



•NOVA•
UCSAL

Universidade Católica do Salvador
Bacharelado em Engenharia de Software

Ícaro Santana Silva
Sérgio Matheus Barbosa Oliveira

Síntese de imagens de Raios-X de problemas respiratórios
utilizando redes neurais adversariais generativas

Salvador

2020

Ícaro Santana Silva
Sérgio Matheus Barbosa Oliveira

**Síntese de imagens de Raios-X de problemas
respiratórios utilizando redes neurais adversariais
generativas**

Trabalho de Conclusão de Curso apresentado à Universidade Católica do Salvador como parte dos requisitos necessários para a obtenção do Título de Engenheiro de Software.

Orientador: Prof. Me. André Brasil Vieira Wyzykowski

Universidade Católica do Salvador

Salvador
2020

Ícaro Santana Silva
Sérgio Matheus Barbosa Oliveira

Síntese de imagens de Raios-X de problemas respiratórios utilizando redes neurais adversariais generativas

Trabalho de Conclusão de Curso apresentado à Universidade Católica do Salvador como requisito parcial para a obtenção do título de Engenheiro de Software.

Comissão Examinadora

Prof. Me. André Brasil Vieira Wyzykowski
Universidade Católica do Salvador
Orientador

Prof. Me. Osvaldo Requião Melo
Universidade Católica do Salvador

Prof. Me. Marcelo Índio
Universidade Católica do Salvador

Salvador, 22 de julho de 2020

Queremos dedicar esse trabalho a todos que contribuíram com nossa trajetória acadêmica, que nos apoiaram e incentivaram a não desistir apesar das dificuldades!

Agradecimentos

Agradecemos primeiramente a Deus por não ter nos desamparado nessa trajetória longa e cansativa que é a faculdade. Queremos agradecer a André Brasil Vieira Wyzykowski, nosso professor e orientador, sem a ajuda dele este trabalho não sairia do lugar. Ele nos guiou em todas as etapas para que não ficássemos perdidos no meio de tanta coisa. Queremos agradecer aos nossos familiares por sempre estarem ao nosso lado nos dando todo o suporte e também aos nossos amigos que mesmo com o isolamento social, não deixaram de nos confortar com palavras e conselhos.

Resumo

Diversos problemas no desenvolvimento de estudos científicos estão relacionados a escassez de informações. Dentro do contexto médico, existem doenças raras onde se têm poucos casos. Além disso, em estágios iniciais de uma nova doença, dados de pacientes ficam contidos em centros médicos, até que seja realizado o trabalho de compilação dos mesmos. Por fim, ainda existem a situação de dados pessoais estão resguardados sob leis de proteção, ocasionando um desbalanceamento de dados.

Atualmente essa dificuldade ficou evidenciada com a *Coronavirus disease 2019 (COVID-19)*. Por se tratar de uma doença nova, a comunidade científica tinha poucos dados precisos para realizar estudos mais detalhados.

Pensando nisso, a finalidade do trabalho é conceber imagens sintéticas de raios-x pulmonares, sendo estas geradas por inteligência artificial, que sejam visualmente realísticas a partir de poucas amostras reais para possibilitar a criação de repositórios que irá subsidiar futuras pesquisas e estudos.

Afim de alcançar o objetivo foi realizado uma pesquisa para atinar trabalhos próximos para o embasamento do desenvolvimento, experimentações e validações que contribuiriam para a etapa de treinamentos de modelos de redes neurais.

Para avaliar as informações providas deste treinamento foi necessário calcular sua qualidade em relação às amostras reais e para isto foi utilizado o cálculo da distância de *Fréchet*, medindo o intervalo das características entre elas, contribuindo com a avaliação do melhor modelo de rede neural aplicado. Contando também com um experimento de percepção humana, aplicado a 81 participantes de forma virtual, para avaliar a partir de suas óticas e julgar entre as imagens apresentadas seu critério ao considerar uma amostra como genuína ou sintética.

Com o resultado obtido do experimento de percepção humana foi possível identificar que as amostras julgadas pelos participantes como as mais próximas da realidade foram as mesmas que a partir da pontuação calculada pela distância de *Fréchet*. Com essas validações foi possível demonstrar que as imagens geradas pelo modelo generativo *Wasserstein generative adversarial network (WGAN)* foram superiores ao *Deep Convolutional Generative Adversarial Network (DCGAN)*.

Palavras-Chave: 1. Modelos Generativos Adversariais, 2. Augmentation. 3. Raio-x Torácico, 4. WGAN, 5. DCGAN.

Abstract

Diverses problems in scientific study development are related to information scarcity. In the medical context, there are rares diseases that have a few cases. Beyond that, in initial stages of a new disease, patient data keep stored in medical centers, until they realize a compilation work of them. Lastly, still there a situation of about personal data are kept safe under law protection, causing a data unbalancing.

Currently, this difficulty was evidenced with *Coronavirus disease 2019 (COVID-19)*. As it's a new disease, the scientific community had little precise data to carry out more detailed studies.

Thinking about this, the work goal is to conceive x-ray lung synthetic images, which were generated by artificial intelligence, being visually realistic from a few real samples to enable the creation of repositories that will support future research and detailed studies.

To reach the objective, one research was carried out to find the next works to support the development, experiments, and validations that contributed to the stage of training neural network models.

To validate the provided information of that training was necessary to calculate the quality related to the original samples, to do that was used the calculation of Fréchet distance, to measure the characteristics interval between them, contributing to the validation to the best neural model applied. Counting too with a human perception experiment was applied in 81 participants in visual form, to validate by his optics and judge between the presented images in your criteria to consider the sample as genuine or synthetic.

With the result obtained from the human perception experiment, was possible to identify that the samples judged by the participants as the closest to reality were the same as from the score calculated by the distance from *Fréchet*. With these validations, it was possible to demonstrate that the images generated by the model *Wasserstein generative adversarial network (WGAN)* were superior to the *Deep Convolutional Generative Adversarial Network (DCGAN)*.

Keywords: 1. Generative Adversarial Models. 2. Augmentation. 3. Chest X-ray, 4. WGAN, 5. DCGAN.

Lista de figuras

Figura 1 – Representação das camadas das <i>CNNs</i> —LeNet-5. (1) representa a imagem crua que é passada, enquanto (2) representa sua forma em matriz que posteriormente é utilizada no processo. Etapa simplificada de uma <i>CNN</i> contendo duas convoluções, o <i>pooling Max-pool</i> e finalização com as camadas totalmente conectadas.	24
Figura 2 – O bloco com D, representa o discriminador e o G, o gerador, o círculo em amarelo com x é a imagem real passada para o D com a finalidade de definir o real na primeira etapa, o treinamento com 3 etapas para ambos e são atribuídos pelas variações de classes em verde por 1,0 e 1. No D sua saída para a próxima etapa possuindo entrada para o G, como na segunda etapa fica responsável de gerar a partir de um ruído o x , levando o D a ter esta imagem gerada como entrada na terceira etapa, caso esta amostra seja confundida com a imagem real o D é “congelado”, não realiza ações e somente o G é treinado.	25
Figura 3 – Autoria própria: Representação básica de camadas dos modelos Gerador e Discriminador do <i>DCGAN</i>	27
Figura 4 – Representação do <i>Fréchet Inception Distance</i> , aplicado durante uma caminhada entre um Cão e seu Dono, evidenciando os pontos e as diferenças de características, sendo estas a posição de ambos.	29
Figura 5 – Presente no artigo Bozorgtabar et al. (2019), geradas pela <i>CAGAN</i> após a aplicação do método para a detecção de imagens.	34
Figura 6 – Presente no artigo Ying et al. (2019), Estes foram os resultados obtidos após a aplicação do modelo X2CTGAN com as suas camadas.	36
Figura 7 – Fluxo de etapas para comparação entre os resultados dos modelos. . . .	38
Figura 8 – Imagens numeradas e seu efeito ou stato respectivo: (1) original sem alterações, (2) invertida horizontalmente, (3) levemente rotacionada e (4) distorções leves	39
Figura 9 – Imagens geradas pelo modelo DCGAN	41
Figura 10 – Gráfico que demonstra a distância de Fréchet, que validam as imagens geradas pelo DCGAN, em comparação com as reais. No eixo y o maior valor é o <i>400</i>	42
Figura 11 – Imagens geradas pelo modelo WGAN	42
Figura 12 – Gráfico que demonstra a distância de Fréchet, que validam as imagens geradas pelo WGAN em comparação com as reais. No eixo y o maior valor é o <i>350</i>	43

Figura 13 – As imagens mostram características apontadas em amarelo, obtidas a partir de análises visuais das GANs. As imagens da primeira fileira, mostram apontam informações de (A1) e (B1) costelas, (A2) e (B2) a área mais clara inferior do parênquima. (A3) e (B3) a região que se aproxima um pouco dos segmentos dos brônquios e possíveis infiltrações do parênquima. (A4) e (B4) as regiões que demonstram os cantos com locais que mostram infiltração do parênquima.	43
Figura 14 – Características apontadas em amarelo, nas imagens reais apontando estruturas que compõem as figuras, como, ossos da costela e fluidos do pulmão	44
Figura 15 – Imagens geradas pelo modelo DCGAN	44
Figura 16 – Imagens geradas pelo modelo WGAN	45
Figura 17 – Comparação das imagens entre <i>dataset original</i> , <i>DCGAN</i> e <i>WGAN</i> . .	46
Figura 18 – Gráfico em pontos espalhados para exemplificar o resultado em percentual dos participantes do experimento, considerando o eixo x como acertos relativos ao experimento do DCGAN e o eixo y como acertos relativos ao experimento do WGAN.	48
Figura 19 – Gráfico semelhante da Figura 18, porém o círculo em vermelho demonstra uma população do experimento de pessoas mais confusas sobre as amostras, as quais não ajudaram muito já que não conseguiram diferenciar às duas classes das imagens.	49
Figura 20 – Gráfico semelhante da Figura 18, porém o círculo em vermelho, desta vez mostra parte da população que notou diferença entre as classes das imagens, ou seja, pessoas que notaram diferenças perceptíveis entre figuras que relacionaram como verdadeiras pelo seu histórico, estas que no que lhe concerne conseguiram mostrar que uma classe de imagem é visivelmente mais realística que as outras.	49

Lista de tabelas

Tabela 1 – Presente no artigo Ganesan et al. (2019), apresentando os resultados presentes no estudo feito sobre augmentation onde suas siglas <i>GA</i> , <i>TA</i> e <i>NA</i> representam aplicações com GAN e tradicional e sem aplicação de augmentation respectivamente.	31
Tabela 2 – Presente no artigo Salehinejad et al. (2017), apresentando a acurácia dos <i>datasets</i> , desbalanceados de imagens reais de radiografia, balanceados de imagens reais de raios-x e balanceados com imagens reais e sintetizadas de raios-x respectivamente como DS1, DS2 e DS3, para cada rótulo adotado está presente o nível de acurácia do modelo de GAN.	32
Tabela 3 – Tabela indicando quantitativamente os acertos para cada GAN relacionada ao experimento e a quantidade de imagens vistas, com a média das pessoas e a quantidade total de imagem que foram vistas.	52

Lista de Siglas e Abreviaturas

COVID-19	<i>Coronavirus disease 2019</i>
SARS-CoV-2	<i>Severe acute respiratory syndrome Coronavirus 2</i>
PLN	<i>Processamento de Linguagem Natural</i>
CNN	<i>Convolutional Neural Network</i>
GAN	<i>Generative Adversarial Network</i>
DCGAN	<i>Deep Convolutional Generative Adversarial Network</i>
HDCGAN	<i>High-resolution Deep Convolutional Generative Adversarial Network</i>
WGAN	<i>Wasserstein Generative Adversarial Network</i>
FID	<i>Frechet Inception Distance</i>
CAGAN	<i>Class Aware Generative Adversarial Network</i>
DCNN	<i>Deep Convolutional Neural Network</i>
GA	<i>GAN Augmentation</i>
TA	<i>Traditional Augmentation</i>
NA	<i>No Augmentation</i>
LGPD	<i>Lei Geral de Proteção de Dados</i>
GDPR	<i>General Data Protection Regulation</i>

Sumário

1	INTRODUÇÃO	17
1.1	Objetivo Geral	18
1.2	Objetivos Específicos	19
2	FUNDAMENTAÇÃO TEÓRICA	20
2.1	Doenças pulmonares	20
2.2	ChestX-ray8	21
2.3	Rede neural convolucional	22
2.4	Redes Adversárias Generativas	24
2.5	Deep Convolutional Generative Adversarial Network	26
2.6	Wasserstein Generative Adversarial Network	27
2.7	Frechet Inception Distance	28
3	TRABALHOS RELACIONADOS	30
3.1	Assessment of Data Augmentation Strategies Toward Performance Improvement of Abnormality Classification in Chest Radiographs	30
3.2	Generalization Of Deep Neural Networks For Chest Pathology Classification In X-rays Using Generative Adversarial Networks	31
3.3	Informative sample generation using class aware generative adversarial networks for classification of chest Xrays	33

3.4	X2CT-GAN: Reconstructing CT from Biplanar X-Rays with Generative Adversarial Networks	35
4	DESENVOLVIMENTO	37
4.1	Preparação do <i>dataset</i>	38
4.2	Tratamento de imagens	38
4.2.1	Augmentation	39
4.3	Treinamento e inferência da rede neural	40
5	EXPERIMENTOS E ANÁLISES	41
5.1	DCGAN versus WGAN	41
5.2	GANs versus real	44
5.3	Validações humanas	46
5.3.1	Hipóteses	47
5.4	Resultados	47
6	CONCLUSÕES	53
6.1	Trabalhos futuros	53
	REFERÊNCIAS BIBLIOGRÁFICAS	55

1 Introdução

A escassez de dados de doenças “raras” é um dos principais problemas enfrentados pelas áreas de estudos computacionais de diagnósticos. Casos de enfermidades pulmonares relatadas por análise possuem poucos exemplares disponíveis para estudo, tanto com finalidade profissional quanto acadêmica (BOZORGTABAR et al., 2019).

A Lei Geral de Proteção de Dados (LGPD), inspirada na lei europeia General Data Protection Regulation (GDPR), entrará em vigor em janeiro de 2021, estabelecendo uma série de proteções às informações sensíveis de uma pessoa, a fim de evitar seu reconhecimento (DEMO; FILHO, 2020). As restrições de privacidade se aplicam também a área de aprendizado de máquina, tornando a falta de dados maior (WACHTER; MITTELSTADT; RUSSELL, 2017).

O paciente pode optar por algum motivo não compartilhar as suas informações, muitas vezes não tendo ciência de que seu dado clínico possa contribuir para pesquisas, ou simplesmente não sabem da possibilidade de autorizar o uso dos seus dados em investigações médicas, no entanto permanecendo no anonimato. Havendo a possibilidade de exposição ao público por causa da doença que contraíram.

Caso o sigilo não seja obedecido, o médico poderá sofrer punições previstas em leis. A principal punição sofrida pelo profissional é a perda do direito de exercer a sua profissão, caso comprovado o vazamento (KOTTOW, 1986; SANKAR et al., 2003).

