

Universidade Federal de Alagoas
Instituto de Computação



Dissertação de Mestrado

**PREDCGAN: Uma abordagem para geração de
nódulos pulmonares sintéticos GAN usando
Pré-Treinamento**

Bruno Henrique Lira dos Anjos
bhla@ic.ufal.br

Orientador:
Prof. Dr. Marcelo Costa Oliveira

Bruno Henrique Lira dos Anjos

**PREDCGAN: Uma abordagem para geração de
nódulos pulmonares sintéticos GAN usando
Pré-Treinamento**

Dissertação apresentada como requisito parcial para
obtenção do grau de Mestre pelo Curso de Mestrado
em Informática do Instituto de Computação da Uni-
versidade Federal de Alagoas.

Orientador:

Prof. Dr. Marcelo Costa Oliveira

Catálogo na fonte
Universidade Federal de Alagoas
Biblioteca Central
Divisão de Tratamento Técnico

Bibliotecário: Marcelino de Carvalho Freitas Neto – CRB-4 - 1767

S599p Anjos, Bruno Henrique Lira dos.
PREDCGAN : uma abordagem para geração de nódulos
pulmonares sintéticos GAN usando pré-treinamento / Bruno Henrique
Lira dos Anjos. – 2021.
37 f. : il.

Orientador: Marcelo Costa Oliveira.
Dissertação (mestrado em Informática) - Universidade Federal de
Alagoas. Instituto de Computação. Maceió, 2020.

Bibliografia: f. 32-37.

1. *Generative Adversarial Network*. 2. PREDCGAN. 3. Pré-treinamento.
4. Neoplasias pulmonares. I. Título.

CDU: 004.932:616.24-006



UNIVERSIDADE FEDERAL DE ALAGOAS/UFAL
Programa de Pós-Graduação em Informática – PPGI
Instituto de Computação/UFAL
Campus A. C. Simões BR 104-Norte Km 14 BL 12 Tabuleiro do Martins
Maceió/AL - Brasil CEP: 57.072-970 | Telefone: (082) 3214-1401



Folha de Aprovação

BRUNO HENRIQUE LIRA DOS ANJOS

PREDCGAN: NA APPROACH TO SYNTHETIC LUNG NODULE GENERATION WITH THE USE OF PRE-TRAINING

Dissertação submetida ao corpo docente do Programa de Pós-Graduação em Informática da Universidade Federal de Alagoas e aprovada em 27 de NOVEMBRO de 2020.

Banca Examinadora:

Prof. Dr. MARCELO COSTA OLIVEIRA
UFAL – Instituto de Computação
Orientador

Prof. Dr. THALES MIRANDA DE ALMEIDA VIEIRA
UFAL – Instituto de Computação
Examinador Interno

Prof. Dr. PAULO MAZZONCINI DE AZEVEDO MARQUES
USP – Universidade de São Paulo
Examinador Externo

Agradecimentos

À minha mãe por tudo. Sem o incentivo e a confiança que ela tem em mim eu jamais teria chegado até aqui. Agradeço especialmente pelos últimos dois anos, pois mesmo após grande mudanças nas nossas vidas ela sempre defendeu que o estudo é sempre o melhor caminho.

À minha avó que me apoiou a mudar totalmente o rumo da minha vida e me fez dar os primeiros passos para a computação.

À minha namorada pelo seu amor, pelo constante incentivo, pela compreensão e pela incansável ajuda ao longo desses sete anos, em especial, esse últimos dois anos em que sua fé auxiliou inúmeras vezes para eu chegar até esse momento.

Aos meus amigos ao longo do curso: Eduardo e Lucas, pelo auxílio em diversos momentos com ideias, correções e empréstimos "na baixa".

Ao meu orientador Marcelo pela oportunidade dada, pela dedicação em extrair o melhor de mim, pela paciência nos ensinamento e perante todos os problemas .

Aos professores que aceitaram o convite de participar da minha banca, Thales e Paulo.

A todos que me ajudaram direta ou indiretamente e torceram por mim.

Resumo

O diagnóstico de câncer de pulmão usando imagens de tomografia computadorizada (CT) é um trabalho complexo e desafiador, mesmo para radiologistas experientes. Para mitigar essa dificuldade, Sistemas de Diagnóstico Auxiliado por Computador (CAD) são desenvolvidos para auxiliar a performance dos especialistas. Uma das principais áreas utilizadas no CAD é a Aprendizagem Profunda (DL). Entretanto, por ser um processo de difícil validação e depender de um profissional bem treinado, a construção de grandes bancos de dados de imagens médicas estruturadas, bem validadas e abrangentes é um dos principais desafios das aplicações de DL no CAD em imagens médicas. A rede adversária generativa (GAN) consiste em um modelo para gerar dados probabilísticos heterogêneos que se apresenta como uma solução adequada para suprir o pequeno número de exames patológicos que compõem os mais diversos bancos de imagens médicas. Entretanto, os treinamentos das GAN para nódulos pulmonares não utilizam de informações condicionais para geração das imagens sintéticas. Assim, neste trabalho propomos a utilização da transferência de conhecimento utilizando um pré-treinamento da GAN para resultar em nódulos pulmonares sintéticos mais eficientes para aumento de base. Para isso, foi realizado o treinamento de uma *Deep Convolutional GAN* (DCGAN) em dois momentos, o primeiro para geração de nódulos segmentados e em um segundo momento foi feito um ajuste fino para gerar nódulos com as estruturas adjacentes. Através do uso das imagens geradas pela nossa proposta como aumento de base e de uma comparação com outras técnicas de aumento de base, obtivemos um aumento na AUC de 0.77% em relação ao aumento de base clássico e quando utilizada nossa proposta em conjunto com aumento de base clássico temos um aumento na AUC de 2.06% em relação ao individual do aumento de base clássico.

Palavras-chaves: Generative Adversarial Network; PREDCGAN; Pré-Treinamento; Câncer Pulmonar.

Abstract

The diagnosis of lung cancer using computed tomography (CT) images is a complex and challenging job, even for experienced radiologists. To mitigate this difficulty, Computer Aided Diagnostic (CAD) systems are developed to assist the performance of specialists. One of the main areas used in CAD is Deep Learning (DL). However, because it is a difficult process to validate and depends on a well-trained professional, the construction of large databases of structured, well-validated and comprehensive medical images is one of the main challenges of DL applications in CAD in medical images. The generative adversarial networks (GAN) consist of a model to generate heterogeneous probabilistic data that presents itself as an adequate solution to supply the small number of pathological exams that make up the most diverse medical image banks. However, GAN training for pulmonary nodules does not use conditional information to generate synthetic images. Thus, in this work it is proposed to use knowledge transfer applied through GAN pre-training to result in more efficient synthetic pulmonary nodules for base increase. For this, the training of a textit Deep Convolutional GAN (DCGAN) was carried out in two moments, the first to generate segmented nodules and in a second moment a fine adjustment was made to generate nodules with the adjacent structures. Through the use of the images generated by our proposal as a base increase and a comparison with other techniques of base increase, we obtained an increase in AUC of 0.77 % in relation to the increase of classic base and when used our proposal in conjunction with increase of classic base we have an increase in AUC of 2.06 % in relation to the individual of the increase of classic base.

Keywords: Generative Adversarial Network; PREDCGAN; Pre-Training; Lung Cancer.

Lista de Figuras

2.1	Espaço <i>voxel</i> . Fonte: Imagem extraída de (FALCÃO, 1993).	4
2.2	Exemplos das categorias de nódulos: sólido (a), semi-sólido (b) e não sólido. Fonte: Imagem extraída de (MOSMANN et al., 2016).	5
2.3	Lesões segundo as três classificações do LIDC-IDRI (ARMATO et al., 2011). Fonte: Imagem extraída de (ARMATO et al., 2011).	6
2.4	Hierarquia de Inteligência Artificial, Aprendizagem de Máquina e <i>Deep Learning</i> (DL). Fonte: elaborado pelo autor.	7
2.5	Funcionamento de um nó, sendo y o resultado calculado após a Função de ativação. Fonte: Imagem extraída de (CHARTRAND et al., 2017a).	8
2.6	Exemplo de extração de características de uma rede profunda. Fonte: Imagem adaptada de (TAIGMAN et al., 2014).	9
2.7	Esquema geral de uma GAN. Fonte: elaborado pelo autor.	9
2.8	Esquema das variantes GAN. (a) representa a GAN original, na qual z é o vetor de ruído, G é a rede geradora, D é a rede discriminante, X_g é a amostra sintética gerada por G , X_r é a amostra real e y_1 é a classificação entre real e sintético. (b) representa a DCGAN e a WGAN. (c) representa a cGAN. (d) representa a TubsGAN. (e) representa a M-GAN. (f) representa a F&BGAN. Fonte: elaborado pelo autor.	11
2.9	Exemplo de curva ROC. Fonte: elaborado pelo autor.	13
3.1	Esquema geral da PREDCGAN. Fonte: Elaborada pelo autor.	20
3.2	Fluxograma das etapas deste trabalho. Fonte: Elaborada pelo autor	21
3.3	Na primeira linha temos ROI dos nódulos não segmentados e na segunda linha a respectiva segmentação manual. Fonte: Elaborada pelo autor.	22
3.4	Arquitetura das GANs. Fonte: Elaborada pelo autor.	23
3.5	Arquitetura da CNN. Fonte: Elaborada pelo autor.	25
3.6	Imagens utilizadas no Teste de Turing Visual. Fonte: Elaborada pelo autor.	26
4.1	Amostragem dos Nódulos	29

Lista de Tabelas

2.1	Matriz de Confusão.	12
2.2	Resumo dos Artigo Mencionados na sessão Trabalhos Relacionados 2.7.	14
3.1	Quantidade de imagens por k-fold	25
4.1	Média dos resultados <i>k-folds</i> para os Conjuntos de Treinamento individuais, sensibilidade (SEN), especificidade (ESP), acurácia (ACC) e AUC-ROC (AUC).	27
4.2	Tabela 4.2 apresenta os resultados da classificação dos nódulos quando utilizamos apenas as técnicas para o aumento de base individualmente. Sensibilidade (Sen), especificidade (ESP) e AUC-ROC (AUC).	28
4.3	Resultados do <i>holdout</i> para os Conjuntos de Treinamentos mistos, sensibilidade (Sen), especificidade (ESP) e AUC-ROC (AUC).	28
4.4	Matriz de Confusão do Teste de Turing Visual	29
4.5	Teste de Turing Visual, em que GM é o grau de malignidade avaliada pelo radiologista e PR é a previsão do radiologista.	30

Lista de Abreviaturas

ACC	Acurácia
AUC	AUC-ROC
AM	Aprendizagem de Máquina
BNP	Banco de Nódulos Pulmonares
CAD	do inglês <i>Computer-Aided Diagnosis</i>
cGan	do inglês <i>Conditional Generative Adversarial Networks</i>
CI	do inglês <i>Confidence Interval</i>
CT	do inglês <i>Computed Tomography</i>
CNN	do inglês <i>Convolutional Neural Networks</i>
DL	do inglês <i>Deep Learning</i>
DCGAN	do inglês <i>Deep Convolutional Generative Adversarial Networks</i>
DCNN	do inglês <i>Deep Convolutional Neural Network</i>
DS	do inglês <i>Dice Score</i>
ESP	Especificidade
F&BGAN	do inglês <i>Forward and Backward Generative Adversarial Networks</i>
FN	Falso Negativo
FP	Falso Positivo
GAN	do inglês <i>Generative Adversarial Networks</i>

IDRI	do inglês <i>Image Database Resource Initiative</i>
LIDC	do inglês <i>Lung Image Database Consortium</i>
LIDC-IDRI	do inglês <i>Lung Image Database Consortium and Image Database Resource Initiative</i>
M-GAN	do inglês <i>Multi-channel Generative Adversarial Networks</i>
MR	do inglês <i>Magnetic Resonance</i>
PET	do inglês <i>Positron Emission Tomography</i>
RNA	Redes Neurais Artificiais
ROC	do inglês <i>Receiving Operating Characteristics</i>
ROI	do inglês <i>Region of Interest</i>
SEN	Sensibilidade
SVM	do inglês <i>Support Vector Machine</i>
TFP	Taxa de Falsos Positivos
TVP	Taxa de Verdadeiros Positivos
VN	Verdadeiro Negativo
VP	Verdadeiro Positivo
WGAN	do inglês <i>Wasserstein Generative Adversarial Networks</i>

Sumário

1	Introdução	1
1.1	Objetivo	3
1.2	Estrutura do Trabalho	3
2	Fundamentação Teórica	4
2.1	Nódulos Pulmonares em Imagens de Tomografia Computadorizada	4
2.2	Diagnóstico Auxiliado por Computador (CAD)	5
2.3	Base LIDC-IDRI	6
2.4	Aprendizagem Profunda	6
2.5	Rede Geradoras Adversárias	9
2.5.1	Outras Arquiteturas GANs	10
2.6	Avaliação de desempenho	12
2.7	Aplicações GANs	14
2.7.1	DCGAN	14
2.7.2	WGAN	16
2.7.3	GAN Clássica	16
2.7.4	cGAN	17
2.7.5	Outras GANs	17
3	Materiais e Métodos	19
3.0.1	PREDCGAN	19
3.0.2	Experimento	21
4	Resultados e Discussão	27
5	Conclusão	31
5.1	Contribuições Científicas	31
	Referências	37

1 Introdução

O câncer de pulmão é a neoplasia maligna mais letal no mundo com uma taxa anual estimada de 2,9 milhões de mortes (FERLAY et al., 2019). De acordo com a *American Cancer Society* (SIEGEL; MILLER; JEMAL, 2020), a taxa de sobrevivência de 5 anos em todos os estágios combinados é de 18,6%, mas diminui para 5% quando a doença se encontra em estágio avançado. Portanto, a detecção precoce do câncer pulmonar pode fornecer melhores alternativas de tratamento aos pacientes e aumentar suas chances de sobrevivência (OUDKERK et al., 2017).

O principal exame utilizado para detecção e diagnóstico do câncer pulmonar é a Tomografia Computadorizada (do inglês *Computed Tomography* - CT) (KALENDER, 2006). O diagnóstico de câncer de pulmão por meio de imagens de CT é um trabalho complexo e desafiador, mesmo para radiologistas experientes. Os nódulos pulmonares podem ser pequenos, ter baixo contraste em comparação com o tecido pulmonar circundante e podem estar ligados a estruturas anatômicas complexas. Além disso, a variação interobservador foi relatada em vários estudos (RIEL S.J., 2019; PINSKY et al., 2013). Isso aconteceu devido a vários aspectos, como exames de aumento de volume, restrições de tempo, falta de treinamento do leitor e fadiga (DEGNAN et al., 2019). Para mitigar essa dificuldade, sistemas de Diagnóstico Auxiliado por Computador (do inglês *Computer-Aided Diagnosis* - CAD) foram propostos para auxiliar a performance dos especialistas (YAMASHITA et al., 2018; PANG et al., 2020; MONTAGNON et al., 2020).

