

Antônio Moreira Pinto

Detecção de Patologias Pulmonares por Imagem usando Deep Learning

São Luís - MA

Setembro - 2021

Antônio Moreira Pinto

Detecção de Patologias Pulmonares por Imagem usando Deep Learning

Monografia apresentada ao curso de Ciência da Computação da Universidade Federal do Maranhão como parte dos requisitos necessários para obtenção do grau de bacharel em Ciência da Computação.

Curso de Ciência da Computação

UFMA

Orientador: Prof. Dr. Geraldo Braz Júnior

São Luís - MA

Setembro - 2021

Ficha gerada por meio do SIGAA/Biblioteca com dados fornecidos pelo(a) autor(a).
Diretoria Integrada de Bibliotecas/UFMA

Moreira Pinto, Antônio.

Detecção de Patologias Pulmonares por Imagem usando
Deep Learning / Antônio Moreira Pinto. - 2021.

40 f.

Orientador(a): Geraldo Braz Junior.

Monografia (Graduação) - Curso de Ciência da
Computação, Universidade Federal do Maranhão, São Luís,
2021.

1. Aprendizagem de máquina. 2. Classificação. 3.
Processamento de imagem. 4. Regressão. I. Braz Junior,
Geraldo. II. Título.

Antônio Moreira Pinto

Detecção de Patologias Pulmonares por Imagem usando Deep Learning

Monografia apresentada ao curso de Ciência da Computação da Universidade Federal do Maranhão como parte dos requisitos necessários para obtenção do grau de bacharel em Ciência da Computação.

Aprovado em: 22 de Setembro de 2021, São Luís - MA

Prof. Dr. Geraldo Braz Júnior
Orientador
Universidade Federal do Maranhão

Prof. 1
Examinador
Universidade Federal do Maranhão

Prof. 2
Examinador
Universidade Federal do Maranhão

São Luís - MA
Setembro - 2021

Agradecimentos

À minha família, por todo o incentivo. Sem eles não estaria graduando.

Ao meu orientador, pelos aprendizados enquanto membro do PETComp e além. Sem seu incentivo não teria me desenvolvido tanto, enquanto acadêmico ou pessoa. Acredito que Geraldo fez bastante diferença não só na minha vida, quanto na vida de várias outras pessoas. Graças à ele eu entendo melhor o que gosto de fazer, algo que para mim é mais importante que escrever artigos.

Ao PETComp. Durante o programa mantive meu primeiro contato com o meio acadêmico e passei a ver propósito nos projetos que desenvolvia, agindo em uma equipe de pessoas que manteve muito apreço entre si. Se eu não tivesse conhecido meus amigos do programa não teria sido capaz de ter me tornar uma pessoa tão aberta ou feliz quanto eu sou. Jogar baralho entre as aulas e ir com aquele grupo enorme ao RU são memórias que guardo com muito carinho.

À Emily, Fernando, Durand, Giovanna, Hannah, Alana, Igor, Gabriel, Pedro, Letícia, Juliane e todos os amigos que tive durante os últimos anos. Os momentos que pude compartilhar com eles foram muito importantes. As conversas, os rolês, estudar às pressas antes de uma prova, ajudar com trabalhos, artigos, entregas de relatórios, eventos, pizzas depois dos eventos. São pessoas me fizeram bastante por mim e me deram força para alcançar os objetivos que conquistei.

Ao meu orientador de PIBIC, João Dallyson, pela oportunidade de atuar como estudante bolsista. Foi uma experiência que me ajudou bastante, ocupando a minha mente durante a pandemia e me desenvolvendo enquanto cientista.

Aos meus professores. Em especial, Simara, Anselmo e novamente ao meu orientador, Geraldo. Com certeza não teria aprendido tanto quanto aprendi durante a graduação sem ter tido a oportunidade de estudar com eles.

À todos que contribuíram de forma direta ou indireta para este trabalho.

“Nós só podemos ver um pouco do futuro, mas o suficiente para perceber que há muito a fazer.” (Alan Turing)

Resumo

A epidemia de Sars-CoV-2 tem sido um dos maiores problemas a serem enfrentados pela saúde pública a nível mundial. Somente no Brasil, de 2020 até Junho de 2021 (DONG; DU; GARDNER, 2020), 482 mil pessoas vieram a óbito em função do vírus. Também, de acordo com a Organização Mundial da Saúde (OMS), a pneumonia é responsável por 15% das mortes de crianças menores de 5 anos, sendo a maior causa de mortalidade infantil no mundo, especialmente em países em desenvolvimento (WHO, 2015). A partir destes dados, vemos como infecções respiratórias podem causar impactos fortes na saúde pública mundial. Portanto, o objetivo deste trabalho consiste em avaliar o desempenho de redes neurais profundas no contexto de detecção e diagnóstico de doenças pulmonares. Serão utilizadas técnicas de otimização e revisão bibliográfica para elaborar modelos de redes neurais convolucionais. Os modelos propostos foram capazes de alcançar *recall* elevado para a detecção de Sars-CoV-2, de 94,89% em redes *Efficient* e 97,44% em redes *Dense*. Para a detecção de doenças, sem discriminação, o modelo *Efficient* alcançou *recall* de 98,94%.

Palavras-chaves: classificação, regressão, aprendizagem de máquina, processamento de imagem.

Lista de ilustrações

| | |
|--|----|
| Figura 1 – Redes Neurais Convolucionais. | 15 |
| Figura 2 – ROC-AUC. | 21 |
| Figura 3 – Metodologia | 23 |
| Figura 4 – Amostras COVIDx-CXR. | 24 |
| Figura 5 – Base COVIDx-CXR 3. Treino e Teste. | 24 |
| Figura 6 – Base COVIDx-CXR 8. Treino e Teste. | 25 |
| Figura 7 – Tree of Parzen Estimators | 26 |
| Figura 8 – Arquitetura <i>Dense</i> | 27 |
| Figura 9 – Arquitetura <i>Efficient</i> | 28 |
| Figura 10 – Experimentos. | 28 |
| Figura 11 – Processo de amostragem. | 29 |
| Figura 12 – Exemplos de Data Augmentation. | 30 |
| Figura 13 – Matrizes de Confusão na Base COVIDx-CXR 3. | 31 |
| Figura 14 – Matrizes de Confusão na Base COVIDx-CXR 8. | 32 |
| Figura 15 – Matrizes por <i>Split</i> | 33 |
| Figura 16 – Curvas de Aprendizado e Perda. | 35 |
| Figura 17 – Exemplos da classe Pneumonia. | 36 |
| Figura 18 – Exemplos da classe COVID. | 36 |
| Figura 19 – Mapa de Ativação em Caso de COVID. | 37 |
| Figura 20 – Mapa de Ativação em Caso de Pneumonia. | 37 |

Lista de tabelas

| | |
|---|----|
| Tabela 1 – Matriz de Confusão | 22 |
| Tabela 2 – Resultados com a Base COVIDx-CXR 3 | 31 |
| Tabela 3 – Resultados com a base COVIDx-CXR 8. | 32 |
| Tabela 4 – Resultados por <i>Split</i> | 33 |
| Tabela 5 – Comparativo com o Estado da Arte. | 34 |
| Tabela 6 – Estado da Arte em Classificação Binária. | 34 |

Lista de abreviaturas e siglas

| | |
|------------|---|
| CNN | Convolutional Neural Networks |
| OMS | Organização Mundial da Saúde |
| AUC | Area Under Curve |
| Sars-CoV-2 | Severe Acute Respiratory Syndrome Coronavirus 2 |
| RT-PCR | Real Time Polymerase Chain Reaction |

Sumário

| | | |
|------------|---------------------------------------|-----------|
| 1 | INTRODUÇÃO | 10 |
| 1.1 | Objetivo | 10 |
| 1.1.1 | Objetivos Específicos | 11 |
| 2 | TRABALHOS RELACIONADOS | 12 |
| 3 | FUNDAMENTAÇÃO TEÓRICA | 14 |
| 3.1 | Aprendizado de Máquina | 14 |
| 3.1.1 | Aprendizagem Profunda | 15 |
| 3.2 | Redes Neurais Convolucionais | 15 |
| 3.2.1 | Arquiteturas DenseNets | 16 |
| 3.2.2 | Arquiteturas EfficientNet | 17 |
| 3.3 | Regularização de Redes Neurais | 19 |
| 3.4 | Avaliação | 20 |
| 4 | METODOLOGIA | 23 |
| 4.1 | Aquisição das Imagens | 23 |
| 4.1.1 | COVIDx-CXR 3 | 24 |
| 4.1.2 | COVIDx-CXR 8 | 25 |
| 4.2 | Otimização das Redes | 25 |
| 4.2.1 | Dense | 25 |
| 4.2.2 | Efficient | 27 |
| 4.3 | Experimentos | 28 |
| 5 | RESULTADOS | 31 |
| 5.1 | COVIDx-CXR 3 | 31 |
| 5.2 | COVIDx-CXR 8 | 32 |
| 5.3 | Melhor Rede | 33 |
| 5.4 | Discussão | 34 |
| 6 | CONCLUSÃO | 38 |
| | REFERÊNCIAS | 39 |

1 Introdução

A epidemia de Sars-CoV-2 se apresenta como um dos maiores desafios a serem superados pela saúde a nível mundial. Foi registrado pela Organização Mundial da Saúde (OMS) que até julho de 2021 a Sars-CoV-2 foi responsável por 4.459.381 mortes (WHO, 2021). Para a prevenção do espalhamento da doença junto às suas variantes algumas técnicas de triagem tem sido aplicadas para que seja possível realizar o diagnóstico de COVID de forma ágil, geralmente por RT-PCR, também conhecido como "teste rápido".

