



**UNIVERSIDADE ESTADUAL DA PARAÍBA - UEPB
NÚCLEO DE TECNOLOGIAS ESTRATÉGICAS EM SAÚDE - NUTES
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA E TECNOLOGIA EM SAÚDE**

AUDEMAR FERNANDES RIBEIRO JÚNIOR

**ATENUAÇÃO DE ARTEFATO METÁLICO EM IMAGEM DE TOMOGRAFIA
COMPUTADORIZADA COM USO DE *DEEP LEARNING***

DISSERTAÇÃO

**CAMPINA GRANDE
2018**

AUDEMAR FERNANDES RIBEIRO JÚNIOR

**ATENUAÇÃO DE ARTEFATO METÁLICO EM IMAGEM DE TOMOGRAFIA
COMPUTADORIZADA COM USO DE *DEEP LEARNING***

Dissertação submetida à Coordenação do Curso de Mestrado em Ciências e Tecnologias em Saúde da Universidade Estadual da Paraíba (UEPB), Campus I, como parte dos requisitos necessários para obtenção do grau de Mestre.

Área de concentração: Tecnologia em Saúde.

Orientador: Prof. Dr. Robson Pequeno de Sousa.

CAMPINA GRANDE, PB - Brasil
2018

É expressamente proibido a comercialização deste documento, tanto na forma impressa como eletrônica. Sua reprodução total ou parcial é permitida exclusivamente para fins acadêmicos e científicos, desde que na reprodução figure a identificação do autor, título, instituição e ano do trabalho.

R484a Ribeiro Junior, Audemar Fernandes.
Atenuação de artefato metálico em imagem de tomografia computadorizada com uso de *Deep learning* [manuscrito] / Audemar Fernandes Ribeiro Junior. - 2018.
81 p. : il. colorido.
Digitado.
Dissertação (Mestrado em Profissional em Ciência e Tecnologia em Saúde) - Universidade Estadual da Paraíba, Pró-Reitoria de Pós-Graduação e Pesquisa , 2019.
"Orientação : Prof. Dr. Robson Pequeno de Sousa , Departamento de Computação - CCT."
1. Aprendizagem profunda. 2. Rede Neural Convolucional.
3. Artefatos metálicos. 4. Deep learning. 5. Tomografias computadorizadas. I. Título
21. ed. CDD 600

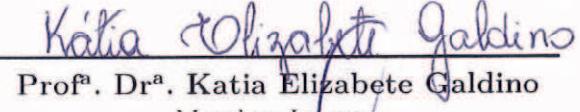
AUDEMAR FERNANDES RIBEIRO JÚNIOR

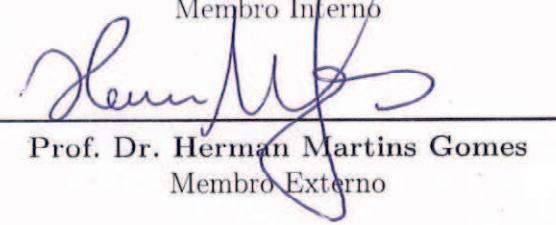
**ATENUAÇÃO DE ARTEFATO METÁLICO EM
IMAGEM DE TOMOGRAFIA COMPUTADORIZADA
COM USO DE *DEEP LEARNING***

Dissertação submetida à Coordenação do
Curso de Mestrado em Ciências e Tecnolo-
gias em Saúde da Universidade Estadual da
Paraíba (UEPB), Campus I, como parte dos
requisitos necessários para obtenção do título
de Mestre.

Aprovada em 26/11/2018


Prof. Dr. Robson Pequeno de Sousa
Orientador


Profª. Drª. Katia Elizabeth Galdino
Membro Interno


Prof. Dr. Herman Martins Gomes
Membro Externo

Campina Grande, PB - Brasil

2018

*Este trabalho é dedicado ao meu falecido pai,
que tanto demonstrou em vida, a importância do conhecimento.*

Agradecimentos

Agradeço ao Professor Dr. Robson Pequeno de Sousa, pelas orientações e direcionamentos de pesquisa e desenvolvimento, bem como, a todo o Programa de Pós-Graduação em Ciência e Tecnologia em Saúde, da UEPB. Aos professores Dra. Katia Galdino e Dr. Herman Gomes, por tão atenciosamente apreciarem este trabalho.

Agradeço principalmente à minha esposa Hilmaria Xavier, pelo incentivo, dedicação e suporte, sendo fundamental para a finalização deste trabalho.

Agradecimentos aos colegas e amigos, que tanto contribuíram para a saúde mental e leveza de espírito. Ao Professor Edmar Candeia, pelo valoroso auxílio, ao jovem amigo Andresso da Silva, cujo olhar externo alertou para o caminho a ser seguido e à família Xavier, que tão carinhosamente me acolheu.

Agradeço aos colegas de trabalho Roberta Pereira, Diego Conserva, Robson Rodrigues e José Marçal, pelo suporte e auxílio nessa jornada.

“Parte da ausência de humanidade do computador deve-se a que, competentemente programado e trabalhando bem, é completamente honesto.“

Isaac Asimov

Resumo

Artefatos metálicos são graves problemas de reconstrução de imagens médicas de tomografias computadorizadas, podendo causar significativas dificuldades de diagnóstico, ou até mesmo inviabilizá-lo. Esta pesquisa, desenvolvida no Núcleo de Tecnologias Estratégicas em Saúde - NUTES, em parceria com o Hospital Regional de Trauma Dom Luiz Gonzaga Fernandes, gerou uma arquitetura de rede neural convolucional que atenua os efeitos causados por artefatos metálicos em imagens produzidas por aparelhos de tomografia computadorizada. A arquitetura proposta, após o treinamento, foi capaz de reconstruir imagens de tomografia computadorizada com atenuação de artefato metálico, com resultado aferido por avaliação qualitativa através de formulários. Os resultados obtidos demonstraram que a técnica é promissora para a atenuação de artefatos metálicos.

Palavras-chave: aprendizagem profunda. rede neural convolucional. aprendizado de máquina. artefato metálico.

Abstract

Metallic artifacts are serious problems of reconstruction of medical images of CT scans, which may cause significant difficulties in diagnosis, or even render it unfeasible. This research, developed in the Nucleus of Strategic Technologies in Health - NUTES, in partnership with the Regional Hospital of Trauma Dom Luiz Gonzaga Fernandes, generated a convolutional neural network architecture that attenuates the effects caused by metallic artifacts in images produced by computed tomography . The proposed architecture, after training, was able to reconstruct computed tomography images with metal artifact attenuation, with a result measured by qualitative evaluation through forms. The obtained results demonstrated that the technique is promising for the attenuation of metallic artifacts.

Keywords: deep learning. convolutional neural network. machine learning. metal artifact.

Listas de ilustrações

Figura 1 – Representação dos níveis de absorção dos raios-x, de acordo com a densidade. (DAFFNER; HARTMAN, 2014)	15
Figura 2 – (i) Esquema de aparelho de raio-x padrão e (ii) foco do feixe de elétrons. Adaptado de (DOUGUERTY, 2009)	18
Figura 3 – Representação de aparelho de tomografia. Adaptado de (LISLE, 2007)	19
Figura 4 – representação normal da reta, de acordo com a equação 2.5 (GONZALEZ; WOODS, 2017b)	22
Figura 5 – Geometria de um feixe paralelo. Adaptado de (GONZALEZ; WOODS, 2017a)	23
Figura 6 – Artigos selecionados pela pesquisa e separados por anos.	24
Figura 7 – Artigos selecionados pela pesquisa e separados pela tecnologia empregada.	24
Figura 8 – a) Diagrama esquemático da aquisição da tomografia computadorizada e b) Sinograma resultante $P(\varphi, s)$. (PARK; CHOI; SEO, 2014)	25
Figura 9 – Arquitetura com uso de U-Net.(JIN et al., 2017)	27
Figura 10 – a) TC pélvica mostrando artefato de anel severo. b) TC principal com artefato de anel sutil, simulando uma lesão (indicada pela seta). c) Alterações nas configurações de nível de intensidade e janela de abertura, evidenciam a região de reconstrução, situada no centro de rotação (anel azul) (adaptado de (BOAS; FLEISCHMANN, 2012))	28
Figura 11 – a) A imagem de TC de baixa dose obtida durante uma biópsia guiada por TC mostra um amplo ruído de Poisson b) A imagem pós-biópsia obtida com uma dose 7,3 vezes maior e apresentou $\sqrt{7,3} = 2,7$ vezes menos ruído. As imagens mostram um nódulo linfático retroperitoneal ampliado (flecha) e infiltração do rim direito em paciente com linfoma de Hodgkin. Adaptado de (BOAS; FLEISCHMANN, 2012)	28
Figura 12 – Simulação de varredura sem endurecimento do feixe (linha superior) e com endurecimento do feixe (linha inferior), mostrando que as faixas escuras ocorrem ao longo das linhas de maior atenuação, e as raias brilhantes ocorrem em outras direções (adaptado de (BOAS; FLEISCHMANN, 2012))	29
Figura 13 – Detecção de projeções afetadas por artefatos metálicos: a) imagem TC original, b) artefato metálico em imagem segmentada através de limiar, c) sinograma original, d) sinograma apenas do metal e e) dados dispersos. (ABDOLI et al., 2011)	32
Figura 14 – Influência na escolha de k no estimador dos k vizinhos mais próximos. (IZBICKI, 2016)	38

Figura 15 – Ilustração dos limites de decisão (SHALEV-SHWARTZ S, 2014).	39
Figura 16 – O hiperplano é representado pela linha sólida ao centro, as margens são representadas pelas linhas pontilhadas e a distância dos pontos ao hiperplano são representadas por setas (JAMES, 2015).	40
Figura 17 – Representação da estrutura de uma Árvore de Regressão (IZBICKI, 2016).	40
Figura 18 – representação de neurônio artificial	43
Figura 19 – Modelo de neurônio artificial (HAYKIN, 1998)	43
Figura 20 – função de ativação limiar (HAYKIN, 1998)	44
Figura 21 – função de ativação linear por partes (HAYKIN, 1998)	45
Figura 22 – função de ativação sigmoide (HAYKIN, 1998)	45
Figura 23 – Rede neural artificial multi-camadas (HAYKIN, 1998)	46
Figura 24 – No quadro da esquerda, uma imagem hipotética representada por um único canal com dimensões 5x5 que recebe a aplicação de um filtro 3x3. No quadro da direita, uma matriz ilustrando o somatório que fornece o resultado da convolução (adaptado de (FERREIRA, 2017))	52
Figura 25 – Representação das amostras da base de dados: cada coluna corresponde a um caso. Adaptado de (ZHANG; YU, 2017)	56
Figura 26 – Modelagem da arquitetura da rede neural convolucional para redução de artefato metálico (ZHANG; YU, 2017))	56
Figura 27 – A figura (a) sem artefato metálico e a figura (b), com artefato inserido artificialmente.	59
Figura 28 – Figura 27 (b), fragmentada em trechos de 64x64 <i>pixels</i>	60
Figura 29 – Arquitetura do modelo proposto de rede neural convolucional	60
Figura 30 – Treinamento do modelo proposto, época 500. a) imagem de treino, b) saída da rede e c), imagem desejada.	61
Figura 31 – Treinamento do modelo proposto, época 8292. a) imagem de treino, b) saída da rede e c), imagem desejada.	62
Figura 32 – Treinamento do modelo proposto, época 15536. a) imagem de treino, b) saída da rede e c), imagem desejada.	62
Figura 33 – Treinamento do modelo proposto, época 17148. a) imagem de treino, b) saída da rede e c), imagem desejada.	63
Figura 34 – Treinamento do modelo proposto, época 17784. a) imagem de treino, b) saída da rede e c), imagem desejada.	63
Figura 35 – Erro Quadrático Médio do treinamento com taxa de aprendizagem $l = 0.0001$	64
Figura 36 – Erro Quadrático Médio do treinamento com taxa de aprendizagem $l = 0.001$	64

Figura 37 – Erro Quadrático Médio do treinamento com taxa de aprendizagem $l = 0.01$	65
Figura 38 – Resultado final da rede neural proposta, com taxa de aprendizagem $l = 0.0001$. a) imagem com artefato real e b), imagem reconstruída pela rede.	66
Figura 39 – Resultado final da rede neural proposta, com taxa de aprendizagem $l = 0.0001$. a) imagem com artefato real e b), imagem reconstruída pela rede.	66
Figura 40 – Resultado final da rede neural proposta, com taxa de aprendizagem $l = 0.001$. a) imagem com artefato real e b), imagem reconstruída pela rede.	67
Figura 41 – Resultado final da rede neural proposta, com taxa de aprendizagem $l = 0.001$. a) imagem com artefato real e b), imagem reconstruída pela rede.	67
Figura 42 – Resultado final da rede neural proposta, com taxa de aprendizagem $l = 0.01$. a) imagem com artefato real e b), imagem reconstruída pela rede.	68
Figura 43 – Detalhe do resultado final da rede neural a) imagem com artefato real e b) imagem reconstruída pela rede	68
Figura 44 – Fonte: (KOEHLER; BRENDEL; BROWN, 2012)	69

Sumário

1	INTRODUÇÃO	14
2	FUNDAMENTAÇÃO TEÓRICA	17
2.1	Aquisição de Imagens Médicas	17
2.2	Pesquisas em Atenuação de Artefatos Metálicos	22
2.3	Caracterização dos Artefatos Metálicos em Tomografia Computadorizada	26
2.3.1	Artefato de Anel (<i>Ring Artifact</i>)	27
2.3.2	Ruído	27
2.3.3	Endurecimento e Dispersão do Feixe	29
2.4	Métodos de Atenuação de Artefatos Metálicos	30
2.4.1	Otimização de Implantes de Metal	30
2.4.2	Melhoria de Aquisição	30
2.4.3	Pré-processamento Baseado em Física	30
2.4.4	Ajuste de Projeção	31
2.4.5	Pós-processamento de Imagem	32
2.4.6	Reconstrução Iterativa	33
3	FUNDAMENTOS DE APRENDIZAGEM DE MÁQUINA	35
3.1	Modelos Supervisionados	36
3.1.1	Régressão	36
3.1.2	Métodos Paramétricos	36
3.1.3	Métodos Não Paramétricos	37
3.1.3.1	K Nearest Neighbor	37
3.1.3.2	Support Vector Machine - SVM	39
3.1.3.3	Árvores de Régressão	39
3.1.4	Deep Learning	41
3.1.4.1	Introdução	41
3.1.4.2	Redes Neurais Convolucionais	47
4	METODOLOGIA	54
4.1	Introdução	54
4.2	Treinamento da RNC	55
4.3	Modelo Proposto	58
5	RESULTADOS OBTIDOS	61

6	CONSIDERAÇÕES FINAIS	73
	REFERÊNCIAS	75
	APÊNDICES	80
	APÊNDICE A –	81

1 Introdução

O uso de computação está diretamente associado aos mais modernos métodos de aquisição de imagens médicas. São conjuntos de técnicas ou procedimentos através dos quais é possível obter imagens de um organismo vivo, ou parte dele. Essas imagens de tecidos e funções orgânicas do corpo humano permitem avaliações médicas não invasivas. As visualizações do funcionamento do corpo humano ocorrem por intermédio do processamento dos sinais obtidos com equipamentos diversos de captura de imagens, sendo as principais modalidades de obtenção de imagens médicas utilizadas atualmente: Ultrassom, Ressonância Magnética, Raios X, Tomografia Computadorizada (CT) e PET (*Positron Emissor Tomography*). Os métodos que utilizam radiação ionizante, têm como precursor o Raio X, descoberto por Roentgen em 1895 (KHAN, 2010), pelo que foi agraciado em 1901, com o Prêmio Nobel de Física. As técnicas mais modernas de processamento de imagens permitem ainda tarefas além da aquisição, realce, armazenamento, análise, compressão e visualização, sendo comum, atualmente, a disposição de imagens em um ambiente de representação 3D. Os aparelhos de aquisição de imageamento médico, como Tomografias Computadorizadas, objeto principal deste trabalho, utilizam em seu processo, a captação do sinal do feixe de raio-x, direcionado a cada fatia do objeto de estudo e a representação dessa imagem é a projeção bidimensional do objeto tridimensional.

No processo de aquisição da imagem resultante de um exame de tomografia computadorizada são capturadas várias imagens de raio-x, alterando-se o plano focal e o ângulo, dependendo do aparelho utilizado. A aquisição dos dados que formarão a imagem se dá pela excitação de um sensor após a emissão de raio-x por uma fonte. O sinal recebido pelo sensor é atenuado tanto pelo caminho percorrido, quanto pela densidade do material do objeto em estudo, conforme representado na Figura 1. As imagens geradas por esse processo possuem seu formato em escala de cinza, pois a intensidade do brilho de cada ponto representa a densidade, variando os valores de acordo com o ponto analisado, por exemplo, ar e tecidos moles (objetos menos atenuantes) possuem representações mais escuras, enquanto ossos e metais (objetos mais atenuantes), representações em tons de cinza mais claros.

No aparelho de tomografia computadorizada, redes de sensores dispostos no plano focal do emissor, captam as projeções dos raios-x que atravessam o mesmo corte em diferentes posições, devido à rotação do conjunto de captação do aparelho, gerando assim, diferentes intensidades do sinal, de acordo com a atenuação que a radiação sofre, ao atravessar o objeto e é captada pelo sensor por determinado tempo, formando uma matriz. Essa matriz passa por um processo de quantização em uma escala discreta, para então poder ser representada como uma imagem bidimensional. Dessa forma, são criadas representações

bidimensionais que são projeções planas de objetos tridimensionais, ocorrendo no entanto, uma considerável perda de informação (DOUGUERTY, 2009). A partir de então, se faz necessária a utilização de um método reconstrutivo da imagem, tendo como origem suas projeções, solução proposta por Johann Radon, em 1917, que mostrou ser possível reconstruir uma imagem bidimensional a partir de conjuntos de perfis de intensidade (BIRKFELLNER, 2015).

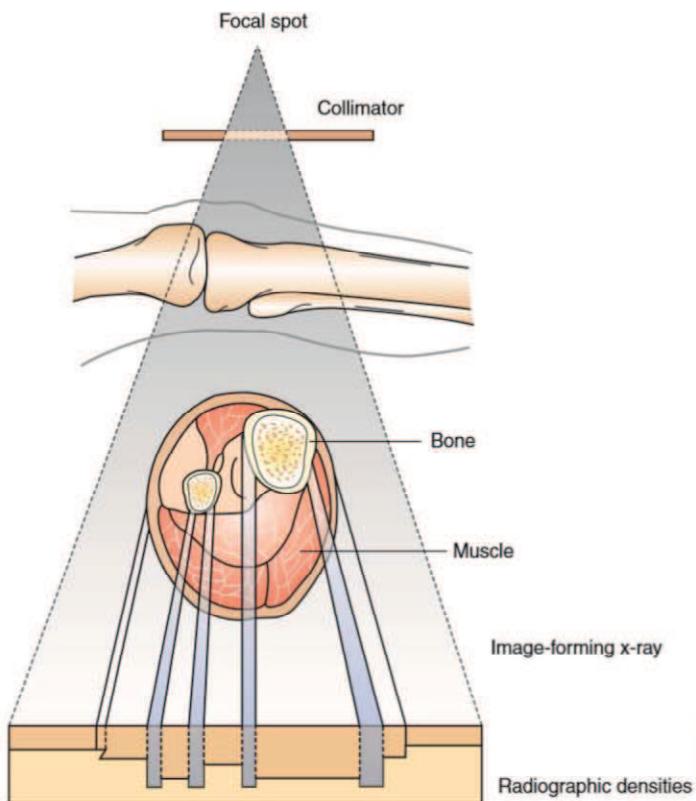


Figura 1 – Representação dos níveis de absorção dos raios-x, de acordo com a densidade. (DAFFNER; HARTMAN, 2014)

Os problemas mais graves na tomografia computadorizada ocorrem com os habituais implantes metálicos e obturações dentárias, pois o metal utilizado nas restaurações causam ruídos na projeção, aqui nomeados como *artefatos*. Na presença de inserções metálicas, fótons de raios-X não podem penetrar o objeto metálico de forma consistente, devido à densidade do metal, ocasionando dispersão e ruído. A inconsistência das projeções introduz graves estrias e sombreamentos nas projeções, prejudicando a reconstrução da imagem.

Desta forma, visando a reduzir os prejuízos causados pela presença de artefatos metálicos nas imagens, este trabalho se desenvolve, buscando a atenuação dos referidos ruídos, utilizando-se de técnicas de aprendizagem profunda. Para isso, uma rede neural profunda será alimentada e treinada com pares de imagens que representem o mesmo objeto de estudo. Uma das imagens possuirá a inconsistência dos efeitos do artefato em sua reconstrução, enquanto a outra será a imagem desejada, formando assim um par

composto por uma imagem ruidosa e uma imagem desejada que irão alimentar a rede neural e treiná-la. Desta forma, após o treinamento, a rede deverá reconstruir imagens de tomografia computadorizada com artefatos atenuados, partindo de imagens inicialmente com ruídos, de acordo com a generalização a que for submetida.

Objetivos Gerais

Projetar um modelo de arquitetura de redes neurais convolucionais que atenuem artefatos metálicos em imagens de tomografia computadorizada.

Objetivos Específicos

- Criar uma base de dados contendo pares de imagens contendo artefatos metálicos e imagens sem artefatos, para alimentar o treinamento da rede neural;
- Desenvolver um modelo de rede neural convolucional para atenuação de artefatos metálicos;
- Comparar os resultados obtidos com o referencial teórico.

2 Fundamentação Teórica

Neste capítulo são abordados conceitos relacionados ao problema cuja pesquisa se propõe a contribuir. São descritos conceitos básicos sobre a aquisição de imagens tomográficas, tipos de artefatos metálicos e suas características, bem como métodos para atenuação desses artefatos.

2.1 Aquisição de Imagens Médicas

Na maioria das situações, imagens médicas são representações das propriedades físicas de um tecido ao ser exposto a um certo tipo de radiação eletromagnética ou mecânica (no caso do ultrassom) e essa radiação eletromagnética é composta por objetos quânticos chamados fótons, cuja energia é descrita pela equação 2.1 (BIRKFELLNER, 2015).

$$E = h\nu \quad (2.1)$$

Temos que h é a Constante de Planck, dada por $6,62607004 \times 10^{-34} Js$ e a variável ν é a frequência dos fótons em Hertz. Interessante observar que como Hertz é medida em tempo, de modo que $1Hz = \frac{1}{s}$, logo, o produto da ação e da frequência é a energia $\frac{Js}{s} = J$. Na tabela abaixo, estão elencadas algumas frequências de radiação eletromagnética.

Tabela 1 – Classificação das radiações eletromagnéticas por frequência

Frequência	Utilização
01 Hz a 10^4 Hz	Corrente Alternada (AC): utilizada apenas em tomografias por impedância elétrica.
10^4 Hz a 10^8 Hz	Rádiofrequência: utilizada em Ressonância Magnética.
10^8 Hz a 10^{12} Hz	Micro-ondas: não tem relevância para aplicações médicas.
10^{12} Hz a 7×10^{14} Hz	Infravermelho: utilizado em aplicações médicas específicas
$4,6 \times 10^{14}$ Hz a $6,6 \times 10^{14}$ Hz	Faixa de luz visível: utilizada em microscopia e endoscopia.
4×10^{14} Hz a 10^{18} Hz	Ultravioleta (UV).
10^{18} Hz a 10^{20} Hz	Raios-x.
acima de 10^{20} Hz	Radiação- γ : faixa utilizada normalmente com produtos radioativos

Fonte: Birkfellner, 2015 - Applied Medical Image Processing

Raios-x são gerados em um tubo a vácuo que possui um catodo e um anodo, conforme Figura 2 (i). O aquecimento de um filamento de tungstênio no catodo libera

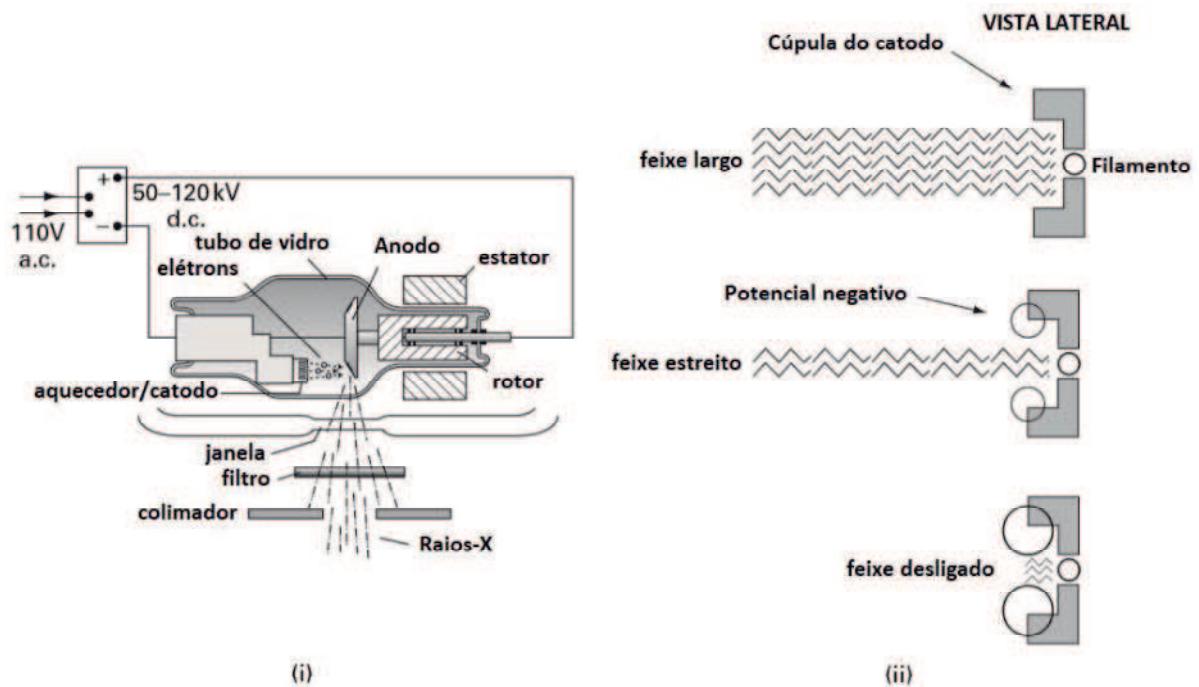


Figura 2 – (i) Esquema de aparelho de raio-x padrão e (ii) foco do feixe de elétrons.
Adaptado de (DOUGUERTY, 2009)

elétrons por excitação térmica. O filamento está localizado dentro de uma depressão ou cúpula com bordas contornadas e pontiagudas que focam eletrostaticamente o feixe de elétrons, de acordo com Figura 2 (ii). Tensões cada vez mais negativas aplicadas à cúpula do cátodo podem focalizar os elétrons em um raio estreito ou mesmo desligar completamente o feixe. Os elétrons são acelerados em direção ao anodo positivo (50-120 kV), onde atingem um alvo de tungstênio incorporado, produzindo os raios-x. A escolha do tungstênio, se deu pelo alto número atômico e de ponto de fusão (KHAN; GIBBONS, 2014). O motivo do tubo ser hermeticamente selado, é para que os elétrons passem para o ânodo em um caminho reto e não sejam espalhados por outras partículas no tubo. A maioria das unidades de raios-x móveis e dentárias de baixa potência, empregam anodos giratórios para que os elétrons atinjam uma área maior ao redor do aro do ânodo e não aqueçam demais o objeto de estudo. A região do alvo a partir da qual os raios X são produzidos é chamada de ponto focal e seu diâmetro é conhecido como o tamanho do ponto (DOUGUERTY, 2009).

