

Universidade de Brasília – UnB
Campus Gama – FGA
Engenharia de Software

**IMAGEAMENTO POR RESSONÂNCIA MAGNÉTICA
COM BASE EM COMPRESSIVE SENSING COM PRÉ-FILTRAGEM
E REDES GENERATIVAS ADVERSÁRIAS**

DAVI ALVES BEZERRA

Orientador: Dr. CRISTIANO JACQUES MIOSSO
Coorientador: BSc. GABRIEL GOMES ZIEGLER



UNB – UNIVERSIDADE DE BRASÍLIA

FGA – FACULDADE GAMA

ENGENHARIA DE SOFTWARE

**IMAGEAMENTO POR RESSONÂNCIA MAGNÉTICA COM BASE EM
COMPRESSIVE SENSING COM
PRÉ-FILTRAGEM E REDES GENERATIVAS ADVERSÁRIAS**

DAVI ALVES BEZERRA

ORIENTADOR: CRISTIANO JACQUES MIOSSO

CO-ORIENTADOR: GABRIEL GOMES ZIEGLER

TRABALHO DE CONCLUSÃO DE CURSO

ENGENHARIA DE SOFTWARE

BRASÍLIA/DF, FEVEREIRO DE 2023

UNB – UNIVERSIDADE DE BRASÍLIA
FGA – FACULDADE GAMA
ENGENHARIA DE SOFTWARE

IMAGEAMENTO POR RESSONÂNCIA MAGNÉTICA COM BASE EM
COMPRESSIVE SENSING COM
PRÉ-FILTRAGEM E REDES GENERATIVAS ADVERSÁRIAS

DAVI ALVES BEZERRA

TRABALHO DE CONCLUSÃO DE CURSO SUBMETIDO À FACULDADE UNB GAMA DA
UNIVERSIDADE DE BRASÍLIA, COMO PARTE DOS REQUISITOS NECESSÁRIOS PARA A
OBTENÇÃO DO GRAU DE BACHAREL EM ENGENHARIA DE SOFTWARE

APROVADA POR:

Cristiano Jacques Miosso

(Orientador)

Vinícius do Couto Pinheiro, Ms.C.

Filipe Emídio Tôrres, Ms.C.

FICHA CATALOGRÁFICA

BEZERRA, DAVI

Imageamento por Ressonância Magnética com base em Compressive Sensing com Pré-Filtragem e Redes Generativas Adversárias, [Distrito Federal], 2018.

44p., 210 × 297 mm (FGA/UnB Gama, Bacharelado em Engenharia de Software, 2018).

Trabalho de Conclusão de Curso, Faculdade UnB Gama, Engenharia de Software

- | | |
|-------------------------------|----------------------------------|
| 1. Imageamento Médico | 2. Aprendizagem de Máquina |
| 3. <i>Compressive Sensing</i> | 4. Redes Generativas Adversárias |
| I. FGA UnB/UnB. | II. Título (série) |

REFERÊNCIA

BEZERRA, DAVI (2023). Imageamento por Ressonância Magnética com base em Compressive Sensing com Pré-Filtragem e Redes Generativas Adversárias. Trabalho de Conclusão de Curso, Engenharia de Software, Faculdade UnB Gama, Universidade de Brasília, Brasília, DF, 44p.

CESSÃO DE DIREITOS

AUTOR: Davi Alves Bezerra

TÍTULO: Imageamento por Ressonância Magnética com base em Compressive Sensing com Pré-Filtragem e Redes Generativas Adversárias

GRAU: Bacharel em Engenharia de Software

ANO: 2023

É concedida à Universidade de Brasília permissão para reproduzir cópias desta monografia de conclusão de curso e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte desta monografia pode ser reproduzida sem a autorização por escrito do autor.

davialvb@gmail.com

Brasília, DF – Brasil

RESUMO ESTENDIDO

A imagem por Ressonância Magnética (do inglês, Magnetic Resonance) (MR) é uma ótima escolha para detecção de doenças pulmonares e tumores com capacidade de diferenciar as características dos tecidos biológicos. Quando comparada com outras modalidades de imagem, tem vantagens pela ausência de radiação e bom contraste de tecidos moles. No entanto, uma desvantagem é o tempo que leva para escanear e reconstruir a imagem, limitando seu uso em algumas aplicações clínicas onde o tempo de imagem é crítico. O uso de Compressed Sensing (CS) na reconstrução acelera a aquisição, porém sofre de um processo iterativo longo e artefatos típicos que dificultam a interpretação da imagem. Quanto menor a taxa de sub amostragem nas imagens usadas para reconstrução, maior é a quantidade de artefato.

Com os avanços no uso de Aprendizagem Profunda (do inglês, Deep Learning) (DL) em processamento de imagens médicas tivemos várias propostas de modelo generativo para encontrar e aprender automaticamente os padrões de uma imagem de MR. Um modelo generativo é usado para gerar novos exemplos extraídos do conjunto de dados original, a partir de um ruído. Atualmente, as arquiteturas Rede Adversaria Generativa (do inglês, Generative Adversarial Network) (GAN) têm apresentado excelente desempenho na modelagem de distribuições, podendo ser usada para aplicação da técnica de informação a priori determinística em CS. A GAN também pode gerar dados sem modelar explicitamente sua função de densidade de probabilidade e gerar dados semelhantes aos da imagem alvo levando menos tempo do que o CS.

Neste trabalho apresentamos uma configuração de GAN que ao invés de receber como entrada apenas um ruído z , ela recebe uma imagem amostrada no espaço k com um número de trajetórias radiais. A versão condicional do modelo generativo tradicional aprende a representação de x condicionada a um rótulo y . Após o Rede Adversaria Generativa Condicional (do inglês, Conditional Generative Adversarial Network) (CGAN) métodos como Image-to-Image foram introduzidos com objetivo de transformar imagens para terem o estilo (ou características) de outra imagem. Isto é, podemos a partir de imagens sub amostradas no domínio da frequência, usando um número de amostras treinar o modelo generativo para aprender a aproximar a distribuição da imagem rotulada, a partir da imagem reconstruída.

Os resultados dos nossos experimentos mostram que houve um ganho nas métricas objetivas quando comparamos imagens geradas pela CGAN e as imagens reconstruídas por CS quando a taxa de amostragem é menor do que 10%. Além disso, nossos testes com diferentes números de trajetórias mostram as limitações encontradas em modelos generativos na tarefa de reconstrução. Durante a execução da pesquisa, levantamos também outras hipóteses de melhorias para o processo de geração das imagens e um melhor uso para as imagens geradas. Para o processo de melhoria na geração das imagens há a possibilidade

de aumentar o número de imagens com técnicas de rotação e deslocamento do conjunto de treinamento para o gerador poder aprender as representações de cada imagem em várias posições diferentes, aprendendo assim a variação que cada imagem de MR pode ter. Já um possível uso para as imagens geradas seria o uso do modelo proposto com objetivo de gerar pontos de suporte na técnica de informação a priori em reconstruções com CS.

ABSTRACT

Magnetic resonance imaging (MRI) is an excellent choice for detecting various diseases in clinical analysis. When compared to other imaging modalities, it has advantages due to the absence of radiation and good contrast of soft tissues. However, a disadvantage is the time it takes to scan and reconstruct the image, limiting its use in some clinical applications where image time is critical. The use of compressed sensing (CS) in reconstruction accelerates acquisition, but suffers from a long iterative process and typical artifacts that make image interpretation difficult. The lower the subsampling rate in the images used for reconstruction, the more artifacts there are.

With advances in the use of deep learning (DL) in medical image processing, several proposals for generative models have been made to automatically find and learn the patterns of an MRI image. A generative model is used to generate new examples extracted from the original data set, from noise. Currently, generative adversarial networks (GANs) have shown excellent performance in modeling prior distributions of images. GANs can also generate data without explicitly modeling their probability density function and generate data similar to the target image in less time than CS.

In this work, we present a configuration of GAN that, instead of taking noise as input, takes a subsampled image in k-space with a number of radial trajectories. Named conditional GAN (CGAN), it uses various techniques to obtain another image from an input image, known as image-to-image. That is, from subsampled images, we trained the generative model to learn to approximate the output image distribution to a target image, i.e., an image that we consider to be the true or target image.

The results of our experiments show a gain in objective metrics when comparing images generated by CGAN and images reconstructed by CS when the subsampling rate is less than 10%. In addition, our tests with different numbers of trajectories show the limitations encountered in generative models in the reconstruction task. During the research, we also raised other hypotheses for improvements to the image generation process and better use of the generated images. To improve the image generation process, there is the possibility of increasing the number of images with rotation and displacement techniques of the training set so that the generator can learn the representations of each image in several different positions, thus learning the variation that each MRI image can have. The better use of the generated images would be to use the CGAN to generate support points as prior information for reconstructions with CS.

SUMÁRIO

Resumo Estendido	v
Abstract	vii
1 Introdução	1
1.1 Contextualização	2
1.2 Definição do Problema Científico e Proposta de Pesquisa	3
1.3 Objetivos	4
1.3.1 Objetivo Geral	4
1.3.2 Objetivos Específicos	5
2 Fundamentação teórica e Estado da Arte	6
2.1 Imageamento por Ressonância Magnética	6
2.2 Amostragem	7
2.2.1 Espaço k	7
2.2.2 Trajetórias de amostragem	8
2.3 Compressed Sensing	9
2.3.1 Introdução	9
2.4 Redes Neurais Artificiais	11
2.4.1 Aprendizado Supervisionado e Não Supervisionado	12

2.4.2	Gradiente Descendente	13
2.4.3	Retro propagação	14
2.5	ConvNets	14
2.6	Modelos Generativos	15
2.6.1	Autoencoder	16
2.6.2	Redes Adversárias Generativas	17
2.6.3	Rede Adversária Generativa Condicional	18
3	 Materiais e Métodos	19
3.1	Informações do equipamento	19
3.2	Aquisição de medidas sub amostradas no espaço k	19
3.3	Base de dados	20
3.3.1	Rede de Pesquisa em Informática Biomédica	20
3.3.2	Preparação da base de dados	21
3.4	Arquitetura Rede Adversaria Generativa Condicional (do inglês, Conditional Generative Adversarial Network) (CGAN)	23
3.4.1	Gerador	24
3.4.2	Discriminador	26
3.4.3	Função de Ativação ReLU	27
3.4.4	Função de Otimização Adam	27
3.4.5	Função de perda BCE	28
3.4.6	Erro Absoluto Médio	28
3.5	Métricas	28
3.5.1	Erro quadrático médio	29

3.5.2	Erro quadrático médio normalizado	29
3.5.3	Relação Sinal-Ruído de Pico	29
3.5.4	Relação sinal-ruído	30
3.6	Método Proposto para Reconstrução de Imagens de Ressonância Magnética (do inglês, Magnetic Resonance) (MR) Usando Compressed Sensing (CS) e CGAN	30
3.7	Experimentos de Avaliação do Modelo de Reconstrução Proposto	31
4	Resultados e discussões	32
4.1	Resultados de reconstrução utilizando base de treinamento e teste com 30 trajetórias radiais	32
4.2	Resultados de reconstrução utilizando 30 trajetórias radiais na base de treinamento e teste com 100	33
4.3	Resultados de reconstrução utilizando 10 trajetórias radiais na base de treinamento e teste com 30	34
4.3.1	Resultados testando o mesmo modelo utilizando 10 trajetórias ra- diais na base de teste	36
5	Conclusão	38
5.1	Trabalhos futuros	38

LISTA DE TABELAS

4.1	Valores das métricas objetivas de avaliação do modelo testado com imagens sub amostradas por 30 trajetórias radiais.	33
4.2	Valores das métricas objetivas de avaliação do modelo testado com imagens sub amostradas por 100 trajetórias radiais.	33
4.3	Valores das métricas objetivas de avaliação do modelo testado com imagens sub amostradas por 10 trajetórias radiais	37

LISTA DE FIGURAS

2.1	Espaço k (esquerda) e imagem original (direita)	8
2.2	Trajétórias do espaço k . (a) Cartesiana, (b) radial, (c) espiral , (d) amostras isoladas no espaço k, conforme a realização de um processo aleatório [1].	9
2.3	Componentes de um neurônio biológico	11
2.4	Elementos Básicos de um Neurônio Artificial de McCulloch e Pitts	12
2.5	Exemplo visual dos métodos de aprendizagem supervisionada e não supervisionada. Figura obtida de: [2]	13
2.6	A imagem de entrada é codificada em uma representação compactada e, em seguida, decodificada gerando a reconstrução [3].	16
3.1	Exemplo 2D dos dados amostrados no domínio da frequência. Trajetória com 10 linhas radiais (Esquerda), Trajetória com 30 linhas radiais (Direita)	20
3.2	Trajétória com 100 linhas radiais	20
3.3	Diferentes cortes de um mesmo exame de Ressonância Magnética	21
3.4	Exemplos de imagens retiradas da décima sexta posição do grupo de cortes de um exame de MR	22
3.5	Trajétória com 10 linhas radiais.	23
3.6	Trajétória com 30 linhas radiais.	23
3.7	Trajétória com 100 linhas radiais.	23

3.8	Treinamento da CGAN para mapear Imagem amostrada \rightarrow Imagem real. O discriminador representado por D , aprende a classificar entre falso (sintetizado pelo gerador), x e $G(x)$ e tuplas reais x e y . O gerador representado por G , aprende a enganar o discriminador. Ao contrário do modelo generativo incondicional, tanto o gerador quanto o discriminador observam o mapa de borda de entrada.	24
3.9	A U-NET recebe como entrada uma imagem, que na CGAN é a imagem sub amostrada e tenta aproximar a saída da imagem original. Cada camada é representada pelas cores com uma ou mais operações a fim de simplificar a visualização.	25
3.10	O modelo PatchGAN recebe as imagens reais e as geradas em um momento e em outro momento as imagens reais e as amostradas e a partir disso ele retorna uma imagem onde por todos os pixels ele pode definir se as entradas são reais ou falsas	26
4.1	Exemplo visual das reconstruções em relação à imagem real onde, (a) Reconstrução CS; (b) Imagem Real; (c) Reconstrução CGAN	32
4.2	Zoom feito na área da imagem que mostra alguns detalhes: (a) Reconstrução CS; (b) Imagem Real; (c) Reconstrução CGAN	33
4.3	Exemplo visual das reconstruções em relação à imagem real, onde: (a) Reconstrução CS; (b) Imagem Real; (c) Reconstrução CGAN.	34
4.4	Zoom feito na área da imagem que mostra alguns detalhes, onde: (a) Reconstrução CS; (b) Imagem Real; (c) Reconstrução CGAN.	34
4.5	Exemplo visual da imagem sub amostrada no espaço k com 30 trajetórias radiais (a) Imagem original; (b) Reconstrução CS a partir da trajetória radial com 10 ângulos;	35
4.6	Exemplo visual das reconstruções em relação à imagem real, onde: (a) Reconstrução CS; (b) Imagem Real; (c) Reconstrução CGAN	35
4.7	Exemplo visual das reconstruções em relação à imagem real, onde: (a) Reconstrução CS; (b) Imagem Real; (c) Reconstrução CGAN	36
4.8	Exemplo visual das reconstruções em relação à imagem real, onde: (a) Reconstrução CS; (b) Imagem Real; (c) Reconstrução CGAN.	36

NOMENCLATURAS E ABREVIACOES

MRI Imageamento por Ressonância Magnética (do inglês, Magnetic Resonance Imaging)

MR Ressonância Magnética (do inglês, Magnetic Resonance)

ANS Agência Nacional de Saúde Suplementar

MOS Pontuação Media de Opinião (do inglês, Mean Opinion Score)

DL Aprendizagem Profunda (do inglês, Deep Learning)

GAN Rede Adversaria Generativa (do inglês, Generative Adversarial Network)

CNN Rede Neural Convolutacional (do inglês, Convolutional Neural Network)

IA Inteligência Artificial

CS Compressed Sensing

RIP Propriedade de Isometria Restrita (do inglês, Restricted Isometry Property)

SNR Relação sinal-ruído (do inglês, Signal-to-Noise Ratio - SNR)

SSIM Semelhança estrutural de escala única (do inglês, Single-Scale Structural Similarity)

PSNR Relação Sinal-Ruído de Pico (do inglês, Peak Signal-to-Noise Ratio)

CGAN Rede Adversaria Generativa Condicional (do inglês, Conditional Generative Adversarial Network)

DCGAN Rede Adversarial Gerativa Convolutacional Profunda (do inglês, Deep Convolutional Generative Adversarial Network)

ReLU Unidade Linear Retificada (do inglês, Rectified Linear Units)

SGD Descida de Gradiente Estocástico (do inglês, Stochastic Gradient Descent)

DCT Transformada Cosseno Discreta (do inglês, Discrete Cosine Transform)

FFT Transformação rápida de Fourier (do inglês, Fast Fourier Transform)

TC Tomografia Computadorizada

NMR Ressonância Magnética Nuclear (do inglês, Nuclear Magnetic Resonance)

TR Tempo de Repetio

BIRN Biomedical Informatics Research Network

GMM Modelo de Mistura de Gaussianas (do inglês, Gaussian Mixture Models)

MSE Erro quadrático médio (do inglês, Mean Squared Error)

1 INTRODUÇÃO

Nesta tese, apresentamos uma Rede Adversaria Generativa Condicional (do inglês, Conditional Generative Adversarial Network) (CGAN), um modelo generativo adversário condicional capaz de criar imagens a partir da amostragem da imagem no espaço transformado com n medidas. Neste método usamos modelos de Aprendizagem Profunda (do inglês, Deep Learning) (DL) estado da arte para compor a arquitetura do gerador e discriminador. Nosso modelo após treinado em uma pequena base de dados alcança uma boa qualidade de reconstrução comparado a estratégia Compressed Sensing (CS) e também é mais rápido em criar a nova imagem a partir da imagem amostrada. Nos experimentos, nossa abordagem fornece uma melhor reconstrução em termos das métricas objetivas: Relação sinal-ruído (do inglês, Signal-to-Noise Ratio - SNR) (SNR), Relação Sinal-Ruído de Pico (do inglês, Peak Signal-to-Noise Ratio) (PSNR), Semelhança estrutural de escala única (do inglês, Single-Scale Structural Similiarity) (SSIM).