Um dos principais fatores que são responsáveis pela transmissão acelerada do vírus está na sua capacidade de se transmitir através do ar, assim como outras formas de pneumonia além de Sars-CoV-2. Dado o seu alto grau de transmissibilidade, é vital isolar de forma eficiente os infectados através de testagem. Com a importância do diagnóstico em massa de doenças respiratórias, é necessário estruturar métodos ágeis e eficientes de diagnóstico para combater de forma mais efetiva a Sars-CoV-2.

A visão computacional, através de aprendizado profundo, conseguiu alcançar o estado da arte em diversas áreas de atuação, como o reconhecimento de imagens (RUSAKOVSKY et al., 2015). Desde o sucesso da rede CNN AlexNet (KRIZHEVSKY; SUTSKEVER; HINTON, 2012), que em sua época marcou o melhor desempenho no desafio *imagenet*, diversas aplicações com aprendizado profundo tem sido desenvolvidas para o reconhecimento de imagem. Passados 9 anos, a visão computacional tem crescido de forma considerável no contexto médico (ESTEVA et al., 2019), acelerando diagnóstico de diversas patologias, como glaucoma (CHEN et al., 2015), pneumonia (KERMANY et al., 2018) e COVID-19 (WANG; LIN; WONG, 2020).

Este trabalho propõe métodos de classificação de pacientes com pneumonia e COVID-19 utilizando radiografias torácicas, processamento de imagem e aprendizado profundo.

1.1 Objetivo

Este trabalho tem como objetivo principal desenvolver uma metodologia de aprendizado de máquina profundo capazes de interpretar, com precisão, radiografias torácicas e realizar o diagnóstico quanto à síndromes respiratórias.

1.1.1 Objetivos Específicos

- Construir bases de dados sólidas que fundamentem o processo de treino e teste de modelos de CNN;
- Delimitar uma variedade de arquiteturas de redes CNN para serem testadas na base de dados COVIDx-CXR em suas versões mais atuais;
- Desenvolver abordagens de treino que contornem problemas de otimização e ajuste de curva para alcançar classificadores ótimos;
- Estruturar classificadores com alto *recall* para a detecção de doenças, COVID ou pneumonia, em um contexto binário;
- Estruturar classificadores com alto *recall* primariamente para a detecção de COVID e em seguida de Pneumonia, em um contexto de três classes;
- Avaliar os impactos do modelo proposto em relação ao desempenho dos trabalhos relacionados, atualmente estado da arte.

2 Trabalhos relacionados

Nesta seção serão apresentados os principais trabalhos relacionados ao tema de diagnóstico de imagens com *deep learning*, obtendo então um referencial do estado da arte na literatura.

Dentro do campo de análise de imagens para diagnóstico de doenças pulmonares existem predominantemente abordagens que trabalham com a análise de tomografias e radiografias. Cada uma das abordagens oferecem suas vantagens. Primeiramente, utilizando tomografias, pode-se obter uma visão ampla de todo o pulmão, em 3D. A abundância de informação pode oferecer ganhos em desempenho para o modelo de aprendizado, dada a alta resolução e diversos ângulos de exposição. Atualmente, trabalhos como a *COVIDNet-CT* (GUNRAJ; WANG; WONG, 2020) mostram que o desempenho destes modelos consegue alcançar resultados impressionantes em um conjunto de imagens amplo, de até 99% de acurácia e *recall* de 99,7% para SARS-CoV-2.

Apesar dos resultados impressionantes alcançados por modelos que analisam tomografias, exames de tomografia são mais custosos que radiografias. Este tipo de acessibilidade torna-se ainda mais problemática quando falamos de países em desenvolvimento, que são os mais afetados por doenças respiratórias como pneumonia, devido a falta de médicos especialistas ou equipamento. Cerca de 90% dos casos de mortalidade infantil em função de pneumonia ocorrem em países situados no sudeste asiático e na África (RAJARATNAM et al., 2010).

Para resolver este problema, trabalhos como os de Rajpurkar et al. (2017) e Rajpurkar et al. (2018) são precursores que buscam classificar imagens clínicas de radiografias e ao mesmo tempo traçam comparativos com o desempenho de médicos especialistas. No segundo trabalho, o modelo proposto de CNN foi capaz de alcançar performance superior a médicos especialistas em 11 das 14 patologias presentes no conjunto de imagens ChestX-ray14 (WANG et al., 2017), elaborado pela universidade de Stanford.

O trabalho de Wang et al. (2017), responsável pelo conjunto de dados ChestX-ray14, contém uma das maiores bases de dados públicas disponíveis para diagnóstico de doenças pulmonares. Neste trabalho, desenvolvido pelo NIH Clinical Center, são fornecidas imagens de 30 mil pacientes, apresentando uma grande variedade de doenças, tal qual o laudo completo de cada imagem. Esta base está se tornando um referencial para o diagnóstico de doenças respiratórias, sendo avaliado no RSNA Pneumonia Detection Challenge, promovido pela plataforma Kaggle (ANDRIOLE et al., 2018).

Considerando a alta mortalidade infantil proporcionada por doenças respiratórias como pneumonia, trabalhos como o de Rajaraman et al. (2018) e (KERMANY et al., 2018)

avaliam o desempenho de arquiteturas de redes neurais profundas para o diagnóstico de radiografias pediátricas, utilizando a base de dados de raios-x do hospital de Guangzhou. Esta base de dados apresenta 5.248 imagens em seu conjunto de treino e 627 em seu conjunto de teste, diferenciando as amostras entre pneumonia viral e bacteriana. A rede CNN proposta por [Rajaraman et al. \(2018\)](#) alcança 96,4% de acurácia para a detecção de pneumonia.

A *COVID-Net* ([WANG; LIN; WONG, 2020](#)), de acordo com o conhecimento do autor, é o trabalho mais eficaz que vem sido desenvolvido com o objetivo de diferenciar adultos com COVID-19, pneumonia e saudáveis. Em 2019 foi formado o dataset público COVIDx ([WANG; LIN; WONG, 2020](#)) abrangendo cerca de 13.975 amostras de 13.870 pacientes, unindo amostras de diversas outras bases de dados públicas, como a do NIH Clinical Center. Também, revisões proporcionadas em sua arquitetura tem continuamente atualizando a base de dados e seus modelos ajustados, disponíveis em um repositório público ([WANG; LIN; WONG, 2021](#)). Atualmente seu melhor resultado, também estado da arte, classifica corretamente 94,0% das amostras normais, 94,0% das amostras com pneumonia e 95,0% das amostras com COVID-19.

Este trabalho então estrutura metodologias de classificação de amostras de raios-x ao invés de tomografias por dois motivos. Primeiramente, devido à falta de melhorias significativas que possam ser aplicadas aos modelos, em termos de acurácia e *recall*, como poder ser visto no trabalho de ([GUNRAJ; WANG; WONG, 2020](#)). Em segundo lugar, para expandir o espaço de aplicação dos modelos no contexto mais importante: em regiões onde o exame de tomografia é menos acessível.

Nosso trabalho, assim como a *COVID-Net* utiliza amostras de adultos com Sars-CoV-2, pneumonia não causada por Sars-CoV-2 e saudáveis. Foram utilizados fontes de dados referências, elicitadas no trabalho de [Wang, Lin e Wong \(2020\)](#), como a base *ChestX-Ray14* ([WANG et al., 2017](#)). Busca-se também realizar uma avaliação mais detalhada dos modelos, com amostras mais atuais, mapas de ativação e *area under curve*.

3 Fundamentação Teórica

Neste capítulo estão presentes os principais fundamentos matemáticos que justificam a metodologia utilizada e explicam os resultados obtidos.