Com os recentes desenvolvimentos de algoritmos de inteligência artificial, um grande número de técnicas de Aprendizado Profundo (do inglês *Deep Learning* - DL) foi usado com sucesso em campos de processamento de imagens médicas (MCBEE et al., 2018). Entretanto, a aplicação de DL na área de imagens médicas é desafiadora, pois a técnica necessita ser treinada usando um grande conjunto de dados heterogênicos para evitar resultados enviesados (LIU et al., 2019). Questões de privacidade relacionadas ao paciente restringem o uso das imagens médicas, dificultando a construção de bases de imagens adequadas para o uso em DL (FRIDADAR et al., 2018). Ainda, gerar imagens médicas, além de ser um processo financeiramente custoso, a depender do exame, submetem os pacientes a altas doses de radiação e a procedimentos invasivos. Além disso, as imagens precisam ser diagnosticadas por profissionais bem treinados (CHARTRAND et al., 2017b).

Proposta por Goodfellow et al. 2014 (GOODFELLOW et al., 2014), a técnica de Redes Adversárias Generativas (do inglês *Generative Adversarial Networks* - GAN) é baseada na teoria

dos jogos e vem revolucionando a área de visão computacional com a sua capacidade de gerar imagens sintéticas de alta qualidade (YI; WALIA; BABYN, 2019). Imagens médicas sintéticas provenientes de uma GAN se apresentam como uma solução para suprir o reduzido número de imagens que compõe os mais diversos bancos de imagens médicas necessários ao treinamento de algoritmos de DL (GUAN; LOEW, 2019; GALBUSERA et al., 2018). Após serem treinadas, as redes GANs são capazes de gerar "imagens novas", ou seja, imagens que não estavam presentes na base de dados, aumentando a variabilidade dos dados. Essa característica representa um avanço em relação as técnicas clássicas de aumento de base em DL que adotam técnicas de transformações geométricas (e.g., rotação e escala) para aumentar o database (ESMAELISHAHMIRZADI; MORTEZAPOUR, 2018; MADANI et al., 2018).

Diversos trabalhos apresentaram melhora em seus resultados ao utilizarem a GAN para gerar imagens sintéticas visando o aumento de base. Por exemplo: Han et al. 2019 (Han et al., 2019) obtiveram 97.48% de sensibilidade na detecção de tumores cerebrais usando imagens sintéticas GAN em conjunto com imagens de aumento de base clássico, o resultado superou em 3.81% a sensibilidade quando os autores usaram apenas o aumento de base clássico. Já no contexto de classificação de nódulos pulmonares com imagens de CT, Onishi et al. 2019 (ONISHI et al., 2019) utilizaram imagens geradas por uma *Wasserstein* GAN (WGAN) no pré-treinamento de uma rede neural convolucional profunda. O pré-treinamento resultou em uma AUC de 0.841, um aumento de 35.2% quando comparado aos resultados sem o uso das imagens da WGAN no pré-treinamento (AUC 0.622). Onishi et al. 2020 (ONISHI et al., 2020) utilizaram imagens geradas por uma WGAN para o pré-treinamento de uma rede convolucional profunda multiplanar para classificação de nódulos pulmonares. Os autores obtiveram sensibilidade de 93.9% melhorando em 11% seus resultados anteriores. Zhao et al. (ZHAO et al., 2018a) utilizaram uma abordagem própria chamada de *Forward and Backward* GAN (F&BGAN) para aumento de base na classificação de nódulos pulmonares. O melhor resultado obtido (AUC 0.980) foi usando as imagens sintéticas geradas da abordagem proposta em conjunto com imagens geradas por transformações clássica de aumento de base.

Uma outra técnica para gerar imagens sintéticas é o uso dos princípios de transferência de aprendizagem na GAN. Modelos de geração que utilizaram os princípios de transferência de aprendizagem no gerador ou no discriminante e, até mesmo em ambos, são capazes de gerar imagens de maior qualidade usando menos imagens no treinamento (WANG et al., 2018). Fora do escopo das imagens médicas, Wang et al. 2018 (WANG et al., 2018) avaliaram o uso da transferência do conhecimento de redes pré-treinadas para um novo domínio. Segundo os autores, a utilização do conhecimento de redes pré-treinadas resulta em menor tempo de convergência e imagens com melhor qualidade. O algoritmo proposto obteve melhora de 36.77% no *Frechet Inception Distance* em relação a geração sem o uso do conhecimento das redes pré-treinadas quando usadas apenas 100 imagens do conjunto de dados LSUN no treinamento. Ghassemi et al. 2020 (GHASSEMI; SHOEIBI; ROUHANI, 2020) apresentaram um novo método para classificação de tumores cerebrais em imagens de Ressonância Magnéticas. A técnica consistiu

em transferir o conhecimento de um discriminador GAN já treinado como ponto partida dos pesos de um classificador de tumores. Os autores obtiveram uma acurácia média de 95.6%. Segundo os autores os seus resultados superam o estado-da-arte. Assim, como evidenciado em trabalhos como de Ghassemi et al. 2020 e Wang et al. 2018 é preciso verificar o impacto de estratégias que direcionem a geração de imagens GAN usando reutilização de conhecimento para diferentes tipos de imagens médicas.

No melhor do nosso conhecimento, apesar dos avanços propostos pelos trabalhos apresentados na classificação de nódulos pulmonares, os trabalhos não fazem uso de informações adicionais que melhorem a geração das imagens sintéticas como estratégia de aumento de base.

1.1 Objetivo

O objetivo principal desse trabalho foi avaliar a hipótese de que a transferência de conhecimento usando um pré-treinamento da GAN resulta em nódulos pulmonares sintéticos mais eficientes para aumento de base.

1.2 Estrutura do Trabalho

A estrutura da dissertação está da seguinte forma:

- **Capítulo 2 - Fundamentação Teórica:** Esse capítulo apresenta os principais conceitos utilizados nesse trabalho, como a base utilizada e algumas aplicações de GAN para geração de imagens médicas sintéticas;
- **Capítulo 3 - Materiais e Métodos:** Esse capítulo apresenta como foi desenvolvido o processo de criação de imagens sintéticas e a classificação dos nódulos pulmonares ;
- **Capítulo 4 - Resultados e Discussão:** Esse capítulo apresenta os resultados obtidos e a discussão em torno deles;
- **Capítulo 5 - Conclusão:** Esse capítulo finaliza esse trabalho apresentando as conclusões.

2 Fundamentação Teórica

2.1 Nódulos Pulmonares em Imagens de Tomografia Computadorizada

Dentro de imagens de CT, a definição para um nódulo pulmonar solitário é de uma opacidade parcialmente delimitada ou arredondada, cujo diâmetro menor mede 3 cm, sendo em geral, com densidade de parte cálcica ou moles (SILVA et al., 2010). A investigação de um nódulo pulmonar solitário por parte de uma especialista, possibilita a análise desse nódulo ser ou se tornar tumor maligno. Sendo necessária a rápida intervenção para evitar o óbito (ALBERTS et al., 2010).

A detecção e classificação dos nódulos é possível com a utilização das imagens de CT, em especial para nódulos pequenos de 1mm e 2mm. Assim, a aplicação de exames de rotina se tornou essencial para detectar a existência de câncer pulmonar (ERASMUS et al., 2010; DICOTTI; AL., 2010). A CT gera uma representação 3D do objeto alvo a partir de emissão de Raios-X em diferentes direções. As imagens geradas pela CT são em tons de cinza e uniformemente espaçadas, sendo essas imagens agrupadas em forma de pilha para simular um volume em 3D do objeto (GONZALEZ; WOODS, 2010).

Mantendo o espaçamento original do empilhamento das imagens de CT, esse empilhamento pode ser visto de forma que cada *pixel* possa representar o volume de um *voxel*. Esse possui valor constante no entorno de um malha de pontos central (OLIVEIRA, 2002). Um conjunto de *voxels* forma o que seria a representação digital de região cuboide, sendo esta denominada "espaço *voxel*". Nela cada *voxel* tem, em geral, um número inteiro associado que representa o tom de nível de cinza do *pixel* da imagem correspondente (Fig. 2.1).

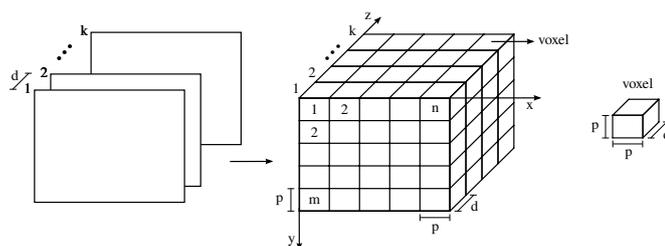


Figura 2.1: Espaço *voxel*. Fonte: Imagem extraída de (FALCÃO, 1993).

Os nódulos pulmonares, dentro das imagens de CT, podem ser observados sobre diferentes tonalidades de intensidades, sendo possível classificar os nódulos em três categorias: sólidos, semi-sólidos ou não sólidos. A Figura 2.2 demonstra exemplos de cada tipo de nódulo.

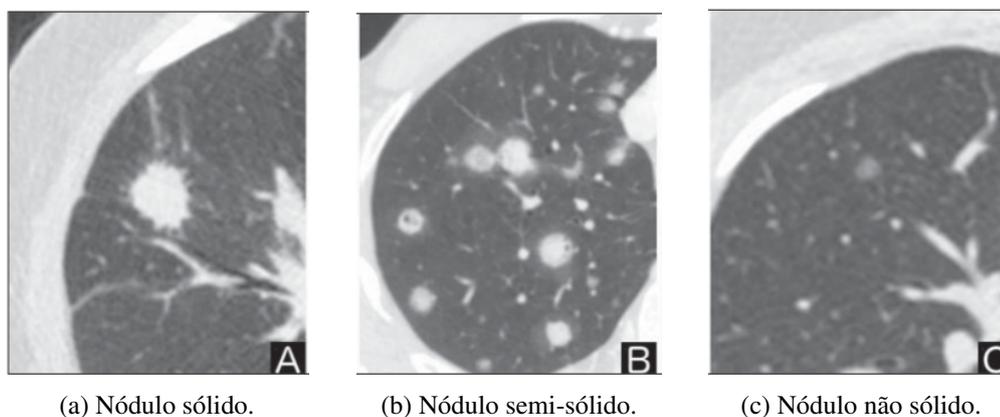


Figura 2.2: Exemplos das categorias de nódulos: sólido (a), semi-sólido (b) e não sólido. Fonte: Imagem extraída de (MOSMANN et al., 2016).

2.2 Diagnóstico Auxiliado por Computador (CAD)

Sistemas CAD tem como propósito dar auxílio aos especialistas na interpretação de imagens médicas, visando contribuir para aumento na acurácia e diminuição no tempo necessário para o diagnóstico. Esses sistemas trabalham como um agente de segunda opinião para os especialistas através do uso de métodos quantitativos que fornecem uma sugestão sem dar o diagnóstico final do exame (FERREIRA-JUNIOR; OLIVEIRA; AZEVEDO-MARQUES, 2018; CHUQUICUSMA et al., 2018; GILLIES; KINAHAN; HRICAK, 2016).

No contexto do câncer pulmonar, os sistemas CAD atuam no auxílio na classificação e na detecção de nódulos, ou na definição do estágio do câncer, mas sempre visando fornecer ao especialista um diagnóstico preliminar que possa ajudar no diagnóstico final. É importante destacar que o papel dos sistemas CAD é de auxílio. Logo devem funcionar apenas como uma ferramenta de melhora nas interpretações médicas através de processamento de informação computacional (AZEVEDO-MARQUES, 2001; DOI, 2007).

A partir das imagens médicas, os sistemas CAD são capazes de extrair informações para realizar análises quantitativas feitas por um computador. Tais informações são extraídas através de modelos matemáticos ou técnicas de processamento de imagens que produzem os atributos matemáticos referentes a imagem processada. O desempenho de um classificador de nódulo pulmonar CAD está diretamente ligado com a escolha dos atributos extraídos e da escolha da técnica para a realização dessa classificação (FILHO; MACHADO; OLIVEIRA, 2016).

2.3 Base LIDC-IDRI

Lung Image Database Consortium e *Image Database Resource Initiative* (LIDC-IDRI) é uma base de dados de CT pública disponível que é resultado da união entre o *Lung Image Database Consortium* (LIDC) e o *Image Database Resource Initiative* (IDRI) (ARMATO et al., 2011).

O LIDC-IDRI contém 1.018 exames e 244.527 imagens de CT torácicas no padrão DICOM, no qual as lesões foram classificadas e identificadas por quatro radiologistas. As imagens encontradas na base são de resolução espacial de 512×512 *pixel*, com profundidade de 16 bits e em níveis de cinza. Os exames da base foram gerados em diferentes instituições, logo as imagens geradas foram feitas em protocolos distintos em tomógrafos distintos. (ARMATO et al., 2011; JUNIOR; OLIVEIRA; AZEVEDO-MARQUES, 2016; SILVA; PAIVA; SILVA, 2016)

A classificação e identificação das lesões são separadas em três categorias: Nódulos ≥ 3 mm de diâmetro (Fig. 2.3a), nódulos < 3 mm de diâmetro (Fig. 2.3b) e não-nódulos ≥ 3 mm de diâmetro (Fig. 2.3c). As classificações realizadas pelos radiologistas foram baseadas na observação de nove características: textura, calcificação, lobulação, margem, sutileza, espiculação, estrutura interna, esfericidade e malignância. No final, 928 lesões foram classificadas como nódulos pelos quatro radiologistas, enquanto 7.731 foram classificados como nódulo por, pelo menos um radiologista, ambos os casos para nódulos com diâmetro ≥ 3 mm. Esses nódulos tiveram seus contornos demarcados por coordenadas cartesianas 2D (x, y) para cada um dos cortes (ARMATO et al., 2011).

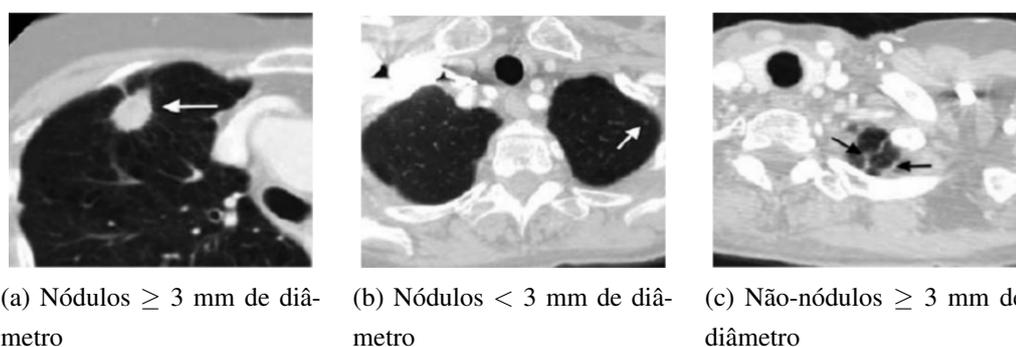


Figura 2.3: Lesões segundo as três classificações do LIDC-IDRI (ARMATO et al., 2011). Fonte: Imagem extraída de (ARMATO et al., 2011).

2.4 Aprendizagem Profunda

No período entre as décadas de 60 e 70, os primeiros sistemas que deram origem a Aprendizagem Profunda (do inglês *Deep Learning* - DL) apareceram. Entretanto foi apenas em 2006 que o termo "Deep Learning" foi criado (SCHMIDHUBER, 2015). A DL tem se tornado mais popular graças a três motivos: avanços nos algoritmos de DL, os custos dos equipamento e o au-

mento da capacidade em conjunto com o poder de processamento paralelo (SCHMIDHUBER, 2015; DENG, 2014).