Atualmente o mundo está a enfrentar o *Sars-CoV-2 (Novel-coronavirus-2019)*, sendo essa uma doença infetocontagiosa de fácil transmissão. A transmissibilidade numa escala para o vírus alcança um valor de 4, sendo o intervalo de 1 a 5. Durante os primeiros meses de ocorrência da doença haviam uma quantidade reduzida de informações confiáveis, resultando em poucas de imagens de raio-x. Apenas alguns meses pós-a disseminação do vírus na China, a doença se propagou para o resto do mundo em 2020, como, por exemplo, a Itália em abril deste ano foi um dos países mais afetados (TARTAGLIONE et al., 2020). Surgindo a partir desta situação, dados confiáveis após tempos de estudo nas áreas de aprendizado de máquina, produzindo *datasets* provenientes dos casos confirmados e posteriormente síntese de imagem¹. Informações de sintetização e reconhecimento, como em Maguolo e Nanni (2020), foram possíveis pela existência destes repositórios, possibilitando outros estudos e diagnósticos com auxílio desses aprendizados.

Casos atuais como o *COVID-19*, foram afetados em partes por conta deste sigilo e vem quebrando este conceito em razão da sua rápida disseminação, se espalhando de maneira alarmante e sem registros prévios, sendo assim caso o paciente opte por não compartilhar os seus dados, pode dificultar estudos e possíveis curas para estas doenças.

¹ <https://github.com/ieee8023/covid-chestxray-dataset>

Conseguir alguns dados relevantes para finalidades de pesquisas em estados críticos é algo muito difícil de se conseguir, além de informações de condições dos mesmos. Os casos iniciais foram tratados como outras doenças, por isso podem não ter entrado para a estatística, tornando assim algumas características sobre o vírus imprecisas (FREITAS; NAPIMOGA; DONALISIO, 2020).

Visualizando a problemática de utilizar-se de um *dataset* desbalanceado para estudos e dados médicos sobre doenças raras, caso imagens sintéticas fossem geradas, estas poderiam ampliar um repositório de imagem. Auxiliando de alguma forma a vida de profissionais e estudantes. A partir de algoritmos é possível criar imagens sintéticas realísticas.

A *Convolutional Neural Network (CNN)* possui entre eles diversos exemplos de redes com múltiplas camadas que conseguem extrair e classificar características de imagens, e após essa extração é possível replicar estes numa estrutura de rede que alcança compendiar dados sintéticos a partir dos reais.

Este modelo de rede neural é constituído de múltiplas camadas inspiradas no córtex visual humano, assim como os neurônios e os conjuntos deles, as redes neurais computacionais são capazes de separar e aplicar filtros visuais para os “*pixels*” ao longo do processamento (VARGAS; PAES; VASCONCELOS, 2016).

Generative Adversarial Network (GAN) é um modelo generativo adversarial capaz de replicar padrões em conjuntos de dados. Utilizando dois conceitos, que são: o gerador, sendo este uma rede neural que gera dados sintéticos e o discriminador, outra rede neural que julga se os dados gerados são realistas ou não. O gerador sempre tentará enganar o discriminador durante o período de execução, sendo que as duas redes iterativamente aprendem com uma base real (METZ et al., 2016).

Utilizando GAN, em específico a *Deep Convolutional Generative Adversarial Network (DCGAN)*, dissertado na Seção 2.5 e o modelo *Wasserstein Generative Adversarial Networks (WGAN)* descrito na Seção 2.6, demonstrando estes modelos de redes neurais generativas possuem um padrão onde suas camadas convolucionais, seja quando as imagens são geradas, ou quando detectam as amostras de um conjunto de figuras com umas das finalidades a de solucionar o problema da falta de informação (ANJOS; JATOBÁ; OLIVEIRA, 2019).

1.1 Objetivo Geral

1. Gerar imagens sintéticas de raio-x pulmonares que sejam visualmente realísticas a partir de poucas amostras de imagens reais.

1.2 Objetivos Específicos

1. Realizar uma pesquisa dentro da área, afim de reconhecer e investigar assuntos análogos.
2. Comparar modelos de *GANs*, o *WGAN* e o *DCGAN*.
3. Realizar um experimento quantitativo para averiguar qual modelo generativo sintetiza imagens com melhor qualidade.
4. Realizar um experimento de validação humana para avaliar qual dos modelos conseguiu gerar imagens mais realísticas na ótica dos participantes.

2 Fundamentação Teórica

Já na Seção 2.3, será dissertado sobre o conceito de redes neurais convolucionais. Em seguida na Seção 2.4, será apresentado os modelos generativos adversarias, assim como os modelos derivados dessas redes como os das Seções 2.5 e 2.6. Por fim, a Seção 2.7, abordará um calculo sobre as distâncias entre características de dados.

2.1 Doenças pulmonares

Uma doença pulmonar é caracterizada pela diminuição das vias respiratórias e destruição do tecido cardiopulmonar. Enfermidades se agravam com o decorrer do tempo e com a chegada do século XX foi possível ter um preciso diagnóstico dessas doenças através de um exame detalhado ou um histórico familiar, porém é passivo de detecções quando se tem padrões muito expostos da enfermidade (MALLOZI; ROZOV, 1998).

O objetivo dos tratamentos é por muitas vezes evitar o contágio de uma doença cardiorrespiratória, não somente com tratamento é possível evitar estas enfermidades, mas também com a simples abstenção do fumo e minimizando da exposição à partículas, como asbesto, carvão, berílio, sílica, mofos, poeira de grãos, poluição e evitando além das citadas anteriormente substâncias que afetam os pulmões (LAB TESTS, 2019).

Com base no estudo de Zonzin et al. (2017), os sinais e sintomas variam entre pessoas e tempo, no entanto, cada doença tem suas próprias características. Existem sintomas comuns entre todas as enfermidades, como por exemplo, tosse persistente, respiração ofegante e dor torácica. Geralmente as pessoas com doenças obstrutivas crônicas têm problemas para expirar, assim como podem ter a pele com uma tonalidade mais azulada por conta da baixa oxigenação no sangue, porém estes sintomas não são tão comuns quanto os citados acima.

Cada doença pulmonar possui um sintoma determinado e também compartilha sintomas parecidos com outros para diagnóstico, no entanto, ainda assim é possível determinar causas e avaliar com seus exames. O diagnóstico precoce é muito importante. Para isso, quando o médico suspeita da doença, exames são solicitados, como a gasometria arterial, avaliando os níveis de oxigênio e dióxido de carbono no sangue e testes de funções pulmonares e radiografias ou tomografias computadorizadas para examinar a estrutura dos pulmões (AGUIAR et al., 1989). Essas doenças podem ser coletadas e separadas normalmente através de uma classificação, por rótulos onde conseguem ser divididas de forma ágil onde a finalidade de estudo ou os objetivos de pesquisa como uma estrutura de banco de dados de imagens.

2.2 ChestX-ray8

O ChestX-ray8 por Wang et al. (2017), é um banco de dados onde constam radiografias de tórax coletadas durante 23 anos, correspondentes as 32.717 pacientes chegando a totalizar 108.948 doenças catalogadas, das quais foram obtidas variações e padrões com nomenclaturas patológicas oriundas de relatórios, gerando diagnósticos através das radiações. Sendo assim, permitiria a visualização de algumas estruturas internas do corpo, como órgãos e ossos, gerando saída do procedimento uma imagem. Os textos das imagens são obtidos através de etiquetas, geradas a partir do Processamento de Linguagem Natural (PLN), no formato interno definido por Annarumma et al. (2019), com o auxílio da inteligência artificial de um algoritmo de sistema. Assim sua arquitetura é composta de duas CNNs, este será maior detalhado na Seção 2.3). Estas com treinamento conseguem detectar determinados padrões nas figuras e registrar como enfermidades.

As principais doenças catalogadas no repositório foram:

1. Atelectasia: Segundo Cunha et al. (2017) ocorre quando um pulmão, ou brônquio têm a sua passagem de ar obstruída por uma secreção, ou objeto, dificultando a entrada de ar e reduzindo a quantidade de alvéolos funcionantes.
2. Cardiomegalia: Caso o índice cardiorácico corporal estiver acima de 0,50, normalmente indicará um aumento do coração (SILVA et al., 2011).
3. Derrame pleural: Ocorre quando há bacilos resultantes de antígenos do *Mycobacterium tuberculosis*, que segundo Seiscento, Conde e Dalcolmo (2006) se na região da membrana, eles conseguem adentrar e após alguma ruptura por lesão pode resultar em água no pulmão.
4. Pneumonias eosinofílicas: Pela definição de Magalhães, Tavares e Chieira (2006) ocorrem quando há uma infiltração no parênquima pulmonar por eosinófilos.
5. Massa pulmonar: Segundo o Mosmann et al. (2016) são pequenas lesões pulmonares de até 30 mm de diâmetro.
6. Nódulo pulmonar: São bem parecidas com a massa no pulmão, sendo difícil de detectar a diferença no momento da classificação e segundo Jr et al. (2002) possuem uma opacidade arredondada na sua forma, podendo apresentar alterações no tamanho da mesma.
7. Pneumonia: Por Costa et al. (2016), caracteriza-se por uma obstrução na área do parênquima pulmonar afetando os alvéolos, brônquios respiratórios e os interstícios.
8. Pneumotórax: Pela definição de Noppen e Keukeleire (2008) ocorre quando há presença de ar no espaço pleural.

Uma técnica de normalização chamada *Disease Normalization*, definida por Leaman, Doğan e Lu (2013) consistindo em um método para determinar as doenças em diagnósticos de textos médicos foi uma das primeiras técnicas a utilizar aprendizado de máquina a identificar as similaridades entre as enfermidades.

Catalogando pela forma por etiquetas e também utilizando a *MetaMap*, sendo este um programa configurável que utiliza a Biblioteca Nacional de Medicina (*National Library Medicine*)¹, como base o PLN, se baseia em símbolos e procedimentos de linguagem de computador, unidos a Inteligência artificial por Aronson e Lang (2010), na aplicação do projeto é posta como comparativo de qualidade usando métricas de medidas.

Segundo Wang et al. (2017) a caracterização de algumas imagens podem ser assemelhadas a doenças de acordo com padrões encontrados e estas características geralmente não são únicas, como, por exemplo, as Pneumonias eosinofílicas que são capazes de serem associadas com Atelectasia e Derrame pleural. Sendo esta uma possibilidade para designar mais de uma doença por imagem, quando fossem classificadas, algumas destas pelos seus respectivos nomes ou caso não fossem identificadas dentre as categorias de enfermidades seriam interpretadas como “Normais”.

O *dataset* pode ser utilizado para criar aplicações significativas, caso seja identificado um padrão das doenças, a criação de relatórios médicos automatizados, proporciona um benefício ao processamento em si tornando-o mais fácil. Esta aplicação pode ser expandida e não somente as doenças específicas do *dataset*.

A partir de uma identificação supervisionada, houve a possibilidade de padrões serem detectados através de redes neurais, para que sejam identificados referências e assim como, humanos os identificam, seja prováveis reproduzi-los com melhor precisão.

2.3 Rede neural convolucional

Rede Neural Convolucional (CNN) é uma arquitetura inspirada em técnicas de aprendizado, representando os padrões das imagens e aprendendo estas características (LECUN; KAVUKCUOGLU; FARABET, 2010).

Esta rede é uma forma de algoritmo como *deep learning* e projetada entre dimensões. A arquitetura destas redes possuem estruturas de multi-dimensões, sendo atualizado em cada nova dimensão uma melhora no aprendizado destas redes através de estímulos para identificar os padrões de cada imagem (JURASZEK et al., 2014).

A convolução é uma operação matemática entre dimensões, representando uma imagem como uma matriz. A figura possui áreas denominadas *pixels* que são separadas como partes da matriz. A estrutura dessas camadas convolucionais são compostas por filtros, também em formato matricial para obter características das imagens. Estas informações

¹ <https://www.nlm.nih.gov/>

da figura original são mantidas, resultando que apenas informações com finalidades classificativas estejam presentes nos filtros. Áreas presentes na imagem são multiplicadas pelos filtros obtendo assim características de partes da imagem.

Caso a figura possua mais de uma camada referente aos canais de cores, o mesmo será aplicado para o filtro. Os pesos, definidos por processos externos de caráter evolutivo para a melhora da rede, são calculados de forma matemática e para cada ponto são feitas conexões da rede para utilizar uma função de ativação do filtro para encontrar maior área, decidindo assim se um neurônio será ou não estimulado (SANTOS, 2019).

Durante a captura de informações da imagem uma técnica benéfica para isso é o *edge detection*. Este é uma forma de se detectar bordas das imagens do dataset. Com base em critérios obtidos através dos filtros, desse modo é possível desenhar estes contornos, de modo a detectá-las melhor. (CANNY, 1986).

“A detecção de borda é uma etapa muito útil de pré-processamento no processamento de imagens e em aplicativos baseados em visão computacional” (ANAND; NAGAJOTHI; NITHYA, 2020)

Isso pode facilitar e acelerar tanto o processo de treinamento com a detecção da imagem e fundo da mesma, assim como pode adiantar alguma parte da detecção das imagens utilizada pós treinamento.

Em sua arquitetura de camadas, após a convolução da rede, as regiões de peso se ligam a camadas denominadas de *Pooling*, em um formato linear tem-se todas as possibilidades e quanto mais utilizado melhor será resultado auxiliando o processo da CNN. A técnica utilizada deve-se extrair características das convoluções sendo passadas de camada em camada. Na etapa de detecção o comportamento desse modelo pode ser refinado e passado adiante de acordo com o núcleo definindo como ponto central da região separada na camada.

A utilização da técnica *Max pooling* têm um papel importante na melhora do desempenho de toda a execução, é considerado toda a matriz da imagem, como demonstrado na Figura 1, representada através de LeNet-5². É benéfico para a memória ter uma saída reduzida de dados, pois evita a incapacidade das amostras de preverem novos padrões fixando resultados nesses filtros. Diminuindo as entradas e tornando as características relevantes e mais nítidas para o filtro (ZANINOVICH, 2019).

Após as camadas convolucionais e de pooling tem-se uma camada denominada totalmente conectada, concebendo a decisão a partir das respostas dos filtros ligados pelas estruturas anteriores, carregando a informação ligando ao neurônio de saída, esta camada é formada por um vetor (VARGAS; PAES; VASCONCELOS, 2016).

² <http://alexlenail.me/NN-SVG/LeNet.html>

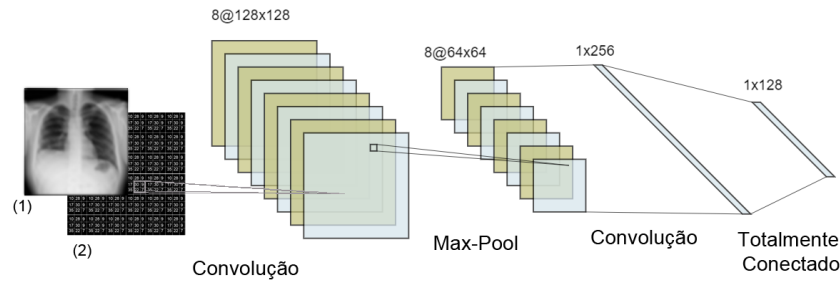


Figura 1 – Representação das camadas das *CNNs* —LeNet-5. (1) representa a imagem crua que é passada, enquanto (2) representa sua forma em matriz que posteriormente é utilizada no processo. Etapa simplificada de uma *CNN* contendo duas convoluções, o *pooling Max-pool* e finalização com as camadas totalmente conectadas.

Durante o processo realizado pela CNN se pode classificar e utilizar as imagens em um vetor, essa classificação é uma parte das camadas de convolução, sendo algo presente e necessário em modelos generativos de rede neural.