3.1 Aprendizado de Máquina

A inteligência artificial é o ramo da computação que lida com a programação de sistemas capazes de identificar contextos com semelhança à interpretação humana, apresentando desempenho ou comportamento que pode ser identificado como inteligência. Não existe, atualmente, um conceito claro que defina inteligência artificial, mas esta definição se alinha com o Teste de Turing (TURING, 2009). Sistemas inteligentes são responsáveis pelo funcionamento de diversas tecnologias, como sistemas de busca à recomendações de produtos.

A aprendizagem de máquina é uma sub-área de inteligência artificial que estrutura procedimentos para que máquinas possam adquirir conhecimento. Esta capacidade, chamada de cognição, é a base para que sistemas sejam capazes de classificar e diferenciar dados. Este processo é alcançada pelo reconhecimento de padrões em conjuntos de dados.

A forma mais comum de aprendizado é o aprendizado supervisionado. Ele consiste em duas etapas: a observação de um conjunto de dados rotulado e o reconhecimento deste rotulo. Por exemplo, pode-se construir um sistema para a identificação de objetos em uma imagem. Cada imagem é rotulada com o objeto que está enquadrado, como carros, pessoas, sapatos ou vestidos. O sistema observa uma determinada amostragem, produzindo probabilidades para cada categoria. Então, é computada uma função capaz de representar o erro do modelo, como na Equação 3.1. Nesta equação y representa as probabilidades reais. \hat{y} as probabilidades dadas pelo modelo. Para reduzir os valores da função objetivo, o modelo ajusta seus parâmetros internos.

$$L(y_i, \hat{y}_i) = \sum_{n=1}^N (y_i - \hat{y}_i) \quad (3.1)$$

Para este trabalho, utilizamos em ambas as redes a função de perda entropia cruzada, definida pela Equação 3.2, tal que y é a classe real, \hat{y} a saída do modelo.

$$L(y_i, \hat{y}_i) = \sum_{n=1}^N y - \log(\hat{y}_i) \quad (3.2)$$

3.1.1 Aprendizagem Profunda

Por muito tempo modelos de aprendizado de máquina estiveram limitados ao dimensionamento meticuloso da informação através de especialistas capazes de interpretar de forma coerente o domínio da aplicação.

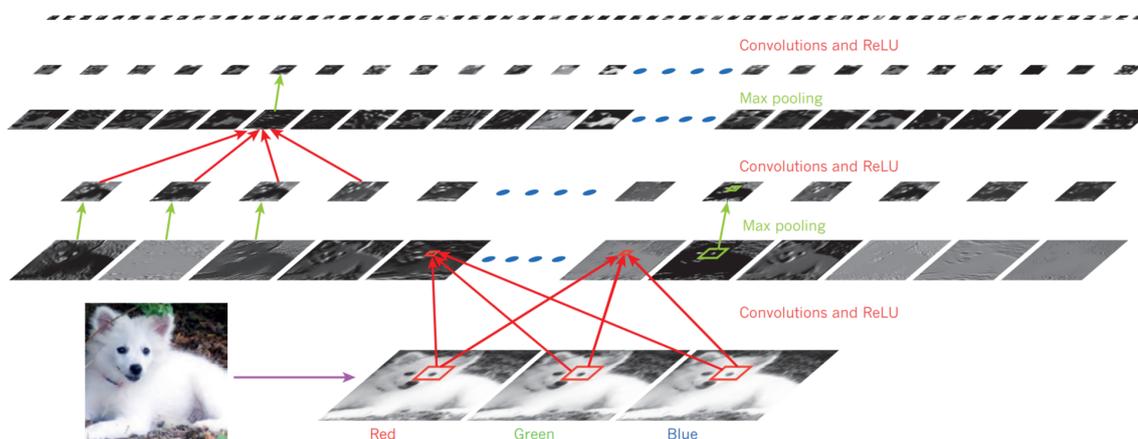
O aprendizado profundo, ao invés de depender do redimensionamento da informação, aprende representações do domínio da aplicação automaticamente. Este aprendizado ocorre através do sequenciamento de operações não-lineares que aditivamente elevam o grau de abstração dos dados. O ponto chave deste procedimento está em que o algoritmo por si só é capaz de criar representações multi-nível da informação com procedimentos gerais de aprendizado (LECUN; BENGIO; HINTON, 2015).

O aprendizado profundo tem elevado sua popularidade em aplicações de alta dimensão, como processamento de linguagem natural e reconhecimento de imagens, devido a sua capacidade de processar informação com mínima reorganização dos dados.

3.2 Redes Neurais Convolucionais

Redes neurais convolucionais integram extração de características e classificação do modelo, aderindo os princípios estabelecidos pelo conceito de aprendizado profundo. O ajuste dos modelos se dá através dos filtros de convolução, que são camadas de processamento que aumentam o grau de não linearidade. Assim, conforme o número de camadas aumenta, eleva-se a dimensão das características extraídas. Em uma CNN, normalmente são utilizadas várias camadas de convolução e *pooling*, onde uma dilata o número de canais e outra os reduz, respectivamente (O'SHEA; NASH, 2015). A Figura 1 ilustra as etapas de processamento de uma CNN.

Figura 1 – Redes Neurais Convolucionais.



Fonte: LeCun, Kavukcuoglu e Farabet (2010)

Os filtros de convolução são baseados em mecanismos biológicos visuais e tem como objetivo de aprender características morfológicas e espaciais dos objetos de interesse. Subsequentes camadas são combinadas através de seus campos receptivos locais e isto permite o reuso de características das imagens, como bordas ou cantos.

Camadas de convolução geram mapas de características a partir da aplicação de uma sucessão de filtros treináveis que estão sujeitos à uma função não-linear. A Equação 3.3 caracteriza o mapa de ativação resultante de uma camada de convolução. k é o índice do neurônio; h_k é o mapa de características com saída k ; w_k é o filtro treinável k ; x é o mapa de entrada do neurônio; b é o bias de k ; o operador $*$ representa a convolução bi-dimensional (LECUN; KAVUKCUOGLU; FARABET, 2010).

$$(h_k)_{ij} = (w_k * x)_{ij} + b_k \quad (3.3)$$

A convolução bi-dimensional é caracterizada pela equação 3.4. Na convolução espacial aplica-se em uma janela N por M , no pixel Z_{ij} novos valores multiplicados pelos pesos da janela de convolução W_{ij} .

$$(Z)_{ij} = \sum_{k=1}^{N*M} Z_k * W_k \quad (3.4)$$

As camadas de *pooling* são as camadas responsáveis por reduzir o mapa de características conforme o crescimento da profundidade da rede. Nem todas as redes modernas utilizam camadas de *pooling*, mas as que utilizam conseguem tornar subseqüentes camadas de ativação mais robustas a variações de taxa de proporção e translação (LECUN; KAVUKCUOGLU; FARABET, 2010) a custo de informação e reuso de características. Um exemplo é a camada *max pooling*, que cobre o mapa de ativação para que somente permaneça o maior valor do campo receptivo filtrado.

As CNNs tem sido um recurso amplamente utilizado, devido a maneira como a propagação da informação pode ser paralelizada, aprendem com muita informação sem gerar modelos superajustados (LECUN; BENGIO; HINTON, 2015) e seus resultados obtidos em competições que são referência, como o desafio *imagenet* (DENG et al., 2009), que só recentemente, em 2020, foi tomado por modelos diferentes, chamados *Vision Transformers* (DOSOVITSKIY et al., 2020).

3.2.1 Arquiteturas DenseNets

Uma das arquiteturas de rede convolucional utilizadas neste está fortemente inspirada na arquitetura de redes do tipo *Dense* (HUANG et al., 2017). A arquitetura *Dense* é

um tipo de rede convolucional que tem como proposta a utilização de *skip connections* para a produção de mapas de características mais eficientes.

Conforme o acréscimo de camadas são perdidas características importantes. Arquiteturas de redes convolucionais como *ResNet* tentam resolver este problema como na Equação 3.5, somando o resultado de camadas anteriores às camadas subsequentes. Nesta equação a saída de uma camada x_l é dada pela aplicação da operação não-linear de convolução H em x_{l-1} com a saída de x_{l-1} .

$$x_l = H_l(x_{l-1}) + x_{l-1} \quad (3.5)$$

O trabalho de [Huang et al. \(2017\)](#) argumenta que a adição não necessariamente consegue persistir a informação das camadas anteriores e então propõe conectividade *Dense*, onde o mapa de características de uma camada l é formado como na equação 3.6, onde a camada x_l conecta-se diretamente com as camadas x_{l-1} , x_{l-2} e em diante.

$$x_l = H_l[x_0, x_1, \dots, x_{l-1}, x_l] \quad (3.6)$$

Devido ao grande número de características acumuladas em cada novo tensor das camadas *Dense*, são propostos blocos de transição para realizar a subamostragem dos campos de ativação, utilizando convolução 1×1 e *max pooling*. No Capítulo de Metodologia estão diagramados os blocos de convolução e transição.