A frequência de raios-x é amplamente utilizada no campo de imagens médicas, sendo utilizada tanto em aparelhos de raios-x, quanto em aparelhos de tomografia computadorizada. Em uma imagem de raio-x, o objeto a ser estudado atenua a emissão da radiação, conforme sua densidade, de forma que uma estrutura mais clara na imagem é fruto de uma atenuação mais intensa (ossos, por exemplo), enquanto que partes mais escuras da imagem, são representações de menor atenuação (tecido mole, por exemplo). Em uma radiografia convencional, o volume tridimensional de um corpo é projetado ao longo

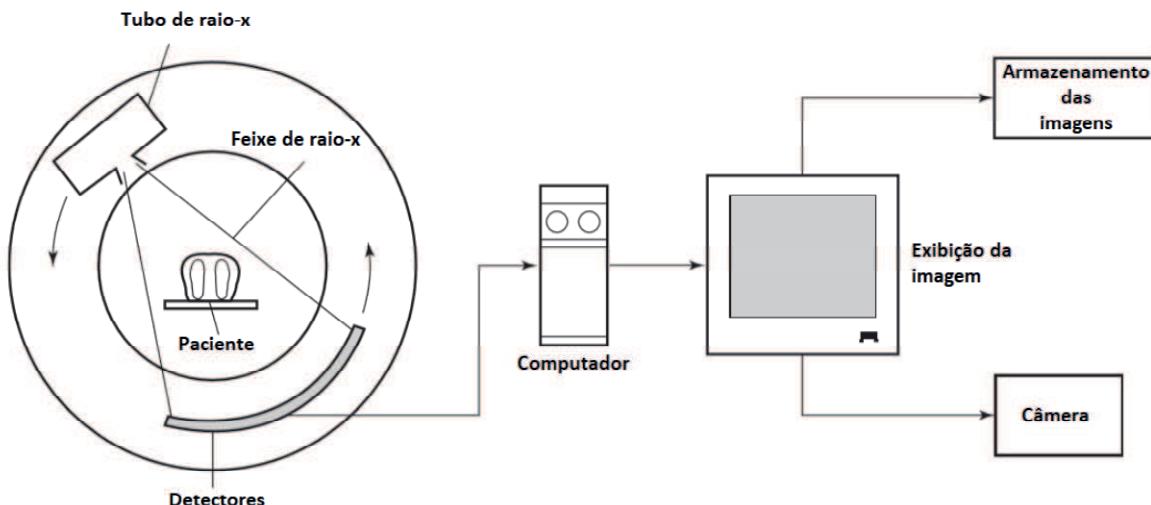


Figura 3 – Representação de aparelho de tomografia. Adaptado de (LISLE, 2007)

da direção do raio-x, sendo representado em uma imagem bidimensional, de forma que tecidos moles e estrutura óssea são sobrepostos na imagem, causando redução dos detalhes dos objetos a serem estudados. Essas limitações da radiografia convencional levaram ao desenvolvimento do tomógrafo.

A Tomografia Computadorizada (TC) é uma técnica de imagem em que imagens transversais são obtidas com o uso de raios-X. O paciente passa pelo aparelho, cujo conjunto emissor/detector gira em torno do nível de interesse de estudo. O conjunto emissor/detector do aparelho é composto por um tubo de raios X de um lado e um conjunto de detectores do outro. A informação dos detectores é analisada por computador e exibido como uma imagem, conforme representação na Figura 3. Devido ao uso de análise por computador, uma variedade muito maior de densidades pode ser exibida do que em filmes de raios-x convencionais, permitindo a diferenciação de órgãos sólidos um do outro e de processos patológicos, como tumores ou fluidos, além de ser sensível à presença de pequenas quantidades de gordura, cálcio ou material de contraste (LISLE, 2007).

Tomografia Computadorizada é descrita por BirkFellner (2015), como o processo de aquisição de imagens através da emissão de um feixe de raio-x colimado, ou seja, várias partículas com trajetórias paralelas, que penetra uma fatia do corpo do paciente, havendo do lado oposto ao emissor, um detector que medirá a intensidade do feixe residual atenuado, conforme a equação de atenuação (2.2). Após a aquisição dessas informações, o aparelho realiza uma pequena rotação, normalmente um grau, de modo que o emissor e o receptor se posicionem novamente para a realização de outras medições (BIRKFELLNER, 2015). Assim, várias projeções em diferentes direções são armazenadas, permitindo sua

reconstrução.

$$I(x) = I_0 \exp(-\mu x) \quad (2.2)$$

onde I_0 é o número de fótons incidentes no material por unidade de tempo, I é a intensidade de fótons que saem do material que é dado em unidades de absorção molar $Lmol^{-1}cm^{-1}$ (ou $M^{-1} \cdot cm^{-1}$), x é a distância percorrida pelo feixe e μ é o coeficiente de atenuação linear em caso do corpo ser homogêneo, ligado diretamente à densidade do material. Nos casos onde o corpo não é homogêneo, a equação (2.2) deverá ser alterada, assumindo a forma descrita na equação (2.3), a qual foi adicionada o somatório das densidades dos materiais que atenuaram o feixe colimado. Na tabela (2.1), estão dispostos os principais coeficientes de atenuações para estruturas do corpo humano.

$$I(x) = I_0 \exp\left(-\sum_{i=1}^n \mu_i x_i\right) \quad (2.3)$$

Tabela 2 – Valores para coeficientes de atenuações

Material	Coeficiente de atenuação linear
μ -ar	0
μ -sangue	$0,48cm^{-1}$
μ -músculo	$0,180cm^{-1}$
μ -osso	$0,178cm^{-1}$

Fonte: (WEBB, 1988)

Em Tomografias Computadorizadas, os valores correspondentes aos níveis de cinza, proporcionais ao coeficiente de atenuação são medidos em *Hounsfield Units* (HU), em homenagem a um dos pioneiros no desenvolvimento da tecnologia de TC, Godfrey Hounsfield. Essa unidade representa os valores de densidade da tomografia, sendo possível quantificar a atenuação do feixe, ao passar por um determinado objeto. Nessa escala, o ar tem valor de -1000 HU, enquanto que a água tem valor de 0 HU, servindo a água, como referência para cálculo do valor em HU de outros materiais com base no valor de seus coeficientes de atenuação (BIRKFELLNER, 2015). A equação 2.4 é a utilizada para esse cálculo de HU (DOUGUERTY, 2009):

$$ValorHU = 1000 \times \frac{(\mu - \mu_{H_2O})}{\mu_{H_2O}} \quad (2.4)$$

em que μ é igual ao coeficiente de atenuação do tecido estudado.

Na tabela abaixo, estão exemplificados valores de HU para alguns tecidos humanos:

Tabela 3 – Valores de HU para vários tecidos

Tecido	Valor em HU
Osso	acima de 1000
Músculo	44-59
Sangue	42-58
Coração	24
Água	0
Gordura	-20 a -100
Pulmão	-300
Ar	-1000

Fonte: (DOUGUERTY, 2009)

Na reconstrução da imagem, após a aquisição pelo Tomógrafo, como exposto anteriormente, é utilizada a Transformada de Radon, proposta pelo matemático Johann Radon, em 1917, sendo a forma mais comum de se realizar a reconstrução dessas imagens (BIRKFELLNER, 2015). O mesmo princípio é utilizado tanto para TC, quanto para PET (*Positron Emission Tomography*), MRI (*Magnetic Resonance Imaging*), SPECT (*Single Photon Emission Tomography*) e algumas modalidades de ultrassom (GONZALEZ; WOODS, 2017b).

O processo matemático demonstrado por Gonzales e Woods (2007), para a transformada de Radon, considera que uma linha reta em coordenadas cartesianas pode ser descrita por um conjunto de linhas compondo inclinações-intersecções na forma $y = ax + b$ ou por sua representação normal, conforme Figura 4:

$$x\cos\theta + y\sin\theta = \rho \quad (2.5)$$

A projeção de um feixe paralelo de raios pode ser descrita como um conjunto de linhas, conforme a Figura 5, onde um ponto arbitrário com as coordenadas (ρ_j, θ_k) , é dado pela soma dos raios ao longo da linha $x\cos\theta_k + y\sin\theta_k = \rho_j$. Considerando a continuidade, a soma dos raios da linha acima é dada pela integral:

$$g(\rho_j, \theta_k) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x\cos\theta_k + y\sin\theta_k - \rho_j) dx dy \quad (2.6)$$

Na equação 2.6, foi utilizada a propriedade do impulso, δ . O lado direito da equação é zero, a menos que o argumento de *delta* seja zero, de modo que a integral é calculada apenas ao longo da linha $x\cos\theta_k + y\sin\theta_k = \rho_j$. Ao se considerar todos os valores de ρ e θ ,

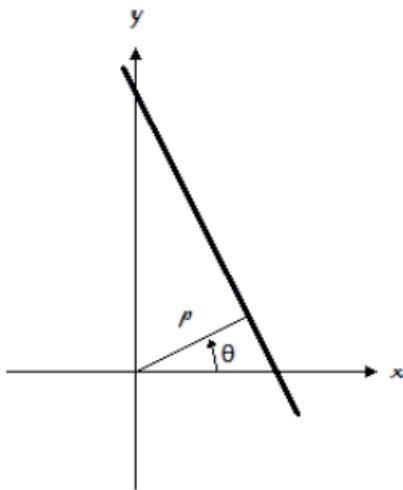


Figura 4 – representação normal da reta, de acordo com a equação 2.5 (GONZALEZ; WOODS, 2017b)

a equação passa a ser:

$$g(\rho, \theta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x \cos \theta + y \sin \theta - \rho) dx dy \quad (2.7)$$

A equação que fornece a projeção, ou integral linha de $f(x, y)$, ao longo de uma linha contínua arbitrária no plano xy , é a Transformada de Radon. Para a utilização da equação 2.7 para a reconstrução de imagens de TC, como aplicação em processamento de imagens, é necessário discretizá-la, conforme abaixo:

$$g(\rho, \theta) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) \delta(x \cos \theta + y \sin \theta - \rho) \quad (2.8)$$

em que x, y são variáveis discretas e M e N , são as dimensões da área retangular sobre a qual a transformada é aplicada. Se θ for fixo, enquanto ρ variar, então a equação 2.8 simplesmente realizará a soma dos *pixels* de $f(x, y)$, ao longo da linha definida e especificada pelos parâmetros θ e ρ .

2.2 Pesquisas em Atenuação de Artefatos Metálicos

No artigo de Gjesteby et al. (2016), há a descrição das principais práticas a respeito da atenuação de artefatos metálicos em imagens de tomografia computadorizada, contribuindo com sua coletânea, para o estado da arte do tema.

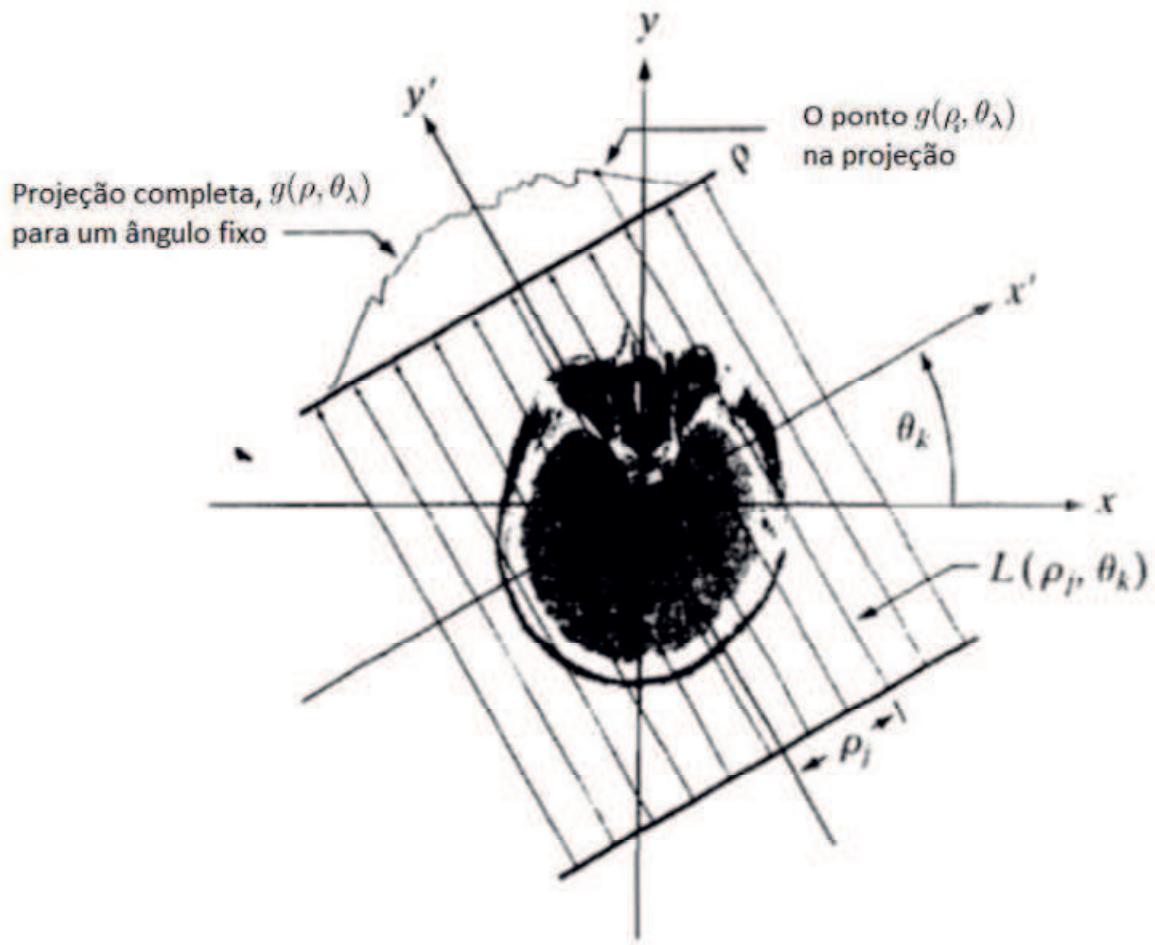


Figura 5 – Geometria de um feixe paralelo. Adaptado de (GONZALEZ; WOODS, 2017a)

Aliado a isso, foi realizada uma revisão bibliográfica com objetivo de se obter as últimas pesquisas na área. Foram pesquisados os repositórios: IEEE Xplore, Science Direct, arXiv, ACM Digital Library e Scopus. As pesquisas foram realizadas com utilização das seguintes palavras chaves: *deep learning*, *convolutional neural network*, *machine learning* e principalmente, *metal artifact*, ocorridos nos últimos 5 anos.

Foram selecionados, dentro desses parâmetros, 55 artigos, dos quais apenas 04 apresentavam contribuições significativas, enquanto 51 artigos foram descartados por apresentarem temática divergente da proposta desse trabalho.

Curioso observar, entre os artigos originalmente selecionados, um crescimento significativo da pesquisa na área, conforme a figura 6.

Os artigos exibidos pelas buscas, foram divididos conforme a tecnologia empregada para atingir seus objetivos, assim, Classificação e Reconstrução permaneceram com os maiores índices de utilização, seguidos pelas demais tecnologias, conforme consta na Figura 7. A utilização de aprendizagem profunda para classificação de imagens foi de 62%, condizendo com a maior utilização dessa tecnologia, inicialmente proposta para esse fim.



Figura 6 – Artigos selecionados pela pesquisa e separados por anos.

Reconstrução de imagens foi detectado em 21% dos artigos, restando as tecnologias de Segmentação e Otimização, com 10% e 7%, respectivamente. Neste trabalho, será utilizada a tecnologia de reconstrução por regressão.

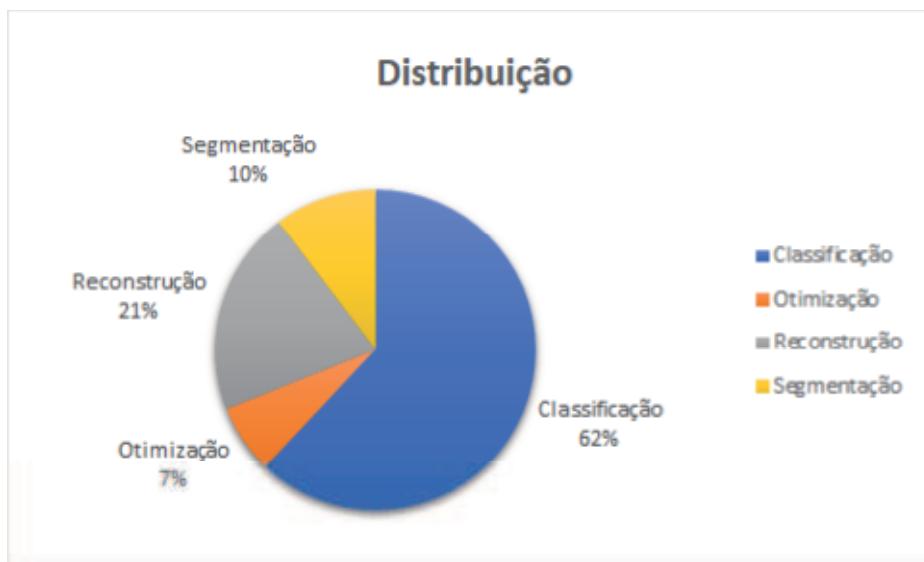


Figura 7 – Artigos selecionados pela pesquisa e separados pela tecnologia empregada.

O artigo de Zhang e Yu (2017) é de fundamental importância para o desenvolvimento deste trabalho. Os autores utilizam o uso de Redes Neurais Convolucionais - RNC (*Convolutional Neural Network - CNN*) para reconstrução de imagens de tomografia com baixas doses de raios-x, método já utilizado para redução de artefatos metálicos, nomeado pelos autores, como CNN-MAR (*Convolutional Neural Network - Metal Artifact Removal*). Para a realização do treinamento, foram utilizadas imagens livres de artefatos metálicos e artificialmente, foram introduzidos esses ruídos através de simulações, alterando essas imagens. Elas foram alteradas para que em alguns trechos representassem o ruído que se desejava tratar, no entanto, não foi utilizado apenas esse par de imagens.

Inicialmente foi aplicado nas imagens com ruído, um método de correção de endurecimento de feixe (*Beam Hardening Correction - BHC*) e interpolação linear (*Linear Interpolation - LI*), para então, juntamente com a imagem corrompida, alimentar a rede neural com três canais e realizar o treinamento supervisionado, possuindo a imagem sem ruído, como referência. A análise desse artigo é mais detalhada na sessão 4.3.

O artigo de Park, Choi e Seo(2017), descreve a caracterizaçāp matemática dos artefatos metálicos, como ocorrem e que efeitos causam no resultado das imagens reconstruídas. Inicialmente discorre sobre os princípios da propagação do raio-x e a forma da aquisição da Tomografia Computadorizada, como demonstra a representação esquemática na figura 8, onde consta em (a), um *phantom*, imagem de teste padrão para utilização em algoritmos de reconstrução (SHEPP; LOGAN, 1974), como objeto a ser estudado sob ação dos raios-x em rotação e em (b), a representação do sinograma criado pela aquisição do sinal. Essa imagem destaca o traço vertical no sinograma, correspondente ao ângulo de incidência do raio-x naquele instante. O trabalho segue demonstrando matematicamente o comportamento das estrias e oclusões dos artefatos, além de citar outras causas para o ruído. Na figura 8, está representada a projeção na tomografia computadorizada, os dados de projeção de raios $P(\varphi, s)$ são coletados para uma fatia após passar raios X em diferentes direções através de um objeto, onde s indica a posição da linha projetada e φ é o ângulo da projeção (PARK; CHOI; SEO, 2014).

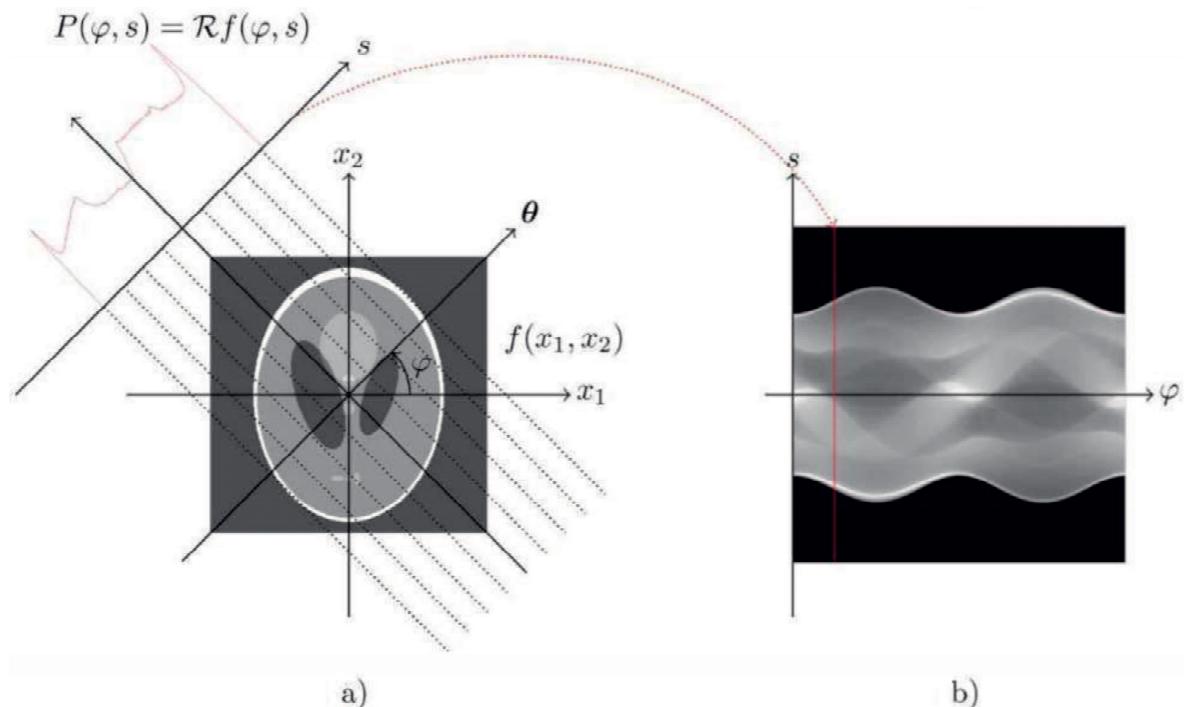


Figura 8 – a) Diagrama esquemático da aquisição da tomografia computadorizada e b)Sinograma resultante $P(\varphi, s)$. (PARK; CHOI; SEO, 2014)

Outros trabalhos contribuem para o tema, como o de Park et al.(2017), no qual foi

proposto um método com aprendizagem profunda, para atenuação de artefatos metálicos, em imagens de TC. No artigo, utilizaram a U-Net (RONNEBERGER; FIISCHER; BROX, 2015), arquitetura de rede convolucional para segmentação, reduzindo assim parte do custo computacional com o treinamento, uma vez que a utilização de uma "rede totalmente convolucional", utilizou poucas imagens, mas com segmentações precisas (PARK et al., 2017). Desta forma, o caminho de expansão (representado pelos trechos na cor roxa) é semelhante em simetria ao caminho de contração (representado pelos trechos de cor verde) e produz uma arquitetura em forma de "U", conforme a figura 9. A rede não possui nenhuma camada totalmente conectada e usa apenas a parte válida de cada convolução, assim, o mapa de segmentação contém apenas os *pixels*, para os quais o contexto completo está disponível na imagem de entrada. Essa estratégia permite a segmentação contínua de imagens arbitrariamente grandes por meio de uma estratégia de sobreposição de blocos (RONNEBERGER; FIISCHER; BROX, 2015).

A rede U-Net mais uma vez foi utilizada por Jin et al. (2017), que a utilizou como ferramenta de arquitetura, seguindo o fluxo de uso das ferramentas de iteração como aprendizagem profunda para alcançar a remoção de artefatos com eficácia. Nota-se que foi utilizada como reconstrução, a transformada inversa de RADON, também conhecida como *FBP* (*Filtered Backprojection*). A arquitetura proposta com uso da U-Net está representada na figura 9 (JIN et al., 2017). É possível notar a utilização nessa arquitetura, também de imagens com 64×64 *pixels*, no entanto, são utilizados 03 canais de entrada e 01 canal de saída, diferente da arquitetura proposta neste trabalho, de 01 canal de entrada e 01 canal de saída.

2.3 Caracterização dos Artefatos Metálicos em Tomografia Computadorizada

Não há um consenso acerca da definição de artefato, no entanto, se percebe um alinhamento conceitual, que define artefato como tudo aquilo que aparece na imagem final mas que não faz parte do objeto real escaneado (SCHULZE et al., 2011), e que artefato é qualquer discrepância entre os valores reconstruídos em uma imagem e os valores reais sobre coeficientes de atenuação dos objetos no corpo estudado (HSIEH, 2009). De um modo geral, os artefatos são induzidos por discrepâncias entre as condições físicas reais do conjunto de medição (ou seja, a composição técnica do aparelho de tomografia mais a composição, posição e comportamento do objeto sob investigação) e os pressupostos matemáticos simplificados utilizados para a reconstrução em 3D.