2.4 Redes Adversárias Generativas

Redes Adversárias Generativas (GAN), são modelos neurais compostas por duas camadas competindo entre si, por isso são chamadas de adversárias. Este modelo é uma nova forma de arquitetura, podendo replicar qualquer distribuição de dados do nosso mundo, bastando ter a forma de treino correta (DEEPLARNINGBOOK, 2015).

A partir da definição de Weng (2019), GAN é uma estrutura com dois modelos, o gerador e o discriminador. O executor das funções responsável por classificar as amostras geradas a partir das reais respeitando os parâmetros do modelo é chamado de discriminador. O gerador concebe saídas sintéticas a partir de uma variável de ruído diversificando as amostragens na saída. Este busca sempre passar uma espécie de amostra para enganar o crítico com o papel de verificar se a imagem é similar às realistas definidas durante o treinamento, definidas pelo *dataset* como base de critério no seu papel. Esta etapa também pode ser representada como na Figura 2 por Pereira (2018), demonstrando este processo de treinamento visualmente.

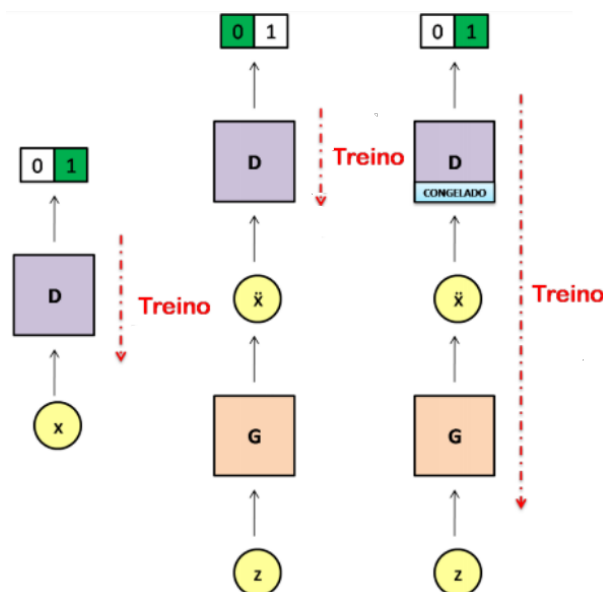


Figura 2 – O bloco com D, representa o discriminador e o G, o gerador, o círculo em amarelo com x é a imagem real passada para o D com a finalidade de definir o real na primeira etapa, o treinamento com 3 etapas para ambos e são atribuídos pelas variações de classes em verde por 1,0 e 1. No D sua saída para a próxima etapa possuindo entrada para o G, como na segunda etapa fica responsável de gerar a partir de um ruído z , levando o D a ter esta imagem gerada como entrada na terceira etapa, caso esta amostra seja confundida com a imagem real o D é “congelado”, não realiza ações e somente o G é treinado.

Goodfellow et al. (2014) diz que uma GAN pode ser usada em vários ramos das ciências, como na astrologia utilizada para a recuperação de informações presentes em imagens de estrelas com ruídos Schawinski et al. (2017). Também pode ser utilizada no ramo dos jogos virtuais no aprimorando métodos de renderização permitindo a utilização de filtros em tempo real e diversas outras aplicações em geral, como em Zhu et al. (2017) é demonstrado a aplicação em imagens e vídeos e *Chitan Trivedi*³, na utilização desse modelo com imagens de dois jogos virtuais para treinar os modelos e conseguir modificar a textura de um dos jogos com base no outro. Possibilidades concebidas também graças à formação das camadas do gerador e discriminador genéricos para poder representar diversas informações do mundo real.

Diversos modelos de GANs foram criados para cada vez mais melhorar o nível para os modelos internos conseguirem gerar uma saída mais realista, como exemplo se pode citar o DCGAN presente na Seção 2.5 e o WGAN presente na Seção 2.6.

³ <https://towardsdatascience.com/turning-fortnite-into-pubg-with-deep-learning-cycle-gan-2f9d339dcd0>

2.5 Deep Convolutional Generative Adversarial Network

Como demonstrado no artigo Radford, Metz e Chintala (2015) nos últimos anos o aprendizado supervisionado das redes neurais tiveram adoções no uso de visão computacional, essa visão pode ser incrementada por grandes quantidades de vídeos e imagens para melhorar o treinamento.

Podemos definir o DCGAN como sendo um modelo de aprendizado supervisionado possuindo uma camada de convolução mais profunda para aprimorar os cálculos das aplicações de novos filtros nas imagens (RADFORD; METZ; CHINTALA, 2015).

O treinamento do modelo de rede supervisionada deve ser feito por um conjunto de imagens, devem mostrar evidências convincentes de algo a se replicar ou conferir se é cópia, a execução destas rotinas é feita de forma hierárquica representando partes de objetos consistentes, tanto no gerador quanto no discriminador este tipo de execução é utilizado para obter um melhor resultado.

Como cita no artigo Mallmann et al. (2019), uma GAN após implementada pode gerar uma grande quantidade de imagens sintéticas. Geralmente no início deste treinamento e na criação dessas representações, são concebidas com bastante ruído e são pouco entendíveis, porém no final do processo se tornam mais compreensivas após a aplicação de filtros.

O DCGAN é uma GAN com diversos usos, e um desses usos foi no japão para treinamento em um *dataset* de tsunamis na cidade de *Hokkaido*, esta rede pós treinamento pode ser utilizado nessa mesma região, onde com base nos acontecimentos naturais anteriores, possa ser previsto futuros desastres (CHENG et al., 2020).

Conforme o sucesso do uso do DCGAN alguns autores utilizaram dessa arquitetura para diversas aplicações sendo algumas dessa, coloração em tempo real de imagens de sensores infravermelhos como em Suarez, Sappa e Vintimilla (2017), ou extração espacial da imagem, sendo esse descobrimento de novos corpos celestes, com base no modelo DCGAN (CHEN et al., 2017).

Como mostrado no artigo Radford, Metz e Chintala (2015) o DCGAN é um modelo modificado, para ficar mais estável com o uso de diversas configurações e treinamentos, as redes DCGAN possuem um gerador, como uma camada propriamente utilizada para gerar as imagens com base no *dataset* treinado e a camada de discriminador que é aproveitada para determinar qual imagem é falsa ou real, correndo em paralelo às duas camadas vão crescendo, como apresentado na Figura 3 onde podemos visualizar uma representação básica de representação dos modelos do DCGAN.

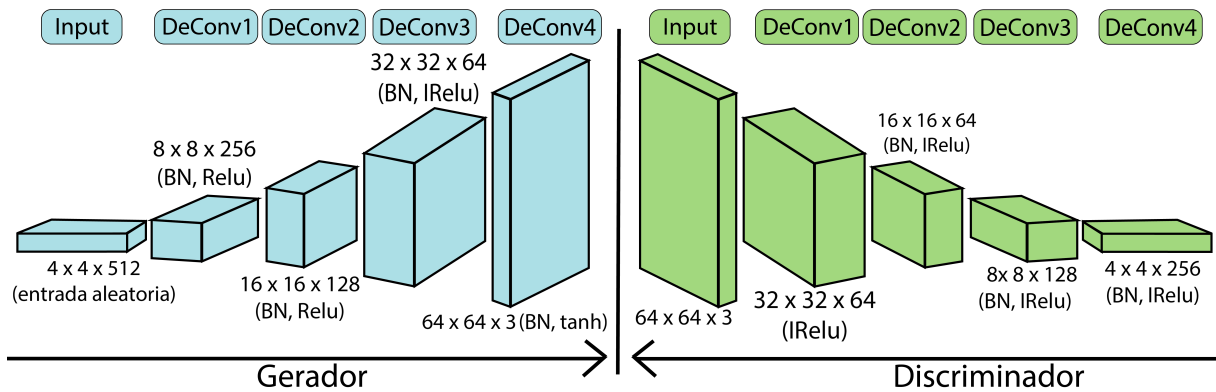


Figura 3 – Autoria própria: Representação básica de camadas dos modelos Gerador e Discriminador do *DCGAN*

A proposta base de implementação do DCGAN gera imagens ligeiramente desfocadas, isso acontece porque o este modelo processa estas já com baixa qualidade para ter um melhor desempenho durante o processo, entretanto quando é buscado processar uma imagem de maior resolução (512x512), nota-se instabilidade durante o treinamento.

Para solucionar esse problema foi introduzido o conceito da *High-resolution Deep Convolutional Generative Adversarial Network (HDCGAN)*, utilizando redes neurais auto-normalizantes possuindo camadas conectadas através de pesos relacionados entre elas. Essa rede além de utilizar uma forma diferente de conectar as camadas, implementa uma nova camada convolucional para poder gerar as imagens em alta resolução (CURTÓ et al., 2017).

Assim como o DCGAN as GANs existem com características e semelhanças presentes entre alguns desses modelos, variando da utilização de cálculos matemáticos para obterem melhores valores para as convoluções e então amostras sintéticas com maior definição.

2.6 Wasserstein Generative Adversarial Network

Uma GAN pode possuir instabilidade durante o treinamento, a forma do modelo base original ou seus derivados semelhantes podem possuir falhas, ou as amostras eram muito instáveis durante as gerações (GULRAJANI et al., 2017). Então foram criados diversos outros modelos para melhorar a síntese dos dados.

O WGAN possui uma métrica com um intervalo entre duas prováveis distribuições, considerando a medida mínima para se transformar uma distribuição em outra provável diversa, definida como distância de *Wasserstein*. Diferente para uma melhoria do modelo base e adaptações foram feitas desde o GAN original, sendo este uma adaptação do DCGAN (ANJOS; JATOBÁ; OLIVEIRA, 2019).

Como proposta para diminuir os problemas encontrados no GAN original, foi substituído a divergência comum utilizada nas métricas pela presente no WGAN, tornando

alinhamento da convergência dos dados mais próximos dos reais para calcular a distância entre as informações sintetizadas e utilizar no critério de evolução dos modelos (DOAN et al., 2020). Pode-se tornar viável a utilização do WGAN já possuindo uma melhora se comparado com modelo base, melhorando sua estabilidade e velocidade (ARJOVSKY; CHINTALA; BOTTOU, 2017).

A utilização do WGAN pura por si ainda apresenta alguns problemas de otimização e também durante a distribuição de características de um treinamento o modelo pode não conseguir captar diferentes padrões de valores de pesos, assim levando o mesmo a perder algumas informações para o seu gerador. Para melhorar o desempenho foi substituído o formato padrão de captação de características durante uma distribuição do modelo um gradiente de penalidade, sendo uma pequena alteração durante a parte do processo da geração dos modelos e como citado é mostrado a modificação não gera comportamentos de alguma falha na obtenção de uma alta quantidade de tais distribuições (GULRAJANI et al., 2017).

Essas informações sintetizadas de GANs criam dados e aprendem de acordo com conhecimentos passados, porém além da análise visual é necessário validá-los de forma matemática e de acordo com as amostras geradas destoam das originais e podemos calcular isso através da Distância de *Fréchet* presente na Seção 2.7.

2.7 Frechet Inception Distance

Apesar do progresso notável com o uso das GANs, a avaliação quantitativa desses modelos geralmente não corresponde ao que se espera, pela falta de métricas para avaliar diferentes propriedades desejáveis, como qualidade da imagem, consistência e diversidade. Nesse cenário, o conjunto e uso de avaliações de grandezas quantitativas são desafiadoras, considerando que cada métrica pode indicar que um modelo é superior a outro com base na dimensão avaliada. Utilizar-se do cálculo da *Fréchet Inception Distance (FID)* é o mais padronizado, pois é considerado o nível de distorção entre as características das imagens (HEUSEL et al., 2017).

A distância de *Fréchet* é uma forma em se medir a intervalo de características entre dois objetos distintos em um cálculo, um formato usado para avaliar as amostras das GANs, é a distância máxima entre um primeiro ponto de duas curvas e se movesse continuamente modificando a segunda curva, por exemplo, um cão e seu dono na sua caminhada, naquelas calçadas que possuem dois caminhos separados por um caminho de grama no meio, sendo na primeira parte da estrada o do proprietário do cachorro e na outra o do próprio animal. A corda deve ser longa para que o cão e o dono andem pelo caminho que formará uma curva ou retas, do ponto inicial até o final do percurso, como demonstrado na Figura 4. Precisamos encontrar a maneira ideal para isso, as medidas que a coleira irá possuir

denominam esse cálculo (HAR-PELED, 2014).

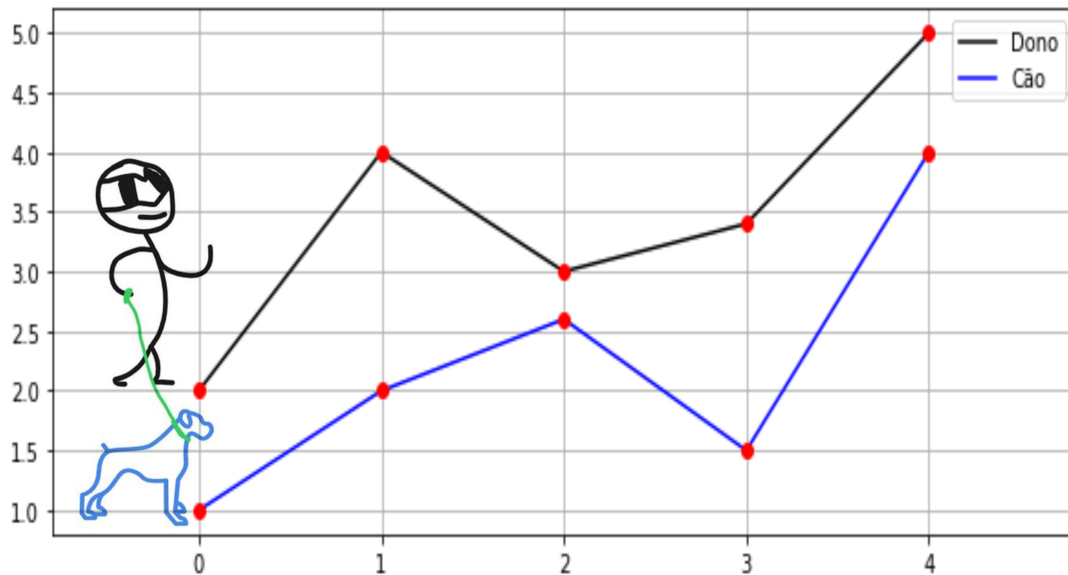


Figura 4 – Representação do *Fréchet Inception Distance*, aplicado durante uma caminhada entre um Cão e seu Dono, evidenciando os pontos e as diferenças de características, sendo estas a posição de ambos.

O FID mais baixo é geralmente sinônimo de correspondência ao modelo generativo que pode gerar imagens com qualidade visual mais realista, possibilitando melhores resultados (DOAN et al., 2020). Segundo Lee e Seok (2020), para avaliar a qualidade e diversidade das amostras geradas pelas GANs, a unidade de medida utilizada é a *Inception* e esta é uma grandeza usada na classificação de imagens, analisando-a como uma métrica de validação de figuras (HORAK; YU; SALIMI-KHORSHIDI, 2020).

Heusel et al. (2017) como mostrado no artigo a FID é o intervalo, enquanto o Inception é uma pontuação. Para compará-los transformamos o Inception para uma distância, que chamamos de inicial (*Inception Distance (IND)*), esta transformação só é possível, caso o Índice da Inception possua o seu valor máximo.