De acordo com os experimentos desempenhados, o trabalho foi capaz de reduzir o erro em relação a arquitetura de redes neurais convolucionais *ResNet*, constituindo então, no ano de 2018, o estado da arte para o reconhecimento de imagens do desafio *imagenet*.

3.2.2 Arquiteturas EfficientNet

As redes *Efficient* são uma família de redes convolucionais baseadas nas redes *MobileNet*. A arquitetura tem como fundamento reduzir o número de pesos utilizados e a complexidade do algoritmo de convolução, garantindo inferência e treino mais rápidos. Em 2019 conseguiu superar os melhores modelos de rede neurais convolucionais com uma arquitetura 8,4x menor e com inferência 6,1x mais rápida ([HOWARD et al., 2017](#)).

Até hoje em dia não existem formas padronizadas de estimar escala de redes para que elas sejam capazes de acomodar mais informações. Alguns trabalhos aumentam o número de parâmetros expandindo o número de camadas, outros através do número de filtros de convolução e alguns pelo aumento da resolução de entrada.

Ao observar que uma rede convolucional N pode ser definida como uma lista de camadas compostas F : $N = F_k \odot F_{k-1} \odot \dots \odot F_1 \odot F(X_1)$. Assumindo que todas as redes compartilham o mesmo tipo de convolução, temos

$$N = \bigodot_{i=1 \dots s} F_i^{L_i}(X_{\langle H_i, W_i, C_i \rangle}) \quad (3.7)$$

Onde $F_i^{L_i}$ denota que F_i é repetida L_i vezes no estágio i . $\langle H_i, W_i, C_i \rangle$ descreve o formato do tensor da camada i , tal que H e W representam as dimensões espaciais e C os canais.

Mesmo com F_i fixo, existem diferentes combinações arquiteturais de L_i , C_i , H_i , W_i em cada camada. Cria-se então uma restrição onde todas as camadas devem ser escaladas de maneira uniforme. A ideia é maximizar a acurácia dada restrições de recursos. O problema de otimização é formulado como nas Equações 3.8, 3.9, 3.10 e 3.11.

$$\max(d, w, r) = \text{Acuracia}(N(d, w, r)) \quad (3.8)$$

$$N(d, w, r) = \bigodot_{i=1 \dots s} \hat{F}_i^{d \cdot L_i}(X_{\langle H_{r^*i}, W_{r^*i}, C_{w^*i} \rangle}) \quad (3.9)$$

$$\text{Memoria}(N) \leq \text{MemoriaAlvo} \quad (3.10)$$

$$\text{FLOPS}(N) \leq \text{FLOPSAlvo} \quad (3.11)$$

O artigo define que $d = \alpha^\phi$, $w = \beta^\phi$ e $r = \gamma^\phi$, impondo as restrições da Equação 3.2.2. Desta forma, espera-se que cada novo ϕ multiplique o número de parâmetros por uma base de 2.

$$\alpha \cdot \beta^2 \cdot \gamma^2 \approx 2 \quad (3.12)$$

Após processos de meta-otimização com ϕ de 0 à 7 dentro do espaço de busca definido pelo autor, através de sua revisão bibliográfica, empiricamente e com experimentos com os hiper-parâmetros d , w e r , foi possível escalar os modelos *MobileNet* utilizando *depthwise separable convolutions* (HOWARD et al., 2017), criando a família de redes *Efficient*.

3.3 Regularização de Redes Neurais

Métodos de regularização de redes convolucionais podem ser aplicados de maneira explícita e implícita aos modelos. O objetivo é permitir que metodologias de ajuste de pesos maximizem métricas de desempenho desejadas, generalizando melhor a informação ou seu domínio de aplicação. Neste trabalho, algumas técnicas de regularização são aplicadas para melhorar o desempenho do método proposto. Entre elas, estão:

- Drop-out
- Data Augmentation
- Pesos de Classe
- Poda
- *Batch-Normalization*

Drop-out é um processo no qual uma determinada região de uma rede neural não tem seus pesos treinados após a aplicação de um algoritmo de otimização. Este processo inibe a co-adaptação de pesos em um detector de características, produzindo tensores que agregam informações menos redundantes (HINTON et al., 2012). O processo de *drop-out* também é responsável por reduzir o número de parâmetros treináveis, acelerando o processo de ajuste dos classificadores.

Com a ideia de proporcionar uma amostragem variada e tornar modelos capazes de produzir características invariantes à transformações geométricas como rotação, translação e distorção, *Data Augmentation* é aplicado. Construir esquemas de augmentation que resultem em algoritmos simples e rápidos é uma questão de implementação, que em estratégias bem-sucedidas variam muito com os dados observados (DYK; MENG, 2001).

O processo de *Data Augmentation* consiste na super-amostragem de pontos do conjunto de treino, através de transformações introduzidas de maneira artificial à amostras já existentes, aumentando a quantidade de dados disponíveis com amostras não observadas. Este tipo de tratamento tende à solucionar problemas como desbalanceamento de dados, tal que permite ajustar a probabilidade de seleção das amostras das classes modificadas.

Um dos problemas a serem solucionados no processo de augmentation em modelos com ajuste estocástico é que as características do mapa de ativação tendem a se ajustar rapidamente às novas amostras, dado que a estrutura dos dados se assemelha. Caso não haja semelhança, a nova amostra não é representativa de sua classe e conseqüentemente pode vir a piorar o procedimento de ajuste. O balanceamento destes fatores define o sucesso da estratégia de *augmentation*.

Outra forma de regularização é a utilização de pesos de classe. Pesos de classe são responsáveis por aumentar valores de perda em amostras de uma determinada classe. Com eles é possível fazer com que o ajuste dos pesos de uma rede neural, ou qualquer modelo que otimiza uma função objetivo, ajuste arbitrariamente à classes importantes. A implementação de pesos de classe para a função objetivo *entropia cruzada* à probabilidades p de uma amostragem y com pesos w pode ser definida pela Equação 3.13.

$$\frac{\sum_{i=1}^n w_{y_i} \log p_{y_i}}{\sum_{i=1}^n w_{y_i}} \quad (3.13)$$

O processo de poda de um classificador trata-se de salvar estágios do treino de seus pesos em avaliações intermediárias. Esta avaliação pode ser da pela função objetivo de treino ou outras métricas que avaliem de forma consistente o desempenho de classificador em seu respectivo estado. Neste trabalho, optamos por gravar pesos a cada época treinada, utilizando como parâmetro acurácia no conjunto de validação.

Um dos problemas ocorridos durante o treino de redes neurais profundas é o fato de que a distribuição das entradas muda a durante o treino, de acordo com o nível das camadas anteriores, o que facilita com que funções de ativação saturem caso não haja dimensionamento correto da taxa de aprendizado durante o treino. Este fenômeno é chamado de "*internal covariate shift*" por Ioffe e Szegedy (2015), que o soluciona com a técnica *batch normalization*.

Batch normalization trata-se de uma camada que refaz o processo de normalização do tensores da rede no eixo de amostragem. Seus experimentos mostram que, com esta técnica, é possível utilizar taxas de aprendizado altas e utilizar diversas inicializações de pesos com impactos mínimos ao desempenho do modelo, amenizando a saturação dos campos de ativação ao centralizar a amostragem. O estudo de (IOFFE; SZEGEDY, 2015) mostra que a técnica é capaz de reduzir o número de ajustes em modelos estado da arte sem impacto à acurácia.

3.4 Avaliação

Algumas das métricas utilizadas durante a avaliação do modelo são: acurácia, *area under curve*, *F1 score* e *recall*, também conhecido como sensibilidade. A Equação 3.16 define recall. A equação 3.18, acurácia. 3.15, *F1 score*.

$$Precisao = \frac{VP}{VP + FP} \quad (3.14)$$

$$F1 = 2 * \frac{Precisao * Sensibilidade}{Precisao + Sensibilidade} \quad (3.15)$$

$$Recall = \frac{VP}{VP + FN} \quad (3.16)$$

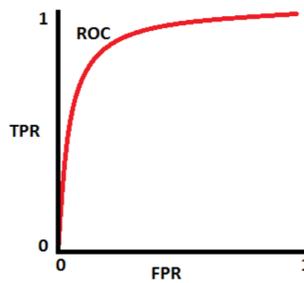
$$Especificidade = \frac{VN}{FP + VN} \quad (3.17)$$

$$Acuracia = \frac{VP + VN}{VP + FP + VN + FN} \quad (3.18)$$

- Verdadeiros Positivos: VP;
- Verdadeiros Negativos: VN;
- Falsos Positivos: FP;
- Falsos Negativos: FN.