A presença das referidas inconsistências podem dificultar ou até mesmo inviabilizar o diagnóstico. Existem diferentes tipos de artefatos, classificados nas seguintes categorias: anéis, estrias, sombreamentos e diversos (HSIEH, 2009), bem como, classificados nas

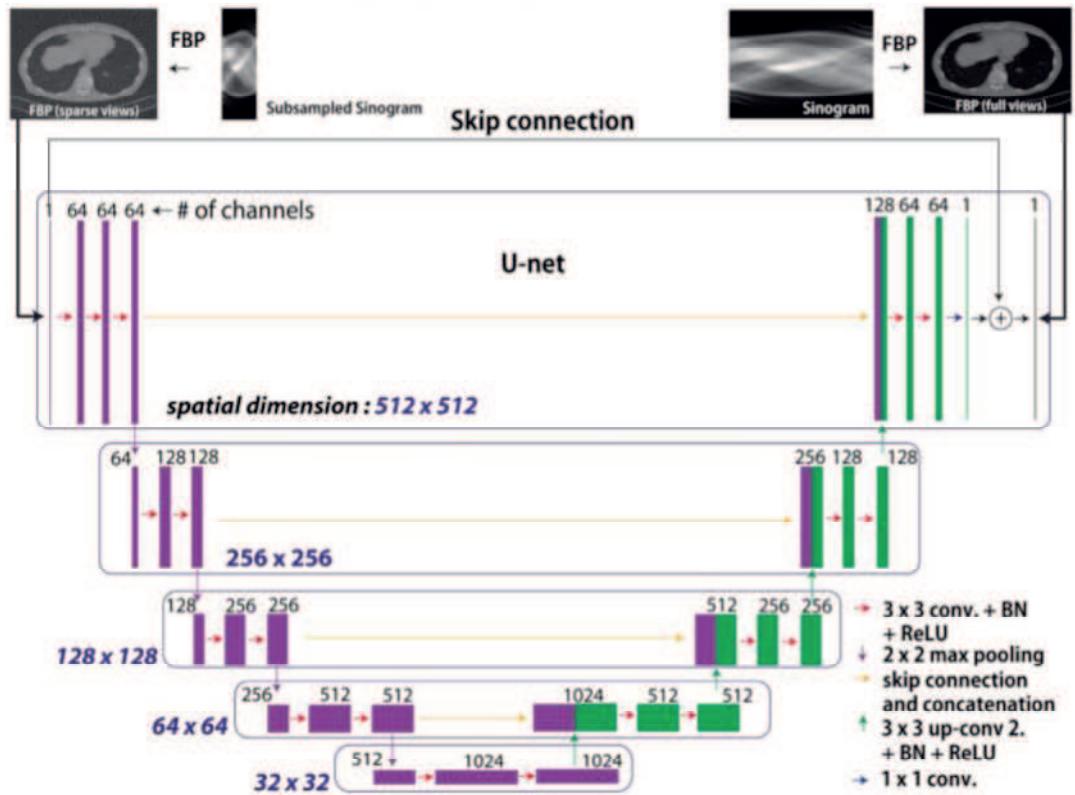


Figura 9 – Arquitetura com uso de U-Net.(JIN et al., 2017)

categorias de anel, ruído e endurecimento e dispersão de feixe (GJESTEBY et al., 2016). Iremos abordar os artefatos segundo essa última classificação.

2.3.1 Artefato de Anel (*Ring Artifact*)

Um defeito ou má calibração do detector do aparelho de tomografia cria um anel brilhante ou escuro situado no centro de rotação, podendo por vezes, simular patologia (Figura 10). Normalmente, recalibração do detector é suficiente para corrigir esse artefato, embora, em situações comuns, o próprio detector necessite de substituição (BOAS; FLEISCHMANN, 2012).

2.3.2 Ruído

O ruído de Poisson ocorre devido ao erro estatístico das baixas contagens de fôtons, durante a aquisição da projeção e resulta em raios finos e brilhantes, que aparecem preferencialmente na direção da maior atenuação (Figura 11). Com o aumento do ruído, objetos de alto contraste, como o osso, ainda podem ser visíveis, mas limites de tecido mole de baixo contraste podem ser escurecidos.

O ruído de Poisson pode ser reduzido ao aumentar a corrente (mA). Os *scanners* modernos podem executar a modulação da corrente do tubo, aumentando seletivamente

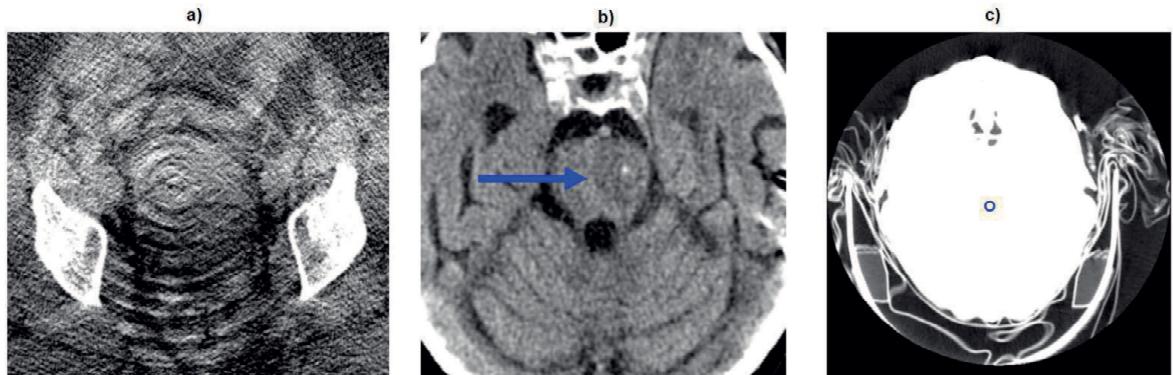


Figura 10 – a) TC pélvica mostrando artefato de anel severo. b) TC principal com artefato de anel sutil, simulando uma lesão (indicada pela seta). c) Alterações nas configurações de nível de intensidade e janela de abertura, evidenciam a região de reconstrução, situada no centro de rotação (anel azul) (adaptado de (BOAS; FLEISCHMANN, 2012))

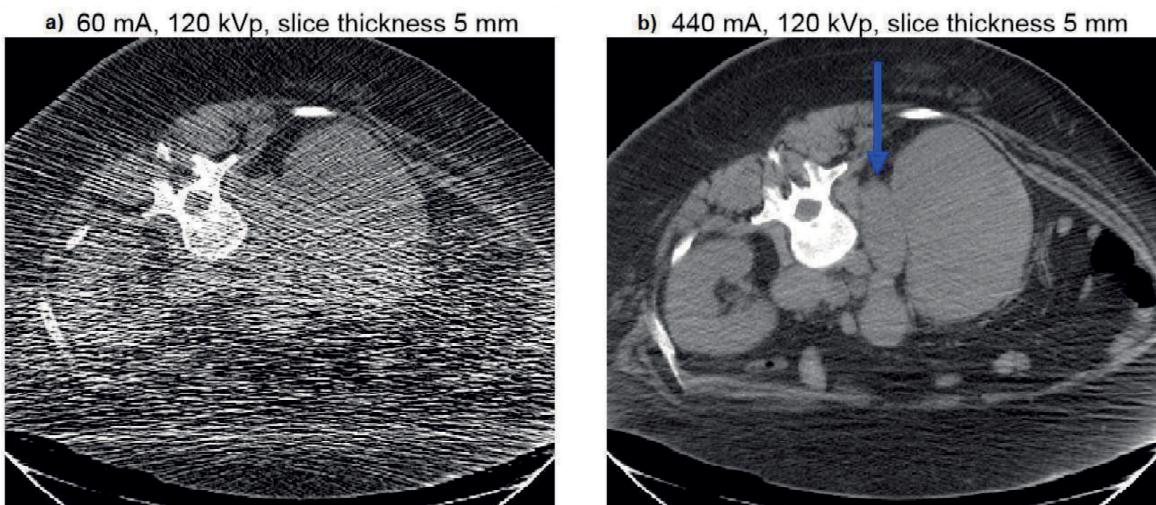


Figura 11 – a) A imagem de TC de baixa dose obtida durante uma biópsia guiada por TC mostra um amplo ruído de Poisson b) A imagem pós-biópsia obtida com uma dose 7,3 vezes maior e apresentou $\sqrt{7,3} = 2,7$ vezes menos ruído. As imagens mostram um nódulo linfático retroperitoneal ampliado (flecha) e infiltração do rim direito em paciente com linfoma de Hodgkin. Adaptado de (BOAS; FLEISCHMANN, 2012)

a dose ao adquirir uma projeção com alta atenuação. Eles também usam tipicamente filtros *bowtie*, que em referência à gravata borboleta, o filtro é utilizado para modular um feixe de raios-x entrante como uma função do ângulo do raio-x em relação a um paciente para equilibrar o fluxo de fôtons em uma matriz de detectores, fornecendo uma dose mais elevada para o centro do campo de visão em comparação com a periferia. Existe um relacionamento entre o ruído e a resolução, de modo que o ruído também pode ser reduzido ao se aumentar a espessura da fatia, utilizando-se uma reconstrução mais suave (kernel de tecido mole em vez de kernel ósseo), ou desfocando a imagem (BOAS; FLEISCHMANN, 2012).

2.3.3 Endurecimento e Dispersão do Feixe

O endurecimento e a dispersão do feixe são mecanismos diferentes que produzem camadas escuras entre dois objetos de alta atenuação, como metal, osso, contraste iodado ou bário. Eles também podem produzir marcas escuras ao longo do eixo longo de um único objeto de alta atenuação (Figura 12). Riscos brilhantes são vistos ao lado das marcas escuras. Esses artefatos são um problema particular na fossa craniana posterior e com implantes metálicos. Quando ocorre a passagem de um feixe de raio-x dentro de uma ampla faixa de energia (raio-x policromático) por um objeto, a atenuação dos raios é inversamente proporcional à energia do fóton, ou seja, quanto maior é a energia do fóton, tanto menor seu coeficiente de atenuação linear, o que implica que raios de baixa energia sofrem maior atenuação do que raios de alta energia. Dessa forma, raios-x de uma ampla faixa de energia, ao passarem por objetos que filtrem/absorvam as baixas energias, lhes sobram apenas as faixas de maior energia, causando o "endurecimento" do feixe (US, 2013).

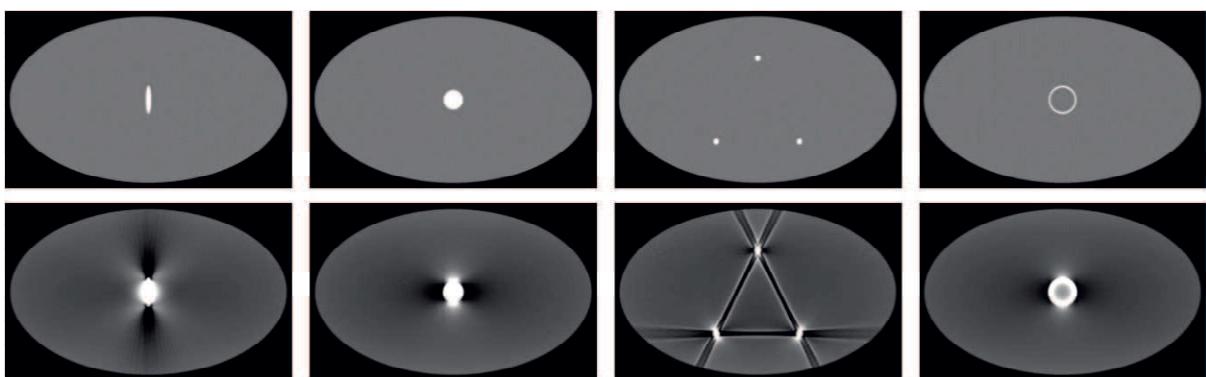


Figura 12 – Simulação de varredura sem endurecimento do feixe (linha superior) e com endurecimento do feixe (linha inferior), mostrando que as faixas escuras ocorrem ao longo das linhas de maior atenuação, e as raias brilhantes ocorrem em outras direções (adaptado de (BOAS; FLEISCHMANN, 2012))

Os artefatos metálicos estão diretamente associados à composição e ao tamanho dos implantes metálicos, pois dependendo desses fatores, ocorrerão diferentes índices de atenuação dos raios-x e dessa forma, diferentes formas de artefatos na imagem.

Os principais métodos de atenuação de artefatos metálicos: Otimização de Implantes de Metal, Melhoria de Aquisição, Pré-processamento baseado em Física, Ajuste de Projeção, Pós-processamento de Imagem e Reconstrução Iterativa (GJESTEBY et al., 2016), serão brevemente explanados a seguir .

2.4 Métodos de Atenuação de Artefatos Metálicos

2.4.1 Otimização de Implantes de Metal

A forma mais óbvia de reduzir o artefato metálico em uma tomografia computadorizada é através da remoção do metal antes do exame, no entanto, na prática não é possível, na maioria das vezes, remover cirurgicamente os implantes metálicos antes dos exames e nem sempre, alternativas não metálicas são viáveis do ponto de vista da eficácia do tratamento, apesar da publicação do estudo de Gray et al.(2012), no qual ele demonstra o aumento da precisão da imagem em pacientes com câncer na faringe, após remover todos os dentes que possuíam restaurações metálicas (GRAY et al., 2012). Assim, Gjesteby (2016) relata que muitos estudos têm se dedicado na avaliação de imagens com artefatos produzidos por vários tipos de metais, com o objetivo de identificar o implante que tenha a menor atuação, para cada caso clínico, de modo que se comprovou, que clipe cirúrgico de Titânio ou aço inoxidável em hemostasia geram menos artefatos metálicos do que clipe cirúrgico feitos de Tântalo.

2.4.2 Melhoria de Aquisição

Uma segunda forma de atenuar os artefatos, é através de alterações nos parâmetros de aquisição de imagens dos aparelhos de tomografia computadorizada, como tensão, corrente, plano focal, deslocamento do eixo, além de empregar ainda, protocolos de energia dupla que adquiram dados usando dois espectros de raios-x para a decomposição do sinal.

Recentemente, detectores de contagem de fótons foram utilizados como uma forma de reduzir os artefatos metálicos. Especificamente, os detectores de contagem de fótons no modo de soma de carga são capazes de discriminar energias de fótons individuais para a reconstrução espectral e são mais eficientes a níveis muito baixos de contagem de fótons (GJESTEBY et al., 2016).

A tabela 2.4.2 contém as principais vantagens e desvantagens das técnicas de melhoria de aquisição.

2.4.3 Pré-processamento Baseado em Física

As técnicas de pré-processamento baseadas em física, visam a modelar a causa física de artefatos metálicos, tais como o ruído, a dispersão, o endurecimento do feixe e o efeito NLPV (*Nonlinear Partial Volume*, também chamado de *gradient edge effect*, que se manifesta como listras escuras tangentes a bordas afiadas (MAN, 2001)) e ainda corrigir os dados atenuados pelo artefato, para melhorar a reconstrução. Essas correções são geralmente aplicadas no domínio de projeção antes da reconstrução, embora os mesmos

Tabela 4 – Principais vantagens e desvantagens das técnicas de melhoria de aquisição.

Ajuste na Aquisição	Vantagem Principal	Desvantagem Principal
aumento da energia do fóton (kVp)	redução de ruído e endurecimento do feixe	aumento da dosagem no paciente
aumento da corrente (mA)	redução de ruído	aumento da dosagem no paciente
maior refinamento	redução de ruído	aumento da dosagem no paciente
redução da espessura da fatia	redução de parte do volume do artefato	aumento do ruído
ampliação da escala da tomografia	aumento da largura e profundidade do sinal	redução da resolução do contraste

Tabela 2.4.2 - Efeito dos ajustes no tomógrafo, para redução dos artefatos metálicos.

Adaptado de (GJESTEBY et al., 2016).

modelos físicos possam, alternativamente, ser incorporados no modelo de um algoritmo de reconstrução iterativa (GJESTEBY et al., 2016)

2.4.4 Ajuste de Projeção

Em alguns casos, as estrias provocadas pelos artefatos metálicos provocam a corrupção da imagem de forma a não ser possível visualizar o objeto de estudo, seja adicionando estrias claras e escuras, ou eclipsando dados importantes para a construção da imagem, de modo que novos dados precisam ser adicionados para completar o sinograma (Figura 13).

Uma forma de se conseguir isso é interpolando valores de substituição, seja de projeções vizinhas ou de um modelo matemático. Uma segunda abordagem de ajuste de projeção que aumenta a precisão é através da incorporação de conhecimento prévio para orientar a estimativa dos dados que substituem projeções corrompidas. Esta imagem anterior é reprojetada para gerar dados para a conclusão do sinograma e, portanto, esses tipos de procedimentos são classificados como reprojeção (GJESTEBY et al., 2016).

Outro método baseado em reprojeção, com baixo custo computacional, capaz de reduzir estrias provocadas por artefatos metálicos está descrito no artigo (CRAWFORD et al., 1988). O processo envolve a identificação de objetos metálicos em uma imagem original reconstruída com segmentação da imagem e posteriormente, o preenchimento dessas regiões com o valor correspondente ao da água. Após isso, a imagem é reprojetada e essas projeções são subtraídas das projeções da imagem original, passando a reconstruir o sinograma corrigido na imagem final.

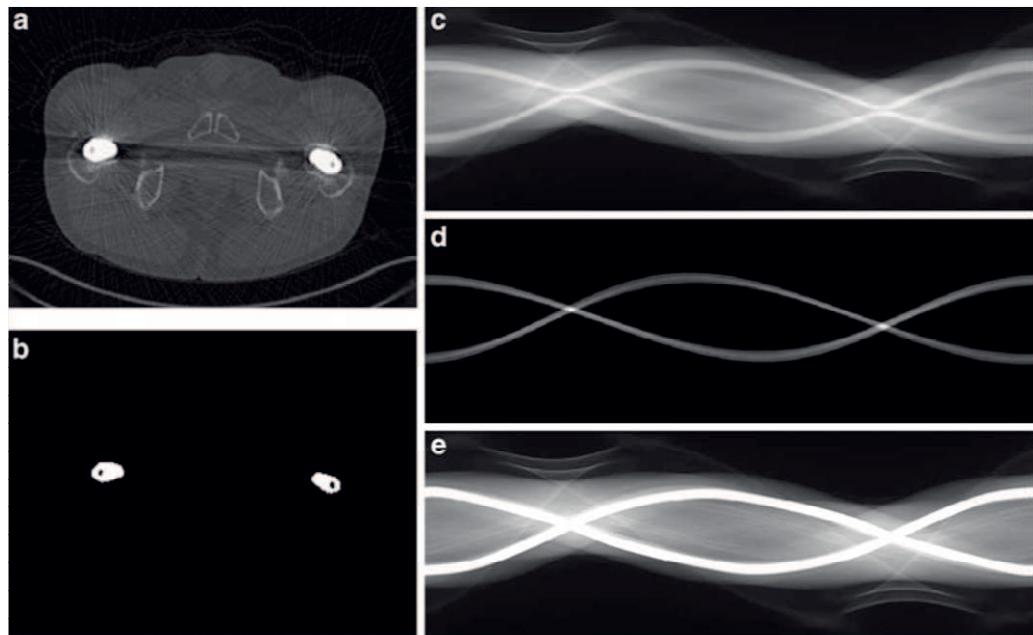


Figura 13 – Detecção de projeções afetadas por artefatos metálicos: a) imagem TC original, b) artefato metálico em imagem segmentada através de limiar, c) sinograma original, d) sinograma apenas do metal e e) dados dispersos. (ABDOLI et al., 2011)

A redução de artefatos metálicos em imagens de tomografia computadorizada dentárias foi proposta produzindo imagens por somas ponderadas de duas imagens: a primeira, uma imagem com estrias por artefatos reduzida e reconstruída a partir dos dados de projeção modificados para se adequar à influência de metal e a outra, uma imagem sem metal reconstruída a partir de dados originais de projeção, utilizados para remoção metal em uma região (HEGAZY; CHO; LEE, 2016). Para fazer a imagem reduzida, foi realizada a segmentação precisa da região do metal com base no limiar local e, em seguida, foram modificados os dados da região do metal, usando a interpolação linear, técnica conhecida também como *inpainting*. Foi realizada a projeção para imagens seguintes, a partir dos dados de projeções anteriores, sendo substituídos os valores dos *pixels* na região do metal nos dados de projeção originais com os dados de projeções anteriores, para finalmente reconstruírem as imagens.

2.4.5 Pós-processamento de Imagem

O pós-processamento de imagem é uma abordagem menos popular e tem objetivo de trabalhar com a imagem no domínio do espaço. Os métodos de pós-processamento reduzem os artefatos após a reconstrução da imagem e não dependem do acesso a dados de projeção, ocorrendo no domínio da imagem. Essas técnicas, utilizadas de forma isolada, não são de grande eficácia se comparadas a abordagens de correção de dados, uma vez que o dano à informação já ocorreu, mas se utilizadas conjuntamente com outras técnicas de

reconstrução, podem aumentar a qualidade do resultado.

2.4.6 Reconstrução Iterativa

A reconstrução iterativa começa com uma imagem ideal e compara suas projeções com os dados de projeção realmente medidos pela aquisição da tomografia computadorizada. O objetivo então, é minimizar o erro entre esses sinogramas, otimizando uma função objetiva que orienta a reconstrução. Exemplos dessa função objetiva incluem erros mínimos quadráticos e máxima verossimilhança, que objetivam encontrar a distribuição de coeficientes de atenuação linear a partir dos dados de projeção com a máxima probabilidade (GJESTEBY et al., 2016).

Uma das abordagens para a reconstrução iterativa é desprezar completamente o subconjunto de dados de projeção que se encontra corrompido por estrias provocadas pelos objetos metálicos. Esses métodos tratam o problema MAR (*Metal Artifact Removal*) como um problema exterior e utilizam as informações fora da estria do artefato metálico para chegar a um parâmetro de reconstrução. Os valores externos às estrias dão uma solução robusta fora do objeto metálico, mas não tão estável perto da superfície metálica.

Foi abordada uma redução de artefato metálico (MAR), tratando o objeto escaneado como um toróide, desprezando as medições dos raios-x que passam pelo orifício desse toróide e realizando a reconstrução da parte externa a partir de linhas não afetadas. O orifício é preenchido com valores de atenuação constantes, opcionalmente com base na densidade conhecida do objeto metálico e as integrais de linha correspondentes são estimadas. Uma atualização iterativa da imagem é realizada até o estado estacionário ser atingido. A qualidade da imagem é melhorada utilizando este método, embora o conteúdo da informação nas bordas do metal seja reduzido (MEDOFF et al., 1983).

Os métodos de reconstrução iterativa que envolvem dados incompletos podem ser aprimorados com associações de outras técnicas, como amostragem compressiva e modelos de componentes conhecidos, assim, modelos atuando em conjunto definem a forma, o tamanho e/ou a informação de densidade dos objetos metálicos na imagem, de modo a auxiliar a estimativa das projeções eventualmente corrompidas. Outra proposta de algoritmo iterativo, usa uma imagem previamente selecionada para casos em que dados de projeção não codificados não estão disponíveis (KALVIN; WILLIAMSON, 1997).

Assim, propostas de combinação de técnicas para a redução de artefatos metálicos são vastas, como no caso de utilizar reconstrução modificada por amostragem compressiva e o método de ajuste por interpolação e substituição de dados corrompidos, (*inpainting*), do sinograma (TANG; HU; ZHANG, 2014). O trabalho proposto começa com a reconstrução da imagem por *total variation*, ou seja, reduzir a variação total do sinal afetado pelo ruído acarreta uma correspondência próxima ao sinal original, removendo-se detalhes indesejados

para obter uma imagem inicial, produzindo assim, uma imagem de fundo sem artefatos. A informação faltando no sinograma original é complementada usando o *inpainting* no sinograma, conforme exposto anteriormente nesta seção. Por fim, as transições entre o sinograma original e artificial são suavizadas para a reconstrução final. O algoritmo é validado usando simulações e dados *phantom* (SHEPP; LOGAN, 1974) e os resultados mostram que as estriadas são eliminadas e as sombras são reduzidas, demonstrando que o método proposto preserva altas frequências e estruturas de baixos contrastes, enquanto remove estriadas de artefatos metálicos.

3 Fundamentos de Aprendizagem de Máquina

Aprendizagem de máquina, de modo geral, é um ramo da Inteligência Artificial onde os modelos computacionais propostos buscam soluções para problemas os quais não foram explicitamente programados. Pode ser descrito como uma série de práticas destinadas à solução de problemas os quais não há uma modelagem inicial capaz de resolvê-los. Essa solução parte através da análise de um conjunto finito de fatos, casos ou exemplos, que descrevem objetos, processos, situações ou ambientes, uma vez que esses algoritmos permitem generalizar modelos através de parâmetros que correspondam à aproximação de uma determinada realidade, possibilitando soluções (PRATI, 2006).

Existem três tipos principais de aprendizado de máquina, sendo eles não-supervisionado, supervisionado e aprendizado por reforço.

- Aprendizado não supervisionado: Não são utilizados reguladores nas variáveis de saída, de modo que as variáveis de entrada têm seus valores associados a grupos de acordo com a proximidade de seus valores e quantidade de grupos a serem divididos. São utilizados também para a descoberta de regularidades ocultas de dados ou para detecção de anomalias nos dados (RÄTSCH, 2004). Essa linha de pesquisa não será aprofundada neste trabalho, por fugir ao seu escopo.
- Aprendizado supervisionado: Neste caso, os valores de saída são comparados com valores ótimos, assim, há um "supervisor" aferindo se a saída do modelo está condizente com o modelo proposto. Os principais algoritmos desta categoria criam modelos que possibilitam a classificação dos valores de saída, variáveis discretas e a regressão, para variáveis contínuas.
- Aprendizado por reforço: O algoritmo pertencente a esta categoria não permanece apenas recebendo rótulos para aferir seu estado atual, ele interage ativamente com o ambiente, adquirindo informações que lhe permitem avaliar seu estado atual e futuro, para atingir um determinado objetivo. Este tipo de algoritmo é utilizado em aprendizado de robôs e sistemas de transporte autônomos e seu aprofundamento foge à proposta deste trabalho.

3.1 Modelos Supervisionados

Dois métodos são essencialmente utilizados em modelos supervisionados: **regressão**, que consiste na utilização de dados de entrada para prever, na melhor forma possível, dados contínuos de saída que correspondam às saídas reais ou corretas e **classificação**, que consiste na atribuição de uma categoria a um dado de entrada, por exemplo, imagens médicas podem ser divididas em imagens contendo tumores ou não.

3.1.1 Regressão

O problema de aprendizagem da regressão pode ser descrito como: seja X o espaço de entrada e Y um subconjunto mensurável de \mathbb{R} . Denota-se por D , uma distribuição desconhecida sobre X , onde os dados de entrada são alocados e por $f : X \rightarrow Y$, a função de rotulagem dos dados de saída. Isso corresponde a um cenário de aprendizado determinístico que pode ser diretamente estendido para um cenário estocástico onde temos uma distribuição sobre os pares $(x, y) \in X \times Y$.

Conforme todo problema de aprendizagem supervisionada, o modelo de aprendizado recebe rótulos para comparar com a saída $S = ((x_1, y_1), \dots, (x_m, y_m)) \in (X \times Y)^m$, com x_1, \dots, x_m de acordo com D e $y_i = f(x_i)$, para todo $i \in [1, m]$.

Como os rótulos são números reais, não é razoável esperar que o modelo de aprendizagem possa prever com precisão o rótulo correto, mas se pode exigir que suas previsões estejam próximas das corretas.

Esta é a principal diferença entre regressão e classificação: na regressão, a medida de erro é baseada na magnitude da diferença entre o rótulo avaliado de valor real e o verdadeiro ou correto e não com base na igualdade ou desigualdade entre esses valores (MOHRI; ROSTAMIZADEH; TALWALKAR, 2012).

3.1.2 Métodos Paramétricos

Os Métodos Paramétricos consideram que a função de regressão pode ser modelada através de um número finito de parâmetros (IZBICKI, 2016).

