



Universidade de Brasília – UnB
Campus Gama – FCTE
Programa de Pós-Graduação em Engenharia Biomédica

**IMPLEMENTAÇÃO E ANÁLISE COMPARATIVA DE RECONSTRUÇÃO
DE IMAGENS DE RESSONÂNCIA MAGNÉTICA UTILIZANDO
REDES ADVERSÁRIAS GERATRIVAS**

AMAURI DA COSTA JÚNIOR

Orientador: GERARDO ANTONIO IDROBO PIZO



UNB – UNIVERSIDADE DE BRASÍLIA
FCTE – FACULDADE DE CIÊNCIAS E TECNOLOGIAS EM ENGENHARIA



Programa de
Pós-Graduação em
ENGENHARIA BIOMÉDICA

**IMPLEMENTAÇÃO E ANÁLISE COMPARATIVA DE RECONSTRUÇÃO DE
IMAGENS DE RESSONÂNCIA MAGNÉTICA UTILIZANDO REDES
ADVERSÁRIAS GENERATIVAS**

AMAURI DA COSTA JÚNIOR

ORIENTADOR: GERARDO ANTONIO IDROBO PIZO

DISSERTAÇÃO DE MESTRADO EM
ENGENHARIA BIOMÉDICA

PUBLICAÇÃO: 201A/2025

BRASÍLIA/DF, MARÇO DE 2025

UNB – UNIVERSIDADE DE BRASÍLIA
FCTE – FACULDADE DE CIÊNCIAS E TECNOLOGIAS EM ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO

**IMPLEMENTAÇÃO E ANÁLISE COMPARATIVA DE RECONSTRUÇÃO DE
IMAGENS DE RESSONÂNCIA MAGNÉTICA UTILIZANDO REDES
ADVERSÁRIAS GENERATIVAS**

AMAURI DA COSTA JÚNIOR

DISSERTAÇÃO DE MESTRADO SUBMETIDA AO PROGRAMA DE PÓS-GRADUAÇÃO EM
ENGENHARIA BIOMÉDICA DA UNIVERSIDADE DE BRASÍLIA, COMO PARTE DOS RE-
QUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE MESTRE EM ENGENHARIA
BIOMÉDICA

APROVADA POR:

Gerardo Antonio Idrobo Pizo
(Orientador)

Dr. Renato Coral Sampaio
(Examinador interno)

Dr. Alexandre Eurípedes Ferreira
(Examinador externo)

FICHA CATALOGRÁFICA

COSTA, AMAURI

Implementação e análise comparativa de reconstrução de imagens de ressonância magnética utilizando redes adversárias generativas
[Distrito Federal], 2025.

50p., 210 × 297 mm (FCTE/UnB Campus Gama, Mestrado em Engenharia Biomédica, 2025).
Dissertação de Mestrado em Engenharia Biomédica, Faculdade UnB Campus Gama, Programa de Pós-Graduação em Engenharia Biomédica.

- | | |
|-----------------------------------|-----------------------------|
| 1. Ressonância Magnética | 2. Processamento de Imagens |
| 3. Redes Generativas Adversariais | 4. Inteligência Artificial |
| I. FCTE UnB/UnB. | II. Título (série) |

REFERÊNCIA

COSTA, AMAURI (2025). Implementação e análise comparativa de reconstrução de imagens de ressonância magnética utilizando redes adversárias generativas. Dissertação de mestrado em engenharia biomédica, Publicação 201A/2025, Programa de Pós-Graduação, Faculdade UnB Campus Gama, Universidade de Brasília, Brasília, DF, 50p.

CESSÃO DE DIREITOS

AUTOR: Amauri da Costa Júnior

TÍTULO: Implementação e análise comparativa de reconstrução de imagens de ressonância magnética utilizando redes adversárias generativas

GRAU: Mestre

ANO: 2025

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

amauri_cj@hotmail.com

Brasília, DF – Brasil

AGRADECIMENTOS

Agradeço ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo apoio financeiro concedido, que foi fundamental para a realização desta pesquisa. O incentivo proporcionado contribuiu de maneira significativa para o desenvolvimento das atividades acadêmicas e científicas vinculadas a este trabalho.

RESUMO

Redes gerativas adversárias (GANs) são um modelo de inteligência artificial muito usado no processamento de imagens. Neste trabalho será abordado o uso dessa ferramenta no auxílio à reconstrução de imagens de ressonância magnética feitas com dados de bobina única sub-amostrados. A GAN foi usada para melhorar as métricas de qualidade das imagens reconstruídas com diferentes trajetórias, e que apresentavam distorções por não respeitarem o teorema da amostragem de Nyquist–Shannon. Testes são realizados alimentando o modelo treinado com imagens feitas com dados sub-amostrados e então gerando novas imagens aprimoradas. O modelo gerador da GAN recebe imagens feitas com dados sub-amostrados e tenta criar imagens aprimoradas baseadas em imagens típicas de ressonância magnética, enquanto o modelo discriminador da GAN recebe as imagens geradas e as imagens típicas e tenta adivinhar quais são verdadeiras e quais são falsas. A cada época os dois modelos se refinam baseando-se no veredito do discriminador, melhorando a qualidade das imagens geradas. Por fim, são extraídas métricas das imagens produzidas como índice de similaridade estrutural e relação sinal ruído. Foram obtidos aumentos de até 100% no índice de similaridade estrutural médio e de até 128,74% na relação sinal ruído média (Teste 78). Os resultados obtidos se equiparam a resultados de outros modelos que usaram o mesmo conjunto de dados e superaram outros modelos GANs nesse mesmo aspecto.

Palavras-chave: Ressonância Magnética; GAN; Reconstrução; Sub-amostrado; Inteligência Artificial.

ABSTRACT

Generative adversarial networks (GANs) are an artificial intelligence model widely used in image processing. This paper will address the use of this tool to aid in the reconstruction of magnetic resonance images made with undersampled single-coil data. The GAN was used to improve the quality metrics of images reconstructed with different trajectories, which presented distortions due to not respecting the Nyquist–Shannon sampling theorem. Tests are performed by feeding the trained model with images made with undersampled data and then generating new enhanced images. The GAN generator model receives images made with undersampled data and tries to create enhanced images based on typical magnetic resonance images, while the GAN discriminator model receives the generated images and the typical images and tries to guess which are true and which are false. At each epoch, the two models refine themselves based on the discriminator’s verdict, improving the quality of the generated images. Finally, metrics such as structural similarity index and signal-to-noise ratio are extracted from the produced images. Increases of up to 100% in the average structural similarity index and up to 128.74% in the average signal-to-noise ratio were obtained (Test 78). The results obtained are similar to results from other models that used the same data set and outperformed other GAN models in this same aspect.

Keywords: MRI; GAN; Reconstruction; Undersampled; Artificial Intelligence.

SUMÁRIO

1	Introdução	1
1.1	Problema	1
1.2	Proposta de Pesquisa	5
1.3	Objetivos	6
1.3.1	Objetivo Geral	6
1.3.2	Objetivos Específicos	6
2	Fundamentação Teórica	7
2.1	Aquisição de Dados e Reconstrução da Imagem	7
2.1.1	Campo Magnético Principal e Vetor Magnetização	7
2.1.2	Pulso de Radiofrequência	9
2.1.3	Codificação de Frequência e de Fase	10
2.1.4	Espaço-k	10
2.1.5	Sub-Amostragem	12
2.2	Redes Generativas Adversariais (GANs)	13
2.2.1	Modelo Generativo	13
2.2.2	Modelo Discriminativo	14
2.3	Levantamento da Literatura Científica	14
2.3.1	Trabalhos Correlatos	15
3	Materiais e Métodos	18
3.1	Banco de Dados	18
3.2	Pré-Processamento	19
3.3	Modelo GAN	20

3.4	Métricas de avaliação	20
4	Resultados e Discussões	23
4.1	Resultados	23
4.1.1	Pré-Processamento	23
4.1.2	Rede Generativa Adversária	27
4.1.3	Rede Generativa	29
4.1.4	Rede Discriminativa	32
4.1.5	Imagens	34
4.1.6	Gráficos de Perda (Loss)	35
4.1.7	Tabelas	37
4.2	Discussão	37
4.2.1	Métricas de Qualidade	38
4.2.2	Completar Espaço-k	41
4.2.3	Comparação com Outros Modelos	42
4.2.4	Comparação Entre Trajetórias	43
5	Conclusão	45
	Lista de Referências	46

LISTA DE TABELAS

2.1	Tempos T1 e T2 para diferentes tecidos do corpo humano para um exame realizado com campo magnético de 3T [36].	10
3.1	Organização do conjunto de dados [42].	19
4.1	Métricas de qualidade para testes da trajetória <i>spiral4</i>	37
4.2	Métricas de qualidade dos melhores resultados de cada trajetória. . . .	37
4.3	SSIM das imagens obtidas com o teste que combina os espaços-k. . . .	41
4.4	Comparação do modelo criado com os da competição do FastMRI. . . .	43

LISTA DE FIGURAS

1.1	Orientação do campo magnético $B_0(z)$ e pulso B_1	2
1.2	Comparação entre imagens produzidas a partir de aquisição com diferentes números de etapas de codificação de fase.	3
1.3	Valor de mercado da inteligência artificial no imageamento por ressonância magnética[32].	4
1.4	Resultados de buscas relacionadas a GANs nos últimos dez anos.	5
2.1	Vetor Magnetização.	8
2.2	Espaço-k.	11
2.3	Trajetórias de obtenção do espaço-k.	12
2.4	Comparação entre imagem reconstruída respeitando critério de Nyquist e imagem feita com dados sub-amostrados apresentando aliasing.	13
2.5	Fluxo de produção e processamento de imagens em um modelo GAN. . .	14
2.6	Fluxograma de triagem dos artigos selecionados.	15
3.1	Áreas de interesse para o cálculo da relação sinal ruído (SNR).	21
4.1	Diferentes trajetórias que podem ser aplicadas ao espaço-k.	25
4.2	Imagens de joelho do banco de dados do FastMRI.	26
4.3	Imagen do conjunto de dados utilizado pela GAN.	27
4.4	Imagens produzidas pelo algoritmo comparando as imagens de entrada, tipicamente amostrada e aprimorada.	29
4.5	Arquitetura do modelo Gerador.	31
4.6	Arquitetura do modelo Discriminador.	33
4.7	Imagens geradas pela GAN durante os testes.	35
4.8	Curva de perda no treino/validação do gerador em função da época . .	36
4.9	Curva de perda do gerador e do discriminador em função da época. . .	36

4.10	Boxplot de SSIM para imagens de entrada e saída do teste 80	38
4.11	Melhores imagens produzidas no teste 80	39
4.12	Boxplot do SSIM de testes que ocorreram melhora nas métricas de qualidade.	40
4.13	Imagens do teste que combina os espaços-k.	41
4.14	Imagens referência, de entrada, geradas e combinadas para comparação.	42
4.15	Imagen produzida pelo teste n° 53.	44

LISTA DE NOMENCLATURAS E ABREVIAÇÕES

GAN Rede Adversária Generativa, do inglês *Generative Adversarial Network*

GPU Unidade de Processamento Gráfico, do inglês *Graphics Processing Unit*

MRI Imageamento por Ressonância Magnética, do inglês *Magnetic Resonance Imaging*

SNR Relação Sinal Ruído, do inglês *Signal Noise Ratio*

SSIM Índice de Similaridade Estrutural, do inglês *Structural Similarity Index Measure*

1 INTRODUÇÃO

Exames de imagem estão entre as ferramentas mais utilizadas na medicina para o diagnóstico e tratamento de diversas patologias. Entre os diversos tipos de exames, a ressonância magnética tem se mostrado um dos mais versáteis exames de imagem e o seu uso cresce a cada ano [30].

A ressonância magnética é capaz de obter imagens de corpos moles ou densos e até mesmo de fluxo, volume e oxigenação do sangue [25]. Também é um exame que utiliza de radiação não ionizante, o que traz riscos显著mente menores para saúde se comparados com outros exames de imagem, como por exemplo tomografias [17], que apesar de se tratar de um exame seguro e com quantidades aceitáveis de radiação não pode ser feito frequentemente para que a dose de radiação absorvida pelo corpo não ultrapasse os limites, que caso ocorra aumentam as chances de surgimento de câncer no paciente [4].

1.1 PROBLEMA

Um dos pontos negativos do exame de ressonância é o desconforto causado pelo procedimento. O exame dura em torno de 30 minutos, o paciente deve ficar sem se movimentar em um ambiente apertado e ruidoso. O ambiente do exame se torna um local difícil de permanecer por muito tempo causando movimentação involuntária em alguns pacientes e essa movimentação atrapalha a aquisição dos sinais, prejudicando a qualidade da imagem final com a criação de artefatos [41]. Outro ponto negativo é que devido ao longo tempo de exame, poucos podem ser realizados por dia, aumentando o seu preço, fazendo com que exames de ressonância magnética estejam entre um dos exames mais caros de imageamento médico [31].

O exame de ressonância magnética consiste em aplicar uma sequência de pulsos magnéticos ao longo do corpo do paciente, fazendo com que os núcleos de hidrogênio dos átomos do corpo se alinhem com o campo imposto. Primeiramente é aplicado um campo magnético gradiente ao longo do corpo do paciente chamado de campo magnético principal ($B_0(z)$), esse campo faz com que os núcleos de hidrogênios se alinhem a ele. Um pulso magnético (B_1) é aplicado perpendicular ao primeiro, como mostrado na Figura 1.1,

fazendo com que os núcleos se alinhem ao novo campo. Quando o segundo campo é desligado, os átomos tendem a se alinhar novamente com o primeiro campo voltando à posição original, porém nesse processo eles emitem um pequeno campo magnético que varia para átomos em diferentes partes do corpo.

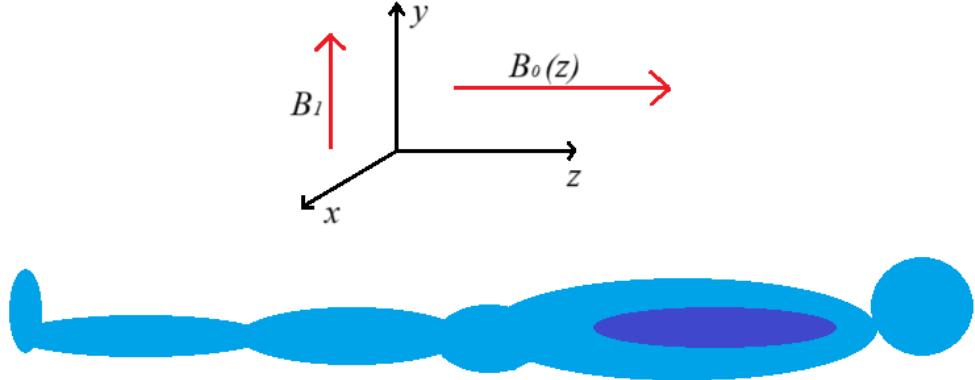


Figura 1.1. Orientação do campo magnético $B_0(z)$ e pulso B_1 .

Bobinas distribuídas ao redor do paciente detectam a intensidade do campo magnético emitido por esses átomos de hidrogênio que voltam a se alinhar ao primeiro campo. São realizadas seguidas alterações na frequência (codificação de frequência) e fase (codificação de fase) do segundo campo emitido, sempre coletando a resposta do corpo a essas alterações no campo. Os sinais obtidos do corpo são organizados em uma matriz complexa, onde cada valor possui uma frequência e uma fase, essa matriz é denominada espaço-k. Por essa matriz se tratar de um espectro de frequências e fases é possível transformá-la em uma imagem de um corte transversal do paciente através de uma transformada inversa de Fourier [27].

Quanto maior o tamanho dessa matriz, ou seja, quanto maior a quantidade de sinais obtidos do corpo, maior a nitidez da imagem final [13], porém para obter uma quantidade maior de sinais, é necessário um maior tempo de exame, o que aumenta o desconforto do paciente e diminui a quantidade de exames feitos por dia o que consequentemente aumenta o preço desse exame.

Uma das formas de diminuir o tempo de exame é adquirindo menos amostras de sinal do corpo, porém com a diminuição do número de amostras diminui-se a qualidade da imagem final obtida. A imagem reconstruída com menos amostras fica borrada, com poucos detalhes, na Figura 1.2 é apresentada uma comparação entre a qualidade das imagens reconstruídas para diferentes quantidades de etapas de codificação de fase realizada na aquisição dos dados. Deve-se então encontrar um equilíbrio entre tempo de exame (quantidade de amostras obtidas) e nitidez desejada para a imagem final.

Estudos são realizados na área de ressonância magnética buscando diminuir o tempo dos exames adquirindo menos sinais do corpo, porém sem comprometer a qualidade

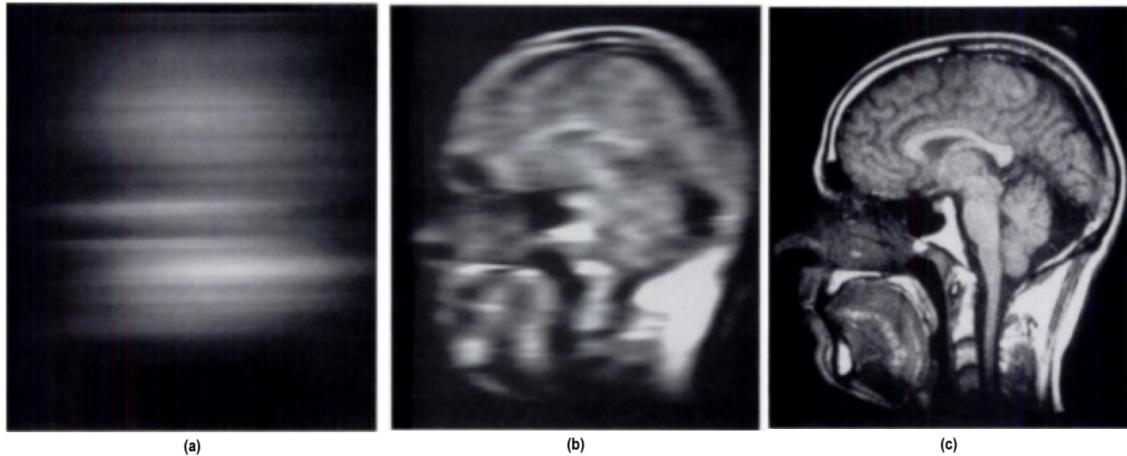


Figura 1.2. Comparação entre imagens produzidas a partir de aquisição com diferentes números de etapas de codificação de fase. (a) 2 etapas de codificação de fase. (b) 16 etapas de codificação de fase. (c) 256 etapas de codificação de fase. Adaptado de [13].

das imagens. Técnicas como imageamento paralelo, *Compressed Sensing*, melhora nos gradientes de campo magnético e melhora na sequência de pulsos magnéticos são apenas algumas das técnicas empregadas buscando a diminuição desse tempo. Reduzir o tempo do exame também traz o benefício de otimizar o tempo da máquina, exames mais velozes permitiriam que mais pacientes sejam examinados diminuindo o custo.