3 Trabalhos Relacionados

Neste capítulo será discorrido sobre 4 importantes trabalhos relacionados com o objetivo deste documento. A Seção 3.1 cita a importância de aplicação do *augmentation* para um repositório de imagem com poucas amostras e seu impacto em um treinamento de um modelo de classificação. Na Seção 3.2 foi proposto um novo modelo da *DCGAN*, o qual gera uma arquitetura base do modelo com mais resolução e nitidez, para isso houve adições de camadas convolucionais as quais obtiveram uma melhora na base, no estado da arte. Na Seção 3.3 é demonstrado em uma aplicação de um *framework* de aprendizado ativo propondo algumas *GANs* e a utilização de um modelo generativo que será explicado mais nesta seção. Já na Seção 3.4 é citado uma nova arquitetura para a *GAN* em que está implementa uma forma de segmentação e classificação que demonstra um melhor desempenho na seleção.

3.1 Assessment of Data Augmentation Strategies Toward Performance Improvement of Abnormality Classification in Chest Radiographs

Neste artigo é relatado a importância de se utilizar *Augmentation*, no seguimento do estudo proposto por Ganesan et al. (2019), se demonstra com clareza a necessidade e notabilidade que é a quantificação de dados em um *dataset* que possui poucas imagens por alguma eventualidade ou carência. Exemplificações no modelo proposto para uma aplicação do modo tradicional de *augmentation* dada sua eficiência e através de uma *GAN* e também uma forma que seguem sem o *augmentation*.

O estudo proposto utiliza a comparação desses três modos de possíveis caminhos a se seguir em uma classificação de imagens de pulmões de raio-x. O modelo implementado para executar o processo foi a *progressively-growing GAN*, marcando a aplicação do *GAN* com o *augmentation* onde deve a imagem caminhar por passos crescentes onde a figura com uma resolução pequena, muito menor que as amostras originais, com determinados valores de pesos ela treine e consiga chegar em um tamanho considerável, com os mesmos pesos, pois estes são fixos, porém, sempre aumentando os *pixels*. Os processos também foram aplicados entre o modelo tradicional de *augmentation* e uma forma que não utilizou o *augmentation*.

Ao ter o resultado dos três modelos de estratégias de augmentation, utilizaram ferramentas de métricas para avaliar o desempenho. A forma sem o augmentation foi menos performática do que os outros dois modos, como demonstrado na Tabela 1 presente no artigo, como resultado do experimento realizado. A forma de aplicação dos aumentos de dados melhora o desempenho do classificador de imagens. A aplicação do modelo tradicional se sobressai ao executado com a GAN, no critério de classificação de imagens por ser levemente mais performático. Assim como os modelos de treinamento *DCGAN* e *WGAN*, foi necessário o uso deste mesmo tipo de arquitetura em que o *augmentation* foi utilizado para aumentar as características, as quais deseja-se replicar para validar a etapa de aplicação do *augmentation* do *ChestX-ray8*.

Method	% Augmentation	ACC	AUC	F-SCORE	MCC
NA (Baseline)	0	79.32	86.03	0.84	0.55
TA	25	81.20	88.01	0.86	0.58
	50	81.38	88.34	0.86	0.59
	75	82.23	88.57	0.86	0.59
	100	84.25	90.70	0.88	0.64
GA	25	80.82	87.84	0.85	0.58
	50	81.34	88.05	0.86	0.58
	75	80.41	87.42	0.85	0.58
	100	83.94	90.29	0.88	0.64

Tabela 1 – Presente no artigo Ganesan et al. (2019), apresentando os resultados presentes no estudo feito sobre augmentation onde suas siglas *GA*, *TA* e *NA* representam aplicações com GAN e tradicional e sem aplicação de augmentation respectivamente.

3.2 Generalization Of Deep Neural Networks For Chest Pathology Classification In X-rays Using Generative Adversarial Networks

Estudo apresentado por Salehinejad et al. (2017), salienta que um conjunto de dados médicos podem ser desbalanceados, com uma escassez de informações ou de uma doença muito rara isso é dado principalmente do sigilo entre o médico e paciente, por isso foi proposto simular uma nova patologia, utilizaram-se de uma GAN e de algumas imagens da região torácica, para poder replicar raios-x do tórax, utilizando-se de uma *Deep Con-*

volutional Neural Network(DCNN), foi treinado com amostras reais e artificiais para obter uma precisão maior na detecção dessas patologias.

Foi proposto um novo modelo de DCGAN o DCNN isso foi feito já que as imagens do treinamento foram embaralhados com dados genuínos, a GAN foi treinada com informações reais e sintéticas para depois poder prever a classificação de patologias torácicas.

Após a execução do trabalho pode ser constada que as DCGAN classificatórias tiveram um potencial maior de acerto após terem sido treinadas por dados reais e fictícios em comparação às outras somente treinadas com informações autênticas sendo estes *datasets* balanceados e desbalanceados, como demonstrado na Tabela 2.

<i>Accuracy (%)</i>	DS1	DS2	DS3
Cardiomegalia	79.15	71.73	95.31
Normal	77.75	72.53	95.02
Derrame pleural	73.64	51.23	91.19
Edema pulmonar	65.86	50.12	89.68
Pneumotórax	57.99	48.92	88.84
Total	70.87±0.47	58.90±0.48	92.10±0.41

Tabela 2 – Presente no artigo Salehinejad et al. (2017), apresentando a acurácia dos *datasets*, desbalanceados de imagens reais de radiografia, balanceados de imagens reais de raios-x e balanceados com imagens reais e sintetizadas de raios-x respectivamente como DS1, DS2 e DS3, para cada rótulo adotado está presente o nível de acurácia do modelo de GAN.

Com esse artigo se pode notar que o uso da *GAN* para crescimento de *datasets* é muito importante e ajuda essas redes a obterem um melhor resultado de desempenho, fazendo com que a acurácia desta tenham um seguimento consideravelmente maior. O aumento da diversidade do conjunto de dados de treinamento permite, que eles obtenham um melhoramento no desempenho da generalização do aprendizado profundo para classificação de informações não vistas.

Este trabalho poderia ser visto como um melhoramento do estado da arte do *DCGAN*, para uma melhor desta aplicação e obtenção de melhores resultados. Através do uso de novas camadas convolucionais e uso de novos formatos de entradas de imagens para poder obter melhores resultados.

3.3 Informative sample generation using class aware generative adversarial networks for classification of chest Xrays

Em Bozorgtabar et al. (2019), o estudo proposto viabiliza a utilização de uma estrutura de aprendizado ativo com um modelo de melhora estatístico de métricas, para possibilitar o enriquecimento dos critérios, o treinamento do modo de aprendizagem profunda durante a geração das amostras é utilizado uma *Class aware generative adversarial network (CAGAN)* com a classificação de uma CNN para área médica, gerando dados para a aplicação do augmentation.

Para uma maior exatidão e certeza durante o aprendizado do modelo, foi proposto um modo de síntese que a partir de um padrão estatístico, esta rede neural profunda separaria as melhores amostras, por testes em incertezas, onde os resultados passariam por classificação e segmentação ou separação das demonstrações excelentes.

Mesmo com escassez de imagens de um repositório o método proposto pela geração de figuras realísticas, pode ser visto como uma aplicação com o *framework* ativo de aprendizado.

Gerando assim uma forma que após o processamento consiga gerar mais amostras para o modelo, focado em treinar o gerador em um direcionamento específico para outras entradas a partir de um encaminhamento de uma rede de segmentação denominada *UNet*, para o melhor resultado e o discriminador que, diferentemente das GANs ocasionais é aplicado configurações para que consiga produzir distribuições de possibilidades entre suas fontes e etiquetas.

Após um treinamento onde foram passados condições lineares de taxas de erros para o discriminador e gerador, com o objetivo de suscitar que o criador conseguisse gerar amostras mais próximas dos modelos reais, sendo constituído de três camadas convolucionais e algumas propriedades básicas que compõem um gerador. Enquanto o discriminador possui seis camadas convolucionais, sendo a última uma camada densa, além de possuir as propriedades básicas relacionadas ao *stride* que seria um redutor do modelo em relação aos modelos comuns.

A geração de imagens sintéticas foi feita com um conjunto de máscaras codificadas para cada entrada de imagem, diversos conjuntos aplicados para gerar figuras. Os resultados adjuntos ao modelo pela condição realística, onde eles poderiam servir como característica ao próprio ciclo de treino de outros modelos pela qualidade de geração das amostras. A acurácia deste modelo deve se também a combinação com as métricas das probabilidades de incertezas tendo um melhor desempenho.

No capítulo de experimento o método foi aplicado no *dataset* do NIH Chest X ray, onde o mesmo possui imagens de raios-x frontais da região torácica, com isso foi detectado algumas doenças, como demonstrado na Figura 5.

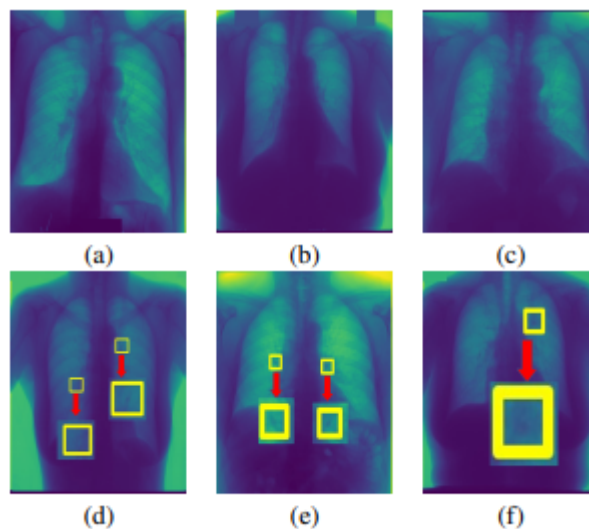


Figura 5 – Presente no artigo Bozorgtabar et al. (2019), geradas pela CAGAN após a aplicação do método para a detecção de imagens.

A rede neural foi implementada com o tensorflow utilizando o otimizador ADAM, e o GAN final foi atualizada com uma taxa de 10^{-3} , os experimentos se concentraram na classificação das imagens das doenças com base em pontos (chaves) detectados nos raios-x, sendo assim foi utilizada a arquitetura *DenseNet* sendo está uma estrutura para seleção de figuras.

O conjunto de dados foi dividido em treinamento (70%), validação (10%) e teste (20%), foram escolhidas imagens igualmente aleatórias dentre as doenças do *dataset* para poder treinar a rede neural. Após rodar por um certo período de épocas foi selecionada as 32 principais amostras e foi utilizada com a CAGAN para poder enfim gerar apresentações explicativas.

Durante a aplicação do modelo do framework de treinamento, foi utilizado 4 modelos de emprego de aprendizado, sendo aplicados em figuras de doenças descritas na sessão do desenvolvimento que foi a base para o *dataset* deste documento. Sendo o CAGAN superior ao GAN, por conseguir gerar uma maior diversidade entre as amostras de classificação, a rede neural *Bayesian* tratada e dita como modelo de probabilidade de melhora, aplicada e com uma melhor relação à certeza das imagens. O treinamento com estas enfermidades foram uma das formas apresentadas como padrão, onde não se é possível notar diferenças entre as imagens reais e as figuras sintéticas. A partir dessas imagens é possível utilizá-las em um novo treinamento para alguma finalidade ou estudos futuros.

As imagens retratadas pela experimentação para uma melhor validação foram também passadas por exame visual de médicos clínicos, essa segmentação foi visualizada por 3

radiologistas com 15 anos de experiência cada, em que os especialistas eram designados à função de dar um diagnóstico de doença sobre a imagem tendo em mãos somente as figuras geradas. Foram entregues no total 90 imagens, em 60% foi possível um diagnóstico exato sobre a doença, as representações que não obtiveram uma análise precisa, os motivos para isso foi explicado pelos especialistas porque a nitidez das imagens geradas tornava difícil obter uma precisão no diagnóstico a partir de uma única visualização. Alguns dos médicos ainda citaram que a doença em algumas imagens eram muito sutil e que necessitam de uma imagem mais nítida para uma resposta mais conclusiva.

Visto que a partir de poucas amostras é possível através de um método proposto de aprendizado ativo gerar demonstrações comparativas com outros modelos, eliminando assim aos poucos o desbalanceamento de um repositório de imagens, onde caso após o balanceamento se necessário uma classificação de múltiplas classes é recomendado uso do modelo CAGAN.

Esta seção apresenta o que seria um melhoramento ao modelo do *WGAN* do estudo executado. Em que este modelo propõe o uso de *framework* de aprendizado ativo para que seja possível ter melhores resultados.

3.4 X2CT-GAN: Reconstructing CT from Biplanar

X-Rays with Generative Adversarial Networks

Com base no artigo Ying et al. (2019), a tomografia computadorizada é uma visualização de órgãos internos utilizada como alicerce para o texto em que facilita o diagnóstico de doenças, porém ocorre maiores doses de radiação no paciente do que em um raio-x. Em comparação aos raios-x as tomografias computadorizadas são mais detalhistas e são construídas em 3D enquanto o raios-x em 2D.

Neste estudo foi utilizado um *dataset* com dados reais e fictícios. No trabalho em específico houve tentativas exaustivas para utilizar a mesma dimensão da imagem, como, por exemplo, ao entrar imagens 2D ter como saídas 2D, porém quando era necessário fazer algum crescimento de nível da figura incorporando dois mecanismos para aumentar o comprimento de dados eram acionados, a arquitetura utilizada para essa geração é a X2CTGAN com camadas para redes 2D e 3D, como o modelo generativo das GANs gerando novas informações, com base no seu treinamento, sendo este supervisionado. O principal objetivo do trabalho é reconstruir tomografias computadorizadas de raios-x biplanares a partir de um conjunto de dados.

Este modelo é composto por 3 componentes individuais dois codificadores e um decodificador, todos com a mesma arquitetura, trabalhando em paralelo para posteriormente

gerar uma radiografia computadorizada melhorada através de uma rede de fusão entre os codificadores e decodificadores. A área dos codificadores busca aprender o mapeamento dos raios-x 2D com entrada em 3D ou vice e versa utilizando-se de técnicas como *Densely Connected Encoder* um *encoder* de conectividade densa para poder gerar os dados de forma conexa, o *Bridging 2D Encoder and 3D Decoder* é uma técnica em que se mistura codificador e decodificador através de convolução e por último o *Feature Fusion of Biplanar X-rays* é feito um codificador biplanar para conseguir todas as dimensões da imagem passada seja ela 2D ou 3D.

As métricas utilizadas para a medição de qualidade foram, PSNR frequentemente utilizado para medir o atributo dos sinais digitais reconstruídos, o que torna o PSNR o mensurador de características das imagens, além do SSIM que é uma métrica utilizada para aferir a semelhança de duas imagens, incluindo brilho, contraste e estrutura. Comparado ao PSNR, ao SSIM, o SSIM pode corresponder a uma avaliação subjetiva humana melhorada.

O objetivo final do trabalho é a reconstrução de uma tomografia computadorizada, porém ao ser utilizar no mundo real mesmo tendo sido treinado com imagens falsas, pôde-se notar pulmões e uma estrutura de superfície plausível, a Figura 6 demonstra os resultados supracitados.