Tomando por base, uma regressão linear, temos:

$$r(x) = \beta^t x = \beta_0 x_0 + \beta_1 x_1 + \dots + \beta_d x_d, \quad (3.1)$$

em que é adotada a convenção $x_0 \equiv 1$, e $\beta = (\beta_0, \dots, \beta_d)$. É necessário expor que x_i não precisa necessariamente ser a i -ésima variável original, pois é possível criar novas covariáveis que são funções das originais (IZBICKI, 2016).

Uma forma de estimar os coeficientes β da regressão linear, é através do método dos mínimos quadrados, que propõe o estimador dado por:

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \sum_{i=1}^n (Y_i - \beta_0 - \beta_1 x_{i,1} - \dots - \beta_d x_{i,d})^2 \quad (3.2)$$

e a solução desse problema é dada por:

$$\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_d) = (\mathbb{X}^t \mathbb{X})^{-1} \mathbb{X}^t Y \quad (3.3)$$

em que

$$\mathbb{X} = \begin{pmatrix} X_{10} & \dots & X_{1d} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ X_{n0} & \dots & X_{nd} \end{pmatrix} \quad (3.4)$$

de modo que a função de regressão é estimada por (IZBICKI, 2016):

$$g(x) = \hat{\beta}^t x \quad (3.5)$$

3.1.3 Métodos Não Paramétricos

Métodos paramétricos costumam apresentar bons resultados para amostras pequenas, pois se o modelo possui poucos parâmetros, se torna possível criar estimadores com baixa variância, até se o número de dados de entrada for pequeno. No caso de uma entrada de dados grande, é possível aumentar o número de parâmetros do modelo, acarretando um viés menor, à custa de uma variância maior. Dessa forma, denota-se a importância dos métodos não paramétricos: "um modelo não paramétrico é um modelo que tem infinitos parâmetros" (IZBICKI, 2016).

3.1.3.1 K Nearest Neighbor

O algoritmo de k-vizinhos mais próximos (*K Nearest Neighbor*, KNN) e suas variações estão entre os algoritmos mais simples e populares na comunidade de aprendizado de máquinas. Utilizando esse método para regressão, para estimar a função $r(x)$ para uma dada configuração de covariáveis x com base nas respostas Y dos k-vizinhos mais próximos a x , temos:

$$\hat{r}(x) = \frac{1}{K} \sum_{x_i \in N_x} Y_i \quad (3.6)$$

em que N_x é o conjunto das k observações mais próximas de x , i.e.,

$$N_x = \{i \in \{1, \dots, n\} : d(x_i, x) \leq d_x^k\} \quad (3.7)$$

onde d_x^k é a distância do k -ésimo vizinho mais próximo de x a x , ou seja, a função de regressão avaliada em x é estimada, utilizando-se de uma média local das respostas dos k vizinhos mais próximos a x no espaço das covariáveis (IZBICKI, 2016).

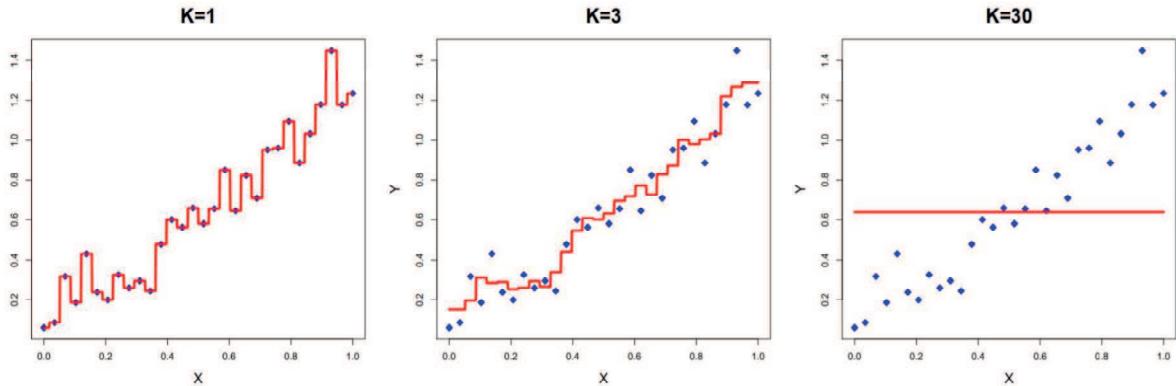


Figura 14 – Influência na escolha de k no estimador dos k vizinhos mais próximos. (IZBICKI, 2016)

Utilizando para classificação, a ideia é memorizar o conjunto de treinamento e depois prever o rótulo de qualquer nova instância com base nos rótulos de seus vizinhos mais próximos no conjunto de treinamento. A lógica existente no método é baseada na suposição de que os pontos de domínio são relevantes para suas vizinhanças de modo que os pontos próximos tenham o mesmo rótulo. Esta técnica permite que em algumas situações, mesmo quando o conjunto de treinamento é muito grande, é possível encontrar o rótulo do vizinho mais próximo com brevidade.

Assumindo χ , dotado com uma função métrica ρ , qual seja, $\rho : \chi \times \chi \rightarrow \mathbb{R}$ é a função que retorna a distância entre dois elementos de χ . Ao tomarmos por exemplo, que $\chi = \mathbb{R}^d$, então ρ pode ser a distância Euclidiana $\rho(x, x') = \|x - x'\| = \sqrt{\sum_{i=1}^d (x_i - x'_i)^2}$.

Seja $S = (x_1, y_1), \dots, (x_m, y_m)$ a sequência de dados de treinamento, para cada $x \in \chi$, seja $\pi_1(x), \dots, \pi_m(x)$ a reordenação de $\{1, \dots, m\}$, de acordo com a distância para x , $\rho(x, x_i)$. Isso, para todo $i < m$, (SHALEV-SHWARTZ S, 2014).

$$\rho(x, x_{\pi_i(x)}) \leq \rho(x, x_{\pi_{i+1}(x)}) \quad (3.8)$$

A figura 15 representa uma distribuição de rótulos no espaço de amostragem dos dados de entrada, simbolizados pelos círculos azuis e as linhas azuis delimitam os espaços de vizinhança dos rótulos.

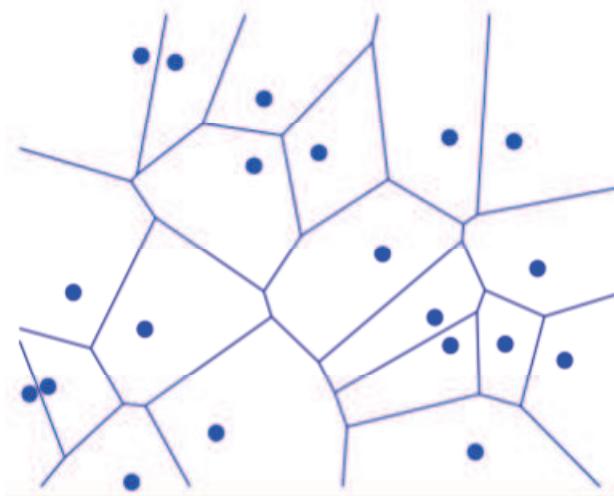


Figura 15 – Ilustração dos limites de decisão (SHALEV-SHWARTZ S, 2014).

3.1.3.2 Support Vector Machine - SVM

O algoritmo original *Support Vector Machine - SVM* foi criado por Vladimir N. Vapnik e Alexey Ya. Chervonenkis em 1963, mas na década de 90, passou-se a utilizar o algoritmo SVM, com a ideia de hiperplano como classificador linear. O melhor hiperplano é aquele que representa a maior separação, ou margem, entre as duas classes.

Existem infinitas soluções de hiperplanos para separar os dados de treinamento e o SVM se propõe a identificar o hiperplano de margem máxima ou de separação ideal, que é o hiperplano separador que está mais distante das observações de treinamento, ou seja, é calculada a distância perpendicular de cada dado de treinamento para um determinado hiperplano de separação e o ideal é um hiperplano com a máxima distância perpendicular até os dados de treinamento, conforme está representado na figura 16 (JAMES, 2015).

3.1.3.3 Árvores de Regressão

Árvores de Regressão são metodologias não paramétricas que permitem resultados interpretáveis, sendo formadas por particionamentos recursivos no espaço das covariáveis, onde cada particionamento recebe o nome de nó e cada resultado final recebe o nome de folha. Uma representação da Árvore de Regressão pode ser visualizada na figura 17 (IZBICKI, 2016).

Uma Árvore de Regressão tem sua estrutura criada através de duas grandes etapas: (I) a criação de uma árvore completa e complexa com seus nós e folhas e (II) a poda desta árvore, com o intuito de evitar o super ajuste.

Uma árvore cria uma partição do espaço das covariáveis em regiões distintas e disjuntas: R_1, \dots, R_j e a resposta Y , é a predição de uma observação com covariáveis x que

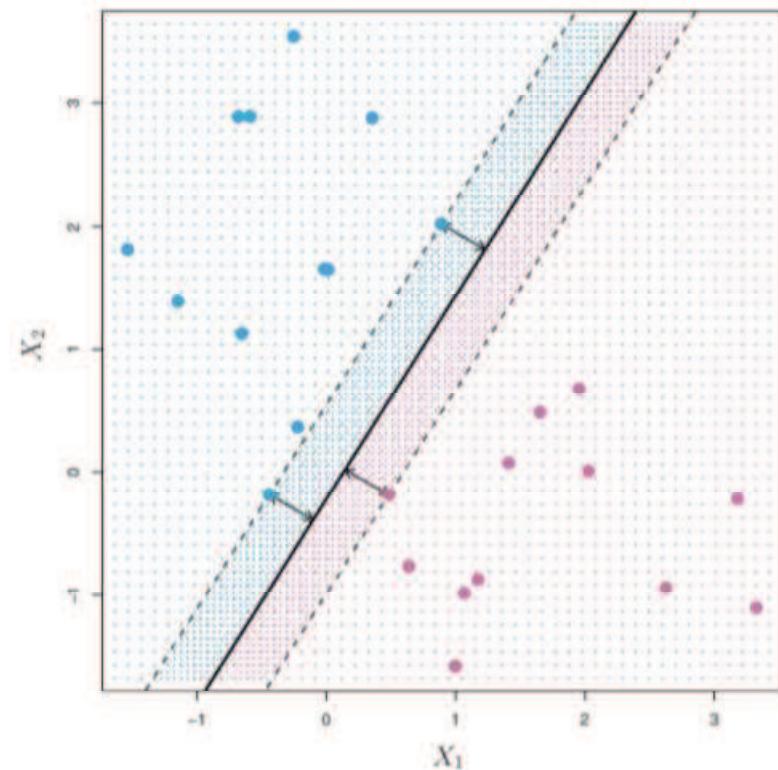


Figura 16 – O hiperplano é representado pela linha sólida ao centro, as margens são representadas pelas linhas pontilhadas e a distância dos pontos ao hiperplano são representadas por setas (JAMES, 2015).

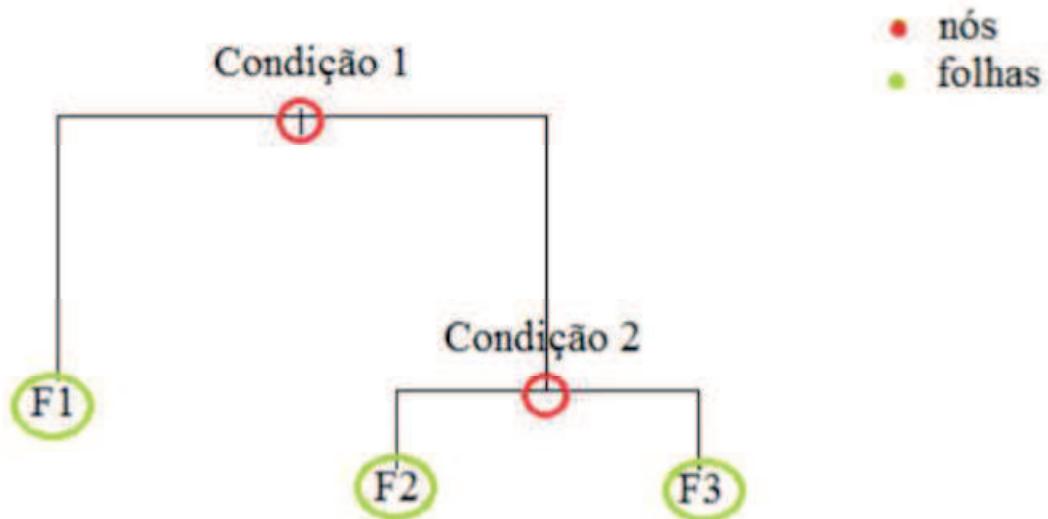


Figura 17 – Representação da estrutura de uma Árvore de Regressão (IZBICKI, 2016).

estão em R_k e é dada por:

$$g(x) = \frac{1}{|\{i : x_i \in R_k\}|} \sum_{i:x_i \in R_k} y_i \quad (3.9)$$

Na criação da árvore, busca-se partições nos quais os valores de Y , nas observações do conjunto de treinamento das folhas seja homogêneo. Assim, para avaliar o erro quadrático da árvore, faz-se uso da equação:

$$P(T) = \sum_R \sum_{k \in R} (y_k - \hat{y}_R)^2 \quad (3.10)$$

em que \hat{y}_R é o valor predito para a resposta de uma observação pertencente à região R . Para escolher qual partição o fluxo lógico deve seguir, busca-se dentre todas as covariáveis x_i e cortes t_1 , a combinação que corresponda a uma partição (R_1, R_2) com as menores previsões de erro quadrático.

$$EQ = \sum_{i: x_i \in R_1} (y_i - \hat{y}_{R_1})^2 + \sum_{i: x_i \in R_2} (y_i - \hat{y}_{R_2})^2 \quad (3.11)$$

em que \hat{y}_{R_k} é a previsão fornecida para a região R_k . Assim, define-se (IZBICKI, 2016):

$$R_1 = \{x : x_{i^*} < t_1^*\}$$

e

$$R_2 = \{x : x_{i^*} \geq t_1^*\} \quad (3.12)$$

em que x_{i^*} e t_1^* é a variável e o corte escolhido, respectivamente.

3.1.4 Deep Learning

3.1.4.1 Introdução

Inicialmente inspirada no neurônio biológico, o primeiro modelo matemático de um neurônio foi proposto pelo neurofisiologista, filósofo e poeta americano Warren McCulloch e pelo cientista cognitivo Walter Pitts, no ano de 1943 (MCCULLOCH; PITTS, 1943). Em 1957, Rosenblatt propôs o primeiro modelo de rede neural, o qual nomeou de *Perceptron*, que consistia em uma rede de duas camadas, sendo uma de entrada e outra de saída, formada por neurônios com função de ativação binária. Em 1969, Minsky e Papert, através de análises das propriedades matemáticas das redes perceptron, apontaram suas limitações, como a impossibilidade de simular o operador "ou-exclusivo" (MINSKY; PAPERT, 1969), bem como, a falta de computadores suficientemente potentes para tratar a complexidade inerente ao método e esses fatores acarretaram o ostracismo da tecnologia na época (BITTENCOURT, 2006).

Houve um ressurgimento do interesse pelas redes neurais artificiais nos anos 80, devido a melhores entendimentos do funcionamento do neurônio biológico, algoritmos de treinamento mais eficientes, além de computadores mais robustos, inclusive paralelos.

As Redes Neurais Artificiais - RNA têm sido utilizadas em diversas demandas na área de visão computacional, como reconhecimento de caracteres (LECUN et al., 1998), classificação de imagens (KRIZHEVSKY; SUTSKEVER; HINTON, 2012), detecção de objetos (REN et al., 2015), colorização de imagens em preto e branco (MARTINS, 2017) e remoção de ruído (MAZZA, 2017), segmentação e reconstrução por regressão, objetivo deste trabalho, que contará com treinamento supervisionado.

Redes Neurais Artificiais

Uma RNA tem seu poder computacional devido à estrutura paralela distribuída e a sua capacidade de aprender e generalizar. Essa generalização permite que a RNA possa resolver problemas complexos para os quais ela não se deparou anteriormente.

Na figura 18, há a representação de um neurônio biológico, sendo possível observar na imagem, desenhos que representam o corpo celular, dentritos e axônio. Os neurônios se comunicam com as outras células através de uma série de breves impulsos de tensão, também chamados de potenciais de ação ou picos. Esses potenciais se originam no corpo celular ou em seu entorno e se propagam através do neurônio em uma velocidade e amplitude constantes. O potencial de ação é baseado na física dos axônios, que por ser muito longo e fino, possui uma resistência elétrica elevada e grande capacidade. Dessa forma, quando uma tensão é aplicada a uma extremidade do axônio, ela decai exponencialmente com a distância, chegando a um nível insignificante no momento que atinge a outra extremidade. Os potenciais de ação são uma maneira que a evolução encontrou para contornar esse problema de transmissão (HAYKIN, 1998).

O neurônio é a unidade de processamento fundamental para o funcionamento de uma rede neural. Na figura 19, está representado um modelo matemático de um neurônio, que é a base da concepção das Redes Neurais Artificiais. É possível identificar três elementos do modelo neuronal (HAYKIN, 1998):

- Um conjunto de sinapses ou elos de ligação, cada um dos quais é caracterizado por um peso ou força própria. Especificamente, um x_p sinal na sinapse p ligada à entrada do neurônio k é multiplicada pelo peso sináptico w_{kp} . O primeiro subscrito refere-se ao neurônio em questão e o segundo subscrito refere-se a entrada na qual a extremidade do peso sináptico se refere. Diferentemente de uma sinapse no cérebro, o peso sináptico de um neurônio artificial pode estar numa gama que inclui tanto valores negativos quanto positivos.
- uma função somatória para somar o sinal de entrada, ponderada pelas respectivas

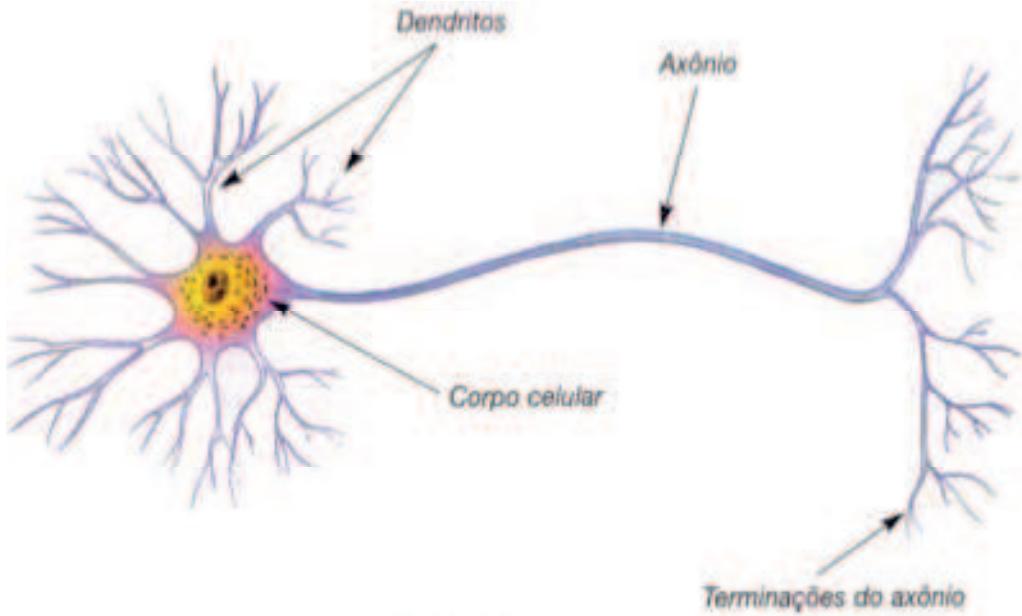


Figura 18 – representação de neurônio artificial

sinapses dos neurônios, constituindo uma combinação linear.

- uma função de ativação para limitar a amplitude da saída de um neurônio.

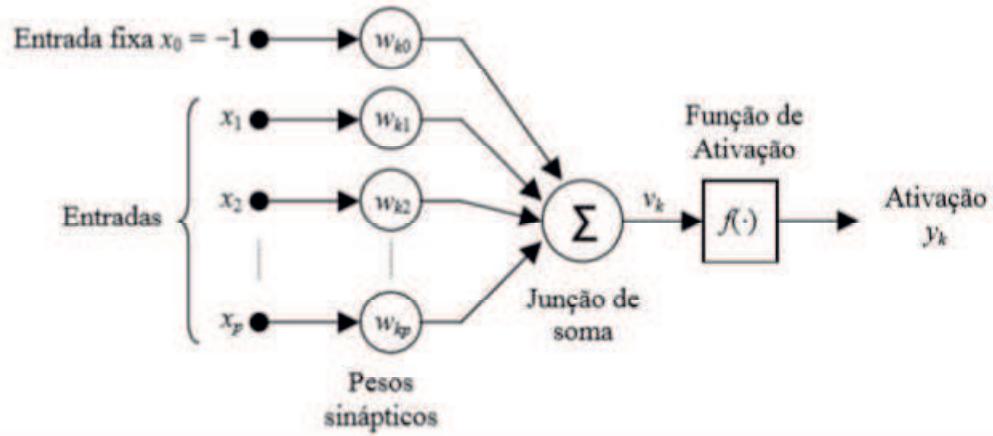


Figura 19 – Modelo de neurônio artificial (HAYKIN, 1998)

Matematicamente, um neurônio k pode ser descrito pelo seguinte par de equações (HAYKIN, 1998):

$$u_k = \sum_{j=1}^m w_{kj} x_j \quad (3.13)$$

e por

$$y_k = f(u_k + w_{k0}) \quad (3.14)$$

em que x_1, x_2, \dots, x_p são os sinais de entrada; $w_{k1}, w_{k2}, \dots, w_{kp}$ são os pesos sinápticos do neurônio k ; u_k é a combinação linear de saída para os sinais de entrada; w_{k0} é um valor fixo; $f(\cdot)$ é a função de ativação e y_k é o sinal de saída.

A polarização tem o efeito de aumentar ou diminuir a entrada do sinal na função de ativação, dependendo se é positivo ou negativo, respectivamente e serve como função de ajuste da rede.

O principal objetivo da função de ativação é trazer não-linearidade ao sistema e assim, permitir resultados de valores de conjuntos de entrada mais complexos, mas a princípio, existem três tipos usuais de função (HAYKIN, 1998):

A função de limiar, representada na figura 20, cuja função é possível visualizar na equação (3.15), descreve um comportamento abrupto, proposto por McCulloch-Pitts (1943).

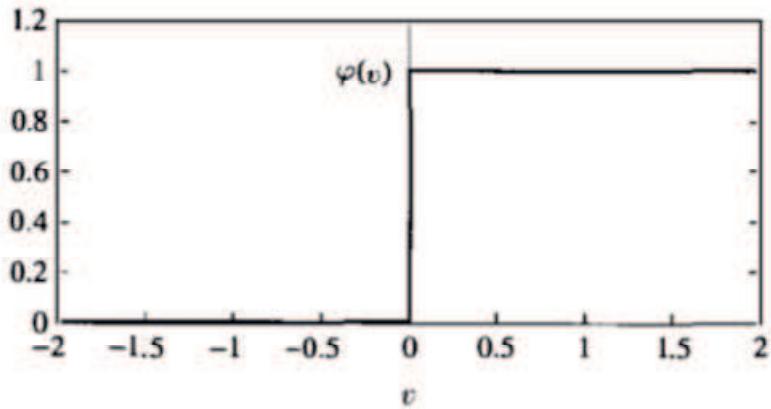


Figura 20 – função de ativação limiar (HAYKIN, 1998)

$$\varphi(v) = \begin{cases} 1, & \text{se } v \geq 0 \\ 0, & \text{se } v < 0 \end{cases} \quad (3.15)$$

A Função linear por partes, representada na figura 21, cuja função está na equação 3.16, resulta num combinador linear se a operação for restrita à região linear e passa a se comportar como a função de limiar se o fator de amplificação a for infinitamente grande.

$$\varphi(v) = \begin{cases} 1, & \text{se } v \geq +0,5/a \\ av + 0,5, & \text{se } -0,5/a > av > +0,5/a \\ 0, & \text{se } v \leq -0,5/a \end{cases} \quad (3.16)$$

A função sigmóide, representada na figura 22, é uma função com forma de "s", descrita em 3.17, monotonamente crescente, que exibe um balanceamento adequado entre comportamento linear e não-linear.

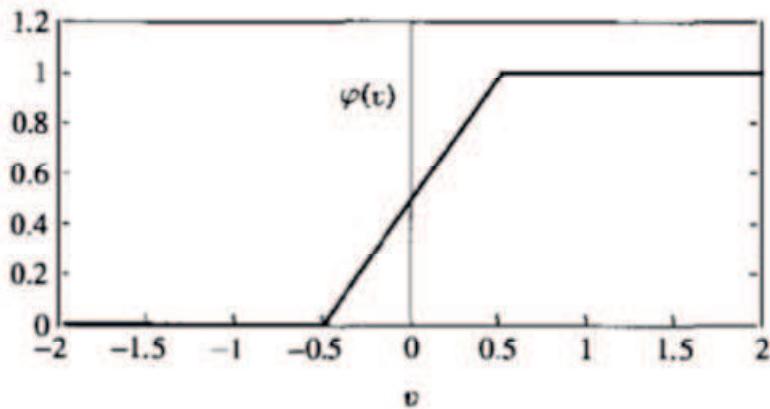


Figura 21 – função de ativação linear por partes (HAYKIN, 1998)

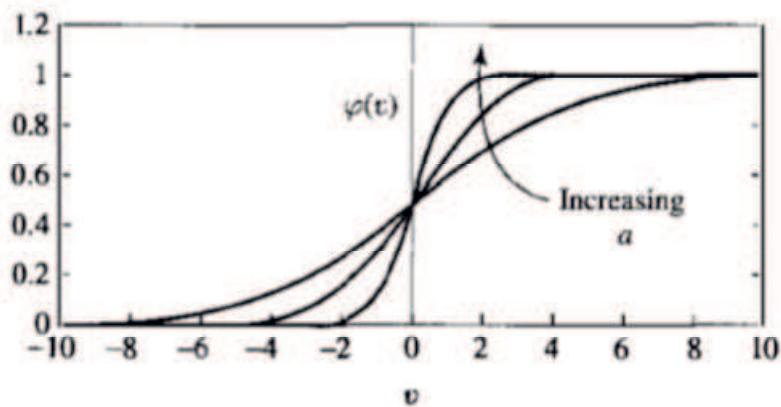


Figura 22 – função de ativação sigmoide (HAYKIN, 1998)

$$\varphi(v) = \frac{1}{1 + \exp(-av)} \quad (3.17)$$

Em uma Rede Neural Artificial, sua arquitetura é fundamental para seu desempenho, diante do problema a ser resolvido, podendo ser de uma única camada, também chamada de perceptron simples, utilizada para demandas resolvidas por linearidade, ou de multi-camadas, indicada para soluções não-lineares. Esse tipo de rede neural se distingue pela presença de uma ou mais camadas escondidas, cujos nós são chamados de neurônios ocultos ou camada oculta. A função dos neurônios ocultos é intervir entre a entrada externa e a saída da rede de alguma forma útil (HAYKIN, 1998).