Existem várias frentes de estudo que buscam obter imagens de qualidade mesmo com amostras abaixo do mínimo necessário para uma reconstrução comum, como por exemplo o uso de *Compressed Sensing*, que consegue recuperar imagens a partir de sinais amostrados abaixo da taxa de Nyquist [34].

Uma dessas frentes de estudo é o uso de aprendizado de máquina na reconstrução e no aprimoramento dessas imagens. Modelos de processamento de imagens que utilizam machine learning e deep learning são amplamente utilizados pois conseguem perceber padrões que normalmente outros modelos matemáticos ou de software não conseguiram [24]. Desde reconhecimento facial até a identificação de tumores em imagens médicas, modelos são treinados com diversas imagens, para então serem aplicados em imagens reais buscando classificá-las, identificando rostos, tumores, ou o que for necessário [2].

O uso de inteligência artificial na indústria médica também tem crescido muito nos últimos anos como forma de auxílio a diagnósticos e otimização de processos. No ano de 2023 o tamanho do mercado de aplicações de inteligência artificial em imageamento por ressonância magnética chegou ao valor de 5,9 bilhões de dólares, e a previsão é que esse valor alcance 16,65 bilhões em 2034, como mostrado na Figura 1.3 [32].

O uso de processamento de imagens através de inteligência artificial para reconstrução de imagens de ressonância ainda está sendo refinado, as pesquisas estão aumentando e

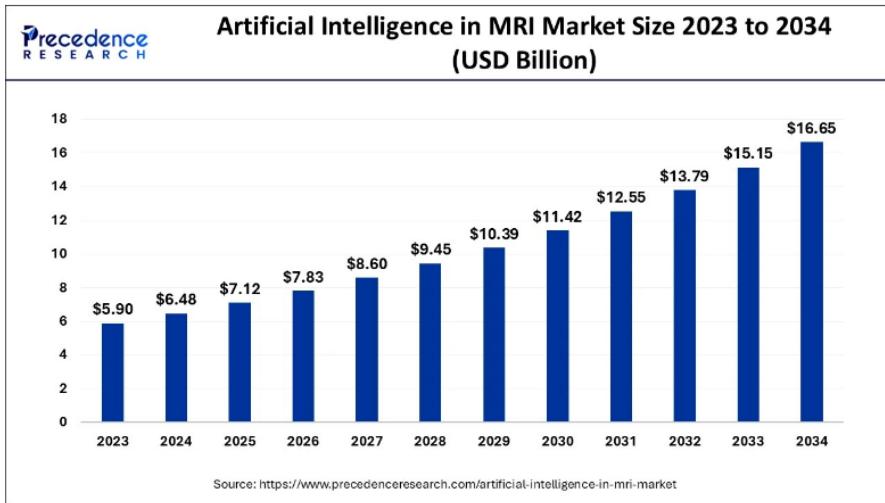


Figura 1.3. Valor de mercado da inteligência artificial no imageamento por ressonância magnética[32].

iniciativas como as competições realizadas pelo grupo FastMRI auxiliam na obtenção de resultados cada vez melhores [42].

A reconstrução da imagem utilizando inteligência artificial busca realçar informações importantes através do processamento da imagem. O algoritmo realiza essa etapa em várias imagens diferentes, através de tentativa e erro, melhorando a cada teste o resultado final. Ao final da análise de todas as imagens, o modelo terá criado regras próprias sobre o valor de cada informação e o que deve ou não analisar, o processo é realizado novamente, só que agora com parâmetros iniciais mais refinados do que a última vez, cada uma dessas análises, chamada de época ou geração, aumenta gradativamente a qualidade da imagem produzida pelo algoritmo. Após realizado todo o treino, o modelo é testado com imagens novas para verificar o seu desempenho e assim chegar em um resultado final [24].

Um dos problemas enfrentados ao treinar modelos de inteligência artificial como GANs é a necessidade de grande quantidade de dados para o modelo obter melhores resultados. Como imagens médicas são recursos limitados são utilizadas formas para expandir o conjunto de dados. Algumas dessas formas são rotacionar as imagens, cortar trechos das imagens gerando novas entre outras. Esse tipo de procedimento é chamado de *data augmentation* e é muito usado na criação de modelos de inteligência artificial [21].

Treinar modelos de inteligência artificial capazes de perceber padrões em imagens criadas a partir de espaços-k sub-amostrados e aprimorá-las tem sido um grande desafio, porém tem gerado bons resultados que podem mudar o futuro do imageamento por ressonância magnética, com exames velozes e com alta qualidade.

1.2 PROPOSTA DE PESQUISA

Redes Generativas Adversariais (GANs) são um dos modelos de inteligência artificial que mais estão sendo estudados para o processamento de imagens. Ao buscar pelo termo “Generative Adversarial Networks” no Google Trends, é possível observar o interesse no assunto ao longo do tempo baseado no número de pesquisas realizadas no Google, como mostrado no gráfico da Figura 1.4. O número de buscas nos últimos anos aumentou e isso se reflete nos artigos publicados, realizando uma busca na base do CAPES por artigos contendo o termo ““Generative Adversarial Networks” or “GAN”” entre os anos de 2016 e 2019 obtém-se 2.307 resultados, ao realizar a mesma busca porém entre os anos de 2020 e 2023 obtém-se 10.641 resultados, o que mostra que a ferramenta está se mostrando promissora e que os estudos na área estão aumentando.



Figura 1.4. Resultados de buscas relacionadas a GANs nos últimos dez anos segundo o Google Trends.

O modelo de uma GAN é separado em duas redes neurais distintas, uma geradora que cria imagens para simular imagens de um conjunto de dados, ou seja, falsas imagens de ressonância no contexto do projeto, e uma rede neural discriminadora que avalia se as imagens criadas pertencem ao conjunto de dados original ou se foram criadas. Caso a imagem criada seja apontada como falsa, a primeira rede neural faz ajustes e repete o processo, dessa forma as duas redes neurais são aperfeiçoadas com o tempo, a primeira em criar imagens cada vez mais semelhantes com as do conjunto de dados, e a segunda cada vez mais refinada em apontar imagens criadas [29]. Modelos desse tipo podem se mostrar capazes de melhorar a qualidade de imagens de ressonância magnética sub-amostradas através do seu processo de aprimorar imagens até que essas se confundam com um grupo controle, que no caso seria de imagens reconstruídas a partir de espaços-k tipicamente amostrados [8].

Redes adversárias generativas poderiam então ser utilizadas na reconstrução de imagens de ressonância magnética no intuito de melhorar as métricas de qualidade final da imagem. A GAN seria utilizada após a reconstrução da imagem feita com dados

sub-amostrados, alimentando o modelo com essas imagens distorcidas com o objetivo de produzir imagens com métricas de qualidade superior às imagens de entrada. Dessa forma podem ser testados diferentes tipos e graus de sub-amostragem procurando aquele em que a GAN gera melhores resultados [8].

Uma das limitações para o uso de GANs é a necessidade de uma grande quantidade de imagens de treinamento, isso facilita que o modelo generalize melhor os dados e não aprenda a melhorar apenas as imagens de treino. Como imagens médicas não são de fácil acesso e por isso a quantidade disponível é limitada, foram aplicadas técnicas de aumento de dados (*data augmentation*) que possibilitam a extração de mais informações do conjunto de dados criando novas imagens a partir deles [21].

Outra limitação é a necessidade de alta capacidade computacional. Computadores necessitam de unidades gráficas fortes para poder realizar o processamento das imagens, porém algumas ferramentas online auxiliam nesse ponto, como o uso do ambiente de programação do Google. O Google Colab possui alto poder computacional de forma gratuita para pesquisa através da GPU Tesla T4 e também a possibilidade de máquinas com maior capacidade através de pagamento de mensalidade. Para esse trabalho foi usada a GPU L4, que requer pagamento de mensalidade, com 22,5 GB e uma memória RAM de 53 GB [16].

1.3 OBJETIVOS

1.3.1 Objetivo Geral

Reconstruir imagens de ressonância magnética a partir de dados sub-amostrados de bobina única utilizando GANs e comparar a qualidade das imagens geradas por meio das métricas SSIM e SNR para diferentes taxas de amostragem e trajetórias.

1.3.2 Objetivos Específicos

- Criar conjuntos de treino e teste com imagens feitas com dados com diferentes parâmetros de sub-amostragem;
- Desenvolver um algoritmo para extrair as métricas de qualidade média(SSIM e SNR) das imagens reconstruídas;
- Ajustar a hiperparametrização da GAN para conjuntos de teste com diferentes trajetórias de sub-amostragem;
- Analisar e apresentar os resultados da pesquisa com base na avaliação experimental.

2 FUNDAMENTAÇÃO TEÓRICA

Para embasar o projeto, foi realizada uma pesquisa teórica sobre a aquisição de imagens de ressonância magnética, para entender o processo de como os dados são obtidos e transformados em imagens.

Em seguida foi realizada um levantamento de literatura científica para criar um acervo de artigos que abordam o que está sendo estudado atualmente no tema, esses artigos formam parte do referencial teórico da pesquisa.

Foi realizado também um estudo sobre GANs, como funcionam e como é o processo de treinamento e produção de imagens, e como essa ferramenta poderia auxiliar na reconstrução e aprimoramento de imagens de ressonância magnética.

2.1 AQUISIÇÃO DE DADOS E RECONSTRUÇÃO DA IMAGEM

A imagem obtida no exame de ressonância magnética é o resultado final de uma série de procedimentos que envolvem aplicação de um campo magnético intenso, a emissão de pulsos de radiofrequência em uma determinada região de estudo do paciente, a coleta de dados e, por fim, a reconstrução da imagem. Os procedimentos serão abordados em detalhes nas seções a seguir.

2.1.1 Campo Magnético Principal e Vetor Magnetização

O equipamento de ressonância magnética consiste em um tubo onde uma maca é posicionada no interior, este tubo possui várias bobinas ocultas dispostas em volta do corpo do paciente que tanto produzem quanto captam campos magnéticos [35].

Com o paciente deitado dentro do equipamento, é aplicado primeiramente um campo magnético gradiente ao longo do corpo do paciente, chamado de direção z, dessa forma o campo magnético é menos intenso no pé do paciente e mais intenso na cabeça [27].

Os spins dos átomos de hidrogênio presentes no corpo do paciente tendem então a alinhar seu próprio campo magnético com o campo magnético gradiente imposto realizando um movimento de precessão. O movimento se caracteriza pela movimentação circular de

um eixo formando um cone. Devido ao campo magnético gradiente os diferentes spins irão precessar em frequências diferentes como descrito pela equação de Larmor (Equação 2.1) , e é através dessa diferença na frequência de cada parte do corpo que é feita a seleção de corte, pois os sinais emitidos por cada parte do corpo vão possuir diferentes frequências [27].

$$\omega = \gamma B_0, \quad (2.1)$$

Na fórmula, ω é a frequência de precessão, B_0 é a intensidade do campo magnético no corte e γ é a razão giromagnética, que para o hidrogênio é de 42,58 MHz/T. Dessa forma o campo magnético gradiente também é chamado de campo magnético principal, ou de gradiente de seleção de corte [27].

Os átomos imersos então no gradiente de seleção de corte tendem a se alinhar paralelamente (direção z) ou antiparalelamente (direção -z) ao campo principal enquanto realizam o movimento de precessão. Porém a maior parte desses átomos se alinham paralelamente gerando assim um campo magnético resultante chamado de vetor magnetização resultante (M_0) que se alinha então ao eixo z, como mostrado na Figura 2.1. [27].

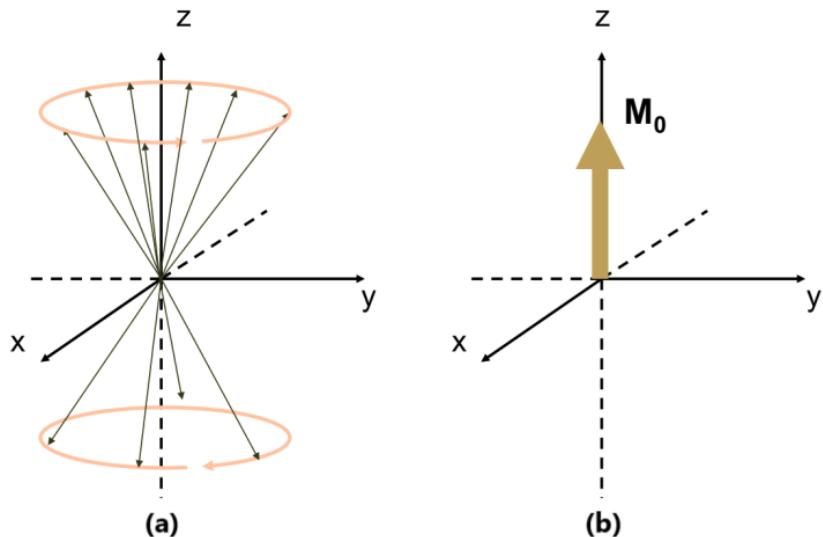


Figura 2.1. (a) Vetores dos campos magnéticos dos átomos de hidrogênio se alinhando paralelamente e antiparalelamente à direção z. (b) Vetor magnetização resultante criado pelo movimento dos átomos de hidrogênio. Imagem retirada de [27] e modificada.

Outros átomos poderiam ser escolhidos para realização do exame, porém o hidrogênio se mostra a melhor escolha por ser o terceiro elemento mais presente no corpo humano, os átomos de hidrogênio possuem alto momento magnético tornando-os mais sensíveis a campos magnéticos externos e moléculas que contém hidrogênio diferem bastante entre tecidos saudáveis e tecidos com alguma patologia [27].

Com os átomos alinhados ao campo magnético principal o exame avança para a etapa de aplicação de pulsos de radiofrequência, nessa etapa são obtidos os dados do corpo do paciente que serão usados na reconstrução da imagem, o procedimento é explicado em detalhes na próxima seção.

2.1.2 Pulso de Radiofrequência

Para obter os sinais, bobinas são posicionadas em volta do paciente. Essas bobinas captam os sinais de radiofrequência induzidos pela precessão dos núcleos de hidrogênio. Elas são posicionadas de forma a captar sinais perpendiculares ao campo magnético principal, não sendo capazes de captar os sinais gerados pelos átomos de hidrogênio que estão paralelos a ele [35].

Para as bobinas em volta do corpo do paciente captarem o sinal é necessário então desviar o vetor magnetização resultante gerado pelos núcleos de hidrogênio de forma perpendicular ao campo magnético principal, para isso são usados pulsos de radiofrequência.

O pulso de radiofrequência (RF) é um campo magnético com sinal perpendicular ao campo magnético principal ele é ligado e desligado logo em seguida, por isso denominado como pulso. Quando o pulso é aplicado todos os núcleos de hidrogênio que estavam paralelos ao campo principal se alinham a ele, assim as bobinas laterais conseguem detectar a magnetização transversal gerada por esses núcleos [27].

Ao desligar o pulso de RF, os átomos de hidrogênio tendem a relaxar e se alinhar novamente com o campo magnético principal gerando novamente uma magnetização longitudinal, porém esse alinhamento não acontece de forma instantânea, átomos em diferentes moléculas terão tempos diferentes de alinhamento. São então estipuladas duas variáveis para agrupar esse tempos de relaxamento, a constante T1 indica o tempo necessário para a magnetização longitudinal recuperar 63% da sua intensidade após o desligamento do pulso de RF, e o tempo T2 indica o tempo necessário para a magnetização transversal atingir 37% da intensidade original do vetor magnetização após o desligamento do pulso de RF [23]. Os tempos T1 e T2 variam de acordo com o tecido do corpo humano como pode ser visto na Tabela 2.1 e é essa diferença que irá gerar diferenças de contraste na imagem final.

São emitidos diversos pulsos de RF ao longo do exame, e a cada repetição de pulso são coletados os sinais vindos do corpo. Um exame de ressonância pode ter várias repetições dessas etapas, de acordo com o tipo de exame e qualidade final desejada para a imagem [13]. Na próxima seção será detalhado o processo de sequência de pulsos para obtenção dos dados.

Tabela 2.1. Tempos T1 e T2 para diferentes tecidos do corpo humano para um exame realizado com campo magnético de 3T [36].

Tecido	T1 [ms]	T2 [ms]
Sangue	1932 \pm 85	275 \pm 50
Massa Cinzenta	1820 \pm 114	99 \pm 7
Matéria Branca	1084 \pm 45	69 \pm 3
Nervo Óptico	1083 \pm 39	78 \pm 5

2.1.3 Codificação de Frequência e de Fase

O sinal emitido pelos núcleos de hidrogênio é detectado pelas bobinas locais. Esse sinal é composto de sinais emitidos por cada um dos núcleos de hidrogênio do corpo do paciente, para selecionar apenas o corte desejado o sinal é filtrado pela frequência de precessão dos spins no corte desejado, como explicado anteriormente [27].

Selecionando o corte é possível captar apenas os sinais emitidos naquela área do corpo. O sinal obtido é a soma da amplitude dos sinais de todos os núcleos de hidrogênio naquele corte, porém esses dados estão desorganizados e para isso foi criada a codificação de fase e codificação de frequência para mapeá-los [27].

A codificação de frequência e de fase mapeiam o corte nos eixos horizontal e vertical respectivamente. A cada etapa do exame é aplicado pulsos com frequência e fase específicos e então medido a amplitude do sinal, essa amplitude então é armazenada em uma matriz onde o eixo horizontal é a frequência e o vertical é a fase dos sinais. O processo é repetido diversas vezes, e as diferentes configurações de frequência e fase são os pares ordenados dessa matriz [27].