Quanto ao nível de qualidade da imagem gerada é necessário aplicar uma métrica como Pontuação Media de Opinião (do inglês, Mean Opinion Score) (MOS), pois visivelmente em pontos mais internos a estrutura da imagem é possível notar que o modelo generativo cria pontos que podem não existir na imagem original. Por esse motivo sugerimos trabalhos futuros capazes de usar o modelo generativo para extrair informação a priori para aprimorar os resultados de reconstrução do CS e não como um modelo de reconstrução, contribuindo com maior qualidade e rapidez na reconstrução de imagens por CS.

Alcançar maior qualidade com um número reduzido de amostras permite procedimentos de exame mais rápidos, tornando o Imageamento por Ressonância Magnética (do inglês, Magnetic Resonance Imaging) (MRI) mais barato, mais rápido e mais conveniente para pacientes e clínicas, sendo nosso objetivo final.

1.1 Contextualização

O exame por MRI é uma forma de diagnóstico por imagem, com capacidade de diferenciar as características dos tecidos biológicos, sendo usada para verificar aspectos anatômicos e funcionais [4]. Por ser um exame que não apresenta riscos biológicos e por sua aplicação se estender a várias partes do corpo humano, ela se torna uma excelente escolha dentre as outras modalidades de imageamento para detecção de algumas doenças. É um dos exames mais indicados em várias aplicações devido ao nível de detalhes presentes nas imagens geradas e ao contraste superior ao das outras técnicas de imageamento médico [5].

Além disso, é considerada uma técnica biologicamente segura, pois não há o envolvimento de nenhum tipo de radiação ionizante [6]. Muitos dos problemas de saúde como câncer, fraturas, infecções no cérebro, identificação de massa ou tumores nos órgãos são facilmente detectados a partir de uma imagem de alta qualidade com inúmeros detalhes. Até mesmo os médicos ortopedistas utilizam o exame para avaliação dos tecidos moles como músculos e cartilagens. Os neurologistas também tendem a solicitar exames para análises com maior precisão para problemas como Alzheimer, atrofas e lesões nos vasos sanguíneos cerebrais e tumores no cérebro[7].

Atualmente o exame por MRI é um dos mais realizados em todo o mundo, e conforme os dados da Agência Nacional de Saúde Suplementar (ANS), no ano de 2020 a cada mil beneficiários de plano de saúde cento e cinquenta e seis foram destinados a exame de ressonância magnética [8], ficando na quarta posição dos exames médicos complementares do plano de saúde mais buscados do país [8]. O exame de ressonância magnética é considerado de grande importância para a saúde e foi um marco na história da medicina. A criação da Ressonância Magnética (do inglês, Magnetic Resonance) (MR) inicia indiretamente em 1888 quando Nikola Tesla descreve o campo magnético rotativo[9], e em 1973 foi sugerida como modalidade de imagem em medicina por Lauterbur [10] e em 1980 John Mallard liderou o desenvolvimento de um escâner de ressonância magnética e examinaram o primeiro paciente. [11]

O processo que nos interessa no exame de MR se inicia com a aquisição dos dados, onde uma amostragem da região é feita pela máquina e precisa ser codificada para enfim ser armazenada. Após o armazenamento podemos processar esses dados digitais e apresentar em uma tela ou imprimir, por exemplo. A imagem final é sensível ao movimento do paciente e por isso depende da velocidade que o exame é conduzido. A informação obtida da MR é comumente representada por amostras individuais no espaço k , que se traduz na transformada de Fourier da imagem a ser reconstruída [12]. Dependendo da trajetória usado para amostrar o espaço k , a imagem reconstruída pode ter alguns

artefatos aparentes que impossibilitam o entendimento.

1.2 Definição do Problema Científico e Proposta de Pesquisa

Apesar de todas as qualidades em MRI, existem uma série de desafios que inviabilizam em muitas situações os exames. Comparando a MR a outras técnicas de exame, ela acaba sendo uma das mais caras e demoradas. O exame completo pode levar horas, tendo em vista que após uma longa sessão ainda seja necessário a adição de um agente de contraste para permitir que certas doenças sejam diagnosticadas.

Em casos de situações de urgência, o médico pode julgar não existir tempo para o exame, ou até mesmo faltar maquinário no momento e partir para um diagnóstico mais rápido para uma ação ser tomada. Muitas das vezes o médico pode optar por uma Tomografia Computadorizada (TC), que possui o risco de exposição a radiação. O ruído sonoro também entra como um dos fatores que tornam o exame por MR não tão agradável assim, sendo necessário um protetor auricular. Devido ao tubo do aparelho fechado, alguns pacientes, mesmo com um ambiente confortável e uma boa experiência dos técnicos durante a realização do exame, não conseguem ir até o fim por serem muito claustrofóbicos. O tempo de aquisição no MRI acaba se tornando um dos principais limitantes, levando pesquisadores a buscarem soluções capazes de acelerar a aquisição mantendo a qualidade e contraste da imagem.

Para entendermos o tempo que leva a reconstrução de uma imagem de MR, é necessário entender como coletamos medidas e armazenamos esses dados para reconstrução. O primeiro passo no MRI é a aquisição dos dados. Nesta etapa é feita uma amostragem dos sinais obtidos do mundo real pela máquina de MR para podermos converter do domínio analógico para uma sequência de bits no domínio digital. Já no segundo passo é realizado a compactação dos dados adquiridos, onde as informações são codificadas usando menos bits do que a representação original. Sem essas etapas não seria possível armazenar, manipular, processar, transmitir e interpretar os dados.

O CS é a junção desses paradigmas, onde a compressão e aquisição de dados são feitas simultaneamente. Visando diminuir o tempo do processo de aquisição e compressão simultânea, para obter o mínimo de informação necessária para a reconstrução de um dado sinal, David Donoho, Emmanuel Candès, Justin Romberg e Terence Tao escreveram trabalhos fundamentais para aprofundar os resultados e divulgá-los nas comunidades de Estatística, Processamento de Sinais e Matemática. Portanto, é importante destacar que as ideias que surgiram das noções de esparsidade, incoerência e aleatoriedade e que agora permeiam o mundo dos dados.

Sabendo que o número de linhas e colunas capturadas durante a aquisição de dados impactam diretamente na qualidade da imagem, sendo assim, para a resolução espacial ser alta é necessário um maior tempo de coleta [13]. Se aumentamos o tempo de coleta, o paciente terá de ficar mais tempo numa seção de imageamento, e quando reduzimos o número de medidas adicionamos na reconstrução uma série de ruídos, que dificultam a análise humana para chegar a um diagnóstico.

Artefatos típicos que dificultam a interpretação da imagem, como falta de homogeneidade, efeitos de volume parcial e artefatos são gerados a partir de uma reconstrução. Para otimizar o tempo no uso de CS, amostragens são feitas abaixo do teorema de Nyquist, o que leva ao fenômeno dos artefatos diretamente relacionado a reconstrução com menos medidas.

A partir das contribuições com uso de técnicas de pré-filtragem de esparsificação e informações prévias que comprovadamente melhoram a eficiência e alcançam melhores reconstruções [14, 15], acreditamos que o uso da CGAN pode aprimorar a extração de informação prévia, portanto também documentamos nossa pesquisa em tais contextos visando uma continuação mais detalhada deste projeto de pesquisa, onde mais experimentação pode ser feita para validar nossa hipótese.

Nos últimos anos, a Inteligência Artificial (IA) tem atraído cada vez mais atenção devido às suas conquistas no campo da medicina. Algumas técnicas de IA podem ser usadas de várias maneiras diferentes para melhorar a análise médica. Neste trabalho aplicaremos CGAN no processo de obtenção de informações prévias aprimoradas para alimentar o algoritmo CS.

1.3 Objetivos

1.3.1 Objetivo Geral

Essa pesquisa tem como objetivo melhorar a reconstrução em qualidade e tempo de execução. Para conseguir isso, propomos a arquitetura de um modelo generativo CGAN treinado com pares de imagens amostradas radialmente no espaço k e a imagem verdadeira. A expectativa é que as imagens geradas pelo modelo possuam qualidade suficiente e assim servir de suporte para uma reconstrução a partir da técnica CS com uso de técnicas de pré-filtragem de esparsificação e informações prévias. Além disso, também apresentamos fundamentos e experimentos preliminares para uso de imagens geradas por modelo generativo para aumento de dados em modelos de classificação de tumor.

1.3.2 Objetivos Específicos

Para alcançar o objetivo geral deste trabalho, estabelecemos os seguintes objetivos específicos:

- Demonstrar experimentos com CS usando pré-filtragem e informação a priori clássica.
- Implementar o pré-processamento dos dados.
- Implementar arquitetura do gerador e discriminador.
- Testar e validar a qualidade do modelo generativo.
- Comparar resultados, mostrar limitações da reconstrução por modelos generativos e propor possíveis melhorias em reconstrução a partir de um modelo generativo.

2 FUNDAMENTAÇÃO TEÓRICA E ESTADO DA ARTE

2.1 Imageamento por Ressonância Magnética

A ressonância magnética é uma técnica de imagem sofisticada que evoluiu como modalidade clínica nos últimos 30 anos. Ao longo do caminho, muitos cientistas de diversas áreas fizeram contribuições notáveis que trouxeram o campo ao seu estado atual. Ela foi proposta pela primeira vez em 1973 por Lauterbur, que apresentou uma imagem transversal de prótons de água em dois tubos de ensaio por Ressonância Magnética Nuclear (do inglês, Nuclear Magnetic Resonance) (NMR)[10]. Com o avanço da tecnologia também foi possível fazer imagens do corpo humano, se inspirando parcialmente na ideia de Damadian de detecção entre tumores malignos e tecidos normais por NMR[16]. A ressonância magnética usa campos magnéticos e pulsos de radiofrequência (radiações não ionizantes), sendo assim uma técnica não invasiva e inofensiva ao paciente.

Um dos avanços em MR é a evolução dos métodos de coleta de dados. A imagem convencional de MR bidimensional é realizada por medições que correspondem a pontos ao longo de um caminho prescrito através do espaço de Fourier multidimensional representando um corpo imageado. Este espaço de Fourier é conhecido como espaço k na comunidade médica. A representação completa no espaço de Fourier de uma região é capturado por uma sequência de amostras que ladrilham o espaço até uma frequência máxima especificada [17].

Portanto, é necessário acelerar o tempo de imagem para adquirir imagens de alta qualidade sobre a região tridimensional do objeto em um tempo de exame limitado. A aquisição rápida de imagens é ainda mais importante no caso de imagens de órgãos em movimento.

2.2 Amostragem

O tempo necessário para adquirir todas as medidas durante um exame de MR em um indivíduo seria extremamente alto. A maneira como as máquinas podem fazer MRI mais rápido é executando *subamostragem*, onde uma menor quantidade de medidas é usada para reconstrução da imagem de MR.

A subamostragem é realizada dando à máquina um caminho prescrito conhecido no qual ela extrairá medições da representação multidimensional do espaço k . Isso permite que as máquinas colem apenas uma fração das medições de dados necessárias para a reconstrução da imagem, acelerando assim o processo de aquisição de dados sem perda crítica de qualidade.

Existem alguns padrões de subamostragem a serem usados e cada um tem seus benefícios dependendo de vários parâmetros, como a extração da região do sujeito, o algoritmo usado para a reconstrução, o tempo de aquisição.

Em MRI, as técnicas de subamostragem são frequentemente usadas para adquirir elementos η onde $\eta < 2\eta < N$, violando o critério de Nyquist.

2.2.1 Espaço k

Espaço k é a informação de frequência espacial em duas ou três dimensões de um objeto. Basicamente ele representa a matriz onde os dados de MRI serão armazenados previamente em uma transformada de Fourier para obtenção da imagem desejada. Cada coluna do espaço k contém os dados obtidos durante a etapa de codificação de frequência. E Cada linha é preenchida repetindo as etapas de codificação de fase.

Os elementos centrais são os de menor frequência, sendo responsáveis por informações como formato da imagem e intensidade dos elementos. Já os elementos mais dispersos são responsáveis pelas bordas e detalhamentos da imagem [18]. O preenchimento linha a linha do espaço k requer uma frequência mínima de amostragem para que não gere o fenômeno conhecido como aliasing ou dobramento, alterando o valor de diversos componentes do sinal.

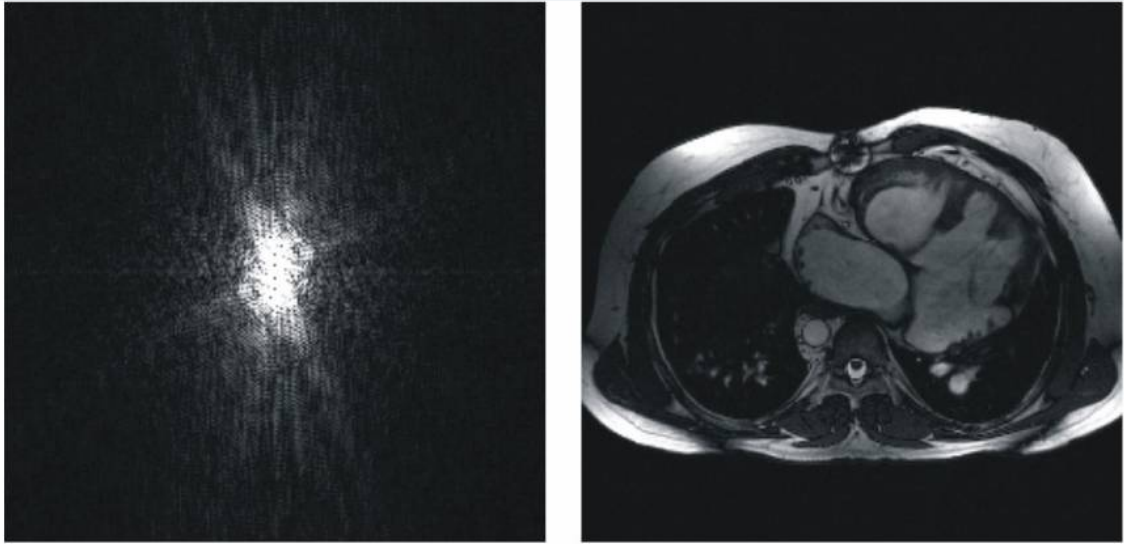


Figura 2.1. Espaço k (esquerda) e imagem original (direita)
[19]

2.2.2 Trajetórias de amostragem

Em exames de ressonância magnética a aquisição de dados é feita a partir de uma trajetória para o preenchimento do espaço k . Existem diversas formas de se realizar a aquisição de dados no espaço k , sendo a mais comum adquirir uma linha do espaço k a cada Tempo de Repetição (TR). A subamostragem deve ser feita de forma rápida, para evitar artefatos de movimento. Dessa forma, para diminuir o tempo de aquisição é necessário reduzir o número de linhas a serem preenchidas no espaço k .