Ao adicionar uma ou mais camadas ocultas, a rede está habilitada a extrair estatísticas ordem superior. De um modo geral, devido ao conjunto extra de conexões sinápticas e a dimensão extra de interações neurais, se consegue informações de ordem mais elevada. A capacidade de neurônios ocultos para extrair estatísticas de ordem mais elevada é particularmente valiosa quando o tamanho da camada de entrada é grande, como ocorre no caso do presente trabalho. Um exemplo de uma rede multi-camadas totalmente

conectado está representada na figura 23, onde se pode notar uma camada de entrada, por onde seriam inseridos os dados, uma camada oculta interconectada com todos os neurônios de entrada e saída e por fim, uma camada de saída, por onde se obtém o resultado do processamento da rede.

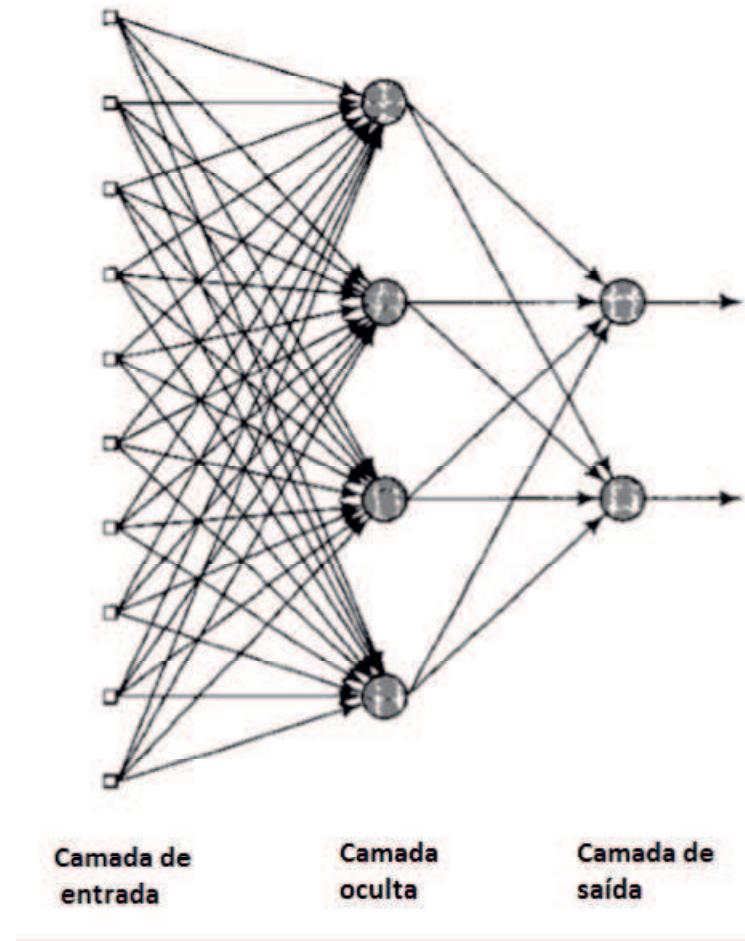


Figura 23 – Rede neural artificial multi-camadas (HAYKIN, 1998)

Apesar de ser conhecida desde a década de 80, redes de aprendizagem profundas (*deep learning*), atualmente diferem das redes neurais artificiais "canônicas" quanto às seguintes características: o maior número de neurônios, aumentando a complexidade dos modelos, maiores níveis de conexão entre as camadas, permitindo maior quantidade de parâmetros para gerenciar, necessitando maior poder computacional para o treinamento e extração automática de características (filtros).

Desta forma, *deep learning* pode ser definida como um tipo de rede neural com um grande número de parâmetros e de camadas e que possui uma ou mais das seguintes arquiteturas (PATTERSON; GIBSON, 2017):

- Redes pré-treinadas não supervisionadas
- Redes neurais convolucionais

- Redes neurais recursivas
- Redes neurais recorrentes

Na próxima seção, será abordada a rede neural convolucional, não sendo abordadas as outras arquiteturas, por não serem abarcadas pela proposta do presente trabalho.

3.1.4.2 Redes Neurais Convolucionais

Uma Rede Neural é dita Convolucional, quando em pelo menos uma das camadas, ocorre uma operação de convolução, que é uma operação matemática entre duas funções f e g , produzindo uma terceira função, que pode ser chamada por função modificada de f (FERREIRA, 2017), utilizada para detecção de bordas, suavização de imagens, extração de atributos, dentre outros usos (PARKER, 2010).

Redes Neurais Convolucionais, diferem das Redes Neurais tradicionais, com relação às seguintes características (MARTINS, 2017):

- **Tensores k -dimensionais de neurônios:** No caso de neurônios com 3 dimensões, cujos tensores possuem largura, altura e profundidade, os neurônios só estão conectados a uma pequena região da camada anterior, normalmente chamada de campo receptivo;
- **Conectividade local:** a RNC usa a correlação espacial local ao garantir um padrão de conectividade entre os neurônios das camadas adjacentes, produzindo uma forte resposta às características espaciais de entrada;
- **Pesos compartilhados:** Todo filtro é replicado no campo visual e possuem as mesmas parametrizações, formando um mapa de características. Isso garante que características de entrada sejam detectadas independentemente de sua posição no campo visual, permitindo às RNCs, a propriedade invariante por translação.

Como em pelo menos uma das camadas, deverá ocorrer uma convolução, para ser chamada de Rede Neural Convolucional, considerando $a_{x,y}$ uma imagem, a convolução em um ponto (x, y) de entrada é dado por:

$$w * a_{x,y} = \sum_l \sum_k w_{l,k} a_{x-l,y-k} \quad (3.18)$$

em que $*$ representa o símbolo operador da convolução e l e k , são as dimensões do kernel. Considerando que w possui tamanho 3 x 3, então a equação 3.18 pode ser expandida para:

$$w * a_{x,y} = w_{1,1}a_{x-1,y-1} + w_{1,2}a_{x-1,y-2} + \cdots + w_{3,3}a_{x-3,y-3} \quad (3.19)$$

ao renomear os índices de w e a , se obtém:

$$w * a_{x,y} = w_1 a_1 + w_2 a_2 + \cdots + w_9 a_9 \quad (3.20)$$

$$w * a_{x,y} = \sum_{j=1}^9 w_j a_j \quad (3.21)$$

Considerando que as equações 3.13 e 3.21 são idênticas, adicionando o viés (*bias*) e chamando o resultado de z , se obtém (GONZALEZ; WOODS, 2017a):

$$z = w * a_{x,y} + b = \sum_{j=1}^9 w_j a_j + b \quad (3.22)$$

Se considerar que z é a entrada de dados do neurônio, utilizando a analogia do modelo representado na figura 19, com h como função de ativação, a saída do neurônio se obtém através de:

$$a = h(z) \quad (3.23)$$

Inserindo os kernels na equação, 3.19, onde os índices sobrescritos são os referentes aos *feature maps*, os valores de l, k, x e y são os mesmos em todas as três equações, porque todos os três kernels possuem o mesmo tamanho e se movem conjuntamente (GONZALEZ; WOODS, 2017a), conforme:

$$\begin{aligned} w_{l,k}^{(1)} * a_{x,y}^{(1)} + w_{l,k}^{(2)} * a_{x,y}^{(2)} + w_{l,k}^{(3)} * a_{x,y}^{(3)} &= \sum_l \sum_k w_{l,k}^{(1)} a_{x-l,y-k}^{(1)} + \\ &+ \sum_l \sum_k w_{l,k}^{(2)} a_{x-l,y-k}^{(2)} + \sum_l \sum_k w_{l,k}^{(3)} a_{x-l,y-k}^{(3)} \end{aligned} \quad (3.24)$$

Como o kernel é formado por pesos no tamanho e formato adequados para que alimente a camada de entrada seguinte, então w e $a_{x,y}$ representam todos os valores correspondentes ao conjunto de dados da imagem ou filtro de entrada para a próxima camada. Durante o treinamento da rede convolucional, seguindo o fluxo adiante, o resultado da convolução de um kernel w e uma matriz de dados cujos valores são $a_{x,y}$, é disposto como:

$$\begin{aligned} z_{x,y} &= \sum_l \sum_k w_{l,k} a_{x-l,y-k} + b \\ z_{x,y} &= w * a_{x,y} + b \end{aligned} \quad (3.25)$$

em que l e k são as dimensões do kernel, x e y são as dimensões da entrada da matriz de dados e b é o viés (*bias*). O valor correspondente para $a_{x,y}$ é dado por (GONZALEZ; WOODS, 2017a):

$$a_{x,y} = h(z_{x,y}) \quad (3.26)$$

no entanto, esse $a_{x,y}$ é diferente do utilizado na equação 3.25, cujo $a_{x,y}$ representa valores das camadas anteriores, sendo necessário então uma notação adicional para diferenciação entre as camadas. Em uma camada totalmente conectada, é utilizado para este propósito, a notação ℓ e as equações 3.25 e 3.26 podem ser escritas como:

$$z_{x,y}(\ell) = \sum_l \sum_k w_{l,k}(\ell) a_{x-l,y-k}(\ell-1) + b(\ell)$$

$$z_{x,y}(\ell) = w(\ell) * a_{x,y}(\ell-1) + b(\ell) \quad (3.27)$$

e

$$a_{x,y}(\ell) = h(z_{x,y}(\ell)) \quad (3.28)$$

para $\ell = 1, 2, \dots, L_c$, onde L_c é a quantidade de camadas convolucionais e $a_{x,y}(\ell)$ são os valores dos filtros da camada convolucional ℓ . Quando $\ell = 1$,

$$a_{x,y}(0) = \text{valores dos pixels da imagem de entrada} \quad (3.29)$$

quando $\ell = L_c$,

$$a_{x,y}(L_c) = \text{valores dos filtros da última camada convolucional} \quad (3.30)$$

Redes neurais profundas utilizam treinamento por retropropagação, que consiste em quatro passos básicos: (1) inserção dos valores de entrada; (2) treinamento da rede por propagação, para definir todos os padrões de treinamento e determinar o erro de classificação ou regressão; (3) o retorno por retropropagação, que avalia a saída de erro através da rede e calcula as mudanças requeridas para atualização dos parâmetros; e (4) atualização dos pesos e viéses (*bias*) da rede. Esses passos são repetidos até que a rede atinja erros em níveis aceitáveis (GONZALEZ; WOODS, 2017a).

A equação para o erro, em função de x e y , é:

$$\delta_{x,y}(\ell) = \frac{\partial E}{\partial z_{x,y}(\ell)} \quad (3.31)$$

Relacionando a equação 3.31 com $\delta(\ell + 1)$, usando a regra da cadeia, se obtém:

$$\delta_{x,y}(\ell) = \frac{\partial E}{\partial z_{x,y}(\ell)} = \sum_u \sum_v \frac{\partial E}{\partial z_{u,v}(\ell + 1)} \frac{\partial z_{u,v}(\ell + 1)}{\partial z_{x,y}(\ell)} \quad (3.32)$$

em que u e v são as dimensões que abrangem os possíveis valores de z .

Por definição, o primeiro termo da dupla soma da equação 3.32 é $\delta_{u,v}(\ell + 1)$, logo, essa equação pode ser reescrita como:

$$\delta_{x,y}(\ell) = \frac{\partial E}{\partial z_{x,y}(\ell)} = \sum_u \sum_v \delta_{u,v}(\ell + 1) \frac{\partial z_{u,v}(\ell + 1)}{\partial z_{x,y}(\ell)} \quad (3.33)$$

Substituindo a equação 3.28 na equação 3.27 e usando esse resultado em 3.33, se obtém:

$$\delta_{x,y}(\ell) = \sum_u \sum_v \delta_{u,v}(\ell + 1) \frac{\partial}{\partial z_{x,y}(\ell)} \left[\sum_l \sum_k w_{l,k}(\ell + 1) h(z_{u-l,v-k}(\ell)) + b(\ell + 1) \right] \quad (3.34)$$

A derivada da expressão dentro dos colchetes será zero, a não ser no caso de $u - l = x$ e $v - k = y$, pelo fato da derivada de $b(\ell + 1)$ em relação a $z_{x,y}(\ell)$ ser zero. Assim, se $u - l = x$ e $v - k = y$, então $l = u - x$ e $k = v - y$, de modo que

$$\delta_{x,y}(\ell) = \sum_u \sum_v \delta_{u,v}(\ell + 1) \left[\sum_{u-x} \sum_{v-y} w_{u-x,v-y}(\ell + 1) h'(z_{x,y}(\ell)) \right] \quad (3.35)$$

Considerando os valores de $u - x$ e $v - y$, constantes, dentro dos colchetes, a dupla soma pode ser escrita como $w_{u-x,v-y}(\ell + 1) h'(z_{x,y}(\ell))$ e a equação 3.35, reescrita como:

$$\begin{aligned} \delta_{x,y}(\ell) &= \sum_u \sum_v \delta_{u,v}(\ell + 1) w_{u-x,v-y}(\ell + 1) h'(z_{x,y}(\ell)) \\ \delta_{x,y}(\ell) &= h'(z_{x,y}(\ell)) \sum_u \sum_v \delta_{u,v}(\ell + 1) w_{u-x,v-y}(\ell + 1) \end{aligned} \quad (3.36)$$

A dupla soma na expressão acima, pode ser reescrita como uma convolução, portanto, obtemos:

$$\delta_{x,y}(\ell) = h'(z_{x,y}(\ell)) [\delta_{x,y}(\ell + 1) * w_{-x,-y}(\ell + 1)] \quad (3.37)$$

Os valores negativos dos índices de w , indicam que ele se encontra "refletido" em ambos os eixos espaciais, o que é o mesmo de w se encontrar rotacionado em 180° . Assim, a expressão para o erro da camada ℓ é equivalente a (GONZALEZ; WOODS, 2017a):

$$\delta_{x,y}(\ell) = h'(z_{x,y}(\ell)) [\delta_{x,y}(\ell + 1) * rot180(w_{x,y}(\ell + 1))] \quad (3.38)$$

mas como os kernels não dependem de x e de y , a equação 3.38 pode ser escrita como:

$$\delta_{x,y}(\ell) = h'(z_{x,y}(\ell)) [\delta_{x,y}(\ell + 1) * rot180(w(\ell + 1))] \quad (3.39)$$

Como o erro E está em função dos pesos e viéses, se obtém:

$$\begin{aligned} \frac{\partial E}{\partial w_{l,k}} &= \sum_x \sum_y \frac{\partial E}{\partial z_{x,y}(\ell)} \frac{\partial z_{x,y}(\ell)}{\partial w_{l,k}} \\ &= \sum_x \sum_y \delta_{x,y}(\ell) \frac{\partial z_{x,y}(\ell)}{\partial w_{l,k}} \\ &= \sum_x \sum_y \delta_{x,y}(\ell) \frac{\partial}{\partial w_{l,k}} \left[\sum_l \sum_k w_{l,k}(\ell) h(z_{x-l,y-k}(\ell - 1)) + b(\ell) \right] \\ &= \sum_x \sum_y \delta_{x,y}(\ell) h(z_{x-l,y-k}(\ell - 1)) \\ &= \sum_x \sum_y \delta_{x,y}(\ell) a_{x-l,y-k}(\ell - 1) \end{aligned} \quad (3.40)$$

Comparando com a equação 3.27, nota-se a inversão do sinal entre as funções de soma e os índices correspondentes. Para colocar na forma de convolução, a última linha é reescrita como:

$$\begin{aligned} \frac{\partial E}{\partial w_{l,k}} &= \sum_x \sum_y \delta_{x,y}(\ell) a_{-(x-l,y-k)}(\ell - 1) \\ &= \sum_x \sum_y \delta_{x,y}(\ell) * a_{-l,-k}(\ell - 1) \\ &= \sum_x \sum_y \delta_{x,y}(\ell) * rot180(a(\ell - 1)) \end{aligned} \quad (3.41)$$

Utilizando as equações 3.40 e 3.41 na equação de gradiente descendente, obtemos:

$$\begin{aligned} w_{l,k}(\ell) &= w_{l,k}(\ell) - \alpha \frac{\partial E}{\partial w_{l,k}} \\ &= w_{l,k}(\ell) - \alpha \delta_{i,k}(\ell) * rot180(a(\ell - 1)) \end{aligned} \quad (3.42)$$

e,

$$\begin{aligned}
b(\ell) &= b(\ell) - \alpha \frac{\partial E}{\partial b(\ell)} \\
&= b(\ell) - \alpha \sum_x \sum_y \delta_{x,y}(\ell)
\end{aligned} \tag{3.43}$$

As equações 3.42 e 3.43 atualizam os pesos e viéses de cada camada convolucional em uma RNC. Como citado anteriormente, $w_{l,k}$ representam todos os pesos da camada, as variáveis l e k são as dimensões do espaço dos kernels e todos têm o mesmo tamanho (GONZALEZ; WOODS, 2017a).

Em uma RNC, a rede é inicializada com um conjunto de pequenos pesos e viéses aleatórios, sendo alimentada em seguida, com os dados de entrada. Na camada em que ocorre a convolução, esta pode ser interpretada como sendo o somatório da multiplicação de cada elemento da imagem e seus vizinhos locais, pelos elementos da matriz filtro da convolução, conforme ilustrado na Figura 24 (FERREIRA, 2017).

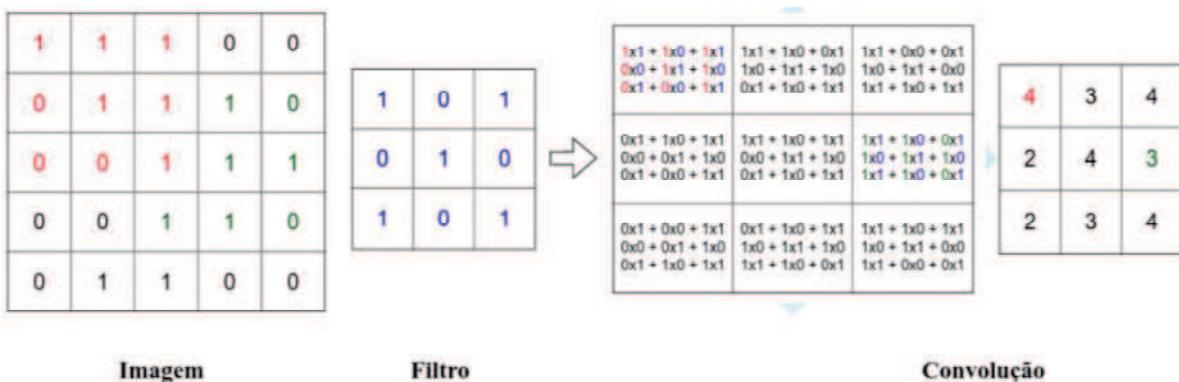


Figura 24 – No quadro da esquerda, uma imagem hipotética representada por um único canal com dimensões 5x5 que recebe a aplicação de um filtro 3x3. No quadro da direita, uma matriz ilustrando o somatório que fornece o resultado da convolução (adaptado de (FERREIRA, 2017))

Uma Rede Neural Convolucional é um tipo especializado de modelo, destinado a extrair de uma função (neste caso, uma imagem), características específicas, se utilizando de filtros, que são pequenos em relação à imagem, mas percorrem toda a extensão do volume da entrada. Assim, cada filtro é convoluído no volume de entrada e calcula os produtos escalares entre os valores do filtro e da entrada em qualquer posição, produzindo um mapa de ativação de duas dimensões. Ao final desse processo, os resultados dos filtros são dispostos em fila e alimentarão a entrada de dados da camada seguinte da rede neural (MARTINS, 2017).

Na tabela seguinte, estão dispostos os passos utilizados no treinamento de uma BNC.

Tabela 5 – Passos do treinamento de uma RNC.

Passo	Descrição	Equações
Passo 1	Entrada de imagens de treinamento	$a(0) = \text{conjunto de pixels da imagem para entrada na camada 1}$
Passo 2	Propagação do treino	Para cada neurônio correspondente (x, y) , em cada <i>feature map</i> , da camada ℓ , calcular $z_{x,y} = w(\ell) * a_{x,y}(\ell - 1) + b(\ell)$ e $a_{x,y(\ell)} = h(z_{x,y}(\ell))$; $\ell = 1, 2, \dots, L_c$
Passo 3	Retropropagação	Para cada neurônio de cada <i>feature map</i> , na camada ℓ , calcular $\delta_{x,y}(\ell) = h'(z_{x,y}(\ell)) [\delta_{x,y}(\ell + 1) * \text{rot180}(w(\ell + 1))]$; $\ell = L_c - 1, L_c - 2, \dots, 1$
Passo 4	Atualização de parâmetros	Atualiza os pesos e viéses para cada <i>feature map</i> , usando $w_{l,k}(\ell) = w_{l,k}(\ell) - \alpha \delta_{l,k}(\ell) * \text{rot180}(a(\ell - 1))$ e $b(\ell) = b(\ell) - \alpha \sum_x \sum_y \delta_{x,y}(\ell)$; $\ell = 1, 2, \dots, L_c$

4 Metodologia

Este capítulo descreve como a solução proposta foi abordada nesta pesquisa, qual arquitetura foi utilizada, como se deu a obtenção da base de dados e como foi realizado o treinamento da rede.

4.1 Introdução

Remoção de ruído é uma linha de pesquisa ampla, por ser um problema recorrente em imagens médicas e no caso de remoção de artefato metálico, o problema ainda se apresenta sem uma solução definitiva, mesmo depois de 40 anos de pesquisa. Implantes médicos como *stents* coronários, implantes ortopédicos, clipes cirúrgicos e restaurações dentárias metálicas, distorcem a visualização de estruturas anatômicas, de modo que as reconstruções das imagens sob esses efeitos têm sua utilidade prejudicada para o diagnóstico (PARK; CHOI; SEO, 2014).

Recentemente, redes neurais com aprendizagem profunda têm demonstrado grande sucesso no processamento de imagens médicas, com resultados qualitativos demonstrando o potencial da tecnologia sobre redução de artefatos e preservação da estrutura da imagem (CHEN et al., 2017).

Neste trabalho, por abordar o tratamento de imagens, visando à redução de ruído de artefato metálico na imagem, no domínio do espaço, foi escolhida a arquitetura de rede neural convolucional, para gerar imagens com artefatos metálicos atenuados a partir de arquivos de Tomografia Computadorizada corrompidos. Duas fases foram utilizadas para o alcance dos objetivos: a primeira fase foi a criação de uma base de dados, contendo imagens sem artefatos metálicos e imagens com artefatos, referentes ao mesmo trecho do objeto/paciente em estudo. A segunda fase, foi a modelagem da arquitetura e o treinamento de uma Rede Neural Convolutacional (RNC), com as informações alimentadas pela base de dados citada.

A base de dados foi formada por imagens oriundas do aparelho de Tomografia Computadorizada Phillips Brilliance 6, utilizado no Hospital Regional de Trauma Dom Luiz Gonzaga Fernandes, situado na cidade de Campina Grande/PB, o qual possui parceria com o Núcleo de Tecnologias Estratégicas em Saúde da Universidade Estadual da Paraíba, sendo escolhidas, inicialmente, as imagens sem qualquer tipo de artefato. Nestas imagens, foram inseridos através de script em MatLab, proposto por Titipong Kaewlek (KAEWLEK, 2012) e devidamente adaptado a este trabalho, pontos com características semelhantes aos metais utilizados em implantes e restaurações dentárias. Em seguida, após as imagens

serem reconstruídas com o algoritmo de retroprojeção filtrada (*Filtered BackProjection - FBP*), as imagens de TC, de 512x512 *pixels*, serão subdivididas em trechos de 64x64 *pixels* e agrupadas em dois arquivos serializados distintos, mas contendo índices idênticos, de modo a manter a integridade da associação entre as imagens.

Os artefatos metálicos inseridos nas imagens, obedecem às estruturas anatômicas aproximadas de HU, uma vez que o coeficiente da água, $\mu = 0.180\text{cm}^{-1}$, corresponde a 0 HU, enquanto que o coeficiente do osso, inicia em $\mu = 0.46\text{cm}^{-1}$, mas varia de acordo com suas características. As estruturas ósseas mais densas, como dentes, possuem coeficientes ainda mais elevados (PAULINO, 2017). Como os valores de densidade óssea variam entre 1000 HU e cerca de 3500 HU (DENNISON; FAHLMAN; MOORE, 2012), o que se comprovou na análise dos dados de imagens utilizadas neste trabalho, foram utilizados como parâmetros de artefatos metálicos para inserção nas imagens, valores em torno de 3900 HU.

Essas duas imagens serviram para treinar a rede neural convolucional, cujo modelo foi adaptado da RNC proposta, conforme (ZHANG; YU, 2017). Nesse trabalho citado, foi modelada uma arquitetura de RNC que realiza atenuação de artefatos metálicos, utilizando como entrada da rede, pacotes com três tipos de imagens: uma com estrias, reconstruída com FBP e oriunda de uma imagem sem metais, mas inseridos computacionalmente, outra entrada com mesma imagem tratada com Interpolação Linear (LI), mas com a ausência dos referidos metais e por fim, aquela mesma imagem, tratada pelo algoritmo *Beam Hardening Correction* (BHC), conforme (VERBURG; SECO, 2012) e como saída, imagens sem implantes metálicos, como parâmetro de treinamento. Na Figura 25, estão representados exemplos das amostras da base de dados utilizada por (ZHANG; YU, 2017), onde se percebe, na primeira linha, imagens de referência, sem ruídos. Na segunda linha, estão a simulação dos implantes metálicos a serem inseridos na imagem sem metais que se encontram na terceira linha, além da quarta linha, que possui a reconstrução da imagem segundo o algoritmo FBP, já com as estrias e sombreamentos provocados pelos artefatos. As linhas seguintes, são resultados das atenuações por *Beam Hardening Correction* e por Interpolação Linear, respectivamente.