A matriz onde os sinais de amplitude são armazenados é chamada de espaço-k. Cada ponto dessa matriz é a amplitude do sinal em determinado momento do exame e por se tratar de uma matriz que está em função de frequência e fase podemos aplicar uma transformada de Fourier de duas dimensões e obter uma nova matriz no domínio espacial, e essa nova matriz é a imagem do corte transversal [27].

2.1.4 Espaço-k

O espaço-k é uma matriz que contém os sinais espaciais obtidos após as diversas sequências de pulsos realizadas durante o exame de ressonância magnética codificado em frequência e fase. Cada ponto contém a amplitude do sinal obtido em determinado momento do exame e dessa forma cada um desses pontos contém informação de todo o corte, não havendo relação de um ponto no espaço-k para um único ponto da imagem final. Uma representação do espaço-k pode ser observada na Figura 2.2 ao lado da image

final produzida [27].

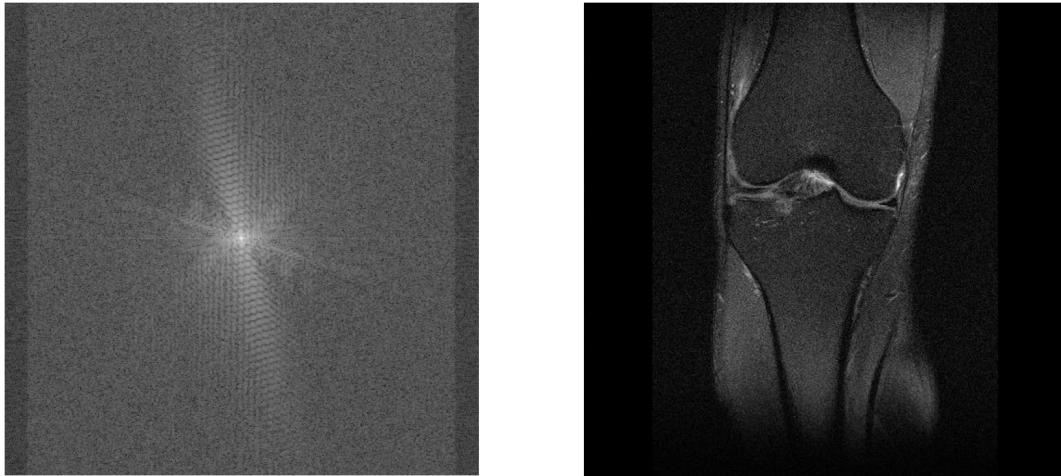


Figura 2.2. Espaço-k. A matriz é composta de tons de cinza, onde os pontos mais claros representam sinais com maior amplitude, e os mais escuros com menor.

É importante lembrar que devido o espaço-k ser em função de frequências e fases ele se trata de uma matriz complexa, possuindo parte real e imaginária, dessa forma, não existe uma representação visível. Para representar o espaço-k é calculado o absoluto do mesmo, podendo assim representar a soma da parte real e imaginária.

Existem várias formas de preencher o espaço-k, essas formas são chamadas de trajetórias e são escolhidas de acordo com o tipo de exame, cada uma tendo seus pontos positivos e negativos. Na Figura 2.3 é possível visualizar algumas dessas trajetórias. Também é possível observar na figura, em sua segunda linha de imagens, como seria a sub-amostragem do espaço-k com cada uma das trajetórias, e na terceira linha, a imagem final produzida por cada um desses espaços-k [10].

A figura contém níveis diferentes de sub-amostragem para cada uma das trajetórias, e por isso as imagens finais possuem diferentes níveis de nitidez. Para a trajetória cartesiana foi apresentado um nível de sub-amostragem de 42,19%, para a radial uma sub-amostragem de 88,65% e para a trajetória espiral uma sub-amostragem de 62,04%.

É interessante notar que diferentes trajetórias vão possuir diferentes impactos na imagem. Por exemplo, a trajetória radial que prioriza manter as baixas frequências do espaço-k, correspondendo ao centro do mesmo, possui um melhor contraste da imagem final, pois são as baixas frequências que impactam nesse contraste[28]. Porém a trajetória espiral que poupa mais as altas frequências em relação a trajetória radial, possui uma maior nitidez nas transições da imagem, possuindo linhas delimitadoras mais nítidas, pois as altas frequências são responsáveis pela noção espacial da imagem final [27].

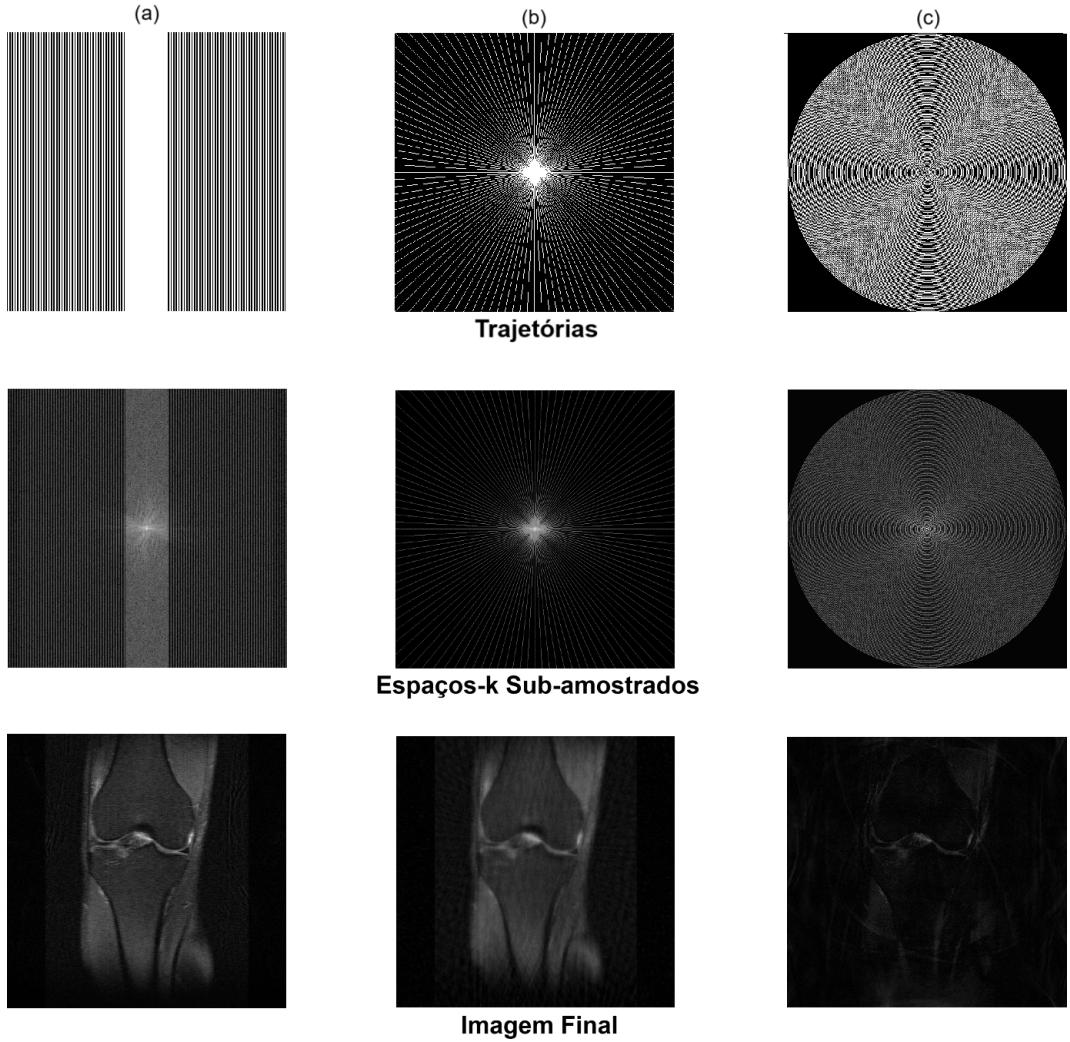


Figura 2.3. Coluna (a): Trajetória cartesiana vertical. Coluna (b): Trajetória radial. Coluna (c): Trajetória espiral.

2.1.5 Sub-Amostragem

A aquisição de sinais em um exame de ressonância magnética deve seguir o Teorema de amostragem de Nyquist–Shannon, que diz que a frequência de amostragem de um sinal deve ser pelo menos o dobro da maior frequência contida naquele sinal. Caso esse critério não seja respeitado, a imagem final irá apresentar “aliasing”, esse fenômeno cria distorções na imagem que afetam a sua qualidade final, como apresentado na Figura 2.4.

Formas de reconstruir imagens com menos amostras são estudadas a vários anos, uma das mais promissoras dessas formas tem sido o uso de *Compressed Sensing*, que aproveita da esparsidade da imagem para reconstruí-la a partir de um espaço-k com um número de amostras abaixo do mínimo segundo o critério de Nyquist–Shannon[14].

Outra forma que tem sido estudada nos anos recentes é o uso de inteligência artificial para a reconstrução de imagens sub-amostradas. Treinar redes neurais para obtenção

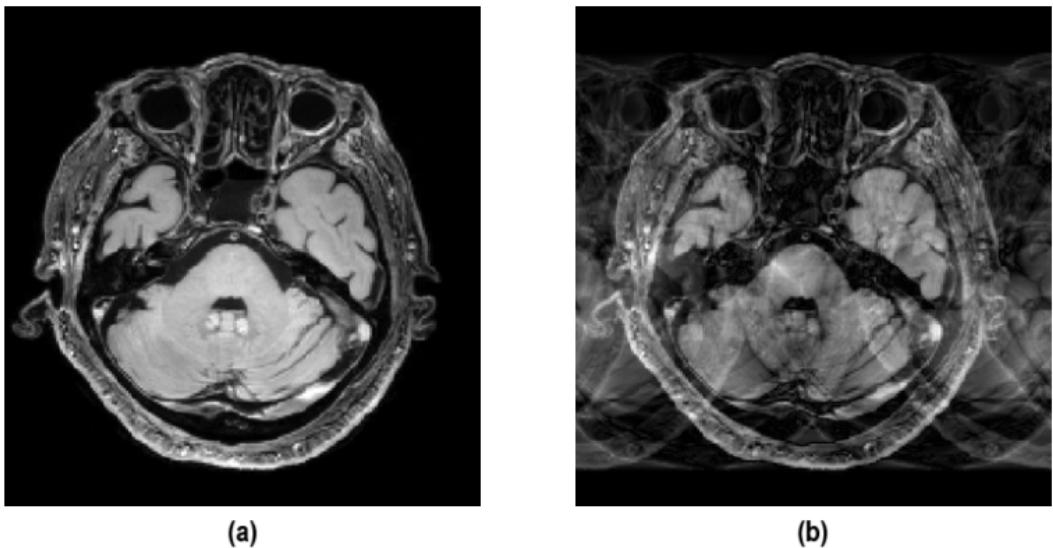


Figura 2.4. Comparação entre imagem reconstruída respeitando critério de Nyquist (a) e imagem feita com dados sub-amostrados apresentando aliasing (b). Imagem retirada de [39] e modificada.

de imagens nítidas tem obtido bons resultados e o número de artigos publicados nessa área tem crescido nos últimos anos conforme verificado no portal da CAPES que será mostrado na seção de levantamento da literatura científica.

2.2 REDES GENERATIVAS ADVERSARIAIS (GANs)

Redes generativas adversárias (do inglês GANs) são um modelo de rede neural de aprendizado profundo que tem como objetivo criar amostras o mais fiéis possíveis a um conjunto de treinamento. Essas amostras podem ser de diversos tipos como por exemplo imagens ou sinais de áudio. No contexto do trabalho de mestrado são criadas imagens utilizando a GAN.

O modelo é separado em duas partes, uma rede neural generativa e outra discriminativa. A parte generativa irá criar imagens que se assemelham a imagens de um conjunto de treinamento, já a parte discriminativa irá julgar se a imagem criada pertence ou não ao conjunto de treinamento [15]. A Figura 2.5 apresenta o fluxo seguido por um modelo GAN genérico em uma etapa, essa etapa é repetida várias vezes aperfeiçoando as imagens geradas.

2.2.1 Modelo Generativo

O modelo generativo tipicamente recebe como entrada diferentes tipos de sinais, como por exemplo outras imagens ou até mesmo ruído, e usa essa informação de entrada como

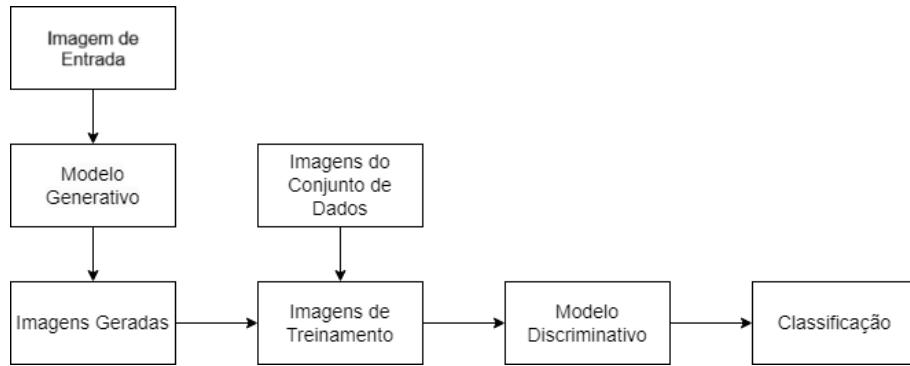


Figura 2.5. Fluxo de produção e processamento de imagens em um modelo GAN.

parâmetros iniciais para gerar imagens, o objetivo do modelo gerativo é enganar o modelo discriminativo, de forma que ele confunda as imagens geradas com as imagens pertencentes ao conjunto de dados. O modelo gerativo não tem acesso às imagens do conjunto de dados para que essas imagens não influenciem nas imagens que serão produzidas por ele e sua única forma de aprendizado é com os veredictos recebidos do discriminador a cada iteração [6]. Nas primeiras etapas o modelo gerativo não conseguirá produzir imagens muito semelhantes com o conjunto de dados, porém a cada nova iteração, o modelo irá se refinar baseando-se nos veredictos do discriminador.

2.2.2 Modelo Discriminativo

O modelo discriminativo é treinado com imagens vindas do conjunto de dados e com imagens produzidas pelo modelo gerativo misturadas e escolhidas de forma aleatória. Ele deve classificar as imagens que receber indicando se elas pertencem ou não ao conjunto de dados. A saída do discriminador é um valor entre 0 e 1, se a imagem analisada for do conjunto de dados ela deve receber um valor próximo a 1, caso contrário um valor próximo de 0. A cada etapa ele recebe o gabarito com a resposta certa da classificação e faz os ajustes necessários para na próxima etapa aumentar a taxa de acerto [6].

2.3 LEVANTAMENTO DA LITERATURA CIENTÍFICA

No intuito de saber o andamento das pesquisas na área foi realizada uma levantamento de artigos na base de dados da CAPES baseada na seguinte pergunta científica:

“É possível reconstruir imagens de ressonância magnética sub-amostradas com nitidez com o auxílio de inteligência artificial?”

Foram então escolhidos os seguintes descritores para as buscas:

- MRI;

- Reconstruction;
- Undersampled;
- Artificial Intelligence.

A busca foi realizada na base de dados da CAPES usando os conectores “and” para filtrar os artigos que possuíssem os quatro temas. A busca então foi realizada da seguinte forma:

““MRI” and “Reconstruction” and “Undersampled” and “Artificial Intelligence””

A Figura 2.6, contém o fluxograma da filtragem realizada para chegar em 48 artigos que auxiliaram na realização do trabalho. Esses artigos contém diferentes abordagens no uso de inteligência artificial aplicada a reconstrução de imagens de ressonância magnética.

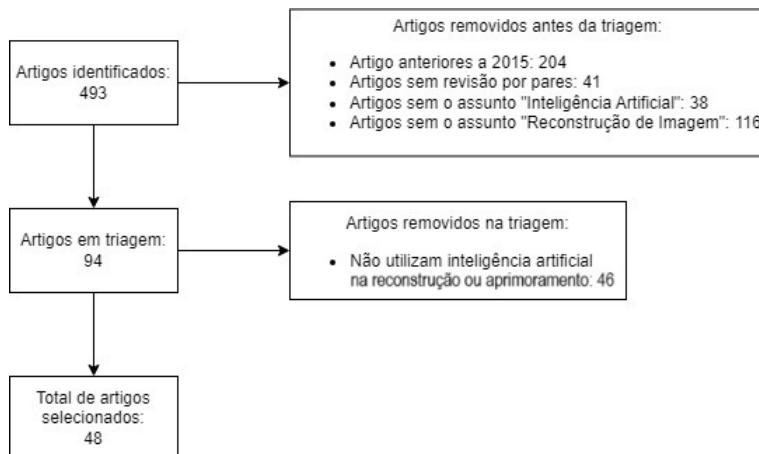


Figura 2.6. Fluxograma de triagem dos artigos selecionados.

Os 48 artigos finais serviram como base teórica para a realização do projeto e auxiliaram na fundamentação teórica e seleção do modelo escolhido para a reconstrução das imagens de ressonância magnética.

2.3.1 Trabalhos Correlatos

Os artigos selecionados possuem diferentes abordagens no uso da inteligência artificial na reconstrução de imagens de ressonância magnética, como uso de CNNs [19] [5], autoencoders [9] [26], dictionary learning [40] [1] entre diversas outras formas de utilizar inteligência artificial na reconstrução de imagens de ressonância magnética. Entretanto, entre os 48 trabalhos filtrados no levantamento realizado no portal de periódicos da CAPES apenas um utilizou de GANs para realizar a tarefa de reconstruir imagens de ressonância a partir de dados sub-amostrados.

O título do trabalho é “Prior-Guided Image Reconstruction for Accelerated Multi-Contrast MRI via Generative Adversarial Networks” [8] e tem como objetivo a melhora na qualidade de imagens de ressonância magnética multi-contraste feitas a partir de dados sub-amostrados. Para isso ele alimenta a rede generativa com imagens feitas a partir de dados sub-amostrados, porém diferente do trabalho de mestrado ele usa diferentes técnicas de sub-amostragem em um único treinamento da GAN, combinando informações de diferentes tipos de dados para obtenção de melhores resultados.

O trabalho apresenta uma técnica que utiliza três tipos de dados de entrada para o treinamento: dados com menos amostras nas baixas frequências, dados com menos amostras nas altas frequências e dados perceptuais, obtidos através de redes neurais classificadoras que evidenciam as características da imagem.

