomical variations between an Alzheimer's disease and a normal brain. In the Alzheimer's brain, notable alterations include considerable shrinkage of the hippocampus and white matter, enlarged ventricles, widening sulci, and cortical thinning. These changes show how Alzheimer's disease profoundly affects brain structure, resulting in memory loss and cognitive decline.[47]

Current therapies, such as NMDA receptor antagonists and acetylcholinesterase inhibitors, only temporarily relieve symptoms; they do not stop the disease from progressing. New therapy modalities, such as brain stimulation methods, neuroprotective drugs, and anti-amyloid and anti-tau therapies, are being investigated through ongoing research. Improvements in these fields are meant to provide better early detection and more efficient illness management.

2.2.1 Stages of Alzheimer disease

Alzheimer's disease progresses through several phases, each of which is distinguished by a decline in cognitive and functional abilities.

2.2.1.1 Cognitively Normal

The cognitively normal (CN) stage is characterized by the absence of memory problems or cognitive deterioration. Their capacity to absorb and retain new information is unaffected, and they carry out daily tasks with ease. Planning, organizing, and making decisions are all examples of executive functions that are performed with ease. Speaking, word retrieval, and comprehension of written and spoken language are among the language skills that are still strong. There is no impact on visuospatial skills, which include navigating and interpreting visual information. Social exchanges and activities are commonplace. There are no abnormalities found on standard cognitive exams, and there are no tau tangles or amyloid plaques, two biomarkers of Alzheimer's or other dementia's.[20]

2.2.1.2 Early Moderate Cognitive Impairment

Early moderate cognitive impairment (EMCI) is a stage in which people begin to experience minor cognitive impairments that are more noticeable than normal changes due to aging but do not substantially interfere with day-to-day functioning.[3] Errors in memory become more frequent, especially when they relate to past occurrences or newly acquired knowledge. There could be little difficulties with assignment planning and organization, as well as sporadic difficulties with word or name finding. With intricate tasks like reading maps, visuospatial skills may exhibit minor difficulties. Though there may be subtle variations, social interactions normally continue as before despite these modifications. Extensive cognitive testing can detect subtle deficiencies, and memory problems are frequently reported by people. Cerebrospinal fluid or PET scans can identify early indicators of tau or amyloid pathology in the brain. [4][3][20]

2.2.1.3 Mild Cognitive Impairment

Significant and quantifiable decreases in cognitive abilities, such as memory and thinking skills, are experienced by people with moderate cognitive impairment (MCI). Names, dates, and recent events are especially prone to frequent memory gaps. Planning, organizing, and finishing complex tasks are noticeably harder. Deterioration in language abilities is accompanied by difficulties finding words and following conversations.[4] When visual-spatial abilities deteriorate, driving and item assembly become more challenging. Close friends and family can observe these changes in cognition. Although these alterations are not yet incapacitating, cognitive tests show significant impairment. Amyloid plaques and tau tangles, two biomarkers for Alzheimer's disease, become more noticeable.[4][3][20]

2.2.1.4 Late Mild Cognitive Impairment

People who are nearing early dementia have severe cognitive deterioration in the late moderate cognitive impairment (LMCI) stage. There is a noticeable rise in memory loss and disorientation, as well as trouble remembering names, faces, and recent events.[3] Language abilities continue to degrade, which causes frequent word finding issues and decreased fluency, making it more difficult to follow complex conversations. Significant deterioration of visual-spatial skills leads to difficulties navigating familiar environments and an increased risk of becoming lost.[4] There are noticeable shifts in conduct and personality, which frequently result in social disengagement. Significant impairment is found in cognitive evaluations, which has a major influence on day-to-day functioning. Imaging or testing on cerebrospinal fluid (CSF) can confirm substantial amounts of Alzheimer's pathology in the brain.[4][3][20]

2.3 Diverse Techniques in Medical Imaging

Medical imaging greatly enhances disease diagnosis, monitoring, and therapy by using a variety of advanced technologies for observing the body's internal structures and operations. These diverse approaches offer unique viewpoints on various tissue kinds and physiological processes, making accurate and non-invasive patient exams possible. Healthcare professionals can obtain precise and comprehensive data by combining the benefits of several imaging modalities. This data is crucial for improving patient outcomes and customizing treatment regimens. Desirable imaging methods include CT, MRI and PET.

2.3.1 Computed Tomography(CT) Scanning

A revolving narrow beam of X-rays is used in computed tomography (CT) imaging to produce numerous cross-sectional images of the body. After being analyzed by a computer, these pictures yield finely defined slices that can be computationally merged to create a three-dimensional representation of the interior organs.[27] Compared to traditional X-rays, this technology offers more comprehensive information, making it easier to identify cancers, abnormalities, and other diseases. The precise diagnostic information provided by CT scans frequently surpasses the possible hazards, even with the little risk of ionizing radiation exposure, especially when diagnosing serious medical conditions.[27][7]

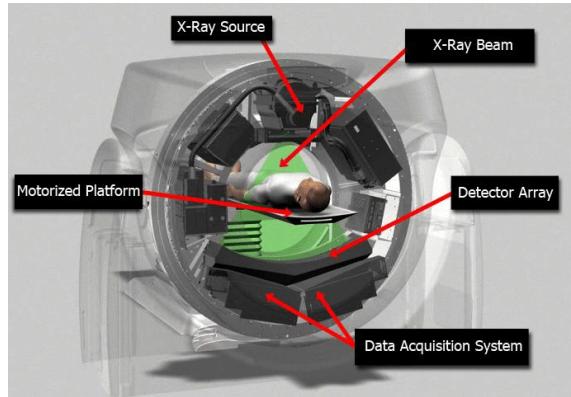


Figure 2.3: An illustration of a CT scanner's main parts, including the motorized platform, detector array, X-ray source, X-ray beam, and data gathering system. The patient is exposed to beams from the X-ray source, and the detector array gathers the rays that are transmitted. With the use of this data, the data gathering system generates intricate cross-sectional photographs of the body, which are necessary for precise medical diagnosis.[2]

2.3.2 Magnetic Resonance Imaging(MRI)

Soft tissues, in particular, can be precisely imaged inside the body using a non-invasive method called magnetic resonance imaging (MRI). Hydrogen protons within the body are aligned by a powerful magnetic field it produces.[49] The protons undergo displacement and emission of signals upon realignment in response to radiofrequency pulses. The MRI system records and interprets these signals to create images. Neurological diseases like Alzheimer's disease are especially beneficial to diagnose and track with MRI.[21] The hippocampus and other regions linked to memory and cognitive function are particularly affected, as are other structural alterations and atrophy in the brain. By combining machine learning with sophisticated MRI methods, Alzheimer's disease can be identified and classified earlier, which allows for prompt management and therapy.[49][21]

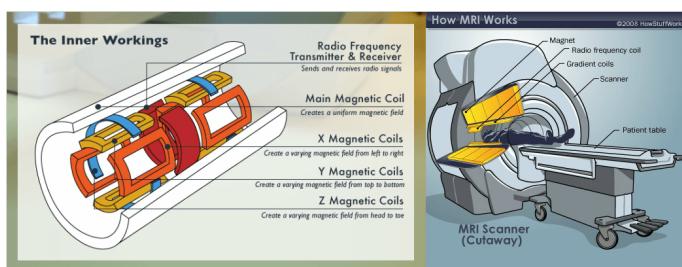


Figure 2.4: An illustration of an MRI scanner highlighting the main parts and how they work. The X, Y, and Z magnetic coils change the magnetic field in different directions, but the main magnetic coil produces a powerful, homogeneous magnetic field. The body's internal architecture can be seen in great detail because to the processing of signals that are sent and detected by the radio frequency transmitter and receiver.[31]

2.3.3 Positron Emission Tomography(PET)

The imaging method known as Positron Emission Tomography (PET) uses gamma rays from a radiotracer that has been put into the circulation to detect metabolic activities. When this radiotracer builds up in tissues that are metabolically active, PET scans can provide finely detailed images. Despite its high cost and low radiation exposure, PET is essential for the diagnosis and monitoring of heart ailments, neurological disorders such as Alzheimer's disease, and malignancies. It also provides unique insights into the metabolic activities of the body.[36]

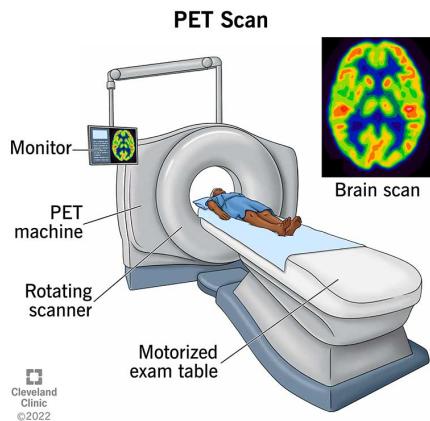


Figure 2.5: Figure illustrating the main parts of a PET scan machine: the motorized exam table moves the patient for scanning, and the PET machine, with its spinning scanner, detects gamma rays released by a radiotracer in the body. The resultant brain scan, which shows metabolic activity in various brain regions and is essential for identifying diseases including cancer, neurological problems, and heart disease, is displayed on the monitor.[8]

2.4 Unveiling PET

2.4.1 Understanding Positron Emission Tomography

Positron emission tomography (PET) is a nuclear medicine treatment that measures the metabolic activity of cells in bodily tissues.[32] It uses nuclear medicine and biochemical analysis to visualize biochemical changes in the body. It is generally utilized for patients with brain or heart disorders, as well as cancer. During the operation, a little amount of a radioactive chemical, known as a radiopharmaceutical or radioactive tracer, is injected to aid analyze the targeted tissue.[32][23]

PET can detect biochemical changes in organs or tissues, allowing for the early detection of disease processes before anatomical alterations are detectable with other imaging techniques such as computed tomography (CT) or magnetic resonance imaging. Gamma camera systems, when

converted for PET procedures, can finish scans faster and at a cheaper cost than typical PET scans, increasing the scope of PET imaging.[32][23]

Furthermore, PET assesses physiological processes that include blood flow, metabolism, neurotransmitters, and radiolabeled medicines, allowing for quantitative examinations of relative variations over time in response to disease progression.[23] The tracer is often injected intravenously, labeled with isotopes such as oxygen-15, fluorine-18, carbon-11, or nitrogen-13, and has a radioactive dose comparable to that used in CT scans.[32] PET generally evaluates the rate of glucose intake in various bodily areas using the radiolabeled glucose analogue 18-fluorodeoxyglucose (FDG). Clinically, this aids in distinguishing between benign and malignant tumors, as malignant tumors metabolize glucose more rapidly.[32][23]

2.4.2 The Science Behind Positron Emission Tomography

PET is effective by using a scanning device with a large central opening to detect photons released by a radionuclide within the target organ or tissue.[23] These radionuclides are generated by connecting a radioactive atom to components that the organ or tissue would naturally consume throughout its metabolic processes. The radionuclide is injected via an intravenous (IV) line and enters a vein, after which the PET scanner glides gently across the area under examination.[36]



Figure 2.6: An illustration showing the intravenous injection of a radioactive tracer to identify the body's areas of metabolic activity during a PET-CT scan. Positioned on a table, the patient passes through the scanner, which combines CT and PET imaging to provide detailed images that are displayed on a monitor. For the detection and assessment of diseases including cancer, heart disease, and neurological disorders, this method is essential.[36]

As the radionuclide decays, it generates positrons that clash with electrons, resulting in annihilation photons. These photons, caught by the scanner's 180-degree apart sensors, are evaluated by a computer to produce a visual map of an organ or tissue. The radiance of the tissue on the image correlates to the amount of radioactivity absorbed, which reflects the organ or tissue's level of function.[32]

2.4.3 Vital Role of Positron Emission Tomography in Medical Imaging

PET scans are crucial for the diagnosis and treatment of dementia, particularly Alzheimer's disease. PET scans provide an image of the metabolic activity of the brain by using radioactive tracers such as fluorodeoxyglucose (FDG).[23] These scans frequently show reduced glucose metabolism in specific regions associated with Alzheimer's disease, such as the temporoparietal and posterior cingulate cortices—regions critical to memory and cognitive function. One important sign of Alzheimer's is a decrease in metabolism, which can be identified early in the illness, even before the disease's main symptoms manifest. As a result, PET imaging plays a crucial role in distinguishing Alzheimer's from other dementias and validating diagnosis in cases where clinical symptoms are not evident. [32]

2.4.4 Fluorodeoxyglucose Positron Emission Tomography

A vital component of our bodies' operation is glucose. It is a reactant in cellular respiration, which produces the necessary energy for our bodies, along with oxygen (O_2). A fluorine-18 atom is used to replace one of the hydroxyl groups (an oxygen atom bound to a hydrogen atom) in the modified glucose molecule known as fluorodeoxyglucose (FDG).[23]

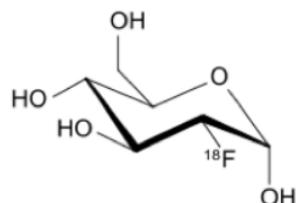


Figure 2.7: Fluorodeoxyglucose (FDG) molecular diagram: Utilized in positron emission tomography (PET) imaging, FDG is a glucose analog that is radiolabeled. Fluorine-18, a radioactive isotope, takes the place of a hydroxyl group in FDG, making it comparable to glucose. PET scans that show metabolic activity can show the accumulation of FDG after injection in regions like the brain that have significant glucose absorption.[12]

FDG PET imaging is critical for detecting metabolic abnormalities in the brain linked with Alzheimer's disease, which is a common form of dementia. This method uses fluorodeoxyglucose (FDG) to evaluate glucose metabolism, with scans often revealing decreased glucose metabolism in brain regions such as the temporoparietal and posterior cingulate cortices, which are critical for memory and cognitive function.[12] These metabolic deficiencies can be seen early in the disease's course, even before substantial symptoms appear, making FDG PET useful for diagnosing and differentiating Alzheimer's from other dementias. FDG PET scans help with accurate diagnosis,

illness progression monitoring, and therapy success by giving precise images of brain metabolism. Using this metabolic imaging method, physicians can better control patients' conditions overall by customizing treatment programs based on each patient's unique metabolic anomalies. Moreover, FDG PET can be used with other imaging modalities, including MRI, to provide a more complete picture of brain health, enhancing its ability to diagnose and predict Alzheimer's disease.[44]

2.5 Manifold Learning

In the fields of machine learning and statistics, manifold learning is an advanced algorithmic method created especially for dimensionality reduction. Finding and representing the underlying low-dimensional structure of high-dimensional data is the main objective of manifold learning.[17] This method can be broadly divided into two categories: linear and nonlinear. Principal Component Analysis (PCA) and Multidimensional Scaling (MDS) are examples of linear methods that have been used for a long time in statistical analysis. The primary components, or initial few coordinates, are where the largest deviations by any projection of the data exist when the data is transformed into a new coordinate system using PCA. MDS, on the other hand, seeks to precisely retain the between-object distances by positioning each object in a low-dimensional space.[50]

Recent years have seen a considerable increase in the use of nonlinear manifold learning techniques, such as Diffusion Maps, Laplacian Eigenmaps, Isomaps, Local Linear Embedding (LLE), and Hessian Eigenmaps. In order to embed high-dimensional points into a lower-dimensional space, these approaches usually include three steps: establishing distances or affinities between points, solving an eigenproblem, and performing a nearest-neighbor search. These methods are especially useful for identifying intricate patterns in data that are missed by linear approaches, which leads to more insightful and efficient data analysis.[50]

2.5.1 Linear

2.5.1.1 Principal Component Analysis(PCA)

In several scientific fields, including medical imaging, Principal Component Analysis (PCA) is a widely used method for reducing the dimensionality of huge datasets while preserving as much variance as feasible. When working with high-dimensional data—like neuroimaging datasets—where there are a lot of variables and observations, this strategy is quite helpful.[22] PCA reduces the number of dimensions in the dataset, which simplifies the data and makes it easier to view and interpret. It also removes redundancy and filters out noise.[46] The method reorients the

initial data into a new coordinate system, in which a direction of maximum variance corresponds to each axis, known as a major component. The most notable variance in the data is captured by the first principle component, which is followed by the second principal component, which captures the next highest variance in a direction perpendicular to the first. This procedure produces a set of uncorrelated components that provide a representation of the original data in a lower dimension.[46]

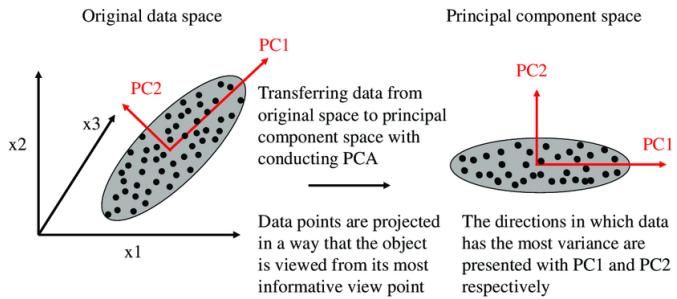


Figure 2.8: Data transformation from the original data space to the principle component space is illustrated in this principle Component Analysis (PCA) illustration. In order to project the data points along the directions of largest variation in the data, the principal components (PC1 and PC2) are used. By lowering the number of dimensions in the data, this method preserves the most important characteristics while simplifying the data and making it simpler to see and analyze.[5]

PCA's mathematical foundation consists of computing the data's covariance matrix first, then eigenvalues and eigenvectors. The amount of variation that each component captures is shown by the eigenvalues, whilst the eigenvectors reflect the primary components.[22] The original data's dimensionality is successfully decreased by projecting it onto these primary components. Since they mostly explain the volatility in the data, only the initial few primary components are typically retained. This reduction in dimensionality makes it easier to store data, removes collinearity, and simplifies visualization—all of which are particularly helpful when working with complicated datasets like those used in medical imaging.[46][22]

PCA has been used in medical imaging for a variety of purposes, including feature extraction, noise reduction, and picture compression. For instance, PCA can be used to remove noise from MRI or PET scans by keeping only the elements that significantly vary, improving the image quality.[46] Additionally, PCA makes it possible to extract the most important features from these pictures, which can enhance analysis and diagnostic precision. Furthermore, exploratory data analysis is made easier by PCA's ability to reduce the dimensionality of the data, which makes it easier to see in two or three dimensions.[22]