A DL é um subconjunto de Aprendizagem de Máquina (AM). Essa por sua vez, é um subconjunto da Inteligência Artificial (Figura 2.4). DL consiste em um conjunto de métodos de Redes Neurais Artificiais (RNA) que utilizam uma representação de aprendizagem em múltiplos níveis e, conforme os níveis se aprofundam, a representatividade se torna mais abstrata (LECUN; BENGIO; HINTON, 2015). A capacidade da DL de trabalhar com imagens para reconhecimento, análise e processamento, permite a técnica obter resultados próximos aos encontrados por humanos (LITJENS et al., 2017a). Tantos métodos supervisionados quanto os métodos não supervisionados podem ser aplicados a DL (SCHMIDHUBER, 2015; ERHAN et al., 2010), sendo que os métodos não supervisionados oferecem vantagens, quando utilizados no campo da medicina, devido a capacidade de trabalho com bases de dados não rotulados (SHIN et al., 2016).

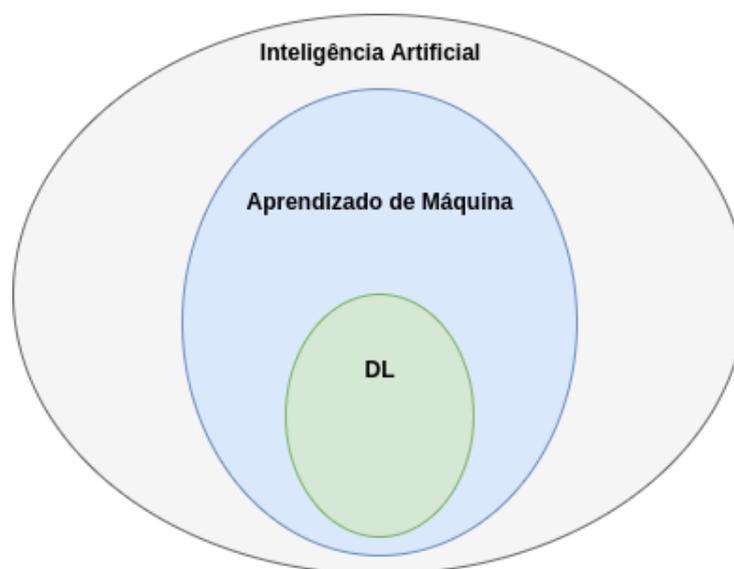


Figura 2.4: Hierarquia de Inteligência Artificial, Aprendizagem de Máquina e *Deep Learning* (DL). Fonte: elaborado pelo autor.

A DL é uma tipo de RNA. Essa é composta por nós que tem como objetivo simular o funcionamento e comportamento de neurônios humanos. Podemos afirmar que uma RNA é descrita por um grafo direcionado, composto por arestas e nós, sendo os nós a representação dos neurônios; e as arestas a representação das sinapses.

Nas RNAs, o treinamento dos nós pode ser visto em duas principais partes: a *feedforward* e a *backpropagation*. A parte do *feedforward* (Figura 2.5) corresponde a uma função Z apresentada na equação 2.1, com w correspondendo aos pesos e x a saída dos nós conectados ao nó "Z". O resultado dessa função pode ser somado a um viés, representado por b . A saída da função Z passa por uma Função de Ativação que determina se o nó vai ou não ser disparado. A parte de *backpropagation* visa otimizar os pesos para que a RNA possa aprender a mapear as entradas e saídas da rede. Essa parte utiliza o erro total encontrado após o *feedforward* ser executado

em todos os nós da RNA. Para isso é calculado o gradiente do erro total em relação a um peso (equação 2.2). O valor encontrado pode ser multiplicado por um valor n correspondente a uma constante de aprendizado e, em seguida, o valor é subtraído do peso original, resultando no novo valor para esse peso. Esse processo é repetido para todos os pesos na RNA (SONODA; MURATA, 2015; LITJENS et al., 2017b).

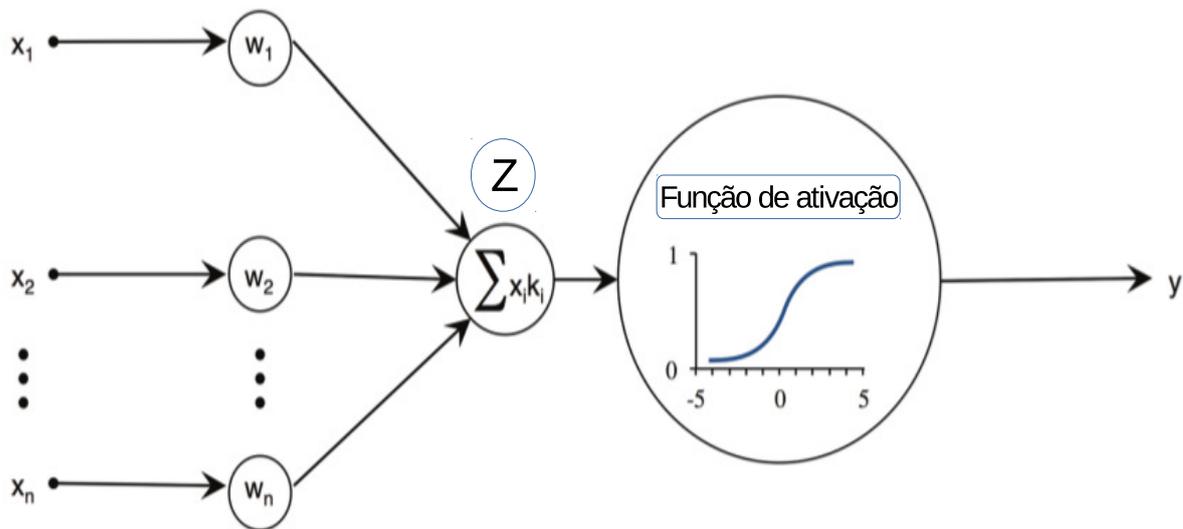


Figura 2.5: Funcionamento de um nó, sendo y o resultado calculado após a Função de ativação
Fonte: Imagem extraída de (CHARTRAND et al., 2017a).

$$Z = \sum_{i=1}^n x_i w_i + b \quad (2.1)$$

$$w_i = w_j - n * \frac{\partial \text{ErroTotal}}{\partial w_j} \quad (2.2)$$

A diferença central entre uma RNA de uma arquitetura DL está no número de camadas ou profundidade da rede que compõem cada arquitetura (SCHMIDHUBER, 2015). Normalmente, uma RNA é composta por apenas uma camada intermediária ou camada oculta, sendo essa camada aquela que fica entre as camadas de entrada e saída. Esse tipo de organização também conhecida como rede *shallow* por apresentar apenas uma camada oculta. Para a DL, temos um mudança mais radical nesse sentido, na qual podemos encontrar n camadas na parte oculta. Cada camada é responsável por detectar representações mais abstratas conforme mais distante da entrada de entrada. Por exemplo, as primeiras camadas detectam formas, arestas, bordas e as camadas mais profundas detectam textura (LECUNN; BENGIO; HINTON, 2015). A Figura 2.6 ilustra um exemplo da extração da características ao longo do aprofundamento das camadas, especificamente para um rosto.

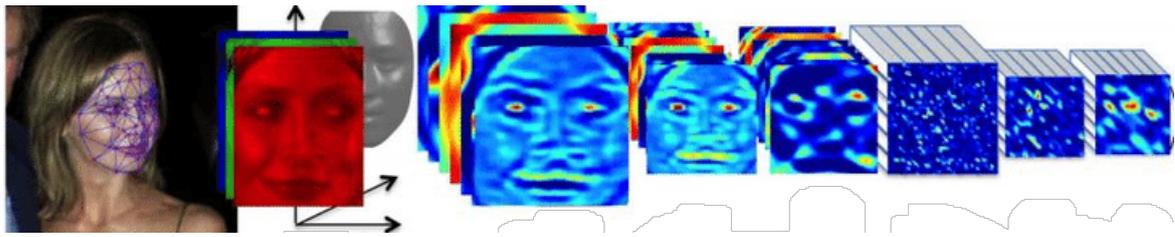


Figura 2.6: Exemplo de extração de características de uma rede profunda. Fonte: Imagem adaptada de (TAIGMAN et al., 2014).

2.5 Rede Geradoras Adversárias

A GAN (Fig. 2.7) é uma abordagem de DL inspirada na teoria dos jogos. Ela é composta por dois componentes principais: um modelo Gerador G , capaz de absorver a distribuição de dados e um modelo Discriminante D que estima a probabilidade de uma amostragem ter vindo do resultado do treinamento de G . O gerador G é treinado para produzir dados de saída, os quais não devem ser considerados distintos dos dados reais, através de um treinamento adversário com o discriminante D . Esse, por sua vez, é treinado para detectar os dados sintéticos criados por G (GOODFELLOW et al., 2014). Assim, o objetivo geral de minimizar a função de perda min-max L definida por:

$$\min_G \max_D L(G, D) = E_x[\log D(x)] + E_z[\log(1 - D(G(z)))] \quad (2.3)$$

em que x é a amostragem real dos dados e z é o vetor ruído de entrada. $D(x)$ representa a probabilidade da amostragem x ser real e $G(z)$ representa o mapeamento das características dos dados reais para a síntese dos dados.

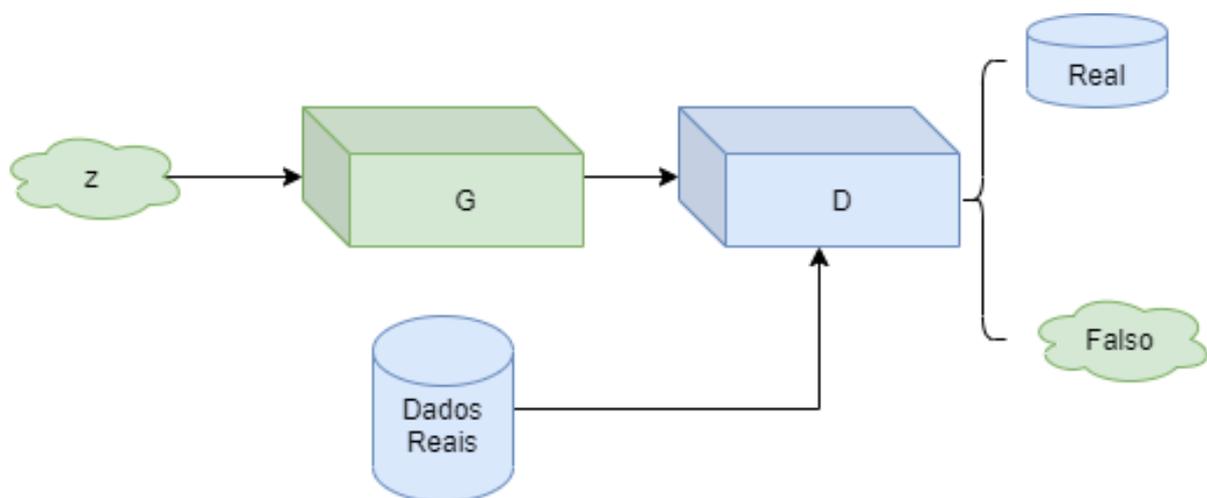


Figura 2.7: Esquema geral de uma GAN. Fonte: elaborado pelo autor.

A analogia diretamente associada ao conceito da GAN é o do falsificador de quadros e o especialista em arte. A cada novo quadro falsificado, o falsificador tenta vendê-lo. O especialista em arte, por sua vez, tenta identificar se o quadro é falso. Caso seja identificado como falso,

o falsificador aprende que sua obra não é fidedigna o suficiente, então aprimora sua técnica de pintura e tenta mais um vez vender o novo quadro. Caso o especialista em arte cometa erros, ele fica sujeito a grande prejuízo, logo ele também tem que melhorar sua capacidade de avaliar os quadros. Conforme os quadros falsos vão se tornando mais próximos dos reais, o especialista procura melhorar cada vez mais sua avaliação, dessa forma entrando em um ciclo.

2.5.1 Outras Arquiteturas GANs

A utilização de GAN para gerar imagens sintéticas aplicadas a diferentes finalidades motivou o desenvolvimento de diversos tipos de arquiteturas como, por exemplo, a *Deep convolutional GAN* (DCGAN) (RADFORD; METZ; CHINTALA, 2015), a *Wasserstein GAN* (WGAN) (Arjovsky; Chintala; Bottou, 2017) e *Conditional GAN* (cGan) (MIRZA; OSINDERO, 2014). Os esquema de algumas arquiteturas são ilustrados na Figura 2.8.

Deep Convolutional GAN (DCGAN). Proposta por Radford et al. (RADFORD; METZ; CHINTALA, 2015), a DCGAN foi a primeira arquitetura GAN na qual suas camadas são do tipo convolucionais tanto na parte da rede discriminante quanto na parte da rede geradora que tem seu real funcionamento voltado para geração de imagens. Esse fato se deu por três principais fatores: a remoção de camadas *fully* conectadas da camada oculta da rede; a substituição de qualquer *pooling layer* por *strided convolutions*, para o discriminante, e *fractional-strided convolutions*, para o gerador; e o uso de *Batch Normalization*, tanto no gerador como no discriminante, para estabilizar o treinamento. (RADFORD; METZ; CHINTALA, 2015). Seu esquema padrão de arquitetura pode ser visto na Figura 2.8b, na qual G_c representa uma rede geradora convolucional, D_c representa uma rede discriminante convolucional.

Wasserstein GAN (WGAN). Proposta por Martin et al. (Arjovsky; Chintala; Bottou, 2017), a WGAN é uma modificação da DCGAN, em que a métrica para a função de perda é definida pela distância de *Wasserstein*, e não pela divergência de *Jensen-Shannon*. Esta alteração é feita sobre o pretexto de diminuição do problema de diversidade identificado nas imagens geradas pela DCGAN.

Conditional GAN (cGAN). Proposta por Mirza e Osidero, a cGAN é uma modificação do modelo GAN, na qual é permitido condicionar o modelo por meio de informações externas adicionais, sendo o gerador e o discriminador alimentado por uma variável condicional. Essa arquitetura é demonstrada através do esquema presente na Figura 2.8c, na qual c são dados condicionais ou *labels* que direcionam o aprendizado.

Tub-Gan. Proposta por Zhao et al. (ZHAO et al., 2018b), a Tub-GAN é uma variação da cGAN e é, segundo os autores, a primeira a incorporar transferência de estilo dentro do *pipeline* de modelos GANs. A transferência de estilo atua sobre as imagens geradas inserindo as texturas das imagens reais sobre as imagens geradas antes de serem avaliadas como reais ou sintéticas pelo discriminante. A incorporação da transferência de estilo pode ser vista na figura 2.8d, na qual VGG representa a rede responsável pela transferência de estilo e L1 a perda para seu

treinamento.

Multi-channel Generative Adversarial Networks (M-GAN). Proposta por Bi et al. (BI et al., 2017), a M-Gan se inspira no modelo condicional da cGAN e aborda um modelo GAN para transformação de CT para Tomografia por Emissão de Pósitrons (do inglês Positron Emission Tomography - PET) no qual a entrada do gerador é composta pela imagem CT que vai ser transformada e um *label* a fim de melhorar o mapeamento para a síntese da imagem PET (BI et al., 2017). O esquema da arquitetura da M-GAN pode ser visto na 2.8e, no qual CT representa as imagens de tomografia computadorizada que deverão ser transformadas, *label* representa o que quer ser inserido, *Petg* representa a imagem PET sintética, *Petr* representa imagem PET real.