O artigo apresenta resultados onde foi atingido SSIM de até 92% usando apenas 2% das amostras necessárias para construir uma imagem típica de ressonância magnética, mostrando o poder das redes generativas[8].

O trabalho em questão, entretanto, necessita de técnicas específicas para obtenção do espaço-k e um dos canais precisa ser pré-processado com uma rede neural classificadora antes de ser utilizado pela GAN. Esse pré-processamento aumenta o tempo de reconstrução das imagens e não é o padrão usado em máquinas de ressonância. Dessa forma para este método ser implementado seria necessário alteração na forma que obtemos as imagens.

O código criado durante o mestrado difere do trabalho em questão pois realiza o processamento sem tratamento prévio, utilizando imagens criadas com dados da forma que eles são gerados por máquinas de ressonância magnética utilizadas hoje na indústria, sendo necessários menos etapas de pré-processamento e aceitando uma maior diversidade de dados para treinamento

Outro artigo que usa um processamento interessante é o “Deep learning for undersampled MRI reconstruction” [20]. Apesar de não usar GAN na reconstrução das imagens, no trabalho é usado um modelo U-Net para processamento dos dados, o mesmo usado neste trabalho na rede generativa.

O trabalho em questão usa a rede generativa para extrair os detalhes da imagem feita com um espaço-k sub-amostrado e depois realçar esses detalhes. Ao final dessa etapa é calculado novamente o espaço-k que é usado para completar os espaços faltantes do espaço-k original, e por fim obter a imagem final a partir desse novo espaço-k.

O estudo consegue atingir melhorias de 0,22 no SSIM da imagem de entrada se comparada com a imagem na saída da U-Net e uma melhora de aproximadamente mais 0,3 com a junção dos espaços-k.

Esse estudo incentivou o teste com a união do espaço-k antes e após o processamento para observar se ocorreriam mais melhorias na imagem.

Outro estudo que foi muito importante como base deste trabalho é o “fastMRI: An Open Dataset and Benchmarks for Accelerated MRI”[42]. O trabalho apresenta os detalhes de diversos conjuntos de dados para desafios de reconstrução de imagens de ressonância magnética, entre eles o conjunto de joelho de bobina única utilizado neste trabalho.

O estudo do FastMRI também propõe a criação de uma tabela de classificação para organizar os algoritmos de reconstrução criados para o desafio. A página da internet com a tabela original não está mais disponível online, porém a página foi arquivada em um repositório online[12].

3 MATERIAIS E MÉTODOS

Serão apresentados nas seções abaixo os materiais e métodos utilizados na realização do trabalho. Serão abordados os temas: Banco de dados, pré-processamento, modelo GAN e métricas de avaliação.

3.1 BANCO DE DADOS

As imagens utilizadas no estudo foram obtidas do banco de dados de espaços-k do grupo de pesquisa FastMRI, este grupo incentiva pesquisas e competições na área de reconstrução e aprimoramento de imagens de ressonância magnética a partir de dados sub-amostrados utilizando inteligência artificial. As competições buscam melhorar cada vez mais o índice de similaridade das imagens reconstruídas e a relação sinal ruído e são realizadas em diversas categorias. O banco de dados é criado especificamente para que pesquisadores possam trabalhar com dados brutos de imagens de ressonância de forma gratuita, dessa forma todas as imagens contidas nele são anonimizadas [18].

O conjunto de dados utilizado será o mesmo que foi utilizado no desafio de reconstrução de imagens de joelho de bobina única (knee singlecoil challenge). Os conjuntos de imagens de uma única bobina não foram diretamente obtidos por uma única bobina, são simulados, obtidos através da combinação de amostras de múltiplas bobinas [42]. A vantagem de se realizar testes com amostras de uma única bobina é poder realizar testes com computadores com menor capacidade de processamento, dessa forma podendo-se obter resultados relevantes para pesquisas na área realizando ensaios que geram resultados mais rapidamente, podendo assim refinar esses resultados de forma mais eficiente realizando várias vezes os treinamentos alterando os seus parâmetros.

O conjunto de imagens utilizado é separado em quatro partes, cada parte é composta de arquivos do tipo .h5, que contém um conjunto de espaços-k, esses conjuntos contém cortes de um mesmo exame realizado em um joelho, formando um volume. A primeira parte é chamada de imagens de treino, contém 973 volumes contendo 34.723 cortes que foram tipicamente amostrados. A segunda é chamada de imagens de validação e contém 199 volumes e 7.135 cortes também tipicamente amostrados. A terceira parte é chamada de imagens de teste e contém 108 volumes e 3.903 cortes sub-amostrados. O quarto e

último é chamado de conjunto desafio, foi utilizado como parâmetro no desafio realizado pela equipe do FastMRI e contém 92 volumes e 3.305 cortes sub-amostrados. As informações acerca do conjunto de dados estão reunidas na Tabela 3.1.

Tabela 3.1. Organização do conjunto de dados [42].

	Volumes	Cortes	Amostragem
Treinamento	973	34.742	Tipicamente amostrado
Validação	199	7.135	Tipicamente amostrado
Teste	108	3.903	Sub-amostrado
Desafio	92	3.305	Sub-amostrado

Para o projeto foram usadas apenas as imagens do conjunto de treinamento, pois por estarem tipicamente amostradas foi possível sub-amostrar as imagens conforme fosse necessário, usando diferentes níveis de sub-amostragem e diferentes trajetórias como será apresentado na seção de pré-processamento.

Para expandir a quantidade de dados utilizados foram extraídos mais de um corte por volume, uma vez que cada volume possui em torno 35 cortes. Porém não é todo corte que se mostra útil, uma vez que as extremidades do volume possuem pouca informação, assim geralmente são extraídos apenas os cortes do meio.

Foram feitos treinamentos de 973 imagens até 6000 imagens, os conjuntos de teste geralmente contém 10% do tamanho dos conjuntos de treinamento e os conjuntos de teste variam de 50 a 100 imagens.

3.2 PRÉ-PROCESSAMENTO

O pré-processamento consiste em uma sequência de etapas para preparar as imagens do conjunto de dados para o processamento que será realizado pela GAN. As imagens utilizadas do conjunto de dados são feitas a partir de espaços-k tipicamente amostrados, dessa forma é necessário simular a sub-amostragem desse espaço-k. Para isso são apagadas algumas amostras da matriz de acordo com a trajetória escolhida e depois reconstruída a imagem, como mostrado na Figura 2.3. Esse processo de apagar amostras é comumente usado em estudos desse tipo pois simula de forma fiel a sub-amostragem gerada por equipamentos de ressonância [42]. Os detalhes do funcionamento do algoritmo que realiza o pré-processamento serão apresentados na seção de Resultados e Discussão.

3.3 MODELO GAN

O código usado é baseado no modelo GAN adaptado para o processamento de imagens. No código feito a entrada do sistema é uma imagem e a saída outra imagem, esse tipo de arquitetura que traduz uma imagem em outra é chamado de pix2pix.

Assim como filtros de redes sociais alteram imagens colocando filtros de cores, redes pix2pix recebem uma imagem de entrada e de acordo com o treinamento feito geram uma outra imagem baseada na entrada. Esse tipo de processamento é realizado no trabalho de forma que as imagens de entrada são as imagens de ressonância magnética feitas com dados sub-amostrados e a saídas são imagens com suas métricas de qualidade aprimoradas.

Phillip Isola apresenta em seu artigo o pix2pix [22] e estabelece métodos para melhor realizar essa tradução de imagem para imagem o que auxiliou na criação do algoritmo GAN que processou as imagens de ressonância. Os detalhes do funcionamento da GAN, assim como das suas redes generativa e discriminativa serão apresentados na seção de Resultados e Discussão.

3.4 MÉTRICAS DE AVALIAÇÃO

As imagens geradas pelo modelo devem ser avaliadas para verificar se o resultado obtido se aproxima de uma imagem feita com dados tipicamente amostrados. Para isso são usados dois valores que indicam a qualidade da imagem: o índice de similaridade estrutural (SSIM) e a relação sinal ruído (SNR).

O índice de similaridade estrutural indica a semelhança entre duas imagens e é dado por

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \quad (3.1)$$

onde uma imagem é dada como padrão e uma segunda recebe um valor de 0% a 100% de semelhança. A imagem usada como padrão é a imagem feita com dados tipicamente amostrados, chamada de imagem x, e a imagem gerada, chamada de imagem y, é a imagem que será comparada. Na equação μ_x e μ_y são a intensidade média dos pixels em cada uma das imagens, σ_x e σ_y são os desvios padrão de cada imagem, σ_{xy} é a covariância entre as duas imagens, C_1 e C_2 são duas constantes para equilibrar a divisão, evitando divisões por zero [3].

A relação sinal ruído apresenta a amplitude de sinal em relação a amplitude de ruído contido na imagem e é dado por

$$\text{SNR} = 10 \log_{10} \left(\frac{\sigma_s^2}{\sigma_r^2} \right), \quad (3.2)$$

quanto maior o valor de SNR, maior a presença de sinal em relação ao ruído. Na fórmula σ_s^2 é o desvio padrão ao quadrado dos valores de amplitude contidos dentro de uma área de interesse representada com um S na Figura 3.1, considerado como sinal, e σ_r^2 é o desvio padrão ao quadrado dos valores de amplitude das áreas representadas com um N na figura, considerados como ruído.

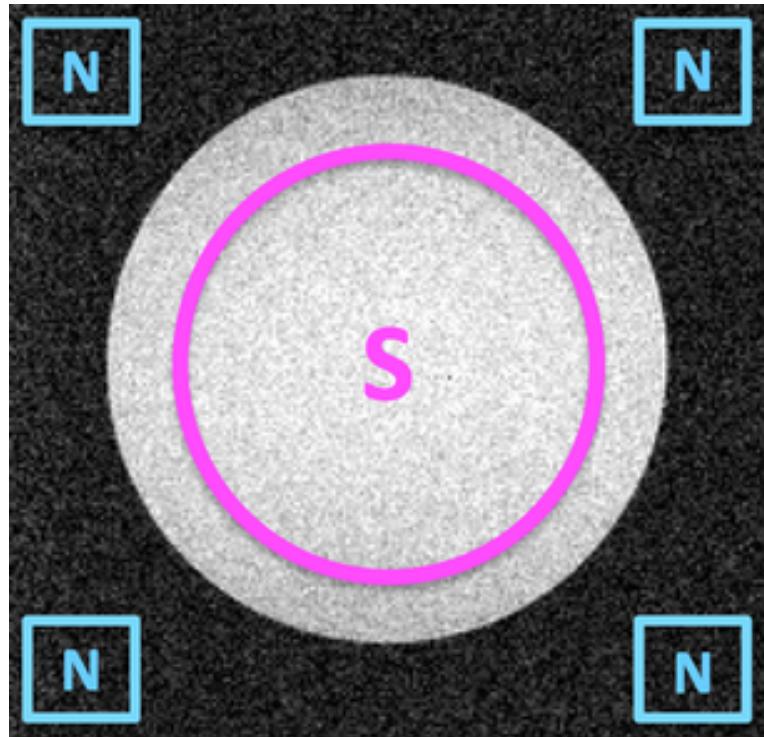


Figura 3.1. Áreas de interesse para o cálculo da relação sinal ruído (SNR). A área assinalada com o S indica a área que contém o sinal de interesse, a área assinalada com o N indica onde se concentra o ruído isolado. Imagem retirada de [11].

É utilizado um algoritmo que calcula essas duas métricas, o algoritmo recebe um número qualquer de imagens, de acordo com a necessidade do projeto, calcula para cada uma dessas imagens o valor de SSIM e SNR e depois calcula a média desses valores para cada imagem. Dessa forma um grupo de imagens gerados juntos recebe apenas dois valores, um de SNR e um de SSIM, esses valores então são usados para avaliar o modelo.

Foram geradas tabelas com os resultados de cada teste feito explicitando os detalhes dos parâmetros utilizados, como tamanho da imagem de entrada, trajetória utilizada, número de épocas, entre outros parâmetros, e também é inserido os resultados na forma de SSIM e SNR.

Os resultados também foram avaliados em função de duas variáveis que ajudam a comparar a melhor no SSIM e no SNR, chamadas de Razão dos SSIM (RSSIM) e Razão

dos SNR (RSNR). Essas variáveis são a razão entre as métricas das imagens produzidas pela GAN e as métricas das imagens de entrada,

$$\text{RSSIM} = \frac{SSIM_{pr}}{SSIM_{in}}, \quad (3.3)$$

$$\text{RSNR} = \frac{SNR_{pr}}{SNR_{in}}, \quad (3.4)$$

na fórmula, $SSIM_{pr}$ e SNR_{pr} são as respectivas métricas das imagens geradas, e $SSIM_{in}$ e SNR_{in} as métricas das imagens de entrada. Logo, caso os valores de RSSIM e RSNR sejam maiores do que 1, ocorreu melhora nas métricas das imagens. Através dessas duas variáveis é possível observar o quanto as métricas aumentaram ou diminuíram em relação às métricas da imagem de entrada.

4 RESULTADOS E DISCUSSÕES

Foram realizados 110 testes com diferentes arquiteturas do modelo da GAN e diferentes parâmetros de sub-amostragem para as imagens. Os parâmetros usados nesses testes como por exemplo trajetórias e nível de sub-amostragem podem ser observados na planilha Resultados.csv no repositório do Github do trabalho [7].

A arquitetura final e detalhamento tanto do código de pré-processamento quanto da GAN serão apresentadas nas seções a seguir. Também serão apresentados os detalhes das imagens produzidas e suas métricas de qualidade. Na discussão será abordada a análise dos dados que permitirão observar o limite encontrado para aprimoramento de imagens de ressonância magnética feitas com dados sub-amostrados.

4.1 RESULTADOS

4.1.1 Pré-Processamento

O algoritmo de pré-processamento possui diversas especificações que podem ser alteradas de acordo com a necessidade, essas especificações alteram os parâmetros da sub-amostragem, como taxa de amostragem, trajetórias e presença ou não de artefatos. No pré-processamento as imagens são modificadas para todas terem o mesmo tamanho e possuírem altura e largura iguais, essa tarefa é realizada sem redimensionar as imagens, ao invés disso, todas as imagens são completadas com zero até terem as mesmas dimensões da maior imagem. O modelo também dá a opção de fixar a dimensão das imagens, cortando as imagens maiores do que o valor, e completando com zero as menores. Essas dimensões são sempre arredondadas para potências de dois, pois no processamento da GAN são realizadas diversas divisões por dois, facilitando esse processo.

O algoritmo possui três tipos de trajetória: *HF*, *random*, *spiral1*, *spiral4* e *radial*. A trajetória do tipo *HF* preserva as colunas com baixas frequências do espaço-k e filtra as colunas com altas frequências em intervalos regulares escolhidos pelo usuário, essa trajetória recebe os parâmetros:

- *start_interval*: Indica o intervalo a partir do centro do espaço-k no qual as colunas

de frequências devem ser preservadas;

- *interval*: Indica o intervalo no qual as frequências acima do *start_interval* serão zeradas. Por exemplo, se *interval* for 3, a cada três colunas a terceira será zerada.

A trajetória *random* filtra aleatoriamente as colunas, de acordo com o valor da variável “*mask_percentage*” que recebe um valor de 0 a 1. Se o valor for 1 nenhuma amostra é removida, a taxa de sub-amostragem é igual a zero e a imagem é mantida inteira, porém se o valor for 0, todas as amostras são removidas e a taxa de sub-amostragem é de 100%

A trajetória *spiral1* e *spiral4* filtram o espaço-k em um formato espiral. A trajetória *spiral1* consiste de apenas uma espiral e a *spiral4* consiste em quatro espirais. Ambas as trajetórias recebem o parâmetro “*turns*”, que indica quantas voltas a espiral deve realizar.

A trajetória *radial* filtra o espaço-k com linhas que atravessam o centro do espaço-k. A função recebe o parâmetro “*number_angles*” que indica quantas linhas serão traçadas, essas linhas são sempre traçadas de forma que o ângulo entre elas sejam iguais.

Na Figura 4.1 é possível observar como são as trajetórias e como a mudança de parâmetros altera seu formato. Foi dado zoom na trajetória *HF* para permitir melhor visualização das colunas.