2.5.2 Non-Linear

2.5.2.1 Uniform Manifold Approximation and Projection(UMAP)

The manifold learning technique known as Uniform Manifold Approximation and Projection (UMAP) is widely used for dimensionality reduction and is well-known for its capacity to preserve both local and global structures in data. This property makes UMAP particularly helpful in fields where maintaining the intrinsic linkages in high-dimensional datasets while reducing their complexity is crucial, such as bioinformatics and medical imaging. UMAP detects patterns beneath higher-dimensional areas that may remain obscured by handling massive datasets effectively.[28]

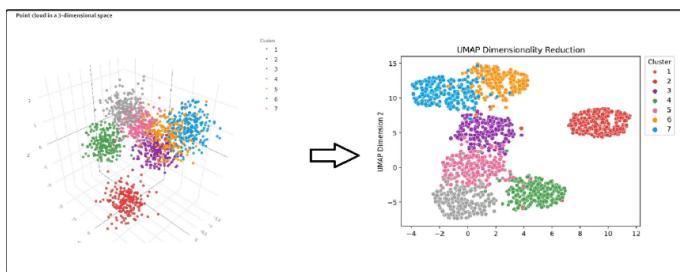


Figure 2.9: An example of dimensionality reduction with UMAP is seen in the left image, which displays a point cloud in three dimensions that represents various data point clusters. The data is reduced to a 2-dimensional space while maintaining the structure and linkages between clusters through the use of UMAP (Uniform Manifold Approximation and Projection), as shown in the image on the right. By improving data presentation and interpretation, this method increases accessibility to complicated, high-dimensional data.[29]

UMAP is based on a solid mathematical basis, which begins with the high-dimensional data being represented in a fuzzy topological manner. In this method, a graph is constructed with edges representing the likelihood of linkages between data points depending on their separations from one another.[28] This graph preserves the integrity of the data during dimensionality reduction by capturing the local structure of the data properly and accounting for global relationships. Stochastic gradient descent (SGD), a method well-suited for handling massive amounts of data, is then used by the algorithm to optimize this high-dimensional graph into a lower-dimensional space. As a result, the original data's structure is accurately represented in a low-dimensional embedding, which facilitates visualization and analysis.[28]

High-Dimensional Similarity Calculation:

$$v_{j|i} = \exp\left(\frac{-d(x_i, x_j) - \rho_i}{\sigma_i}\right)$$

where $d(x_i, x_j)$ is the distance between data points x_i and x_j , ρ_i is the distance to the nearest neighbor of i , and σ_i is a normalizing factor.[28]

Symmetrization of Fuzzy Set Union:

$$v_{ij} = (v_{j|i} + v_{i|j}) - v_{j|i} \cdot v_{i|j}$$

This equation is used to symmetrize the fuzzy set membership in UMAP.[28]

Low-Dimensional Similarities:

$$w_{ij} = \left(1 + a\|y_i - y_j\|^{2b}\right)^{-1}$$

Here, a and b are user-defined positive values, and $\|y_i - y_j\|$ represents the distance between points y_i and y_j in the low-dimensional space. [28][50]

The ability of UMAP to reduce dimensionality while preserving data structures has important applications in medical imaging. For instance, it is essential to preserve both local variations and global patterns in PET imaging for Alzheimer's disease in order to accurately diagnose the condition and track its course. Because of its efficiency, UMAP can process enormous amounts of imaging data quickly, which makes it a useful tool for academics working with complicated datasets. Furthermore, researchers can fine-tune the embedding process with its adjustable parameter adjustment, guaranteeing that the most pertinent elements of the data are highlighted.[28]

2.5.2.2 t-Distributed Stochastic Neighbor Embedding(t-SNE)

For dimensionality reduction, t-Distributed Stochastic Neighbor Embedding (t-SNE) is a well-liked manifold learning technique that works especially well for displaying high-dimensional datasets. Its exceptional ability to maintain the local structure of data makes it an invaluable instrument for identifying intricate patterns in data-intensive domains such as neuroimaging. When

applied to medical situations, such as PET imaging for Alzheimer's disease, t-SNE can help researchers make sense of the underlying structures in high-dimensional data, which may lead to better methods to diagnosis and treatment.[41]

To begin, the t-SNE algorithm converts the pairwise similarities in the high-dimensional space between data points into conditional probabilities. These probabilities show the chance that a point i would choose a point j as its neighbor, with the expectation that points that are closer together will likely be neighbors more frequently. In terms of math, this relationship is represented as:

$$p_{j|i} = \frac{\exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma_i^2}\right)}{\sum_{k \neq i} \exp\left(-\frac{\|x_i - x_k\|^2}{2\sigma_i^2}\right)},$$

A variance term that accounts for the local density of the data points is the parameter σ_i . This computation makes sure that the data's local structure is appropriately reflected, which is crucial for the t-SNE process's subsequent phases.[41]

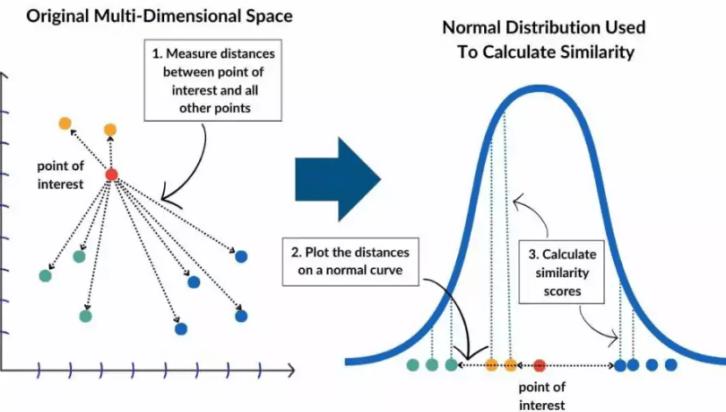


Figure 2.10: An example of the t-Distributed Stochastic Neighbor Embedding (t-SNE) method is displayed on the left, where distances are measured from a location of interest to every other point in the original high-dimensional space. To determine similarity scores, these distances are then displayed on a normal distribution curve (right side). t-SNE preserves the local structure and facilitates the visualization of the relationships between data points by projecting these similarities onto a lower-dimensional space.[43]

Following the computation of the probabilities in the high-dimensional space, t-SNE uses a t-distribution to project these probabilities into a lower-dimensional environment, typically two or three dimensions. Because of its larger tails, which serve to disperse the data points in the

lower-dimensional space and lessen the "crowding problem" that frequently occurs in other dimensionality reduction techniques, this distribution is preferred over a Gaussian distribution. In the lower-dimensional space, the likelihood that point i will choose point j as a neighbor is provided by:

$$q_{j|i} = \frac{(1 + \|y_i - y_j\|^2)^{-1}}{\sum_{k \neq l} (1 + \|y_k - y_l\|^2)^{-1}}.$$

By highlighting the distinctions between adjacent points in the lower-dimensional space, this formulation allows t-SNE to preserve both local and global structures in an equilibrium[41]. The Kullback-Leibler (KL) divergence between the high-dimensional probability distribution P and the low-dimensional probability distribution Q is then minimized by the t-SNE method to maximize the placement of points in the lower-dimensional space. The following is the goal function that t-SNE seeks to minimize:

$$C = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}},$$

where the joint probabilities generated are represented, respectively, by p_{ij} and q_{ij} . Gradient descent is typically used for this reduction, in which the locations of the points in the low-dimensional space are repeatedly changed to best preserve the relationships seen in the high-dimensional space.[41]

The significance of t-SNE in medical imaging stems from its ability to lower the dimensionality of big datasets while maintaining important local correlations. For example, by classifying PET pictures based on the various phases of the disease or by highlighting minute changes in brain activity, t-SNE can be used to visualize and evaluate the data related to Alzheimer's disease. By providing insightful information on the course of the disease, these clusters may make it possible to diagnose patients earlier and provide more individualized treatment plans.[41]

2.5.2.3 Isometric Mapping(Isomap)

The term "Isomap," which stands for "Isometric Mapping," refers to a well-known nonlinear dimensionality reduction method that applies MDS concepts to nonlinear manifolds.[40] There are three primary phases in the isomap algorithm. First, it determines the pairwise Euclidean distances

between each data point in order to create a neighborhood graph. Next, a weighted graph is created with edges representing the distances between surrounding points, connecting each point to its nearest neighbors based on a predetermined radius (ϵ) or a fixed number of neighbors (K).[40][24]

Finding the shortest paths within this graph is the second stage in estimating the geodesic distances between every pair of points on the manifold. Usually, Dijkstra's algorithm or Floyd-Warshall algorithm are used for this.[40] The matrix of these geodesic distances is then subjected to classical MDS in order to create a low-dimensional embedding that maintains the intrinsic geometry of the manifold. This entails calculating the geodesic distance matrix's eigenvalues and eigenvectors, which yield the coordinates in the reduced space.[24]

Isomap's computing load may become unaffordable for larger datasets. This problem is solved by Landmark Isomap, which computes geodesic distances solely between a selection of landmark points and all other points.[24] This efficiently approximates the overall manifold structure while reducing the computational cost to a great extent.[48]

Among its many advantages is the ability to maintain global geometric qualities, which makes isomap appropriate for data on a smooth, convex manifold. Nevertheless, it is not without flaws. Unless techniques such as Landmark Isomap are used, Isomap struggles with very big datasets and performs badly on manifolds with holes or non-convex shapes. Furthermore, the neighborhood size parameter that is selected has a substantial impact on the quality of the resulting embedding and affects the performance of the system.[48]

2.6 Generative Adversarial Networks

Generative Adversarial Networks (GANs) are generative modeling techniques that are widely used for image production. In order to produce fresh samples that closely resemble the initial training dataset, this kind of modeling is an unsupervised machine learning technique that finds patterns and contextual aspects in input data.[14]

Mathematical Foundation of GANs

A GAN is made up of two primary parts:

A neural network called generator G is used to create a synthetic data sample called $G(z;\theta_g)$ from a noise vector z that was sampled from a previous noise distribution $p_z(z)$. In order to make

the generated samples as near to the real data as feasible, the generator's goal is to understand the mapping from the noise distribution to the data distribution.[14]

A neural network called the discriminator D receives a data sample as input (which can come from the generator G or the real data distribution $p_{\text{data}}(z)$), and it outputs a probability $D(x)$ that indicates the chance that the input sample x comes from the real data rather than the generator. Being able to accurately discriminate between created and genuine data is the aim of the discriminator.[14]

GAN training can be structured as a two-player minimax game, in which the discriminator seeks to maximize its accuracy in differentiating generated from real data, and the generator seeks to minimize its loss. The function $V(D, G)$ for the GAN is given as:[14]

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))]$$

The expression $\mathbb{E}_{x \sim p_{\text{data}}(x)} [\log D(x)]$ is the log-likelihood of the discriminator accurately determining the actual data distribution from real samples, while $\mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))]$ is the log-likelihood that the discriminator will incorrectly recognize produced samples as genuine.[14]

Iteratively updating both the generator and the discriminator is how GAN training works:[14]

Update the discriminator D for a fixed generator G by optimizing the chance that generated samples will be categorized as fake and actual samples as real. The loss function (L_D) of the discriminator is provided by:[14]

$$L_D = -\mathbb{E}_{x \sim p_{\text{data}}(x)} [\log D(x)] - \mathbb{E}_{z \sim p_z(z)} [\log(1 - D(G(z)))]$$

Update the discriminator by ascending its stochastic gradient:[14]

$$\nabla_{\theta_d} \frac{1}{m} \sum_{i=1}^m \left[\log D(x^{(i)}) + \log(1 - D(G(z^{(i)}))) \right]$$

Update generator G for a fixed discriminator D so as to reduce the likelihood that the samples produced will be labeled as fraudulent. The formula for the generator's loss function, L_G , is as

follows:[14]

$$L_G = -\mathbb{E}_{z \sim p_z(z)} [\log D(G(z))]$$

2.6.1 Deep Convolutional Generative Adversarial Network(DCGAN)

The Deep Convolutional GAN (DC-GAN), a popular kind of GAN used in picture synthesis, is made up of two separate Convolutional Neural Networks (CNNs) that were trained concurrently for distinct objectives. The first CNN, known as the generator, generates new images, whilst the second, known as the discriminator, distinguishes between genuine and synthetic images. These two networks compete in a minimax game in which the discriminator constantly improves its capacity to distinguish between fake and real images, while the generator refines its ability to make realistic images.[14] In order to trick the discriminator, the generator modifies its settings with each iteration during training. When the discriminator is unable to distinguish between artificial and genuine images, the generator is said to be operating at peak efficiency. On the other hand, the discriminator likewise stops getting better if the generator is unable to produce images that are convincing. The strength of both networks as well as the caliber and volume of training data determine how realistic the generated images are; the result might be artificial images that are visually identical to real ones.[14][13]

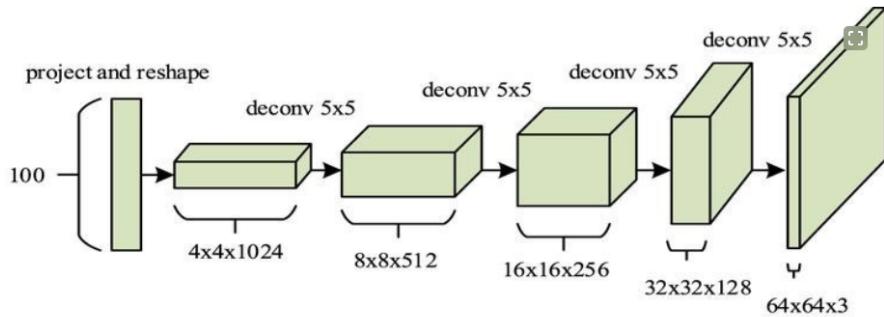


Figure 2.11: An illustration of the generator architecture of a DCGAN (Deep Convolutional Generative Adversarial Network): An input vector that is 100 dimensions in size is first projected and reshaped to begin the process. The data is then incrementally upsampled by the network as it goes through a sequence of deconvolutional (deconv) layers with increasing spatial dimensions (4x4x1024 to 64x64x3). After applying a 5x5 filter to every deconvolution layer, a 64x64 image with three color channels is produced. The creation of realistic images from random noise in GAN models relies heavily on this architecture.[9]

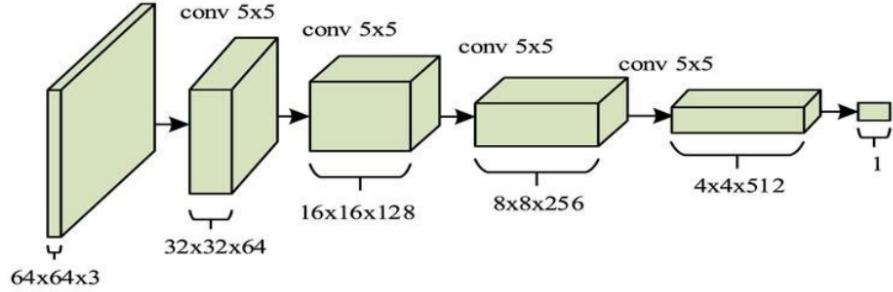


Figure 2.12: An example of a discriminator architecture using a DCGAN (Deep Convolutional Generative Adversarial Network): A $64 \times 64 \times 3$ image is entered at the start of the process and is routed via several convolutional (conv) layers. From $32 \times 32 \times 64$ to $4 \times 4 \times 512$, each layer decreases the spatial dimensions while increasing the depth. Subsequently, the network reduces the final feature map to a single output value, which indicates the likelihood that the input image is real instead of contrived. In order to differentiate between authentic and fraudulent images in GAN models, this architecture is necessary.[9]

2.6.2 Super Resolution Generative Adversarial Network(SRGAN)

A complex deep learning model called Super-Resolution Generative Adversarial Network (SRGAN) aims to improve image resolution by transforming low-resolution (LR) inputs into high-resolution (HR) outputs. Based on Generative Adversarial Networks (GANs), the SRGAN architecture consists of two primary parts: the discriminator and the generator. These networks function in an adversarial manner, with the discriminator attempting to distinguish between created and actual HR images and the generator attempting to produce convincing HR images.