Forward Backward GAN (F&BGAN). Proposta por Zhao et al. (ZHAO et al., 2018a), a F&BGAN se apresenta como a solução para eventuais problemas de colapso do gradiente e borrões que possam aparecer nas imagens sintéticas proveniente da GAN. A F&BGAN se caracteriza por ser uma arquitetura em duas partes: a *Forward GAN* e a *Backward GAN*. A *Forward GAN* trabalha em cima da geração de imagens e extração das *features* das imagens. Essas imagens geradas são a entrada da *Backward GAN*. Senda esta composta por uma rede discriminante que tem um papel de *encoder* seguido de uma rede geradora atuando com um *decoder*, o que proporciona eliminações de eventuais ruídos nas imagens geradas. A Figura 2.8f demonstra o esquema da arquitetura da F&BGAN, no qual D' é uma rede *encoder* e G' é uma rede *decoder*.

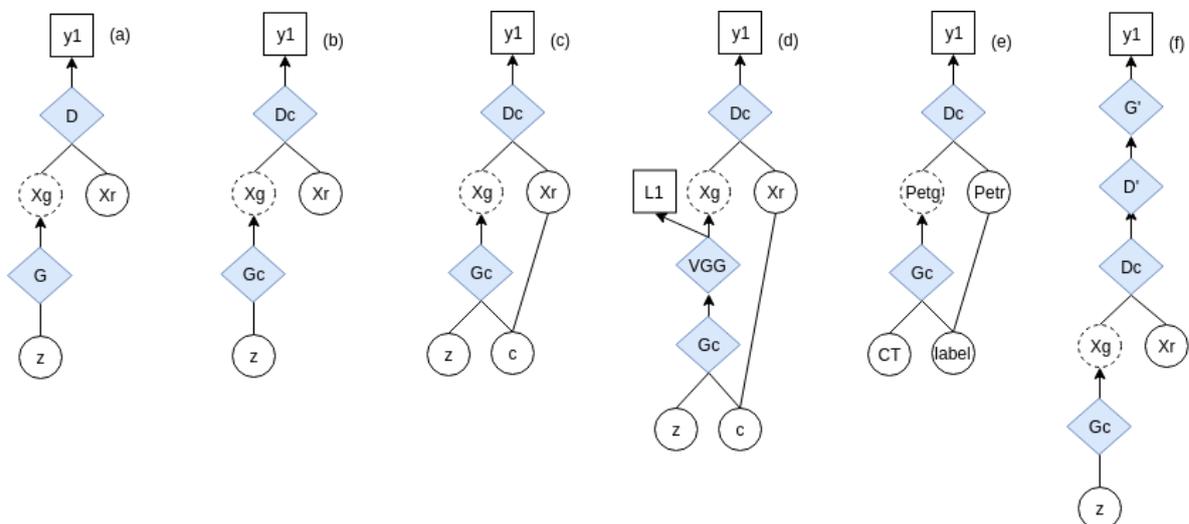


Figura 2.8: Esquema das variantes GAN. (a) representa a GAN original, na qual z é o vetor de ruído, G é a rede geradora, D é a rede discriminante, X_g é a amostra sintética gerada por G , X_r é a amostra real e y_1 é a classificação entre real e sintético. (b) representa a DCGAN e a WGAN. (c) representa a cGAN. (d) representa a Tubs-GAN. (e) representa a M-GAN. (f) representa a F&BGAN. Fonte: elaborado pelo autor.

2.6 Avaliação de desempenho

Um modelo computacional para classificação binária na área da saúde deve ser avaliado e, a partir desta avaliação será determinado o quão bom o modelo computacional é em discriminar uma determinada tarefa.

Em técnicas de AM e DL, um modelo de classificação prediz se amostras de um conjunto de teste são verdadeira ou falsa. É possível depreender dessa predição, quatro classificações a partir da matriz de confusão (Tabela 2.1):

		Valor Predito	
		Benigno (-)	Maligno (+)
Valor Verdadeiro	Benigno (-)	VN	FP
	Maligno (+)	FN	VP

Tabela 2.1: Matriz de Confusão.

- Verdadeiro Negativo (VN): a amostra foi classificada corretamente como benigna;
- Falso Positivo (FP): a amostra foi classificada erroneamente como maligna, mas era benigna;
- Falso Negativo (FN): a amostra foi classificada erroneamente como benigna, mas era maligna;
- Verdadeiro Positivo (VP): amostra foi classificada corretamente como maligna.

Utilizando a matriz de confusão, é possível calcular algumas métricas, destacando-se para esse trabalho: a sensibilidade (SEN) (Equação 2.4), especificidade (ESP) (Equação 2.5), acurácia (ACC) (Equação 2.6):

$$\text{Sensibilidade} = \frac{VP}{VP + FN} \quad (2.4)$$

$$\text{Especificidade} = \frac{VN}{VN + FP} \quad (2.5)$$

$$\text{Acurácia} = \frac{VP + VN}{VP + VN + FP + FN} \quad (2.6)$$

A sensibilidade tem como função avaliar a proporção de resultados VP em relação ao conjunto total de nódulos malignos; a especificidade tem como função avaliar a proporção de VN em relação ao conjunto total de nódulos benignos; a acurácia demonstra a proporção de nódulos classificados de forma correta em relação ao total de nódulos.

Uma outra maneira de avaliar um classificador binário é pelo uso da curva ROC (do inglês *Receiving Operating Characteristics*). Nela, é demonstrado como as amostragens positivas corretas (Taxa de Verdadeiros Positivos (TVP)) (Equação 2.7) variam em relação a amostragens negativas incorretas que foram classificadas como positivas (Taxa de Falsos Positivos (TFP)) (Equação 2.8) (LANTZ, 2015). Ou seja, o número de vezes que um a predição de um classificador foi correta contra as vezes que a predição foi incorreta.

$$TVP = \frac{VP}{TotaldeCasosPositivos} \quad (2.7)$$

$$TFP = \frac{FP}{TotaldeCasosNegativos} \quad (2.8)$$

Assim, conforme ocorre a predição do classificador, as coordenadas (TFP, TVP) para um gráfico bidimensional são geradas. Através desse gráfico bidimensional (Figura 2.9), com eixos das abscisas representado a *TFP*, e eixo da ordenadas representado a *TVP*, a curva ROC pode ser construída. Neste gráfico, a área sob a curva (AUC) será um valor no intervalo de 0 e 1. Esse valor é usado para comparar o desempenho de diferentes modelos de classificação, no qual quanto mais próximo o valor for de 1, melhor é considerado o desempenho do modelo. Observando a Figura 2.9 temos algumas características gerais do espaço ROC. Tomando como base a curva ROC azul, quanto maior a proximidade do ponto (1,0) pior será considerado o classificador. Em contrapartida, quanto maior a proximidade do ponto (0,1) mais perfeito o classificador será considerado (LANTZ, 2015).

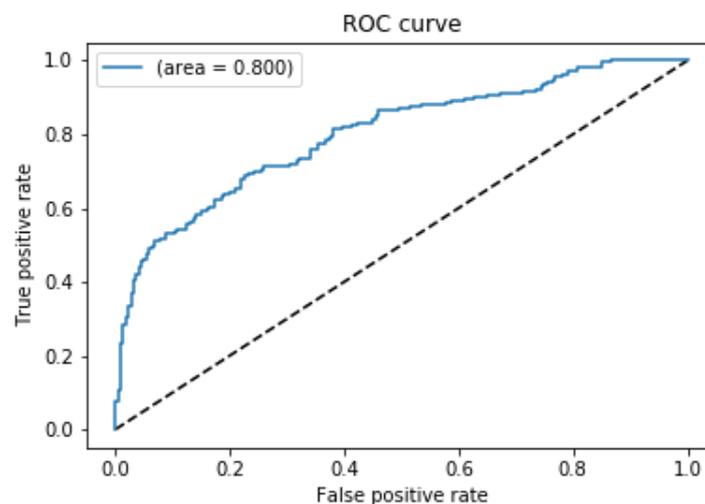


Figura 2.9: Exemplo de curva ROC. Fonte: elaborado pelo autor.

Dentre as forma de amostragens existentes, há a forma de validação cruzada *k-fold*, caracterizada por separar os dado em *k* subconjunto de dados menores que são mutualmente exclusivos, em que um desses subconjuntos é separado para teste e o restante são utilizados no treinamento e validação. Esse processo é repetido *k* vezes alternando os grupo que fazem parte do teste e

do treino e validação. Com isso, o classificador faz seu aprendizado a partir do conjunto de treinamento, utiliza os dados da validação para validar seus resultados buscando aprimorar os resultados para enfim usar o conjunto de teste para avaliação do classificador. A rotatividade dos conjuntos pelo uso do *k-fold* garante uma medida da capacidade de generalidade do modelo de classificação utilizado.

2.7 Aplicações GANs

Imagens médicas sintéticas podem ser utilizadas em diversas aplicações. A seguir será apresentado alguns trabalhos agrupados por arquitetura GAN utilizada. A Tabela 2.2 demonstra um resumo evidenciando as aplicações, arquiteturas e os tipos de imagens utilizadas nos trabalhos.

Tabela 2.2: Resumo dos Artigos Mencionados na sessão Trabalhos Relacionados 2.7.

Artigos	Ano	Arquitetura	Imagem	Aplicação
Salehinejad et al.	2018a	DCGAN	Raio-X	Classificação
Frid-adar et al.	2018	DCGAN	CT	Classificação
Salehinejad et al.	2018b	DCGAN	Raio-X	Classificação
Togo et al.	2018	DCGAN	Raio-X	Classificação
Chuquicusma et al.	2018	DCGAN	CT	Geração
Bermudez et al.	2018	DCGAN	MR	Geração
Onishi et al.	2018	WGAN	CT	Classificação
Han et al.	2018	DCGAN/WGAN	MR	Geração
Esmailishahmirzadi	2018	GAN	CT	Detecção
Loew	2018	GAN	Mamografia	Detecção
Moradi	2018	GAN	Raio-X	Classificação
Costa et al.	2018	cGAN	Retina	Segmentação
Senaras et al.	2018	cGAN	Histológica	Geração
Zhao et al.	2018a	F&BGAN	CT	Classificação
Chartsias et al.	2017	CycleGAN	CT	Segmentação
Zhao et al.	2018b	Tub-Gan	Retina	Segmentação
Bi et al.	2017	M-GAN	PET	Detecção

2.7.1 DCGAN

Salehinejad et al. 2018 (Salehinejad et al., 2018a) trabalha com a geração de imagens de Raio-X torácicas que foram treinadas independentemente para cinco casos de anomalia. As imagens sintéticas geradas por uma DCGAN foram utilizadas para balancear as bases de dados de uma Rede Neural Convolutacional Profunda (do inglês Deep Convolutional Neural Network - DCNN) que classificam as anomalias. Após testes sobre cinco conjunto de dados em seis diferentes DCNNs, foi obtido, no melhor caso, uma acurácia média 96,14% a partir do uso de

imagens sintéticas em conjunto com imagens reais, mas de forma desbalanceada; 89,72% com imagem reais balanceadas.

Frid-Adar et al. 2018 (FRID-ADAR et al., 2018) gerou imagens sintéticas de CT de lesões hepáticas de alta qualidade com intuito de melhorar o desempenho de uma Rede Neural Convolutiva (do inglês Convolutional Neural Network - CNN) com a função de classificar lesões hepáticas. Utilizando imagens geradas por uma DCGAN, em conjunto com o aumento base clássico, foi obtida uma acurácia de 85,7%, sendo superior aos 78,6% obtido apenas com o aumento de base clássico. Também foi feita uma avaliação das imagens geradas sob a perspectiva de dois especialistas, aos quais, foi perguntado se as imagens eram reais. O primeiro especialista obteve 62,5% de acurácia e o segundo especialista obteve 58,6% de acurácia.

Salehinejad et al. 2018 (Salehinejad et al., 2018b) treinou uma CNN profunda para classificação de patologia de tórax utilizando imagens sintéticas de Raio-X geradas por uma DCGAN como aumento de base. O treinamento foi feito em três conjuntos dados, sendo o primeiro apenas com imagens reais desbalanceado, o segundo com imagens sintéticas para balancear o primeiro conjunto e o terceiro com as imagens sintéticas sendo utilizadas para aumento de base. A acurácia média encontrada foi de 92,10% no terceiro conjunto, sendo superior aos 70,87% encontrado apenas com as imagens reais

Togo et al. 2018 (Togo et al., 2018) utilizou uma DCGAN modificada capaz de gerar *pachts* de imagens de Raio-X. Foram feitos dois experimentos visando a usabilidade de imagens sintéticas. O primeiro, o treinamento de uma Máquina de Vetores de Suporte (do inglês Support Vector Machine - SVM), foi feito com imagens reais e o teste foi feito com imagens sintéticas geradas pela DCGAN. São alcançados 0,960 (sensibilidade), 0,726 (especificidade), 0,827 (média harmônica). No segundo experimento, o treinamento da SVM foi feito apenas com imagens geradas e o conjunto de teste foi composto apenas por imagens reais, sendo alcançados 0,771 (sensibilidade), 0,846 (especificidade), 0,807 (média harmônica).

Chuquicusma et al. 2018 (CHUQUICUSMA et al., 2018) utilizou DCGAN para gerar imagens de nódulos pulmonares, assumindo que as imagens geradas são suficientemente realistas, a ponto de passarem despercebidas por especialistas. Através de um Teste de *Turing* Visual, o primeiro radiologista foi enganado em 67% dos casos e o segundo em 100%.

Bermudez et al. 2018 (BERMUDEZ et al., 2018) utilizou uma DCGAN para gerar imagens de Ressonância Magnética (do inglês Magnetic Resonance - MR) em alta resolução. Essas imagens passam por um processo de *denoised* através de uma rede *autoencoder*. A avaliação foi feita através de um teste *Turing* visual no qual um neurorradiologista e um especialista em imagens deram pontuações de 0 a 5 em imagens reais e sintéticas mostradas aleatoriamente. As imagens sintéticas avaliadas pelo neurorradiologista apresentaram pontuação 2 em aproximadamente 60% dos casos; já, a avaliação do especialista em imagem apresentou pontuação 4 em quase 90% dos casos.

2.7.2 WGAN

Onishi et al. 2019 (ONISHI et al., 2019) utilizou uma WGAN para gerar imagens sintéticas de nódulos pulmonares usadas para pré-treinamento de uma DCNN para classificar nódulos pulmonares, sendo o ajuste fino feito com a utilização das imagens originais dos nódulos. Quando o número de imagens geradas pela WGAN foi de 60000, a acurácia na classificação de nódulos benignos foi de 66,7% e de nódulos malignos foi de 93,9% no seu melhor caso. A AUC do método proposto alcançou 0,841, sendo superior aos 0,622 alcançados pelo método sem pré-treinamento.

Han et al. 2018 (Han et al., 2018) demonstrou que é possível obter imagens artificiais realísticas de MR de cérebro utilizando abordagem baseadas em GAN. Uma DCGAN e uma WGAN foram treinadas sobre seis conjuntos de imagens de MR. As imagens geradas após o treinamento foram avaliadas por um médico especialista através de um teste *Turing Visual*. O conjunto de imagens da base Concat de 64x64 *pixels* gerado pela WGAN foi aquele que o especialista cometeu mais erros, com um total de 15 de 50 imagens.