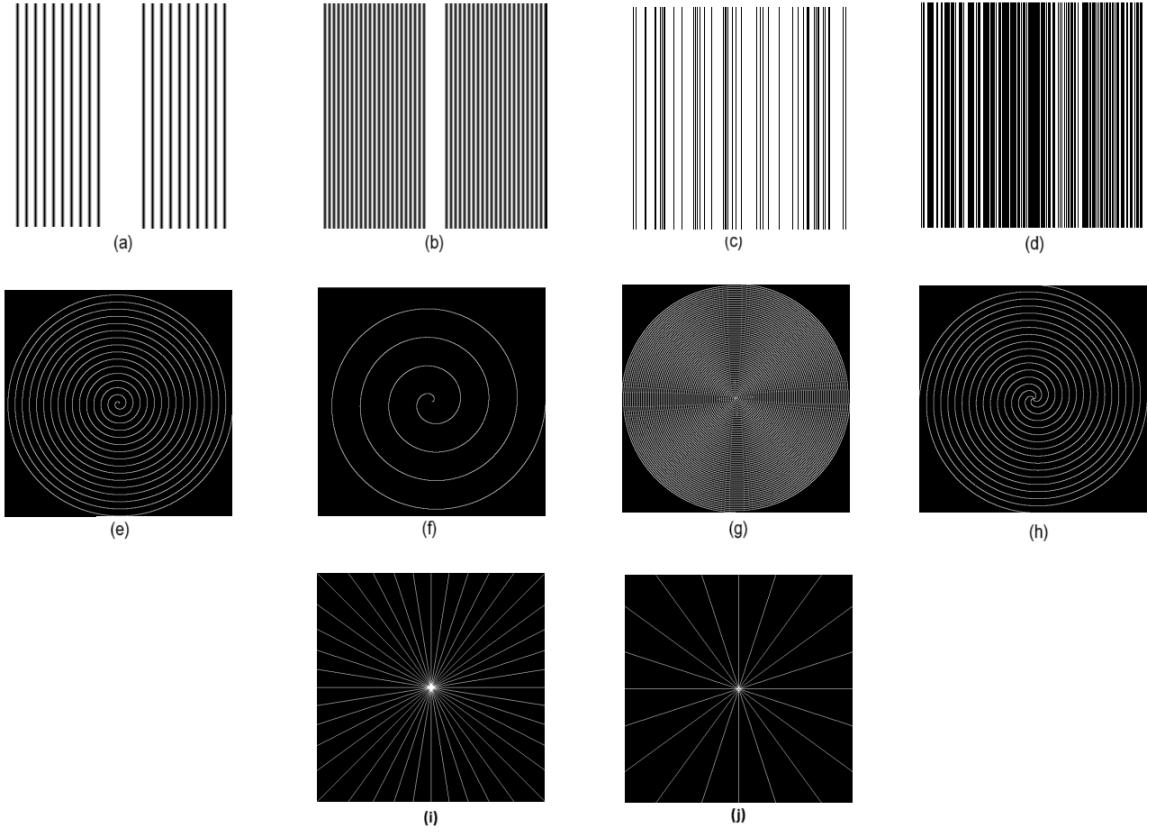


Figura 4.1. Diferentes trajetórias que podem ser aplicadas ao espaço-k. (a) Trajetória tipo *HF* com “*start_interval*” igual a 10 e “*interval*” igual a 4, a Trajetória corresponde a 24,22% de sub-amostragem. (b) Trajetória tipo *HF* com “*start_interval*” igual a 5 e “*interval*” igual a 2, a Trajetória corresponde a 49,22% de sub-amostragem. (c) Trajetória tipo *random* com “*mask_percentage*” igual a 0,8, a trajetória corresponde a 19,92% de sub-amostragem. (d) Trajetória tipo *random* com “*mask_percentage*” igual a 0,3, a Trajetória corresponde a 69,92% de sub-amostragem. (e) Trajetória tipo *spiral1* com “*turns*” igual a 16, a Trajetória corresponde a aproximadamente 93,79% de sub-amostragem. (f) Trajetória tipo *spiral1* com “*turns*” igual a 4, a Trajetória corresponde a 98,43% de sub-amostragem. (g) Trajetória tipo *spiral4* com “*turns*” igual a 16, a trajetória corresponde a 75,16% de sub-amostragem. (h) Trajetória tipo *spiral4* com “*turns*” igual a 4, a trajetória corresponde a 93,72% de sub-amostragem. (i) Trajetória tipo *radial* com “*number_angles*” igual a 20, a trajetória corresponde a 96,13% de sub-amostragem. (j) Trajetória tipo *radial* com “*number_angles*” igual a 10, a trajetória corresponde a 98,06% de sub-amostragem.

As trajetórias utilizadas pelo grupo FastMRI no conjunto de dados de teste e desafio são chamadas por eles de “*Equispaced mask*”, que é igual a trajetória *HF*, e “*Random mask*”, que se assemelha a trajetória *random*, sendo a única diferença que a trajetória usada no conjunto de dados poupa as baixas frequências e apaga aleatoriamente as outras frequências, enquanto a trajetória usada no mestrado apaga todas as freqüências com mesma probabilidade.

O algoritmo também permite a inserção de artefatos. Esses artefatos são inseridos na

imagem reconstruída sem sub-amostragem no espaço-k, depois é realizada a transformada direta para recriar o espaço-k e realizar a sub-amostragem se necessário, dessa forma o artefato aparece na imagem como se tivesse aparecido no espaço-k durante a aquisição. Os artefatos são círculos brancos, seu tamanho, intensidade de cor e posição são escolhidos de forma aleatória, para não criar nenhum padrão que possa influenciar o treinamento. A Figura 4.2 apresenta imagens com artefatos inseridos.

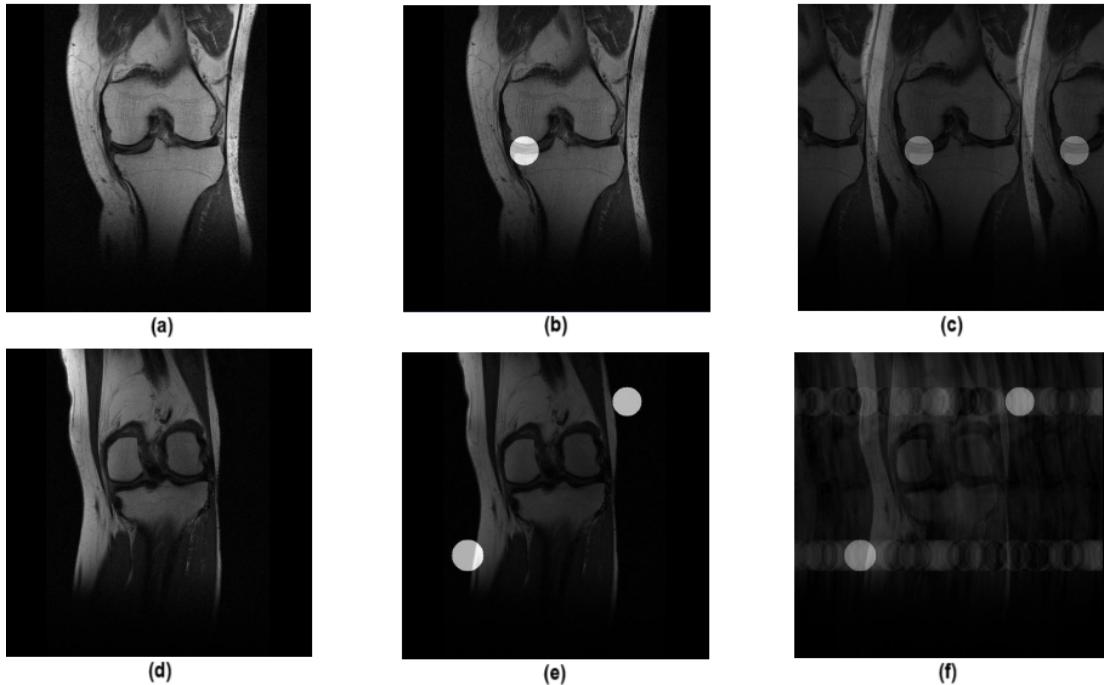


Figura 4.2. Imagens de joelho do banco de dados do FastMRI. (a) Imagem de joelho do conjunto de dados de treino. (b) Artefato colocado na imagem final sem sub-amostragem. (c) Artefato colocado na imagem final depois realizado sub-amostragem no espaço-k da imagem com artefato. Sub-amostragem tipo *HF* com *start_interval* igual a 5 e *interval* igual a 2. (d) Imagem de joelho do conjunto de dados de treino. (e) Dois artefatos colocados na imagem final sem sub-amostragem. (f) Dois artefatos colocados na imagem final depois realizado sub-amostragem no espaço-k da imagem com artefatos. Sub-amostragem tipo *random*, com *mask_percentage* igual a 0,5.

Após realizar o pré-processamento em todas as imagens que serão usadas pela GAN, essas imagens são separadas em três grupos: grupo de treinamento, grupo de validação e grupo de teste. Os três grupos podem passar por pré-processamentos diferentes dependendo da aplicação que será estudada e o algoritmo permite a seleção de parâmetros diferentes para os conjuntos de treino, validação e de teste para cada uma das trajetórias.

As imagens são armazenadas em pares contendo duas imagens do mesmo exame posicionadas lado a lado. Do lado esquerdo é colocada a imagem que será usada no treinamento do modelo discriminador da GAN, são imagens feitas a partir de dados tipicamente amostradas, do lado direito é posicionada uma imagem feita com dados sub-

amostrados de acordo com as especificações usadas no pré-processamento, essas serão as imagens usadas pelo modelo gerativo e discriminativo da GAN como entrada para criar novas imagens. Na Figura 4.3 é possível observar um exemplo de como são armazenadas as imagens.



Figura 4.3. Imagem do conjunto de dados utilizado pela GAN. A imagem é separada em duas partes, na esquerda a imagem feita com dados tipicamente amostrados e a direita a imagem feita com o espaço-k sub-amostrado..

4.1.2 Rede Generativa Adversária

Finalizado o pré-processamento inicia-se a fase de aprimoramento das imagens pela GAN. As imagens são recebidas pela GAN conforme apresentado na Figura 4.3 e o número de imagens pode variar de acordo com as necessidades do projeto. As imagens são separadas em três grupos: imagens de treino, de validação e de teste. As imagens de treino são usadas no treinamento dos modelos gerador e discriminador, as imagens de validação são usadas para checar o estado do treinamento a cada 100 épocas, e as imagens de teste são usadas depois que os modelos já estão treinados como conjunto de imagens a serem aprimoradas.

Para a criação das funções geradora e discriminadora foram utilizadas funções da biblioteca *tensorflow*, a biblioteca foi escolhida por ser a com maior suporte para criação de algoritmos de inteligência artificial.

O algoritmo possui alguns parâmetros modificáveis que permitem a visualização das imagens conforme é realizado o processamento, essas opções são:

- *show_intermediate_images* - Apresenta imagens intermediárias, como por exemplo as imagens que entram no modelo gerativo e qual a saída esperada;
- *show_predicted_images_after_training* - Após a finalização do treinamento apresenta

as imagens criadas em comparação com as imagens de entrada.colunas a terceira será zerada.

O algoritmo também possui configurações que foram ajustadas para o seu melhor funcionamento, são elas:

- *BUFFER_SIZE* - Argumento usado na função “tf.data.Dataset.shuffle”, indica o número de imagens de treinamento que serão embaralhadas para depois serem escolhidas, idealmente deve ser maior ou igual ao número de imagens de treinamento. Foi configurado para sempre ter o mesmo valor que a quantidade de imagens de treinamento.
- *BATCH_SIZE* - O modelo é treinado com conjunto de imagens, o número de imagens nesse conjunto é dado pelo *BATCH_SIZE*. Para modelos U-Net, o ideal é que o tamanho de batch seja igual a um[33].
- *LAMBDA*. Controla o impacto da perda L1 (Diferença absoluta média entre valores preditos e reais) em relação a perda do discriminador. Valor que foi testado empiricamente até chegar no ideal de 150.

No código também é possível ajustar a quantidade de etapas de repetição que serão realizadas na variável “*step*”, o valor normalmente ajustado é de 1000 etapas, mas esse valor foi alterado de acordo com o desejado durante a realização dos testes.

O treinamento dos modelos usando a GPU L4 do Google colab citada anteriormente levava em torno de 30 minutos a cada 1500 épocas do modelo usando as configurações ideais citadas acima. O tempo de aplicação do modelo nas imagens de teste dependia da quantidade de imagens. Para 100 imagens o modelo levava em torno de dez minutos.

As imagens de teste aprimoradas são armazenadas em uma pasta. Para cada imagem aprimorada são geradas quatro imagens: A imagem feita com dados sub-amostrados, a imagem feita com dados tipicamente amostrados, a imagem aprimorada e uma imagem contendo as três primeiras lado a lado para comparação, essa quarta pode ser vista na Figura 4.4.

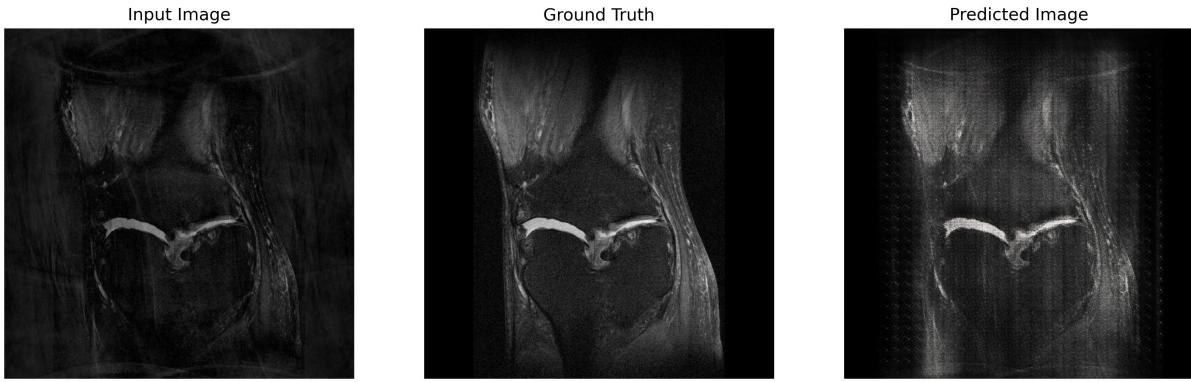


Figura 4.4. Imagens produzidas pelo algoritmo comparando as imagens de entrada (reconstruída a partir de dados sub-amostrados com a trajetória *spiral1*), tipicamente amostrada e aprimorada.

Nas seções a seguir será explicado de forma mais detalhada o funcionamento dos modelos gerador e discriminador contidos dentro da GAN.

4.1.3 Rede Generativa

A rede gerativa foi baseada no modelo U-Net, usado em algoritmos pix2pix. O modelo U-Net é um modelo de rede neural convolucional muito usado na segmentação de imagens médicas pois é eficiente em extrair informações de imagens mesmo com uma quantidade limitada de amostras. A arquitetura da U-Net consiste em um encoder onde ocorre a redução da resolução espacial com ênfase nas características da imagem (downsampling), e um decoder que realiza o aumento da resolução da imagem com as características realçadas (upsampling)[33].

No encoder é realizada a operação de downsampling, que diminui as dimensões da imagem. São utilizadas oito camadas convolucionais com diferentes quantidades de filtros de tamanho 4x4 e salto de dois pixels para os filtros(strides = 2), a cada salto é realizada a soma ponderada dos valores e salvo em uma nova matriz, dessa forma as dimensões da imagem são reduzidas pela metade. No primeiro downsample são 64 filtros e cada um desses filtros realiza as operações de saltos e somas ponderadas para gerar novas matrizes, dessa forma gerando 64 matrizes.

No primeiro downsample também não é aplicada a normalização da batch aos dados pois os dados ainda não passaram por transformações e aplicar a normalização nesse momento pode excluir dados úteis, porém a normalização é aplicada nas próximas sete camadas.

A cada camada as dimensões das matrizes são reduzidas pela metade e, para não serem perdidas informações, a quantidade de filtros aumenta a cada downsample, começando

com 64 na primeira camada e terminando com 512 na última.

No decoder é realizado o upsampling, que aumenta as dimensões da imagem. São usadas sete camadas que realizam convolução transposta, com diferentes quantidades de filtros e salto de dois pixels para os filtros ($\text{strides} = 2$). A cada camada são usadas diferentes quantidades de filtros de tamanho 4x4 para recuperar a resolução da imagem, começando com 512 para poder processar melhor os diversos detalhes, e terminando com 64 filtros. Também as três primeiras camadas são implementadas contendo dropout, essas camadas zeram metade dos seus resultados, forçando a rede a não depender totalmente das partes que são zeradas, evitando que o modelo não consiga generalizar os dados de treinamento, ou seja, evitando o overfitting[37].

O decoder também possui camadas de concatenação, entre as camadas de upsampling, essas camadas recebem, através das skip connections imagens intermediárias das camadas de downsampling e as concatena. Skip connections são conexões entre camadas do encoder com a camada correspondente do decoder, dessa forma, possíveis detalhes que tenham sido perdidos durante a etapa de donwsampling podem ser recuperados durante o upsampling. Após cada camada de upsampling a matriz sendo processada é concatenada com a matriz do bloco donwsampling correspondente, dessa forma são realizadas sete operações do tipo, uma para cada skip connection.

A arquitetura completa do gerador pode ser observada na Figura 4.5, nela é possível observar também o formato em "U" que dá nome a arquitetura U-Net.

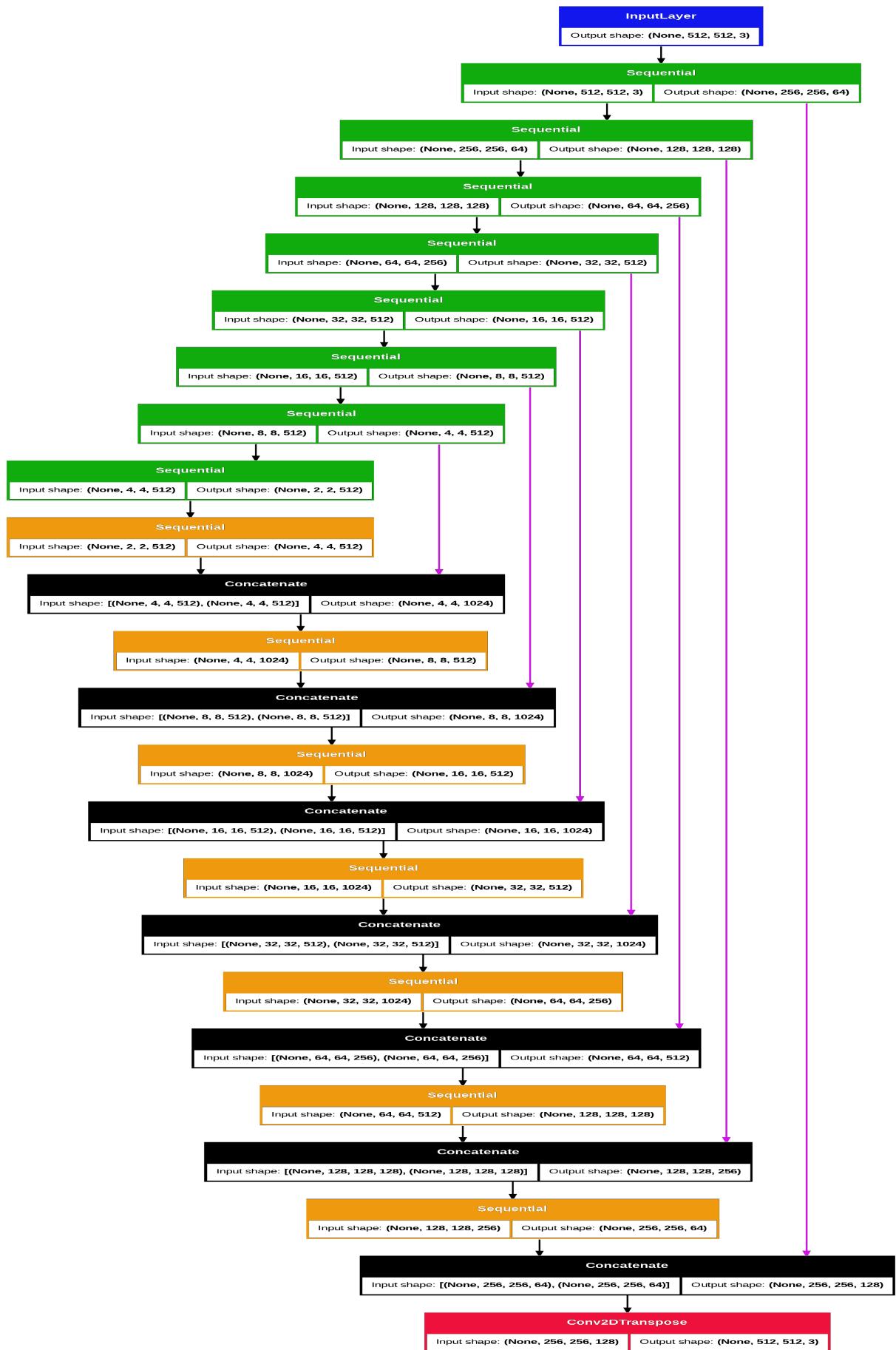


Figura 4.5. Arquitetura do modelo Gerador.