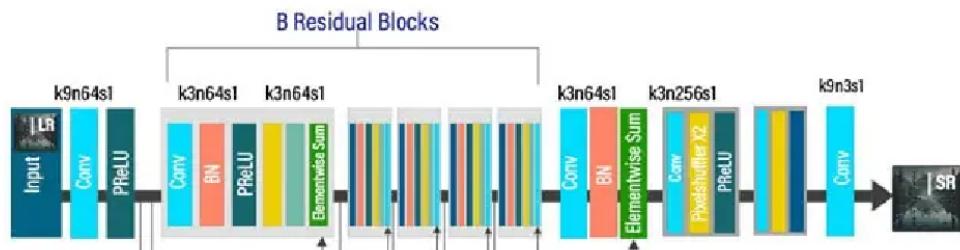


Figure 2.13: Architecture of the SRGAN generator: A low-resolution input is fed into the first convolutional layer of the process, which is activated by PReLU. After that, it processes a number of residual blocks (B blocks) that have PReLU, batch normalization, and convolutional layers in them. The network employs pixel shufflers for upsampling and a final convolutional layer to produce the high-quality output image, improving image details and resolution, following residual learning.[11]

Upscaling low-resolution photos into high-resolution counterparts is the generator's job in SRGAN. Remaining blocks are incorporated into this deep convolutional neural network (CNN) component, which is crucial for identifying and preserving high-frequency information in the image. Convolutional layers make up these residual blocks, which are then followed by batch normalization and activation functions—typically ReLU—that help learn the conversion from LR to HR while maintaining image quality. These blocks' skip connections facilitate training and enhance the model's capacity to pick up complex information. The generator employs upsampling layers, which are often implemented as sub-pixel convolution layers and rearrange pixels to form a bigger, higher-resolution image, to improve image resolution.[30]

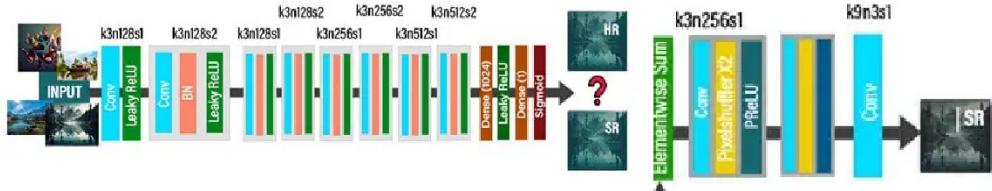


Figure 2.14: SRGAN discriminator architecture: An input image is first processed by batch normalization, multiple convolutional layers with leaky ReLU activations, and subsequent down-sampling of the picture. After that, the features are flattened and sent through a series of thick layers, culminating in a final sigmoid activation that yields a probability that either the input image is created (super-resolved) or real (high-resolution). To distinguish between created and genuine images, this output is then compared to the comparable super-resolved (SR) image from the generator.[11]

The discriminator, a second deep CNN, is intended to discern between images produced by the generator and actual high-resolution (HR) images. While the generator tries to produce images realistic enough to fool the discriminator, its job is to maximize the accuracy of distinguishing genuine images. Because of this adversarial training, the generator and discriminator get better over time, which eventually produces very realistic images. A mix of loss functions, such as adversarial loss, perceptual loss, and content loss, direct the training process. The resulting images closely resemble the original HR photos in terms of pixel values thanks to the content loss, which is commonly computed using Mean Squared Error (MSE).[42]

SRGAN has demonstrated significant promise in medical imaging, particularly in enhancing the resolution of PET and MRI images. In diseases like Alzheimer's disease, where minute details in brain scans are essential for monitoring the illness's course, high-resolution images are

essential for thorough analysis and precise diagnosis. SRGAN is particularly well-suited for this application because of its capacity to preserve textures and small details. Even though SRGAN doesn't always produce the best results on conventional metrics like Peak Signal-to-Noise Ratio (PSNR) or Structural Similarity Index (SSIM), it does produce images that are visually closer to the true high-resolution images when compared to traditional methods like bicubic interpolation or simple CNN-based models. This preference for perceived quality over pixel-by-pixel precision is especially important in medical settings, where the fidelity and clarity of images that can directly influence clinical decisions.[30][42]

2.7 Summary

Dementia and Alzheimer's disease are thoroughly examined in Chapter 2, along with the use of medical imaging—specifically, Positron Emission Tomography, or PET—in their diagnosis. The course of Alzheimer's disease is described in detail in this chapter, along with the vital significance of early identification using cutting-edge imaging methods. Additionally, manifold learning techniques—which are necessary for assessing high-dimensional neuroimaging data—are introduced, including Principal Component Analysis (PCA), Uniform Manifold Approximation and Projection (UMAP), t-Distributed Stochastic Neighbor Embedding (t-SNE), and Isometric Mapping (Isomap). In order to get around the drawbacks of conventional imaging techniques, the chapter also explores the potential of Generative Adversarial Networks (GANs), including Deep Convolutional GANs (DCGANs) and Super Resolution GANs (SRGANs), in producing lifelike synthetic PET pictures. Though large-scale, high-quality PET datasets are scarce, the chapter highlights important knowledge gaps that need to be filled. These include the need for more thorough integration of manifold learning techniques in PET image analysis, the need for additional validation of GANs in producing clinically relevant synthetic PET images for Alzheimer's diagnosis, and more. These voids prepared the ground for the novel contributions that the next chapters will address.

3 EXTENSIVE ANALYSIS OF PET IMAGE DATASET USING MANIFOLD LEARNING

3.1 Overview of the Dataset

A variety of classifications are represented in the dataset: 'CN' (Cognitively Normal), 'EMCI' (Early Mild Cognitive Impairment), 'LMCI' (Late Mild Cognitive Impairment), 'MCI' (Mild Cognitive Impairment) and 'AD' (Alzheimer's Disease). To provide a balanced dataset for analysis, each class has been complemented by 50 NIfTI (.nii) format images. Studying the development of cognitive decline and Alzheimer's disease requires accurate brain scans, which are captured in these photographs. A thorough analysis of various stages and kinds of cognitive impairments is made possible by the dataset's diversity, which also makes it possible to build reliable machine learning models for diagnosis and progression prediction. It is also possible to conduct a thorough and precise analysis because the NIfTI format guarantees that the images maintain their high quality and are compatible with a variety of neuroimaging analysis tools.

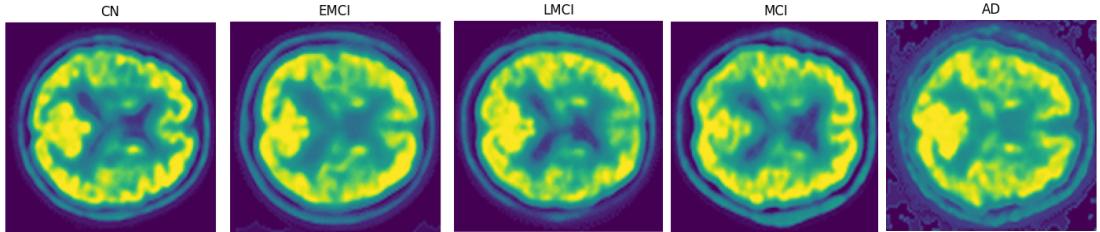


Figure 3.1: Each categorization in the dataset—Cognitively Normal (CN), Early Mild Cognitive Impairment (EMCI), Late Mild Cognitive Impairment (LMCI), Mild Cognitive Impairment (MCI), and Alzheimer's Disease (AD)—is represented by one example brain image in this figure.

3.2 Data Preprocessing

3.2.1 Standardization

The standardization process for PET brain images involves converting the pixel intensity values of each image to a common scale. This is done using the min-max standardization method, which rescales the pixel values to a range of [0, 1]. The process involves subtracting the global minimum pixel value (calculated across all images) from each pixel value in an image, and then dividing the result by the difference between the global maximum and minimum pixel values. This ensures

that the intensity values of all images are on a comparable scale, enabling more consistent and meaningful analysis.

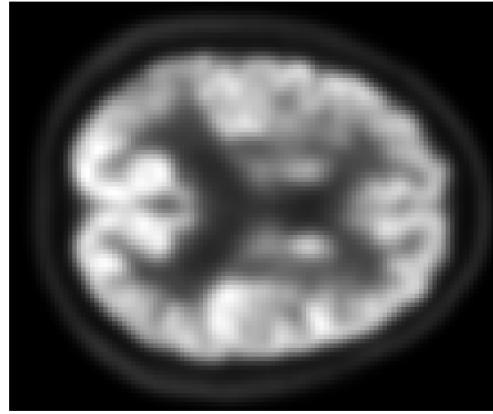


Figure 3.2: A clearer depiction of the structure of the brain has been achieved by the standardizing process. Differentiating between distinct brain areas is made simpler by the standardized image's more consistent distribution of intensity values. Various shades of gray represent intermediate levels in the intensity values, which range from 0 (black) to 1 (white). A proper study and interpretation of the brain depend on the ability to visually discriminate between its functional and anatomical parts, which is made possible by this improved contrast.

The efficiency of the min-max standardization procedure is demonstrated by the constant, standardized values of 0 and 1 for the minimum and maximum across all photographs. This consistency shows that the entire dataset's pixel values have been correctly rescaled to the intended range.

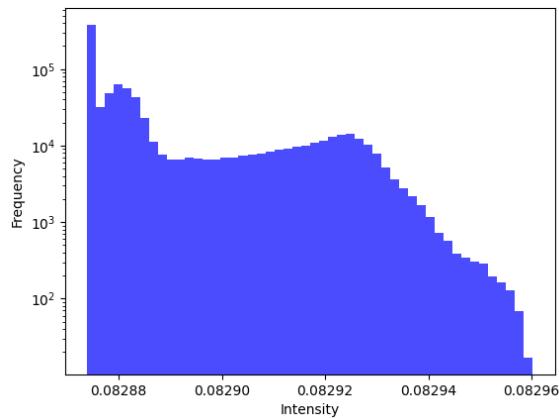


Figure 3.3: With a notable peak at the lowest intensities, the histogram shows the normal distribution of the intensity values of pixels in the standardized PET brain image. The frequency of pixels gradually drops as intensity values rise, suggesting that higher intensity values are less frequent. The large range of pixel frequencies is effectively visualized by the logarithmic scale on the y-axis.

The standardized PET brain image histogram displays a variance of pixel intensity values. The intensity values are plotted on the x-axis and range roughly from 0.08288 to 0.08296. The frequency of these values is plotted on the y-axis, which is logarithmic in scale. A notable peak can be seen in the histogram at the lowest intensity values, indicating a large number of low-intensity pixels that most likely correspond to less dense or less active brain regions. The frequency of pixels gradually drops as intensity values climb, indicating that greater intensity values—which can indicate denser or more active regions—are less frequent.

3.2.2 Enhancement

Adaptive histogram equalization is used in the enhancement process to boost contrast in the PET brain pictures. By dispersing pixel intensity values, this technique—known as Contrast Limited Adaptive Histogram Equalization, or CLAHE—adjusts the contrast of the image, improving the visibility of minute details and features.

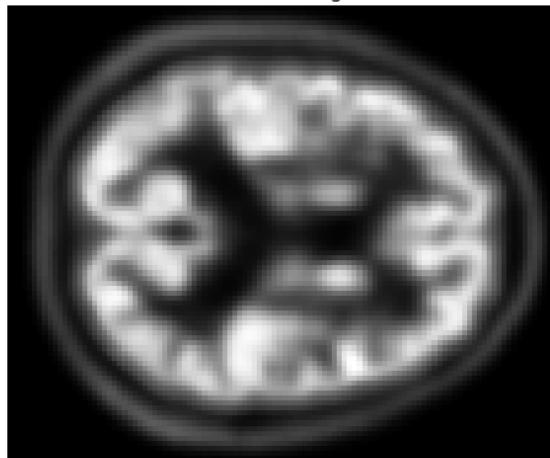


Figure 3.4: This image depicts the PET brain scan following adaptive histogram equalization, which results in increased contrast. The improved image exposes more distinct anatomical and functional brain regions, allowing for greater visualization and analysis.

When compared to the first standardized image, the contrast in the enhanced PET brain image has been greatly improved. Different brain regions are now easier to discern from one another thanks to adaptive histogram equalization, which has also brought to light previously hidden finer details. This improvement is especially helpful for medical imaging since it makes the anatomical and functional regions of the brain easier to see and distinguish.

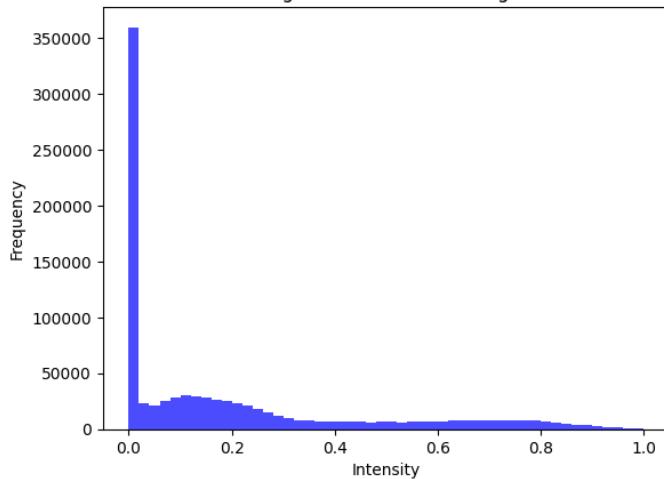


Figure 3.5: The histogram has a prominent peak at low intensities and a wider spread throughout the range, showing better contrast and visibility of brain characteristics following adaptive histogram equalization.

The enhanced PET brain image's histogram shows how adaptive histogram equalization affects the distribution of pixel intensities. The improved histogram shows a wider distribution of intensity values over the whole range from 0 to 1, in contrast to the previous histogram, which had a noticeable clustering of intensity values at the lower end.

At the lowest possible intensity values, a noticeable peak still exists, suggesting a large number of pixels with low intensity. Nonetheless, more pixels now exhibit medium to high intensity levels, as seen by the rise in the frequencies of higher intensity values. Due to the improved contrast brought about by this redistribution, finer details and distinct brain areas may be seen in the enhanced image. Overall, this augmentation makes the PET brain image more balanced and informative, making it easier to analyze and comprehend anatomical and functional features.

3.3 Principal Component Analysis

While extracting features from an array of enhanced images, multiple metrics—like mean intensity, intensity standard deviation, skewness, and kurtosis—are computed for each image iteratively. Next, for additional analysis, these features are kept in a DataFrame. Relationships between various attributes are shown by computing a correlation matrix using the Data Frame and visualizing it with a heatmap.

Table 3.1: Table of intensity statistics.

mean_intensity	std_intensity	skewness	kurtosis
0.198892	0.252417	1.313273	0.562488
0.230304	0.255406	1.059260	0.102874
0.222768	0.256225	1.144955	0.247051
0.206482	0.254171	1.274105	0.485003
0.212942	0.261697	1.201248	0.308728

To reduce the data to three dimensions, Principal Component Analysis (PCA) is done to the features DataFrame. The most notable variance in the data is represented by the principal components that are produced, which are shown in a three-dimensional scatter plot. This decrease of dimensionality helps to provide a more comprehensible visual representation of complex data.

Table 3.2: Table of principal component analysis results.

Principal Component 1	Principal Component 2	Principal Component 3
0.009044	-0.016386	-0.002587
-0.510018	0.069566	-0.003222
-0.345295	0.037365	0.000540
-0.077364	-0.004868	0.000859
-0.268058	0.004171	-0.001028

Each improved image's mean intensity, intensity standard deviation, skewness, and kurtosis are shown in the first few rows of the extracted features table. These characteristics, which highlight significant facets of the images' intensity distributions, offer a statistical synopsis of the pictures. The distribution's skewness, kurtosis, and mean intensity all quantify how asymmetrical the distribution is, standard deviation shows how widely distributed the intensity values are, and mean intensity provides an overall average brightness.

By retaining as much variance as feasible, PCA is a statistical approach that helps a dataset become less dimensional. For the purpose of comprehending the fundamental structure of the data, the table offers a succinct illustration of the most substantial variance represented by among the initial three primary components.

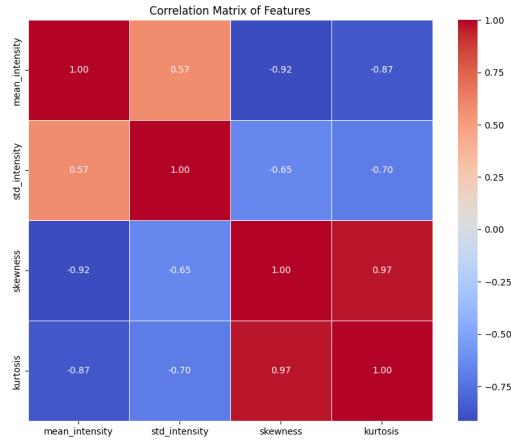


Figure 3.6: Correlation matrix illustrating the connections between kurtosis, skewness, standard deviation, and mean intensity. Negative correlations are shown in blue, and strong positive correlations are shown in red. While mean intensity and skewness have a strong negative correlation (-0.92), skewness and kurtosis have a substantial positive correlation (0.97).

The correlation matrix heatmap depicts the correlations between the dataset's numerous properties, including mean intensity, standard deviation of intensity, skewness, and kurtosis. The heatmap employs color coding to show the intensity and direction of these associations. For example, skewness and kurtosis have a substantial positive association (0.97), indicating that as skewness increases, so does kurtosis. In contrast, there is a large negative association between mean intensity and skewness (-0.92), implying that greater mean intensity levels are related with lower skewness. This matrix is vital for determining how various traits interact and influence one another.

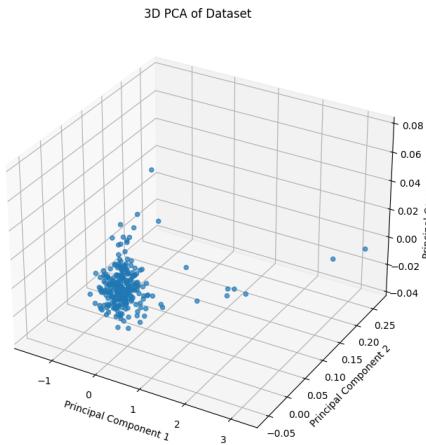


Figure 3.7: Dataset's 3D PCA plot: The data points mapped onto the first three major components are visualized in this scatter plot. Plotting the variance that each component captures aids in illuminating the linkages and underlying structure of the data in a lower dimensional space.

Using the first three main components, the 3D PCA scatter plot represents the dataset in three dimensions. This graphic aids in understanding the variation captured by the PCA as well as how the data points relate to one another in the reduced dimension. The visualization is especially useful for finding clusters, patterns, and outliers in data, revealing insights that would be difficult to distinguish in higher dimensions.

Table 3.3: Table of principal component analysis results for intensity statistics.

	mean_intensity	std_intensity	skewness	kurtosis
PC1	-0.037453	-0.012868	0.335540	0.941193
PC2	0.155474	-0.033000	-0.928131	0.336619
PC3	0.986259	0.046681	0.157635	-0.016313

High loadings for skewness (0.336) and kurtosis (0.941) are seen for Principal Component 1 (PC1), indicating that these characteristics have a significant impact on PC1. On the other hand, the standard deviation and mean intensity have very small negative loadings, suggesting that they have less of an impact. Skewness (-0.928) and kurtosis (0.337), which have opposite effects on Principal Component 2 (PC2), are the primary factors influencing PC2. This indicates an inverse link between these traits in this component. Principal Component 3 (PC3) is substantially impacted by mean intensity (0.986), with standard deviation, skewness, and kurtosis contributing insignificantly. This suggests that mean intensity is the primary cause of fluctuations in the data that PC3 represents.