2.7.3 GAN Clássica

Esmailshahmirzadi e Mortezapour 2018 (ESMAELISHAHMIRZADI; MORTEZAPOUR, 2018) fizeram uso de *patches* de imagens de CT geradas através de uma GAN para equilibrar uma base de nódulos e não-nódulos pulmonares. As novas imagens foram usadas para melhorar a performance de uma CNN que é focada em detecção de nódulos pulmonares. Foi encontrado um total de 95,13% de acurácia quando utilizado aumento de base baseado em imagens sintéticas GANs, sendo esse valor superior aos 87,31% utilizando aumento de base clássico.

Guan e Loew 2018 (GUAN; LOEW, 2019) encontraram resultados divergentes da maioria da literatura, ao utilizar GAN para gerar imagens sintéticas de mamografia utilizada como aumento de base para uma CNN de detecção de câncer de mama. Os resultados após o treinamento são que, com as imagens originais, foram alcançados 88,85% de acurácia. No melhor caso, com uso apenas de imagens reais e aumento de base clássico foi encontrado 89,62% de acurácia; a utilização de aumento de base por imagens GAN em conjunto com as imagens reais, foi de apenas 87,31%, sendo esta o menor resultado entre as três.

Moradi et al. 2018 (MADANI et al., 2018) utilizou imagens sintéticas de Raio-X torácico geradas através de uma GAN para aumentar o conjunto de dados de uma CNN classificadora de anomalias cardiovasculares. O treinamento da CNN foi realizado em três cenários, sendo o primeiro cenário composto de imagens reais, o segundo cenário no qual imagens reais com aumento de base clássico e o terceiro com reais e aumento de base com GAN. A acurácia obtida na classificação do terceiro cenário foi o melhor resultado com 84,19%.

2.7.4 cGAN

Costa et al. 2018 (Costa et al., 2018) transformou imagens de vasos de retina em sua associada imagem de fundo de olho através da utilização da uma cGAN em conjunto de uma *Adversarial Autoencode*. Utilizando a métrica *Image Structure Clusterin* sobre as imagens dos vasos da retina em relação as imagens de fundo de olho respectivamente associada, foram obtidos um média de 0.9832 para imagens reais e 0.9671 em imagens sintéticas.

Senaras et al. 2018 (SENARAS et al., 2018) desenvolveu uma abordagem de criação de lâminas histopatológicas digitais sintéticas de câncer de mama através da geração proporcionada de um cGAN. Para validar a imagens sintéticas, foi proposta uma avaliação perante seis profissionais, compostos de três patologista e três analistas de imagens. Eles foram apresentados a três conjuntos de imagens, um contendo dez imagens sintéticas, um segundo contendo 10 reais, e, um terceiro contendo 5 reais e 5 sintéticas. A acurácia obtida pelo primeiro patologista foi de 33,3%, a do segundo foi de 60,0% e a do terceiro, 50,0%. Já para os analistas de imagens temos que o primeiro teve uma acurácia de 46,7%, o segundo, 46,7% e o terceiro, 53,3%.

2.7.5 Outras GANs

Zhao et al. 2018 (ZHAO et al., 2018a) apresentou a F&BGAN que foi utilizada para gerar imagens de CT de nódulos pulmonares malignos e benignos demonstrando a capacidade das imagens sintéticas na melhora da performance de uma classificador M-VGG16. Para isso, foram gerados sete conjuntos de dados misturando técnicas de aumento de base distintas, incluindo a F&BGAN. Cada base e algumas combinações entre si foram utilizadas no treinamento para um classificador de nódulos M-VGG16. O melhor resultado obtido foi a junção das imagens originais com aumento de base clássico e com as imagens geradas pela F&BGAN, obtendo um total de 95,24% de acurácia e uma AUC de 0,980. Segundo os autores, esse resultado é superior aos anteriores encontrados no estado-da-arte até então.

Chartsias et al. 2017 (CHARTSIAS et al., 2017) utilizou da CycleGAN para gerar imagens de MR cardíacas a partir de imagens CT cardíacas. As imagens sintéticas geradas são utilizadas para aumento de base para melhorar a performasse de um modelo de segmentação de arquitetura U-Net. A combinação entre imagens sintéticas e imagens com aumento de base clássico, obteve o melhor resultado com 0,712 de *dice score* (DS) médio, sendo superior aos 0,613 de DS encontrados utilizando apenas dados reais.

Zhao et al. 2018 (ZHAO et al., 2018b) apresentou a Tub-GAN e sua variante Tub-sGAN, uma nova abordagem para síntese de imagens retinianas e neuronais, ambas com aparência realista, e com capacidade de melhorar o desempenho em módulos de segmentação. As imagens geradas foram de cinco base de dados: DRIVE, STARE, HRF, NeuB1 e Kaggle. Utilizando apenas as imagens sintéticas criadas do base DRIVE, foram obtidos 0,8924 e 0,8980 de *structural similarity* para a Tub-GAN e Tub-sGAN, respectivamente. Foi utilizado o método de segmentação CNN *pacth-based* treinado com três conjuntos de dados diferentes. O primeiro

conjunto composto de imagens reais, o segundo apenas por imagens sintéticas e o terceiro pela junção dos conjuntos anteriores. O conjunto misto obteve melhor resultado para todas as bases de imagens sobre a métrica *F1-score*, obtendo 80,33 para imagens do DRIVE, 79,02 para imagens do STARE, 79,50 para imagens do HRF e 85,06 para imagens do NeuB1.

Bi et al. 2017 (BI et al., 2017) propôs um método de produzir imagens de PET sintéticas usando M-GAN. As imagens geradas foram utilizadas para o treinamento de uma CNN totalmente conectada na detecção de tumores pulmonares. Os resultados utilizando apenas as imagens da abordagem proposta alcançou um total de 63,84 *F-score*, sendo menor do que quando utilizadas apenas imagens reais, que alcançou 66,38.

3 Materiais e Métodos

3.0.1 PREDCGAN

Nesse trabalho, fizemos transferência de conhecimento em um modelo arquitetural DCGAN através um pré-treinamento, que iremos denominar de PREDCGAN. Na PREDCGAN (Figura 3.1) assumimos que o aprendizado para a geração das imagens de nódulos segmentados pode ser utilizado para proporcionar uma melhora na capacidade de geração de imagens sintéticas de nódulos com suas estruturas adjacentes. Para esse funcionamento ocorrer, nossa proposta implementa um pré-treinamento em uma DCGAN. Ou seja, primeiro temos o treinamento da DCGAN para geração de imagens de nódulos segmentados e, após o fim deste treinamento, um novo treinamento é feito, com o intuito de gerar nódulos com suas estruturas adjacentes. Dessa forma, o conhecimento para geração de nódulos segmentados é aproveitado como ponto de partida para geração dos nódulos com as estruturas adjacentes.

A Figura 3.1a representa o pré-treinamento, no qual a DCGAN é treinada para gerar as imagens dos nódulos segmentados (domínio de origem). Ao término do pré-treinamento, os pesos da DCGAN estão ajustados para a geração de nódulos segmentados. Na Figura 3.1b, temos a representação do segundo momento, no qual ocorre o treinamento para geração da imagens dos nódulos com estruturas adjacentes (domínio alvo). Nessa etapa, os pesos começam ajustados para geração do nódulo segmentado. Dessa forma, nas primeiras épocas, temos a geração apenas do nódulos segmentados e, com o decorrer dessas épocas, o aprendizado das estruturas adjacentes.

Do ponto de vista de implementação, o treinamento é dividido em dois momentos na mesma rede DCGAN, não sendo necessário carregar os pesos em uma segunda rede. Logo, não existe alteração na resolução da imagem e, como pode ser observado na Figura 3.1b, a geração das imagens com estruturas adjacentes tem como foco preencher o fundo preto dos nódulos segmentados aprendidos anteriormente.

Esse formato de treinamento levou a criação de duas funções a serem minimizadas em diferentes momentos. A primeira representada pela função de perda min-max L_s :

$$\min_G \max_D L_s(G, D) = E_x[\log D(xs)] + E_v[\log(1 - D(G(v)))] \quad (3.1)$$

em que v é o vetor ruído, xs é a amostragem real de nódulos segmentados. $D(xs)$ representa

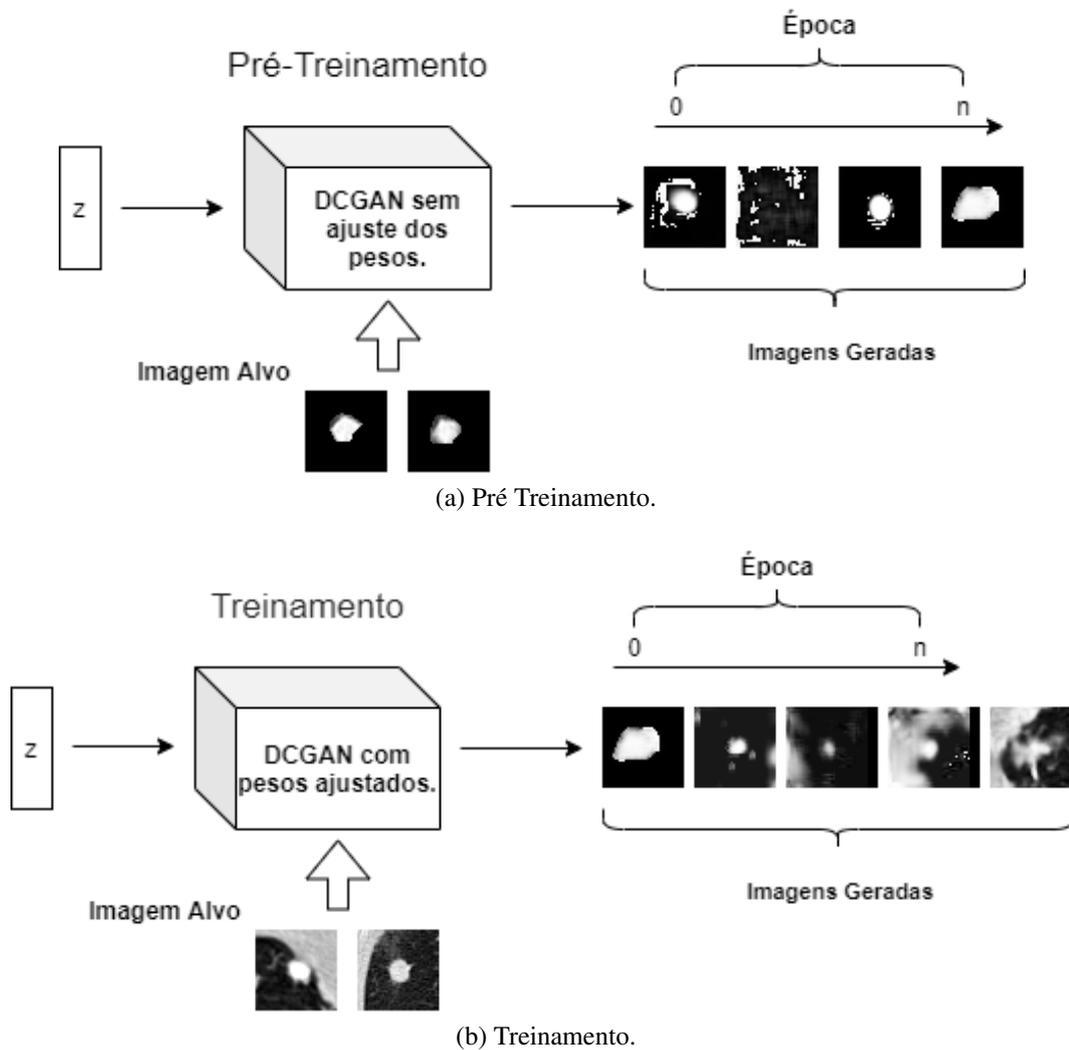


Figura 3.1: Esquema geral da PREDCGAN. Fonte: Elaborada pelo autor.

a probabilidade da amostragem xs ser real e $G(v)$ representa o mapeamento das características dos dados reais (nódulos segmentados) para a síntese dos dados. Essa função representa o aprendizado para geração de nódulos segmentados e também para discriminação desses nódulos. Assim, ela representa o processo de pré treinamento para a próxima etapa.

Com utilização dos pesos encontrados no treinamento, a partir da função 3.1, é definida uma segunda função que contém similaridades com a anterior pelo fato de reutilizar o gerador e o discriminador anteriormente treinados. Assim, mudando o alvo, temos essa função de perda min-max L :

$$\min_G \max_D L(G, D) = E_x[\log D(x)] + E_v[\log(1 - D(G(v)))] \quad (3.2)$$

em que v é o vetor ruído, x é a amostragem real de nódulos. $D(xs)$ representa a probabilidade da amostragem x ser real e $G(v)$ representa o mapeamento das características dos dados reais para a síntese dos dados.

Os detalhes dos hiperparâmetros e das estruturas da arquitetura DCGAN utilizadas no funcionamento da PREDCGAN estão na secção 3.0.2.

3.0.2 Experimento

Nosso experimento teve como objetivo demonstrar que as imagens sintéticas da PREDCGAN são mais eficientes para o uso em aumento de base quando comparadas com arquiteturas GANs clássicas e também quando comparadas ao aumento de base clássico. O ambiente para realizar os experimentos foi desenvolvido utilizando a biblioteca *Tensorflow v2.0*, sendo o processamento executado em uma GPU GeForce GTX 1060. O trabalho foi dividido em duas etapas (Fig. 3.2).

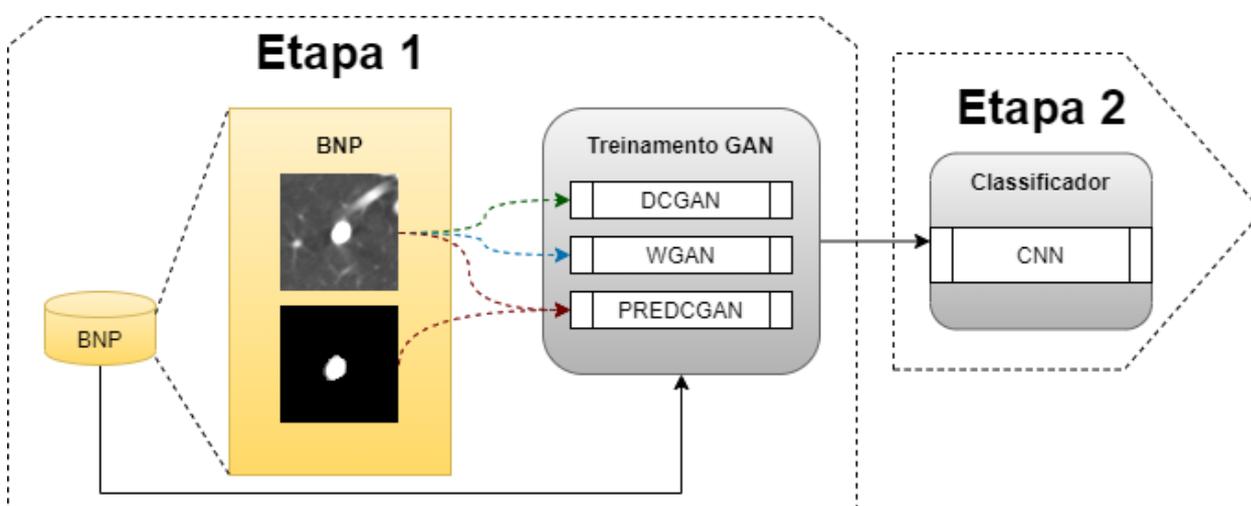


Figura 3.2: Fluxograma das etapas deste trabalho. Fonte: Elaborada pelo autor

Etapa 1: consistiu em gerar as imagens sintéticas de nódulos pulmonares usando três GANs: Deep Convolutional GAN (DCGAN), Wasserstein GAN (WGAN) e PREDCGAN, usando imagens do Banco de Nódulos Pulmonares (secção 3.0.2). Os nódulos segmentados são utilizados apenas pela PREDCGAN, e os nódulos com estruturas adjacentes são utilizados pelas três GANs. Para cada GAN foram treinadas duas redes, uma para gerar os nódulos benignos sintéticos e outra para os nódulos malignos sintéticos.