4.2 Treinamento da RNC

A base de dados proposta em (ZHANG; YU, 2017) contou como sinal de entrada, uma imagem de três canais, formada pela imagem original sem correção, a imagem BHC e a LI. As amostras foram divididas em dois grupos, sendo um de treinamento e outro de teste, na proporção de 80% e 20%, contabilizando 8000 imagens de treinamento e 2000 de testes, respectivamente. Pequenos trechos das imagens alimentaram a rede neural, ou seja, o algoritmo seleciona trechos das imagens de tamanho $s \times t \times 3$, nos três canais

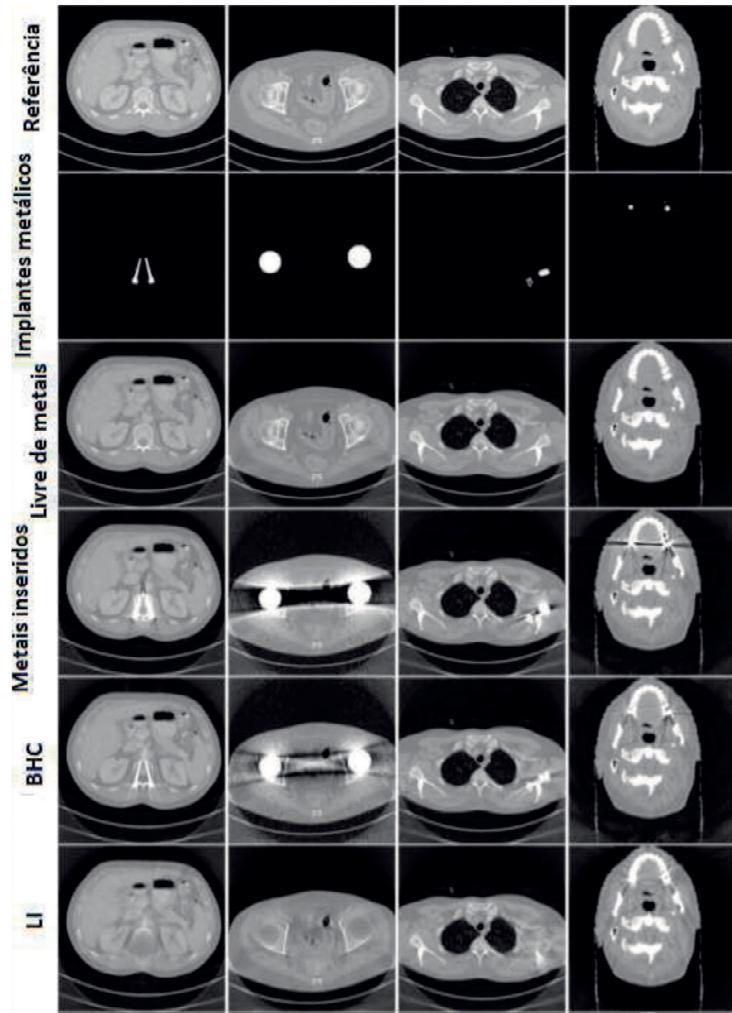


Figura 25 – Representação das amostras da base de dados: cada coluna corresponde a um caso. Adaptado de (ZHANG; YU, 2017)

de entrada, tomando o cuidado do trecho $s \times t$ ser o mesmo utilizado na imagem sem implantes metálicos (*metal-free*), como saída da RNC, durante o treinamento. Na prática, das imagens de 512×512 pixels, os trechos tinham tamanho de 64×64 , artifício utilizado para reduzir o custo computacional que se mostraria proibitivo caso fosse utilizada a imagem em seu tamanho original.

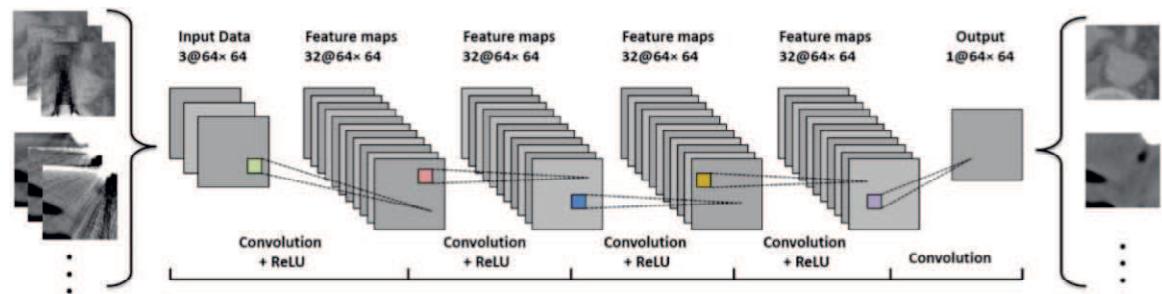


Figura 26 – Modelagem da arquitetura da rede neural convolucional para redução de artefato metálico (ZHANG; YU, 2017)

A figura 26 representa a estrutura da rede neural convolucional proposta (ZHANG; YU, 2017), onde se percebe a camada de entrada (*input data*) com 3 canais de 64 x 64 pixels, passando por uma função de ativação. As redes convolucionais normalmente utilizam a não-linearidade ReLU para garantir que a ativação dos *feature maps* seja sempre positiva e em alguns casos, esparsa. Para se obter reconstruções válidas, o sinal reconstruído é passado através de uma não-linearidade (PROTAS, 2017), definida como $\text{ReLU}(x) = \max(0, x)$. Essa função de ativação é executada após a primeira camada de entrada de dados.

Esta função de ativação é a recomendada para uso com a maioria das redes neurais *feedforward*. A aplicação dessa função à saída de uma transformação linear gera uma transformação não linear. No entanto, a função permanece muito próxima de linear, de modo que é uma função linear por partes com duas peças lineares. Como as unidades lineares retificadas são quase lineares, elas preservam muitas das propriedades que facilitam a otimização de modelos lineares com métodos baseados em gradiente. Eles também preservam muitas das propriedades que tornam os modelos lineares bem generalizados (GOODFELLOW et al., 2016).

A saída da camada de convolução juntamente com a função de ativação ReLU pode ser definida como:

$$C_l(\mathbf{u}) = \text{ReLU}(\mathbf{W}_l * C_{l-1}(\mathbf{u}) + \mathbf{b}_l),$$

para as camadas

$$l = 1, \dots, L - 1 \quad (4.1)$$

onde $*$ corresponde à convolução, \mathbf{W}_l e \mathbf{b}_l correspondem aos pesos (*weight*) e viéses (*biases*) da primeira camada, respectivamente, bem como, foi definido que $C_0(\mathbf{u}) = \mathbf{u}$. \mathbf{W}_l pode ser assumido como um n_l kernel convolutivo, com tamanho fixo de c_l x c_l . $C_l(\mathbf{u})$ gera novos *feature maps*, baseado na saída da camada $(l - 1)^{\text{ésima}}$. Na camada de saída, os *feature maps* são então utilizados para gerarem uma imagem assemelhada à imagem livre de artefatos metálicos usada anteriormente para alimentar a saída da RNC, como treinamento supervisionado. Assim, se obtém:

$$C_L(\mathbf{u}) = \mathbf{W}_L * C_{L-1}(\mathbf{u}) + \mathbf{b}_L \quad (4.2)$$

Depois da configuração da RNC, os parâmetros $\Theta = \mathbf{W}_1, \dots, \mathbf{W}_L, \mathbf{b}_1, \dots, \mathbf{b}_L$ são atualizados durante o treinamento e a estimativa desses parâmetros pode ser obtida pela minimização da função de perda (*loss*):

$$\text{loss}(\mathbf{U}, \mathbf{V}, \Theta) = \frac{1}{R} \sum_{r=1}^R \|C_L(\mathbf{u}_r) - \mathbf{v}_r\|_F^2 \quad (4.3)$$

em que $\mathbf{U} = (\mathbf{u}_1, \dots, \mathbf{u}_R)$ e $\mathbf{V} = (\mathbf{v}_1, \dots, \mathbf{v}_R)$, são os dados de entrada e saída desejada, respectivamente.

4.3 Modelo Proposto

A arquitetura proposta por este trabalho tem semelhanças com (ZHANG; YU, 2017), no entanto, não foram utilizados os canais os quais foram aplicadas tecnologias de atenuação por Interpolação Linear (LI) e *Beam Hardening Corretion* (BHC), logo, a entrada da rede neural convolucional foi um fragmento de 64 x 64 pixels, retirado das imagens de treinamento, cuja posição desse recorte é a mesma do fragmento da saída desejada, de modo a viabilizar a associação e correta reconstrução.

A base de dados para treinamento seguiu o disposto na Introdução, com a inserção artificial de valores compatíveis aos identificados em metais utilizados em implantes e reconstruídos com algoritmo FBP, ou seja, a imagem original teve seus dados, em determinados trechos, acrescidos com valores acima do valor máximo da imagem original, representando locais de alta densidade, para posteriormente serem reconstruídas. Após a reconstrução das imagens já com os artefatos inseridos artificialmente, cada uma foi sub-dividida em 64 pequenas imagens de 64x64 *pixels*, uma vez que o tamanho de cada imagem original é de 512x512 *pixels*. A Figura 27, é um exemplo da inserção artificial de artefatos metálicos.

As imagens foram distribuídas em duas pastas, uma contendo as imagens sem artefato e a outra contendo imagens com artefatos simulados. Imagens originais que já continham artefatos metálicos, não foram utilizadas para o treinamento, mas destinadas a testes.

A figura 28, é um exemplo de divisão da imagem de 512x512 *pixels* em 64 imagens de 64x64 *pixels*. Essas divisões foram utilizadas em todas as imagens, para se adequarem ao modelo da Rede Neural Convolucional proposto.

A arquitetura da RNC proposta neste trabalho está demonstrada na Figura 29. Inicialmente os dados de treinamento alimentam a rede no sinal de entrada, passando por 05 camadas convolucionais e com função de ativação ReLU, contendo 32 filtros de 3x3 *pixels* cada, passando para uma última camada convolucional, com 64 filtros de 3x3 e terminando com 02 camadas totalmente conectadas de tamanho 4096 (64x64), para enfim, a saída de 64x64 *pixels*. Nenhuma camada de *pooling* (camada de agrupamento ou sub-amostragem) foi utilizada entre as camadas convolucionais nesta arquitetura, objetivando evitar a perda de dados decorrente do processo, uma vez que neste caso, por se tratar de uma técnica de reconstrução por regressão e não de classificação, toda informação é preciosa, não sendo interessante realizar uma redução de escala dos filtros, perdendo dados (PROTAS, 2017). Diferente do trabalho de Yanbo Zhang e Hengyong Yu (2017), foram utilizadas 05

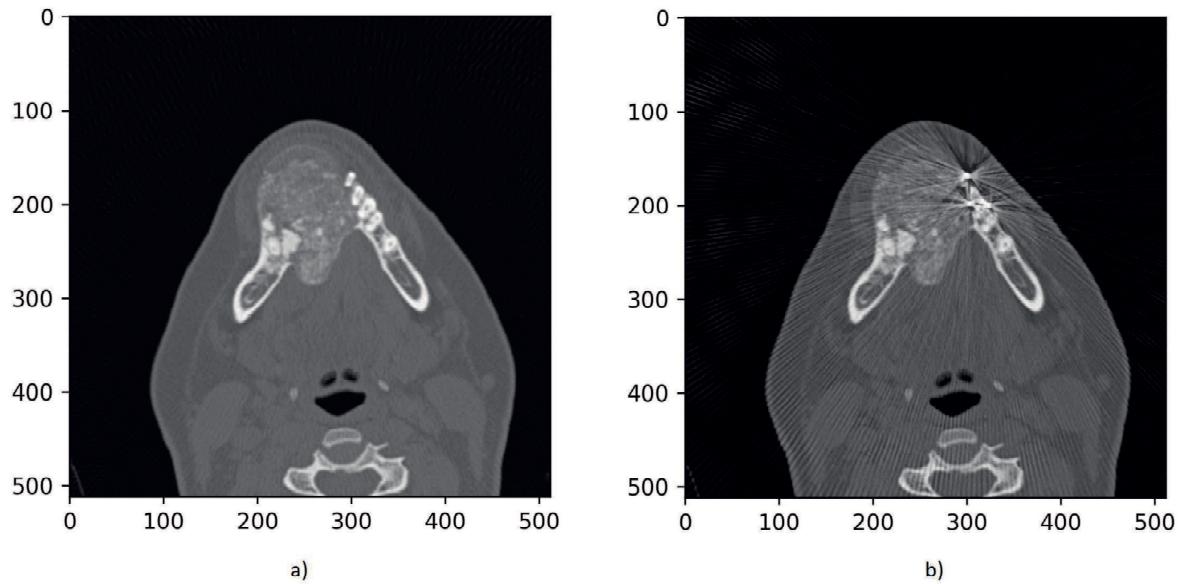


Figura 27 – A figura (a) sem artefato metálico e a figura (b), com artefato inserido artificialmente.

camadas em vez de 04, pois durante a fase de testes, os resultados com 04 camadas não foram satisfatórios. Foram utilizados 33088 pares de imagens de 64×64 pixels durante o treinamento, divididas conforme já citado.

De acordo com a metodologia já descrita, a imagem de saída foi comparada com a imagem sem artefato, fornecida durante o treinamento, sendo realizado o gradiente de otimização *AdamOptimizer* (KINGMA; BA, 2014), com taxa de aprendizado de $l = 0.0001$. Outras taxas de aprendizado, quais sejam: $l = 0.001$ e $l = 0.01$ foram utilizadas em treinamento, com resultados de qualidade inferior. A validação da imagem se deu através do cálculo do erro quadrático médio (*mean squared error - MSE*), conforme equação 4.4.

$$MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \quad (4.4)$$

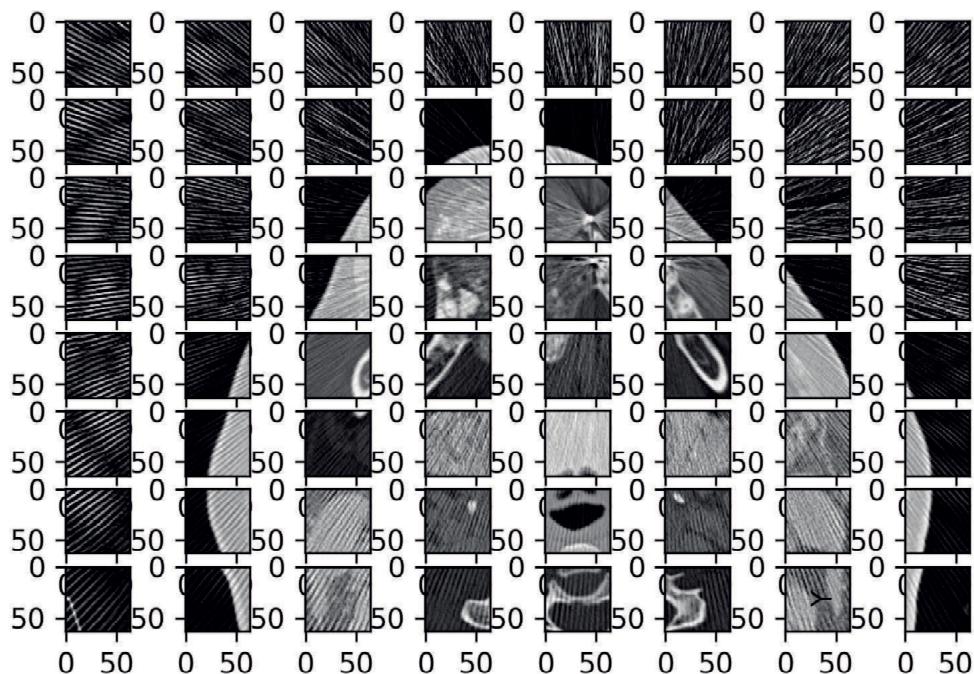


Figura 28 – Figura 27 (b), fragmentada em trechos de 64x64 pixels

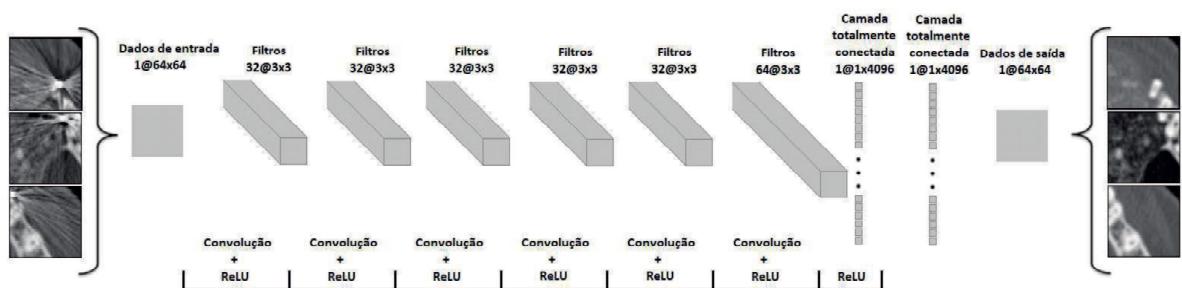


Figura 29 – Arquitetura do modelo proposto de rede neural convolucional

5 Resultados obtidos

Neste trabalho, foi proposto um modelo de Rede Neural Convolucional para reconstrução por regressão de imagens oriundas de tomografia computadorizada, com atenuação de ruídos provocados por artefatos metálicos. A abordagem escolhida, de buscar uma nova arquitetura se distanciando das redes pré-treinadas, tem por finalidade ampliar a discussão sobre modelos de reconstrução de imagens.

Após o treinamento da rede, modelada na linguagem Python e com uso da biblioteca de código aberto Tensorflow (ABADI et al., 2016), foi gerado um conjunto de pesos e filtros, que foram salvos e podem ser recuperados e aplicados na reconstrução da imagem, conforme a demanda. A esses arquivos, podem se somar novos dados, oriundos de novas etapas de treinamento, aumentando sua acurácia dentro do limite de treinamento, ou seja, com o cuidado de evitar *overfitting* ou sobreajuste, quando a rede não mais aprende, mas memoriza os dados. Essa característica é detectável quando ela passa a ter alto grau de acurácia nos dados de treinamento, mas não tão eficiente durante os testes, de modo que a rede passa apenas a repetir o comportamento do treinamento, sem que de fato passe a reproduzir um modelo de regressão, geração ou classificação.

Durante o treinamento, a cada 100 épocas (iterações), um trecho da imagem de treinamento foi exibido para se ter noção do progresso do treinamento. Na figura 30, temos um exemplo dessa avaliação, ocorrida na época 500, bem como, outros exemplos de avaliações que se seguem.

Apesar do alto ruído existente, nota-se claramente que a rede está alinhando os

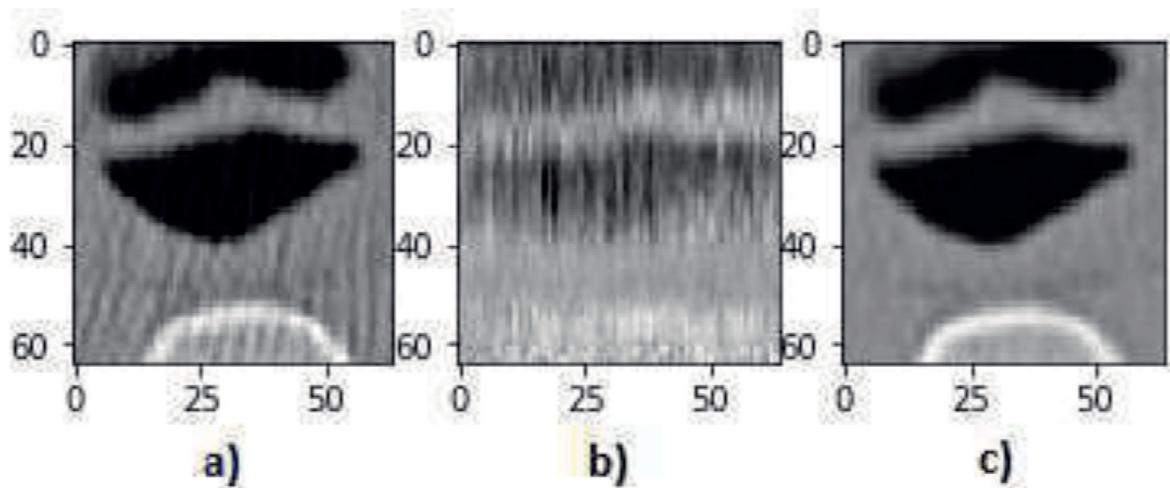


Figura 30 – Treinamento do modelo proposto, época 500. a) imagem de treino, b) saída da rede e c), imagem desejada.

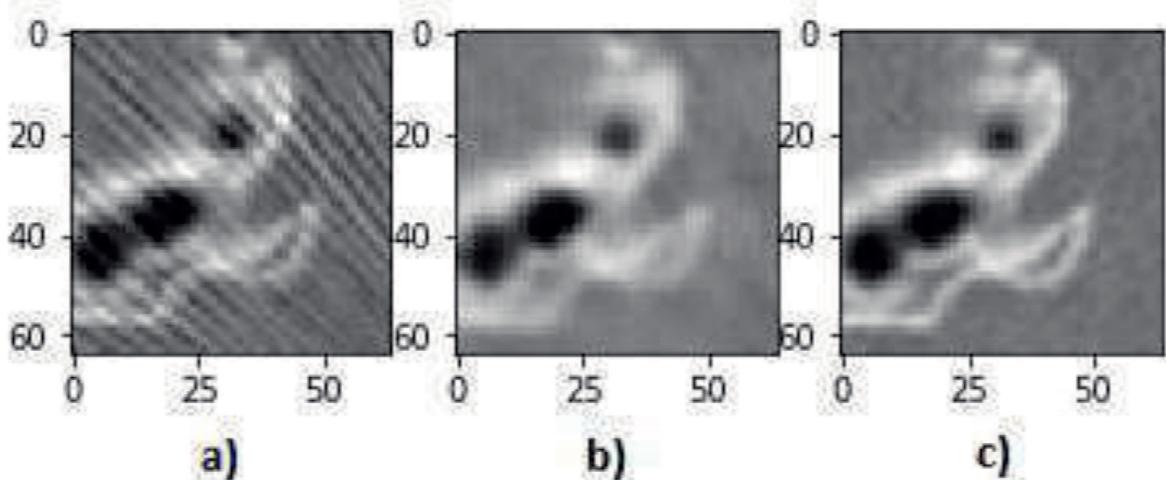


Figura 31 – Treinamento do modelo proposto, época 8292. a) imagem de treino, b) saída da rede e c), imagem desejada.

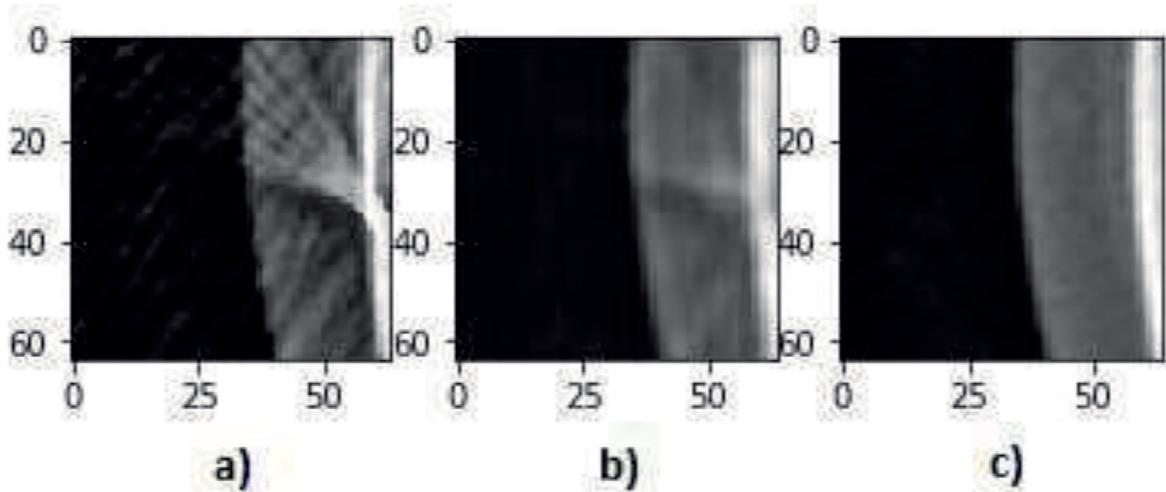


Figura 32 – Treinamento do modelo proposto, época 15536. a) imagem de treino, b) saída da rede e c), imagem desejada.

resultados com a imagem desejada, mesmo em uma época de iteração ainda inicial. Nas figuras 31, 32 e 33, já se visualiza uma maior semelhança da imagem de saída da rede b), com a imagem de saída desejada c), enquanto que na figura 34, se percebe a redução das estriadas provocadas pelo artefato metálico, que se encontra evidenciado em a) e reduzido em b), conforme a saída desejada c).

O erro médio quadrático, seguiu o comportamento descrito na figura ??, com taxa de aprendizagem em $l = 0.0001$. Inicialmente se nota alta taxa de erro, que decaiu conforme o treinamento, apesar de ainda haver uma oscilação considerável, ocorrida toda vez que era inserido um conjunto de imagens de um novo paciente. Foram utilizadas 33.088 pares de imagens para treinamento da rede, que durou 1 hora e 48 minutos. O computador

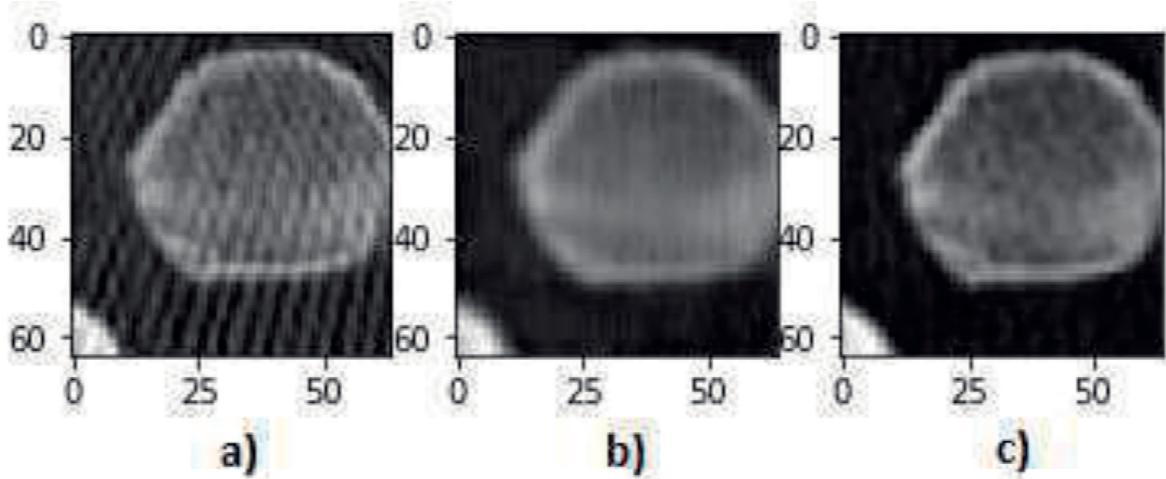


Figura 33 – Treinamento do modelo proposto, época 17148. a) imagem de treino, b) saída da rede e c), imagem desejada.