Area under curve, é a métrica que mensura a capacidade do modelo de distinguir entre classes. Ela é a curva que define a interseção da taxa de positivos verdadeiros (TPR) em relação a taxa de falsos positivos (FPR) de acordo com cortes nas probabilidades de classificação de 0 à 1, como ilustra a Figura 2. A integral desta curva é a métrica referenciada nas tabelas do Capítulo Resultados.

Figura 2 – ROC-AUC.



Fonte: [Narkhede \(2018\)](#)

$$TPR = \frac{VP}{VP + FN} \quad (3.19)$$

$$FPR = \frac{VN}{VN + FP} \quad (3.20)$$

Tabela 1 – Matriz de Confusão

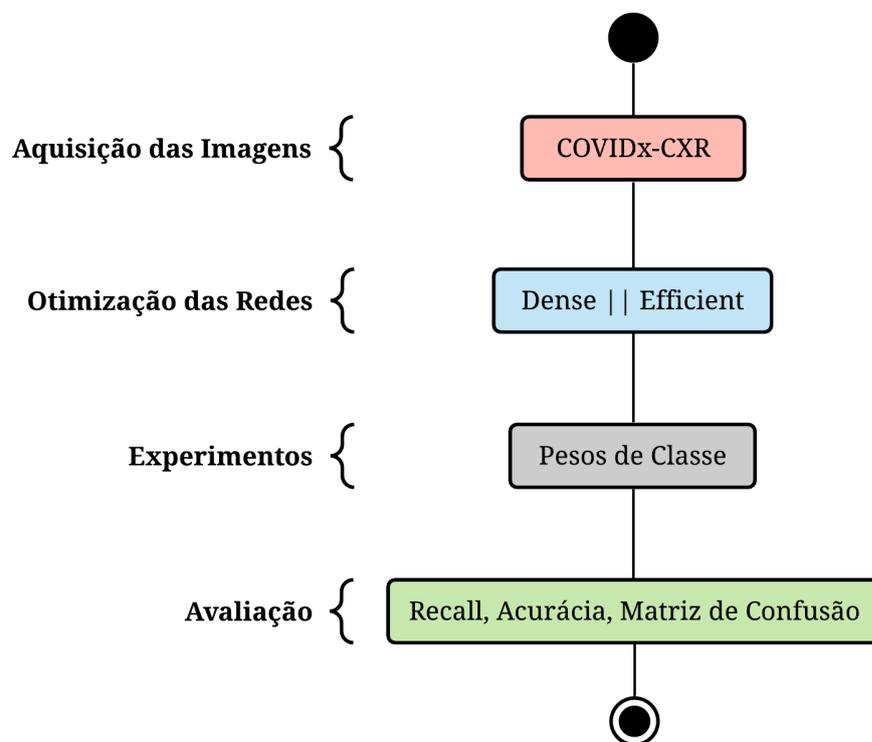
| | Saudável | Pneumonia | COVID |
|------------------|-----------------|------------------|--------------|
| Saudável | VP | FP | FP |
| Pneumonia | FP | VP | FP |
| COVID | FP | FP | VP |

Uma matriz de confusão é uma tabela de visualização do desempenho de algoritmos de aprendizagem supervisionada. As linhas representam como as classes foram selecionadas pelo algoritmo, enquanto as colunas representam as classes reais das amostras (POWERS, 2020). Um exemplo pode ser visto na tabela 1. Matrizes como esta estarão detalhando os experimentos citados no Capítulo Resultados.

4 Metodologia

Nesta seção é detalhada a metodologia utilizada, sequenciando o processo da construção das bases de dados, a escolha das arquiteturas de CNN, como foram executados os experimentos e o processo de avaliação. O passo a passo é apresentado na Figura 3.

Figura 3 – Metodologia



Fonte: Autor.

4.1 Aquisição das Imagens

Durante nosso trabalho são utilizados splits da base de dados COVID_x-CXR (WANG; LIN; WONG, 2020). Esta base de radiografias reúne, em sua terceira versão, as imagens das coleções: COVID-19 Image Data Collection (COHEN et al., 2020); COVID-19 Chest X-ray Dataset Initiative (CHUNG et al., 2021b); ActualMed COVID-19 Chest X-ray Dataset Initiative (CHUNG et al., 2021a); RSNA Pneumonia Detection Challenge (ANDRIOLE et al., 2018); COVID-19 radiography database (RAHMAN et al., 2021).

Desde sua primeira publicação a base de dados COVIDx-CXR tem passado por alterações. Neste trabalho foram executados experimentos com as versões 3 e 8, sendo 3 a versão do trabalho de Wang, Lin e Wong (2020) e 8 a edição mais atual. Além de amostras adicionadas às bases já existentes, uma nova base de dados, RICORD COVID-19 (TSAI et al., 2021), foi agregada em sua oitava versão.

Na Figura 4, em a, b e c, podem ser visualizadas as respectivas amostras de cada tipo de paciente. Infectados por COVID, infectados por pneumonia não causada por COVID e pacientes saudáveis.

Figura 4 – Amostras COVIDx-CXR.



a) COVID.

b) Pneumonia.

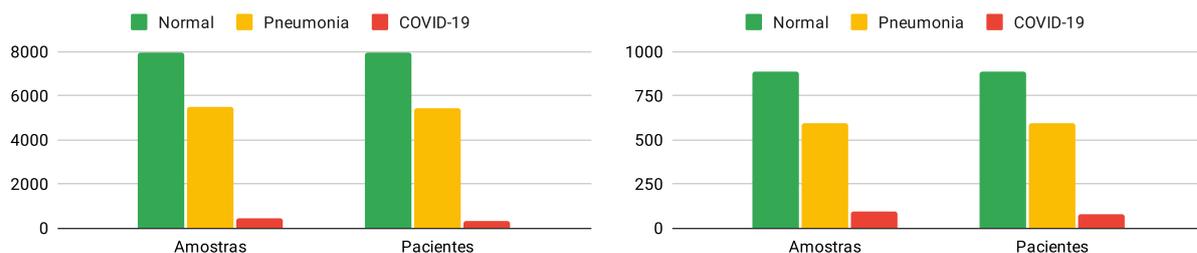
c) Normal.

Fonte: Wang, Lin e Wong (2020).

4.1.1 COVIDx-CXR 3

No split da base de dados COVIDx-CXR 3 temos, no conjunto de treino, 461 amostras que apresentam infecções por COVID-19, 5.458 com pneumonia causada por outros agente infecciosos e 7.966 que não apresentam qualquer tipo de patologia. Também, existem 100 imagens de teste para COVID, 594 com pneumonia e 885 saudáveis. Pode-se visualizar a distribuição da base de dados na Figura 5.

Figura 5 – Base COVIDx-CXR 3. Treino e Teste.

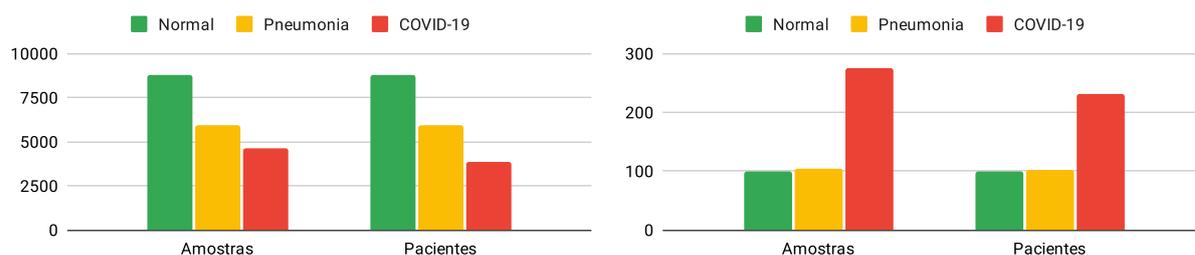


Fonte: Acervo do Autor.

4.1.2 COVIDx-CXR 8

A base de dados COVIDx-CXR 8 tem em seu conjunto de treino 4.649 radiografias que apresentam infecção por COVID, 5.964 amostras com a presença de pneumonia causada por outros agentes infecciosos e 8.751 amostras saudáveis. No conjunto de teste temos 274 amostras com COVID, 105 amostras de pneumonia e 100 amostras saudáveis. Pode-se visualizar a distribuição da base de dados na Figura 6.