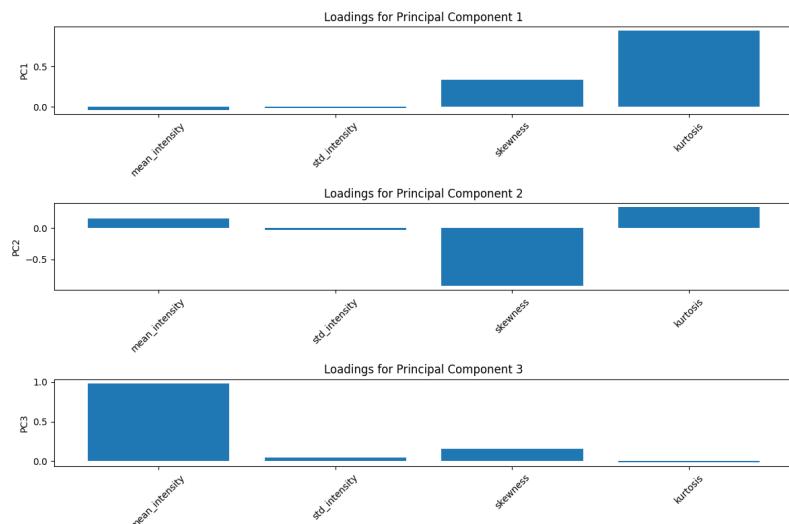


Figure 3.8: Every primary component loading is displayed in the bar plots. For PC1, skewness and kurtosis have the greatest positive loading, and kurtosis is the most influential. Skewness (negative loading) and kurtosis (positive loading) have an inverse connection that PC2 depicts. Other variables have negligible impact on PC3, which is mostly determined by mean intensity.

When combined, the loadings and bar graphs provide information about the data's underlying structure. The greatest substantial variance is captured by Principal Component 1, which is influenced by kurtosis and skewness. Elevated PC1 levels are associated with images that exhibit greater skewness and kurtosis, signifying an uneven and peaked distribution of intensity. The second most important variance is represented by Principal Component 2, which is marked by an inverse connection between kurtosis and skewness. Images with peaked distributions and less asymmetry are suggested by high PC2 values, which are linked to high kurtosis and low skewness. Principal Component 3 represents the variance mostly associated with mean intensity. Greater mean intensity in the photos is indicated by high PC3 values, which also indicate the images' overall brightness.

3.3.1 K-means Clustering

Principal component extraction from the dataset is the first step in the procedure. The most significant variance is captured and data presentation and analysis are made simpler when PCA reduces the dataset to three primary components. After that, the dataset has been reduced by PCA and split into three clusters using the K-means technique. To reduce variance within each cluster, K-means iteratively groups data points into clusters according to how close they are to the cluster centroids. The cluster label for each data point is then assigned and added to the PCA DataFrame so that it may be seen.

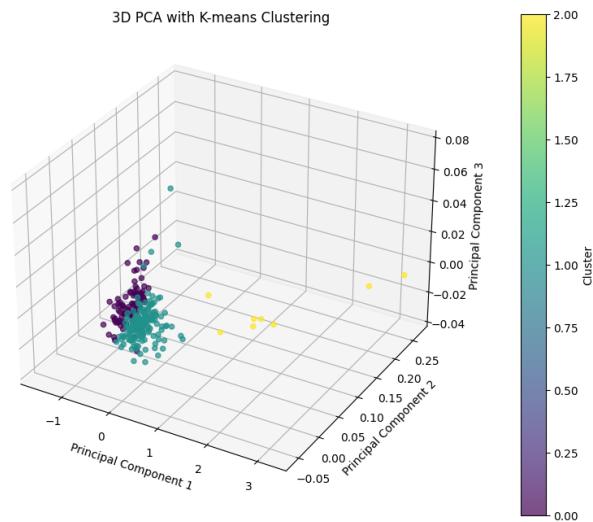


Figure 3.9: The 3D scatter plot shows the results of K-means clustering on the PCA-reduced dataset. Data points are colored according to their assigned clusters, revealing distinct groupings in the principal component space. The plot illustrates the separation of clusters based on the principal components, highlighting the inherent structure and variability within the dataset.

Within the area delineated by the first three main components, the data points are arranged into three discrete clusters as seen in the three-dimensional scatter plot. These primary components, which capture the primary causes of variance in the dataset, are represented by the axes. Each data point's cluster is shown by a color bar on the plot's right side, which makes it easy to identify between the various clusters.

The plot demonstrates that while some data points are dispersed to create different clusters, the majority of data points are clustered in a single area of the PCA space. This indicates the K-means algorithm is effective in identifying innate groupings within the data. The clusters' reasonable degree of separation in three dimensions indicates that the dataset's many groups can be distinguished from one another by the features that were employed, which include skewness, kurtosis, mean intensity, and standard deviation of intensity. The principle component loadings of each cluster can be used to interpret it and get insight into the features of each group. For example, mean intensity may have an impact on clusters distributed along PC3, but skewness and kurtosis may have an impact on those distributed along PC1.

3.4 Manifold Learning

3.4.1 t-Distributed Stochastic Neighbor Embedding(t-SNE)

The dataset was initially reduced to three principal components using PCA, which captured the majority of the variance. After that, the PCA-reduced dataset was sent to the K-means method, which separated the data into three groups according to how close the cluster centroids were. The three-dimensional PCA-reduced dataset was then further reduced into two dimensions using the t-SNE technique. t-SNE is very helpful for showing clusters and neighborhood interactions since it maintains the local structure of the data.

The data points are displayed in a two-dimensional space on the t-SNE plot, with colors indicating the clusters that the K-means clustering assigned them to. In order to maintain the local structure of the data, the two dimensions that are represented by the axes are those that are obtained from the t-SNE algorithm. These dimensions do not match the original features; rather, they reflect the connections and separations between the data points.

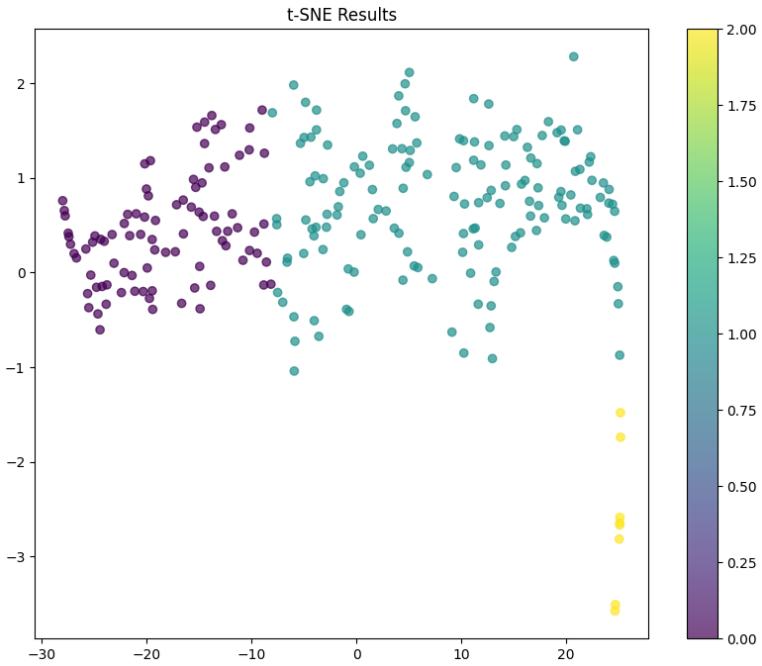


Figure 3.10: The PCA-reduced dataset’s two-dimensional embedding is displayed in the t-SNE plot, where data points are colored based on their K-means cluster allocations. The display highlights the underlying linkages and groupings in the data by revealing distinct clusters and maintaining local structures.

Plotting the data points according to their cluster labels reveals discrete clusters. This illustrates how well the clustering structure found by the K-means algorithm has been captured by the t-SNE method. The distinct separation of clusters and the clustering of points within each cluster demonstrate how well t-SNE preserves local associations between data points. In the t-SNE graphic, data points that are near together in the high-dimensional space stay close together, preserving neighborhood ties. Additionally, the plot sheds light on each cluster’s unique qualities. A discrete set of data points with comparable qualities is indicated by the cluster on the left (purple), for instance, which is clearly differentiated from the others. The more dispersed cluster on the right (cyan) indicates variation within this group.

3.4.2 Uniform Manifold Approximation and Projection(Umap)

Following PCA and K-means clustering, the dataset is reduced to two dimensions, which is visualized using the UMAP (Uniform Manifold Approximation and Projection) plot. Visualizing high-dimensional datasets in a low-dimensional space is made possible by UMAP, a potent manifold learning technique that maintains the data’s global and local structures.

To identify the most significant variance, PCA was first used to reduce the dataset to three principal components. The data was then subjected to PCA reduction and the K-means method, which separated it into three groups according to how close the data were to the cluster centroids. Ultimately, the dataset that was reduced from three dimensions by PCA was reduced to two dimensions using the UMAP technique. Complex datasets can be effectively visualized using UMAP since it prioritizes maintaining both the general topological structure of the data and the links between nearby neighborhoods.

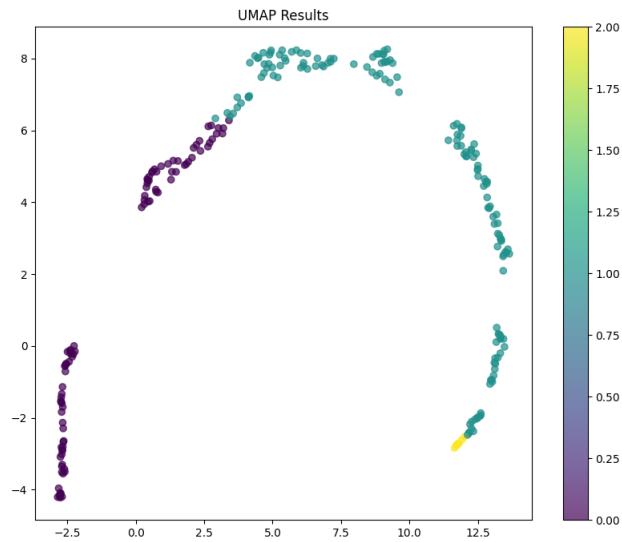


Figure 3.11: The UMAP map displays the PCA-reduced dataset's two-dimensional embedding, with data points colored based on their K-means cluster assignments. The display highlights the innate linkages and groupings within the data by revealing unique clusters while maintaining local and global structures.

Plotting separate groups of data points colored by their cluster labels reveals that the K-means algorithm's identified clustering structure was successfully captured by the UMAP method. Both the global structure of the dataset and the local associations among data items are successfully preserved using UMAP. The clusters' distinct division and the way the points are grouped within each cluster demonstrate this. The overall structure and topology of the data are preserved, and data points that are close to one another in the high-dimensional space stay close in the UMAP graphic. The plot sheds light on each cluster's unique qualities. As an example, the purple cluster on the left is clearly distinguished from the rest, suggesting that it is a unique set of data points with related characteristics. Given its greater dispersion, the cluster on the right (cyan) suggests that there is variation within this group. Another distinct collection of data points is represented by the little cluster at the bottom (yellow).

3.4.3 Isometric Mapping(Isomap)

A two-dimensional dataset that has undergone PCA and K-means clustering is visualized using the Isomap (Isometric Mapping) display. An effective method for visualizing high-dimensional datasets in a lower-dimensional environment is isomap, a manifold learning methodology that seeks to maintain the global geometric organization of the data while lowering its dimensionality.

First, principal component analysis (PCA) was used to reduce the dataset to three components, which represented the majority of the variation. After that, the PCA-reduced dataset was sent to the K-means method, which divided the data into three clusters according to how close the data were to the cluster centroids. Lastly, the three-dimensional PCA-reduced dataset was reduced to two dimensions using the Isomap technique. With the goal of keeping the distances between every pair of data points as constant as possible, isomap concentrates on maintaining the overall structure of the data.

The K-means clustering's cluster assignments are indicated by color in the Isomap plot, which shows the data points in a two-dimensional space. The Isomap technique yielded two dimensions, represented by the axes, which do not match the original characteristics but rather depict the distances and relationships between data points while maintaining the overall structure.

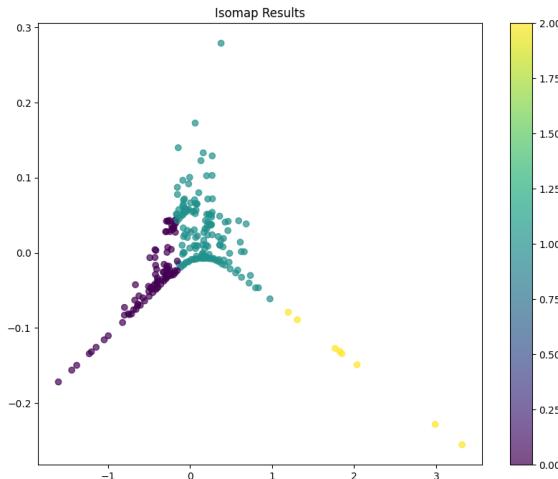


Figure 3.12: The PCA-reduced dataset's two-dimensional embedding is displayed in an Isomap plot, where data points are colored based on their K-means cluster assignments. By emphasizing discrete clusters and retaining the separations between data points, the display maintains the overall geometric structure of the data. This representation sheds light on the relationships and underlying groups within the collection, offering an understanding of its general topological structure.

Different data point clusters are visible in the plot, and these cluster labels are used to color-code the data points. This suggests that the grouping structure found by the K-means algorithm has been effectively captured by the Isomap technique. Maintaining the global links between data points is effectively accomplished by isomap. The general layout and arrangement of the clusters in the plot make this clear. In the Isomap graphic, data points that are close together stay close together, while those that are situated far apart in the high-dimensional space stay far apart. The plot sheds light on each cluster's unique qualities as well. For instance, the purple cluster on the left is clearly distinguished from the other clusters, signifying a unique collection of data points with related characteristics.

4 IMPLEMENTATION OF GAN

4.1 Preprocessing and Fixing Hyperparameters

Beginning with the creation of output folders for each of the following conditions (AD, CN, EMCI, LMCI, MCI), the method systematically gathers images from assigned folders. To process each NIfTI (.nii) file, the middle slice of the 3D picture data is extracted, transformed into a 2D image, then normalized to fall inside the [0, 255] range so that it may be converted to a PNG file. Normalization guarantees that the photos are appropriate for training machine learning models. Maintaining a consistent framework, the photos are arranged according to their respective conditions. The procedure is made more efficient by the inclusion of performance measures, which are essential for expanding to larger datasets. With consistency, quality control, and ease of use in mind, this method shows how to effectively prepare and visualize medical imaging data for machine learning and analysis.

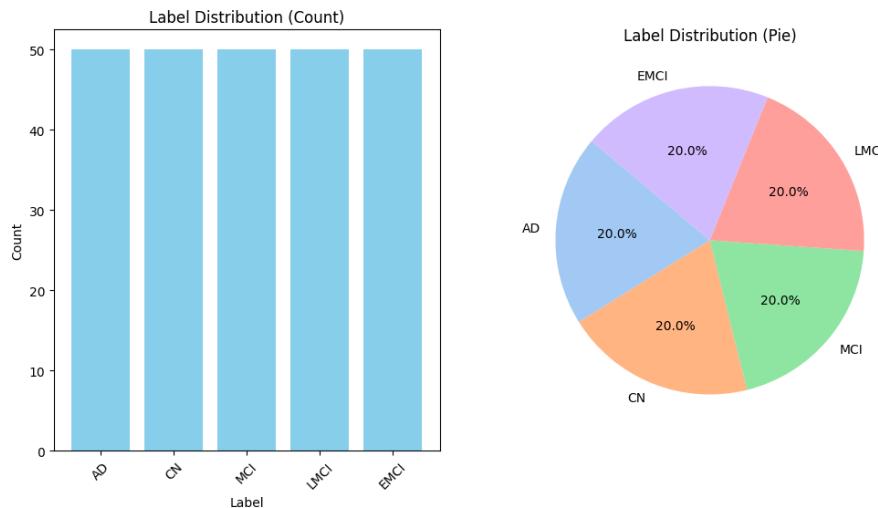


Figure 4.1: Visualization of the label distribution: The dataset's counts for each label category (AD, CN, MCI, LMCI, and EMCI) are shown in the left bar chart, which exhibits an equal distribution with 50 cases for each label. The same distribution is shown in the pie chart on the right, which shows that each category makes up 20% of the entire dataset, indicating that the label distribution is balanced.

Effective preprocessing and visualization of medical imaging data—presumably brain scans from a range of conditions—are required in this procedure. Pictures are collected in a methodical manner from designated folders, and then they are resized and converted to RGB to guarantee con-

sistency in dimensions and color channels. In order to effectively train machine learning models, the images are standardized to a [0, 1] range. By exposing the model to a range of data distributions, random shuffling improves generalization. A grid of photos is also created for visual inspection, enabling quality control and initial pattern analysis. Efficiency is highlighted by performance measurements, which is important for scaling up to larger datasets. This technique is a good example of how to prepare and visualize medical picture data for use in machine learning and analysis.