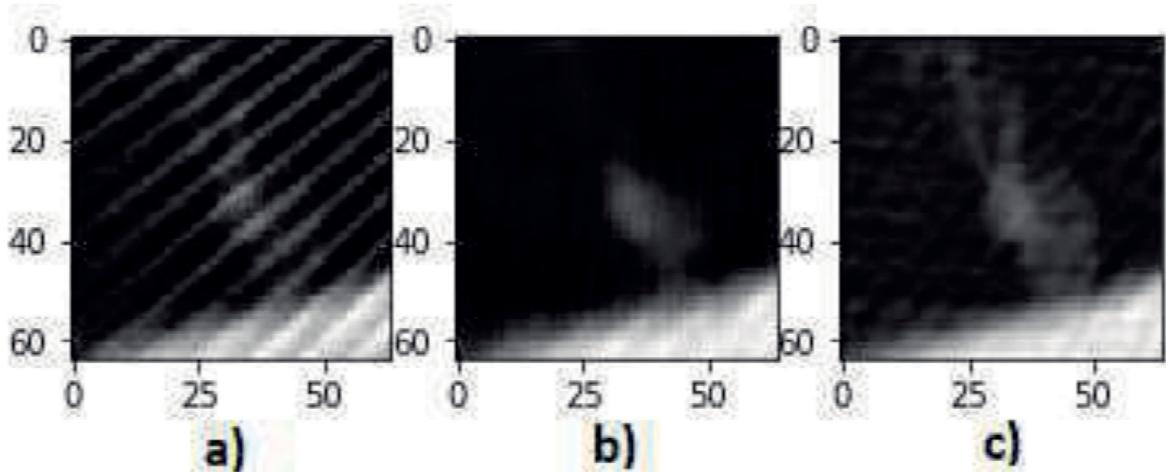


Figura 34 – Treinamento do modelo proposto, época 17784. a) imagem de treino, b) saída da rede e c), imagem desejada.

utilizado para o treinamento, foi um Samsung Core I3, com 4GB de memória RAM e 1 TB de espaço em disco. Os arquivos de recuperação da rede neural ocupam 3,8MB.

Na Figura 35, contém o gráfico do Erro Quadrático Médio durante a fase de treinamento. Observa-se algumas oscilações durante o decorrer do treinamento, que coincidem com as mudanças de pacientes, uma vez que para uma maior generalização, as imagens de vários pacientes alimentaram a rede em intervalos diferentes.

Após concluir o treinamento, quando todos os 33.088 pares de imagens foram submetidos ao treinamento, foi realizada avaliação, inserindo para reconstrução na rede uma imagem de tomografia computadorizada contendo artefato real e não o artefato simulado, utilizado pelo treinamento. Essas imagens de artefatos reais, presentes em algumas das fatias da tomografia computadorizada não foram utilizadas no treinamento,

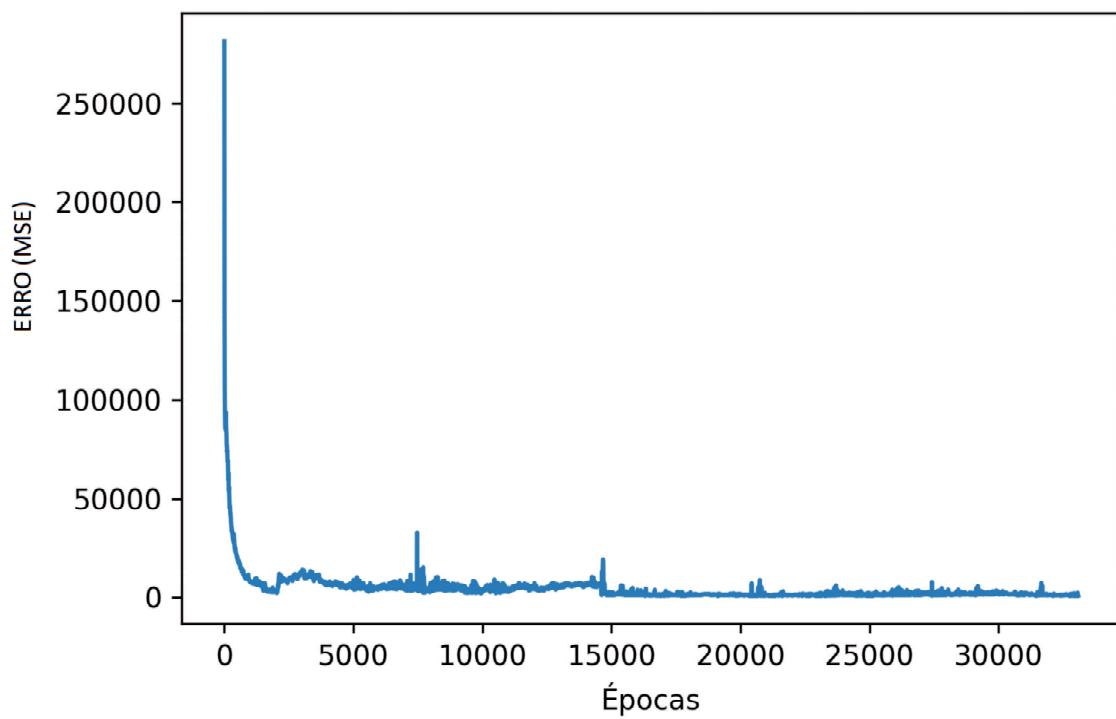


Figura 35 – Erro Quadrático Médio do treinamento com taxa de aprendizagem $l = 0.0001$

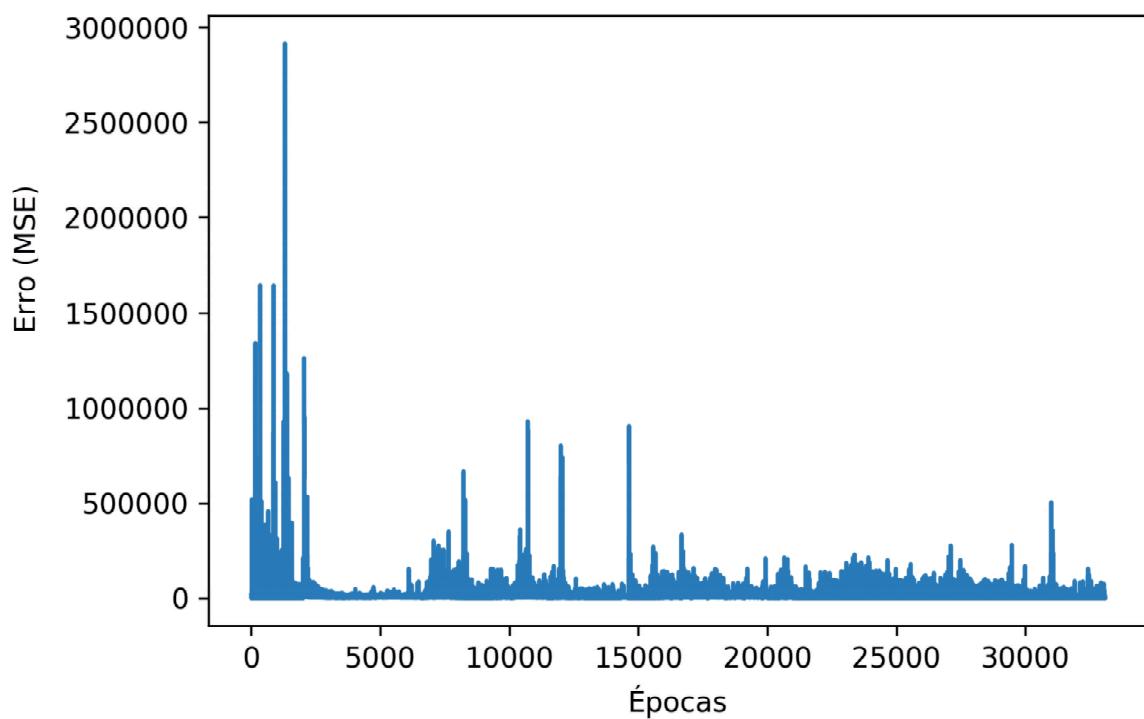


Figura 36 – Erro Quadrático Médio do treinamento com taxa de aprendizagem $l = 0.001$

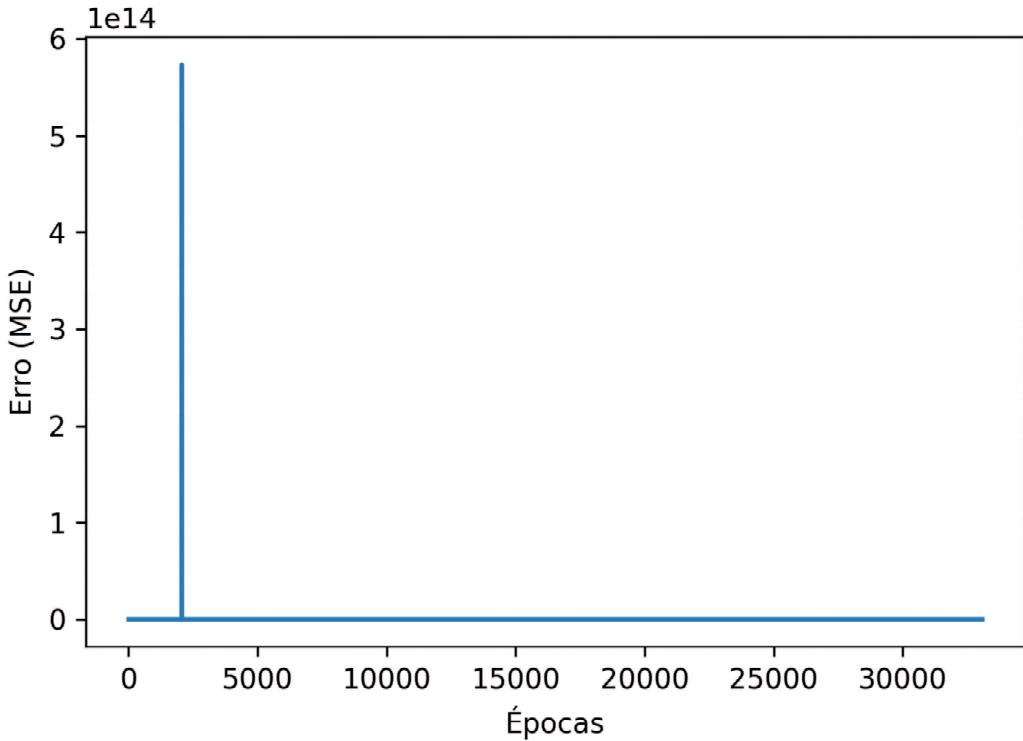


Figura 37 – Erro Quadrático Médio do treinamento com taxa de aprendizagem $l = 0.01$

para evitar a exposição da rede a valores anômalos, sendo utilizadas como validação ao final da execução. A seguir, algumas imagens com artefato real submetidas à reconstrução pela rede proposta, como consta na figura 38, demonstrando o resultado final da rede. Na figura 39, percebe-se que os artefatos de maior diâmetro, tem menor sucesso na construção, uma vez que na construção da base de dados, se deu ênfase a pontos reduzidos de artefatos. A rede apresenta bons resultados em estrias e sombreamentos dispersos na imagem.

Nas imagens de saída da RNC proposta, com taxa de aprendizagem $l = 0.001$, nota-se que cada trecho da imagem de tamanho 64x64, é visivelmente delimitado, causando um acréscimo de artefato na junção das bordas dos trechos, ao reconstruir a imagem completa.

Na Figura 42, o resultado exibido pela rede com taxa de aprendizagem em $l = 0.01$, se mostrou incompatível com o objetivo, exibindo dados totalmente anômalos, em comparação com o desejado.

Na figura 43 está exibido o detalhe do artefato metálico existente na imagem (a) e sua reconstrução na imagem (b). É possível visualizar que estrias e sombreamentos que estavam propagados por toda a imagem, foram visivelmente reduzidos. O metal existente no dente, também teve sua imagem suavizada, demonstrando a eficiência da técnica como reconstrução de imagens com artefatos metálicos. O processo de leitura da imagem, subdivisão em 64 imagens de 64x64 *pixels*, a leitura dos pesos e filtros da rede neural

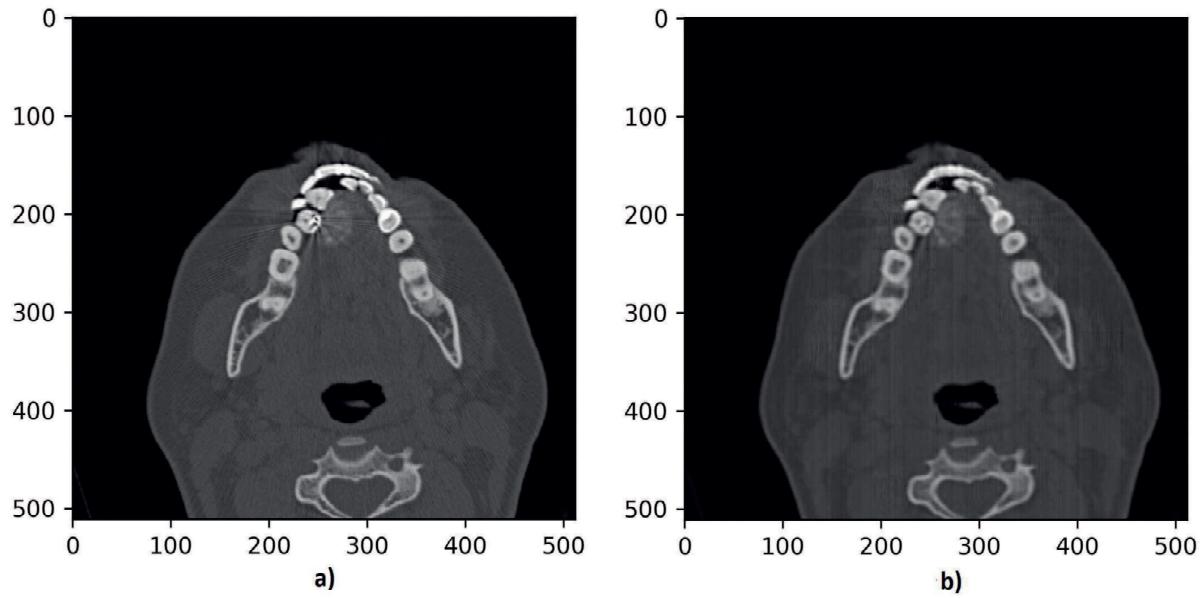


Figura 38 – Resultado final da rede neural proposta, com taxa de aprendizagem $l = 0.0001$.
 a) imagem com artefato real e b), imagem reconstruída pela rede.

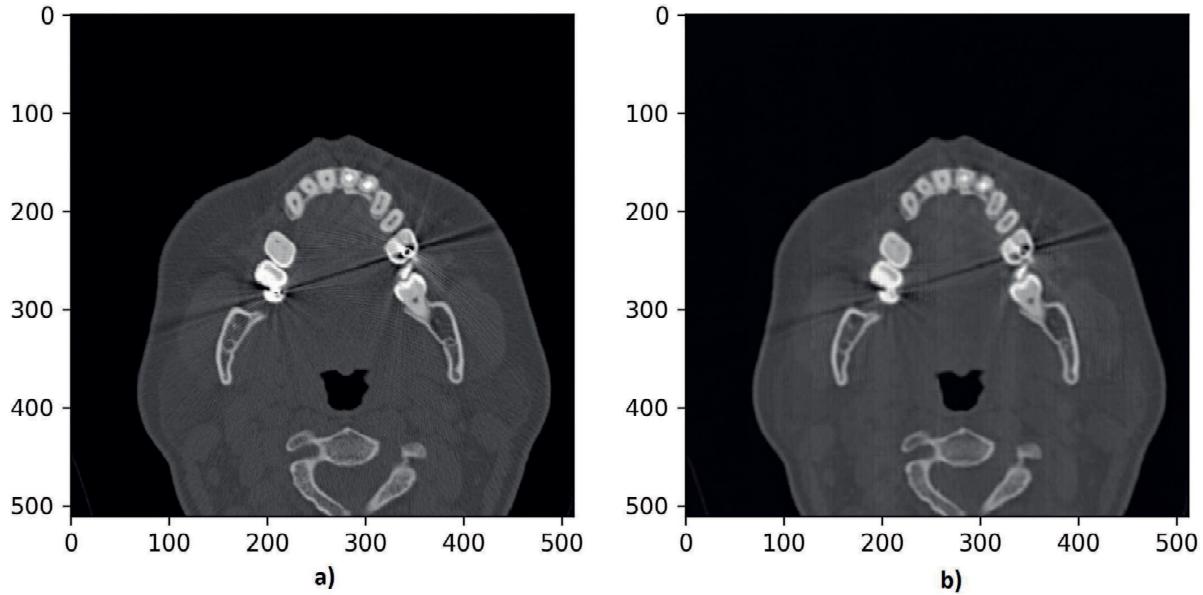


Figura 39 – Resultado final da rede neural proposta, com taxa de aprendizagem $l = 0.0001$.
 a) imagem com artefato real e b), imagem reconstruída pela rede.

convolucional proposta, a submissão da imagem à rede e sua respectiva saída, para a posterior reorganização dos 64 trechos em uma única imagem de 512x512, durou 5,91 segundos.

Os resultados deste trabalho foram comparados com os resultados da pesquisa também desenvolvida no NUTES, elaborada pelo pesquisador José Alberto Souza Paulino (2017). Nessa pesquisa, foi implementada uma solução para atenuação de artefatos metálicos

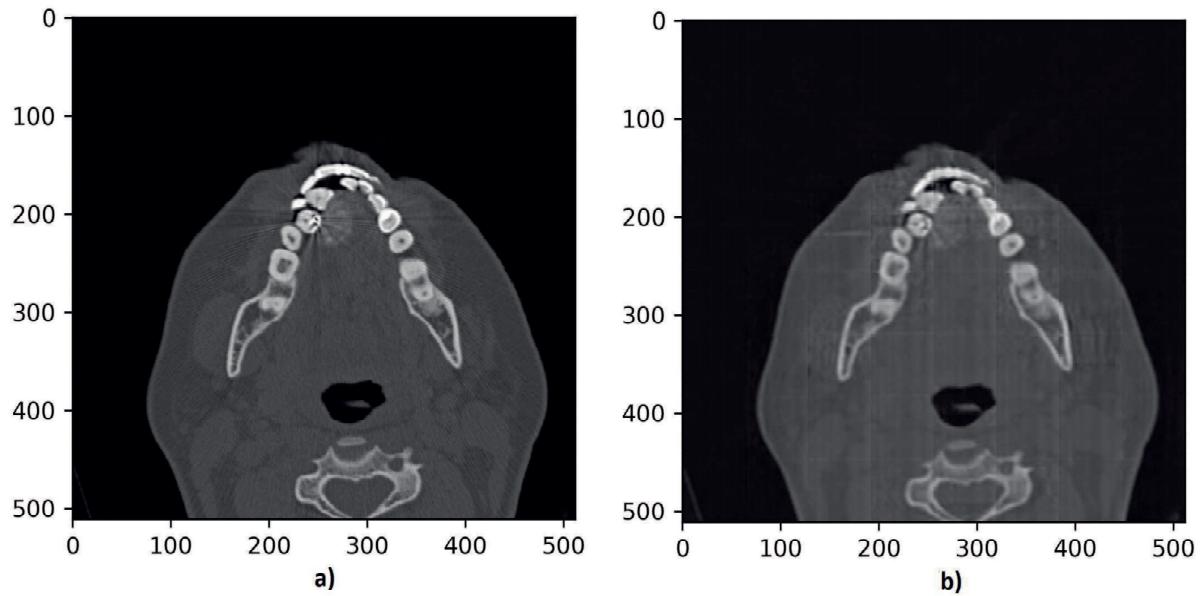


Figura 40 – Resultado final da rede neural proposta, com taxa de aprendizagem $l = 0.001$.
 a) imagem com artefato real e b), imagem reconstruída pela rede.

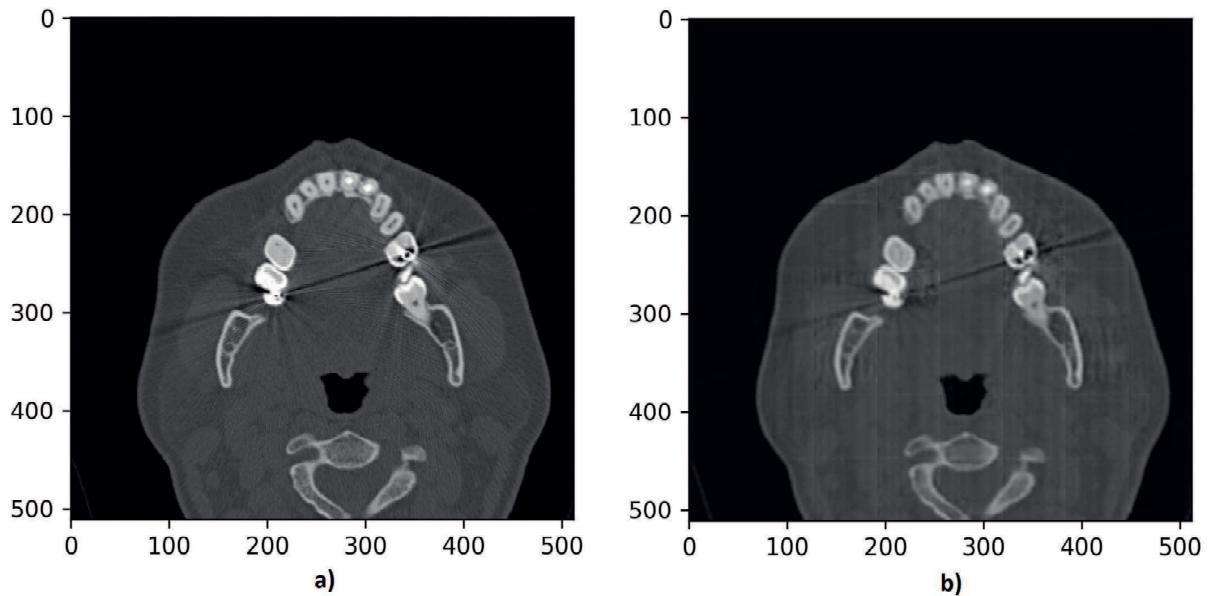


Figura 41 – Resultado final da rede neural proposta, com taxa de aprendizagem $l = 0.001$.
 a) imagem com artefato real e b), imagem reconstruída pela rede.

em tomografias computadorizadas, através do método de correção de sinograma com interpolação linear.

A solução proposta por (PAULINO, 2017), aplica a correção de sinograma por meio de interpolação dos dados corrompidos, substituindo os dados corrompidos nos traços metálicos por dados mais significativos com base em valores da vizinhança. Uma técnica semelhante foi proposta por (KOEHLER; BRENDEL; BROWN, 2012), conforme a Figura

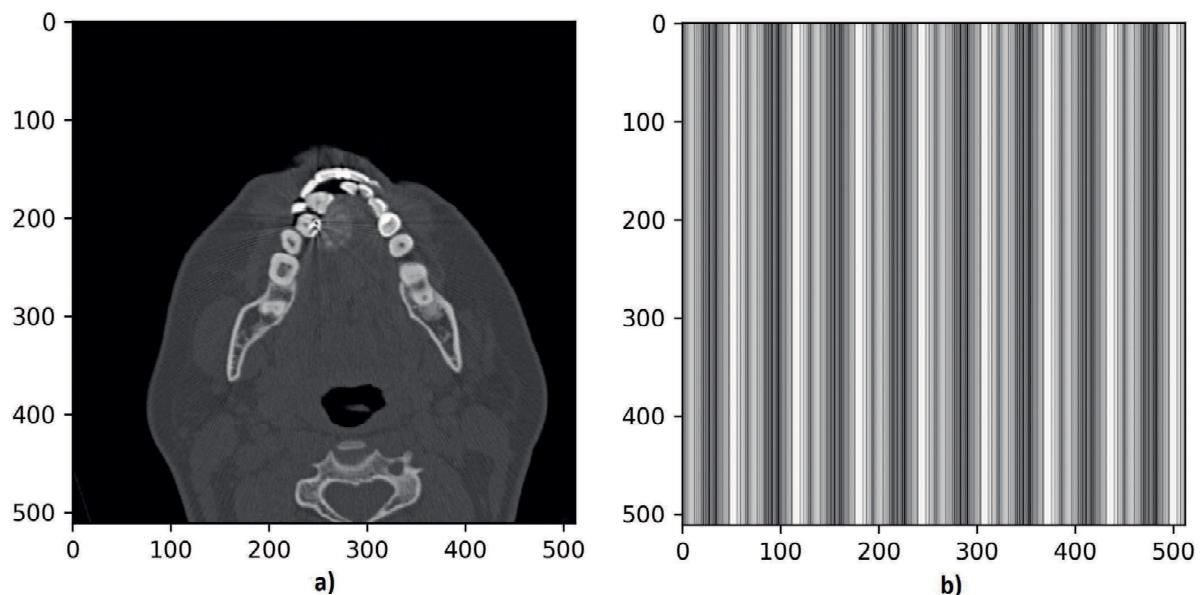


Figura 42 – Resultado final da rede neural proposta, com taxa de aprendizagem $l = 0.01$.
a) imagem com artefato real e b), imagem reconstruída pela rede.

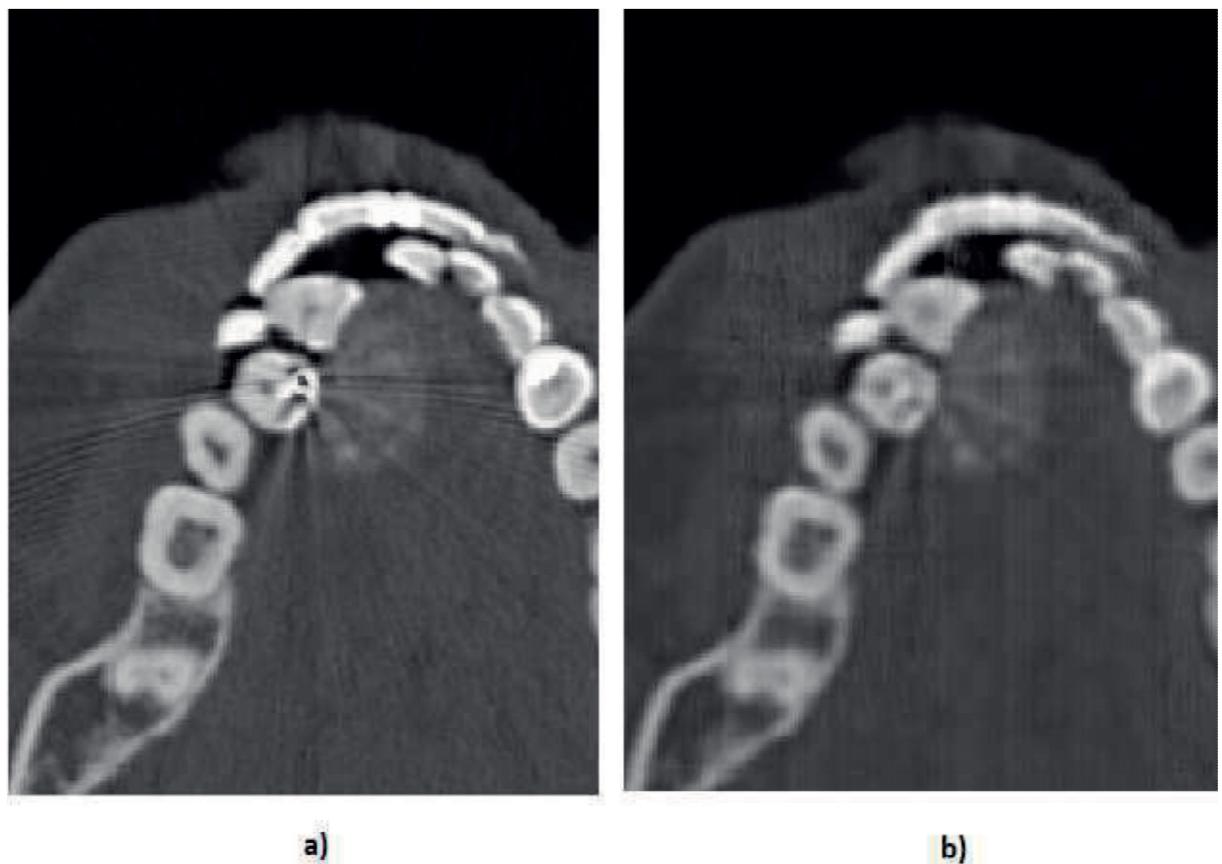


Figura 43 – Detalhe do resultado final da rede neural a) imagem com artefato real e b) imagem reconstruída pela rede

44.