Na imagem, os retângulos representam as camadas do modelo gerador. De azul temos a camada de entrada que recebe as imagens de tamanho 512 por 512 e com os três canais correspondendo ao RGB, mesmo que a imagem seja preta e branca ela possui os três canais de cor com a mesma intensidade. As setas pretas representam as conexões entre as camadas.

Os retângulos verdes representam as oito camadas do encoder, cada uma possuindo suas próprias dimensões de entrada e saída, podemos ver que a cada camada, as dimensões da imagem diminuem e o número de imagens aumenta, pois a cada camada o número de filtros cria aquele correspondente número de imagens.

As setas rosa representam as skip connections, que ligam as camadas do encoder as camadas do decoder. Os retângulos laranja correspondem às sete camadas do decoder e são associadas às camadas de concatenação, representadas pelos retângulos pretos. As camadas de concatenação recebem matrizes das skip connections e agrupa todas elas juntas, dobrando o número de imagens. As camadas do decoder então diminuem o número de imagens enquanto aumenta o tamanho da imagem.

Por fim, é realizada uma convolução transposta, que aumenta a resolução da imagem, possui saída também com três canais RGB e utiliza a função de ativação tanh que normaliza os valores da matriz entre -1 e 1.

4.1.4 Rede Discriminativa

A rede discriminativa é um modelo convolucional classificatório que deve distinguir pares de imagens. Existem dois tipos de pares. O primeiro tipo de par contém a imagem de entrada da GAN, feita usando dados sub-amostrados, em conjunto com a imagem feita com dados tipicamente amostrados, chamada de imagem alvo. Esse primeiro par é chamado de par verdadeiro. O segundo tipo de par é composto pela imagem de entrada e a imagem gerada pelo modelo gerativo, chamada de par falso.

Ele recebe duas imagens de dimensões 512x512x3 na sua camada de entrada, e as concatena, gerando um novo conjunto de tamanho 512x512x6. São então realizadas uma série de downsampleings em sete camadas que diminuem as dimensões da imagem evidenciando suas características. Assim como no modelo gerador, esses downsampleings possuem filtros de tamanho 4x4, saltos de tamanho 2 (strides = 2) e possuem filtros que vão aumentando a cada camada, começando com 64 e finalizando com 512. Também é aplicada a regularização do L2, que penaliza pesos muito grandes na rede, forçando o modelo a possuir pesos menores evitando o overfitting^[38].

Assim como na rede geradora, em algumas das camadas de downsample é aplicado dropout, que desativa metade das unidades da rede convolucional de forma aleatória

a cada atualização de pesos. Isso ajuda a rede a não ficar completamente dependente daquelas conexões evitando assim o overfitting, ou seja, conseguindo generalizar melhor os dados de treinamento[37].

O modelo possui uma camada que ajusta o tamanho dos dados adicionando duas camadas de zeros aos dados (zeropadding), seguida de uma camada de convolução com 512 filtros de tamanho 4x4, saltos de tamanho 1 e regularização do L2.

Em seguida é utilizada uma camada de normalização de batch para normalizar a camada anterior para melhorar a estabilidade do treino, seguida de uma camada que aplica a função de ativação LeakyReLU, que mantém os valores positivos e diminui os negativos sem zerá-los, evitando neurônios mortos e problemas de convergência.

Por fim, é aplicada mais uma camada de zeropadding e então uma última convolução que faz com que a saída dos dados tenha dimensões bem reduzidas, de tamanho (None, 2, 2, 1), sendo a primeira dimensão o tamanho de batch, o segundo e terceiro o tamanho da imagem final, e o quarto, que contém apenas um valor, o veredito do discriminador, que vai de 0 a 1, sendo zero um par de imagens falsas, e um, um par de imagens verdadeiras.

A Figura 4.6 contém a arquitetura do modelo discriminador, com todas as suas camadas e dimensões dos dados.



Figura 4.6. Arquitetura do modelo Discriminador.

Na Figura 4.6, os blocos representam as camadas da rede, com as respectivas dimensões dos dados de entrada e de saída. Os blocos vermelhos se tratam das camadas de entrada, seguidas pelo bloco azul que representa a camada de concatenação. Os blocos verdes representam os downsamplings e os blocos rosa representam os dropouts. No final do modelo, os blocos laranja representam os zeropadding, os amarelos as convoluções e os dois blocos pretos representam primeiro a normalização de batch e função de ativação LeakyReLU.

4.1.5 Imagens

As imagens produzidas pelo algoritmo são salvas de forma que possibilite a comparação entre a imagem de entrada, a imagem tipicamente amostrada e a imagem produzida conforme mostrado na Figura 4.4.

A Figura 4.7 apresenta alguns resultados para as trajetórias *spiral4* (teste 78), *radial* (Teste 94) e *spiral1* (Teste 103). A figura apresenta as melhores imagens produzidas pelos seus respectivos testes. As métricas com os resultados de todos os testes foram organizadas em tabelas que podem ser observadas no repositório do Github do projeto[7]. As imagens e modelos treinados podem ser requisitadas por e-mail pois consistem em mais de 20 Gb de material.

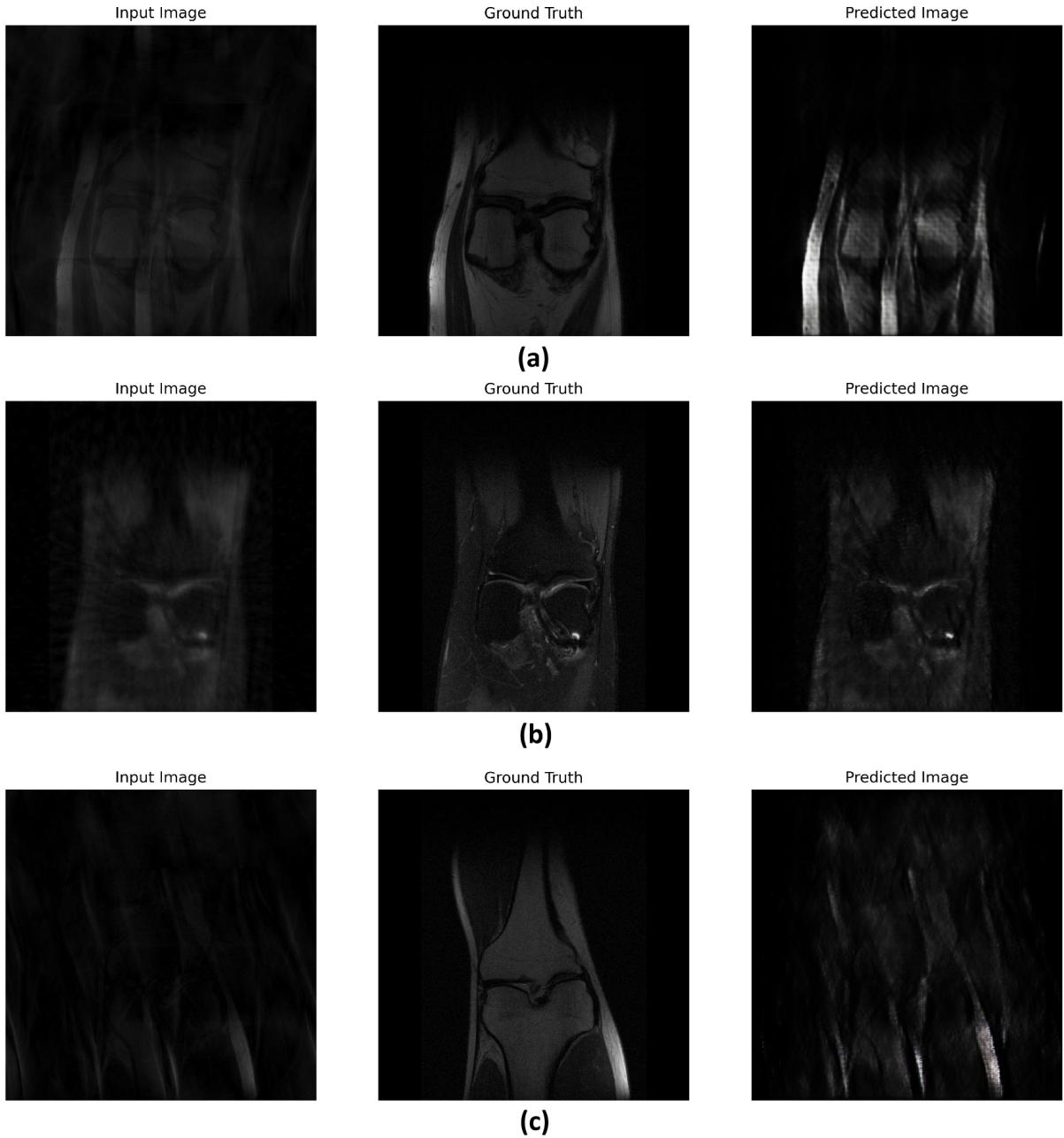


Figura 4.7. Imagens geradas pela GAN durante os testes. (a) Imagens geradas pelo teste n° 78. (b) Imagens geradas pelo teste n° 94. (c) Imagens geradas pelo teste n° 103.

4.1.6 Gráficos de Perda (Loss)

Nos testes mais avançados do estudo foram extraídos as curvas da perda da rede gerativa durante o treino e a validação em função da época, para poder observar se o modelo estava generalizando de forma satisfatória os dados evitando assim o overfitting, que acontece quando o modelo só consegue bons resultados para as imagens de treinamento. Na Figura 4.8 é possível observar um desses gráficos, referentes ao teste de

número 78, onde apesar do erro da validação estar oscilando ele não está crescendo e está estabilizado abaixo de 100, enquanto o erro do treinamento fica próximo a zero.

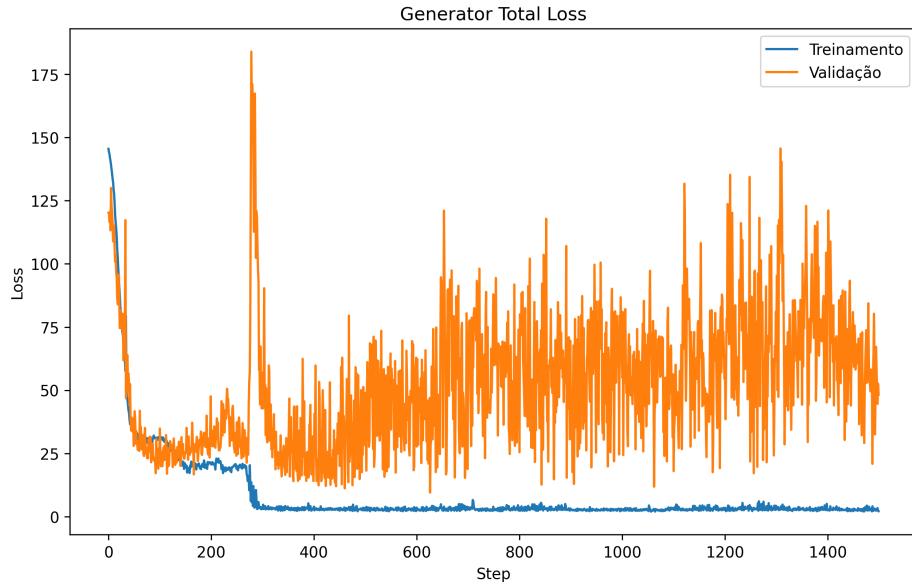


Figura 4.8. Curva de perda no treino/validação do gerador em função da época.

Também foram extraídos gráficos da perda da rede generativa e da rede discriminativa em função da época. Esse tipo de gráfico permite observar se as redes estão equilibradas. Caso as curvas se afastem muito, pode significar que uma rede está aprendendo mais a cada época do que a outra, impactando nos resultados finais. A curva referente ao teste número 78 pode ser observada na Figura 4.9.

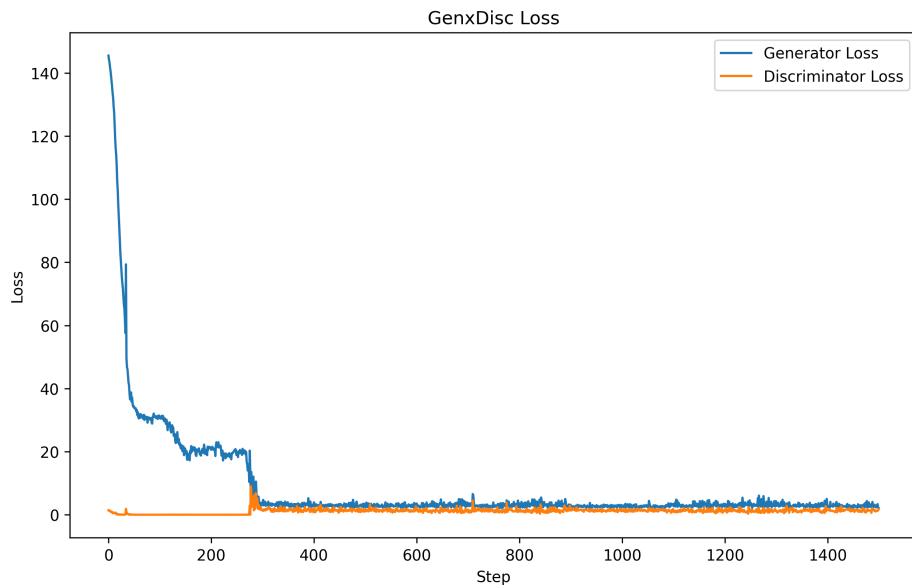


Figura 4.9. Curva de perda do gerador e do discriminador em função da época.

Os gráficos extraídos dos testes citados neste trabalho poderão ser observados no repositório do github do projeto[7].

4.1.7 Tabelas

Os resultados dos testes foram organizados em tabelas contendo os parâmetros da GAN, e também as especificações de sub-amostragem. Na Tabela 4.1 é possível observar alguns resultados para a trajetória *spiral4*.

Tabela 4.1. Métricas de qualidade para testes da trajetória *spiral4*.

Nº	Nº Épocas	Tamanho de Buffer	SSIM In	SNR In (dB)	SSIM Pr	SNR Pr (dB)	RSSIM	RSNR
65	1500	1000	$0,28 \pm 0,06$	$4,02 \pm 2,08$	$0,58 \pm 0,07$	$8,02 \pm 3,94$	2,067	1,996
69	1500	1000	$0,27 \pm 0,05$	$4,14 \pm 1,42$	$0,56 \pm 0,06$	$7,17 \pm 1,35$	2,107	1,732
75	500	1000	$0,28 \pm 0,06$	$2,92 \pm 1,63$	$0,58 \pm 0,08$	$5,55 \pm 2,05$	2,096	1,900
78	1500	1000	$0,33 \pm 0,05$	$5,15 \pm 1,65$	$0,66 \pm 0,05$	$11,53 \pm 5,90$	1,977	2,241
80	1500	1500	$0,29 \pm 0,06$	$4,02 \pm 2,15$	$0,63 \pm 0,08$	$7,84 \pm 3,85$	2,180	1,948

Na Tabela 4.1 todos os testes foram realizados com 1500 imagens de treino de tamanho 512 por 512 com exceção do teste 78 que foram utilizadas 3000 imagens. O número de voltas da espiral da trajetória *spiral4* foi de 16 voltas tanto para o treino quanto para o teste, poupando 24,8% do espaço-k. As variáveis precedidas de Pr se tratam daquelas geradas pelo modelo (*predicted*), e as precedidas de In são obtidas das imagens de entrada (*input*).

A Tabela 4.2 contém os melhores resultados para cada trajetória, esses resultados foram os que obtiveram melhor relação de melhora no SSIM e maior SSIM final. Para chegar nessa relação foi multiplicado o valor do SSIM final pelo RSSIM, e os testes com os maiores valores da multiplicação foram considerados os melhores.

Tabela 4.2. Métricas de qualidade dos melhores resultados de cada trajetória.

Teste	Espaço-k Poupado	Trajetória	SSIM In	SNR In	SSIM Pr	SNR Pr	RSSIM
89	50,39%	HF	$0,53 \pm 0,10$	$4,74 \pm 3,13$	$0,63 \pm 0,09$	$7,43 \pm 5,31$	1,184
53	30,07%	Random	$0,23 \pm 0,10$	$0,47 \pm 2,33$	$0,50 \pm 0,04$	$6,98 \pm 3,19$	2,124
103	24,49%	Spiral1	$0,08 \pm 0,04$	$2,33 \pm 1,30$	$0,38 \pm 0,06$	$6,68 \pm 3,15$	4,774
80	24,84%	Spiral4	$0,29 \pm 0,06$	$4,02 \pm 2,15$	$0,63 \pm 0,08$	$7,84 \pm 3,85$	2,180
94	6,14%	Radial	$0,53 \pm 0,11$	$5,38 \pm 3,38$	$0,65 \pm 0,09$	$7,40 \pm 4,98$	1,246

4.2 DISCUSSÃO

Após os 110 testes foram obtidos resultados que melhoraram as métricas de qualidade das imagens. Serão discutidos sobre os detalhes, assim como as métricas de qualidade, imagens produzidas e estatísticas que podemos extrair desses testes.

4.2.1 Métricas de Qualidade

Analisando os resultados obtidos foi possível observar padrões nas métricas de qualidade obtidas e nas características das imagens. Tomemos o teste número 80 citado na Tabela 4.1, a média de SSIM das imagens geradas foi de 0,63, porém foram analisadas as métricas individuais de cada imagem de teste, e os dados organizados no diagrama boxplot da Figura 4.10.