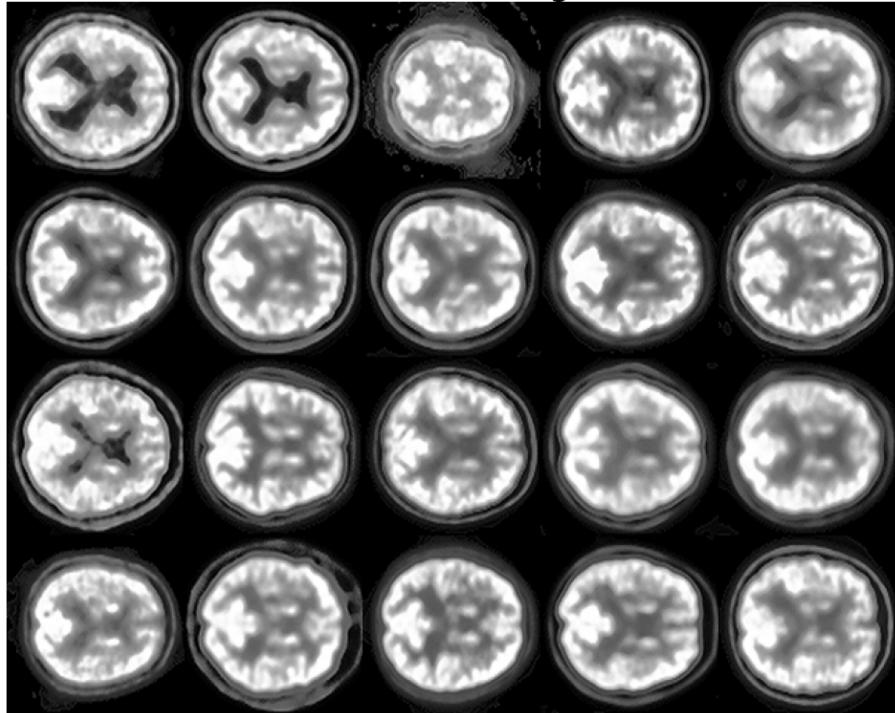


Figure 4.2: Grid of 20 preprocessed brain scan pictures: A collection of brain slices that have undergone preprocessing, like normalization, are displayed in this grid. The consistent appearance of the images makes analysis and comparison easier, which helps with jobs like medical imaging research machine learning model training and feature extraction.

4.2 Deep Convolution Generative Adversarial Network (DCGAN)

When creating and training a Deep Convolutional Generative Adversarial Network (DCGAN), hyperparameters are essential. To balance training effectiveness, model stability, and the quality of output images, the DCGAN's hyperparameters are carefully chosen. Enough learning opportunities are guaranteed by a high number of epochs (500), and a batch size of 64 strikes a balance between computational efficiency and steady training. The latent dimension of 100 is not unduly high-dimensional, but it offers enough complexity to generate a variety of images. To capture im-

portant features while keeping computing requirements reasonable, images are scaled to 128x128 pixels. Realistic RGB color visuals can be produced by using three channels. Training progresses steadily and without instability with a learning rate of 0.0002, and mode collapse is avoided and training is stabilized with a beta1 value of 0.5 for the Adam optimizer.

Table 4.1: Hyperparameters for the DCGAN model.

Hyperparameter	Value	Description
n_epoch	500	Number of training epochs.
batch_size	64	Size of each training batch.
latent_dim	100	Dimensionality of the latent space for the GAN.
cols	128	Width of the input images (in pixels).
rows	128	Height of the input images (in pixels).
channels	3	Number of color channels in the images (RGB).
dim	(128, 128)	Dimensions of the input images (width, height).
in_shape	(128, 128, 3)	Shape of the input images (height, width, channels).
lr	0.0002	Learning rate for the Adam optimizer.
beta1	0.5	Beta1 parameter for the Adam optimizer (momentum term).
nrows	3	Number of rows in the plot for generated images.
ncols	4	Number of columns in the plot for generated images.

The DCGAN generator model uses a number of neural network layers to transform a latent vector—random noise—into realistic images. Initially, a Dense layer is used to project the input noise vector of size latent_dim onto a high-dimensional space of 128x8x8 nodes. An activation function called LeakyReLU follows this projection, adding non-linearity and assisting in avoiding the dying ReLU issue. After that, the output is reshaped into an 8x8x128 tensor.

Four Conv2DTranspose layers are applied on this reshaped tensor. By gradually upsampling the tensor and doubling its spatial dimensions, each of these transposed convolutional layers preserves or improves the feature maps. To add non-linearity and guarantee stable training, these layers employ a 4x4 kernel with a 2x2 stride, and are succeeded by LeakyReLU activations.

Using a 5x5 kernel and the tanh activation function, the final Conv2D layer employs three filters that correspond to the RGB channels of the image, normalizing the output pixel values to the range of [-1, 1]. Since it fits the normalized range of the training images, this normalization is essential for stabilizing the training process.

The architecture of the generator, which consists of numerous transposed convolutional layers, reshape, and dense layers, guarantees that the input noise may be gradually refined and upscaled

Table 4.2: Layer summary for the DCGAN generator.

Layer (type)	Output Shape	Param #
dense_1 (Dense)	(None, 8192)	827392
leaky_re_lu_5 (LeakyReLU)	(None, 8192)	0
reshape (Reshape)	(None, 8, 8, 128)	0
conv2d_transpose (Conv2DTran)	(None, 16, 16, 128)	262272
leaky_re_lu_6 (LeakyReLU)	(None, 16, 16, 128)	0
conv2d_transpose_1 (Conv2DTran)	(None, 32, 32, 128)	262272
leaky_re_lu_7 (LeakyReLU)	(None, 32, 32, 128)	0
conv2d_transpose_2 (Conv2DTran)	(None, 64, 64, 128)	262272
leaky_re_lu_8 (LeakyReLU)	(None, 64, 64, 128)	0
conv2d_transpose_3 (Conv2DTran)	(None, 128, 128, 128)	262272
leaky_re_lu_9 (LeakyReLU)	(None, 128, 128, 128)	0
conv2d_5 (Conv2D)	(None, 128, 128, 3)	9603
Total params:		1,886,083
Trainable params:		1,886,083
Non-trainable params:		0

into a high-resolution 64x64x3 image. Through training this generator in an adversarial environment with a discriminator, the model acquires the ability to produce incredibly lifelike images. The generator becomes more adept at producing visuals that are identical to genuine images over time.

In order to distinguish between actual and fake images, the discriminator model in the DCGAN uses a convolutional neural network. Starting with a Conv2D layer, it applies 128 filters with a 5×5 kernel size, using 'same' padding to preserve the input dimensions, to an input picture of shape 128×128×3 (height, width, and channels). Introducing non-linearity and avoiding the dying ReLU problem, this layer is followed by a LeakyReLU activation function with an alpha of 0.2.

The network then consists of many Conv2D layers, each of which applies 128 filters with 5×5 kernels and gradually downsamples the image using 2×2 strides. To preserve non-linearity, a LeakyReLU activation comes after every Conv2D layer. By reducing the spatial dimensions of the image, this downsampling preserves and learns crucial elements that aid in differentiating between authentic and fraudulent photos.

A 1D tensor is created by flattening the output following multiple convolutional layers. A Dropout layer with a rate of 0.4 is applied to this tensor during training in order to help prevent overfitting by arbitrarily setting a portion of the input units to zero. The last layer is a Dense layer

that outputs a probability indicating whether the input image is real or fraudulent. It consists of a single neuron and a sigmoid activation function.

Table 4.3: Layer summary for the DCGAN discriminator.

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 128, 128, 128)	9728
leaky_re_lu (LeakyReLU)	(None, 128, 128, 128)	0
conv2d_1 (Conv2D)	(None, 64, 64, 128)	409728
leaky_re_lu_1 (LeakyReLU)	(None, 64, 64, 128)	0
conv2d_2 (Conv2D)	(None, 32, 32, 128)	409728
leaky_re_lu_2 (LeakyReLU)	(None, 32, 32, 128)	0
conv2d_3 (Conv2D)	(None, 16, 16, 128)	409728
leaky_re_lu_3 (LeakyReLU)	(None, 16, 16, 128)	0
conv2d_4 (Conv2D)	(None, 8, 8, 128)	409728
leaky_re_lu_4 (LeakyReLU)	(None, 8, 8, 128)	0
flatten (Flatten)	(None, 8192)	0
dropout (Dropout)	(None, 8192)	0
dense (Dense)	(None, 1)	8193
Total params:		3,313,666
Trainable params:		1,656,833
Non-trainable params:		1,656,833

Using the typical loss function of binary cross-entropy for binary classification problems, the model is created using the Adam optimizer, set up with a learning rate of 0.0002 and beta_1 = 0.5. With the help of its design, which consists of activation's, convolutional layers, and a final classifier, the discriminator can successfully learn to distinguish between real and created images. This ability to do so gives the generator crucial feedback during adverse training.

Table 4.4: Layer summary for the DCGAN.

Layer (type)	Output Shape	Param #
sequential_1 (Sequential)	(None, 128, 128, 3)	1886083
sequential (Sequential)	(None, 1)	1656833
Total params:		3,542,916
Trainable params:		1,886,083
Non-trainable params:		1,656,833

The train function alternates between training the generator and the discriminator in a methodical manner to train a DCGAN. Setting the number of batches per epoch and initializing parameters are the first steps. The function uses both real and fake samples to train the discriminator during each epoch. Real samples are taken from the dataset and marked as real, while fake samples are

created by the generator and marked as fake. On both the fake and real samples, the discriminator's losses are noted. The generator is then trained using the GAN model in order to generate fictitious samples that the discriminator will identify as authentic. The loss of the generator is also noted.

The function records the time taken, the number of epochs, the losses for the generator, the genuine and false samples, and the epochs. Furthermore, the function summarizes and visualizes the model's performance every 10% of the total epochs. When every epoch has finished, the total training time is displayed, and the training process is visualized by creating a plot of the recorded losses. Over time, more realistic images are produced because to this alternating training procedure, which guarantees simultaneous improvement of the discriminator and the generator.

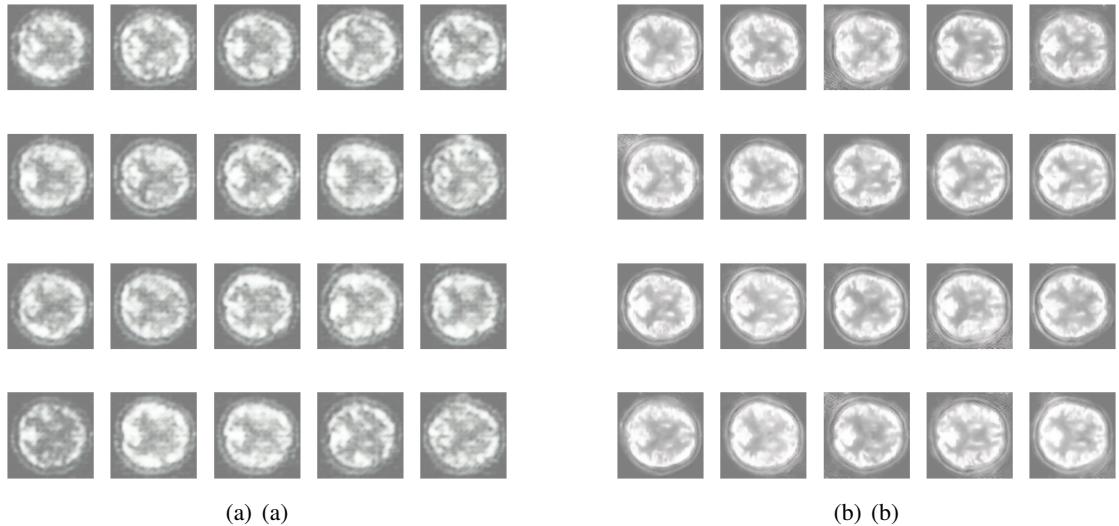


Figure 4.3: Comparison of images at different epochs: (a) The brain images generated at epoch 100 show the early phases of the model's training, when the fundamental structural patterns of the brain are starting to take shape. However, the high levels of noise and blurring suggest that the model has not yet fully caught the intricate details of the architecture of the brain. (b) The brain images generated at epoch 300 show improved structural correctness and decreased noise, suggesting that as training goes on, the model has significantly improved its capacity to recreate the minute intricacies of brain architecture.

GAN (Generative Adversarial Network) training for 500 epochs resulted in considerable increases in image quality. During the initial phases, such as epoch 300, the generated images revealed fundamental structures but lacked fine details. As training went to epoch 350, image clarity and structure improved significantly, demonstrating that the generator was learning to produce more refined images. Around epoch 400, the images had become more refined, with more constant

patterns and details, but the discriminator's ability to discern between real and fake images began to deteriorate.

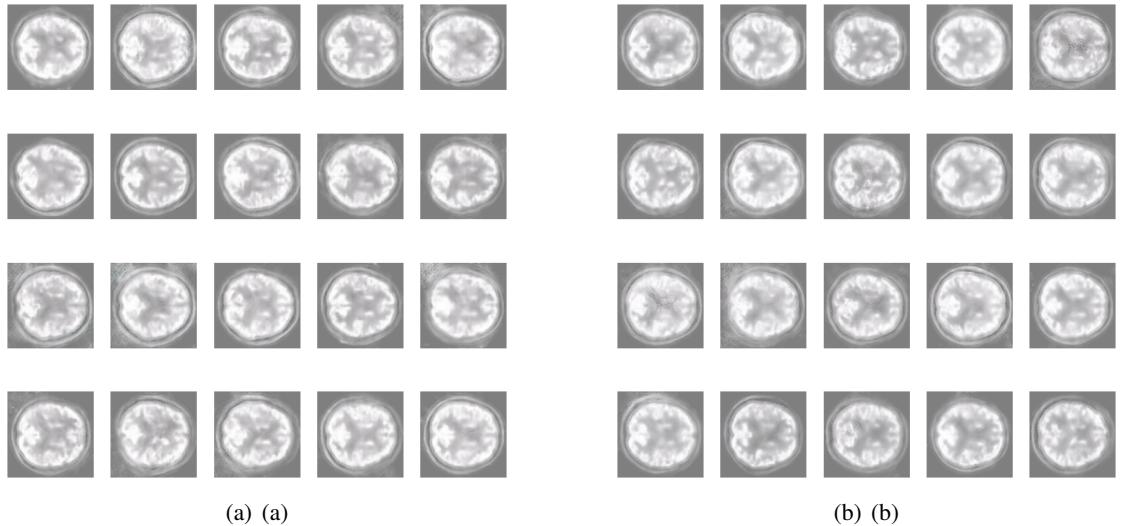


Figure 4.4: Comparison of images at different epochs: (a) Images of the brain generated at epoch 400 show additional structural and lucid refinement along with notable enhancements in brain feature representation, suggesting that the model is becoming more and more capable of producing realistic images. (b) Brain images generated at epoch 500 show reduced artifacts and great structural fidelity, indicating that the model has successfully learned to capture and duplicate the fine aspects of brain anatomy as training achieves maturity.

With better quality and clearly identifiable structures, the generated images had significantly improved by epoch 450. The generator's images followed this trajectory, and by epoch 500, they were nearly identical to real photographs in terms of clarity and quality. As a result of the generator's better loss values during training, the discriminator's low accuracy rates showed that it was having more difficulty identifying false images.

As a result, the GAN training was effective, and the generator's ability to produce realistic images kept getting better. The generator proved to be useful as the discriminator's accuracy declined over time, although it was still effective at differentiating between actual and phony images at first. The training procedure demonstrates the ability of GANs to produce high-quality images and the necessity of thorough training to get the best outcomes.

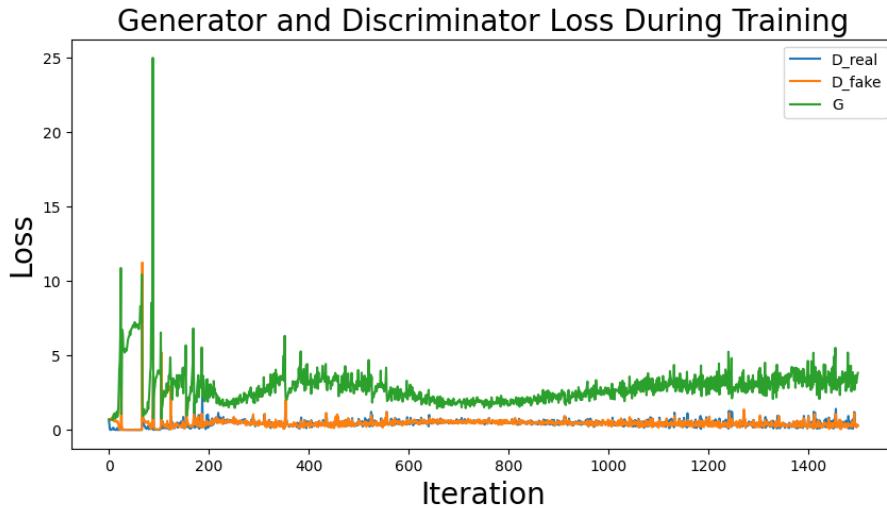


Figure 4.5: Plot showing discriminator and generator loss during training: The generator (G) and discriminator (D_real and D_fake) loss values are plotted on the graph as they increase with each repetition. As a result of the adversarial nature of GAN training, there is a notable initially in both losses. With time, the discriminator retains its capacity to discern between actual and created data, and the generator becomes more adept at producing realistic images, suggesting that the model is gradually getting closer together.

The generator struggles to create realistic images at first, as seen by its relatively high loss. But as training goes on, this loss drastically drops, indicating the generator's growing capacity for producing realistic images. The generator's loss spikes early on are a sign of the initial instability that is typical of GAN training.

In contrast, with sporadic spikes, the discriminator's losses (D_real and D_fake) stay comparatively modest and consistent during the training. These spikes indicate times when the discriminator detects phony images produced by the generator. These losses tend to remain stable throughout the course of training, suggesting a balance between the discriminator and generator performances. Even as the generator gets better, its effectiveness in separating real from phony images is further demonstrated by the continually low discriminator loss.

The generator gradually gains the ability to produce more realistic images, as seen by the graph's overall characteristic GAN training pattern characterized by a declining generator loss. As evidenced by its continued success in detecting phony photos, the discriminator keeps a comparatively low loss. The generator and discriminator appear to be improving simultaneously, indicating that the GAN training is going as planned based on the convergence of these losses.

4.3 Super Resolution Generative Adversarial Network (SRGAN)

Advanced deep learning models called Super-Resolution Generative Adversarial Networks (SRGAN) are designed to enhance image resolution by transforming low-quality images into high-resolution counterparts. This method works particularly well in fields like digital forensics, medical imaging, and satellite photography that require extremely precise and detailed images. With the SRGAN implemented, a generator, discriminator, and combined GAN model make up a typical structure. Perceptual loss, which is computed employing a pre-trained VGG19 network, is added to further optimize it.