Etapa 2: consistiu na etapa de validação dos resultados, na qual uma arquitetura CNN (secção 3.0.2) é treinada para classificar nódulos pulmonares utilizando as imagens geradas pelas GANs como aumento de base em diferentes conjuntos de treinamento. As imagens sintéticas geradas pelas GANs foram usadas como técnica de aumento de base para evitar possíveis enviesamentos nos resultados da classificação, as imagens usadas no treinamento das GANs são também as utilizadas no treinamento da CNN. Não há fatias do mesmo nódulo que possam estar tanto no treino quanto no teste, por exemplo. Também foi realizado um Teste de Turing Visual com um profissional da radiologia apenas das imagens de PREDCGAN(secção3.0.2).

Para avaliar a significância estatística da diferença do desempenho apresentada pelos conjuntos de treinamento, optamos por utilizar o nível do Intervalo de Confiança (do inglês confidence interval - CI) em 95% e por utilizar o teste de Mann-Whitney Wilcoxon Rank-Sum. Nesse contexto, a existência de um valor de $p < 0.05$ na comparação entre duas populações indica que essa diferença tem significância estatística.

Banco de Nódulos Pulmonares (BNP)

A base utilizada para o treinamento da CNN e das GANs é o BNP criado por Ferreira Junior et al. 2016 (JUNIOR; OLIVEIRA; AZEVEDO-MARQUES, 2016), sendo esse um recorte do LIDC-IDRI (ARMATO et al.,). O BNP contém imagens de nódulos e seus respectivos nódulos segmentados manualmente pelos radiologistas do LIDC-IDRI. As imagens DICOM de CT de tórax foram definidas em *level* de -600 e nível de janela de 1,600. A Figura 3.3 demonstra a Região de Interesse (ROI) dos nódulos e as suas respectivas segmentações.

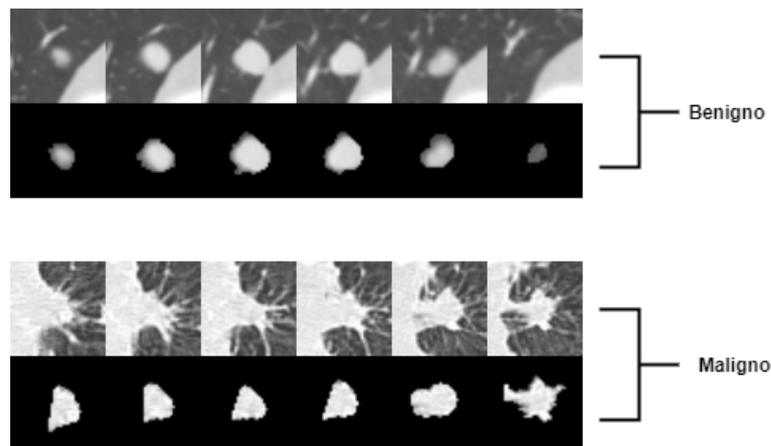


Figura 3.3: Na primeira linha temos ROI dos nódulos não segmentados e na segunda linha a respectiva segmentação manual. Fonte: Elaborada pelo autor.

O BNP possui 698 nódulos benignos provendo um total de 2,517 fatias rotuladas como benignas e 307 nódulos malignos provendo um total de 1,994 fatias rotuladas como malignas. Cada fatia foi reduzida ao tamanho 40×40 *pixels*, para diminuir a quantidade de estruturas adjacentes ao nódulos sem eventualmente corta-lo. Cada nódulo está centralizado na imagem e o tamanho do nódulo varia entre 3mm a 30mm.

Comparação de Desempenho

Na PREDCGAN, aplicamos um pré-treinamento sobre a arquitetura DCGAN buscando um melhora nas imagens. Dessa forma optamos por comparar nossa proposta com arquiteturas GAN diretamente relacionados a DCGAN. Assim comparamos nossa abordagem com os seguintes métodos do estado-da-arte: DCGAN e WGAN.

A Figura 3.4 apresenta a organização das camadas da rede geradora e da rede discriminante para todas as redes GAN. As arquiteturas das GANs treinadas são similares, tanto para os geradores quanto para os seus respectivos discriminantes. As alterações são feitas dentro de suas peculiaridades. Nos geradores das GANs, todas as camadas, após o *reshape* do vetor ruído, são de convolução transposta, contendo filtros de *kernel* de tamanho 5 e ativação *LeakyRelu*. Em todas as camadas foi utilizado o *batch normalization* com *momentum* de 0.8, com exceção da última camada. Já para os discriminantes, todas as camadas são de convolução contendo filtros de *kernel* de tamanho 5 e ativação *LeakyRelu*. Com exceção da camada de *flatten*, em todas as outras camadas é utilizado o *batch normalization* com *momentum* de 0.8 e também é utilizado *dropout* de 0.5.

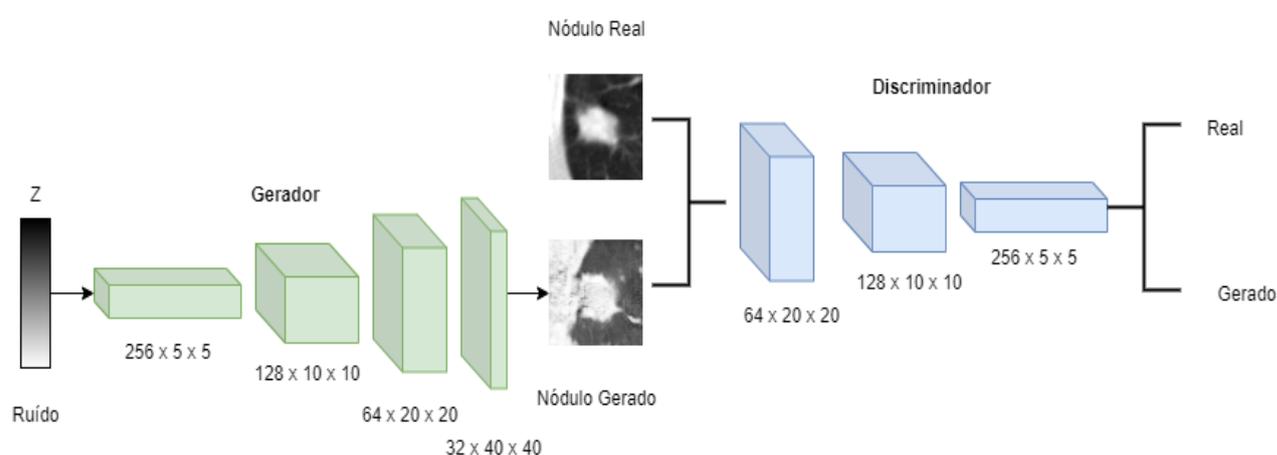


Figura 3.4: Arquitetura das GANs. Fonte: Elaborada pelo autor.

A WGAN utiliza em seu discriminante e seu gerador o *loss* baseado na distância de *wasserstein* e o otimizador utilizado é RMSprop, com um *learning rate* de 0.00005. Para a DCGAN, a perda é sobre a *binary cross entropy* e o otimizador utilizado é Adam com um *learning rate* de 0.0001 para o discriminante e 0.001 para o gerador. Ambas, DCGAN e WGAN foram treinadas durante 4000 épocas. A PREDCGAN utilizou perda *binary cross entropy* e o otimizador utilizado é Adam, com um *learning rate* de 0.0001 para o discriminante e 0.001 para o gerador. O treinamento foi de 2000 épocas para o pré-treinamento com os nódulos segmentados e 2000 épocas para o treinamento final da ROI do nódulo.

Classificador

Com o treinamento das GANs da etapa 1, criamos os conjuntos de treinamento que foram avaliados no treino da CNN:

1. Composto por imagens reais com aumento clássico;
2. Composto de imagens reais e sintéticas da PREDCGAN;

3. Composto por imagens reais com aumento de base clássico e imagens sintéticas PREDCGAN;
4. Composto por imagens reais e sintéticas DCGAN;
5. Composto por imagens reais com aumento de base clássico e imagens sintéticas DCGAN;
6. Composto por imagens reais e sintéticas WGAN;
7. Composto por imagens reais com aumento de base clássico e imagens sintéticas WGAN.

O processo de classificação foi realizado nos diferentes conjuntos de treinamento utilizando uma validação cruzada *k-fold* ($k=6$) e *holdout*. A Tabela 3.1 apresenta os valores para cada ciclo do *k-fold*, em que Real representa as imagens reais utilizadas no treinamento das GANs (etapa 1) e da CNN (etapa 2), Aumento representa a quantidade do aumento de base das fatias dos nódulos para a etapa 2 com base nas imagens do Real. Validação representa as imagens que foram utilizadas pra a validação do treinamento (etapa 2); Teste representa as imagens utilizadas nos testes de cada iteração do *k-fold* (etapa 2). Para o caso dos conjuntos de treinamento 3, 5 e 7, as imagens do aumento de base é dividida pela metade para cada técnica. Por exemplo, no conjunto de treinamento 3 com $k=0$, são 4428 fatias benignas geradas pelos PREDCGAN e 4428 fatias gerados pelo aumento de base clássico.

Para a validação *holdout*, foram utilizados 2517 fatias de nódulos benignos e 1994 fatias nódulos malignos para o treino das GANs (etapa 1) e da CNN (etapa 2), 220 fatias benignas e 220 malignas para o teste (etapa 2) e 220 fatias benignas e 220 malignas para a validação (etapa 2). O aumento de base na validação *holdout* foi de 31848 imagens de fatias benignas e 31649 imagens de fatias malignas, utilizando como base as imagens do treino. Para o caso dos conjuntos de treinamento 3,5 e 7 a quantidade das imagens do aumento de base é dividido pela metade para cada técnica, por exemplo, no conjunto de treinamento 3 são 15924 fatias benignas gerados pelos PREDCGAN e 15924 fatias geradas pelo aumento de base clássico.

Para o treinamento da CNN, além do aumento de base, são utilizadas as imagens reais separadas para o treinamento do respectivo *fold*, assim como no *holdout*. Para o tipo de aumento de base clássico, foi usado um zoom range de 0.15 nas imagens, uma rotação de 10 graus positiva com interpolação do tipo *nearest neighbor*.

A arquitetura do classificador desenvolvida pode ser vista na Figura 3.5. Ela é composta por três camadas de convolução de *kernel size* de tamanho 3 com ativação *LeakRelu* seguida de uma camada *max pooling* de *kernel size* de tamanho 2. Após a ultima *max pooling* temos uma camada *flatten*, esse seguida de duas camadas totalmente conectadas de tamanho 2048, finalizando com uma camada de saída com ativação *sigmoid*.

Tabela 3.1: Quantidade de imagens por k-fold

Tipo	Fold	Real	Validação	Teste	Aumento
Benigno	k = 0	2214	449	447	8856
Benigno	k = 1	2214	447	447	8856
Benigno	k = 2	2147	447	447	8588
Benigno	k = 3	2137	447	447	8548
Benigno	k = 4	2119	447	447	8476
Benigno	k = 5	2035	447	447	8140
Maligno	k = 0	1815	307	307	7260
Maligno	k = 1	1815	356	307	7260
Maligno	k = 2	1757	356	307	7028
Maligno	k = 3	1643	356	307	6572
Maligno	k = 4	1681	356	307	6724
Maligno	k = 5	1592	356	307	6368

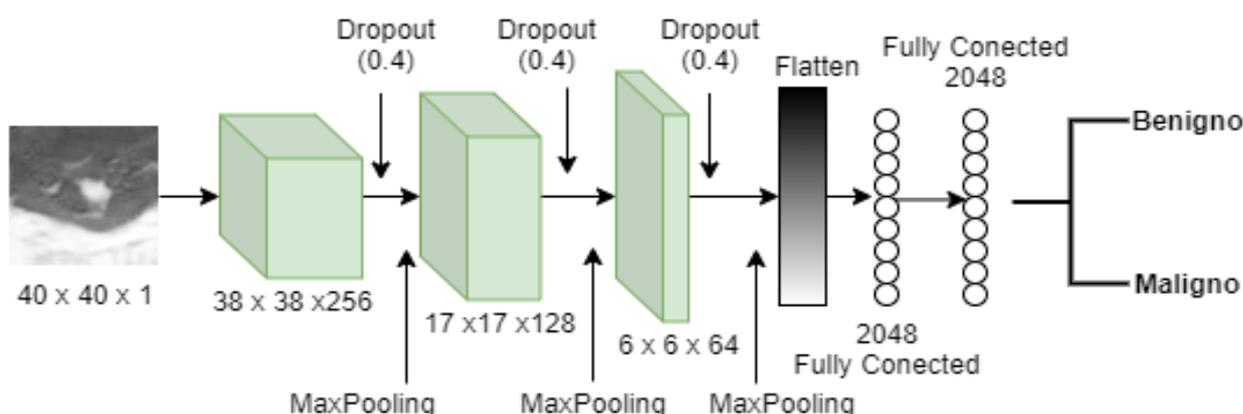


Figura 3.5: Arquitetura da CNN. Fonte: Elaborada pelo autor.

Teste de Turing Visual

Para observar a similaridade com imagens reais das imagens sintéticas geradas pela PREDCGAN, foi realizado um teste de Turing visual com um radiologista. Nesse teste, foram apresentadas ao radiologista fatias de nódulos sintéticos e reais escolhidos aleatoriamente, nos quais dentro das limitações do formato do teste, o radiologista foi questionado se os nódulos eram reais e a malignidade em um intervalo de 1 a 5, no qual 1 representaria um tipo garantido benigno e 5 um tipo garantido maligno. Ao todo, foram apresentados 20 fatias escolhidas aleatoriamente, sendo 10 reais (5 benignos e 5 malignos) e 10 sintéticos (5 benignos e 5 malignos). Cada fatia de tamanho 40x40 *pixels* no formato de imagem PNG. Para as imagens reais foram usadas as fatias dos nódulos contidos na fatia principal. O teste foi executado com o uso de um formulário e não houve limite de tempo para execução do teste. Na Figura 3.6, são apresentadas as imagens utilizadas no teste e suas respectivas classes.