O tipo de trajetória pode ser escolhido conforme o tipo de exame que será executado, o algoritmo de reconstrução usado e por fim seguindo a tolerância a artefatos na imagem. As trajetórias em espiral são comumente usadas em aquisições dinâmicas, ou seja, exame em que se deseja observar o movimento de órgãos. As aquisições espirais ainda são mais eficientes ao nível de hardware e podem ser usadas para imageamento em tempo real [20]. As aquisições radiais são menos susceptíveis a artefatos de movimento e evitam contraste e descontinuidades e preservam a suavidade da imagem [21].

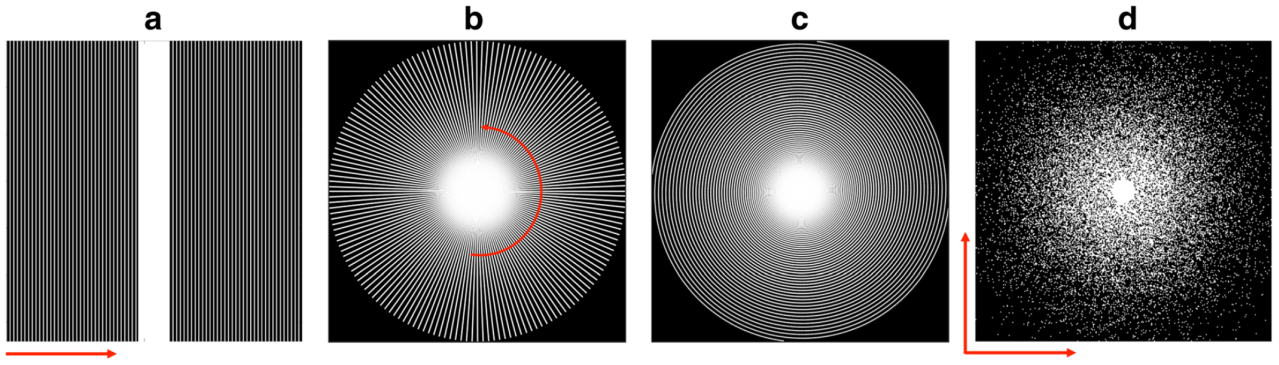


Figura 2.2. Trajetórias do espaço k . (a) Cartesiana, (b) radial, (c) espiral, (d) amostras isoladas no espaço k , conforme a realização de um processo aleatório [1].

2.3 Compressed Sensing

2.3.1 Introdução

A velocidade na qual os dados podem ser coletados em MRI é limitada e suscetível a artefatos relacionados ao movimento, que podem até exigir sedação ou anestesia. Os métodos para acelerar a MRI geralmente envolvem a redução do número de segmentos no espaço k , ou seja, uma sub amostragem. Porém, quando o espaço k é sub amostrado, o critério de Nyquist é violado e as reconstruções exibem artefatos de aliasing que podem enviesar a análise médica.

A teoria do CS pode ser resumida a reconstrução de sinais a partir de menos amostras, violando o critério de Nyquist. Nas técnicas tradicionais de processamento de sinal, o dado é amostrado uniformemente na taxa de Nyquist, gerando n amostras. Essas amostras são então comprimidas em m amostras, por fim descartando $n - m$ amostras. Seguindo a taxa de Nyquist, o sinal pode ser perfeitamente reconstruído se a taxa do sinal de amostragem for pelo menos duas vezes a frequência máxima presente no sinal[22].

Ao contrário do paradigma Nyquist tradicional, o CS consegue encontrar soluções esparsas para sistemas lineares subdeterminados, reconstruindo o sinal a partir de menos amostras do que é possível, usando a taxa de amostragem de Nyquist e apresentando excelentes resultados. O CS foi uma teoria revolucionária, e ao ser apresentada em seus primeiros artigos no ano de 2004, foi inicialmente rejeitada pelos revisores por parecer algo impossível de substanciar segundo Candés [23].

CS parte do princípio de que dado um x , ou seja, uma imagem ou sinal digital com uma representação esparsa em uma base ortonormal (por exemplo, wavelet, Transformação

rápida de Fourier (do inglês, Fast Fourier Transform) (FFT), Transformada Cosseno Discreta (do inglês, Discrete Cosine Transform) (DCT), etc.), o N de coeficientes mais importantes nessa expansão permitem a reconstrução com o erro ℓ_2 $O(N^{1/2-1/p})$ [24].

Um sinal x é conhecido por ter uma representação esparsa se houver uma matriz determinística e invertível $T_{N \times N}$ de modo que o vetor transformado \hat{x} seja composto pela maioria de seus N componentes iguais a zero $\hat{x} = Tx$ [14]. No MRI as l medidas extraídas do escâner de MR correspondem a alguns dos coeficientes em $\hat{x} = Tx$.

Dado um sinal x complexo de tempo discreto com dimensão N , compressível por uma transformação linear com uma representação esparsa, diz-se que este sinal x pode ser totalmente reconstruído a partir de um vetor l dimensional de medições b definidas como $b = Mx$, onde M tem forma $l \times N$. Assim, a mesma equação também pode ser representada como $b = MT^{-1}\hat{x}$ e então como $MT^{-1}\hat{x} - b = 0$.

Para que a teoria CS realize com sucesso a reconstrução do sinal x , este sinal precisa ser esparsa e também deve corresponder com a condição de incoerência, onde sinais esparsos no domínio transformado devem ser bem distribuídos no domínio das medidas, ou seja, as linhas do vetor de medidas b e o domínio definido por T devem ser incoerentes [25].

Como o sistema no cenário MRI é subdeterminado, $l \ll N$, o sistema admite infinitas soluções, portanto informações adicionais sobre a natureza do sinal são necessárias. Na teoria CS, isso é feito usando um algoritmo de otimização com restrições que exploram a esparsidade no domínio definido por T .

O Propriedade de Isometria Restrita (do inglês, Restricted Isometry Property) (RIP), introduzido por Candès e Tao [26] em compressive sensing [27] confirma que a solução para a reconstrução tem estabilidade garantida se a matriz de medição b satisfizer a propriedade RIP, com relação à transformação esparsificante [12]. Para um sinal x com elementos η diferentes de zero na representação esparsa, diz-se que a matriz $A = MT^{-1}$ satisfaz o RIP somente se qualquer vetor v com dimensão N e um máximo de 3η elementos diferentes de zero,

$$1 - \epsilon \leq \frac{\|MT^{-1}v\|_2}{\|v\|_2} \leq 1 + \epsilon, \quad (2.1)$$

onde ϵ é uma constante de tolerância que deve utilizar valores menores para garantir maiores estabilidades no procedimento de reconstrução.

2.4 Redes Neurais Artificiais

Redes Neurais Artificiais são técnicas computacionais que simulam o mecanismo de aprendizado e apresentam um modelo matemático inspirado na estrutura neural de organismos inteligentes, que adquirem conhecimento através da experiência[28].

O sistema nervoso é formado por um conjunto extremamente complexo de células, os neurônios. Eles têm um papel essencial na determinação do funcionamento e comportamento do corpo humano e do raciocínio. Os neurônios são formados pelos dendritos, os quais são um conjunto de terminais de entrada, pelo corpo central, e pelos axônios que são longos terminais de saída como pode ser visto na figura 2.3.

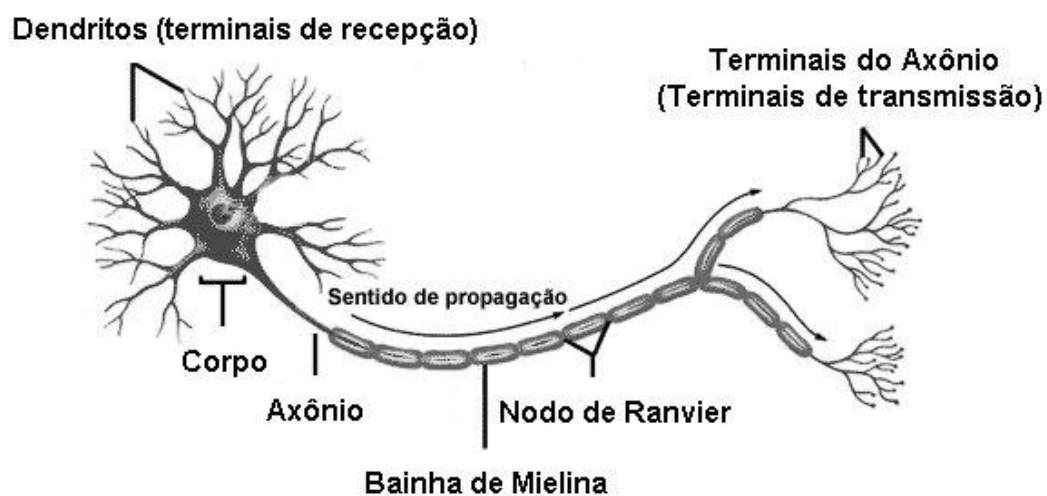


Figura 2.3. Componentes de um neurônio biológico [29].

Em 1943 a operação de uma unidade de processamento foi proposta por McCulloch e Pitts no artigo *A logical calculus of the ideas immanent in nervous activity* [30], combinando neurofisiologia e lógica matemática, onde eles mostraram como excitação, inibição e limiar podem ser usados para construir uma grande variedade de “neurônios”. Foi proposto um modelo chamado modelo de McCulloch-Pitts onde dendritos do neurônio biológico k são representados pelas m entradas x_i e o axônio é representado pela saída y_k . As sinapses nervosas são representadas por um peso w_{ki} associado a cada entrada, podendo este ser excitatório (valor positivo) ou inibitório (valor negativo). O corpo celular é responsável por processar os sinais de entrada e transformá-lo em um sinal de saída.

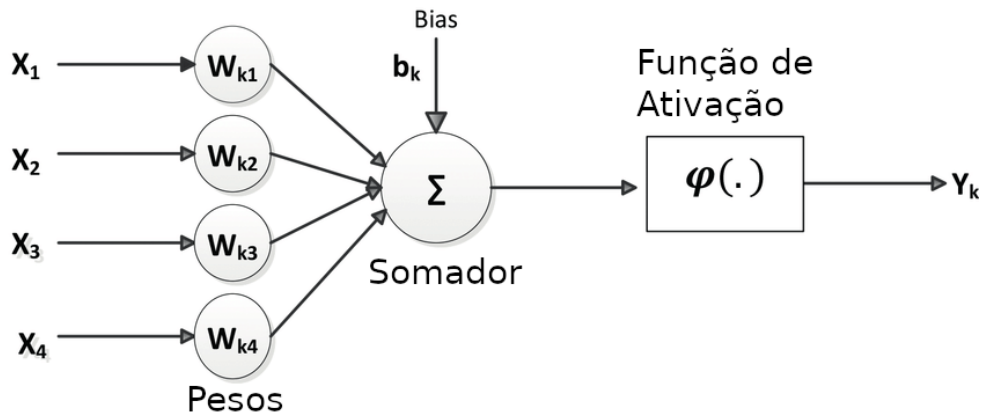


Figura 2.4. Elementos Básicos de um Neurônio Artificial de McCulloch e Pitts [30].

Como mostra na figura 2.4 a saída do neurônio é o resultado da aplicação da função de ativação, responsável pelo limiar de disparo do neurônio, sobre a soma dos sinais de entradas ponderados pelos seus respectivos pesos sinápticos somados à polarização externa ou viés b_k . Esse processo pode ser descrito pelas seguintes equações:

$$u_k = \sum_{i=1}^m x_i w_{ki} \quad (2.2)$$

$$y_k = f(u_k + b_k) \quad (2.3)$$

onde x_1, x_2, \dots, x_m são os sinais de entradas; $w_{k1}, w_{k2}, \dots, w_{km}$ são os pesos sinápticos do neurônio k ; u_k é o resultado da função soma; b_k é o viés que tem o efeito de aumentar ou diminuir a polarização na função de ativação; $\varphi(.)$ é a função de ativação do neurônio e y_k é a saída do neurônio k .

2.4.1 Aprendizado Supervisionado e Não Supervisionado

No processo de aprendizagem das redes neurais há dois tipos diferentes de abordagem na modelagem. Eles diferem na forma como os modelos são treinados e na condição inicial dos dados de treinamento. O aprendizado de máquina supervisionado requer dados rotulados de entrada e saída durante a fase de treinamento. Uma vez que o modelo tenha aprendido a relação entre os dados de entrada e saída, ele pode ser usado para classificar conjuntos de dados novos e não vistos e prever resultados. A razão pela qual este método é chamado de supervisionado é porque pelo menos parte dessa abordagem requer uma

supervisão humana.

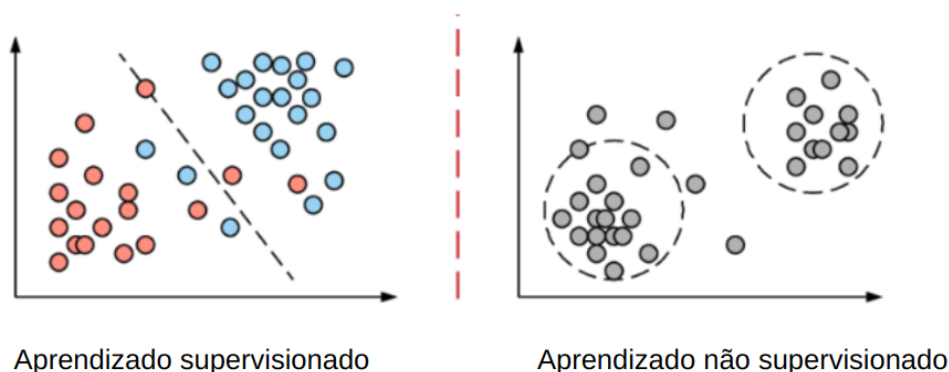


Figura 2.5. Exemplo visual dos métodos de aprendizagem supervisionada e não supervisionada. Figura obtida de: [2]

O aprendizado de máquina não supervisionado é o treinamento de modelos em dados de treinamento brutos e não rotulados. Geralmente é usado para identificar padrões e tendências em conjuntos de dados brutos ou para agrupar dados semelhantes em um número específico de grupos. Também costuma ser uma abordagem usada na fase exploratória inicial para entender melhor os conjuntos de dados. Devido à menor supervisão humana, esse método exige um bom pré-processamento e maior conhecimento a respeito das técnicas usadas para agrupar os dados [31]. Como mostra na figura 2.5 o método de aprendizado supervisionado possui as classes de dados representadas por cores e pode nesse caso ser parcialmente separadas por uma linha. Já no caso do método de aprendizagem não supervisionada, os dados não possuem uma classificação, então podemos agrupar por distância ou alguma característica específica.

2.4.2 Gradiente Descendente

Gradiente descendente é um dos algoritmos mais usados para otimizar redes neurais e outros diversos algoritmos de aprendizagem de máquina [32]. O método Gradiente descendente se baseia em encontrar iterativamente os parâmetros que minimizam uma função. Devido à grande massa de dados a serem processadas existe uma maior variância nas atualizações, sendo assim usada a variação conhecida por Gradiente Descendente Estocástico. Este método tem sua convergência de forma rápida e por ter uma implementação vetorizada é possível paralelizar os cálculos.

Neste caso, o Gradiente descendente utiliza apenas uma observação por cada iteração e atualiza os parâmetros de treinamento, isto é, dado um ponto inicial escolhido pelo chute β_1 , e uma taxa de aprendizagem α responsável pela distância que será percorrida

por cada iteração, temos então o β_1 como sendo,

$$\beta_1 = \beta_0 - \alpha \nabla f(\beta_0; x_i, y_i) \quad (2.4)$$

$\nabla f(\beta_0; x_i, y_i)$ é o gradiente da função de custo no ponto β_0 para cada exemplo de treinamento x_i e uma label y_i . Ou seja, escrevemos a fórmula geral,

$$\beta_{n+1} = \beta_n - \alpha \nabla f(\beta_n; x_i, y_i). \quad (2.5)$$

2.4.3 Retro propagação

A retro propagação do inglês Backpropagation ou propagação reversa de erros é um método que monitora o aprendizado utilizando a média do erro quadrático e gradiente descendente para ser feita uma modificação nos pesos da rede neural[33].