Table 4.5: Hyperparameters for the SRGAN model.

Hyperparameter	Value
Learning Rate	0.0002 (Adam optimizer)
Batch Size	32
Epochs	500
Latent Dimension	100
Generator Filters	[256, 128, 64, 3]
Discriminator Filters	[64, 128]
Loss Function (Discriminator)	Binary Cross-Entropy
Loss Function (Generator)	Binary Cross-Entropy (GAN) Mean Squared Error (SRGAN)
Activation Function (Generator)	LeakyReLU, Tanh
Activation Function (Discriminator)	LeakyReLU, Sigmoid
Optimizer (Generator)	Adam
Optimizer (Discriminator)	Adam
Dropout Rate	Not used
Image Size	128x128
Initial Learning Rate (VGG)	1e-4
Weight Decay (VGG)	Not used
Normalization (Discriminator)	Batch Normalization
Normalization (Generator)	Batch Normalization
Rescaling (Generated Images)	[-1, 1] to [0, 1]

In order to transform an image from low resolution to high resolution, the first step in the process is to use the Generator Network. A feature map of size 32,768 is produced by the generator after it first processes a latent input vector of size 100 through a dense layer. The resulting 16x16x128 tensor is then used to modify the feature map, creating the basic framework for the image that will be gradually enhanced and brightened. Multiple up-sampling blocks, which double the feature map's spatial dimensions each, are part of the generator's architecture. Employing BatchNormalization and LeakyReLU activation layers improving training stability and adding

non-linearities for learning intricate patterns, a convolutional layer refines the image further after each up-sampling step. A three-channel RGB image with pixel values scaled between -1 and 1 is produced by the generator's last layer, a Conv2D layer with a tanh activation function.

Table 4.6: Layer summary for the SRGAN generator.

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	(None, 100)	0
dense_2 (Dense)	(None, 32768)	3,309,568
leaky_re_lu_10 (LeakyReLU)	(None, 32768)	0
reshape_1 (Reshape)	(None, 16, 16, 128)	0
up_sampling2d (UpSampling2D)	(None, 32, 32, 128)	0
conv2d_6 (Conv2D)	(None, 32, 32, 128)	147,584
batch_normalization (BatchNormalization)	(None, 32, 32, 128)	512
leaky_re_lu_11 (LeakyReLU)	(None, 32, 32, 128)	0
up_sampling2d_1 (UpSampling2D)	(None, 64, 64, 64)	0
conv2d_7 (Conv2D)	(None, 64, 64, 64)	73,792
batch_normalization_1 (BatchNormalization)	(None, 64, 64, 64)	256
leaky_re_lu_12 (LeakyReLU)	(None, 64, 64, 64)	0
up_sampling2d_2 (UpSampling2D)	(None, 128, 128, 64)	0
conv2d_8 (Conv2D)	(None, 128, 128, 32)	18,464
batch_normalization_2 (BatchNormalization)	(None, 128, 128, 32)	128
leaky_re_lu_13 (LeakyReLU)	(None, 128, 128, 32)	0
conv2d_9 (Conv2D)	(None, 128, 128, 3)	867
Total params:		3,551,171
Trainable params:		3,550,723
Non-trainable params:		448

When determining whether a picture is real (from the dataset) or produced (from the generator), the discriminator works as a binary classifier. First, there is an input layer that can hold 128x128x3 images. Following the processing of these pictures, a LeakyReLU activation function is used after each convolutional layer. As the spatial dimensions of the image are further reduced, such convolutional layers extract features that are more and more complex. After the last convolutional layer's output is flattened into a single vector and sent through a dense layer, a single value that represents the likelihood that the input image is real is produced. By being trained to distinguish between actual and created images, the discriminator gives the generator vital input that helps it improve with time.

The generator and discriminator are combined to form the GAN model. To focus on training the generator, the discriminator's weights are maintained fixed during GAN training. By using this method, the generator is able to improve the realistic images it produces and fool the discriminator.

Table 4.7: Layer summary for SRGAN discriminator.

Layer (type)	Output Shape	Param #
input_2 (InputLayer)	(None, 128, 128, 3)	0
conv2d_10 (Conv2D)	(None, 128, 128, 64)	1792
leaky_re_lu_14 (LeakyReLU)	(None, 128, 128, 64)	0
conv2d_11 (Conv2D)	(None, 128, 128, 128)	73,856
leaky_re_lu_15 (LeakyReLU)	(None, 128, 128, 128)	0
flatten_1 (Flatten)	(None, 2097152)	0
dense_3 (Dense)	(None, 1)	2,097,153
Total params:		4,345,602
Trainable params:		2,172,801
Non-trainable params:		2,172,801

The generator and discriminator are updated in turn throughout the training process, gradually raising the caliber of the images that are produced.

Table 4.8: Layer summary for the SRGAN.

Layer (type)	Output Shape	Param #
model (Model)	(None, 128, 128, 3)	3,551,171
model_1 (Model)	(None, 1)	2,172,801
Total params:		5,723,972
Trainable params:		3,550,723
Non-trainable params:		2,173,249

Utilizing perceptual loss, which is based on a VGG19 model that has already been trained, is a crucial component of your SRGAN implementation. Rather than evaluating the produced and real images' pixel-by-pixel differences—differences that might not be well aligned with human sense of image quality—the perceptual loss evaluates the high-level features that the VGG19 model collected from both images. In particular, the model computes the mean squared error (MSE) between the feature maps from the VGG19 network that correspond to the real and produced images. This loss function forces the generator to create images that are both visually and structurally authentic, preserving elements that are vital to human perception.

In training, random noise is fed into the generator to create fake images, and the discriminator is used to compare the output with the genuine images. The discriminator's objective is to correctly discern between actual and bogus images, while the generator seeks to minimize both adversarial and perceptual losses. During the training process, a number of metrics are monitored, such as Discriminator Loss (D Loss), which measures the discriminator's accuracy, and Generator Loss

(G Loss), indicating how well the generator can trick the discriminator. Furthermore, pixel-wise discrepancies are measured using the Mean Squared Error (MSE) calculation, and the perceptual similarity between the produced and real images is assessed using the Structural Similarity Index (SSIM).

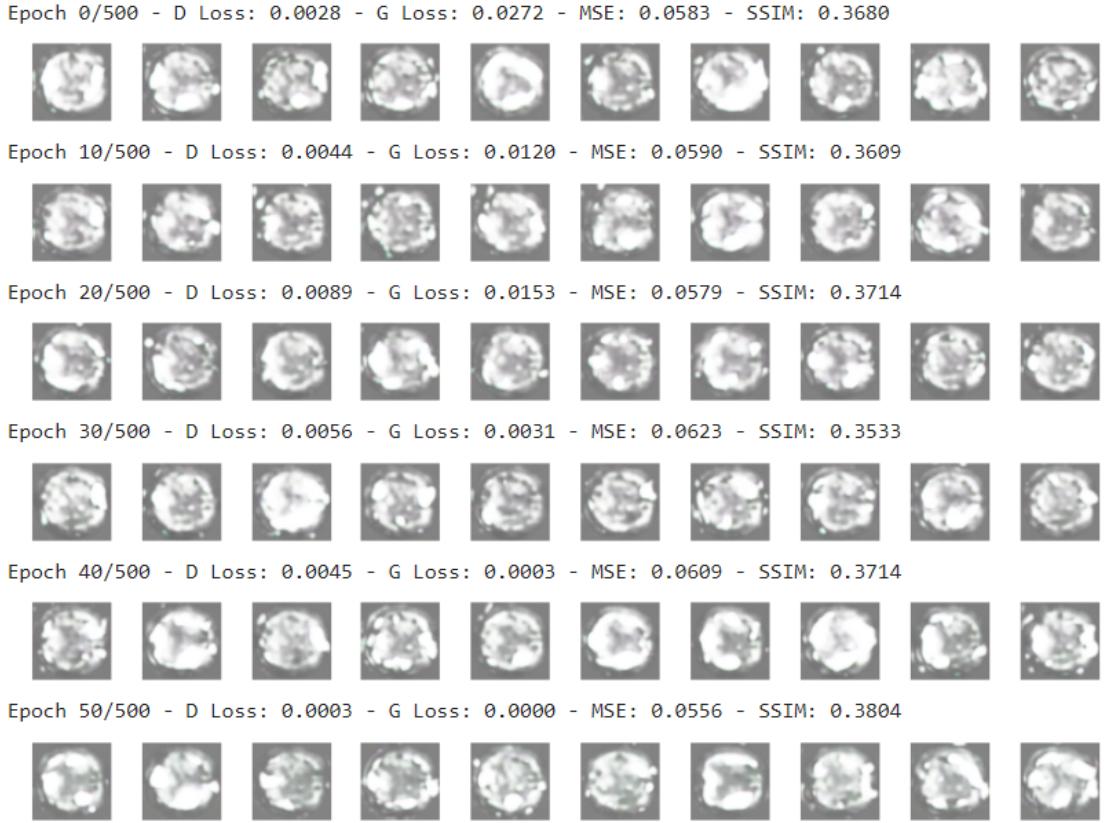


Figure 4.6: Depiction of generated images across training phases, accompanied by relevant measurements: Along with generated brain pictures at various training epochs (0 to 50), the grid also shows mean squared error (MSE), structural similarity index (SSIM), discriminator loss (D Loss), and generator loss (G Loss). The model's rising picture quality is demonstrated by the images and metrics, which show how training improves structural correctness and similarity to the target images, with dropping MSE and increasing SSIM.

The images in the sequence across the training epochs show how the SRGAN model's capacity to produce high-resolution images from low-resolution inputs gradually improved. The generated images at Epoch 0/500 are primarily noise with little discernible pattern or detail. The generator's inability to provide credible outputs is shown by the raised Discriminator Loss (D Loss) and Generator Loss (G Loss). This is corroborated by the low Structural Similarity Index (SSIM) and high Mean Squared Error (MSE). There are slight gains once training reaches Epoch 10/500; the

images start to show faint structures. But the MSE is still high and the SSIM is still low, indicating that the images still don't closely resemble the intended high-resolution targets.

The produced images begin to exhibit more identifiable patterns and structures by Epoch 20/500 and later, with notable advancements by Epoch 30/500. Improved image quality is indicated by the decreasing MSE and the continuously rising SSIM. On the other hand, the D Loss and G Loss approach near-zero levels by Epochs 40/500 and 50/500, indicating a possible problem with the training procedure. This suggests that the discriminator could not be efficiently testing the generator any more, which could cause overfitting or cause the training process to stall. Notwithstanding these reservations, the image quality has increased dramatically, showing more recognizable structures and a closer likeness to the target images.

The training of the SRGAN model from Epoch 60 to Epoch 490 shows a discernible increase in its capacity to generate high-resolution images. The resulting photos show severe color aberrations and noise in the early stages, especially between Epochs 60 and 100. Many of them have strange greenish tones. The generator is still learning the complex patterns needed to produce coherent high-resolution images, as evidenced by the hardly recognisable image structures. The generator struggles to fool the discriminator during this phase of adversarial training, which leads to a significant generator loss and noticeable differences between the generated and actual images.

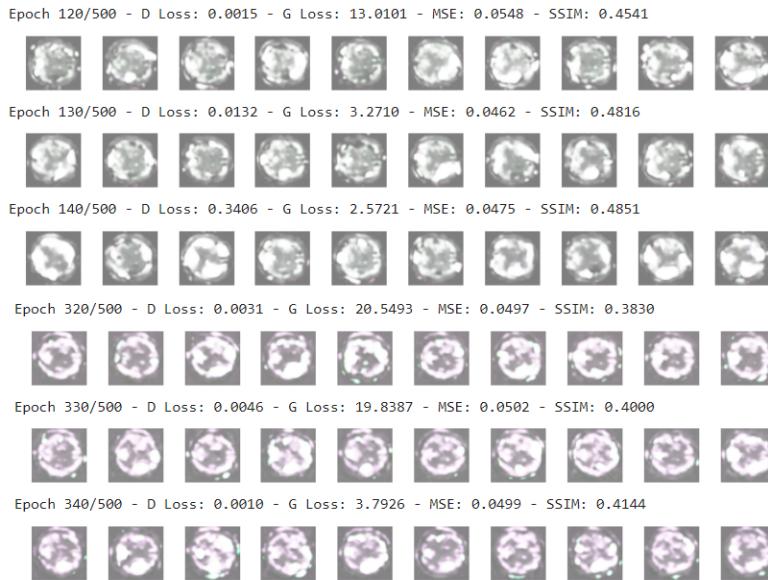


Figure 4.7: Gradually, the clarity and realism of early visuals (Epochs 120–140) improve. But, as seen by the rise in G Loss and fall in SSIM, the generated images start to look less realistic about Epochs 320–330. Even if the image quality has somewhat improved by Epoch 340, more tweaks are need to guarantee consistent progress.

The model's outputs' clarity and structure significantly increase as training moves toward Epochs 100–200. With the noise levels going down and the general coherence getting better, the images start to take on more distinct patterns. While some color distortions still exist, they are not as noticeable as they were in previous eras. This phase's variations in discriminator and generator losses show that, although the model is improving, it also goes through periods of instability during which the produced images' quality briefly declines before rising once more. These oscillations are probably caused by the complex dynamics between the discriminator and generator, as the generator sometimes fails to provide images of a high enough quality.

The model becomes more and more adept at improving the structure and quality of the images it generates in the later phases of training, especially from Epochs 200 to 400. Through the use of realistic textures and sharply defined features, the outputs become crisper and more faithfully mimic the target images. Some epochs exhibit unexpected color tints, suggesting that despite these advances, the model still has issues with color constancy. While the generator has come a long way, these challenges show that it still has troubles from time to time. These drawbacks include overfitting and mode collapse, which cause the generator to produce less diversified outputs with consistent but defective properties.

Epoch 460/500 - D Loss: 0.0075 - G Loss: 2.6461 - MSE: 0.0410 - SSIM: 0.4734



Epoch 470/500 - D Loss: 0.0001 - G Loss: 0.9459 - MSE: 0.0453 - SSIM: 0.4809



Epoch 480/500 - D Loss: 0.0000 - G Loss: 0.0016 - MSE: 0.0412 - SSIM: 0.4790



Epoch 490/500 - D Loss: 0.0009 - G Loss: 0.0175 - MSE: 0.0442 - SSIM: 0.4749



Figure 4.8: The grid displays the generated brain images at training epochs 460 to 490, together with the accompanying Mean Squared Error (MSE), Generator Loss (G Loss), Discriminator Loss (D Loss), and Structural Similarity Index (SSIM). With consistent SSIM and MSE values, these images and metrics demonstrate the stabilization of the model and show that, as the training process approaches its conclusion, the model has successfully captured the structural properties of the brain images.

The model continues to improve its outputs between Epochs 400 and 490, producing images with significant structural accuracy and detail. The resulting photos from these epochs are highly similar to the real high-resolution images, indicating that the model has essentially learned all of the necessary elements to generate realistic images. This stability is still there in the model, even with current improvements: loss spikes and small color inconsistencies are signs of the occasional instability. Given these problems, it appears that while the model can generate high-quality photos, it still lacks overall stability and consistency, meaning more work is necessary to guarantee consistent performance throughout.

The SRGAN model's increasing capacity to produce high-resolution images with more accuracy and detail is demonstrated overall by the training from Epoch 60 to 490. Noise-filled, distorted outputs give way to more refined, structurally accurate images, demonstrating a clear improvement in image quality. For the model to continuously produce high-quality results throughout all epochs, more fine-tuning is necessary, as seen by persistent issues with stability, color consistency, and sporadic mode collapse.

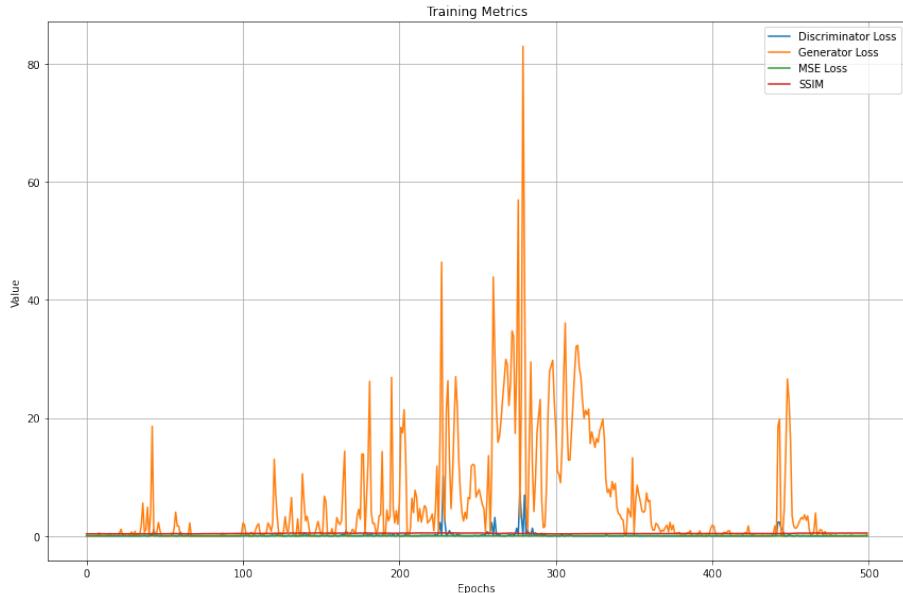


Figure 4.9: Over 500 epochs, the training metrics plot: The Structural Similarity Index (SSIM), Mean Squared Error (MSE), Generator Loss, and Discriminator Loss progressions during the training process are displayed on the graph. An adversarial dynamic between the discriminator and generator is seen in the large oscillations in the Generator Loss, particularly around the halfway point of the training. A comparatively low discriminator loss indicates that actual and produced images may be distinguished from one another effectively. The model appears to be getting better at producing realistic and structurally accurate images as evidenced by the stabilization of MSE and SSIM measures over time.