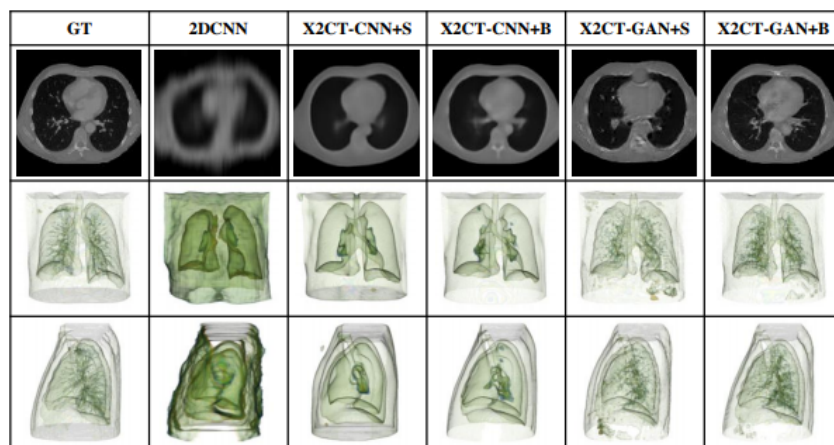


Figura 6 – Presente no artigo Ying et al. (2019), Estes foram os resultados obtidos após a aplicação do modelo X2CTGAN com as suas camadas.

Em futuros trabalhos essa GAN pretende ajudar médicos a avaliar tomografias computadorizadas e com isso medir tamanho de órgãos ou aplicar dosagem específica de remédios a partir de uma possível doença com essa melhor visualização. Assim como os nossos modelos generativos, este artigo cita uma nova forma de se implementar a *GAN*, com segmentações e melhores formulas de classificação, o qual pode ser utilizado junto com os modelos no estado da arte para obter melhores resultados.

4 Desenvolvimento

Durante a criação de um repositório de imagens, podem ocorrer diversas situações inesperadas como o caso de algumas enfermidades raras ou de difícil classificação, ou ainda durante o início dos casos recorrentes da doença. Havendo poucas informações ou poucas amostras tornando o estudo proveniente desses dados inviável.

O atual caso do coronavírus tem a possibilidade em ser utilizado para estudo com uma inteligência artificial para conseguir detectar os casos quando a quantidade de pacientes disponíveis para a pesquisa foi aceitável como mostrado no *TheNextWeb*¹. Pela concentração obtém-se uma maior quantidade de imagens em um *dataset*, ao qual precisa ser considerável para poder ter um estudo e servir de foco para pesquisas e afins.

É possível estudar e treinar CNNs com imagens de casos de doenças. Uma detecção astuta pode auxiliar e controlar o contágio da doença, porém existe a grande problemática de poucos casos que tenham feito diagnósticos tornando o número de amostras precário. Tornando este um desafio para a área médica (ABBAS; ABDELSAMEA; GABER, 2020).

Segundo Narin, Kaya e Pamuk (2020) casos onde pacientes com determinadas doenças não são identificados rapidamente e receberem o tratamento correto, podem infectar outros disseminando o vírus na população. Criaram uma CNN com o propósito de detectar o coronavírus com base em algumas redes convolucionais já existentes. Este trabalho demonstra que estes casos foram utilizadas nestas redes em conjunto com transformações em imagens, afim de viabilizar a geração de mais amostras para ocorrências de detecção e estudo de doenças.

Para conceber imagens sintéticas, dividimos a abordagem do trabalho proposto em 4 etapas. A primeira etapa consiste em preparar o dataset, o qual a partir do catálogo de doenças, foram separadas em pastas. A segunda etapa consiste no tratamento das imagens utilizando *data augmentation* com o objetivo de expandir o *dataset* escolhido. Na terceira etapa é a realizado o treinamento da rede neural utilizando dois modelos geradores: WGAN e DCGAN. A quarta e última etapa consiste em realizar a comparação entre esses modelos utilizando o *FID*.

¹ <https://thenextweb.com/neural/2020/03/02/alibabas-new-ai-system-can-detect-coronavirus-in-seconds-with-96-accuracy/>

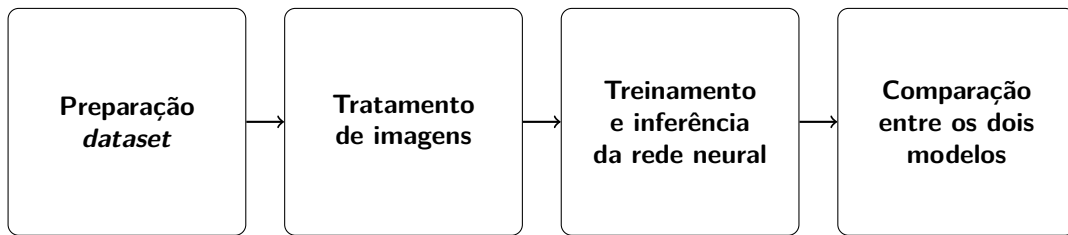


Figura 7 – Fluxo de etapas para comparação entre os resultados dos modelos.

4.1 Preparação do *dataset*

Inicialmente foi definido um índice em um arquivo *Comma-separated values (CSV)* ou arquivos separados por ponto e vírgula, onde haveriam as doenças do *dataset*. A separação sendo atribuída cada doença pode ser estudada, foram separadas em uma pastas com imagens, somente definidas de acordo com seu rótulo proveniente do CSV disponibilizado pela *National Institutes of Health (NIH)* ² ao publicar o *dataset* escolhido como *ChestX-ray8* ³, disponível um documento contendo seus nomes e rótulos.

A partir dos índices definidos foi feito um *script* em *Python*⁴, consistindo em ler o arquivo e separar de acordo com a etiqueta definida as imagens de cada respectiva doença na sua pasta com o nome da mesma. Assim seguiu para todas as imagens, para separar o *dataset* em pastas, gerando índices numéricos para cada pasta de arquivos, onde foi observado a relação da arquitetura na organização das figuras contendo apenas 123 fotos. Caso um treinamento fosse feito daquela pasta poderia não gerar amostras válidas, pois se tratava de uma doença rara de uma complicação no pulmão, chamada *Pneumonias eosinofílicas*, então a mesma foi escolhida como *dataset* para iniciar o gerenciamento e tratamento das imagens.

4.2 Tratamento de imagens

Para o tratamento, a primeira etapa foi o redimensionamento das imagens do *dataset*, sendo utilizado uma interface dentro da biblioteca do *OpenCV*, denominada *cv2*, para ler e redimensionar a imagem, pelos métodos *imread* e o *resize*, onde respectivamente foi necessário informar o caminho da mesma e informar a dimensão desejada para modificá-la. Após o procedimento, o tamanho da imagem foi alterado para 128 x 128 *pixels*, sendo este tamanho determinado por limitações no hardware utilizado, sendo ainda passível utilizar uma prova de conceito para determinar o valor otimizado em *pixels*. Tais representações anteriormente eram de resolução 1024 x 1024, isso foi feito para viabilizar o

² <https://www.nih.gov/>

³ <https://nihcc.app.box.com/v/ChestXray-NIHCC/file/256057377774>

⁴ <https://www.python.org/>

processamento, pois caso as figuras fossem grandes demais, para o desenvolvimento seria necessário um alto desempenho e hardware dedicado a isso no uso e aplicabilidade da metodologia estabelecida. Este redimensionamento teve como finalidade validar o conceito imposto para obter melhores variações pelos modelos generativos. Foi necessário aplicar algoritmos de aumento de imagens para ampliar o *dataset*.

4.2.1 Augmentation

Após o redimensionamento o próximo uso seria do *Augmentation* ou *Data augmentation* que a tradução significa aumento de dados Oliveira (2019). A técnica é utilizada para obter mais figuras para o *dataset*, disponível em uma biblioteca denominada *OpenCV*⁵. Utilizada com pequenas modificações nas imagens, aumentando a quantidade de fotos. Todos os passos, gerando novas imagens e assim a cada passo aplicando uma nova etapa a todas as amostras.

A inversão das imagens, virando a imagem do lado oposto. A segunda parte consiste em pequenas rotações em angulações entre 5 e 25° graus. Na terceira parte foi utilizado uma distorção, considerando que a imagem fosse uma matriz com retângulos horizontais e verticais, seguido de um valor definindo o tamanho da distorção aplicada no objeto de valor 5 em uma escala de 1 a 10, para que tivessem modificações variadas diferenciando as mesmas. Durante a quarta parte foi aplicado uma ampliação aleatória nas imagens, gerando amostras ligeiramente mais aproximadas. Por último no quinto passo foi aplicado contraste e brilho para modificar as imagens levemente.

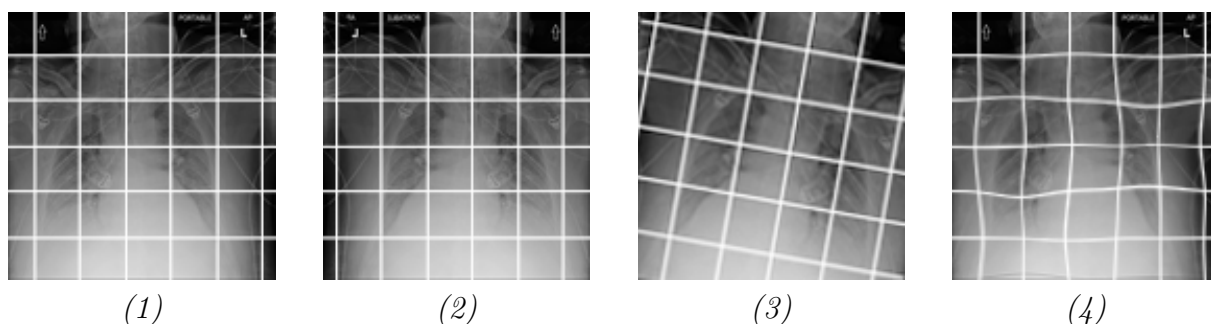


Figura 8 – Imagens numeradas e seu efeito ou stato respectivo: (1) original sem alterações, (2) invertida horizontalmente, (3) levemente rotacionada e (4) distorções leves

Construindo assim um *dataset* mais robusto e variado da doença, para na próxima etapa, a qual se faz necessário um certo volume de imagens para que uma rede neural possa aprender e ter bons resultados e criar boas amostras sintéticas.

⁵ <https://opencv.org/>

4.3 Treinamento e inferência da rede neural

Para prever padrões durante uma classificação de dados em um treinamento de rede neural é necessário um nível de aprendizado já realizado. Essa predição que é feita em duas etapas para aumentar o nível da sua confiabilidade, utilizando diferentes métricas como mostrado por Mallmann et al. (2019), sendo estas a acurácia e a taxa de categorização. Esses valores são retirados da *Confusion Matrix*, que é uma nova técnica de seleção de recursos como introduzido por Visa et al. (2011), comparando um conjunto de dados reais, a outros com base nos critérios citados acima. Possibilitando assim a comparação entre as formas de inferências diferentes de soluções para problemas semelhantes.

As redes neurais anteriormente citadas, após o período de treinamento, conseguem inferir um novo conjunto de dados geralmente com uma qualidade aceitável. Dessa forma, se pode gerar um conjunto de dados maior, com as imagens originais do *dataset* e as figuras geradas pelas *GANs*.

O *Google Colab*⁶ foi utilizado e este consiste em uma plataforma *online* e gratuita que hospeda um ambiente computacional, o Jupyter notebook, sendo utilizado para execução dos códigos em python que é uma opção que torna bibliotecas acessíveis, como, por exemplo, *Tensor Flow*, *PyTorch* e *Keras*.

Os modelos DCGAN e WGAN utilizados, foram provenientes de repositórios do *github*^{7,8}, possuindo estas estruturas de padrões generativos, possibilitando a execução dos módulos para efetuar o treinamento das redes com pesos de inicialização aleatórios, utilizando o otimizador *Adam*. Durante o treinamento foi necessário definir um tamanho ou número de lote adotado a partir de exemplos utilizados durante cada treino de iterações, foi por padrão 64. Sendo ajustado experimentalmente entre valores iniciais e sendo fixados em 100, este foi um valor suficiente de épocas para atingir um resultado considerável em ambas dos *GANs*, após muitas épocas uma *GAN* pode começar a replicar ou não assimilar mais conhecimento e acaba memorizando não tendo nenhum avanço.

As imagens geradas foram utilizadas para montar o *dataset* utilizado, onde cada modelo gerou uma pasta com uma quantidade de figuras equivalentes à quantias das pastas das amostras originais para que pudessem corresponder ao teste.

No Capítulo 5 será elucidado a comparação entre os dois modelos, e também uma analogia entre as amostras geradas e as reais. Contando com experimentos realizados, possibilitando a geração de uma medida para a distância entre pontos ou características das imagens de ambos modelos, quantificando seu desempenho visualmente por uma percepção humana.

⁶ <https://colab.research.google.com/>

⁷ <https://github.com/ReidWilliams/GANs>

⁸ <https://github.com/sinanatra/DCGAN-Art-Tensorflow>

5 Experimentos e análises

Neste capítulo serão apresentadas amostras obtidas e para as comparações de tal amostra será utilizada a distância de Fréchet, diferenças e peculiaridade entre os resultados das GANs e entre cada um dos modelos generativos, como também as imagens originais. Além de demonstrar um experimento sob uma ótica humana com 81 indivíduos sobre a classificação para verificação das imagens.

Na Seção 5.1 serão mostradas análises das imagens, comparações entre os modelos generativos e a distância de Fréchet entre eles. Já na Seção 5.2 será feita uma comparação com as amostras geradas pelos modelos com as reais, para comparar e validar suas características visuais. Por fim a Seção 5.3 discorrerá sobre um experimento realizado com indivíduos, para validar as imagens de acordo com a percepção de quais mais se aproximam das amostras reais.

5.1 DCGAN versus WGAN

Os modelos generativos treinados produziram imagens realísticas e variadas, conseguiram reproduzir características de raio-x pulmonares. O DCGAN conseguiu reconhecer um pulmão e reproduzir características próximas às figuras, porém embaçadas.

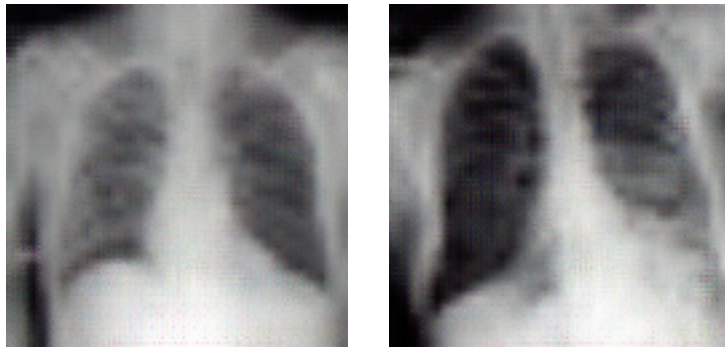


Figura 9 – Imagens geradas pelo modelo DCGAN

Apesar de aprender bem as partes importantes das estruturas torácicas vistas em um raio-x, O DCGAN não foi capaz de gerar imagens com nitidez. Mediante as imagens, conseguimos captar a estrutura base de um pulmão, porém em métrica quantitativa foi utilizado a distância de Fréchet para validar a mesma.