A retro propagação calcula as derivadas parciais de erro de uma função de aproximação $F(w, x)$ calculado pela Redes Neural Artificial em relação ao peso e vetor de entrada para cada exemplo de treinamento.

Para avaliar as derivadas da função $F(w, x)$ em relação a todos os elementos no vetor de peso w para um vetor de entrada $x = [x_1, x_2, \dots, x_{m_0}]^T$ e para um Mutli-Layer Perceptron com camada $l = 2$, temos a seguinte equação onde φ é a função de ativação, w é o vetor de peso ordenado e x é o vetor de entrada [34].

$$F(\mathbf{w}, \mathbf{x}) = \sum_{j=0}^{m_1} w_{oj} \varphi \left(\sum_{i=0}^{m_0} w_{ji} x_i \right) \quad (2.6)$$

2.5 ConvNets

A Rede Neural Convolutacional (do inglês, Convolutional Neural Network) (CNN) é um tipo de rede neural, introduzida pela primeira vez na década de 1980. Eles foram originalmente desenvolvidos para tarefas de reconhecimento de imagens, sendo inspirados na estrutura do sistema visual do cérebro, composto por camadas hierárquicas de neurônios que extraem características cada vez mais complexas dos estímulos visuais.

O primeiro modelo convolutacional foi desenvolvido por Kunihiko Fukushima em 1980 e chamado de Neocognitron [35]. Consistia em várias camadas de unidades convolucionais

e de agrupamento, que foram treinadas usando aprendizado não supervisionado para reconhecer dígitos manuscritos.

Em 1989, Yann LeCun e seus colegas da Bell Labs estenderam o modelo Neocognitron e introduziram a arquitetura LeNet [36], que foi uma das primeiras ConvNets bem sucedidas para reconhecer dígitos manuscritos. Este modelo foi treinado no conjunto de dados MNIST e alcançou resultados impressionantes para a época, com uma precisão de reconhecimento de mais de 99%[36].

Desde então, as ConvNets foram aplicadas a uma ampla gama de tarefas, incluindo reconhecimento de objetos, reconhecimento de fala, processamento de linguagem natural e muitas outras. Avanços recentes em computação paralela permitiram treinar modelos ConvNets muito grandes e complexos como a AlexNet de 2012 já possuem oito camadas e milhões de parâmetros, levando a melhorias significativas no desempenho em muitas tarefas. Hoje, as ConvNets são uma parte fundamental do campo de aprendizado profundo, sendo usadas em muitos modelos de aprendizado de máquina.

2.6 Modelos Generativos

Os modelos generativos são do método de aprendizagem não supervisionada e surgiram partindo do pré-suposto onde a partir de um conjunto finito de amostras de uma distribuição subjacente, é possível aproximar essa distribuição de dados com acesso ao conjunto de dados de destino. Esses modelos são usados para gerar novos exemplos de dados. Eles foram desenvolvidos a partir da década de 1990, e têm sido aplicados em uma ampla variedade de tarefas, incluindo a geração de imagens, áudio e texto.

Os primeiros modelos generativos foram baseados em Modelo de Mistura de Gaussianas (do inglês, Gaussian Mixture Models) (GMM), que assumiam que os dados seguiam uma distribuição gaussiana e usavam o algoritmo Expectation Maximization (EM) para estimar os parâmetros desta distribuição [37]. Estes modelos foram usados principalmente para análise de dados e clusterização. Também foram desenvolvidos modelos generativos baseados em redes neurais, como o modelo de máquina de Boltzmann restrita [38] que ainda tínhamos a limitação das dos neurônios serem bidirecionais e simétricos.

Em 2014, foi introduzido Rede Generativa Adversaria (GAN)[39], que utiliza um par de redes neurais, um gerador e um discriminador, para gerar novos exemplos de dados. O gerador tenta produzir exemplos que são indistinguíveis dos dados reais, enquanto o discriminador tenta classificar corretamente os exemplos como reais ou sintéticos. O treinamento é feito por um jogo minimax, onde o gerador busca minimizar a função de

perda e o discriminador busca maximizá-la.

Desde então, foram desenvolvidos vários tipos de modelos Rede Adversaria Generativa (do inglês, Generative Adversarial Network) (GAN), como o Rede Adversarial Gerativa Convolutacional Profunda (do inglês, Deep Convolutional Generative Adversarial Network) (DCGAN) e o CGAN, que utilizam arquiteturas de redes neurais mais avançadas e novas técnicas de treinamento para melhorar a desempenho dos modelos. Estes modelos têm sido aplicados com sucesso em uma ampla variedade de tarefas, sendo considerados um dos principais avanços recentes na área.

2.6.1 Autoencoder

Autoencoder é uma arquitetura de aprendizado profundo não supervisionado que manipula a proximidade dos dados no espaço latente e atua como uma função identidade. Sua estrutura simétrica é composta por dois blocos principais: Codificador f_ϕ que comprime a entrada em uma representação compacta e de baixa dimensão com o conteúdo informativo e relevante dos dados, desconsiderando o ruído [40]. O Decodificador f'_θ aprende as representações dos dados, e por um processo de decodificação reconstrói a entrada. Sendo ϕ e θ parâmetros das funções do codificador e do decodificador, respectivamente.

O autoencoder visa minimizar a diferença entre a entrada x e uma entrada reconstruída x' , e o erro quadrático pode ser representado por,

$$L(x, x') = \|x - x'\|^2 = \|x - f'_\theta(f_\phi(x))\|^2, \quad (2.7)$$

onde o espaço latente representado por $z = f_\phi(x)$ possui um espaço de dimensão muito inferior do que a entrada original, e contém informações suficientes para reconstruir a entrada original o mais próximo possível.

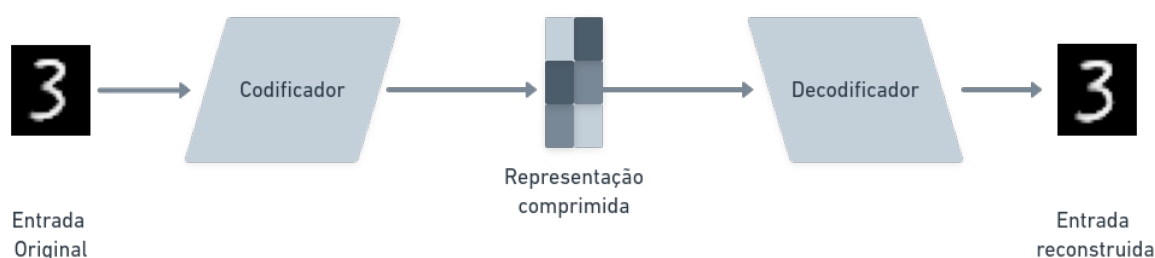


Figura 2.6. A imagem de entrada é codificada em uma representação compactada e, em seguida, decodificada gerando a reconstrução [3].

2.6.2 Redes Adversárias Generativas

A Rede Adversaria Generativa foi proposta em 2014 por Ian Goodfellow como uma técnica para sintetizar imagens baseadas em redes neurais artificiais. As Redes Adversárias Generativas fornecem uma maneira de aprender representações sem dados de treinamento anotados. Isso é possível por um processo competitivo envolvendo duas redes neurais. Uma GAN pode ser usada em vários tipos de aplicações, incluindo edição, super-resolução e classificação de imagens.

A GAN possui arquitetura similar a de um codificador-decodificador onde o *Gerador* $G(z)$ cria amostras a partir de um ruído z enquanto o *Discriminador* $D(x|G(z))$ é treinado para distinguir entre amostras reais e as geradas. Os dois modelos são treinados juntos e por uma função de perda, e competem entre si, assim melhorando a cada ciclo a qualidade dos dados gerados[39]. O *Gerador* $G(z)$ não acessa diretamente as imagens reais, então aprende por meio de interação com o *Discriminador* $D(x|G(z))$. Por outro lado, o *Discriminador* $D(x|G(z))$ possui acesso aos dados reais e as amostras sintéticas geradas. Ambos *Gerador* e *Discriminador* podem ser uma função de mapeamento não linear, como uma multicamada perceptron.

O papel do gerador é aproximar uma distribuição $p_g(x)$ que representa uma função de densidade de probabilidade com uma variável aleatória x , por exemplo, valores de pixel que se assemelha aos dados reais $p_{data}(x)$. Quando tal distribuição pode ser aproximada, é possível extrair amostras que parecem reais. Para definir matematicamente o objetivo de uma GAN, temos $G(z; \theta_g)$ sendo uma rede neural usada como gerador, onde z é o ruído aleatório obtido de uma distribuição anterior $z \sim p_z$ e θ_g como os parâmetros da rede. O objetivo do gerador é realizar o mapeamento de um determinado vetor de ruído para os dados reais ($G : z \rightarrow x$). Isso equivale a treinar uma rede que recebe como entrada um vetor de ruído aleatório para simular, por exemplo, imagens de MR.

A rede discriminadora $D(x; \theta_d)$ é parametrizada por θ_d e objetiva distinguir dados sintéticos dos dados reais $x \sim p_{data}$, enquanto o gerador tenta maximizar a probabilidade de detectar amostras falsas como reais. Após cada iteração de treinamento, se a qualidade dos dados falsos for insuficiente, o que significa que o discriminador pode facilmente detectá-los como falsos, o gerador recebe um retorno negativo ajudando a melhorar o processo de geração. Este processo é repetido até que as imagens falsas do gerador não possam ser facilmente diferenciadas das reais pelo discriminador. Em essência, essas duas redes competem de forma adversária por meio de um jogo minimax para dois jogadores, conforme mostrado na equação,

$$\min_G \max_D V(G, D) = \min_G \max_D \mathbb{E}_{x \sim p_{\text{data}}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [1 - \log D(G(z))]. \quad (2.8)$$

2.6.3 Rede Adversária Generativa Condicional

A CGAN é uma configuração de uma GAN com objetivo de aprender a mapear e aproximar uma distribuição x e uma imagem alvo z para um determinado y [41]. O objetivo de uma CGAN pode ser expresso como,

$$CGAN(G, D) = \mathbb{E}_{x,y} [\log D(x, y)] + \mathbb{E}_{x,z} [1 - \log D(x, G(x, z))], \quad (2.9)$$

a partir disso temos então nossa função objetivo G^* , onde G tenta minimizar o erro e D o maximizar,

$$G^* = \arg \min_G \max_D CGAN(G, D). \quad (2.10)$$

A partir dos excelentes resultados apresentados no principal artigo para tradução Image-to-Image[41], usaremos o erro L1 calculado a partir da média absoluta do erro entre as classificações das imagens reais e as imagens geradas que pode ser expressa como,

$$L1(G) = \mathbb{E}_{x,y,z} [\|y - G(x, z)\|_1]. \quad (2.11)$$

3 MATERIAIS E MÉTODOS

As seções a seguir são uma visão geral detalhada dos experimentos realizados. Os tópicos abordados são automação de treinamento, geração e pré-processamento de dados. Em seguida, uma olhada passo a passo nas arquiteturas de modelo usadas, com suas respectivas funções de perda e o algoritmo para treinar esses modelos.

3.1 Informações do equipamento

Todos os experimentos mencionados foram conduzidos usando um computador Lenovo ThinkPad T15g Gen 2 com as seguintes características:

- CPU: 11th Gen Intel(R) Core(TM) i7-11800H @ 2.30GHz
- GPU: NVIDIA GeForce RTX 3070 Laptop 8Gb
- Memória: 64GB at 3200 MHz
- OS: Ubuntu 22.04.1 LTS x86_64
- Kernel: 5.15.0-48-generic
- Python: 3.8.13
- GCC: 11.2.0

3.2 Aquisição de medidas sub amostradas no espaço k

O conjunto de dados foi criado a partir de imagens sub amostradas no domínio da frequência de Fourier. Seleccionamos pontos usando trajetórias de amostragem radial em dados do espaço k com um número n de linhas. Foram usadas 10, 30 e 100 linhas para gerar diferentes espaçamentos em cada trajetória. Na trajetória radial com 10 linhas

obtemos 3% do total de pontos, com 30 linhas 11% e 100 linhas 35%. A região em branco na figura 3.5 mostra de onde serão adquiridas as medidas no domínio transformado.

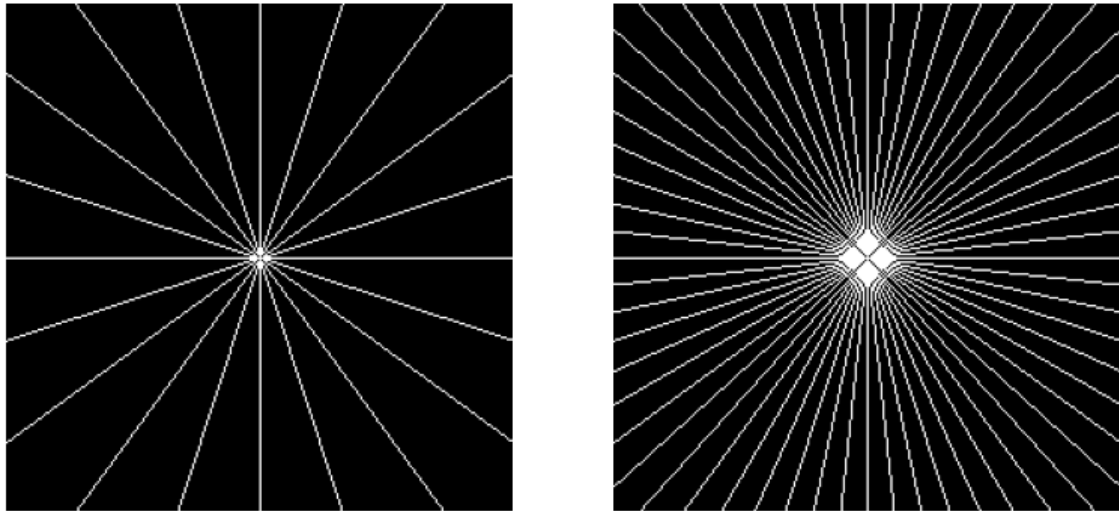


Figura 3.1. Exemplo 2D dos dados amostrados no domínio da frequência. Trajetória com 10 linhas radiais (Esquerda), Trajetória com 30 linhas radiais (Direita)

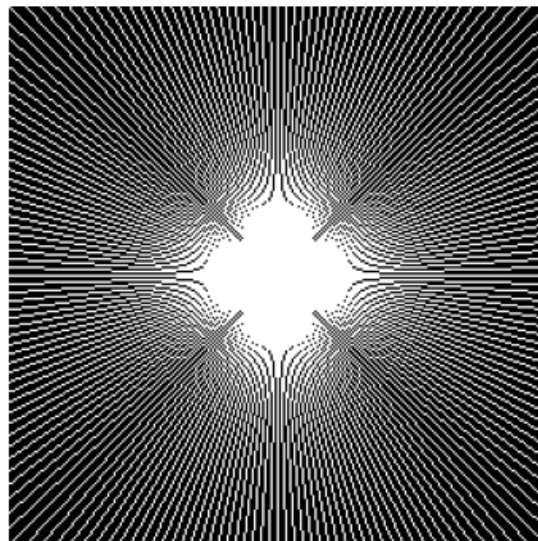


Figura 3.2. Trajetória com 100 linhas radiais

3.3 Base de dados

3.3.1 Rede de Pesquisa em Informática Biomédica

O conjunto de dados usados para treinamento e teste do modelo generativo adversário condicional foram adquiridos a partir do Biomedical Informatics Research Network (BIRN). Essa comunidade virtual, focada em pesquisa biomédica, concentrou recursos e distribui

para contribuir no avanço de diagnóstico e tratamento de doenças. A base de dados é composta por imagens de ressonância magnética funcional de pacientes saudáveis, onde cada exame possui 26 imagens. Cada imagem é como uma fatia da cabeça, ou seja, são escaneamentos estruturais de cortes axiais do paciente de cima para baixo.