Figura 6 – Base COVIDx-CXR 8. Treino e Teste.



Fonte: Acervo do Autor.

4.2 Otimização das Redes

Ao longo de nosso trabalho tomamos a decisão de elaborar dois tipos de arquitetura de rede neural, utilizando como referência diferentes tipos de rede amplamente aplicadas na literatura. Redes do tipo *Dense* (IANDOLA et al., 2014) e *Efficient* (TAN; LE, 2019).

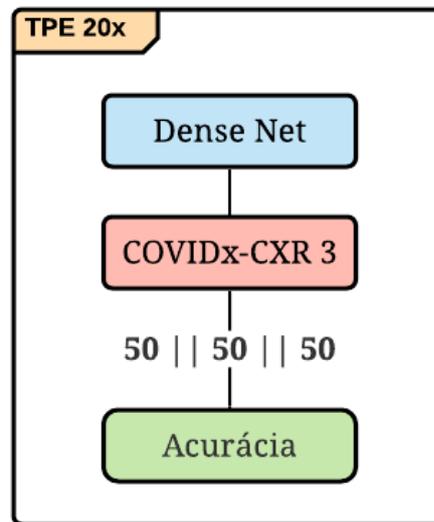
Ambas as redes foram originalmente propostas com o objetivo de solucionar o desafio de classificação *imagenet* (DENG et al., 2009) e são redimensionadas para a tarefa de detecção de Sars-CoV-2. Para a rede *Dense*, foram adotados métodos de *meta-otimização*, enquanto a rede *Efficient* foi redimensionada de acordo com a metodologia de escala do seu artigo original.

4.2.1 Dense

Para realizar o processo de meta-otimização, optamos pelo algoritmo *Tree of Parzen Estimators* (BERGSTRA et al., 2011) implementado pela biblioteca *hyperopt* (BERGSTRA; YAMINS; COX, 2013). Nele foi possível especificar diferentes fluxos de treino para obter a solução otimizada em nosso problema de classificação, tomando como referência a base de dados COVIDx-CXR 3, como ilustra a Figura 7.

Durante o treino foram utilizadas duas redes *Dense*. Uma rede *Dense 121* com entrada em escala de cinza e uma rede do tipo *Dense* com menos camadas, mas com uma resolução de entrada maior. O objetivo foi se aproximar à arquitetura de rede neural *COVID-Net*, assimilando a resolução de entrada de 480 por 480 pixels.

Figura 7 – Tree of Parzen Estimators



Fonte: Autor.

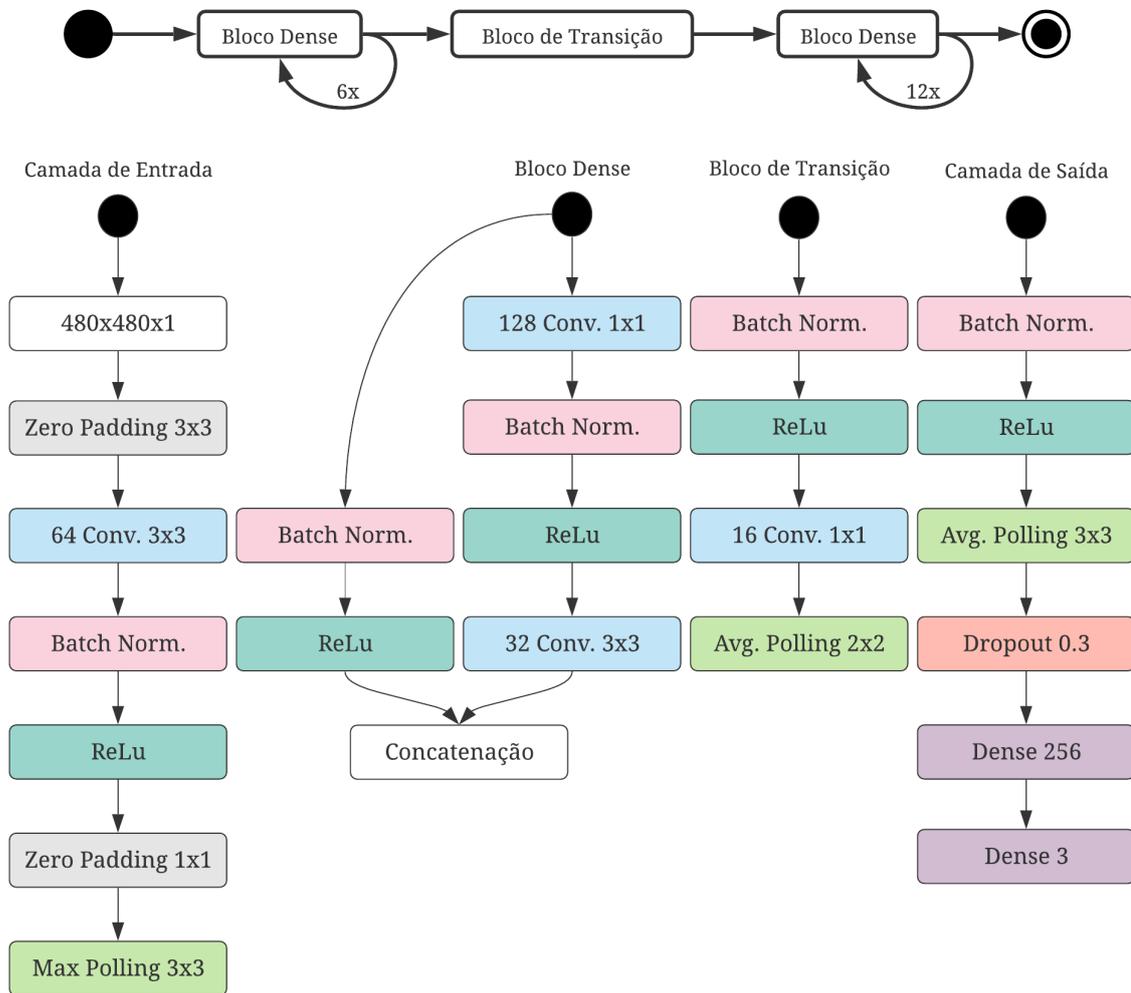
Assim, como parâmetros do processo de meta-otimização, definimos no espaço de busca ambas as redes, taxa de aprendizado e tipo de otimizador. A função objetivo do algoritmo então buscou maximizar a acurácia em um conjunto de validação com 50 amostras saudáveis, 50 amostras apresentando pneumonia sem COVID e 50 amostras de pacientes infectados por COVID.

Após 20 iterações do algoritmo *Tree of Parzen Estimators* o modelo que apresentou a melhor acurácia foi a *Dense Net* com:

- 2 blocos de convolução do tipo *Dense*;
- Gradiente Descendente Estocástico;
- *Batch* de tamanho 8;
- Taxa de aprendizado 0,0076.

A arquitetura resultante pode ser visualizada no diagrama da Figura 8. A estrutura, a nível macro, é destacada pelo fluxograma horizontal. Cada um dos blocos pode ser expandido nas estruturas verticais e suas operações.

Figura 8 – Arquitetura *Dense*.



Fonte: Acervo do Autor.

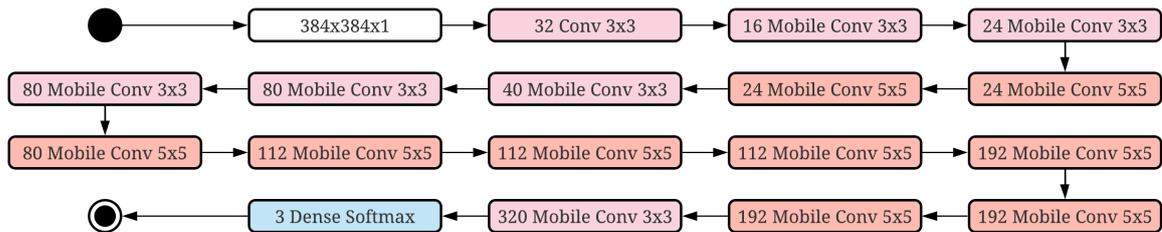
4.2.2 Efficient

Assim como a rede *Dense*, parametrizamos uma rede convolucional *Efficient*. A proposta das redes do tipo *Efficient* segue a linha de manter uma quantidade de parâmetros com escala ajustável através da otimização de ϕ , que mantém uma razão quadrática para resolução, linear para profundidade e número de camadas de convolução (TAN; LE, 2019).