Significant variations in the Generator Loss are evidently visible in the training metrics distribution graph throughout 500 epochs, particularly between epochs 200 and 350. These abrupt spikes point out times throughout the training process when the generator is unstable and finds it difficult to generate realistic images. This instability shows that the generator sometimes experiences a transient loss in image quality due to difficulties maintaining the competitive balance with the discriminator. The sharp increases in Generator Loss highlight how difficult it is for the model to continuously increase production, with periods of significant difficulties interspersed with periods of recovery.

On the other hand, the Discriminator Loss remains consistent and low during the training phase. This steady low loss indicates that the discriminator learns to discriminate between produced and genuine images early on. Particularly when contrasted with the extreme volatility of Generator Loss, the Discriminator Loss shows no change, suggesting that the discriminator may have dominated during these epochs. Due to the generator's constant attempts to adapt but unable to match the discriminator's strict requirements, which results in the unpredictable loss patterns, this dominance most certainly adds to the instability in the generator's performance.

The generated images' pixel-wise accuracy and structural similarity have steadily improved over time, according to the MSE Loss and SSIM metrics, making them more stable than the Generator Loss. Despite the generator's sporadic instability, these measurements do not exhibit similar volatility of the Generator Loss, indicating that overall image quality is trending upward. But as the steep spikes in Generator Loss show, these gains are occasionally hampered by unstable spells. This shows that even while the model is improving, more work has to be done to make it more dependable and consistent in its performance during the training period.

5 COMPARATIVE ANALYSIS OF DCGAN AND SRGAN USING EVALUATION METRICS

Four main metrics are used to compare the SRGAN and DCGAN models: Mean Squared Error (MSE), Structural Similarity Index (SSIM), and the losses related to the generator and discriminator. These measures are chosen because they provide a thorough evaluation of the models' performance for both perceptual quality and pixel-level precision. The mean squared error (MSE) between the anticipated and actual pixel values is a widely used metric that shows how well the generated images resemble the real ones pixel by pixel. Conversely, SSIM assesses the degree of structural similarity between produced and real images, which makes it useful for determining how well perceptual quality and the preservation of image details are preserved. The generator and discriminator losses, which demonstrate how well the generator can produce convincing images and how successfully the discriminator can discern between produced and actual images, are crucial for comprehending the mechanism of the training process. Taken together, these criteria provide a fair assessment of both models, incorporating not only the technical correctness but also the visual appeal and the efficacy of the competitive training of the generated images.

Table 5.1: Comparison of DCGAN and SRGAN Performance Metrics

Metric	Model	
	DCGAN	SRGAN
Discriminator Loss on Real Images	0.436	N/A
Discriminator Loss on Fake Images	0.198	0.0439
Mean SSIM	0.588	-0.6162
Mean MSE	0.018	2.9408
Generator Loss	2.626	2.9292

Metrics that are presented for both the SRGAN and DCGAN models show a notable variation in their performance, especially with regard to image quality and the efficiency of their generative processes. A stable and well-balanced performance is displayed by the DCGAN model. Real photos and those produced by the model may be distinguished from one another with an effective discriminator loss of 0.436 on real images. Since the discriminator's loss on fake photos is far smaller—0.198—it is possible that the generator is generating results that effectively challenge the discriminator, advancing the process of learning even as it is able to detect created images. There

is a moderate degree of similarities in structure between the generated images and the real images, as indicated by the Mean SSIM of 0.588 obtained by the DCGAN. Furthermore, the average mean square error (MSE) of 0.018 suggests that the generated images have a high level of pixel precision. The results show that the DCGAN can generate visually accurate and structurally cohesive images, together with a Generator Loss of 2.626, highlighting its efficacy in jobs demanding complex image generation.

On the other hand, the metrics of the SRGAN model indicate that it poses considerable difficulties. The lack of discriminator Loss on Real Images implies that this particular component may not have been sufficiently stressed or captured throughout the training process. The discriminator is overconfident in its ability to recognize generated photos as false, though, as evidenced by the extremely low Discriminator Loss on false photos of 0.0439. This implies that the generator is having difficulty creating outputs that accurately replicate high-resolution images, which at first is important for applications requiring the creation of high-fidelity images. The produced images' Mean SSIM of -0.6162 is especially concerning because a negative SSIM suggests that the images are not only structurally different from the real images but also introduce notable distortions. The Mean MSE of 2.9408, which is significantly greater than that of the DCGAN and indicates significant pixel-level differences between the produced and real images, lends more credence to this worry. The problems indicated by the SSIM and MSE measurements are not mitigated by the Generator Loss of 2.9292, while being marginally less than the DCGAN. These findings point to underlying issues with the SRGAN's capacity to produce high-quality images, most likely as a result of unbalanced training or inefficient adversarial dynamics.

Critical distinctions in training processes and overall performance are found between DCGAN and SRGAN. The negative SSIM and high MSE of the SRGAN point to possible problems such mode collapse, in which the generator is unable to provide a variety of high-quality outputs, which results in excessively confident discriminator output and inadequate generator improvement. The extremely low Discriminator Loss on Fake Images in SRGAN could indicate that the discriminator is not receiving enough challenges, which could lead to a feedback loop in which the generator finds it difficult to pick up new skills. On the other hand, the DCGAN's mild SSIM and more balanced discriminator losses point to a more robust adversarial training dynamic in which the discriminator and generator both steadily improve with time.

A further study of these metrics suggests that overfitting or instability may be present in the SRGAN during training. The generator appears to be overfitting to certain patterns in the training data, as indicated by the negative SSIM score and high MSE. This leads to outputs that are significantly distorted and do not generalize well. The overfitting of the model might result in pixel-by-pixel incorrect and architecturally incoherent images, hence significantly reducing its practical use. However, the metrics of the DCGAN indicate to a more generalized learning process, in which the generator can continue to produce outputs that are both structurally comparable to the target images and accurately represent the pixels.

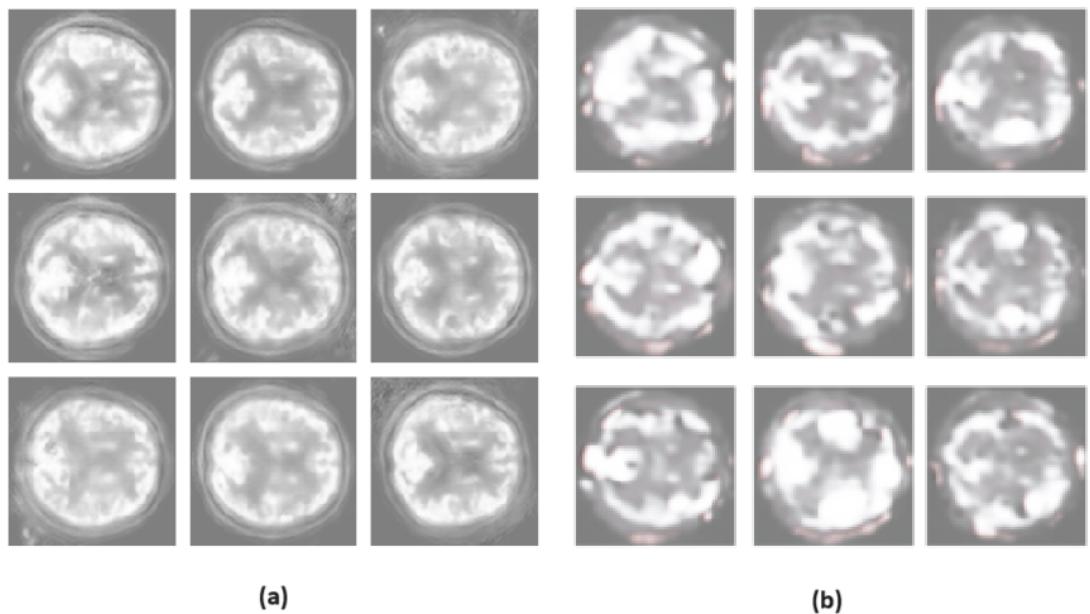


Figure 5.1: Comparing the brain pictures produced by DCGAN and SRGAN When it comes to capturing the finer elements of brain anatomy, DCGAN outperforms SRGAN in terms of structural details and clarity with the images (a) it generates. DCGAN generates images that are higher quality and more realistic than SRGAN, while the SRGAN-generated images (b) are more blurry and less detailed, emphasizing an unexpected scenario.

It is evident from a comparison of the images produced by DCGAN and SRGAN that DCGAN generates images with better clarity and structural integrity. DCGAN produces images with minimal distortion and blurring, consistent texturing, and well-defined details. This shows that DCGAN produces outputs that are more visually consistent and accurate by better maintaining the important details and patterns of the original images. The fact that the structural aspects of the generated images are consistently consistent suggests that DCGAN is more capable of accurately replicating the target data.

The images generated by SRGAN, on the other hand, are significantly less sharp and more blurry than those produced by DCGAN. SRGAN produces outputs that are less aesthetically pleasing and structurally sound because to severe distortion and information loss in their images. The aforementioned negative SSIM and high MSE values are consistent with this drop in image quality, highlighting the challenge SRGAN faces in generating high-resolution images that faithfully capture the original input. In terms of quantitative measures and visual evaluation, DCGAN performs significantly better than SRGAN, exhibiting a more resilient capacity to produce high-quality images with superior structural fidelity.

6 CONCLUSIONS

Using manifold learning approaches, PET scans associated with Alzheimer's illness were analyzed first. Various techniques were applied to decrease the dimensionality of the dataset, identify underlying patterns, and facilitate data clustering. These techniques included Principal Component Analysis (PCA), t-Distributed Stochastic Neighbor Embedding (t-SNE), and Uniform Manifold Approximation and Projection (UMAP). Understanding the high-dimensional linkages within the dataset was made easier by these studies, and this knowledge was crucial for the next phases of the study.

After the manifold learning analysis, Super-Resolution GAN (SRGAN) and Deep Convolutional GAN (DCGAN) were the two GAN architectures that were implemented. The major objective was to evaluate their ability to produce high-quality synthetic PET images that might take the place of actual PET scans, potentially solving issues like the high expense and restricted accessibility of standard PET imaging.

Then, the Mean Squared Error (MSE) and Structural Similarity Index (SSIM), two qualitative and quantitative indicators, were used to assess the performance of these models. The results demonstrated that DCGAN performed noticeably better than SRGAN. DCGAN produced synthetic images with improved clarity, more structural integrity, and details that were maintained, and it demonstrated more robust and consistent training dynamics. On the other hand, SRGAN struggled with problems like mode collapse, overfitting, and instability during training, which led to lower-quality outputs even though its goal was to improve image resolution.

In terms of maintaining the PET images' structural integrity, DCGAN performed better. With its deep convolutional layers, the architecture of DCGAN was created with the express purpose of capturing spatial hierarchies in images, which makes it very skilled at preserving delicate details and textures. This led to the creation of artificial images with clarity and detail that were very similar to actual PET scans—an important characteristic in medical imaging where precision can have a direct impact on the results of a diagnosis.

On the other hand, because of its increased complexity, SRGAN faced considerable difficulties during training even though it was designed to improve image resolution. SRGAN's super-

resolution-focused architecture made the model more prone to instability and overfitting, which reduced its performance. In comparison to DCGAN, SRGAN produced images with worse quality, as evidenced by its greater Mean Squared Error (MSE) and negative Structural Similarity Index (SSIM) values.

To conclude, it was shown that DCGAN outperformed SRGAN in generating PET images for studies on Alzheimer's disease. The success of DCGAN was largely due to its higher stability, greater preservation of structural information, and lesser complexity handling.

6.1 Evaluation

Designing and implementing a GAN architecture appropriate for synthesizing PET pictures associated with Alzheimer's disease was the main goal of this research. The creation of both Deep Convolutional GAN (DCGAN) and Super-Resolution GAN (SRGAN) successfully achieved this goal. In particular, DCGAN proved to be quite successful in generating realistic-looking synthetic PET pictures with excellent quality. The initiative provided important insights into each model's suitability for medical image synthesis by carefully examining its strengths and weaknesses. The research effectively illustrated the potential of GANs in producing realistic medical images, even if SRGAN encountered issues with instability and image quality.

Optimising model parameters to mitigate underfitting and overfitting was a crucial component of the research. Through methodical experimentation with learning rates, batch sizes, and epoch counts, the models were optimized for best results. To guarantee an equilibrium between training and validation performance, regularization approaches and rigorous model complexity optimization were used. By employing these techniques, DCGAN was able to efficiently address the issues of underfitting and overfitting that frequently occur in GAN models, while maintaining stable training dynamics and consistently producing high-quality images.

A professional approach and a well-structured methodology in line with the set objectives were maintained throughout the project. The superiority of DCGAN in preserving image fidelity was validated by a qualitative examination of synthetic image quality employing measures such as Mean Squared Error (MSE) and Structural Similarity Index (SSIM). The clinical utility of synthetic PET pictures was proven through a successful comparison with real scans, especially for research on Alzheimer's disease. Overall, the study shown great potential for enhancing diagnostic

tools using sophisticated GAN structures, not only fulfilling its main goals but also laying a solid platform for future work in the field of medical picture synthesis.

6.2 Future Work

Future research will concentrate on improving SRGAN's training dynamics in order to overcome the instability that is typified by irregular image quality and unpredictable generator loss spikes. Sophisticated strategies will be used to stabilize adversarial training, including the incorporation of adaptive learning rates, gradient penalty techniques, and enhanced loss functions designed especially for high-resolution medical imaging. Further research into other GAN designs such as StyleGAN or CycleGAN will guarantee better performance in producing PET images of high quality and will successfully tackle the drawbacks of SRGAN. Better and more accurate diagnostic tools for Alzheimer's disease will result from the integration of multi-modal data, such as MRI or CT scans, into the synthesis images. A unique loss function that more closely correlates with medical picture quality will either replace or refine the perceptual loss function now used in SRGAN, which is generated from a pre-trained VGG19 network. In addition, the Variational Autoencoder GAN (VAEGAN) will be implemented, since it combines the regularization advantages of VAEs with the generative power of GANs, leading to more stable training and better image synthesis, especially for complicated medical imaging tasks that are crucial for furthering research on Alzheimer's disease.

BIBLIOGRAPHY

- [1] Ahmad, W. , Ali, H. , Shah, Z. , and Azmat, S. . A new generative adversarial network for medical images super resolution. *Scientific Reports*, 12, 12 2022. ISSN 20452322. doi: 10.1038/s41598-022-13658-4.
- [2] Al-Ameen, Z. and Sulong, G. . Prevalent degradations and processing challenges of computed tomography medical images: A compendious analysis. *International Journal of Grid and Distributed Computing*, 9:107–118, 11 2016. doi: 10.14257/ijgdc.2016.9.10.10.
- [3] ALZORG. Stages of Alzheimer’s Disease. <https://www.alz.org/alzheimers-dementia/stages>. Accessed: 01-07-2024.
- [4] ALZUK. Progression and Stages of Dementia. <https://www.alzheimers.org.uk/about-dementia/symptoms-and-diagnosis/how-dementia-progresses/progression-stages-dementia>. Accessed: 01-07-2024.
- [5] AnalyticsVidhyaPCA. PCA and Its Underlying Mathematical Principles. <https://www.analyticsvidhya.com/blog/2021/09/pca-and-its-underlying-mathematical-principles/>. Accessed: 19-08-2024.
- [6] Carreira-Perpiñán, M. A. , Perpiñán, P. , and Zemel, R. S. . Proximity graphs for clustering and manifold learning.
- [7] clevel clininc. CT (Computed Tomography) Scan. <https://my.clevelandclinic.org/health/diagnostics/4808-ct-computed-tomography-scan>. Accessed: 02-07-2024.
- [8] ClevelandClinicPET. PET Scan: What It Is, Types, Purpose, Procedure Results. <https://my.clevelandclinic.org/health/diagnostics/10123-pet-scan>. Accessed: 21-08-2024.
- [9] DCGANIdiotDeveloper. What is deep convolutional generative adversarial networks (dcgan)? <https://idiotdeveloper.com/>

what-is-deep-convolutional-generative-adversarial-networks-dcgan/. Accessed: 12-07-2024.

- [10] DeTure, M. A. and Dickson, D. W. . The neuropathological diagnosis of Alzheimer's disease. *Molecular Neurodegeneration*, 14(32), 2019. doi: 10.1186/s13024-019-0333-5.
- [11] EducbaSRGAN. Super Resolution Generative Adversarial Network (SRGAN). <https://www.educba.com/super-resolution-generative-adversarial-network-srgan/>. Accessed: 15-08-2024.
- [12] FDG. Fluorodeoxyglucose (18F). [https://en.wikipedia.org/wiki/Fluorodeoxyglucose_\(18F\)](https://en.wikipedia.org/wiki/Fluorodeoxyglucose_(18F)). Accessed: 26-08-2024.
- [13] Goodfellow, I. . Nips 2016 tutorial: Generative adversarial networks. 12 2016. URL <http://arxiv.org/abs/1701.00160>.
- [14] Goodfellow, I. J. , Pouget-Abadie, J. , Mirza, M. , Xu, B. , Warde-Farley, D. , Ozair, S. , Courville, A. , and Bengio, Y. . Generative adversarial networks. 6 2014. URL <http://arxiv.org/abs/1406.2661>.
- [15] Guerreiro, J. , Tomás, P. , Garcia, N. , and Aidos, H. . Super-resolution of magnetic resonance images using generative adversarial networks, 9 2023. ISSN 18790771.
- [16] HOPKINS. Stages of Alzheimer Disease. <https://www.hopkinsmedicine.org/health/conditions-and-diseases/alzheimers-disease/stages-of-alzheimer-disease>. Accessed: 01-07-2024.
- [17] Izenman, A. J. . Introduction to manifold learning. *Wiley Interdisciplinary Reviews: Computational Statistics*, 4:439–446, 9 2012. ISSN 19395108. doi: 10.1002/wics.1222.
- [18] Izenman, A. J. . Introduction to manifold learning. *WIREs Computational Statistics*, 4 (5):439–446, 2012. doi: <https://doi.org/10.1002/wics.1222>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/wics.1222>.
- [19] JavatpointDementia. Dementia Definition - Javatpoint. <https://www.javatpoint.com/dementia-definition>. Accessed: 12-07-2024.
- [20] John, R. , Penning, J. , Chandler, H. , Fielding, P. , Marshall, C. , and Smith, R. . Quantitative evaluation of synthesized brain pet using a variational autoencoder. In *2021 IEEE*

Nuclear Science Symposium and Medical Imaging Conference Record, NSS/MIC 2021 and 28th International Symposium on Room-Temperature Semiconductor Detectors, RTSD 2022. Institute of Electrical and Electronics Engineers Inc., 2021. ISBN 9781665421133. doi: 10.1109/NSS/MIC44867.2021.9875443.