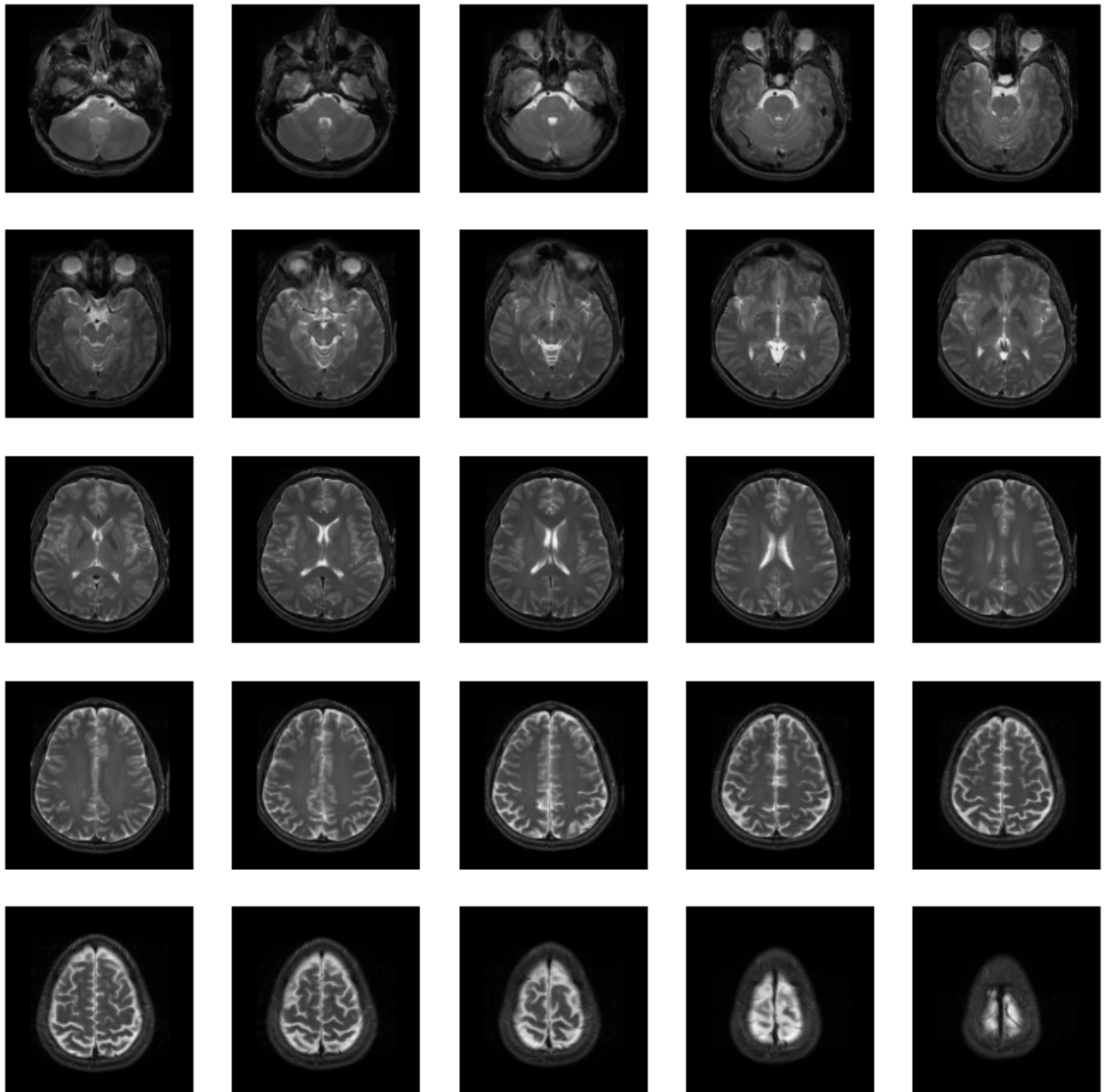


Figura 3.3. Diferentes cortes de um mesmo exame de Ressonância Magnética

3.3.2 Preparação da base de dados

A base de dados completa possui 158 escaneamentos, sendo cada fatia uma imagem 2D de 256x256 e totalizando 4.266 imagens. Selecionamos a décima sexta posição do corte

para padronizar as imagens de treinamento e teste, o exemplo da figura 3.4. Separamos 80% da base para treinamento e 20% para teste, onde cada amostra é composta por um par de imagens. A primeira imagem $x_{sampled}$ do par vem da aquisição de medidas sub amostradas no espaço k da imagem x . A segunda imagem do par é a imagem x , ou seja, é a imagem alvo que gostaríamos de ter numa reconstrução.

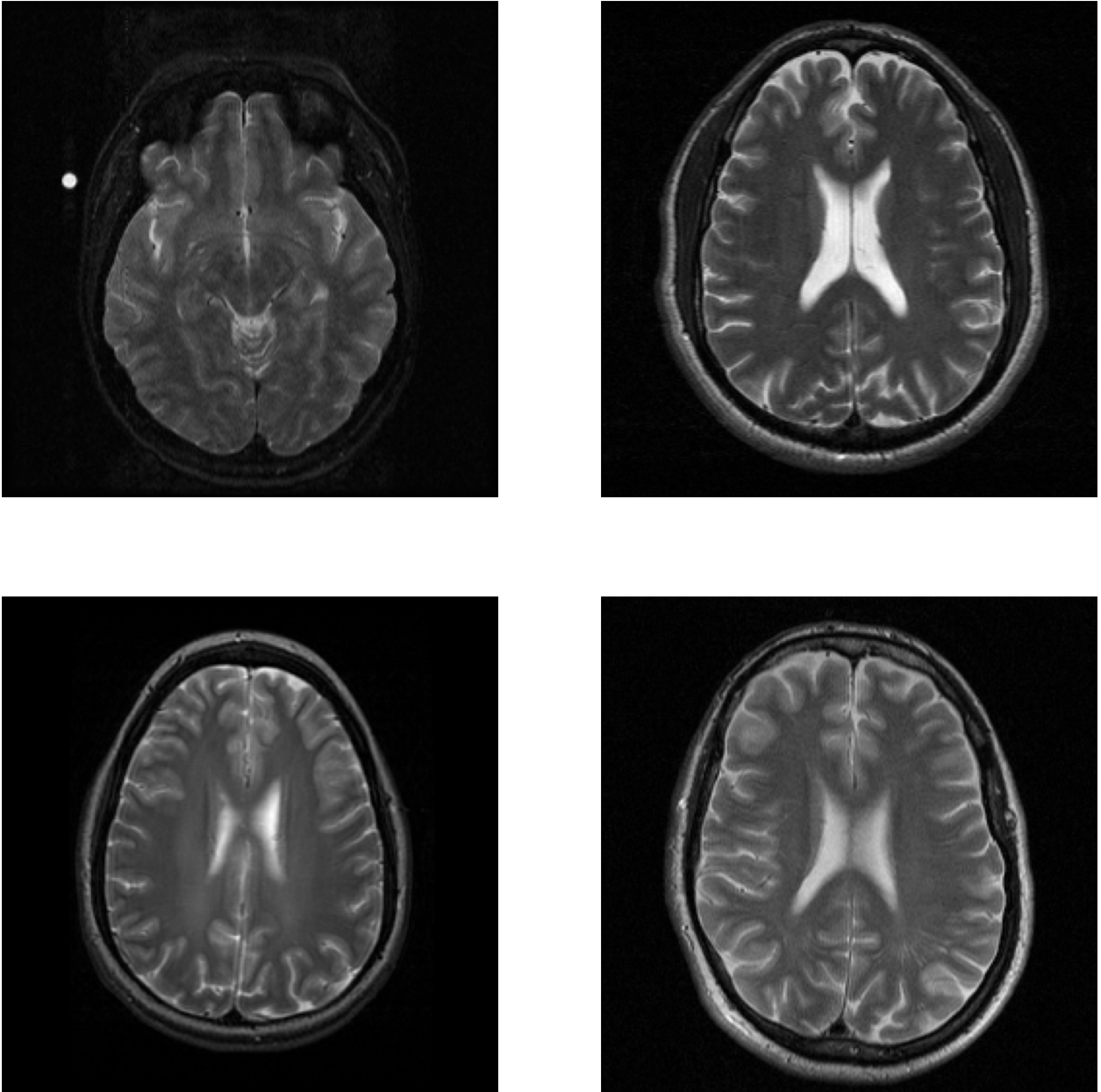


Figura 3.4. Exemplos de imagens retiradas da décima sexta posição do grupo de cortes de um exame de MR

Por fim, reconstruímos a imagem a partir do CS com um número n de medidas, a partir de um filtro radial. O filtro radial recebe como parâmetro um número de ângulos para realizar a amostragem da imagem. As imagens de treinamento e teste são armazenadas

em pares contendo a imagem original e a imagem reconstruída.

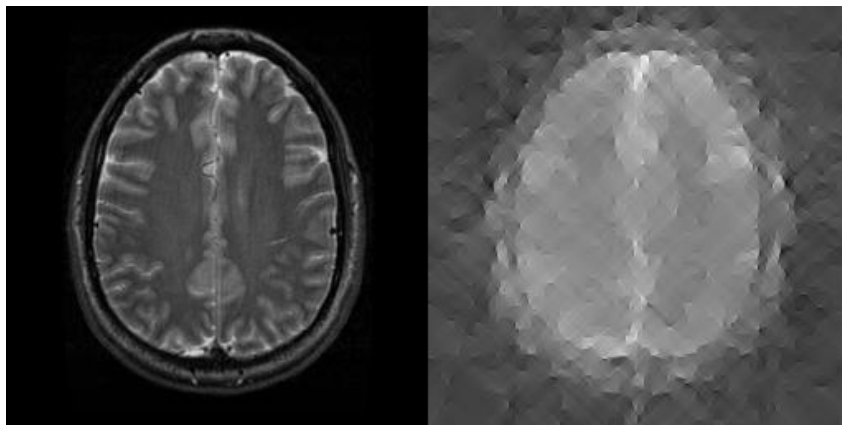


Figura 3.5. Trajetória com 10 linhas radiais.

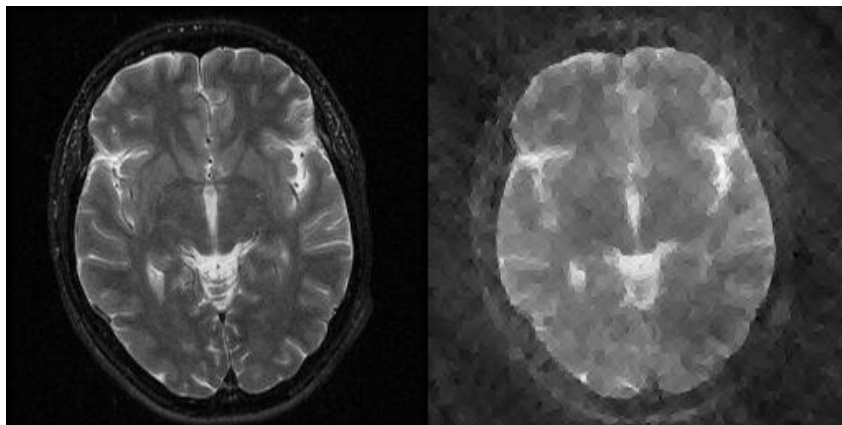


Figura 3.6. Trajetória com 30 linhas radiais.

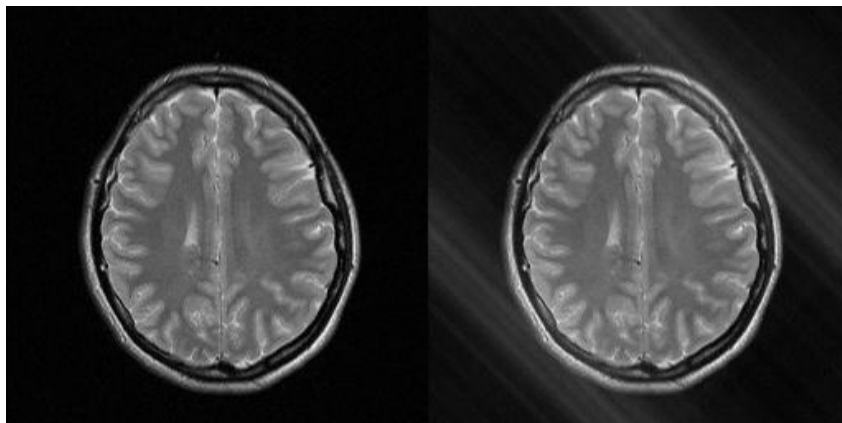


Figura 3.7. Trajetória com 100 linhas radiais.

3.4 Arquitetura CGAN

A arquitetura do modelo generativo condicional usada será baseada na Pix2Pix, que possui uma grande semelhança com uma GAN tradicional. Elas se diferem na em sua

entrada de dados para o gerador e discriminador.

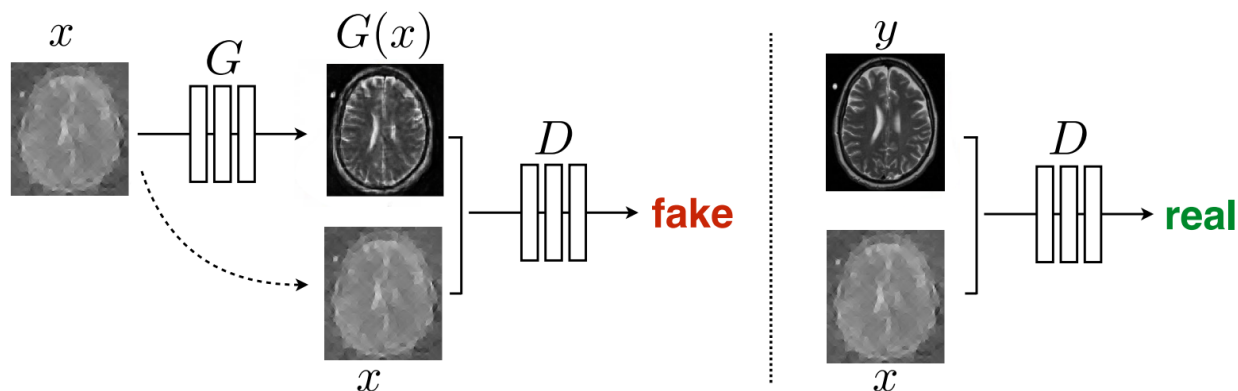


Figura 3.8. Treinamento da CGAN para mapear Imagem amostrada \rightarrow Imagem real. O discriminador representado por D , aprende a classificar entre falso (sintetizado pelo gerador), x e $G(x)$ e tuplas reais x e y . O gerador representado por G , aprende a enganar o discriminador. Ao contrário do modelo generativo incondicional, tanto o gerador quanto o discriminador observam o mapa de borda de entrada.

3.4.1 Gerador

Para gerar as imagens usamos a rede U-net, sendo uma arquitetura de rede neural convolucional encoder-decoder. Usada inicialmente para segmentação de imagens médicas onde seu grande diferencial é performar usando um pequeno conjunto de dados. A arquitetura da rede U-Net possui dois passos, o primeiro chamado de contração ou decodificador que é onde encodamos a imagem sendo capturado o contexto na imagem. Consiste na aplicação repetida de duas convoluções 3×3 sem preenchimento, seguidas por uma Unidade Linear Retificada (do inglês, Rectified Linear Units) (ReLU) e uma operação de max pooling 2×2 com passo 2 para downsampling. A diferença de uma CNN tradicional é que a cada etapa de downsampling é dobrado o número de canais. O segundo passo é o expansivo ou decodificador que consiste em um up-sampling do mapa de características do passo anterior, seguido por uma convolução 2×2 (up-convolution ou convolução transposta) que reduz pela metade o número de canais de características. Durante a expansão são feitas concatenações com o mapa de características recortado correspondente do passo contração e duas convoluções 3×3 , cada uma seguida por um ReLU[42].