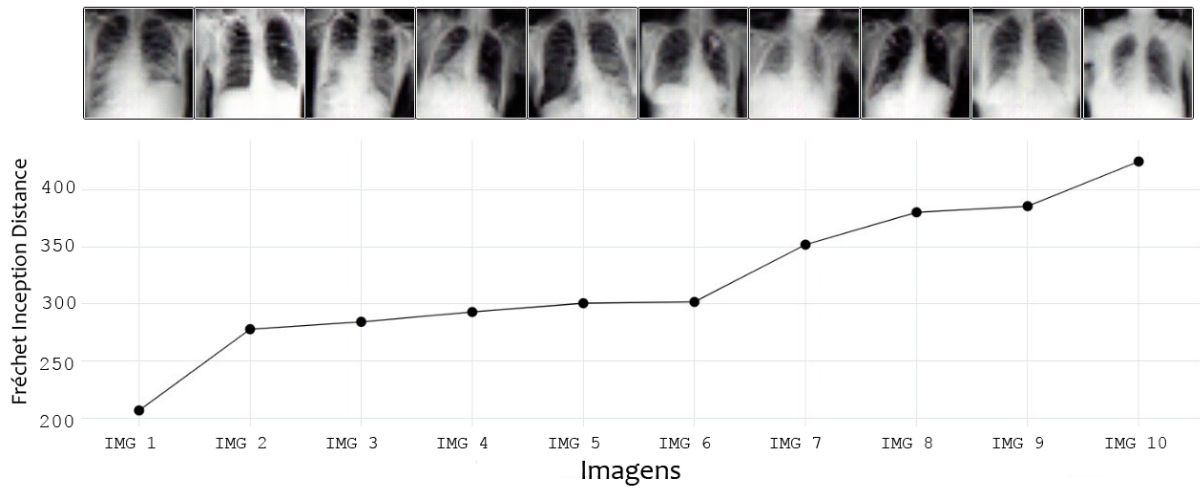


Figura 10 – Gráfico que demonstra a distância de Fréchet, que validam as imagens geradas pelo DCGAN, em comparação com as reais. No eixo y o maior valor é o 400

O WGAN também conseguiu reconhecer o padrão das imagens e reproduzir características próximas às figuras, porém com um pouco mais de qualidade que as do DCGAN e com uma leve textura como se estivessem levemente tremidas.

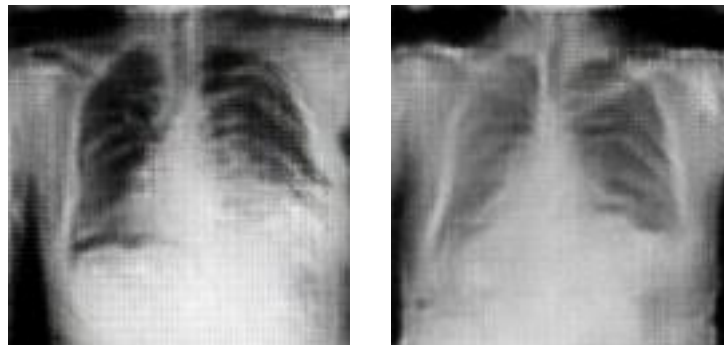


Figura 11 – Imagens geradas pelo modelo WGAN

Este modelo conseguiu se aproximar em relação à nitidez da figura, se comparado ao DCGAN, porém algumas informações como costelas e afins não ficaram muito nítidas. Também para validar as todas as amostras, foi utilizado o Fréchet.

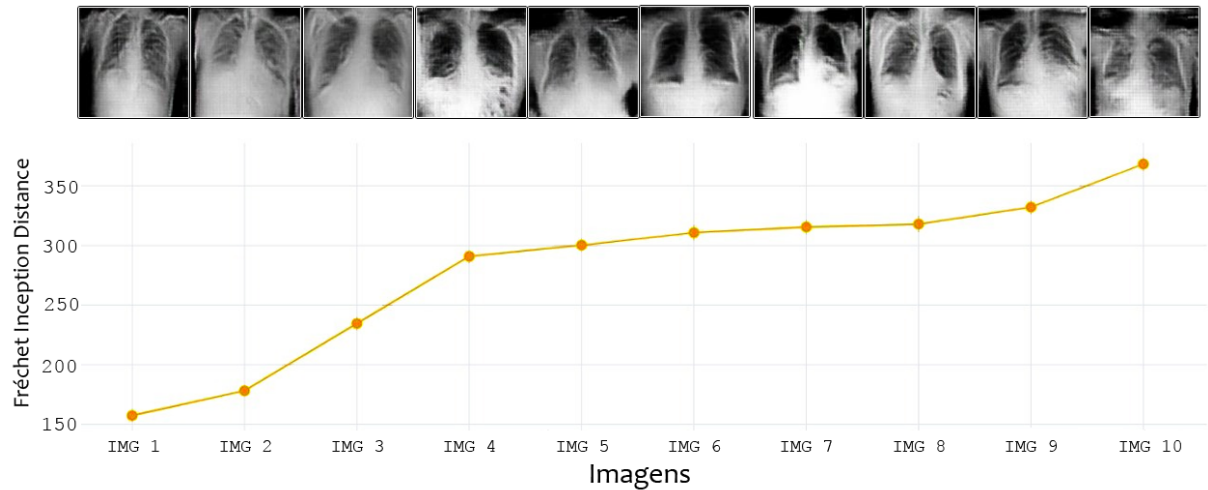


Figura 12 – Gráfico que demonstra a distância de Fréchet, que validam as imagens geradas pelo WGAN em comparação com as reais. No eixo y o maior valor é o 350

Ambos modelos sofreram um pouco pela baixa qualidade, o tamanho da imagem e o poder de processamento do ambiente. Algumas características foram notadas e assemelhadas entre os modelos, como ossos da costela e particularidade da infiltração do parênquima pulmonar.

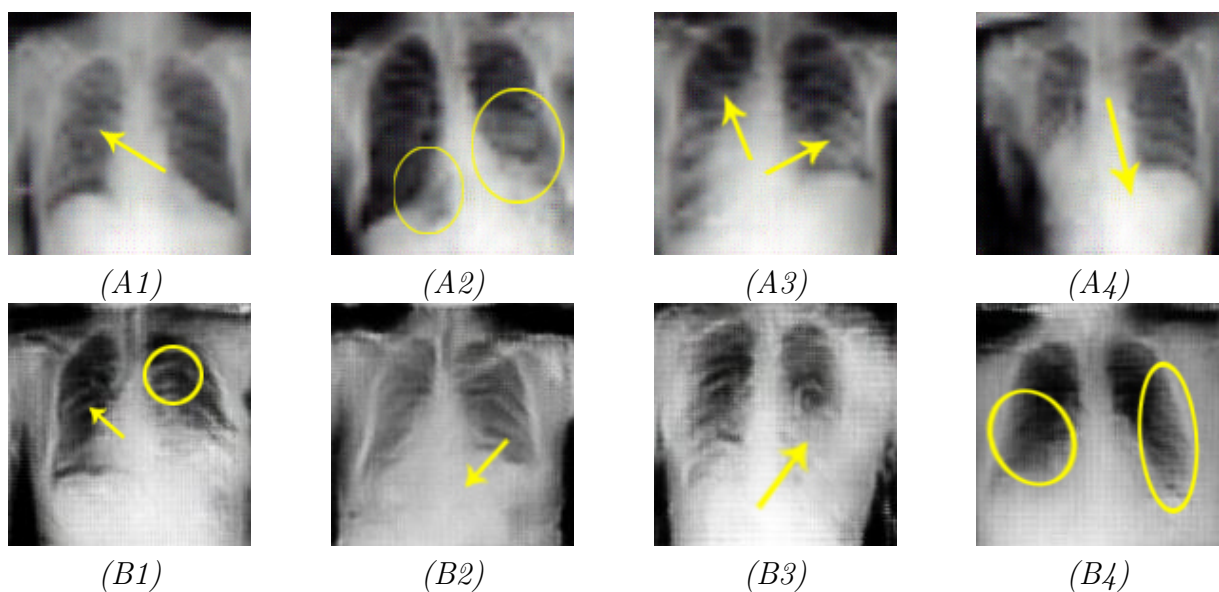


Figura 13 – As imagens mostram características apontadas em amarelo, obtidas a partir de análises visuais das GANs. As imagens da primeira fileira, mostram apontam informações de (A1) e (B1) costelas, (A2) e (B2) a área mais clara inferior do parênquima. (A3) e (B3) a região que se aproxima um pouco dos segmentos dos brônquios e possíveis infiltrações do parênquima. (A4) e (B4) as regiões que demonstram os cantos com locais que mostram infiltração do parênquima.

Características como formatos que estão presentes em raio-x nas amostras geradas que definem uma melhor qualidade das imagens e pelas análises visuais de conseguir extrair

mais propriedades e pela distância de Fréchet demonstrada em gráfico o WGAN teve um desempenho maior que o DCGAN.

A necessidade de aplicar e fazer comparações visuais com ambos dos modelos se faz presente no dado momento que possuímos imagens que podem ser comparadas com as reais e uma métrica como a distância de Fréchet entre elas. Na seção a seguir serão feitas análises visuais comparando com os resultados da figura anterior das GANs assemelhando com as reais.

5.2 GANs versus real

Assim como representado na Figura 14, os modelos generativos durante o seu treinamento reconhece as partes mais escuras, além das outras partes, no entanto, estas partes mais escuras são detectados como algo a ser demonstrado nas imagens geradas.

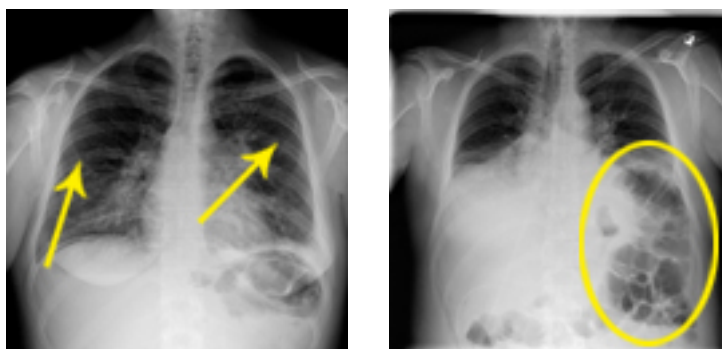


Figura 14 – Características apontadas em amarelo, nas imagens reais apontando estruturas que compõem as figuras, como, ossos da costela e fluidos do pulmão

As imagens presentes na Figura 15, são sínteses do modelo *DCGAN*. E é possível identificar visualmente alguns aspectos do pulmão, que seriam as partes mais escuras supracitadas, no entanto, outras determinadas características estão levemente embaçadas ou distorcidas.

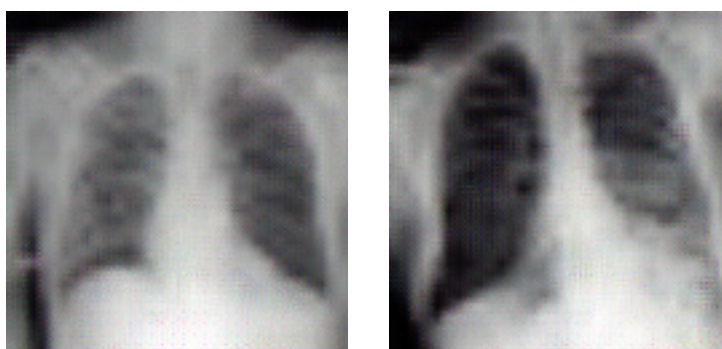


Figura 15 – Imagens geradas pelo modelo DCGAN

As imagens da Figura 16, são as representações do modelo *WGAN*. Estas conseguem representar com mais nitidez as partes escuras das figuras originais além de melhor representar as partes do objeto estudado.

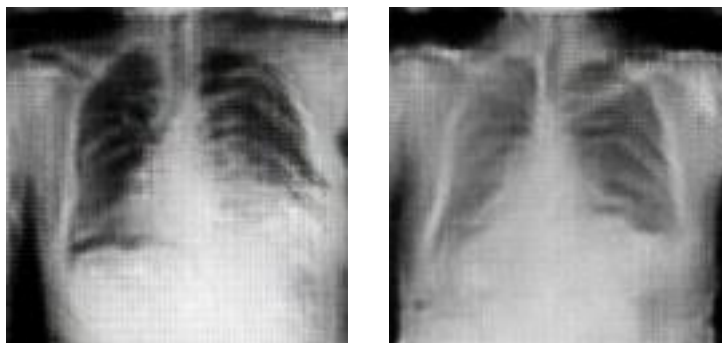


Figura 16 – Imagens geradas pelo modelo *WGAN*

As peculiaridades que podem ser reconhecidas nas imagens reais são informações relevantes e pontos de marco para as GANs.

A Figura 17 representam as amostras do *dataset* original, assim como as amostras geradas pelo *WGAN* e *DCGAN*, entre elas podemos perceber de forma clara a área do pulmão, costelas, região inferior do parênquima, que seriam as áreas mais escuras da imagem original.

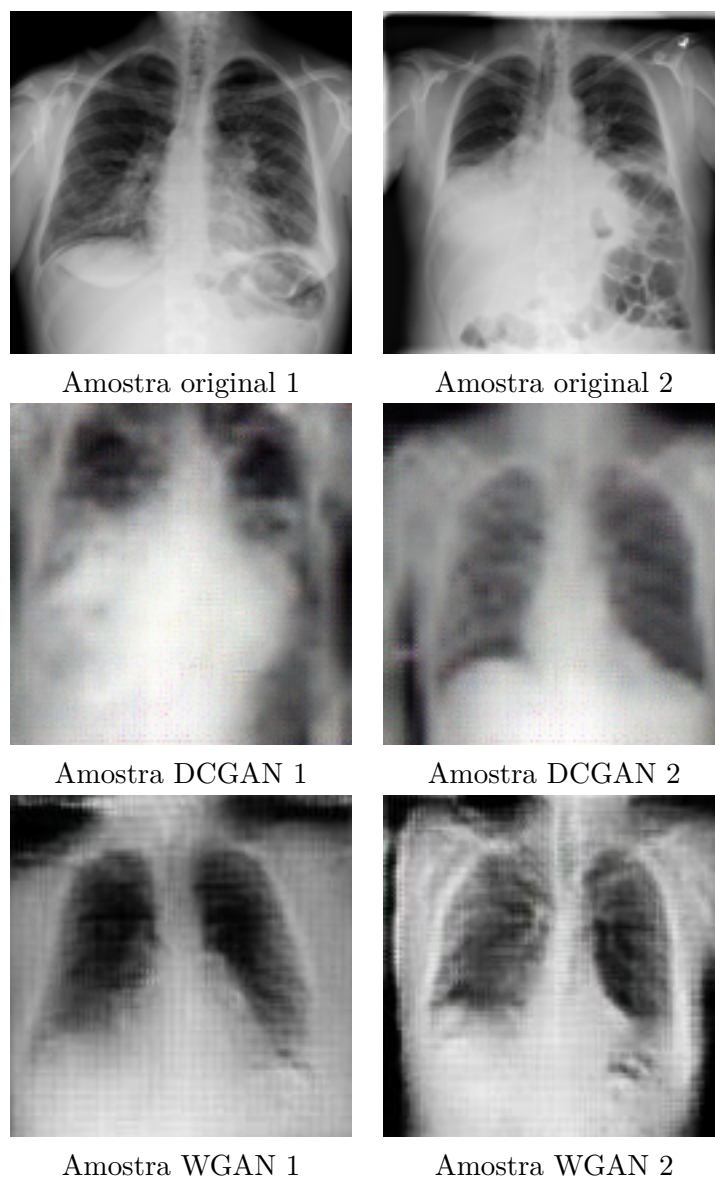


Figura 17 – Comparação das imagens entre *dataset original*, *DCGAN* e *WGAN*

5.3 Validações humanas

Esta seção demonstrará as hipóteses de um experimento e seus resultados sob a ótica e investigação humana sobre as amostras geradas por redes generativas, após o treinamento das redes propostas nesse estudo. A experimentação contou com 81 voluntário, dentre estes estudantes do ensino superior e profissionais de áreas diversas, ambos com idades variadas. Cada participante foi submetido a ambas das condições, totalizando em 236 imagens sintéticas.