- [21] Johns Hopkins Medicine. Magnetic Resonance Imaging (MRI). <https://www.hopkinsmedicine.org/health/treatment-tests-and-therapies/magnetic-resonance-imaging-mri>. Accessed: 02-07-2024.
- [22] Kherif, F. and Latypova, A. . *Principal component analysis*, pages 209–225. Elsevier, 1 2019. ISBN 9780128157398. doi: 10.1016/B978-0-12-815739-8.00012-2.
- [23] Lameka, K. and Farwell, M. D. . Positron emission tomography, 2016.
- [24] Li, Y. , Xiao, D. , and Wei, X. . Isomap Algorithm for Adaptive Neighborhood Selection. *Journal of Hebei Institute of Architecture and Civil Engineering*, (2):110–112, 2014.
- [25] Lin, T. and Zha, H. . Riemannian manifold learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(5):796–809, 2008. doi: 10.1109/TPAMI.2007.70735.
- [26] Ma, Y. and Fu, Y. . *Manifold learning theory and applications*. CRC ; Taylor Francis [distributor], 2012. ISBN 9781439871096.
- [27] Mayo clinic. CT scan: About. <https://www.mayoclinic.org/tests-procedures/ct-scan/about/pac-20393675>. Accessed: 02-07-2024.
- [28] McInnes, L. , Healy, J. , and Melville, J. . Umap: Uniform manifold approximation and projection for dimension reduction. 2 2018. URL <http://arxiv.org/abs/1802.03426>.
- [29] MediumUMAP. UMAP: An Alternative Dimensionality Reduction Technique. <https://medium.com/mcd-unison/umap-an-alternative-dimensionality-reduction-technique-7a5e77e80982>. Accessed: 19-08-2024.
- [30] Mohamed, A. O. , Salih, K. H. , and Ali, H. H. . Super resolution of medical images using generative adversarial networks. In *Proceedings of: 2020 International Conference on Computer, Control, Electrical, and Electronics Engineering, ICCCEEE 2020*. In-

- stitute of Electrical and Electronics Engineers Inc., 2 2021. ISBN 9781728191119. doi: 10.1109/ICCCEEE49695.2021.9429656.
- [31] MRI. How does MRI work? <https://magneticresonance.weebly.com/how-does-mri-work.html>. Accessed: 18-08-2024.
- [32] Muehllehner, G. and Karp, J. S. . Positron emission tomography, 7 2006. ISSN 00319155.
- [33] NCRI. UK PET Core Lab . http://www.ncri-pet.org.uk/about_pet_fdg.php. Accessed: 08-06-2024.
- [34] NHS. Causes of Dementia. <https://www.nhs.uk/conditions/dementia/about-dementia/causes/>. Accessed: 13-07-2024.
- [35] NIA. National Institute on Aging. <https://www.nia.nih.gov/health/alzheimers-causes-and-risk-factors/what-happens-brain-alzheimers-disease#:~:text=At%20first%2C%20Alzheimer's%20usually%20damages,%2C%20reasoning%2C%20and%20social%20behavior>. Accessed: 13-07-2024.
- [36] OklahomaPETCT. What is a PET/CT Scan? <https://www.oklahomapetscan.com/pet-ct-scan/what-is-a-pet-ct-scan.php>. Accessed: 19-08-2024.
- [37] Parmar, H. , Nutter, B. , Long, R. , Antani, S. , and Mitra, S. . Visualizing temporal brain-state changes for fmri using t-distributed stochastic neighbor embedding. *Journal of Medical Imaging*, 8, 8 2021. ISSN 23294310. doi: 10.1117/1.jmi.8.4.046001.
- [38] Radford, A. , Metz, L. , and Chintala, S. . Unsupervised representation learning with deep convolutional generative adversarial networks, 2016. URL <https://arxiv.org/abs/1511.06434>.
- [39] S, V. , N, L. , M, S. , Manmatti, S. , K P, A. R. , and S, G. . Enhancing medical imaging resolution: Exploring srgan for high-quality medical image reconstruction. In *2024 IEEE International Conference for Women in Innovation, Technology Entrepreneurship (ICWITE)*, pages 1–8, 2024. doi: 10.1109/ICWITE59797.2024.10503215.
- [40] Scikit-Learn Developers. Manifold Learning. <https://scikit-learn.org/stable/modules/manifold.html>. Accessed: 02-07-2024.

- [41] Shi, S. , Xu, Y. , Xu, X. , Mo, X. , and Ding, J. . A preprocessing manifold learning strategy based on t-distributed stochastic neighbor embedding. *Entropy*, 25, 7 2023. ISSN 10994300. doi: 10.3390/e25071065.
- [42] Sood, R. , Topiwala, B. , Choutagunta, K. , Sood, R. , and Rusu, M. . An application of generative adversarial networks for super resolution medical imaging. In *Proceedings - 17th IEEE International Conference on Machine Learning and Applications, ICMLA 2018*, pages 326–331. Institute of Electrical and Electronics Engineers Inc., 7 2018. ISBN 9781538668047. doi: 10.1109/ICMLA.2018.00055.
- [43] SpotIntelligenceTSNE. Practical Guide to t-SNE. <https://spotintelligence.com/2023/12/22/practical-guide-t-sne/>. Accessed: 19-08-2024.
- [44] Talbot, J. N. , Haioun, C. , Rain, J. D. , Meignan, M. , Wioland, M. , Misset, J. L. , Grahek, D. , Kerrou, K. , and Montravers, F. . [18f]-fdg positron imaging in clinical management of lymphoma patients, 2001. URL www.elsevier.com/locate/critrevonc.
- [45] Tharwat, A. . Principal component analysis-a tutorial. *International Journal of Applied Pattern Recognition*, 3(3):197–240, 2016.
- [46] Weaver, J. B. . Principal components analysis in medical imaging.
- [47] WIKIPEDIA. Alzheimer’s Disease. https://en.wikipedia.org/wiki/Alzheimer%27s_disease. Accessed: 13-07-2024.
- [48] Wikipedia contributors. Isomap. <https://en.wikipedia.org/wiki/Isomap>. Accessed: 02-07-2024.
- [49] Wikipedia contributors. Magnetic Resonance Imaging. https://en.wikipedia.org/wiki/Magnetic_resonance_imaging, 2024. Accessed: 02-07-2024.
- [50] Yi, W. , Bu, S. , Lee, H. H. , and Chan, C. H. . Comparative analysis of manifold learning-based dimension reduction methods: A mathematical perspective. *Mathematics*, 12, 8 2024. ISSN 22277390. doi: 10.3390/math12152388.
- [51] Zhu, B. , Liu, J. Z. , Cauley, S. F. , Rosen, B. R. , and Rosen, M. S. . Image reconstruction by domain-transform manifold learning. *Nature*, 555:487–492, 3 2018. ISSN 14764687. doi: 10.1038/nature25988.

APPENDIX

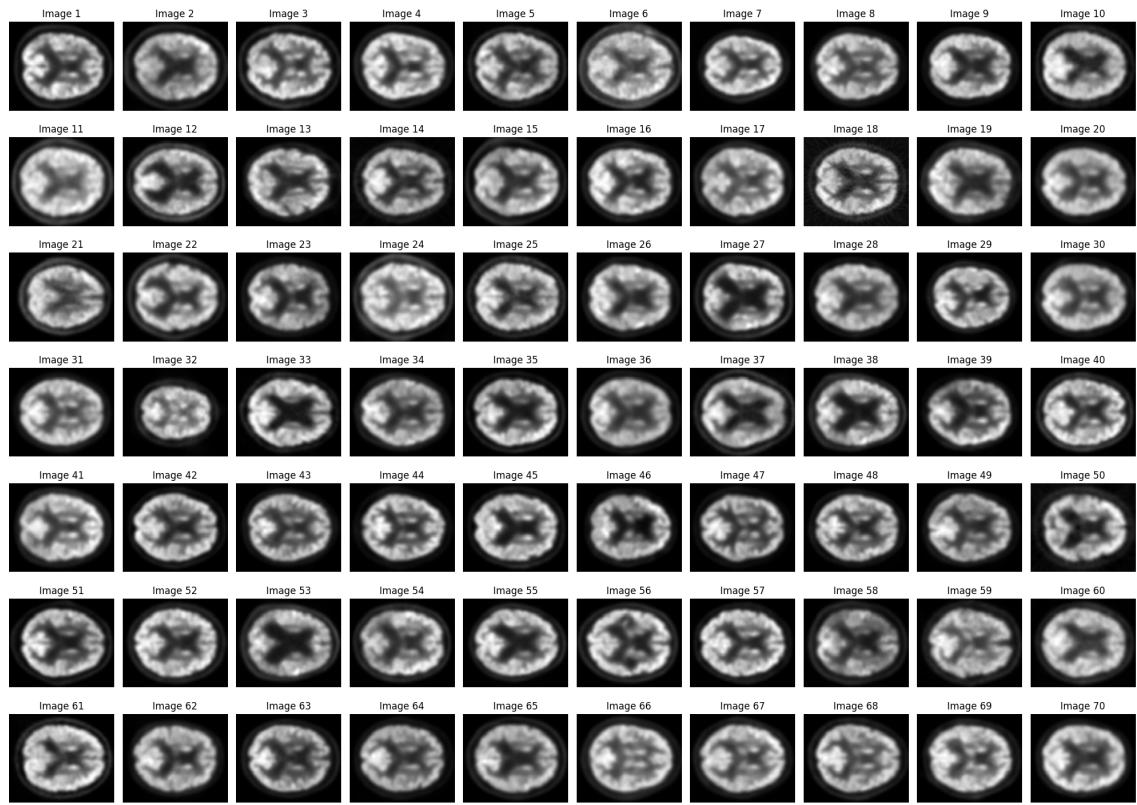


Figure 6.1: A few real images from the dataset.

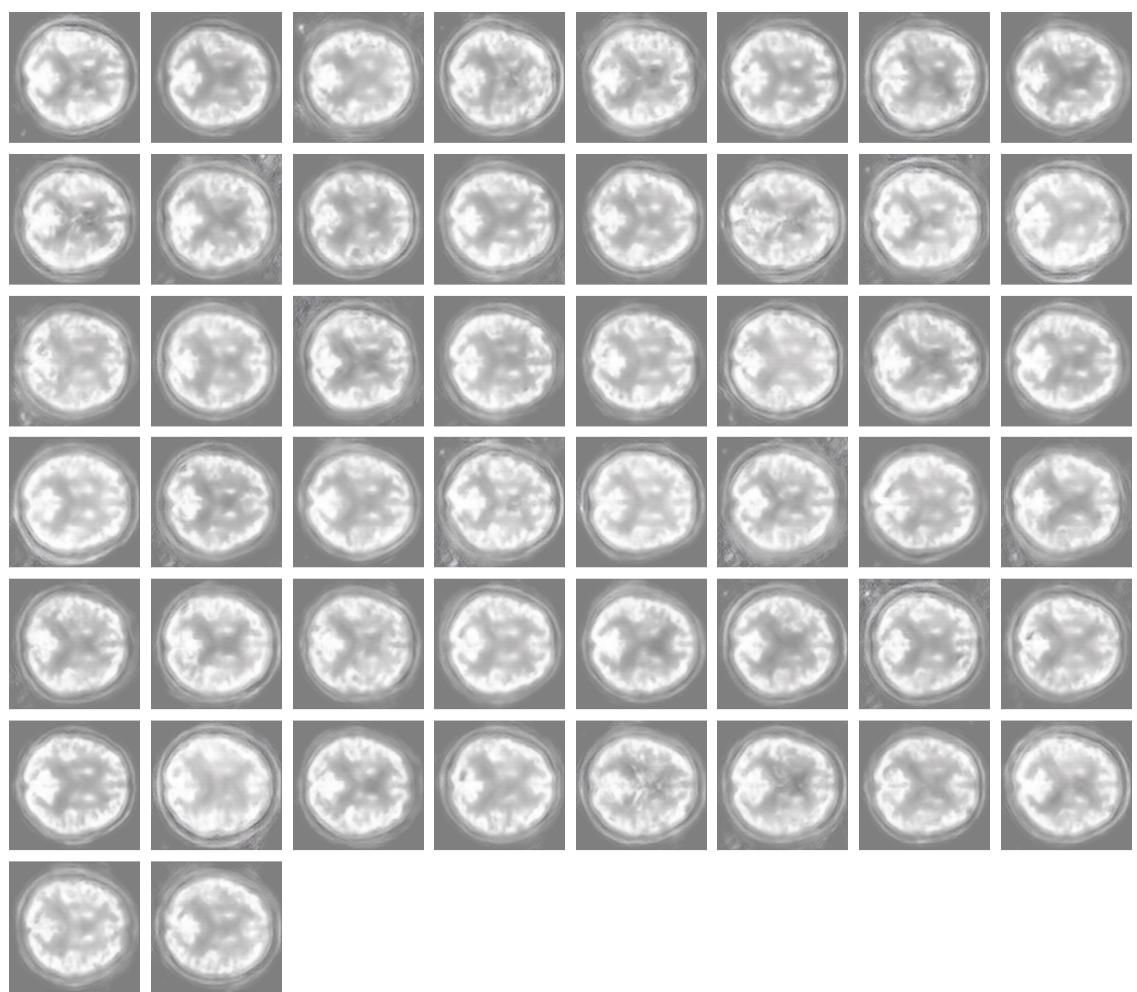


Figure 6.2: Generated DCGAN images.

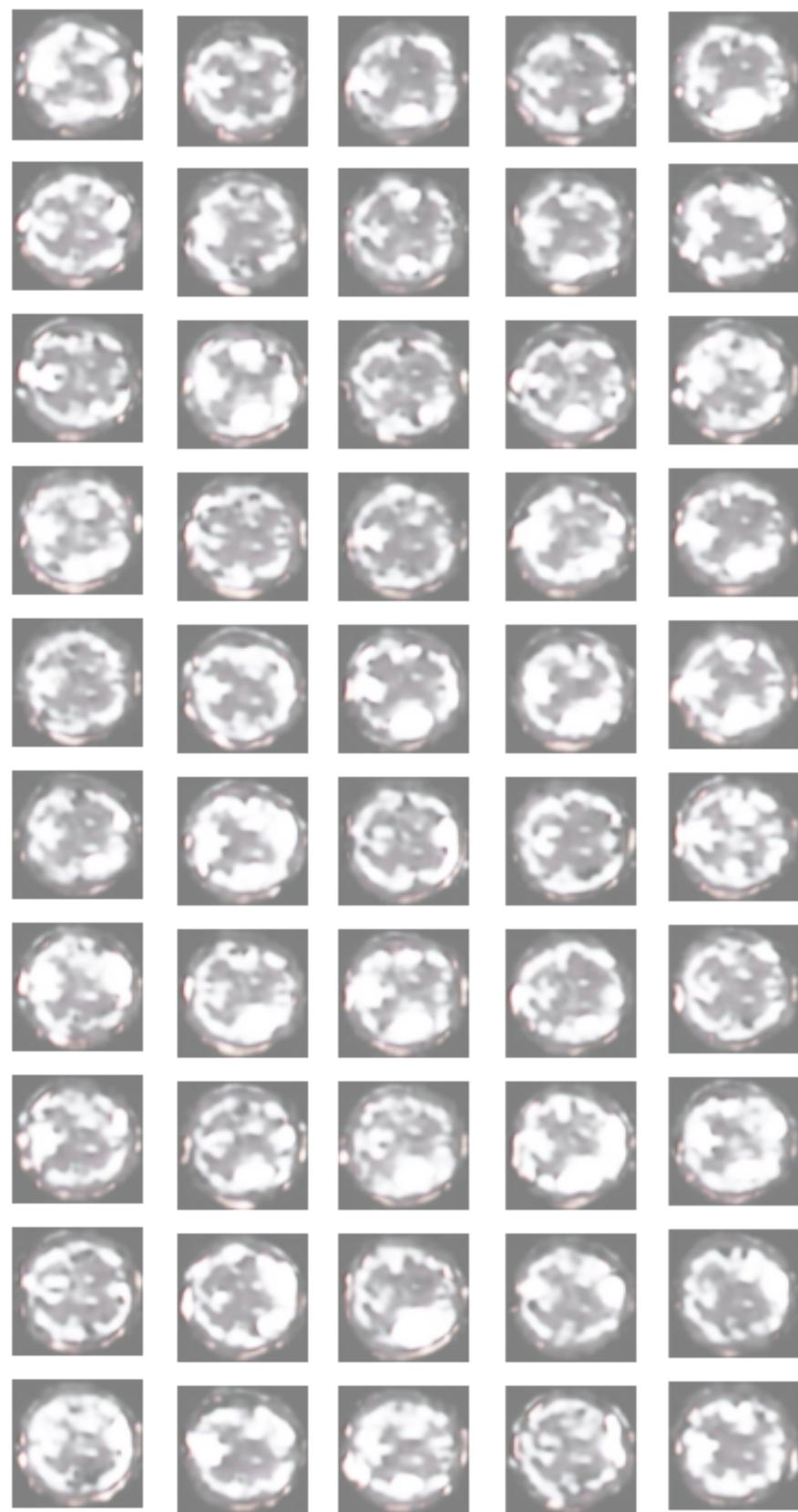


Figure 6.3: Generated SRGAN images.

Link for the codes:

To access the dataset(.nii format) on Google Drive: Original Dataset.

To access the processed dataset(.png format) on Google Drive: Processed Dataset.

To access the programming codes on Google Drive: Final dissertation.