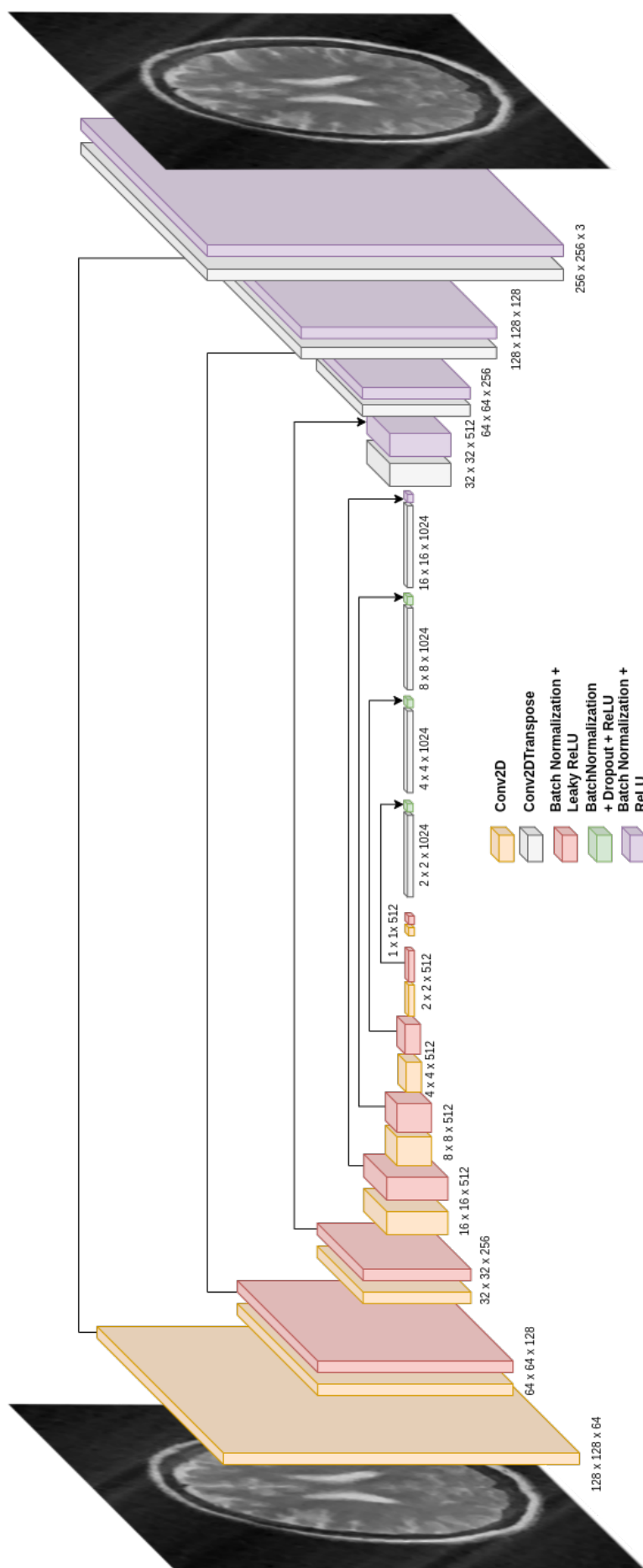


Figura 3.9. A U-NET recebe como entrada uma imagem, que na CGAN é a imagem sub amostrada e tenta aproximar a saída da imagem original. Cada camada é representada pelas cores com uma ou mais operações a fim de simplificar a visualização.

3.4.2 Discriminador

PatchGAN [41] é um tipo de discriminador que penaliza apenas a estrutura na escala de patches locais de imagem. O discriminador PatchGAN tenta classificar se cada patch em uma imagem é real ou falso. Esse discriminador é executado convolucionalmente em toda a imagem, calculando a média de todas as respostas para fornecer a saída final. Tal discriminador modela efetivamente a imagem como um campo aleatório de Markov, assumindo independência entre píxeis separados por mais do que o diâmetro do patch. Ele pode ser entendido como um tipo de perda de textura/estilo.

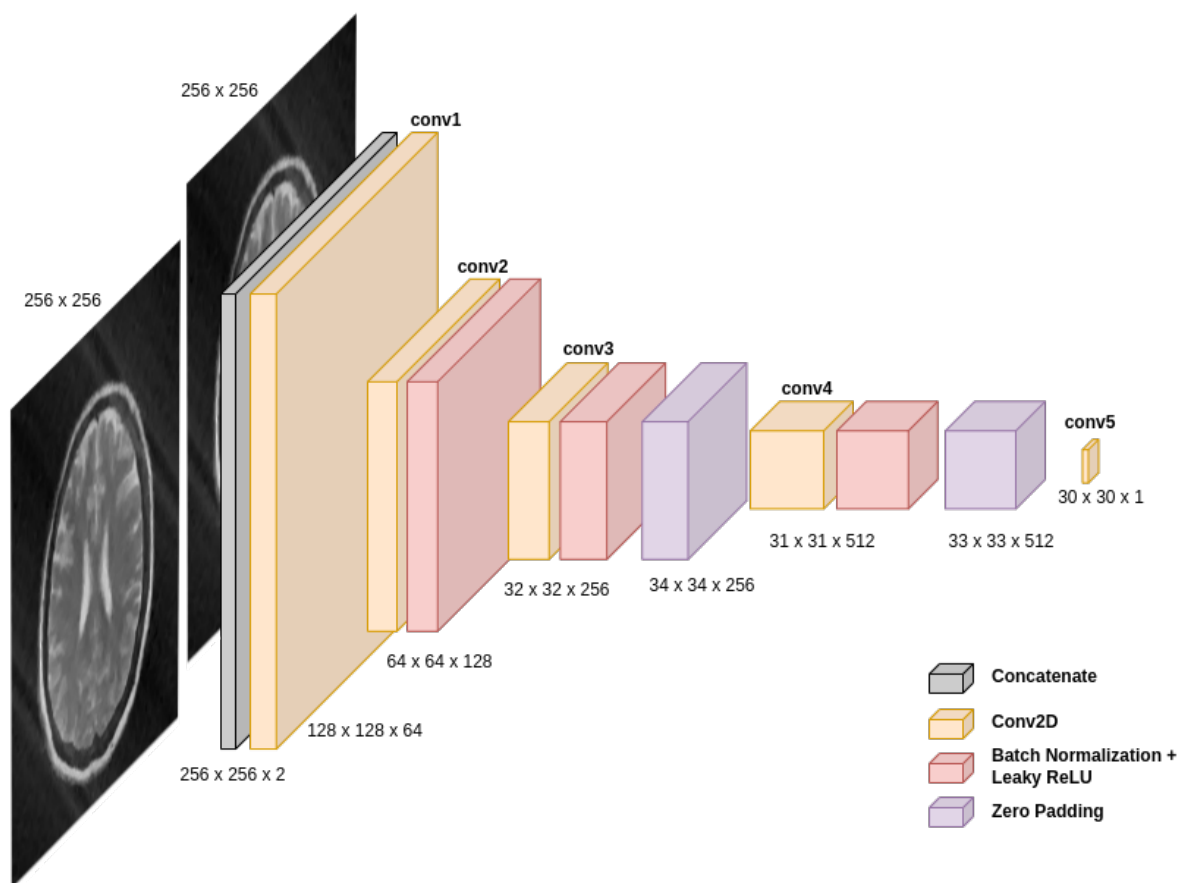


Figura 3.10. O modelo PatchGAN recebe as imagens reais e as geradas em um momento e em outro momento as imagens reais e as amostradas e a partir disso ele retorna uma imagem onde por todos os píxeis ele pode definir se as entradas são reais ou falsas

3.4.3 Função de Ativação ReLU

Em uma rede neural, a função de ativação é responsável por transformar a entrada do nó em uma ativação do próximo nó. A grosso modo, ela ativa o neurônio quando a saída atinge um valor limite definido na função. O neurônio recebe a soma do produto das entradas e pesos inicializados aleatoriamente com um viés constante para cada camada.

A função de ativação é aplicada a esta soma e uma saída é gerada. As funções de ativação introduzem uma não linearidade, com objetivo de fazer com que a rede aprenda padrões complexos nos dados. Sem uma função de ativação, nosso modelo vai se comportar como um modelo de regressão linear com capacidade de aprendizado limitada.

A ReLU é uma função de ativação R que retorna zero para entradas negativas e mantém na saída as entradas positivas, ou seja,

$$R(x) = \max(0, x), \quad \forall x \in \mathbb{R}. \quad (3.1)$$

Trata-se da função de ativação mais comumente usada em redes neurais, especialmente em redes neurais convolucionais e perceptrons multicamadas [43].

3.4.4 Função de Otimização Adam

A otimização de Adam é uma extensão da otimização baseada em gradiente estocástica desenvolvida para aprendizado profundo em processamento de imagem e linguagem natural. Esta técnica, primeiramente introduzida em 2015 por Diederik Kingma [44], usa a média do primeiro e segundo momento dos gradientes, referidos como média e variância descentralizada. O nome Adam é derivado da estimativa de momento adaptativo no inglês Adaptive Moment Estimation. Basicamente o Adam é uma adaptação do Adaptive Gradient Algorithm (AdaGrad) [45] e é importante no treinamento do GAN, pois substituí-lo por métodos não adaptativos (por exemplo, Descida de Gradiente Estocástico (do inglês, Stochastic Gradient Descent) (SGD)) deteriorariam significativamente o desempenho[46].

O AdaGrad mantém uma taxa de aprendizado por parâmetro que melhora o desempenho geral em problemas com gradientes esparsos, como em problemas de linguagem natural e visão computacional. Já o RMSProp é outra adaptação do AdaGrad que também mantém o aprendizado por parâmetro, usando taxas que podem ser adaptadas com base na média de novas magnitudes dos gradientes para o peso [47]. Adam junta os principais benefícios do AdaGrad e do RMSProp. Ao invés de adaptar as taxas de aprendizado dos

parâmetros com base na média do primeiro momento como no RMSProp, Adam também usa a média dos segundos momentos dos gradientes [48]. O otimizador Adam foi usado tanto no gerador quanto no discriminador.

Como proposto no artigo [48], uma boa configuração inicial em problemas de machine learning são $\alpha = 0.001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$. Para o uso do Adam no modelo generativo a melhor configuração foi $\alpha = 0.0005$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 10^{-8}$.

3.4.5 Função de perda BCE

A entropia cruzada binária compara cada uma das probabilidades previstas com a saída da classe real, que pode ser 0 ou 1. Em seguida, calcula a pontuação que penaliza as probabilidades com base na distância do valor esperado. Isso significa quão perto ou longe do valor real. Na classificação binária, A entropia cruzada binária (BCE, do inglês *binary cross-entropy*) é dada por

$$B_{CE} = -(y \log(p) + (1 - y) \log(1 - p)), \quad (3.2)$$

onde y é o indicador binário (0 ou 1) e p é a observação de probabilidade prevista.

3.4.6 Erro Absoluto Médio

Usamos o Erro Absoluto Médio para calcular a média de todos os erros absolutos entre as classificações do Discriminador quando classifica as imagens geradas pelo Gerador. O MAE pode ser representado por,

$$MAE = \frac{1}{n} \sum_{t=1}^n |e_t|, \quad (3.3)$$

onde n é o número de erros e $|e_t|$ são os erros absolutos que serão somados.

3.5 Métricas

Comparar os resultados da restauração requer uma medida da qualidade da imagem. Existem diversas métricas que podem ser usadas para essa comparação e cada uma contem sua característica.

3.5.1 Erro quadrático médio

Erro quadrático médio é uma função de risco que corresponde ao valor esperado do erro quadrático. Ela é a média do erro entre a imagem reconstruída e a imagem original. O valor Erro quadrático médio (do inglês, Mean Squared Error) (MSE) denota a diferença média dos pixels em toda a imagem. Quanto maior o valor de MSE maior a diferença entre a imagem reconstruída e a imagem original. O MSE entre duas imagens, $v(n, m)$ sendo a original e $\hat{v}(n, m)$ a reconstruída, temos:

$$MSE = \frac{1}{N * M} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} [\hat{v}(i, j) - v(i, j)]^2 \quad (3.4)$$

onde, M é a largura e N altura da imagem. O i e j são os pixels das linhas e das colunas de ambas as imagens originais e reconstruída[49].

Um problema com o erro quadrático médio é que ele depende fortemente da escala de intensidade da imagem. A relação sinal-ruído de pico (PSNR) evita este problema ao dimensionar o MSE conforme o intervalo da imagem.

3.5.2 Erro quadrático médio normalizado

O erro quadrático médio normalizado é uma medida da dispersão relativa média e reflete os erros aleatórios entre a imagem original representada pelo vetor v e a reconstruída \hat{v} [50].

$$NMSE_{(\hat{v}, v)} = \frac{\|\hat{v} - v\|_2^2}{\|v\|_2^2} \quad (3.5)$$

Onde $\|\cdot\|_2^2$ é a norma euclidiana quadrada, e a subtração é feita por cada entrada. Quanto menor o valor de erro quadrático médio normalizado melhor é a imagem reconstruída.

3.5.3 Relação Sinal-Ruído de Pico

PSNR é a razão entre potência máxima possível e ruído corruptor que afetam a representação da imagem. PSNR é expresso em escala de decibéis(dB). O sinal é representado pelos dados originais e o ruído é o erro introduzido. Quanto maior o valor de PSNR melhor é a qualidade da imagem reconstruída[51].

$$PSNR_{(\hat{v},v)} = 10 \log_{10} \frac{\max(v)^2}{MSE(\hat{v},v)} \quad (3.6)$$

Onde \hat{v} é a imagem reconstruída e v a imagem original. $\max(v)$ é o maior valor do vetor v , $MSE(\hat{v},v)$ é o erro quadrático médio entre \hat{v} e v . Valores altos de PSNR representam uma melhor qualidade na imagem reconstruída.

3.5.4 Relação sinal-ruído

SNR representa o ruído total presente na borda de saída detectado em uma imagem, em comparação com o ruído no nível do sinal original[52]. A relação sinal-ruído é definida pela relação de potência entre um sinal e o ruído.

$$SNR = \frac{P_{sinal}}{P_{ruído}} \quad (3.7)$$

Onde P é a potência media do sinal, é esperado que P_{sinal} e $P_{ruído}$ tenham a mesma largura de banda.

3.6 Método Proposto para Reconstrução de Imagens de MR Usando CS e CGAN

No método apresentado, aplicamos a arquitetura apresentada como pix2pix como alternativa de procedimento de reconstrução de imagens de ressonância magnética. Através dos pares de imagens x e $x_{sampled}$ divididos em lotes de tamanho 32, geramos novas imagens passando como entrada para o gerador, o $x_{sampled}$. Logo após o discriminador recebe a imagem x com a imagem $x_{sampled}$ e como saída temos a classificação das imagens para saber se as imagens são reais ou falsas. Além disso, o discriminador recebe novamente um par de imagens para dizer se são ou não reais, mas dessa vez recebe como entrada as imagens reais x e as geradas pelo gerador $G(x)$.

A partir da saída da rede discriminadora quando aplicada nas imagens x e $G(x_{sampled})$ é calculado o erro do gerador, ou seja, calculamos o erro do gerador $G(x_{sampled})$ a partir de, $BCE(D(x_{sampled}, G(x_{sampled})))$. Esse erro ainda é somado ao produto entre a função de perda l1 e um λ pré setado com valor 5. Podemos representar a perda l1 por, $MAE(x - G(x_{sampled}))$. Por fim, temos o erro total do gerador, que é maximizado pela maximização do erro do discriminador quando dados gerados são fornecidos a ele. Ou seja, o objetivo

do gerador é gerar dados rotulados que o discriminador classifique como "reais".

Para maximizar o desempenho do discriminador, a perda do discriminador é minimizada quando fornecidos lotes de imagens reais x e gerados $G(x_{sampled})$. O objetivo do discriminador é não ser "enganado" pelo gerador, e aprender as representações dos recursos que são característicos dos dados de treinamento.

O modelo é treinado por 150000 passos, onde cada passo possui um lote composto por um par de imagens para treinamento. Após esse processo de treinamento, o modelo é testado e avaliado a partir de métricas objetivas. Como trabalho futuro é interessante avaliar as imagens do modelo a partir de uma análise clínica.

3.7 Experimentos de Avaliação do Modelo de Reconstrução Proposto

A partir do treinamento do modelo com imagens sub-amostradas no espaço k e reconstruídas a partir do CS, avaliamos as saídas usando o conjunto de teste. Com as imagens reconstruídas pelo modelo generativo podemos medir por métricas objetivas a qualidade de ambas reconstruções (CS vs CGAN), comparando de modo estatístico se há ou não uma melhora em SSIM, PSNR ou SNR.