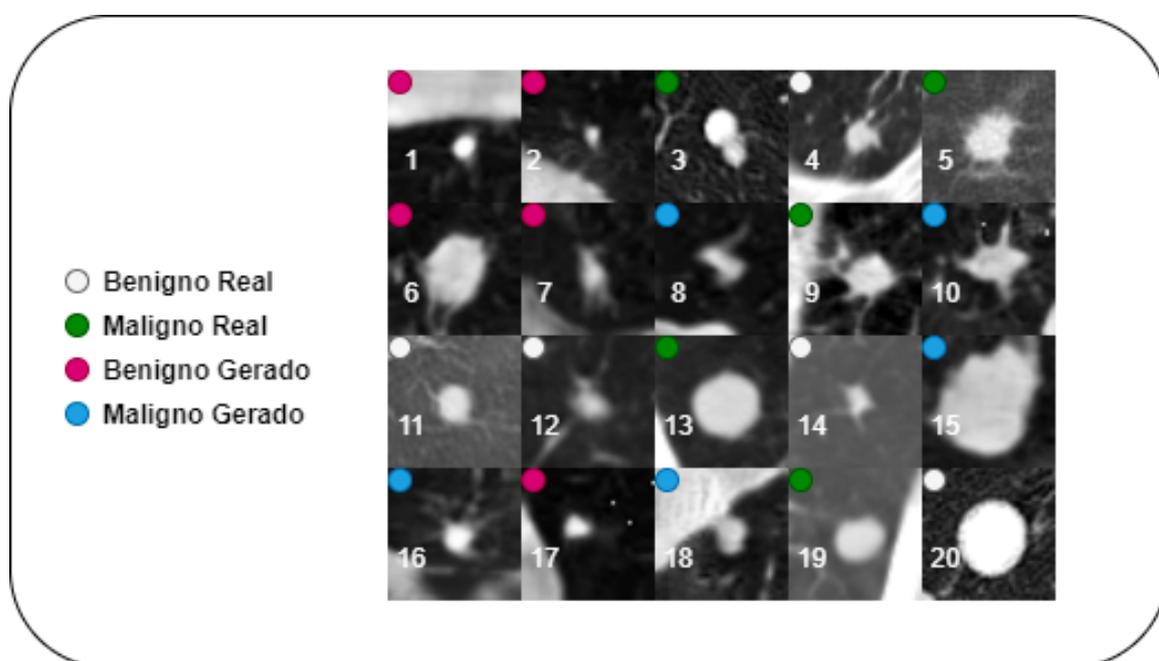


Figura 3.6: Imagens utilizadas no Teste de Turing Visual. Fonte: Elaborada pelo autor.

4 Resultados e Discussão

A Tabela 4.1 apresenta os resultados dos conjuntos de treinamento pela validação cruzada *k-fold*, observando as médias das métricas de sensibilidade, especificidade, acurácia e AUC-ROC, os valores encontrados do *p value* foram calculados apenas relacionando os resultados entre o conjunto de treinamento 3 (Classic + PREDCGAN) com os demais. Para as métricas sensibilidade, especificidade e AUC-ROC os valores de *p* obtidos foram de $p > 0.05$, o que implica que não existe diferença significativa entre o conjunto de treinamento 3 com os outros conjuntos de treinamento avaliados. Portanto, baseados nos resultados encontrado na classificação da CNN, esse resultado sugere que a PREDCGAN é pelo menos tão eficiente quanto os outros métodos avaliados no aumento de base de nódulos pulmonares quando observadas essas métricas.

Tabela 4.1: Média dos resultados *k-folds* para os Conjuntos de Treinamento individuais, sensibilidade (SEN), especificidade (ESP), acurácia (ACC) e AUC-ROC (AUC).

Conjunto de Treinamento	SEN	ESP	ACC	AUC
Classic (1)	0.736 ± 0.054	0.780 ± 0.070	0.747 ± 0.038	0.824 ± 0.030
PREDCGAN (2)	0.750 ± 0.083	0.700 ± 0.095	0.691 ± 0.024	0.808 ± 0.030
Classic + PREDCGAN (3)	0.760 ± 0.036	0.720 ± 0.064	0.728 ± 0.029	0.816 ± 0.023
DCGAN (4)	0.750 ± 0.121	0.750 ± 0.144	0.685 ± 0.093	0.806 ± 0.02
Classic + DCGAN (5)	0.750 ± 0.065	0.750 ± 0.092	0.730 ± 0.027	0.821 ± 0.023
WGAN (6)	0.67 ± 0.129	0.76 ± 0.121	0.642 ± 0.067	0.801 ± 0.030
Classic + WGAN (7)	0.72 ± 0.090	0.770 ± 0.106	0.708 ± 0.027	0.818 ± 0.024

Ainda na Tabela 4.1, ao analisarmos os resultados da acurácia obtivemos $p < 0.05$ ao compararmos os conjuntos de treinamento 3 e 2 (PREDCGAN), apresentando uma melhora de 5%, e os conjuntos de treinamento 3 e 6 (WGAN), apresentando uma melhora de 8,5%. A melhora do conjunto de treinamento 3 em comparação 2 sugere a importância em usar o aumento de base clássico para uma melhor desempenho.

Nesse trabalho também avaliamos os resultados da classificação utilizando a validação *holdout* (tabela 4.2 e tabela 4.3). A tabela 4.2 apresenta os resultados da classificação dos nódulos quando utilizamos apenas as técnicas para o aumento de base individualmente. Ao analisarmos os resultados das técnicas para aumento de base, a PREDCGAN obteve o melhor resultado encontrado nesse trabalho com uma AUC-ROC de 0.781 e uma sensibilidade de 84.26% (CI: 76.00, 90.55). Entretanto, apesar do melhor resultado para as duas métricas citadas, a PREDCGAN apresenta o menor valor de especificidade com 61,14% (CI: 55.67, 66.42). Dentre as

técnicas, a DCGAN obteve a menor AUC-ROC (0.723) e menor sensibilidade (63.27%). Contudo, foi a técnica que obteve o maior valor de especificidade com 72.12% (CI: 64.62, 78.81).

Os resultados da utilização das diferentes GAN em conjunto com as imagens de aumento de base clássico, são apresentadas na Tabela 4.3. A utilização das técnicas conjuntas elevou os valores encontrados na AUC-ROC de todos os conjuntos de treinamento individuais. O conjunto de treinamento PREDCGAN e aumento de base clássico obteve a maior AUC-ROC (0.791) desse trabalho quando usada a validação *holdout*. Esse conjunto de treinamento ainda apresentou 75.00% (CI: 76.10, 81.08) de sensibilidade, sendo o melhor valor observando apenas os conjunto de treinamento composto com validação *holdout*. Além disso, esse conjunto de treinamento apresentou o melhor valor de sensibilidade 75.00% (CI: 76.10, 81.08) na validação *holdout*.

Tabela 4.2: Tabela 4.2 apresenta os resultados da classificação dos nódulos quando utilizamos apenas as técnicas para o aumento de base individualmente. Sensibilidade (Sen), especificidade (ESP) e AUC-ROC (AUC).

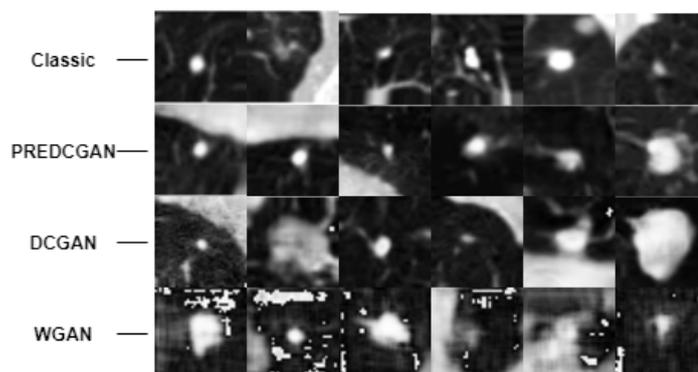
Conjuntos de Treinamento	Sen (%)	ESP (%)	AUC
Classic (1)	70.32	70.14	0.775
PREDCGAN (2)	84.26	61.14	0.781
DCGAN (4)	63.27	72.12	0.723
WGAN (6)	70.52	63.30	0.764

Tabela 4.3: Resultados do *holdout* para os Conjuntos de Treinamentos mistos, sensibilidade (Sen), especificidade (ESP) e AUC-ROC (AUC).

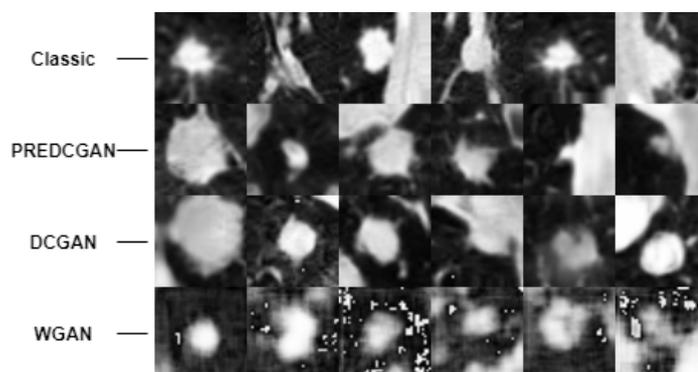
Conjunto de Treinamento	Sen (%)	ESP (%)	AUC
Classic (1)	70.32	70.14	0.775
Classic + PREDCGAN (3)	75.00	67.97 2	0.791
Classic + DCGAN (5)	70.62	66.26	0.771
Classic + WGAN (7)	70.73	68.09	0.781

A demonstração das imagens geradas por cada técnica de aumento de base pode ser visualizada nas figuras 4.1a e 4.1b. As imagens das fatias dos nódulos apresentadas nesse trabalho foram selecionadas aleatoriamente. Na Figura ?? temos exemplos dos resultados das imagens de nódulos benignos geradas por cada uma das técnicas utilizadas. Analogamente temos na Figura 4.1b exemplos dos resultados da imagens para nódulos malignos. É visível um problema na convergência das imagens geradas pela WGAN em ambos os casos, em que é encontrado um forte grau de borrão e ruído evidenciados pela cor branca. Para as imagens geradas pela PREDCGAN e a DCGAN, também são visíveis os ruídos em branco, entretanto, são mais discretos, demonstrando uma melhor convergência para as imagens alvo dentro do conjunto de imagens utilizadas.

Para uma validação qualitativa das imagens sintéticas geradas pela da PREDCGAN e demonstração da semelhança com imagens as reais, as imagens sintéticas foram submetidas a um



(a) Amostragem dos nódulos benignos



(b) Amostragem dos nódulos malignos.

Figura 4.1: Amostragem dos Nódulos

Teste de Turing Visual com um radiologista que foi questionado se as fatias apresentadas eram reais ou sintéticas e o grau de malignidade. A Matriz de Confusão do Teste de Turing Visual é encontrada na Tabela 4.4. É possível observar que o radiologista obteve uma acurácia média de 70%, sendo: uma acurácia de 80% para detecção de fatias reais e de 60% para fatias sintéticas. Observando a tabela 4.5, temos que, das imagens sintéticas classificadas como reais pelo radiologista, três delas foram nódulos avaliados com malignidade 3.

Tabela 4.4: Matriz de Confusão do Teste de Turing Visual

Radiologista	Real	Sintético
Real	8	2
Sintético	4	6

Ainda observando a Tabela 4.5, o radiologista submetido ao Teste de Turing Visual classificou grande parte das imagens sintéticas com um grau de malignidade 3, sendo esse grau um indicio de uma indeterminação. No conjunto das imagens sintéticas, foram duas imagens benignas e uma maligna classificadas corretamente, uma benigna classificada como maligna e seis imagens de nódulos classificadas com grau 3. Somando as imagens de nódulos reais classificadas com grau 3, temos oito nódulos indeterminados, sendo cinco deles de malignidade benigna e três deles maligna. Excluindo os nódulos com classificação de grau 3, temos que o

radiologista rotulou o grau de malignidade corretamente para 58.33% dos casos, mesmo com as limitações do Teste de Turing Visual aplicado.

Tabela 4.5: Teste de Turing Visual, em que GM é o grau de malignidade avaliada pelo radiologista e PR é a previsão do radiologista.

	Malignidade	Tipo de Imagem	GM	PR
Nódulo 1	Benigno	Sintético	1	Sintético
Nódulo 2	Benigno	Sintético	3	Real
Nódulo 3	Maligno	Real	2	Real
Nódulo 4	Benigno	Real	4	Real
Nódulo 5	Maligno	Real	5	Real
Nódulo 6	Maligno	Sintético	4	Sintético
Nódulo 7	Benigno	Sintético	3	Real
Nódulo 8	Maligno	Sintético	3	Sintético
Nódulo 9	Maligno	Real	4	Real
Nódulo 10	Maligno	Sintético	3	Sintético
Nódulo 11	Benigno	Real	2	Real
Nódulo 12	Benigno	Real	3	Sintético
Nódulo 13	Maligno	Real	2	Sintético
Nódulo 14	Benigno	Real	3	Real
Nódulo 15	Maligno	Sintético	3	Sintético
Nódulo 16	Benigno	Sintético	3	Real
Nódulo 17	Benigno	Sintético	2	Sintético
Nódulo 18	Maligno	Sintético	2	Real
Nódulo 19	Maligno	Real	2	Real
Nódulo 20	Benigno	Real	1	Real

5 Conclusão

O objetivo principal desse trabalho foi avaliar a hipótese de que a transferência de conhecimento aplicada através de um pré-treinamento da GAN resulta em nódulos pulmonares sintéticos mais eficientes para aumento de base. Nossos resultados evidenciaram que a técnica proposta (PREDCGAN) é melhor do que arquiteturas GANs clássicas e também que técnicas de aumento de base clássico. Adotar a transferência de conhecimento, através da utilização de um pré-treinamento na DCGAN como aumento de base, resultou em uma AUC de 0.781 na classificação de nódulos pulmonares. Esse resultado foi superior aos obtidos pela DCGAN (0.723), WGAN (0.764) e pelo método clássico (0.775). O nosso melhor resultado quando a PREDCGAN é usada em conjunto com aumento de base clássico resultando em uma AUC de 0.791 na classificação de nódulos pulmonares. Esse resultado foi superior aos obtidos pela DCGAN em conjunto com aumento de base clássico (0.771) e a WGAN em conjunto com aumento de base clássico (0.781). Além disso, as imagens geradas pela PREDCGAN demonstraram melhor convergência e qualidade visual quando comparadas a outros métodos GAN observados nesse trabalho.

5.1 Contribuições Científicas

O trabalho científico aceito foi:

- ANJOS, B. H. L. ; JATOBA, A. E. A. ; OLIVEIRA, M. C. . Generative Adversarial Networks: A Renewal for Data Augmentation in Lung Nodule Classification. In: XV Workshop on Computational Vision (WVC'2019), 2019, São Bernado do Campo. Anais do XV Workshop on Computational Vision (WVC'2019), 2019.

Referências Bibliográficas

ALBERTS, B. et al. *Biologia molecular da célula*. [S.l.: s.n.], 2010.

Arjovsky, M.; Chintala, S.; Bottou, L. Wasserstein GAN. *arXiv e-prints*, p. arXiv:1701.07875, Jan 2017.

ARMATO, S. G. et al. The lung image database consortium (lidc) and image database resource initiative (idri): A completed reference database of lung nodules on ct scans. *Medical Physics*, v. 38, n. 2, p. 915–931. Disponível em: <<https://aapm.onlinelibrary.wiley.com/doi/abs/10.1118/1.3528204>>.

ARMATO, S. G. et al. The lung image database consortium (lidc) and image database resource initiative (idri): A completed reference database of lung nodules on ct scans. *Medical Physics*, v. 38, n. 2, p. 915–931, 2011.

AZEVEDO-MARQUES, P. M. *Diagnóstico auxiliado por computador na radiologia*. [S.l.]: Radiologia Brasileira, 2001. v. 34. (285-293, 5).