A percepção visual do ser humano consegue distinguir a veracidade das informações de imagens, mesmo não sendo peritos no assunto abordado pelo contexto.

Os humanos conseguem avaliar o realismo visual conforme renderizações computadori-

zadas acontecem em sua tela, no entanto as formas de renderização atual buscam uma melhor validação parecida com a humana (FAN et al., 2017).

O objetivo era considerar quais imagens as pessoas iriam julgar como verdadeiras, em um período de três minutos, onde a plataforma mostrava aleatoriamente duas figuras, sendo possível estas serem de ambos os modelos *DCGAN* ou *WGAN*, ou apenas de um mesmo modelo, e quando fossem julgadas como real o usuário marcava a imagem. Após a classificação das imagens anteriores ou ao final do tempo o usuário recebia uma mensagem de agradecimento, após a distribuição do experimento foi possível analisar os dados e determinar uma porcentagem de escolhas sobre as imagens geradas de cada modelo.

5.3.1 Hipóteses

Antes do experimento foram levantadas as hipóteses H_0 e H_1 , dado o modelo das redes neurais era esperado um desempenho levemente maior dentre os modelos durante os resultados do experimento com a percepção humana de indivíduos selecionados, onde:

Hipótese 0 (H0) : *As amostras geradas pelo DCGAN terem uma taxa de acerto menor considerando como falsas, concluindo que podem ser melhores que as do modelo WGAN.*

Hipótese 1 (H1) : *As amostras do WGAN são as que possuem resultados superiores principalmente pelo seu período de treinamento ser mais estável, ao terem uma porcentagem maior e serem mais realistas.*

Os critérios de julgamento utilizados foram que estatísticas relacionadas aos modelos, pois possuíam características semelhantes considerandos as amostras. No modelo WGAN, considerando as suas amostras como as genuínas com média de 39,55% de acerto, enquanto na arquitetura DCGAN seriam não originais com o percentual de 63.35% de acerto, a partir das pessoas do experimento.

5.4 Resultados

Esta seção é onde nos iremos demonstrar graficamente os dados relatados e apurados pelo experimento de percepção humana, onde pessoas julgaram imagens pela sua detecção no que conheciam sobre figuras de raio-x, escolhendo qual imagem seriam reais, sem saber que eram representações sortidas das GANs.

O experimento de validação de percepção humana com amostras sintéticas e reais realizadas por Wyzykowski (2020) em sua qualificação de doutorado, serviu de base para a condução da experimentação presente neste documento. A utilização da forma de abordagem sobre o Figura 18 e Tabela 3, tiveram objetivos semelhantes em avaliar a influência humana sobre figuras sintéticas, trabalho este que culminou em Wyzykowski, Segundo e Lemes (2020).

Para demonstrar os percentuais dos participantes, na Figura 18, demonstra a escolha considerada verdadeira da pessoa em relação ao eixo x para a porcentagem do WGAN e o centro y o percentual do DCGAN. Cada círculo simboliza um participante e a localização do ponto demonstra uma área correspondente a seu resultado de classificação.

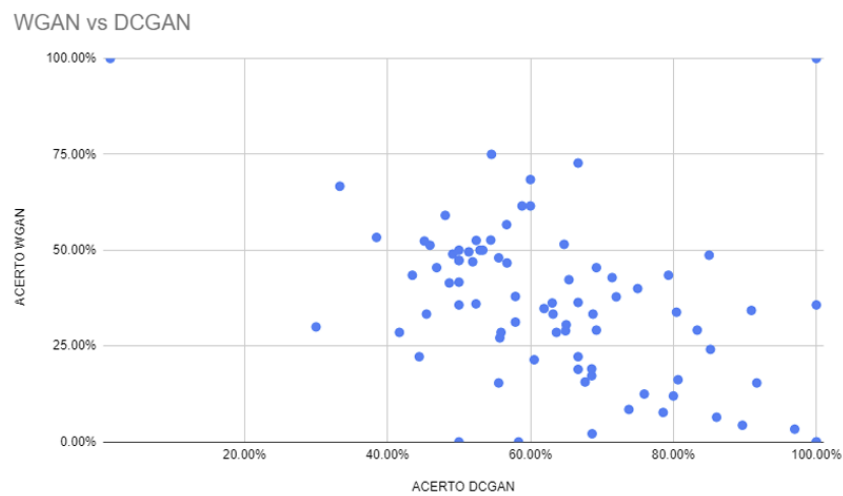


Figura 18 – Gráfico em pontos espalhados para exemplificar o resultado em percentual dos participantes do experimento, considerando o eixo x como acertos relativos ao experimento do DCGAN e o eixo y como acertos relativos ao experimento do WGAN.

Os resultados demonstrados que pessoas tiveram foram semelhantes, porém alguns demonstraram diferentes áreas de escolha descritos a seguir.

Nota-se que a região do círculo em vermelho na Figura 19 no gráfico demonstra a interseção entre as pessoas que tiveram mais dúvidas de qual das imagens demonstradas eram as mais reais pelos seus critérios.

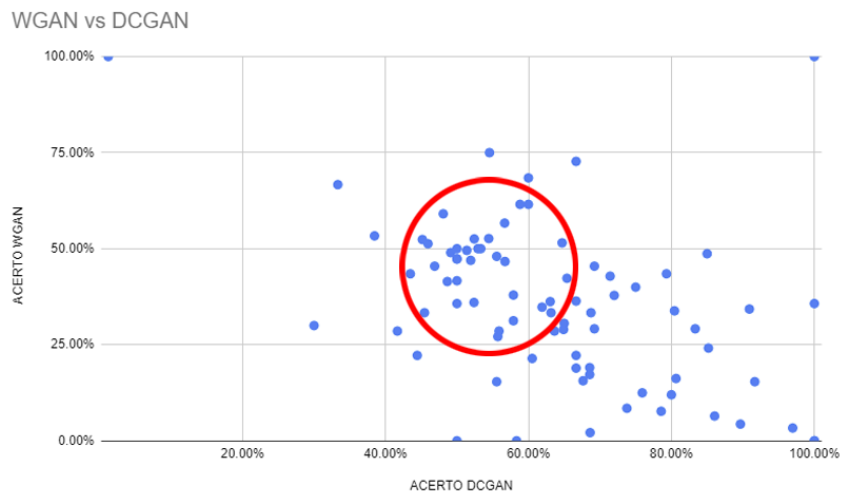


Figura 19 – Gráfico semelhante da Figura 18, porém o círculo em vermelho demonstra uma população do experimento de pessoas mais confusas sobre as amostras, as quais não ajudaram muito já que não conseguiram diferenciar às duas classes das imagens.

Assim como houve pessoas confusas tendo um percentual bem próximo de acerto em relação às amostras dos modelos, houve participantes que notaram uma diferença entre as imagens demonstrado na Figura 20 em prol a sua percepção de figuras de raios-x reais.

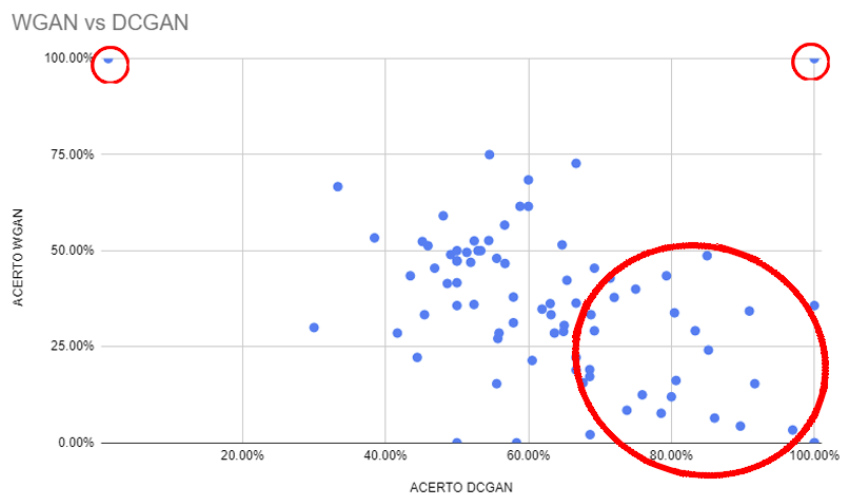


Figura 20 – Gráfico semelhante da Figura 18, porém o círculo em vermelho, desta vez mostra parte da população que notou diferença entre as classes das imagens, ou seja, pessoas que notaram diferenças perceptíveis entre figuras que relacionaram como verdadeiras pelo seu histórico, estas que no que lhe concerne conseguiram mostrar que uma classe de imagem é visivelmente mais realística que as outras.

Os resultados obtidos nas Figuras 18, 19, 20 e na Tabela 3, são representações que auxiliam a forma de visualizar as soluções do experimento, confirmando os dados das

gravuras com algumas informações quantitativas.

Indivíduo	DCGAN % Acerto	WGAN % Acerto	Imagens Visualizadas
1	52.94%	50.00%	78
2	100.00%	100.00%	2
3	85.19%	24.14%	56
4	66.67%	72.73%	26
5	53.33%	50.00%	62
6	46.88%	45.45%	54
7	78.57%	7.69%	54
8	83.33%	29.17%	42
9	51.92%	46.97%	118
10	48.08%	59.09%	118
11	66.67%	18.92%	88
12	55.71%	27.14%	140
13	96.97%	3.33%	126
14	51.38%	49.56%	222
15	43.48%	43.48%	46
16	58.33%	0.00%	22
17	44.44%	22.22%	18
18	91.67%	15.38%	50
19	50.00%	41.67%	66
20	66.67%	22.22%	24
21	45.95%	51.28%	76
22	89.66%	4.35%	52
23	55.88%	28.57%	76
24	41.67%	28.57%	26
25	71.43%	42.86%	28
26	54.44%	52.63%	166
27	68.57%	19.05%	56
28	63.64%	28.57%	18
29	60.53%	21.43%	66
30	79.31%	43.48%	52
31	54.55%	75.00%	46
32	64.71%	51.52%	50
33	73.77%	8.47%	120
34	55.56%	15.38%	62
35	85.00%	48.68%	136
36	50.00%	47.37%	70

37	90.91%	34.29%	68
38	38.46%	53.33%	28
39	65.38%	42.31%	52
40	57.89%	31.25%	70
41	75.00%	40.00%	22
42	57.89%	37.93%	48
43	86.05%	6.45%	74
44	45.45%	33.33%	46
45	75.93%	12.50%	102
46	69.23%	45.45%	24
47	80.43%	33.82%	114
48	100.00%	0.00%	2
49	60.00%	68.42%	34
50	100.00%	0.00%	58
51	65.00%	30.56%	56
52	80.00%	12.00%	40
53	52.38%	36.00%	46
54	72.00%	37.84%	62
55	56.67%	56.67%	60
56	30.00%	30.00%	20
57	68.63%	2.13%	98
58	58.82%	61.54%	30
59	67.65%	15.63%	66
60	50.00%	50.00%	4
61	68.75%	33.33%	80
62	56.72%	46.67%	142
63	48.65%	41.46%	78
64	33.33%	66.67%	6
65	64.91%	28.99%	126
66	100.00%	35.71%	20
67	66.67%	36.36%	80
68	60.00%	61.54%	46
69	50.00%	100.00%	6
70	45.16%	52.38%	52
71	68.57%	82.76%	64
72	63.16%	66.67%	34
73	69.23%	70.83%	50
74	75.00%	64.29%	22

75	61.90%	34.78%	88
76	50.00%	47.30%	138
77	52.42%	52.54%	242
78	55.56%	48.00%	52
79	49.12%	48.98%	106
80	1.22%	100.00%	154
81	80.65%	16.22%	68
Média	63.35%	39.55%	-
Total	-	-	5390

Tabela 3 – Tabela indicando quantitativamente os acertos para cada GAN relacionada ao experimento e a quantidade de imagens vistas, com a média das pessoas e a quantidade total de imagem que foram vistas.

Os valores apresentados no gráfico, são provenientes do experimento executado com pessoas para que fosse escolhida as opções que elas considerassem mais próximas da realidade, com isso foi possível validar a hipótese inicial levantada:

- **H1:** Comparando os resultados entre os valores gerados pelo DCGAN e WGAN, foi possível considerar determinadas diferenças a partir das respostas do experimento. Considerando as amostras do DCGAN como falsas obtendo assim, um alto índice em não serem escolhidas, critério este do experimento, que exibia as imagens para seleção como genuínas ou sintéticas. Em comparação com as amostras do WGAN, ou quando apareceram duas vezes imagens do mesmo modelo as pessoas selecionaram poucas como verdadeiras. O valor presente relacionado ao WGAN verdadeiras, foi baixo por parte da população do experimento.

Esta hipótese se confirma pela questão da porcentagem de pessoas que marcaram as imagens, pela sua percepção como DCGAN como verdadeira é menor que as que marcaram como o WGAN.

Considerando um bom e melhor resultado durante a curva demonstrada nos resultados do Fréchet, podemos considerar que em casos futuros experimentos sejam utilizados nos modelos, com um objetivo de ponto de partida, se terá um grau de classificação e geração por meio desses padrões.

6 Conclusões

As GANs são um tipo de assunto relativamente recente, porém bem estudado no ramo computacional, como mostrado neste documento onde comparamos os modelos DCGAN e WGAN, como julgar o melhor modelo para gerar novas imagens para fomentar um *dataset* de raios-x de região torácica.

Esse modelo de visão computacional vem sendo bastante estudado durante as últimas décadas, porque com base na forma de implementação de uma rede neural generativa é possível replicar diversas formas da atualidade como reconstrução facial que pode ser utilizado para identificação de rosto em crimes Hao et al. (2020), ou melhoramento de imagens para aumento de resolução, melhorando algumas figuras em baixa qualidade para melhorar o reconhecimento a olho humano Zeng e Zhang (2020).

Com os experimentos, foi possível notar que o WGAN por possuir uma camada extra de convolução acaba possuindo uma melhor resolução em relação ao DCGAN por isso este termina desfrutando de uma menor distância *Frechet*, além de uma maior estabilidade durante o treinamento Gulrajani et al. (2017).

Sendo assim o WGAN por ter uma menor distância *Frechet* foi mais assertivo no experimento com pessoas reais, havendo mais escolhas deste como próximo aos raio-x originais, além do DCGAN ser mais escolhido como uma imagem falsa (FATHI, 2019).

Como o principal objetivo era determinar, porque o WGAN seria melhor que o DCGAN, e este foi possível através da visão fornecida pelo experimento e da distância de *Frechet*, deliberando qual modelo generativo conseguiu um desempenho superior no critério de gerar amostras sintéticas. Ainda assim a diferença em relação ao DCGAN não foi muito grande, porém visualmente foi significativo a ponto de ter uma resolução e definição superior.