- PSNR: representa a razão entre a potência do valor máximo de magnitude na imagem e a potência do ruído que impacta a fidelidade da reconstrução. PSNR é representado em dB e quanto maior, maior a qualidade da imagem percebida.
- SSIM: varia de 0 a 1 e mede a semelhança entre duas imagens. Quanto mais próximo de 1, mais os dois sinais são semelhantes.
- SNR: assim como o PSNR, o SNR mede o nível do sinal desejado ao nível de ruído no sinal reconstruído. Também é representado em dB e quanto maior, melhor.

A partir de dados pareados, ou seja, que provêm da mesma população, aplicamos teste de hipótese. Nos casos onde a distribuição é normal usamos o teste paramétrico t-student. Teste para a hipótese nula de que duas amostras relacionadas ou repetidas têm valores médios (esperados) idênticos. Caso uma das distribuições dos vetores de métricas não seja normal, usamos o teste não paramétrico Wilcoxon signed-rank. Teste para hipótese nula de que duas amostras pareadas relacionadas vêm da mesma distribuição. Em particular, ele testa se a distribuição das diferenças $x - y$ é simétrica em relação a zero. É uma versão não paramétrica do t-student pareado [53].

4 RESULTADOS E DISCUSSÕES

4.1 Resultados de reconstrução utilizando base de treinamento e teste com 30 trajetórias radiais

Após o treinamento do modelo CGAN com imagens sub amostradas no espaço k , geramos imagens a partir da base de teste que também possuíam 30 trajetórias radiais. Visualmente há uma redução dos artefatos externos e internos ao cérebro, onde temos uma melhor noção das áreas.

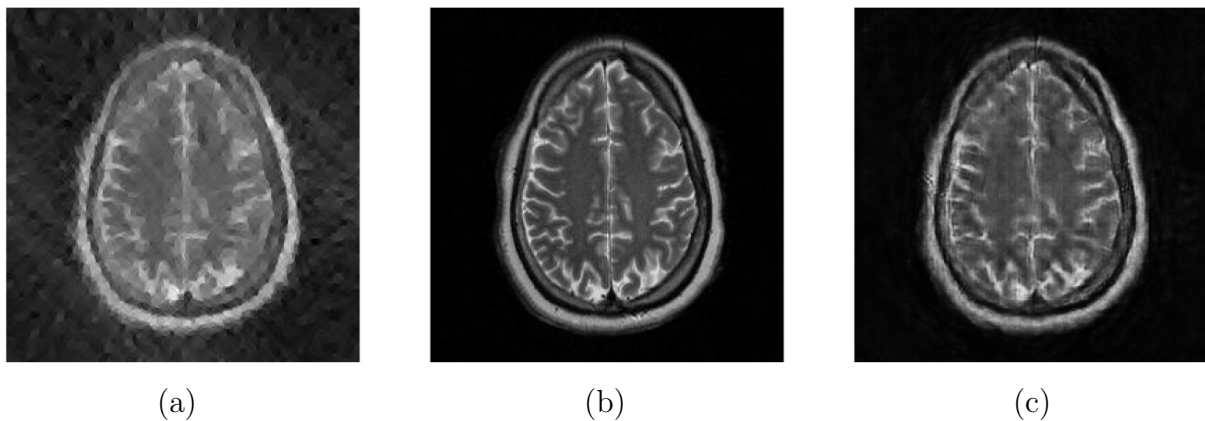


Figura 4.1. Exemplo visual das reconstruções em relação à imagem real onde, (a) Reconstrução CS; (b) Imagem Real; (c) Reconstrução CGAN

Ainda que visualmente melhor a imagem gerada, podemos observar na comparação da figura 4.2 que muitos detalhes foram perdidos. O modelo generativo com objetivo de reconstruir a imagem original pode criar partes que não existem e gerar confusão numa tomada de decisão clínica, por exemplo.

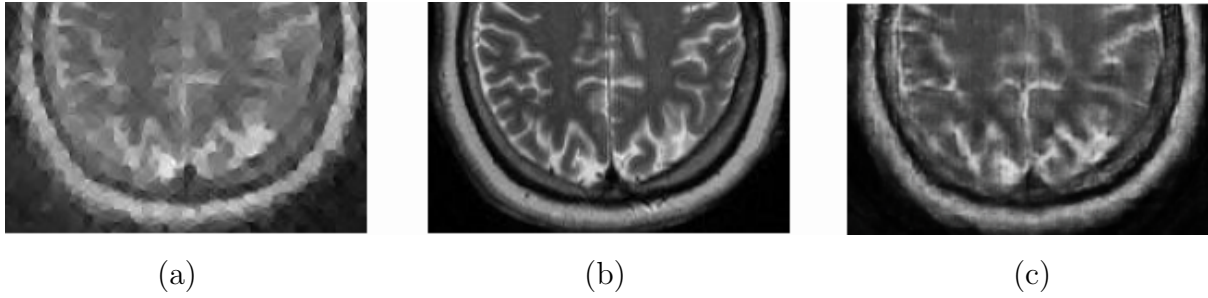


Figura 4.2. Zoom feito na área da imagem que mostra alguns detalhes: (a) Reconstrução CS; (b) Imagem Real; (c) Reconstrução CGAN

Além da avaliação visual que compara regiões importantes do cérebro nas reconstruções, podemos avaliar métricas objetivas das imagens reconstruídas em relação à imagem original. Mostramos na tabela 4.1 a comparação da média de PSNR, SNR e SSIM.

Tabela 4.1. Valores das métricas objetivas de avaliação do modelo testado com imagens sub amostradas por 30 trajetórias radiais.

Métrica	Media CS	Media cGAN	Teste	p
SNR	2.821	11.939	wilcoxon	4.66e-10
SSIM	0.357	0.592	t-student	1.36e-19
PSNR	16.126	25.244	t-student	2.30e-29

4.2 Resultados de reconstrução utilizando 30 trajetórias radiais na base de treinamento e teste com 100

O teste desse modelo treinado com 30 trajetórias radiais possui um valor maior de SSIM para CGAN do que para a reconstrução do CS. Na tabela 4.2 é possível observar que a média dos valores de SNR e PSNR são maiores na reconstrução feita pelo CS. O modelo treinado com poucas medidas e testado com muitas medidas não apresenta melhores resultados por conta da fidelidade da reconstrução no CS. Como o modelo generativo cria alguns pixels a partir da distribuição aprendida, ele modifica algumas regiões importantes para uma análise visual.

Tabela 4.2. Valores das métricas objetivas de avaliação do modelo testado com imagens sub amostradas por 100 trajetórias radiais.

Métrica	Media CS	Media cGAN	Teste	p
SNR	14.956	13.183	t-student	0.0001
SSIM	0.676	0.710	wilcoxon	0.001
PSNR	27.802	26.030	wilcoxon	0.0001

Apesar do valor da métrica objetiva SSIM ser mais alto para a imagem gerada pela CGAN, podemos observar na figura 4.3 que há regiões onde a informação não é a mesma visualmente. Há inúmeros artefatos no em toda imagem que mostram que o modelo generativo gerou partes inconsistentes.

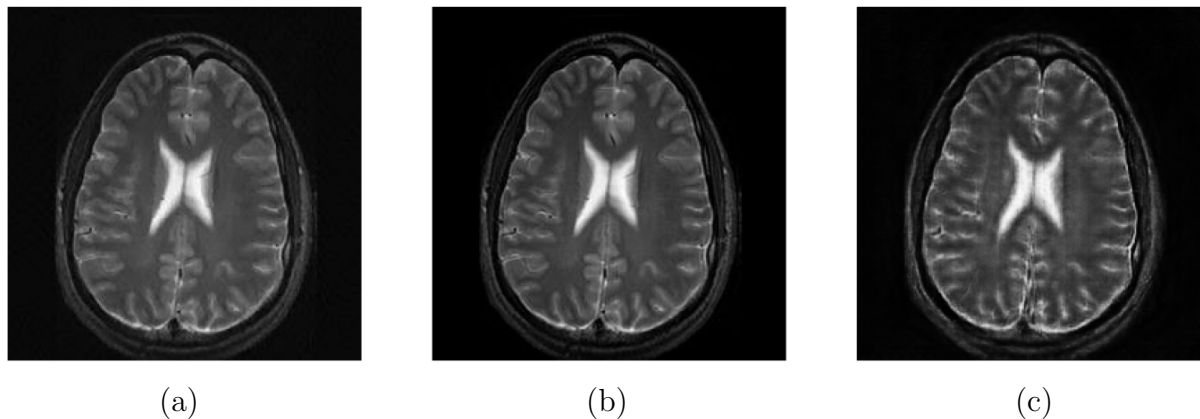


Figura 4.3. Exemplo visual das reconstruções em relação à imagem real, onde: (a) Reconstrução CS; (b) Imagem Real; (c) Reconstrução CGAN.

A partir de um zoom nas imagens mostrada na figura 4.4 podemos notar a diferença entre as regiões observadas. Os pontos escuros na imagem reconstruída e na imagem original não aparecem na imagem gerada. Há também um grande artefato que separa regiões que não são separadas na imagem original.

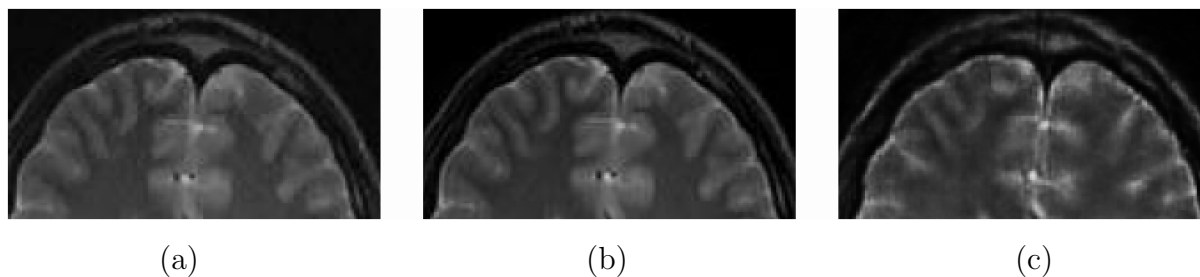


Figura 4.4. Zoom feito na área da imagem que mostra alguns detalhes, onde: (a) Reconstrução CS; (b) Imagem Real; (c) Reconstrução CGAN.

4.3 Resultados de reconstrução utilizando 10 trajetórias radiais na base de treinamento e teste com 30

Treinamos o modelo CGAN com imagens sub amostradas com 10 trajetórias radiais no espaço k para avaliar a qualidade das imagens geradas. As imagens têm 0.039% da informação total, ou seja, com apenas 2527 pontos de 65536 tentamos aproximar a imagem reconstruída da imagem real. Com menos de 10% da informação, o CS apresenta

uma série de artefatos que é impossível visualmente definir regiões da imagem de forma assertiva.

Um exemplo da imagem reconstruída a partir da trajetória de amostragem da figura 3.1 em comparação com a imagem original pode ser vista na figura 4.5. Não há como identificar visualmente nenhum detalhe quando a imagem é amostrada com pouquíssimas medidas.

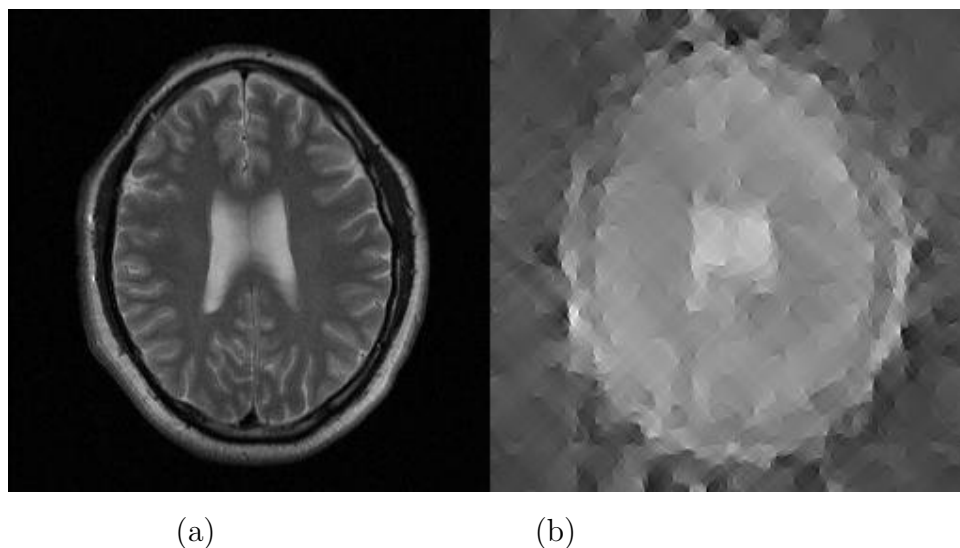


Figura 4.5. Exemplo visual da imagem sub amostrada no espaço k com 30 trajetórias radiais (a) Imagem original; (b) Reconstrução CS a partir da trajetória radial com 10 ângulos;

Após treinar o modelo com esses pares de imagens mostrados na figura 4.5 e testar em imagens sub amostradas com 30 ângulos radiais temos o resultado visual mostrado na 4.6.

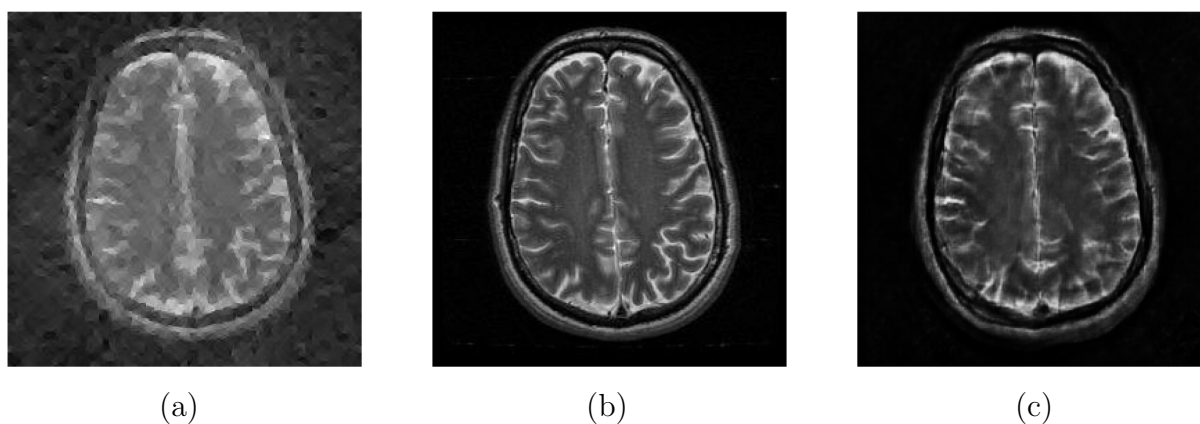


Figura 4.6. Exemplo visual das reconstruções em relação à imagem real, onde: (a) Reconstrução CS; (b) Imagem Real; (c) Reconstrução CGAN

4.3.1 Resultados testando o mesmo modelo utilizando 10 trajetórias radiais na base de teste

O caso mais interessante foi quando testamos o mesmo modelo com imagens também sub amostradas com 10 ângulos radiais. Algumas imagens geradas se parecem muito com as imagens reais nas regiões principais, ou seja, onde temos uma informação mais grosseira do cérebro. Apesar de perder detalhes, a imagem gerada é bem mais próxima da imagem reconstruída, tanto nas métricas objetivas quanto na avaliação visual como mostra na figura 4.7.