BERMUDEZ, C. et al. Learning implicit brain MRI manifolds with deep learning . In: ANGELINI, E. D.; LANDMAN, B. A. (Ed.). *Medical Imaging 2018: Image Processing*. SPIE, 2018. v. 10574, p. 408 – 414. Disponível em: <<https://doi.org/10.1117/12.2293515>>.

BI, L. et al. Synthesis of positron emission tomography (pet) images via multi-channel generative adversarial networks (gans). In: CARDOSO, M. J. et al. (Ed.). *Molecular Imaging, Reconstruction and Analysis of Moving Body Organs, and Stroke Imaging and Treatment*. Cham: Springer International Publishing, 2017. p. 43–51. ISBN 978-3-319-67564-0.

CHARTRAND, G. et al. Deep learning: A primer for radiologists. *Radiographics : a review publication of the Radiological Society of North America, Inc*, v. 37 7, p. 2113–2131, 2017.

CHARTRAND, G. et al. Deep learning: A primer for radiologists. *RadioGraphics*, v. 37, n. 7, p. 2113–2131, 2017. PMID: 29131760. Disponível em: <<https://doi.org/10.1148/rg.2017170077>>.

CHARTSIAS, A. et al. Adversarial image synthesis for unpaired multi-modal cardiac data. In: TSAFTARIS, S. A. et al. (Ed.). *Simulation and Synthesis in Medical Imaging*. Cham: Springer International Publishing, 2017. p. 3–13. ISBN 978-3-319-68127-6.

CHUQUICUSMA, M. J. M. et al. How to fool radiologists with generative adversarial networks? a visual turing test for lung cancer diagnosis. *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, p. 240–244, 2018.

Costa, P. et al. End-to-end adversarial retinal image synthesis. *IEEE Transactions on Medical Imaging*, v. 37, n. 3, p. 781–791, March 2018. ISSN 0278-0062.

- DEGNAN, A. J. et al. Perceptual and interpretive error in diagnostic radiology—causes and potential solutions. *Academic Radiology*, v. 26, n. 6, p. 833–845, 2019. ISSN 1076-6332. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S107663321830521X>>.
- DENG, L. A tutorial survey of architectures algorithms and applications for deep learning. *APSIPA Transactions on Signal and Information Processing*, v. 3, n. e2, p. 1–29, 2014. Disponível em: <<https://doi.org/10.1017/atsip.2013.9>>.
- DICIOTTI, S.; AL. et. The LoG Characteristic Scale: A Consistent Measurement of Lung Nodule Size in CT Imaging. *IEEE Trans. Med. Imaging*, v. 29, n. 2, p. 397–409, 2010.
- DOI, K. Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. *Computerized Medical Imaging and Graphics*, v. 31, n. 4-5, p. 198–211, 2007.
- ERASMUS, J. J. et al. Solitary pulmonary nodules: Part i. morphologic evaluation for differentiation of benign and malignant lesions. *RadioGraphics*, v. 20, p. 43–58, 2010. PMID: 10682770. Disponível em: <<http://dx.doi.org/10.1148/radiographics.20.1.g00ja0259>>.
- ERHAN, D. et al. Why does Unsupervised Pre-training Help Deep Learning? *Journal of Machine Learning Research*, p. 625–660, 2010. Disponível em: <<http://jmlr.csail.mit.edu/papers/v11/erhan10a.html>>.
- ESMAEILISHAHMIRZADI, N.; MORTEZAPOUR, H. A novel method for enhancing the classification of pulmonary data sets using generative adversarial networks. *Biomedical Research*, v. 29, 01 2018.
- FALCÃO, A. X. *Visualização de Volumes Aplicada à Área Médica*. [S.l.]: Tese de Mestrado - UNICAMP, 1993.
- FERLAY, J. et al. Estimating the global cancer incidence and mortality in 2018: Globocan sources and methods. *International Journal of Cancer*, v. 144, n. 8, p. 1941–1953, 2019.
- FERREIRA-JUNIOR, J. R.; OLIVEIRA, M. C.; AZEVEDO-MARQUES, P. M. de. Characterization of Pulmonary Nodules Based on Features of Margin Sharpness and Texture. *J. Digital Imaging*, v. 31, n. 4, p. 451–463, 2018.
- FILHO, A. L.; MACHADO, A. P.; OLIVEIRA, M. *Modelo para Classificação de Nódulos Pulmonares Pequenos usando Descritores Radiomics*. Dissertação (Mestrado) — University of Alagoas - (UFAL), 2016.
- FRID-ADAR, M. et al. Gan-based synthetic medical image augmentation for increased cnn performance in liver lesion classification. *Neurocomputing*, v. 321, p. 321–331, 2018. ISSN 0925-2312. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0925231218310749>>.
- GALBUSERA, F. et al. Exploring the potential of generative adversarial networks for synthesizing radiological images of the spine to be used in in silico trials. *Frontiers in Bioengineering and Biotechnology*, Frontiers Research Foundation, v. 6, n. MAY, 5 2018. ISSN 2296-4185.

- GHASSEMI, N.; SHOEIBI, A.; ROUHANI, M. Deep neural network with generative adversarial networks pre-training for brain tumor classification based on mr images. *Biomedical Signal Processing and Control*, v. 57, p. 101678, 2020. ISSN 1746-8094. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1746809419302599>>.
- GILLIES, R. J.; KINAHAN, P. E.; HRICAK, H. Radiomics: Images are more than pictures, they are data. *Radiology*, v. 278, n. 2, p. 563–577, 2016. PMID: 26579733.
- GONZALEZ, R. C.; WOODS, R. E. *Processamento Digital De Imagens*. 3. ed. [S.l.]: Pearson Education, 2010.
- GOODFELLOW, I. et al. Generative adversarial nets. In: GHAHRAMANI, Z. et al. (Ed.). *Advances in Neural Information Processing Systems 27*. Curran Associates, Inc., 2014. p. 2672–2680. Disponível em: <<http://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf>>.
- GUAN, S.; LOEW, M. Breast cancer detection using synthetic mammograms from generative adversarial networks in convolutional neural networks. *Journal of Medical Imaging*, SPIE, v. 6, n. 3, p. 1–10, 2019. Disponível em: <<https://doi.org/10.1117/1.JMI.6.3.031411>>.
- Han, C. et al. Gan-based synthetic brain mr image generation. In: *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. [S.l.: s.n.], 2018. p. 734–738. ISSN 1945-8452.
- Han, C. et al. Combining noise-to-image and image-to-image gans: Brain mr image augmentation for tumor detection. *IEEE Access*, v. 7, p. 156966–156977, 2019.
- JUNIOR, J. R. F.; OLIVEIRA, M. C.; AZEVEDO-MARQUES, P. M. de. Cloud-based nosql open database of pulmonary nodules for computer-aided lung cancer diagnosis and reproducible research. *Journal of Digital Imaging*, v. 29, n. 6, p. 716–729, Dec 2016. ISSN 1618-727X. Disponível em: <<https://doi.org/10.1007/s10278-016-9894-9>>.
- KALENDER, W. A. X-ray computed tomography. *Physics in Medicine Biology*, v. 51, n. 13, p. R29, 2006. Disponível em: <<http://stacks.iop.org/0031-9155/51/i=13/a=R03>>.
- LANTZ, B. *Machine Learning with R*. 2. ed. [S.l.: s.n.], 2015.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *Nature*, Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved. SN -, v. 521, p. 436 EP –, May 2015. Disponível em: <<https://doi.org/10.1038/nature14539>>.
- LECUNN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *Nature*, v. 521, p. 436–444, 2015.
- LITJENS, G. et al. A survey on deep learning in medical image analysis. *Medical Image Analysis*, v. 42, p. 60–88, 2017. ISSN 1361-8415. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1361841517301135>>.
- LITJENS, G. J. S. et al. A survey on deep learning in medical image analysis. *CoRR*, abs/1702.05747, 2017. Disponível em: <<http://arxiv.org/abs/1702.05747>>.
- LIU, Y. et al. Wasserstein gan-based small-sample augmentation for new-generation artificial intelligence: A case study of cancer-staging data in biology. *Engineering*, v. 5, n. 1, p. 156–163, 2019. ISSN 2095-8099. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S2095809918301127>>.

- MADANI, A. et al. Chest x-ray generation and data augmentation for cardiovascular abnormality classification. In: ANGELINI, E. D.; LANDMAN, B. A. (Ed.). *Medical Imaging 2018: Image Processing*. SPIE, 2018. v. 10574, p. 415–420. Disponível em: <<https://doi.org/10.1117/12.2293971>>.
- MCBEE, M. P. et al. Deep learning in radiology. *Academic Radiology*, Elsevier, v. 25, n. 11, p. 1472–1480, Nov 2018. ISSN 1076-6332. Disponível em: <<https://doi.org/10.1016/j.acra.2018.02.018>>.
- MIRZA, M.; OSINDERO, S. Conditional generative adversarial nets. *CoRR*, abs/1411.1784, 2014. Disponível em: <<http://arxiv.org/abs/1411.1784>>.
- MONTAGNON, E. et al. Deep learning workflow in radiology: a primer. *Insights into Imaging*, v. 11, n. 1, p. 22, Feb 2020. ISSN 1869-4101. Disponível em: <<https://doi.org/10.1186/s13244-019-0832-5>>.
- MOSMANN, M. P. et al. Solitary pulmonary nodule and 18F-FDG PET/CT. Part 1: epidemiology, morphological evaluation and cancer probability. *Radiologia Brasileira*, scielo, v. 49, p. 35–42, 02 2016. ISSN 0100-3984. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-39842016000100035&nrm=iso>.
- OLIVEIRA, M. C. *Visualização de volumes em morfometria de estruturas cerebrais a partir de imagens de ressonância magnética nuclear*. Tese (Doutorado) — Universidade de São Paulo (USP), 2002.
- ONISHI, Y. et al. Automated pulmonary nodule classification in computed tomography images using a deep convolutional neural network trained by generative adversarial networks. *BioMed Research International*, v. 2019, p. 1–9, 01 2019.
- ONISHI, Y. et al. Multiplanar analysis for pulmonary nodule classification in ct images using deep convolutional neural network and generative adversarial networks. *International Journal of Computer Assisted Radiology and Surgery*, v. 15, n. 1, p. 173–178, Jan 2020. ISSN 1861-6429. Disponível em: <<https://doi.org/10.1007/s11548-019-02092-z>>.
- OUDKERK, M. et al. European position statement on lung cancer screening. *The Lancet Oncology*, Elsevier, v. 18, n. 12, p. e754–e766, Dec 2017. ISSN 1470-2045. Disponível em: <[https://doi.org/10.1016/S1470-2045\(17\)30861-6](https://doi.org/10.1016/S1470-2045(17)30861-6)>.
- PANG, T. et al. Deep learning radiomics in breast cancer with different modalities: Overview and future. *Expert Systems with Applications*, v. 158, p. 113501, 2020. ISSN 0957-4174. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0957417420303250>>.
- PINSKY, P. F. et al. National lung screening trial: Variability in nodule detection rates in chest ct studies. *Radiology*, v. 268, n. 3, p. 865–873, 2013.
- RADFORD, A.; METZ, L.; CHINTALA, S. Unsupervised representation learning with deep convolutional generative adversarial networks. *CoRR*, abs/1511.06434, 2015. Disponível em: <<http://arxiv.org/abs/1511.06434>>.
- RIEL S.J., J. C. S. E. van. Observer variability for lung-rads categorisation of lung cancer screening cts: impact on patient management. *Eur Radiol*, v. 29, n. 1, p. 924–931, 2019.

Salehinejad, H. et al. Synthesizing chest x-ray pathology for training deep convolutional neural networks. *IEEE Transactions on Medical Imaging*, p. 1–1, 2018. ISSN 0278-0062.

Salehinejad, H. et al. Generalization of deep neural networks for chest pathology classification in x-rays using generative adversarial networks. In: *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. [S.l.: s.n.], 2018. p. 990–994. ISSN 2379-190X.

SCHMIDHUBER, J. Deep Learning in Neural Networks: An Overview. *Neural Networks*, v. 61, p. 85–117, 2015. Disponível em: <<http://dx.doi.org/10.1016/j.neunet.2014.09.003>>.

SENARAS, C. et al. Optimized generation of high-resolution phantom images using cgan: Application to quantification of ki67 breast cancer images. *PLOS ONE*, Public Library of Science, v. 13, n. 5, p. 1–12, 05 2018. Disponível em: <<https://doi.org/10.1371/journal.pone.0196846>>.

SHIN, H. et al. Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *CoRR*, abs/1602.03409, 2016. Disponível em: <<http://arxiv.org/abs/1602.03409>>.

SIEGEL, R. L.; MILLER, K. D.; JEMAL, A. Cancer statistics, 2020. *CA: A Cancer Journal for Clinicians*, v. 70, n. 1, p. 7–30, 2020. Disponível em: <<https://acsjournals.onlinelibrary.wiley.com/doi/abs/10.3322/caac.21590>>.

SILVA, C. I. S. et al. Consenso brasileiro ilustrado sobre a terminologia dos descritores e padrões fundamentais da TC de tórax. *J. Bras. Pneumol*, v. 36, n. 1, p. 99–123, 2010. Disponível em: <<http://dx.doi.org/10.1590/S1806-37132010000100016>>.

SILVA, G. L. F. da; PAIVA, A. C. de; SILVA, A. C. *Diagnóstico de nódulos pulmonares em imagens de tomografia computadorizada usando redes neurais convolucionais evolutivas*. Dissertação (Mestrado) — University of Maranhão - (UFMA), 2016. Disponível em: <<https://tedebc.ufma.br/jspui/handle/tede/1534>>.

SONODA, S.; MURATA, N. Neural network with unbounded activations is universal approximator. *CoRR*, abs/1505.03654, 2015. Disponível em: <<http://arxiv.org/abs/1505.03654>>.

TAIGMAN, Y. et al. Deepface: Closing the gap to human-level performance in face verification. In: . [S.l.: s.n.], 2014.

Togo, R. et al. Anonymous gastritis image generation via adversarial learning from gastric x-ray images. In: *2018 25th IEEE International Conference on Image Processing (ICIP)*. [S.l.: s.n.], 2018. p. 2082–2086. ISSN 2381-8549.

WANG, Y. et al. Transferring gans: Generating images from limited data. In: FERRARI, V. et al. (Ed.). *Computer Vision – ECCV 2018*. Cham: Springer International Publishing, 2018. p. 220–236. ISBN 978-3-030-01231-1.

YAMASHITA, R. et al. Convolutional neural networks: an overview and application in radiology. *Insights into Imaging*, v. 9, n. 4, p. 611–629, Aug 2018. ISSN 1869-4101. Disponível em: <<https://doi.org/10.1007/s13244-018-0639-9>>.

YI, X.; WALIA, E.; BABYN, P. Generative adversarial network in medical imaging: A review. *Medical Image Analysis*, v. 58, p. 101552, 2019. ISSN 1361-8415. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1361841518308430>>.

ZHAO, D. et al. Synthetic medical images using fbgan for improved lung nodules classification by multi-scale vgg16. *Symmetry*, v. 10, p. 519, 2018.

ZHAO, H. et al. Synthesizing retinal and neuronal images with generative adversarial nets. *Medical Image Analysis*, v. 49, p. 14–26, 2018. ISSN 1361-8415. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1361841518304596>>.