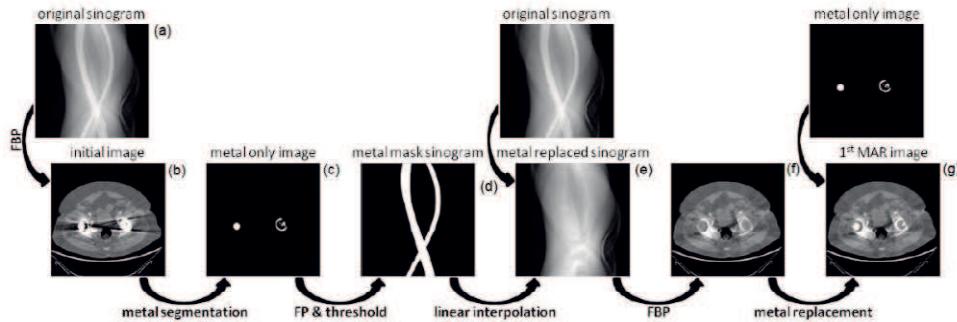


Figura 44 – Fonte: (KOEHLER; BRENDEL; BROWN, 2012)

Na Figura 44, é possível identificar as seguintes etapas da proposta utilizada por (PAULINO, 2017), quais sejam:

1. **Geração da imagem inicial:** realizada com base no sinograma original, Figura 44(a).
2. **Segmentação do objeto metálico:** por meio da segmentação por limiarização da imagem original, Figura 44(b), isola-se apenas os objetos metálicos, gerando uma nova imagem denominada de máscara, Figura 44(c).
3. **Geração de sinograma da máscara:** a partir da Transformada de Radon, busca-se gerar o sinograma do objeto metálico, Figura 44(d).
4. **Geração de sinograma da imagem original:** por meio da Transformada de radon, gera-se o sinograma da imagem inicial, Figura 44(a).
5. **Substituição dos traços metálicos:** reconstrução dos dados corrompidos no sinograma original por meio da interpolação linear para criação de um novo sinograma corrigido, Figura 44(e).
6. **Visualização da imagem:** após obter o sinograma interpolado, Figura 44(e), faz-se uso da Transformada Inversa de Radon para se obter a imagem reconstruída, Figura 44(f).
7. **Inclusão do objeto metálico:** a soma da imagem resultante, Figura 44(f), com a imagem do objeto metálico segmentado, Figura 44(c), resulta na imagem final corrigida pelo algoritmo MAR, Figura 44(g)

Foi realizada uma avaliação com imagens reais, ou seja, foram aplicadas em uma fatia de uma imagem de Tomografia Computadorizada contendo artefato metálico real, tanto a técnica proposta nesta pesquisa, quanto a proposta por (PAULINO, 2017) e disposta

para avaliadores compostos por estudantes de vários cursos da Universidade Estadual da Paraíba. Essa avaliação foi espontânea e os avaliadores receberam um formulário contendo a imagem original de TC, contendo artefato metálico e duas outras imagens, resultados do tratamento da imagem original, pelas técnicas já descritas, conforme Apêndice A. A avaliação qualitativa por meio do formulário foi a escala Likert(1932) que está exposta na Tabela 6 e os resultados estão dispostos no Apêndice B.

Tabela 6 – Critérios de Avaliação em escala de Likert

Valor	Avaliação	Descrição
6	Excelente	Uma imagem com altíssima qualidade tão boa quanto a original.
5	Boa	Imagen de alta qualidade, visualização agradável e interferência não desagradável.
4	Aceitável	Uma imagem com qualidade aceitável e interferência não desagradável.
3	Não apresenta melhora	imagem pobre em qualidade e que você deseja melhorar. A interferência desagrada de alguma forma.
2	Ruim	imagem muito pobre em qualidade, mas ainda poderia visualizar. Desagradável interferência presente.
1	Não utilizável	Imagen extremamente ruidosa, com alto nível de interferência e irrecuperável.

A Tabela 7, contém os principais dados estatísticos que descrevem os dados coletados durante a avaliação das técnicas utilizadas para atenuação de artefato metálico em imagens de Tomografia Computadorizada e neste trabalho comparadas. Nota-se uma sutil assimetria da distribuição dos resultados em ambas as tecnologias, uma vez que para a Interpolação Linear a assimetria foi positiva à direita, com valor de 0,5, enquanto que nos dados que representam a Rede Neural Convolucional, a assimetria foi negativa à esquerda, com valor de -0,052. A curtose ou o “achatamento” da curva em relação ao eixo central e consequentemente o espalhamento dos valores ao longo da variação dos dados, mostrou que a Interpolação Linear possui uma curtose de -0,241 e a Rede Neural Convolucional possui uma curtose de -1,098 e com ambos os valores menores do que 0,263, pode-se afirmar que a distribuição dos dados é leptocúrtica, ou seja, mais alongada e concentrada ao centro (BULMER, 1979).

Inicialmente se optou por verificar se os dados correspondem a uma distribuição normal, ou seja, se a maior frequência dos valores das amostras estaria próxima ao valor médio de todas as amostras, ou não. Para isso, foi realizado o teste de (SHAPIRO; WILK, 1965) e o teste Kolmogorov-Smirnov (KOLMOGOROV, 1933), cujos resultados estão dispostos na Tabela 8.

Tabela 7 – Descrição estatística dos dados dos formulários de avaliação qualitativa

		Estatística	Desvio Padrão
Interpolação Linear	Média	3,98	0,116
	95% de Intervalo de Confiança para Média	Limite inferior Limite superior	3,75 4,21
	5% da média aparada		3,93
	Mediana		4,00
	Variância		0,673
	Erro Padrão		0,820
	Mínimo		3
	Máximo		6
	Amplitude		3
	Amplitude interquartil		1
	Assimetria		0,500
	Curtose		-0,241
Rede Neural Convolucional	Média	4,14	0,118
	95% de Intervalo de Confiança para Média	Limite inferior Limite superior	3,90 4,38
	5% da média aparada		4,13
	Mediana		4,00
	Variância		0,694
	Erro Padrão		0,833
	Mínimo		3
	Máximo		6
	Amplitude		3
	Amplitude interquartil		2
	Assimetria		-0,052
	Curtose		-1,098

Tabela 8 – Testes de Normalidade

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Estatística	gl	Sig.	Estatística	gl	Sig.
Interpolação Linear	0,250	50	0,000	0,844	50	0,000
Rede Neural Convolucional	0,229	50	0,000	0,837	50	0,000

Considerando o nível de significância de 5% para avaliar as suposições do modelo estatístico proposto, verificou-se conforme a Tabela 8, que o valor de significância para ambos os testes, foi 0,000, menor do que 0,005, logo, os dados coletados não seguem um comportamento normal. Diante da não normalidade dos dados, o teste T de Student não seria o mais adequado, sendo utilizado o teste não paramétrico de (WILCOXON, 1945), assim, permitindo que se determine se duas amostras pareadas foram selecionadas a partir de populações que têm a mesma distribuição, no caso, as amostras colhidas das avaliações das técnicas abordadas. Os resultados do teste de Wilcoxon apontam para a equivalência das tecnologias utilizadas, conforme Tabela 9.

Tabela 9 – Sumarização de Teste de Hipótese

	Hipótese nula	Teste	Sig.	Decisão
1	A mediana de diferenças entre a Interpolação Linear e a Rede Neural Convolucional é igual a 0	Teste dos pontos sinalizados de Wilcoxon	0,387	Reter a hipótese nula

6 Considerações finais

Foi proposta uma arquitetura de Rede Neural Convolucional para atenuar artefatos metálicos em imagens de Tomografia Computadorizada, sendo elaborado um modelo com 06 camadas convolucionais e 02 camadas totalmente conectadas com o fim de produzir imagens com atenuações de artefatos metálicos no tamanho de 64×64 *pixels* a partir de dados de treinamento previamente fornecidos. Os dados de treinamento foram criados em pares de imagens de tamanho 64×64 *pixels* e desses pares, um continha simulação de artefato metálico e o outro, era uma imagem de boa qualidade, sem ruído. Assim, a RNC teve como sinal de entrada, trechos de imagens com artefatos metálicos e o sinal de saída desejado, era o par equivalente ao da entrada, mas sem qualquer artefato metálico. Essas imagens foram criadas a partir de fatias de tomografia computadorizada, sem ruído provocado por artefatos, para a correta associação sobre o que a rede deveria aprender e assim, reproduzir a reconstrução da imagem, sem o ruído provocado pelo artefato metálico. O treinamento ocorreu em tempo consideravelmente reduzido, consumindo cerca de quase 1 hora e 48 minutos, sendo utilizados 33088 pares de imagens. As diferentes taxas de treinamento apresentaram resultados com qualidades diferentes, sendo o melhor resultado, com $l = 0.0001$, que permitiu uma junção homogênea entre os trechos de imagens de 64×64 *pixels*. A rede treinada com a taxa de aprendizagem $l = 0.001$, gerou imagens com artefatos secundários entre os trechos citados e a rede treinada com taxa $l = 0.01$, não conseguiu realizar treinamento.

Os artefatos simulados foram inseridos nos pontos de maior densidade da imagem, para simularem o mesmo comportamento, o que ocorreu com sucesso, no entanto, para trabalhos futuros, é interessante que se faça uma variação maior de tamanhos, formatos e densidades, permitindo assim, que a rede neural tenha maior capacidade de generalização, pois para o treinamento de pequenos pontos e pequenas estrias ao longo da imagem, a RNC se mostrou eficiente, no entanto, para artefatos de maior diâmetro, ainda permaneceram artefatos consideráveis.

A avaliação qualitativa dos dados foi realizada através de formulário contendo comparativo entre duas técnicas desenvolvidas no NUTES e mensuradas conforme escala Likert(1932). Os resultados dos formulários foram avaliados estatisticamente, ficando evidenciada a equivalência das técnicas utilizadas.

Aprendizado de máquina tem se mostrado eficaz em solucionar problemas com redes neurais, devido à capacidade de generalização e assim, encontrar soluções complexas as quais não havia se deparado antes, mas necessita de uma grande variedade de dados na etapa de treinamento, de modo a poder modelar e resolver as demandas às quais forem

submetidas.

O modelo proposto foi capaz de ser treinado em um computador doméstico, demonstrando que a arquitetura utilizada, com pequenos filtros de tamanho 3x3, é escalável e eficaz para a proposta.

Referências

- ABADI, M. et al. Tensorflow: a system for large-scale machine learning. In: *OSDI*. [S.l.: s.n.], 2016. v. 16, p. 265–283. Citado na página 61.
- ABDOLI, M. et al. Reduction of artefacts caused by hip implants in ct-based attenuation-corrected pet images using 2-d interpolation of a virtual sinogram on an irregular grid. *European journal of nuclear medicine and molecular imaging*, Springer, v. 38, n. 12, p. 2257–2268, 2011. Citado 2 vezes nas páginas 9 e 32.
- BIRKFELLNER, W. *Applied medical image processing: a basic course*. [S.l.]: CRC Press, 2015. Citado 5 vezes nas páginas 15, 17, 19, 20 e 21.
- BITTENCOURT, G. *Inteligência artificial: ferramentas e teorias*. [S.l.: s.n.], 2006. Citado na página 41.
- BOAS, F. E.; FLEISCHMANN, D. Ct artifacts: causes and reduction techniques. *Imaging Med*, v. 4, n. 2, p. 229–240, 2012. Citado 4 vezes nas páginas 9, 27, 28 e 29.
- BULMER, M. G. *Principles of statistics*. [S.l.]: Courier Corporation, 1979. Citado na página 70.
- CHEN, H. et al. Low-dose ct via convolutional neural network. *Biomed. Opt. Express*, OSA, v. 8, n. 2, p. 679–694, Feb 2017. Disponível em: <<http://www.osapublishing.org/boe/abstract.cfm?>> Citado na página 54.
- CRAWFORD, C. R. et al. High speed reprojection and its applications. *Proc.SPIE*, v. 0914, 1988. Citado na página 31.
- DAFFNER, R. H.; HARTMAN, M. *Clinical radiology: the essentials*. [S.l.]: Lippincott Williams & Wilkins, 2014. Citado 2 vezes nas páginas 9 e 15.
- DENNISON, S.; FAHLMAN, A.; MOORE, M. The use of diagnostic imaging for identifying abnormal gas accumulations in cetaceans and pinnipeds. *Frontiers in physiology*, Frontiers, v. 3, p. 181, 2012. Citado na página 55.
- DOUGUERTY, G. *Digital image processing for medical applications*. [S.l.]: Cambridge University Press, 2009. Citado 5 vezes nas páginas 9, 15, 18, 20 e 21.
- FERREIRA, A. d. S. *Redes Neurais Convolucionais Profundas na Detecção de Plantas Daninhas em Lavoura de Soja*. Dissertação (Mestrado), 2017. Citado 3 vezes nas páginas 10, 47 e 52.
- GJESTEBY, L. et al. Metal artifact reduction in ct: where are we after four decades? *IEEE Access*, IEEE, v. 4, p. 5826–5849, 2016. Citado 5 vezes nas páginas 27, 29, 30, 31 e 33.
- GONZALEZ, R. C.; WOODS, R. E. *Digital image processing fourth edition*. [S.l.]: Pearson Education Limited, 2017. Citado 6 vezes nas páginas 9, 23, 48, 49, 50 e 52.

- GONZALEZ, R. C.; WOODS, R. E. *Digital image processing third edition*. [S.l.]: Pearson Education Limited, 2017. Citado 3 vezes nas páginas 9, 21 e 22.
- GOODFELLOW, I. et al. *Deep learning*. [S.l.]: MIT press Cambridge, 2016. v. 1. Citado na página 57.
- GRAY, W. et al. Implications of dental artifacts on radiotherapy planning for head and neck cancer. *Departments of Oral and Maxillofacial Surgery and Radiation Medicine*, Proc. Multidisciplinary Head Neck Symp, 2012. Citado na página 30.
- HAYKIN, S. *Neural Networks: A Comprehensive Foundation*. 2nd. ed. Upper Saddle River, NJ, USA: Prentice Hall PTR, 1998. ISBN 0132733501. Citado 6 vezes nas páginas 10, 42, 43, 44, 45 e 46.
- HEGAZY, M. A.; CHO, M. H.; LEE, S. Y. A metal artifact reduction method for a dental ct based on adaptive local thresholding and prior image generation. *Biomedical engineering online*, BioMed Central, v. 15, n. 1, p. 119, 2016. Citado na página 32.
- HSIEH, J. Computed tomography: principles, design, artifacts, and recent advances. SPI press, 2009. Citado na página 26.
- IZBICKI, R. *Machine Learning sob a ótica estatística: Uma abordagem preditivista para estatística com exemplos em R*. Departamento de Estatística - UFSCar, 2016. Disponível em: <<https://rizbicki.files.wordpress.com/2016/09/main2.pdf>>. Citado 8 vezes nas páginas 9, 10, 36, 37, 38, 39, 40 e 41.
- JAMES, G. e. a. *An Introduction to Statistical Learning*. [S.l.]: Springer, 2015. Citado 3 vezes nas páginas 10, 39 e 40.
- JIN, K. H. et al. Deep convolutional neural network for inverse problems in imaging. *IEEE Transactions on Image Processing*, IEEE, v. 26, n. 9, p. 4509–4522, 2017. Citado 3 vezes nas páginas 9, 26 e 27.
- KAEWLEK, T. [s.n.], 2012. Disponível em: <https://www.mathworks.com/matlabcentral/fileexchange/34589-created-metal-artifact-of-computed-tomography?s_tid=prof_contriblnk>. Acesso em: 8 mar. 2017. Citado na página 54.
- KALVIN, A. D.; WILLIAMSON, B. Using scout images to reduce metal artifacts in ct with application to revision total hip surgery. *Proc.SPIE*, v. 3034, p. 3034 – 3034 – 12, 1997. Disponível em: <<http://dx.doi.org/10.1117/12.274080>>. Citado na página 33.
- KHAN, F. M. *The Physics of Radiation Therapy*. [S.l.]: Lippincott Williams & Wilkins, 2010. Citado na página 14.
- KHAN, F. M.; GIBBONS, J. P. *Khan s the physics of radiation therapy*. [S.l.]: Lippincott Williams & Wilkins, 2014. Citado na página 18.
- KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. Citado na página 59.
- KOEHLER, T.; BRENDEL, B.; BROWN, K. A new method for metal artifact reduction. In: *The Second International Conference on Image Formation in X-ray Computed Tomography, June 24-27, 2012, Salt Lake City, Utah, USA; authors version*. [S.l.: s.n.], 2012. Citado 3 vezes nas páginas 11, 67 e 69.

- KOLMOGOROV, A. Sulla determinazione empirica di una legge di distributione). *Giornale dell' Istituto Italiano degli Attuari* 4, p. 93–91, 1933. Citado na página 70.
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: *Advances in neural information processing systems*. [S.l.: s.n.], 2012. p. 1097–1105. Citado na página 42.
- LECUN, Y. et al. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, IEEE, v. 86, n. 11, p. 2278–2324, 1998. Citado na página 42.
- LISLE, D. A. *Imaging for Students Third Edition*. [S.l.]: CRC Press, 2007. Citado 2 vezes nas páginas 9 e 19.
- MAN, B. D. *Iterative reconstruction for reduction of metal artifacts in computed tomography, Phd dissertation*. [S.l.]: KU Leuven, Belgium, 2001. Citado na página 30.
- MARTINS, P. B. M. Aplicação de redes neurais geradoras adversárias para colorização de imagens em preto e branco. 2017. Citado 3 vezes nas páginas 42, 47 e 52.
- MAZZA, L. O. Aplicação de redes neurais convolucionais densamente conectadas no processamento digital de imagens para remoção de ruído gaussiano. 2017. Citado na página 42.
- MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, Springer, v. 5, n. 4, p. 115–133, 1943. Citado na página 41.
- MEDOFF, B. P. et al. Iterative convolution backprojection algorithms for image reconstruction from limited data. *J. Opt. Soc. Am.*, OSA, v. 73, n. 11, p. 1493–1500, Nov 1983. Disponível em: <<http://www.osapublishing.org/abstract.cfm?URI=josa-73-11-1493>>. Citado na página 33.
- MINSKY, M.; PAPERT, S. *Perceptrons*. MIT press, 1969. Citado na página 41.
- MOHRI, M.; ROSTAMIZADEH, A.; TALWALKAR, A. *Foundations of machine learning*. [S.l.]: MIT press, 2012. Citado na página 36.
- PARK, H.; CHOI, J. K.; SEO, J. K. Characterization of metal artifacts in x-ray computed tomography. *Communications on Pure and Applied Mathematics*, Wiley Online Library, v. 70, n. 11, p. 2191–2217, 2014. Citado 3 vezes nas páginas 9, 25 e 54.
- PARK, H. et al. Machine-learning-based nonlinear decomposition of ct images for metal artifact reduction. *arXiv preprint arXiv:1708.00244*, 2017. Citado na página 26.
- PARKER, J. R. *Algorithms for image processing and computer vision*. [S.l.]: John Wiley & Sons, 2010. Citado na página 47.
- PATTERSON, J.; GIBSON, A. *Deep Learning: A Practitioner's Approach*. [S.l.]: O'REILLY, 2017. Citado na página 46.
- PAULINO, J. A. S. Desenvolvimento de uma ferramenta para automatizar redução de artefato metálico em imagens de tomografias computadorizadas. Universidade Estadual da Paraíba, 2017. Citado 3 vezes nas páginas 55, 67 e 69.

- PRATI, R. C. *Novas abordagens em aprendizado de máquina para a geração de regras, classes desbalanceadas e ordenação de casos*. Tese (Doutorado) — Universidade de São Paulo, 2006. Citado na página 35.
- PROTAS, E. V. Visualização de camadas intermediárias de redes neurais convolucionais de transformação de imagem. 2017. Citado 2 vezes nas páginas 57 e 58.
- REN, S. et al. Faster r-cnn: Towards real-time object detection with region proposal networks. In: *Advances in neural information processing systems*. [S.l.: s.n.], 2015. p. 91–99. Citado na página 42.
- RONNEBERGER, O.; FIISCHER, P.; BROX, T. U-net: Convolutional networks for biomedical image segmentation. In: SPRINGER. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. [S.l.], 2015. p. 234–241. Citado na página 26.
- RÄTSCH, G. A brief introduction into machine learning. 21st Chaos Communication Congress., 2004. Citado na página 35.
- SCHULZE, R. et al. Artefacts in cbct: a review. *Dentomaxillofacial Radiology*, The British Institute of Radiology. 36 Portland Place, London, W1B 1AT, v. 40, n. 5, p. 265–273, 2011. Citado na página 26.
- SHALEV-SHWARTZ S, B.-D. S. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, 2014. Disponível em: <<https://www.cs.huji.ac.il/~shais/UnderstandingMachineLearning/understanding-machine-learning-theory-algorithms.pdf>>. Citado 3 vezes nas páginas 10, 38 e 39.
- SHAPIRO, S. S.; WILK, M. B. An analysis of variance test for normality (complete samples). *Biometrika*, JSTOR, v. 52, n. 3/4, p. 591–611, 1965. Citado na página 70.
- SHEPP, L. A.; LOGAN, B. F. The fourier reconstruction of a head section. *IEEE Transactions on nuclear science*, IEEE, v. 21, n. 3, p. 21–43, 1974. Citado 2 vezes nas páginas 25 e 34.
- TANG, Z.; HU, G.; ZHANG, H. Efficient metal artifact reduction method based on improved total variation regularization. *Journal of Medical and Biological Engineering*, v. 34, n. 3, p. 261–268, 2014. Citado na página 33.
- US, D. *Metal Artifact Reduction in Sinograms of Dental Computed Tomography*. [S.l.]: Tampere University of Technology, 2013. Citado na página 29.
- VERBURG, J. M.; SECO, J. Ct metal artifact reduction method correcting for beam hardening and missing projections. *Physics in medicine and biology*, IOP Publishing, v. 57, n. 9, p. 2803, 2012. Citado na página 55.
- WEBB, S. *The physics of medical imaging*. [S.l.]: CRC Press, 1988. Citado na página 20.
- WILCOXON, F. Individual comparisons by ranking methods. *Biometrics bulletin*, JSTOR, v. 1, n. 6, p. 80–83, 1945. Citado na página 72.

ZHANG, Y.; YU, H. Convolutional neural network based metal artifact reduction in x-ray computed tomography. *arXiv preprint arXiv:1709.01581*, 2017. Citado 5 vezes nas páginas 10, 55, 56, 57 e 58.

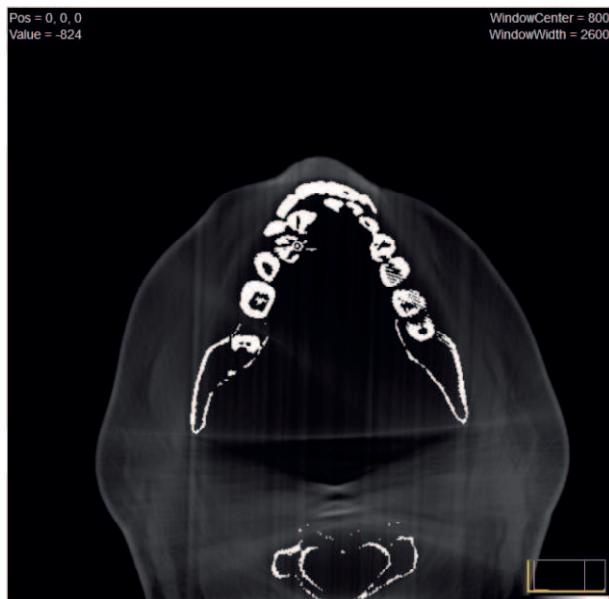
Apêndices

Apêndice A -

Fatia de tomografia computadorizada afetada por artefatos metálicos ocasionados pela presença de restaurações do tipo amálgama.

A imagem em questão foi submetida a dois métodos para redução de artefatos metálicos. Avalie o resultado de cada método com as notas de 1 a 6, de acordo com a legenda abaixo:

Valor	Avaliação
6	Excelente
5	Boa
4	Aceitável
3	Não apresenta melhora
2	Ruim
1	Não utilizável



6
5
4
3
2
1

Imagen corrigida através do método de Interpolação Linear.



Imagen corrigida através do método aprendizagem profunda com Rede Neural Convolucional.

6
5
4
3
2
1

Apêndice B -

	Interpolação Linear	Rede Neural Convolucionial
Avaliador 01	3	4
Avaliador 02	3	5
Avaliador 03	4	5
Avaliador 04	4	5
Avaliador 05	4	5
Avaliador 06	4	5
Avaliador 07	4	3
Avaliador 08	3	5
Avaliador 09	4	5
Avaliador 10	5	4
Avaliador 11	3	5
Avaliador 12	4	5
Avaliador 13	5	4
Avaliador 14	5	3
Avaliador 15	6	3
Avaliador 16	5	4
Avaliador 17	5	6
Avaliador 18	5	4
Avaliador 19	4	5
Avaliador 20	6	3
Avaliador 21	5	5
Avaliador 22	4	3
Avaliador 23	4	4
Avaliador 24	4	5
Avaliador 25	3	4
Avaliador 26	4	4
Avaliador 27	3	4
Avaliador 28	4	5
Avaliador 29	4	4
Avaliador 30	5	3
Avaliador 31	4	5
Avaliador 32	3	4
Avaliador 33	4	5
Avaliador 34	3	3
Avaliador 35	4	4
Avaliador 36	3	4
Avaliador 37	5	3
Avaliador 38	5	3
Avaliador 39	4	5
Avaliador 40	3	3
Avaliador 41	3	4
Avaliador 42	4	4
Avaliador 43	3	4
Avaliador 44	4	3
Avaliador 45	3	5
Avaliador 46	4	4
Avaliador 47	4	5
Avaliador 48	3	3
Avaliador 49	3	3
Avaliador 50	4	4