A principal alteração durante a parametrização da CNN foi reduzir os canais de entrada, de três para um, em escala de cinza. Em seguida, compensando a perda de parâmetros da alteração de profundidade, a resolução de entrada foi expandida em uma proporção aproximada à raiz de 3, adotando a resolução de 384 por 384. Assim foi possível manter uma quantidade semelhante de pesos à arquitetura *EfficientNet-b0*, mantendo a proposta de rede escalável (TAN; LE, 2019).

A Figura 9 apresenta a topologia da rede. As células com o nome "Mobile Conv" representam a *Depthwise Separable Convolution*, proposta com as arquiteturas *Mobile Net* (HOWARD et al., 2017), precursoras das redes *Efficient*. As células "Conv", convolução convencional com ativação *ReLU*.

Figura 9 – Arquitetura *Efficient*.

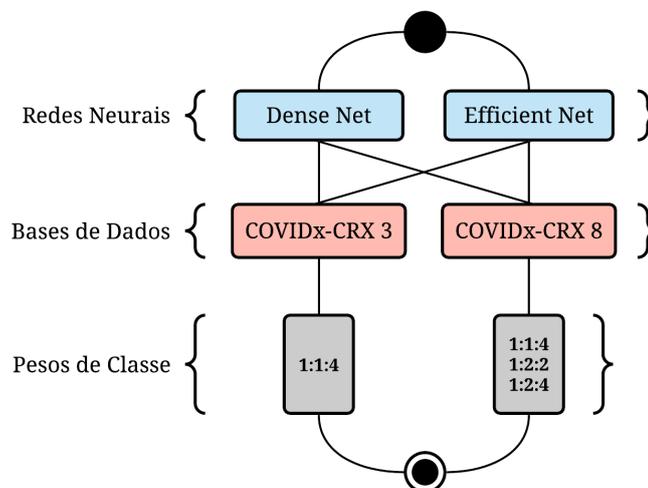


Fonte: Autor.

4.3 Experimentos

Entre os experimentos desempenhados com as técnicas de aprendizado profundo, temos variações de pesos de classe para contornar problemas de desbalanceamento, perceptíveis nas Figuras 5 e 6. Enquanto para a base COVIDx-CXR 3 adotamos os pesos de classe utilizados pelo trabalho de Wang, Lin e Wong (2020), base COVIDx-CXR 8 foram utilizadas 3 variações de pesos de classe, como na Figura 10.

Figura 10 – Experimentos.

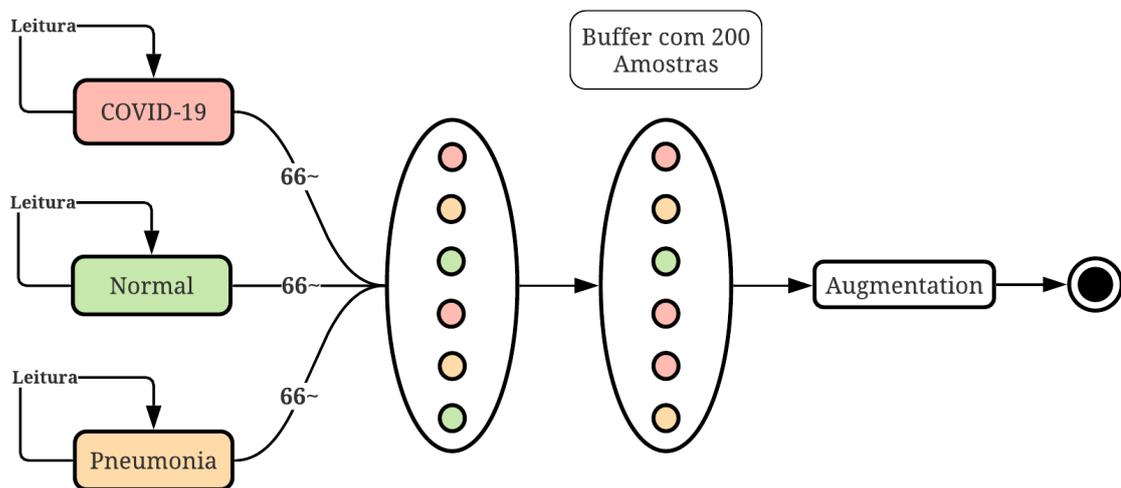


Fonte: Autor.

A Figura 10 ilustra, em suas células cinzas, quais pesos de classe foram utilizados para o cálculo da perda. Na célula esquerda, temos 1 para amostras saudáveis, 1 para pneumonia não Sars-CoV-2 e 4 para Sars-CoV-2. Para a direita, os diferentes pesos de classe aplicados com a base COVIDx-CXR 8.

Durante o processo de treino estabelecemos uma abordagem diferente de amostragem também para evitar problemas de super-ajuste em função do desbalanceamento da base de dados. Um dos problemas da amostragem aleatória de uma base de dados é que o algoritmo gradiente descendente tende a favorecer a classe majoritária, dada a sua natureza iterativa. Com a intenção de proporcionar uma amostragem mais equilibrada, nosso fluxo de treino então segue o passo a passo diagramado na Figura 11.

Figura 11 – Processo de amostragem.



Fonte: Autor.

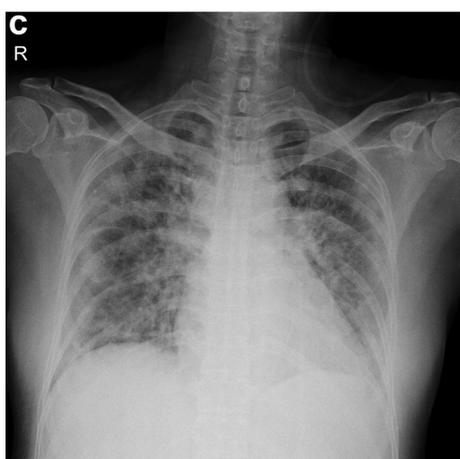
Primeiramente são extraídas sequencialmente amostras de cada base de dados. Então, tuplas formadas por trios de amostras para cada classe são sequenciadas. A partir daí, as 200 próximas amostras são embaralhadas entre si, formando um *buffer* que será consumido pelo processo de *data augmentation*.

Por último, são aplicadas técnicas de *Data Augmentation*. Foram aplicadas as seguintes transformações geométricas:

- Rotação variando entre 10° e -10° ;
- Translação variando entre $+10\%$ e -10% da resolução da imagem;
- Zoom entre $+15\%$ e -15% .

Um exemplo pode ser visualizado na Figura 12.

Figura 12 – Exemplos de Data Augmentation.



a) Sars-CoV-2.



b) Sars-CoV-2, Pós-*Augmentation*.

Fonte: Acervo do Autor.

5 Resultados

Neste capítulo são apresentadas visualizações e métricas que especificam os resultados de nosso fluxo metodológico, começando pela COVIDx-CXR 3 à COVIDx-CXR 8. Por último, na subseção de discussão são desenvolvidas as informações de cada experimento. Foram utilizadas, como referência, Acurácia, *ROC AUC*, *F-1 score* e *Recall*.

5.1 COVIDx-CXR 3

Para a base de dados COVIDx-CXR temos a matriz de confusão da Figura 13, apresentando os resultados obtidos pelas CNNs *Dense* e *Efficient*. Na Tabela 2 podem ser visualizadas as métricas de avaliação de ambas as redes.

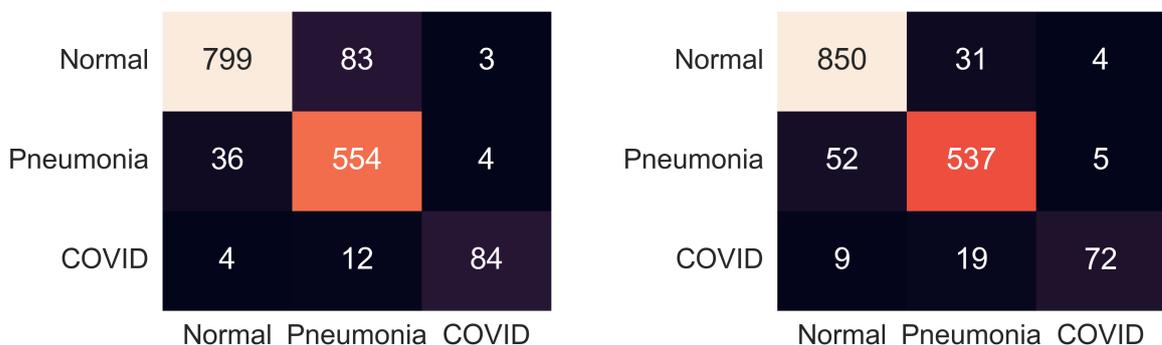
Tabela 2 – Resultados com a Base COVIDx-CXR 3

| | Acurácia | AUC | F-1 | Normal | Pneumonia | COVID |
|------------------|-----------------|------------|------------|---------------|------------------|--------------|
| Dense | 91,00% | 97,92% | 91,05% | 90,28% | 93,26% | 84,00% |
| Efficient | 92,40% | 86,14% | 92,30% | 96,04% | 90,40% | 72,00% |

Fonte: Acervo do Autor.