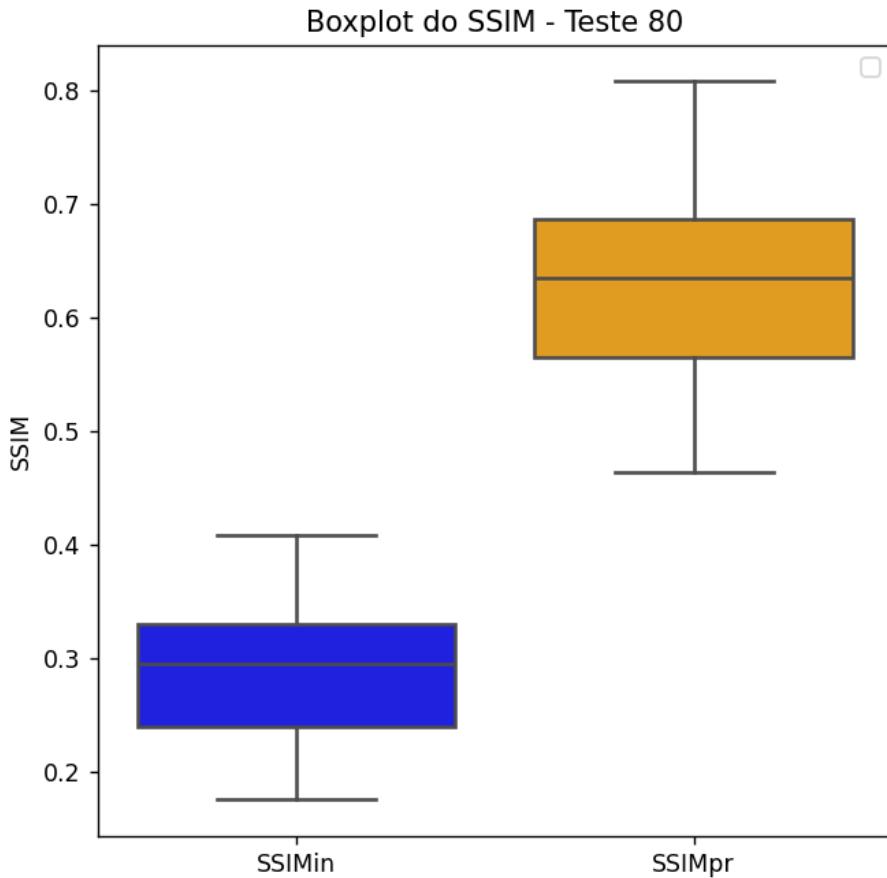


Figura 4.10. Boxplot de SSIM para imagens de entrada e saída do teste 80.

Pelo boxplot é possível observar que todas as imagens do conjunto foram aprimoradas, e apesar da média de SSIM das imagens aprimoradas ser de 0,63, algumas imagens possuem SSIM acima de 0,7. Na Figura 4.11 estão todas as imagens do teste 80 que obtiveram SSIM acima de 0,7.

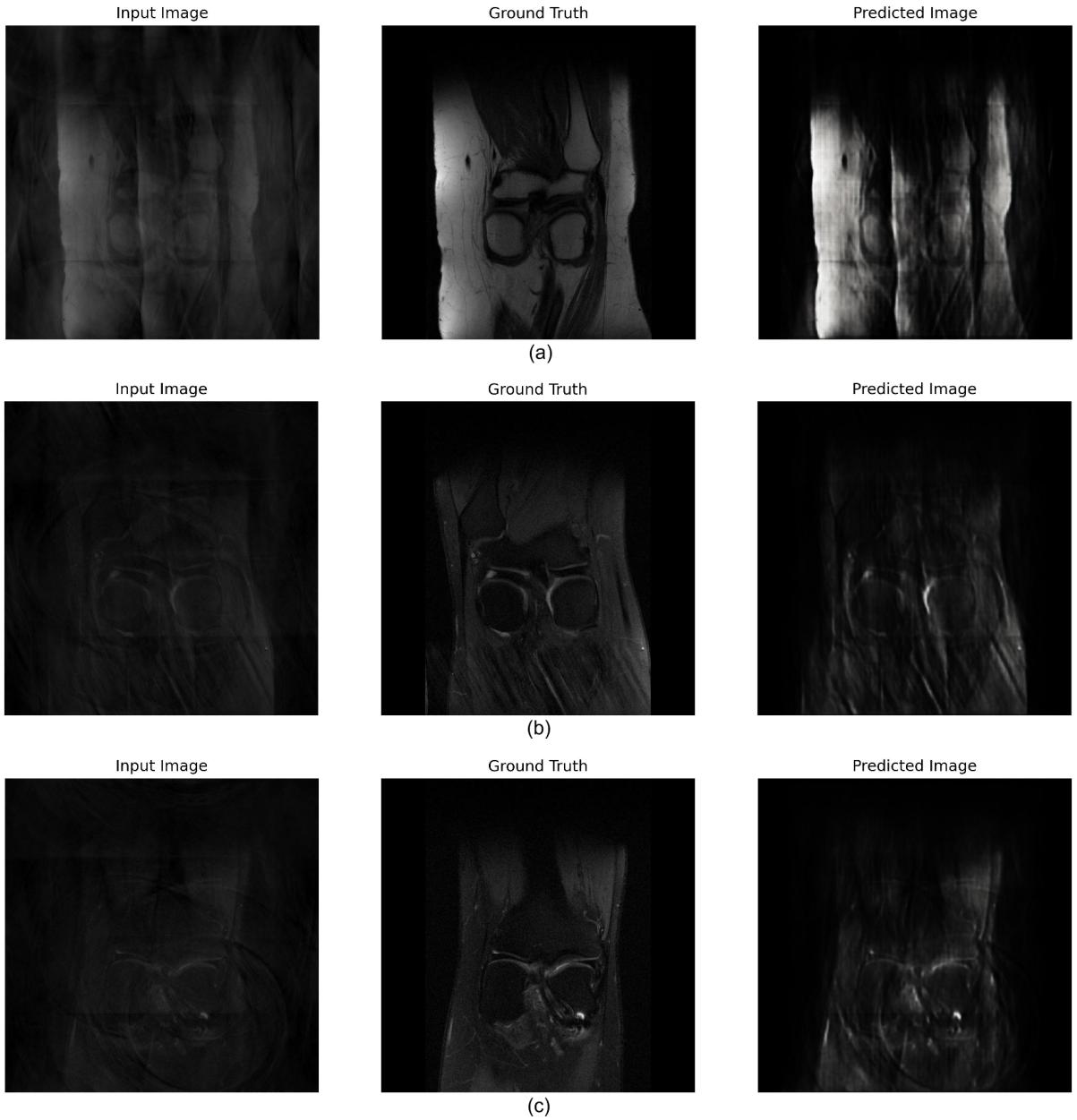


Figura 4.11. Melhores imagens produzidas no teste 80. (a) SSIM = 0,72. (b) SSIM = 0,80. (c) SSIM = 0,81.

Assim como nas imagens (b) e (c), que foram as melhores imagens produzidas nesse teste, as melhores imagens produzidas em outros testes também tinham a característica de serem mais escuras, com alguns poucos detalhes claros, como mostrado anteriormente na Figura 4.7

Isso mostra um padrão, onde as imagens mais escuras obtém maiores valores de SSIM pois devido ao fato de que as imagens sendo comparadas possuem poucos detalhes e baixa amplitude dos pixels faz com que as imagens geradas se assemelhem mais com as imagens referência. Porém essas imagens possuem baixo SNR devido a baixa amplitude dos pixels. Considerando que o SNR médio para o teste 80 foi de 7,84 dB todas as

imagens da Figura 4.7 conseguiram resultados abaixo da média possuindo: (a) $\text{SNR} = 4,50 \text{ dB}$, (b) $\text{SNR} = 2,66 \text{ dB}$ e (c) $\text{SNR} = 4,63 \text{ dB}$. Esse aspecto mostra a importância de analisar SSIM e SNR sempre em conjunto

Outro ponto interessante a ser observado é o boxplot da média do SSIM das imagens geradas nos testes onde ocorreu melhora nas métricas de qualidade, correspondente a 71 dos 103 primeiros testes realizados. A Figura 4.12 apresenta o boxplot da média do SSIM das imagens de entrada e das imagens geradas por testes onde o RSSIM foi maior do que um, ou seja, ocorreu melhora nas imagens.

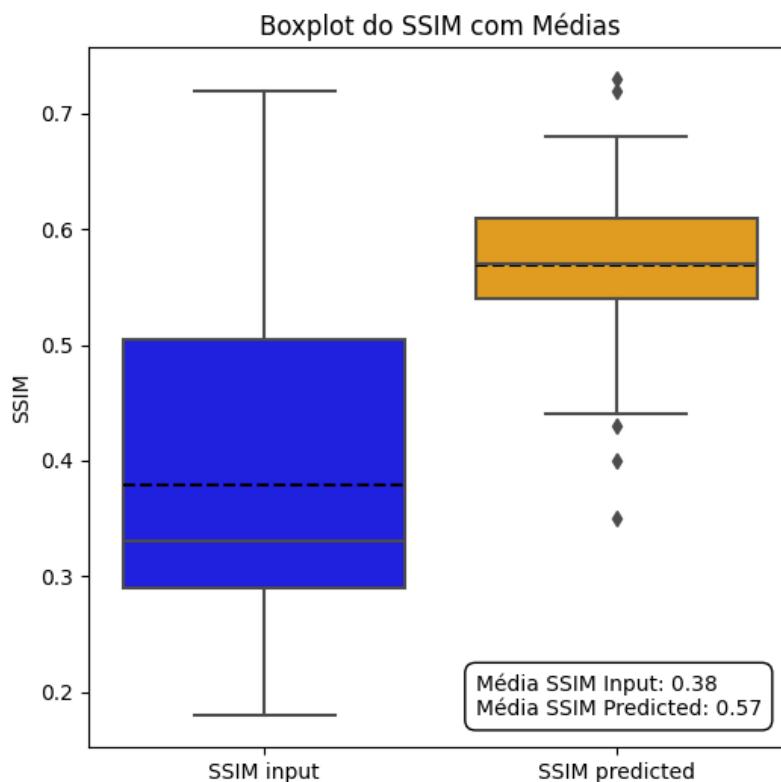


Figura 4.12. Boxplot do SSIM de testes que ocorreram melhora nas métricas de qualidade.

Na Figura 4.12 é possível observar que as métricas das imagens de entrada tendem a ser bem dispersas, e ao passarem pela GAN elas tendem a ter um menor desvio padrão. E além disso o algoritmo tende a convergir em volta da média de 0,57 de SSIM, com imagens de entrada com diferentes SSIM passando a ter um SSIM perto dessa média após a GAN.

4.2.2 Completar Espaço-k

Durante os testes foi testada a abordagem citada anteriormente na seção de Trabalhos Correlatos do estudo “Deep learning for undersampled MRI reconstruction” [20]. O teste foi realizado de duas formas diferentes, na primeira os espaços-k das imagens geradas pela GAN eram usados para completar o espaço-k das imagens de entrada, assim gerando um novo espaço-k e consequentemente uma nova imagem, as métricas então eram comparadas com as imagens referência e também com as imagens antes da fusão dos espaços-k. Nos testes realizados dessa primeira forma não foram observadas melhorias, e sim uma leve queda no SSIM das imagens. Na Figura 4.13 é possível observar as imagens de um dos testes, e na Tabela 4.3 a média das métricas de cem imagens testadas.

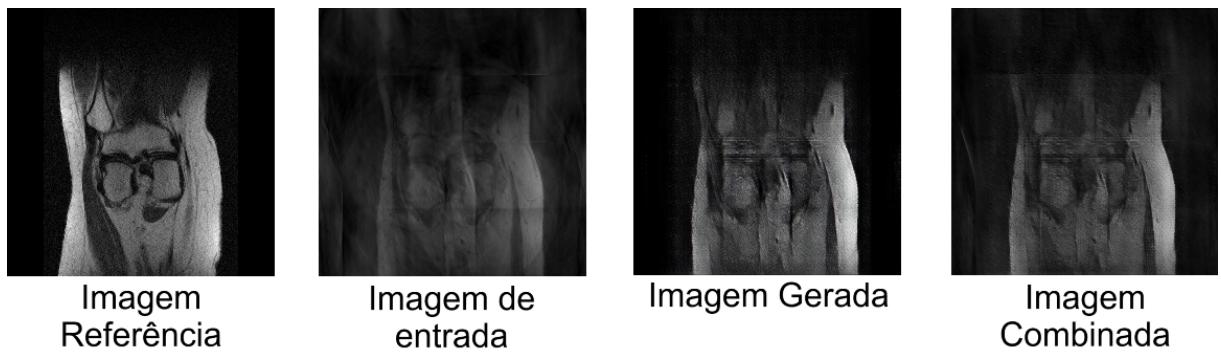


Figura 4.13. Imagens do teste que combina os espaços-k.

Tabela 4.3. SSIM das imagens obtidas com o teste que combina os espaços-k.

SSIM In	SSIM Pr	SSIM Comb
$0,290 \pm 0,02$	$0,493 \pm 0,03$	$0,295 \pm 0,03$

A segunda forma de testar o método de junção do espaço-k foi usar a imagem produzida a partir da junção como nova imagem de entrada para o modelo gerador repetindo o processo várias vezes. Testou-se dessa forma pois esperava-se que o modelo treinado para melhorar imagens seria capaz de melhorar outras imagens mesmo que essas tivessem parâmetros de amostragem diferentes, porém também não foram obtidos bons resultados, como mostrado na Figura 4.14.

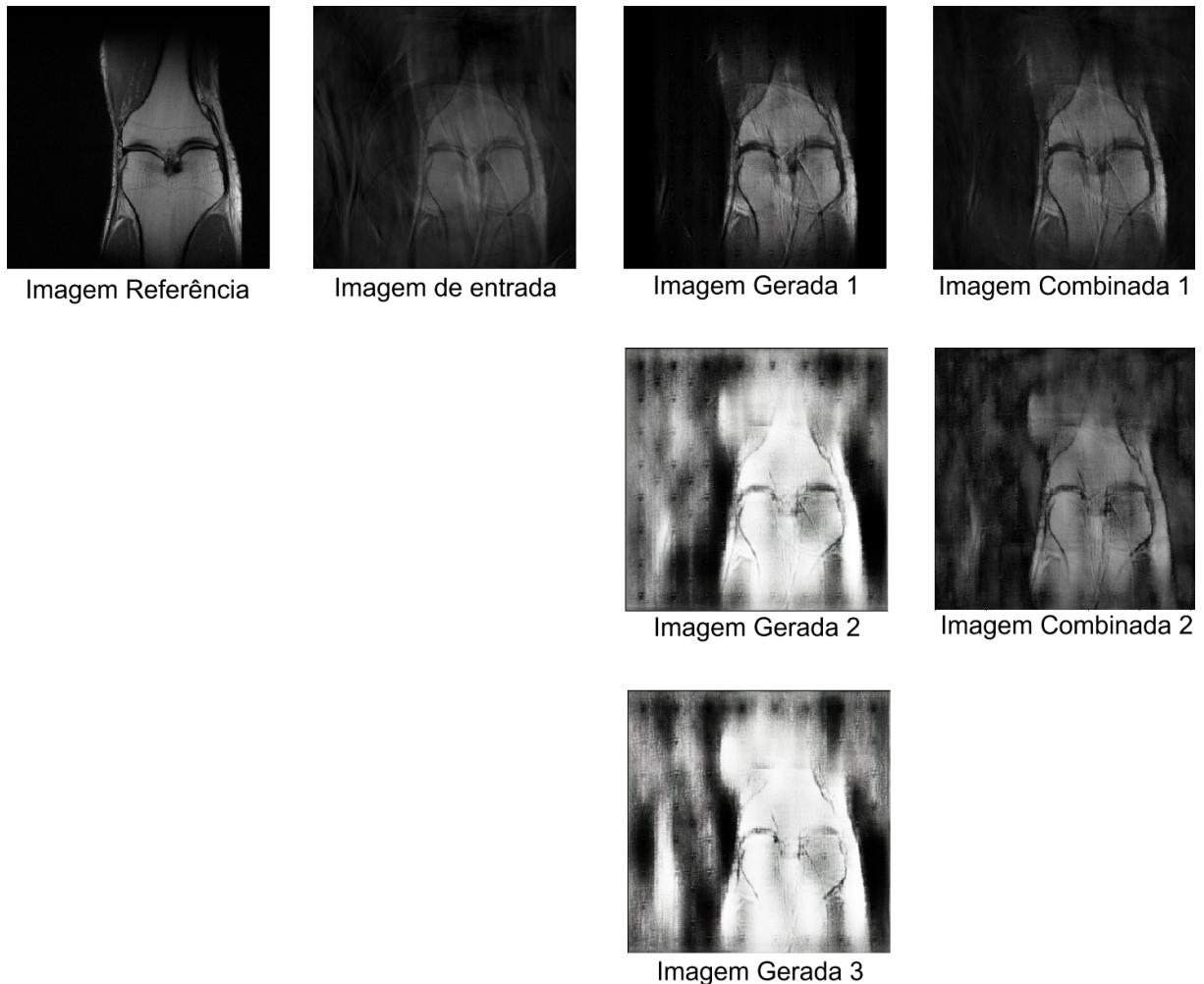


Figura 4.14. Imagens referência, de entrada, geradas e combinadas para comparação. A imagem referência é a feita a partir do espaço-k tipicamente amostrado, a imagem de entrada é a feita com o espaço-k sub-amostrado, as imagens geradas são aquelas produzidas pelo modelo e as combinadas são as feitas a partir da combinação do espaço-k da imagem gerada com a imagem de entrada.

4.2.3 Comparação com Outros Modelos

Os resultados e métricas obtidas durante os testes foram comparados com aqueles disponibilizados pelo grupo FastMRI em sua tabela de classificação. Apesar das trajetórias usadas neste trabalho não serem as mesmas usadas pelos competidores na época, o conjunto de dados foi o mesmo, o de imagens de joelho com bobina única.

Para realizar a comparação foram selecionados os algoritmos criados na época que processavam espaços-k com 4% das suas baixas frequências preservadas e com diferentes taxas de amostragem no restante do espaço-k. Para a comparação, foi selecionado o melhor resultado obtido no algoritmo feito para este trabalho com a sub-amostragem no centro do espaço-k semelhante. Para isso foi escolhido um teste usando a trajetória *spiral4* que possui 2,11% das baixas frequências poupadadas.