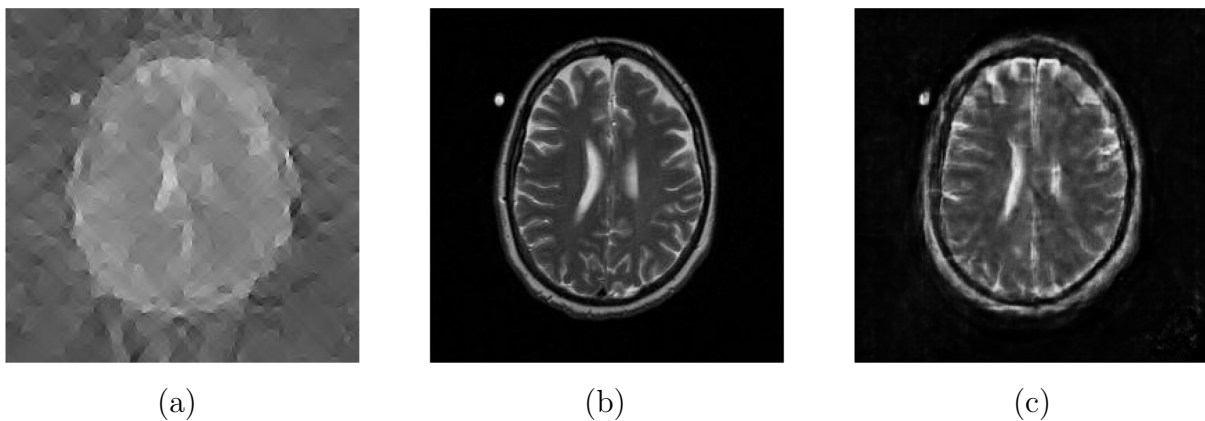


Figura 4.7. Exemplo visual das reconstruções em relação à imagem real, onde: (a) Reconstrução CS; (b) Imagem Real; (c) Reconstrução CGAN

Há casos onde a estrutura mais detalhada do cérebro é completamente criada pelo modelo generativo, criando uma divergência com a imagem original como mostra na figura 4.8. Como visto na figura anterior, a estrutura principal da imagem gerada é ainda bem melhor do que a da imagem reconstruída pelo CS.

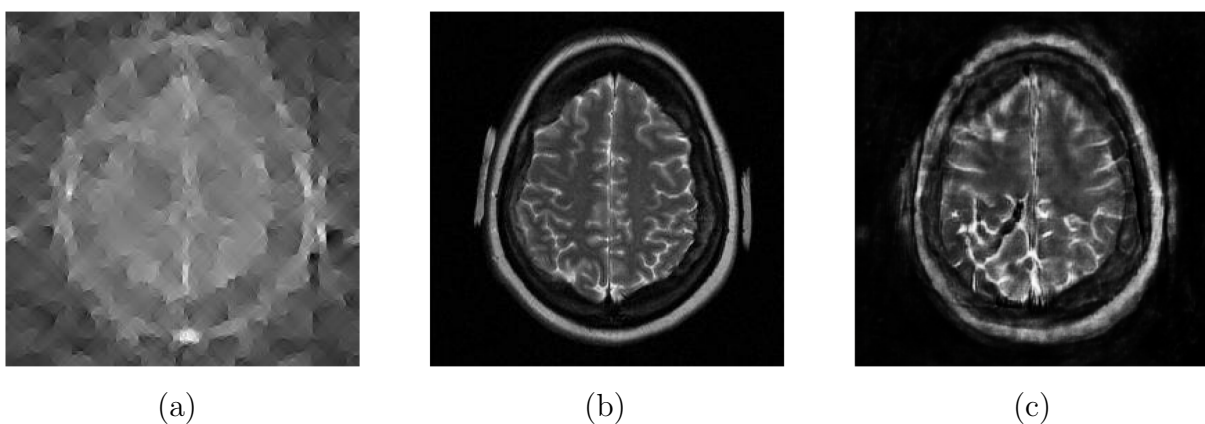


Figura 4.8. Exemplo visual das reconstruções em relação à imagem real, onde: (a) Reconstrução CS; (b) Imagem Real; (c) Reconstrução CGAN.

Podemos ver na tabela 4.3 que para o caso do SNR na media das imagens reconstruídas por CS temos um valor negativo, que significa que o poder do sinal é menor que o poder do ruído. As médias de SSIM e PSNR também foram favoráveis ao modelo generativo CGAN.

Tabela 4.3. Valores das métricas objetivas de avaliação do modelo testado com imagens sub amostradas por 10 trajetórias radiais

Métrica	Media CS	Media cGAN	Teste	p
SNR	-1.535	8.963	t-student	2.03e-31
SSIM	0.205	0.519	wilcoxon	4.66e-10
PSNR	11.436	21.934	t-student	2.03e-31

5 CONCLUSÃO

Através da avaliação da qualidade objetiva das imagens, onde temos a média para cada uma das métricas, é possível notar que a média de todas as métricas ficaram melhor usando o modelo CGAN para tentar reconstruir a imagem sub amostrada. Porém, como observamos no exemplo da figura 4.7, o modelo generativo criou a partir do que ele aprendeu. Além das métricas objetivas, para afirmar que o modelo generativo reconstrói uma imagem com maior proximidade da imagem original devemos considerar avaliação mais confiável da qualidade perceptual que é a MOS, que basicamente pede a radiologistas experientes para avaliar as imagens reconstruídas. Tipicamente, as imagens são classificadas de 0 a 4, dependendo da qualidade da imagem reconstruída (ou seja, não-diagnóstica, ruim, regular, boa e excelente), e a MOS final é calculada como a média aritmética das pontuações de todos os avaliadores. Em muitos casos tivemos a sensação visual de que o modelo generativo criou partes do cérebro que não existia. Os treinamentos levam cerca de 40 minutos com a configuração de máquina descrita no Capítulo 3, então é possível parametrizar melhor o modelo para achar os melhores hiper parâmetros do gerador e discriminador. Os resultados apresentados para 10 trajetórias de sub amostragens tanto no teste quanto no treino mostrou que para extração de informação a priori, a partir de uma imagem com poucas medidas podemos definir as regiões que não possui muitos artefatos. A velocidade para gerar uma imagem com a CGAN leva uma fração de segundos, diferente do CS que possui um tempo maior.

5.1 Trabalhos futuros

A partir dos resultados notamos haver uma série de outros trabalhos extras para melhorar a imagem gerada e usar ela não como reconstrução, mas sim como informação a priori para uma reconstrução a partir de CS.

- Adicionar aumento de dados nas imagens sub amostradas para treinar o modelo CGAN: Tendo em vista o objetivo da U-NET em segmentar imagens com menos amostras apenas deslocando ou rotacionando a mesma imagem.

- Usar as imagens geradas como informação a priori: É possível validar a hipótese se usando a imagem gerada pela CGAN como informação a priori no processo de reconstrução do CS temos algum ganho expressivo.

Referências Bibliográficas

- [1] Jong Chul Ye. Compressed sensing MRI: a review from signal processing perspective. *BMC Biomedical Engineering*, 1(1):8, December 2019.
- [2] Shaily jain. Supervised learning, 2021. Available at: <https://codatalicious.medium.com/supervised-learning-6564e32fc79e>. Last access on January 10th, 2023.
- [3] Dor Bank, Noam Koenigstein, e Raja Giryes. Autoencoders, 2021.
- [4] C. Nimsky e R. Fahlbusch. *Medical Technologies in Neurosurgery*. Acta Neurochirurgica Supplement. Springer Vienna, 2007.
- [5] Eugene Lin e Adam Alessio. What are the basic concepts of temporal, contrast, and spatial resolution in cardiac ct? *Journal of cardiovascular computed tomography*, 3:403–8, 08 2009.
- [6] F.H. Doyle, J.M. Pennock, J.S. Orr, J.C. Gore, G.M. Bydder, R.E. Steiner, I.R. Young, H. Clow, D.R. Bailes, M. Burl, D.J. Gilderdale, e P.E. Walters. Imaging of the brain by nuclear magnetic resonance. *The Lancet*, 318(8237):53–57, 1981. Originally published as Volume 2, Issue 8237.
- [7] Justin E. Vranic, Thien J. Huynh, Peter Fata, Jason Barber, Robert H. Bonow, Michael R. Levitt, Nancy Temkin, Ryan Morton, Cory Kelly, Dean K. Shibata, e Mahmud Mossa-Basha. The ability of magnetic resonance black blood vessel wall imaging to evaluate blunt cerebrovascular injury following acute trauma. *Journal of Neuroradiology*, 47(3):210–215, 2020.
- [8] Agência Nacional de Saúde Suplementar. ANS publica dados sobre assistência prestada pelos planos de saúde. <https://www.gov.br/ans/pt-br/assuntos/noticias/numeros-do-setor/ans-publica-dados-sobre-assistencia-prestada-pelos-planos-de-saude-2>, 2022. [Online; accessed 30-Dezembro-2022].

- [9] C Mackechnie Jarvis. Nikola tesla and the induction motor. *Physics Education*, 5(5):280, sep 1970.
- [10] P. C. Lauterbur. Image Formation by Induced Local Interactions: Examples Employing Nuclear Magnetic Resonance. , 242(5394):190–191, March 1973.
- [11] Slavik D. Tabakov. Professor john mallard, 1927 – 2021. *Health and Technology*, 11(3):443–444, May 2021.
- [12] Cristiano Jacques Miosso. *Compressive Sensing with Prior Information Applied to Magnetic Resonance Imaging*. PhD thesis, Department of Electrical and Computer Engineering, University of Texas at El Paso (UTEP).
- [13] Jeffrey A. Fessler. Model-based image reconstruction for mri. *IEEE Signal Processing Magazine*, 27(4):81–89, 2010.
- [14] C.J. Miosso, R. von Borries, M. Argaez, L. Velazquez, C. Quintero, e C.M. Potes. Compressive Sensing Reconstruction With Prior Information by Iteratively Reweighted Least-Squares. *IEEE Transactions on Signal Processing*, 57(6):2424–2431, June 2009.
- [15] C. J. Miosso, R. von Borries, e J. H. Pierluissi. Compressive sensing method for improved reconstruction of gradient-sparse magnetic resonance images. In *2009 Conference Record of the Forty-Third Asilomar Conference on Signals, Systems and Computers*, pages 799–806, Pacific Grove, CA, USA, 2009. IEEE.
- [16] Raymond Damadian. Tumor detection by nuclear magnetic resonance. *Science*, 171(3976):1151–1153, 1971.
- [17] Jure Zbontar, Florian Knoll, Anuroop Sriram, Tullie Murrell, Zhengnan Huang, Matthew J. Muckley, Aaron Defazio, Ruben Stern, Patricia Johnson, Mary Bruno, Marc Parente, Krzysztof J. Geras, Joe Katsnelson, Hersh Chandarana, Zizhao Zhang, Michal Drozdal, Adriana Romero, Michael Rabbat, Pascal Vincent, Nafissa Yakubova, James Pinkerton, Duo Wang, Erich Owens, C. Lawrence Zitnick, Michael P. Recht, Daniel K. Sodickson, e Yvonne W. Lui. fastmri: An open dataset and benchmarks for accelerated mri, 2018.
- [18] R Mezrich. A perspective on k-space. *Radiology*, 195(2):297–315, 1995. PMID: 7724743.
- [19] Min Oh e Liqing Zhang. Deepmicro: deep representation learning for disease prediction based on microbiome data. *Scientific Reports*, 10, 04 2020.

- [20] Craig H. Meyer, Bob S. Hu, Dwight G. Nishimura, e Albert Macovski. Fast spiral coronary artery imaging. *Magnetic Resonance in Medicine*, 28(2):202–213, 1992.
- [21] Klaus Scheffler e Jürgen Hennig. Reduced circular field-of-view imaging. *Magnetic Resonance in Medicine*, 40(3):474–480, 1998.
- [22] C.E. Shannon. Communication in the presence of noise. *Proceedings of the IRE*, 37(1):10–21, jan 1949.
- [23] Kirk L. Kroeker. Rethinking signal processing. *Commun. ACM*, 52(5):13–15, May 2009.
- [24] D.L. Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, 2006.
- [25] Filipe Emídio Tôrres. Avaliação de representações transformadas para compressão de sinais de eletroencefalografia, com base em análise de componentes principais, decomposições wavelet, transformada discreta de cossenos e compressive sensing. Master’s thesis, Universidade de Brasília, 2018.
- [26] E.J. Candes, J. Romberg, e T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489–509, February 2006.
- [27] Emmanuel J Candes. The restricted isometry property and its implications for compressed sensing. *Comptes rendus mathématique*, 346(9-10):589–592, 2008.
- [28] Miroslav Kubat. Neural networks: a comprehensive foundation by simon haykin, macmillan, 1994, isbn 0-02-352781-7. *The Knowledge Engineering Review*, 13(4):409–412, 1999.
- [29] Robson Alcântara Santana. *Roteamento em Redes Ópticas Transparentes Utilizando Redes Neurais de Hopfield.[Sl]*. PhD thesis, 01 2010.
- [30] Warren S McCulloch e Walter Pitts. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, 5(4):115–133, 1943.
- [31] Gao Huang, Shiji Song, Jatinder N. D. Gupta, e Cheng Wu. Semi-supervised and unsupervised extreme learning machines. *IEEE Transactions on Cybernetics*, 44(12):2405–2417, 2014.
- [32] Stephen Whitelam, Viktor Selin, Sang-Won Park, e Isaac Tamblyn. Correspondence between neuroevolution and gradient descent. *CoRR*, abs/2008.06643, 2020.

- [33] David E. Rumelhart, Geoffrey E. Hinton, e Ronald J. Williams. *Learning Representations by Back-Propagating Errors*, page 696–699. MIT Press, Cambridge, MA, USA, 1988.
- [34] Simon S. Haykin e Simon S. Haykin. *Neural networks and learning machines*. Prentice Hall, New York, 3rd ed edition, 2009. OCLC: ocn237325326.
- [35] Kunihiro Fukushima e Sei Miyake. Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition. In Shun-ichi Amari e Michael A. Arbib, editors, *Competition and Cooperation in Neural Nets*, pages 267–285, Berlin, Heidelberg, 1982. Springer Berlin Heidelberg.
- [36] Y. Lecun, L. Bottou, Y. Bengio, e P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.
- [37] Douglas Reynolds. *Gaussian Mixture Models*, pages 659–663. Springer US, Boston, MA, 2009.
- [38] David E. Rumelhart e James L. McClelland. *Information Processing in Dynamical Systems: Foundations of Harmony Theory*, pages 194–281. 1987.
- [39] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, e Yoshua Bengio. Generative adversarial networks, 2014.
- [40] Dor Bank, Noam Koenigstein, e Raja Giryes. Autoencoders, 2020.
- [41] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, e Alexei A. Efros. Image-to-image translation with conditional adversarial networks, 2016.
- [42] Olaf Ronneberger, Philipp Fischer, e Thomas Brox. U-net: Convolutional networks for biomedical image segmentation, 2015.
- [43] Abien Fred Agarap. Deep learning using rectified linear units (relu). 03 2018.
- [44] Diederik P. Kingma e Jimmy Ba. Adam: A method for stochastic optimization, 2014.
- [45] John Duchi, Elad Hazan, e Yoram Singer. Adaptive subgradient methods for on-line learning and stochastic optimization. *Journal of Machine Learning Research*, 12(61):2121–2159, 2011.
- [46] Mingrui Liu, Youssef Mroueh, Jerret Ross, Wei Zhang, Xiaodong Cui, Payel Das, e Tianbao Yang. Towards better understanding of adaptive gradient algorithms in generative adversarial nets, 2019.

- [47] Geoffrey Hinton, Nitish Srivastava, e Kevin Swersky. Neural networks for machine learning lecture 6a overview of mini-batch gradient descent. *Cited on*, 14(8):2, 2012.
- [48] Diederik P. Kingma e Jimmy Ba. Adam: A method for stochastic optimization, 2014.
- [49] Dominic Asamoah, Emmanuel Oppong, Stephen Oppong, e Juliana Danso. Measuring the performance of image contrast enhancement technique. *International Journal of Computer Applications*, 181:6–13, 10 2018.
- [50] Attilio A. Poli e Mario C. Cirillo. On the use of the normalized mean square error in evaluating dispersion model performance. *Atmospheric Environment. Part A. General Topics*, 27(15):2427–2434, 1993.
- [51] Jure Zbontar, Florian Knoll, Anuroop Sriram, Matthew J. Muckley, Mary Bruno, Aaron Defazio, Marc Parente, Krzysztof J. Geras, Joe Katsnelson, Hersh Chandarana, Zizhao Zhang, Michal Drozdal, Adriana Romero, Michael G. Rabbat, Pascal Vincent, James Pinkerton, Duo Wang, Nafissa Yakubova, Erich Owens, C. Lawrence Zitnick, Michael P. Recht, Daniel K. Sodickson, e Yvonne W. Lui. fastmri: An open dataset and benchmarks for accelerated MRI. *CoRR*, abs/1811.08839, 2018.
- [52] Rumelhart, E. David, James McClelland, e James L. *Parallel distributed processing: explorations in the microstructure of cognition. Volume 1. Foundations*. 01 1986.
- [53] Frank Wilcoxon. Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6):80–83, 1945.