6.1 Trabalhos futuros

Para os próximos passos do trabalho assim como para incrementa-lo seria interessante os seguintes itens:

1. Adicionar novas métricas para melhor avaliação do trabalho final. Como por exemplo em Khrulkov e Oseledets (2018), aplicando o *Geometry Score* sendo essa uma forma de avaliar e comparar as propriedades geométricas de cada imagem gerada com as imagens originais. Utilizar também a *Discriminator Rejection Sampling*, introduzida por Azadi et al. (2018), sendo uma forma utilizar as imagens descarta-

das pelo discriminador para melhorar a qualidade dos geradores, reprocessando as imagens com as informações do discriminador.

2. Utilização do HDCGAN no estudo proposto, em contraposição as GANs abordadas. A GAN desenvolvida por Curtó, Zarza e Kim (2017) é demonstrado um desempenho considerável em relação síntese de rostos humanos, interpretando a diferença entre as dimensões da imagem algo a analisar e também atrelar a utilização de mais épocas de aprendizado.
3. Utilização do CAGAN no estudo proposto, em contraposição as GANs abordadas, possui uma implementação superior à implementação da divergência do WGAN com o gradiente de penalidade demonstrado por Bozorgtabar et al. (2019), pois se utiliza o mesmo como base o que melhoraria seu desempenho, concebendo um resultado melhor ao experimento e síntese de imagens.
4. Melhorar a forma em que foi aplicada o experimento para evitar ruídos relacionados a dispersão humana. A realização do experimento se conduziu de maneira virtual, inclusive o contato e explicação sobre o mesmo. Caso o experimento fosse realizado em um ambiente controlado, considerando a percepção também sobre as pessoas envolvidas, poderia ocasionar em resultados diferentes.
5. Aumento da resolução das imagens e aumento de épocas de treinamento, para obtenção de amostras mais realistas. Durante o experimento conduzido neste documento, é possível notar que em estudos como Curtó, Zarza e Kim (2017), que a resolução da imagem a quantidade de épocas empregadas no treinamento é um fator que influência no resultado, porém exige um grande poder de processamento. Dado a resolução utilizada no documento a possibilidade com mais processamento, maiores imagens e mais épocas poderiam impactar em resultados diferentes.
6. A utilização de um *Adaptive discriminator augmentation* por Karras et al. (2020), têm como objetivo de gerar imagens mais próximas da realidade a partir de um *dataset* escasso. Esta implementação diminuiu o *FID*, comparando as amostras sintetizadas com as reais. Este discriminador se fez necessário, por possuir uma maior assertividade em suas escolhas, com poucos dados precedentes ao treinamento das redes e então aplicamos ampliações versáteis que resultou em um novo modelo de discriminador.

Estes próximos passos são citados para fomento deste trabalho, para que pessoas interessadas pelo mesmo assunto possam ter resultados ainda melhores que os obtidos.

Referências Bibliográficas

- ABBAS, A.; ABDELSAMEA, M. M.; GABER, M. M. Classification of covid-19 in chest x-ray images using detrac deep convolutional neural network. *arXiv preprint arXiv:2003.13815*, 2020. 37
- AGUIAR, V. et al. Prevalência de sintomas, doenças pulmonares e tabagismo em populações universitárias vivendo em ambientes com níveis diferentes de poluição atmosférica. *J. pneumol*, n. 15, p. 61–68, 1989. 20
- ANAND, S.; NAGAJOTHI, K.; NITHYA, K. *Edge Detection using Stationary Wavelet Transform, HMM, and EM algorithm*. 2020. 23
- ANJOS, B. H. dos; JATOBÁ, A. E.; OLIVEIRA, M. C. Generative adversarial networks: a renewal for data augmentation in lung nodule classification. In: SBC. *Anais do XV Workshop de Visão Computacional*. [S.l.], 2019. p. 61–66. 18, 27
- ANNARUMMA, M. et al. Automated triaging of adult chest radiographs with deep artificial neural networks. *Radiology*, Radiological Society of North America, v. 291, n. 1, p. 196–202, 2019. 21
- ARJOVSKY, M.; CHINTALA, S.; BOTTOU, L. Wasserstein gan. *arXiv preprint arXiv:1701.07875*, 2017. 28
- ARONSON, A. R.; LANG, F.-M. An overview of metamap: historical perspective and recent advances. *Journal of the American Medical Informatics Association*, BMJ Group BMA House, Tavistock Square, London, WC1H 9JR, v. 17, n. 3, p. 229–236, 2010. 22
- AZADI, S. et al. *Discriminator Rejection Sampling*. 2018. 53
- BOZORGTABAR, B. et al. Informative sample generation using class aware generative adversarial networks for classification of chest xrays. *Computer vision and image understanding*, Elsevier, v. 184, p. 57–65, 2019. 11, 17, 33, 34, 54
- CANNY, J. A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence*, Ieee, n. 6, p. 679–698, 1986. 23
- CHEN, L. et al. Deep spectral-spatial feature extraction based on dcgan for hyperspectral image retrieval. In: IEEE. *2017 IEEE 15th Intl Conf on Dependable, Autonomic and Secure Computing, 15th Intl Conf on Pervasive Intelligence and Computing, 3rd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech)*. [S.l.], 2017. p. 752–759. 26
- CHENG, M. et al. *Data-driven modelling of nonlinear spatio-temporal fluid flows using a deep convolutional generative adversarial network*. 2020. 26
- COSTA, J. B. et al. Os principais fatores de risco da pneumonia associada à ventilação mecânica em uti adulta. 2016. 21

- CUNHA, C. S. et al. Atuação da fisioterapia na reversão das atelectasias: Um relato de caso na unidade de terapia intensiva the performance of the physiotherapy in the reversion of the atelectasis: A report of a case in the intensive therapy unit. *Cadernos UniFOA*, v. 2, n. 4, p. 81–87, 2017. 21
- CURTÓ, J. D.; ZARZA, H.; KIM, T. High-resolution deep convolutional generative adversarial networks. *arXiv preprint arXiv:1711.06491*, 2017. 54
- CURTÓ, J. D. et al. *High-resolution Deep Convolutional Generative Adversarial Networks*. 2017. 27
- DEEPLARNINGBOOK. *Capítulo 54 – Introdução às Redes Adversárias Generativas (GANs – Generative Adversarial Networks)*. 2015. Disponível em: <<http://deeplearningbook.com.br/introducao-as-redes-adversarias-generativas-gans-generative-adversarial-networks/>>. 24
- DEMO, F.; FILHO, W. P. Gestão de privacidade no armazenamento de dados do paciente em registros médico-hospitalares eletrônicos. In: SBC. *Anais da XVII Escola Regional de Redes de Computadores*. [S.l.], 2020. p. 8–12. 17
- DOAN, K. D. et al. Image generation via minimizing fréchet distance in discriminator feature space. *arXiv preprint arXiv:2003.11774*, 2020. 28, 29
- FAN, S. et al. Image visual realism: From human perception to machine computation. *IEEE transactions on pattern analysis and machine intelligence*, IEEE, v. 40, n. 9, p. 2180–2193, 2017. 47
- FATHI, M. *Optimization in Large Scale Problems: Industry 4.0 and Society 5.0 Applications*. [S.l.]: Springer Nature, 2019. 53
- FREITAS, A. A. R. R.; NAPIMOGA, M.; DONALISIO, M. R. Análise da gravidade da pandemia de Covid-19. *Epidemiologia e Serviço de Saúde*, scielo, v. 29, Abril 2020. ISSN 2237-9622. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S2237-96222020000200900&nrm=iso>. 18
- GANESAN, P. et al. Assessment of data augmentation strategies toward performance improvement of abnormality classification in chest radiographs. In: IEEE. *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. [S.l.], 2019. p. 841–844. 13, 30, 31
- GOODFELLOW, I. J. et al. *Generative Adversarial Networks*. 2014. 25
- GULRAJANI, I. et al. Improved training of wasserstein gans. In: *Advances in neural information processing systems*. [S.l.: s.n.], 2017. p. 5767–5777. 27, 28, 53
- HAO, H. et al. *FaR-GAN for One-Shot Face Reenactment*. 2020. 53
- HAR-PELED, S. Fréchet distance: How to walk your dog. *Geometric Approximation Algorithms*, 2014. 29
- HEUSEL, M. et al. *GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium*. 2017. 28, 29

- HORAK, D.; YU, S.; SALIMI-KHORSHIDI, G. *Topology Distance: A Topology-Based Approach For Evaluating Generative Adversarial Networks*. 2020. 29
- JR, A. S. S. et al. Terminologia para a descrição de tomografia computadorizada do tórax: sugestões iniciais para um consenso brasileiro. *Radiologia Brasileira*, SciELO Brasil, v. 35, n. 2, p. 125–128, 2002. 21
- JURASZEK, G. D. et al. Reconhecimento de produtos por imagem utilizando palavras visuais e redes neurais convolucionais. Universidade do Estado de Santa Catarina, 2014. 22
- KARRAS, T. et al. Training generative adversarial networks with limited data. *Research Nvidia*, June 2020. 54
- KHRULKOV, V.; OSELEDETS, I. *Geometry Score: A Method For Comparing Generative Adversarial Networks*. 2018. 53
- KOTTOW, M. H. Medical confidentiality: an intransigent and absolute obligation. *Journal of medical ethics*, Institute of Medical Ethics, v. 12, n. 3, 1986. 17
- LAB TESTS, O. Doenças pulmonares. *SBPC/ML | SOCIEDADE BRASILEIRA DE PATOLOGIA CLÍNICA/MEDICINA LABORATORIAL*, 2019. 20
- LEAMAN, R.; DOĞAN, R. I.; LU, Z. Dnorm: disease name normalization with pairwise learning to rank. *Bioinformatics*, Oxford University Press, v. 29, n. 22, p. 2909–2917, 2013. 22
- LECUN, Y.; KAVUKCUOGLU, K.; FARABET, C. Convolutional networks and applications in vision. In: IEEE. *Proceedings of 2010 IEEE international symposium on circuits and systems*. [S.l.], 2010. p. 253–256. 22
- LEE, M.; SEOK, J. *Score-Guided Generative Adversarial Networks*. 2020. 29
- MAGALHÃES, E.; TAVARES, B.; CHIEIRA, C. Pneumonias eosinofílicas. *Rev Port Imunoalergologia*, v. 14, n. 3, p. 196–217, 2006. 21
- MAGUOLO, G.; NANNI, L. A critic evaluation of methods for covid-19 automatic detection from x-ray images. *arXiv preprint arXiv:2004.12823*, 2020. 17
- MALLMANN, J. et al. Mitigando os efeitos de gan em classificação de imagens com cnn. 2019. 26, 40
- MALLOZI, M. C.; ROZOV, T. O laboratório nas doenças pulmonares. *J pediatr (Rio J.)*, v. 74, p. S125–S132, 1998. 20
- METZ, L. et al. *Unrolled Generative Adversarial Networks*. 2016. 18
- MOSMANN, M. P. et al. Solitary pulmonary nodule and 18f-fdg pet/ct. part 1: epidemiology, morphological evaluation and cancer probability. *Radiologia brasileira*, SciELO Brasil, v. 49, n. 1, p. 35–42, 2016. 21
- NARIN, A.; KAYA, C.; PAMUK, Z. Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. *arXiv preprint arXiv:2003.10849*, 2020. 37

- NOPPEN, M.; KEUKELEIRE, T. D. Pneumothorax. *Respiration*, Karger Publishers, v. 76, n. 2, p. 121–127, 2008. 21
- OLIVEIRA, W. D. G. d. Data augmentation via generative adversarial networks aplicado em classificação de imagens. Universidade Federal de São Paulo, 2019. 39
- PEREIRA, I. V. d. S. T. Uma representação geral da fala para reconhecimento de emoção com uma rede semi-supervisionada utilizando gan. Escola Politécnica de Pernambuco, 2018. 24
- RADFORD, A.; METZ, L.; CHINTALA, S. *Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks*. 2015. 26
- SALEHINEJAD, H. et al. *Generalization of Deep Neural Networks for Chest Pathology Classification in X-Rays Using Generative Adversarial Networks*. 2017. 13, 31, 32
- SANKAR, P. et al. Patient perspectives of medical confidentiality: a review of the literature. *Journal of general internal medicine*, Wiley Online Library, v. 18, n. 8, 2003. 17
- SANTOS, R. R. d. Detecção de ataques de apresentação facial utilizando redes neurais convolucionais e informações de contexto. 2019. 23
- SCHAWINSKI, K. et al. Generative adversarial networks recover features in astrophysical images of galaxies beyond the deconvolution limit. *Monthly Notices of the Royal Astronomical Society: Letters*, Oxford University Press, v. 467, n. 1, p. L110–L114, 2017. 25
- SEISCENTO, M.; CONDE, M. B.; DALCOLMO, M. M. P. Tuberculose pleural. *Jornal Brasileiro de Pneumologia*, SciELO Brasil, v. 32, p. S174–S181, 2006. 21
- SILVA, J. D. D. d. et al. Repercussões da cardiomegalia na função pulmonar de indivíduos adultos com insuficiência cardíaca crônica: uma revisão sistemática. *Fisioterapia e Pesquisa*, SciELO Brasil, v. 18, n. 1, p. 84–91, 2011. 21
- SUAREZ, P. L.; SAPP, A. D.; VINTIMILLA, B. X. Infrared image colorization based on a triplet dcgan architecture. In: *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*. [S.l.: s.n.], 2017. 26
- TARTAGLIONE, E. et al. Unveiling covid-19 from chest x-ray with deep learning: a hurdles race with small data. *arXiv preprint arXiv:2004.05405*, 2020. 17
- VARGAS, A. C. G.; PAES, A.; VASCONCELOS, C. N. Um estudo sobre redes neurais convolucionais e sua aplicação em detecção de pedestres. In: *Proceedings of the XXIX Conference on Graphics, Patterns and Images*. [S.l.: s.n.], 2016. v. 1, n. 4. 18, 23
- VISA, S. et al. Confusion matrix-based feature selection. *MAICS*, v. 710, p. 120–127, 2011. 40
- WACHTER, S.; MITTELSTADT, B.; RUSSELL, C. Counterfactual explanations without opening the black box: Automated decisions and the gdpr. *Harv. JL & Tech.*, HeinOnline, v. 31, p. 841, 2017. 17

- WANG, X. et al. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2017. p. 2097–2106. 21, 22
- WENG, L. From gan to wgan. *arXiv preprint arXiv:1904.08994*, 2019. 24
- WYZYKOWSKI, A. B. V. Level three synthetic fingerprint generator. *Universidade Federal da Bahia Instituto de Matemática*, Janeiro 2020. Qualificação de Doutorado. 48
- WYZYKOWSKI, A. B. V.; SEGUNDO, M. P.; LEMES, R. d. P. Level three synthetic fingerprint generation. *arXiv preprint arXiv:2002.03809*, 2020. 48
- YING, X. et al. *X2CT-GAN: Reconstructing CT from Biplanar X-Rays with Generative Adversarial Networks*. 2019. 11, 35, 36
- ZANINOVICH, Y. Mamografia classi fi cação usando convolucionais redes neurais. 2019. 23
- ZENG, S.; ZHANG, B. *Noise Homogenization via Multi-Channel Wavelet Filtering for High-Fidelity Sample Generation in GANs*. 2020. 53
- ZHU, J.-Y. et al. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: *Computer Vision (ICCV), 2017 IEEE International Conference on*. [S.l.: s.n.], 2017. 25
- ZONZIN, G. A. et al. O que é importante para o diagnóstico da dpoc? *DIRETORIA DA SOPTEJ-BIÊNIO 2015/2017*, v. 26, n. 1, p. 5–14, 2017. 20