Na Tabela 4.4 é possível observar a comparação entre os resultados obtidos pelo algoritmo feito neste trabalho (Mestrado), o algoritmo com melhor resultado do desafio do FastMRI (AIRS-Net) e o melhor algoritmo GAN (UnetGan) na tabela da competição. O teste do mestrado considerado para a comparação é o teste 78. Não foram comparadas as relações sinal ruído pois neste estudo foi usado SNR por área de interesse, e na competição do FastMRI foi usada a relação sinal ruído de pico.

Tabela 4.4. Comparação do modelo criado com os da competição do FastMRI.

Algoritmo	SSIM
AIRS-Net	0,6927
UnetGan	0,6061
Mestrado	0,6569

Os modelos criados para a competição não conseguiram valores maiores do que 0,7. Durante os testes do mestrado a maior parte das imagens acima de 0,65 possuem baixo contraste e a maior média obtida em um teste é de 0,73. Esse comportamento dos resultados pode ser creditado ao fato de que o conjunto de dados usado é de bobina única, nele o espaço-k obtido vêm de várias bobinas que são combinados em um único espaço-k para diminuir o tamanho ocupado em memória porém perdendo alguns dados.

A tabela de classificação possui os resultados da competição de imagens de cérebro, que usa um conjunto de dados obtido de várias bobinas. Nos modelos em questão os resultados conseguiram alcançar valores de até 0,96. Porém o uso desse tipo de conjunto de dados é mais limitado devido ao espaço ocupado em memória. O conjunto completo de imagens de cérebro multi-bobina passa de 500Gb.

Outro ponto interessante de se notar é que em estudos onde se obteve melhores resultados, como o estudo com GANs citado nos trabalhos correlatos, os dados são obtidos de forma que evidencie alguma característica na frequência, não sendo o intuito do mestrado, que era usar dados da forma como são obtidos em exames hoje.

4.2.4 Comparação Entre Trajetórias

A Tabela 4.2 apresentada anteriormente apresentou os melhores resultados para cada uma das trajetórias. Foi multiplicado os valores de SSIM Pr e RSSIM de cada teste, aqueles com os maiores resultados foram considerados os melhores, por combinar alto valor de SSIM da imagem gerada e alta taxa de crescimento do SSIM.

Nesses testes foi observado que as trajetórias que pouparam mais as baixas frequências

(*HF*, *spiral4* e *radial*) obtiveram os maiores valores de SSIM da imagem gerada, isso acontece pois a maior parte da informação de um espaço-k está contido nas baixas frequências. Essa informação se confirma ao observar o teste com a trajetória *radial*, a que mais poupa as baixas frequências, em que apenas 6,14% do espaço-k foi poupado, porém a imagem de entrada possui SSIM de 0,53. Esse comportamento mostra que a GAN consegue melhores resultados quando é treinada com imagens que possuem mais informações nas baixas frequências do seu espaço-k.

Porém é interessante notar que o maior crescimento de SSIM foi obtido com a trajetória *spiral1*, aumentando o valor de 0,08 da imagem de entrada para 0,38. Com um conjunto de dados mais vasto, poderia-se testar a possibilidade de gerar um novo conjunto de dados de treino e treinar um novo modelo com essas imagens, e assim poder investigar se ocorreria uma melhoria ainda maior nas imagens finais. Essa forma de repetição difere da testada durante o mestrado, pois realiza um novo treinamento com um novo conjunto de dados ao invés de repetir usando o modelo previamente treinado, porém por realizar um novo treinamento precisaria de mais imagens para gerar um conjunto de treinamento novo.

Outro ponto interessante de se notar na tabela com os melhores resultados é o aumento do SNR para o teste da trajetória *random*. Apesar de não conseguir atingir um SSIM acima de 0,6 como os melhores testes, o teste com a máscara *random* conseguiu diminuir de forma significante o ruído da imagem de entrada como mostrado na Figura 4.15, aumentando assim o SNR de 0,47 dB para 6,98 dB.

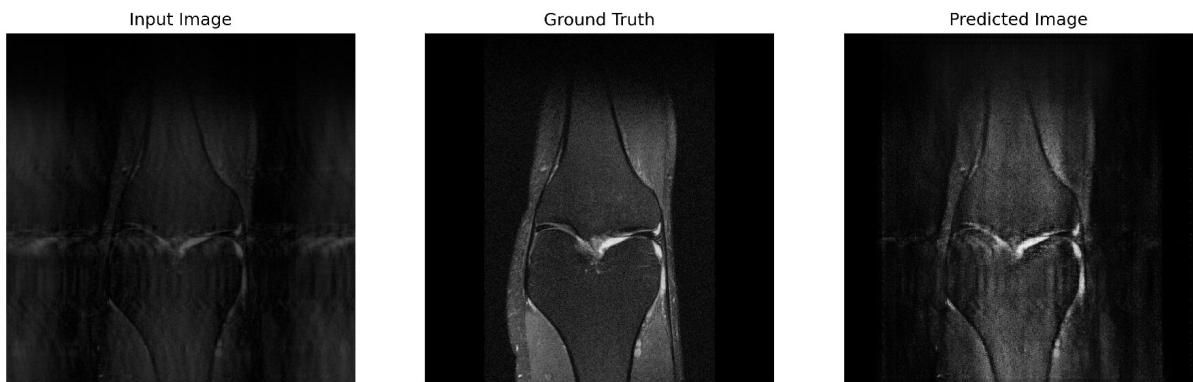


Figura 4.15. Imagem produzida pelo teste n° 53.

5 CONCLUSÃO

O estudo realizou a reconstrução de imagens de ressonância magnética a partir de dados sub-amostrados de bobina única usando redes gerativas adversárias. GANs são ferramentas poderosas no processamento de imagens e foi investigado o uso dessa ferramenta para o aprimoramento de imagens de ressonância feitas a partir de dados em diferentes cenários de sub-amostragem.

O trabalho foi capaz de reconstruir imagens de ressonância magnética feitas a partir de dados sub-amostrados, melhorando as métricas de qualidade das imagens de entrada. O modelo GAN foi capaz de obter a melhora com diferentes tipos de trajetórias e diferentes padrões de sub-amostragem apresentando um limite do quanto poderia melhorar as imagens.

Os melhores resultados foram obtidos para as trajetórias *HF*, *spiral4* e *radial*. Essas trajetórias são as que mais pouparam as baixas frequências do espaço-k. Porém, bons resultados também foram obtidos com as trajetórias *random*, que obteve alta melhora no SNR e *spiral1* que conseguiu melhorar de forma significativa o SSIM, ambos inicialmente muito baixos nas imagens de entrada do modelo.

Foi possível atingir os hiperparâmetros ideais para a GAN através de testes e do estudo da literatura, baseando as decisões em trabalhos correlatos que abordaram temas como GANs, U-Net e pix2pix.

Ao comparar o modelo criado com aqueles feitos para a competição do FastMRI, foi possível observar que os resultados obtidos conseguiram se equiparar aos resultados obtidos por modelos trabalhando com o mesmo conjunto de dados, e quando comparado com outras GANs foi capaz de superá-las.

Apesar da obtenção de resultados positivos, o fato do uso de um conjunto de dados de bobina única limita a obtenção de métricas próximas a 100%. Esse tipo de conjunto de dados combina os espaços-k de várias bobinas em um só espaço-k, simulando uma bobina única, porém nesse processo muita informação é perdida. Para efeitos de comparação, o conjunto de dados de treino de bobina única de imagens de joelho possui 70 GB enquanto o multi-bobina possui 455 GB. Conjuntos de dados de múltiplas bobinas, poderiam permitir melhorias mais expressivas na qualidade das imagens geradas, porém são mais exigentes

em termos computacionais.

Em trabalhos futuros serão testados conjuntos de dados mais amplos, com múltiplas bobinas e um pré-processamento auxiliado por inteligência artificial, uma vez que modelos classificatórios conseguiram bons resultados ao preparar os conjuntos de dados para treinamento [8]. Também serão realizados testes em que as imagens geradas pelo modelo possam ser usadas como conjunto de treinamento para um outro modelo, fazendo assim um aprimoramento em múltiplos estágios de treinamento.

Outra abordagem que será estudada em trabalhos futuros é o processamento dos espaços-k no aprimoramento das imagens. Duas formas serão abordadas, a primeira treinando a GAN diretamente com os espaços-k ao invés das imagens reconstruídas, gerando espaços-k aprimorados.

A segunda forma, usando espaços-k de conjuntos multi-bobina reconstruindo as imagens a partir de cada um dos espaços-k. Dessa forma, um exame usando quatro bobinas geraria quatro imagens diferentes incompletas, porém cada uma possuindo aspectos únicos devido às diferenças nos espaços-k.

LISTA DE REFERÊNCIAS

- [1] Emmanuel Ahishakiye, Martin Bastiaan Van Gijzen, Julius Tumwiine, e Johnes Obungoloch. Adaptive-size dictionary learning using information theoretic criteria for image reconstruction from undersampled k-space data in low field magnetic resonance imaging. *BMC medical imaging*, 20:1–12, 2020.
- [2] Parameswari Alagarsamy, Bhavani Sridharan, e Vinoth Kumar Kalimuthu. A convolutional deep neural network based brain tumordiagnoses using clustered image and feature-supported classifier (cifc) technique. *Brazilian Archives of Biology and Technology*, 66:e23230012, 2023.
- [3] Alireza Nasiri Avanaki. Exact global histogram specification optimized for structural similarity. *Optical review*, 16:613–621, 2009.
- [4] Denise Bos, Nika Guberina, Sebastian Zensen, Marcel Opitz, Michael Forsting, e Axel Wetter. Radiation exposure in computed tomography. *Deutsches Ärzteblatt International*, 120(9):135, 2023.
- [5] Elizabeth Cole, Joseph Cheng, John Pauly, e Shreyas Vasanawala. Analysis of deep complex-valued convolutional neural networks for mri reconstruction and phase-focused applications. *Magnetic resonance in medicine*, 86(2):1093–1109, 2021.
- [6] Antonia Creswell, Tom White, Vincent Dumoulin, Kai Arulkumaran, Biswa Sengupta, e Anil A Bharath. Generative adversarial networks: An overview. *IEEE signal processing magazine*, 35(1):53–65, 2018.
- [7] Amauri da Costa Júnior. Repositório do trabalho de mestrado. <https://github.com/AmauriCJr/Mestrado-MRI>. Accessed: 2025-02-21.
- [8] Salman UH Dar, Mahmut Yurt, Mohammad Shahdloo, Muhammed Emrullah Ildız, Berk Tinaz, e Tolga Çukur. Prior-guided image reconstruction for accelerated multi-contrast mri via generative adversarial networks. *IEEE Journal of Selected Topics in Signal Processing*, 14(6):1072–1087, 2020.
- [9] Nikhil Dhengre e Saugata Sinha. K sparse autoencoder-based accelerated reconstruction of magnetic resonance imaging. *The Visual Computer*, 38(3):837–847, 2022.

- [10] Allen D. Elster. Methods for filling k-space. <https://mriquestions.com/k-space-trajectories.html>. Accessed: 2024-03-01.
- [11] Allen D. Elster. MR quality control: SNR. How do you measure signal-to-noise in an image? <https://mriquestions.com/signal-to-noise.html>. Accessed: 2024-03-01.
- [12] FastMRI. Public leaderboard. https://web.archive.org/web/20230324102125mp_/https://fastmri.org/leaderboards. Accessed: 2025-02-21.
- [13] JP Felmlee, RL Morin, JR Salutz, e GB Lund. Magnetic resonance imaging phase encoding: a pictorial essay. *radiographics*, 9(4):717–722, 1989.
- [14] Li Feng, Thomas Benkert, Kai Tobias Block, Daniel K Sodickson, Ricardo Otazo, e Hersh Chandarana. Compressed sensing for body MRI. *Journal of Magnetic Resonance Imaging*, 45(4):966–987, 2017.
- [15] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, e Yoshua Bengio. Generative adversarial nets. *Advances in neural information processing systems*, 27, 2014.
- [16] Google. GPU architecture. https://colab.research.google.com/github/d2l-ai/d2l-tvm-colab/blob/master/chapter_gpu_schedules/arch.ipynb. Accessed: 2024-05-07.
- [17] Kjell Hansson Mild, Ronnie Lundström, e Jonna Wilén. Non-ionizing radiation in swedish health care—exposure and safety aspects. *International journal of environmental research and public health*, 16(7):1186, 2019.
- [18] NYU Langone Health. Welcome to the fastmri dataset. <https://fastmri.med.nyu.edu/>. Accessed: 2025-04-04.
- [19] Seyed Amir Hosseini, Chi Zhang, Sebastian Weingärtner, Steen Moeller, Matthias Stuber, Kamil Ugurbil, e Mehmet Akçakaya. Accelerated coronary mri with sraki: A database-free self-consistent neural network k-space reconstruction for arbitrary undersampling. *Plos one*, 15(2):e0229418, 2020.
- [20] Chang Min Hyun, Hwa Pyung Kim, Sung Min Lee, Sungchul Lee, e Jin Keun Seo. Deep learning for undersampled mri reconstruction. *Physics in Medicine & Biology*, 63(13):135007, 2018.
- [21] Tauhidul Islam, Md Sadman Hafiz, Jamin Rahman Jim, Md Mohsin Kabir, e MF Mridha. A systematic review of deep learning data augmentation in medical

- imaging: Recent advances and future research directions. *Healthcare Analytics*, page 100340, 2024.
- [22] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, e Alexei A Efros. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, páginas 1125–1134, 2017.
- [23] M. C. Keuken, P.-L. Bazin, K. Backhouse, S. Beekhuizen, L. Himmer, A. Kandola, J. J. Lafeber, L. Prochazkova, A. Trutti, A. Schäfer, R. Turner, e B. U. Forstmann. Effects of aging on T1, T2, and QSM MRI values in the subcortex. *Brain Structure and Function*, 222(6):2487–2505, 2017.
- [24] Hadelin de Ponteves Kirill Eremenko. Machine Learning A-Z: AI, Python and R. <https://www.udemy.com/course/machinelearning/>. Accessed: 2024-01-24.
- [25] Steffen N Krieger, Laurentius Huber, Benedikt A Poser, Robert Turner, e Gary F Egan. Simultaneous acquisition of cerebral blood volume-, blood flow-, and blood oxygenation-weighted MRI signals at ultra-high magnetic field. *Magnetic Resonance in Medicine*, 74(2):513–517, 2015.
- [26] Qiegen Liu, Qingxin Yang, Huitao Cheng, Shanshan Wang, Minghui Zhang, e Dong Liang. Highly undersampled magnetic resonance imaging reconstruction using autoencoding priors. *Magnetic resonance in medicine*, 83(1):322–336, 2020.
- [27] Alessandro A Mazzola. Ressonância magnética: princípios de formação da imagem e aplicações em imagem funcional. *Revista brasileira de física médica*, 3(1):117–129, 2009.
- [28] Reuben Mezrich. A perspective on k-space. *Radiology*, 195(2):297–315, 1995.
- [29] Javier Montalt-Tordera, Vivek Muthurangu, Andreas Hauptmann, e Jennifer Anne Steeden. Machine learning in magnetic resonance imaging: image reconstruction. *Physica Medica*, 83:79–87, 2021.
- [30] OECD. *Health at a Glance 2023*. 2023.
- [31] Karen Pallarito. 6 key medical scans and what they should cost. <https://www.health.com/mind-body/6-key-medical-scans-and-what-they-should-cost>. Published by Health Magazine. Accessed: 2024-01-24.
- [32] Precedence Research. Artificial intelligence in MRI market. <https://www.precedenceresearch.com/artificial-intelligence-in-mri-market>. Accessed: 2025-04-01.

- [33] Olaf Ronneberger, Philipp Fischer, e Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III* 18, páginas 234–241. Springer, 2015.
- [34] Adriana Schulz, Luiz Velho, e Eduardo AB Da Silva. On the empirical rate-distortion performance of compressive sensing. In *2009 16th IEEE international conference on image processing (ICIP)*, páginas 3049–3052. IEEE, 2009.
- [35] Suraj D Serai, Mai-Lan Ho, Maddy Artunduaga, Sherwin S Chan, e Govind B Chavhan. Components of a magnetic resonance imaging system and their relationship to safety and image quality. *Pediatric radiology*, 51:716–723, 2021.
- [36] Greg J Stanisz, Ewa E Odrobina, Joseph Pun, Michael Escaravage, Simon J Graham, Michael J Bronskill, e R Mark Henkelman. T1, T2 relaxation and magnetization transfer in tissue at 3T. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 54(3):507–512, 2005.
- [37] TensorFlow. Dropout. https://www.tensorflow.org/api_docs/python/tf/keras/layers/Dropout. Accessed: 2025-03-03.
- [38] TensorFlow. Regularizer. https://www.tensorflow.org/api_docs/python/tf/keras/Regularizer. Accessed: 2025-03-03.
- [39] Angel Torrado-Carvajal, Joaquin L Herraiz, Juan A Hernandez-Tamames, Raul San Jose-Estepar, Yigitcan Eryaman, Yves Rozenholc, Elfar Adalsteinsson, Lawrence L Wald, e Norberto Malpica. Multi-atlas and label fusion approach for patient-specific MRI based skull estimation. *Magnetic resonance in medicine*, 75(4):1797–1807, 2016.
- [40] Matthias Utzschneider, Nicolas GR Behl, Sebastian Lachner, Lena V Gast, Andreas Maier, Michael Uder, e Armin M Nagel. Accelerated quantification of tissue sodium concentration in skeletal muscle tissue: quantitative capability of dictionary learning compressed sensing. *Magnetic Resonance Materials in Physics, Biology and Medicine*, 33(4):495–505, 2020.
- [41] Maxim Zaitsev, Julian Maclare, e Michael Herbst. Motion artifacts in MRI: A complex problem with many partial solutions. *Journal of Magnetic Resonance Imaging*, 42(4):887–901, 2015.
- [42] Jure Zbontar, Florian Knoll, Anuroop Sriram, Tullie Murrell, Zhengnan Huang, Matthew J Muckley, Aaron Defazio, Ruben Stern, Patricia Johnson, Mary Bruno, et al. fastMRI: An open dataset and benchmarks for accelerated MRI. *arXiv preprint arXiv:1811.08839*, 2018.