Aqui pode-se observar o desempenho superior da rede do tipo *Dense*, dada a diferença na métrica *area under curve*, proporcionada pelo tendenciamento da rede *Efficient* à classe majoritária.

Figura 13 – Matrizes de Confusão na Base COVIDx-CXR 3.



a) Matriz de Confusão *Dense*.

b) Matriz de Confusão *Efficient*.

Fonte: Acervo do Autor.

5.2 COVID_x-CXR 8

Com a base COVID_x-CXR 8 houveram testes com 3 variações de peso de classe. A Tabela 3 apresenta as métricas dos experimentos. Em todos os casos, com exceção à rede Dense com pesos 1:1:4, foram obtidos *recalls* elevados para as três classes.

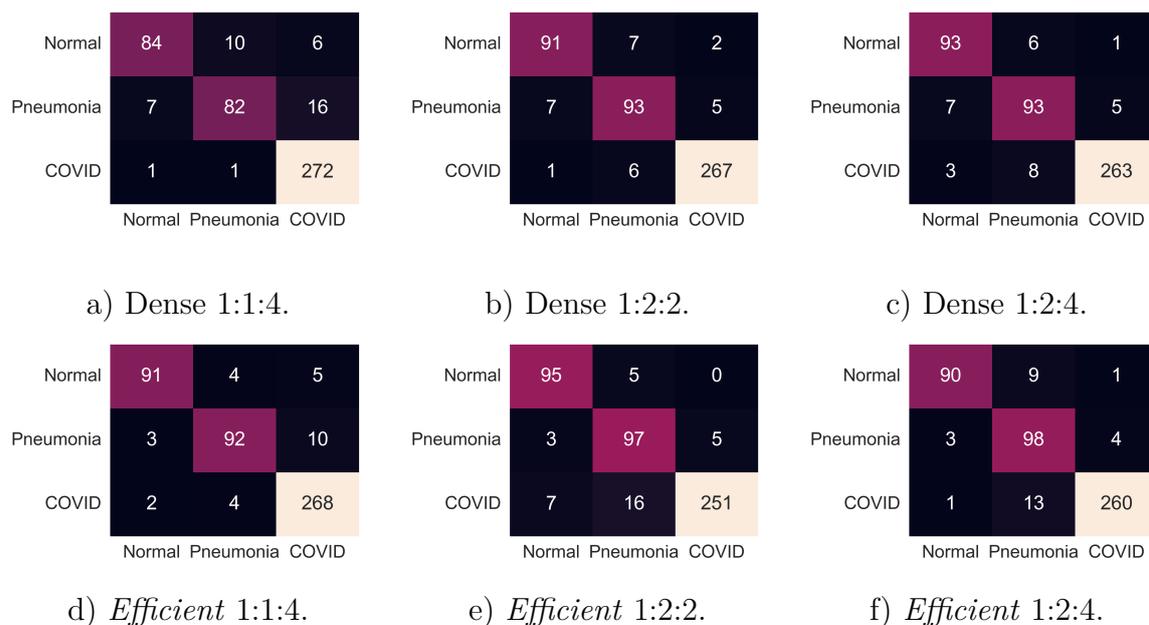
Tabela 3 – Resultados com a base COVID_x-CXR 8.

| Pesos | Acurácia | AUC | F-1 | Normal | Pneumonia | COVID | Rede |
|-------|----------|--------|--------|--------|-----------|--------|-----------|
| 1:1:4 | 91,44% | 99,15% | 91,20% | 84,00% | 78,09% | 99,27% | Dense |
| 1:2:2 | 94,15% | 98,91% | 94,15% | 91,00% | 88,57% | 97,44% | Dense |
| 1:2:4 | 93,73% | 99,01% | 93,77% | 93,00% | 88,57% | 95,98% | Dense |
| 1:1:4 | 94,15% | 98,42% | 94,10% | 91,00% | 87,61% | 97,81% | Efficient |
| 1:2:2 | 92,48% | 98,77% | 92,59% | 95,00% | 92,38% | 91,60% | Efficient |
| 1:2:4 | 93,52% | 99,13% | 93,65% | 90,00% | 93,33% | 94,89% | Efficient |

Fonte: Acervo do Autor.

As matrizes de confusão que apresentam os resultados das CNNs Dense e *Efficient* podem ser visualizadas na Figura 14. Através destas matrizes, pode-se concluir que o melhor resultado é alcançado através da combinação de pesos 1:2:4 e rede *Efficient*. Caso hajam falsos positivos a classificação estará distribuída entre casos semelhantes, pneumonia e Sars-CoV-2. Somente 4 amostras com doenças foram classificadas como saudáveis, consolidando um *recall* de 98,94% em um contexto de classificação binária.

Figura 14 – Matrizes de Confusão na Base COVID_x-CXR 8.



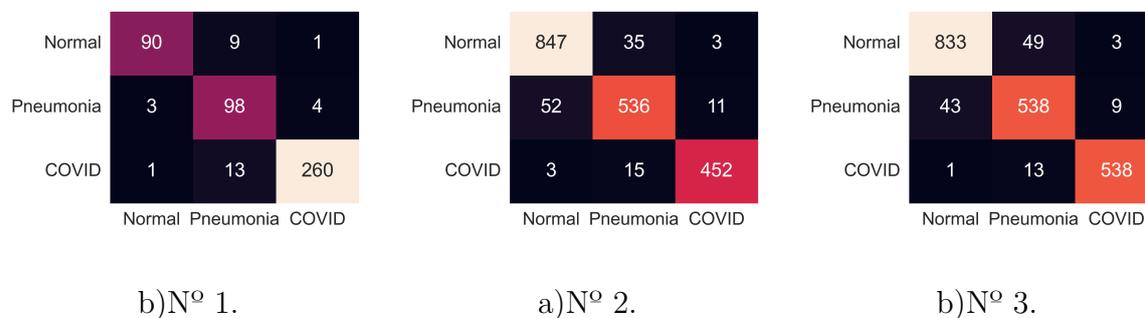
Fonte: Acervo do Autor.

5.3 Melhor Rede

A partir dos resultados obtidos na seção anterior, realizamos experimentos para validar a consistência dos resultados obtidos através da rede *Efficient* com pesos de classe 1:2:4. A rede foi selecionada devido ao seu alto *recall* para a detecção de doenças, Pneumonia e COVID, de 98,90%. O modelo errou três amostras positivas para pneumonia como normais e uma amostra positiva para COVID como normal.

Foram geradas, nesta abordagem, mais dois *splits* da base de dados COVIDx-CXR 8 em *hold-out*, onde separamos 90% dos pacientes para treino e os 10% restantes para teste. A Figura 15 mostra as matrizes de confusão na distribuição original, número 2 e número 3.

Figura 15 – Matrizes por *Split*.



Fonte: Acervo do Autor.

A Tabela 4 apresenta as principais métricas extraídas dos experimentos executados com a melhor rede, *Efficient*, com pesos por classe 1:2:4. Nela é possível visualizar que o modelo de aprendizado proposto com a rede *Efficient* consegue manter alto *recall* mesmo com distribuições de pacientes diferentes no conjunto de teste, ainda que a distribuição original tenha sido proposta para apresentar um conjunto de teste representativo. Este tipo de teste não foi desempenhado nas redes CNN propostas por Wang, Lin e Wong (2021), além da falta da métrica *area under curve*.

Tabela 4 – Resultados por *Split*

| Split | Acurácia | AUC | F-1 | Normal | Pneumonia | COVID |
|-------|----------|--------|--------|--------|-----------|--------|
| 1º | 93,52% | 99,13% | 93,65% | 90,00% | 93,33% | 94,89% |
| 2º | 93,90% | 98,73% | 93,89% | 95,70% | 89,48% | 96,17% |
| 3º | 94,17% | 99,00% | 94,19% | 94,12% | 91,18% | 97,46% |

Fonte: Acervo do Autor.

5.4 Discussão

Entre os experimentos realizados neste Capítulo, destacam-se os modelos de aprendizado *Efficient* com pesos de classe 1:2:4 e a rede convolucional Dense com pesos de classe 1:2:2. Estas redes conseguiram, através de abordagens diferentes alcançar *recall* elevado para a detecção de COVID, conforme pode ser visto na tabela 3. Em relação ao estado da arte, mantém *recall* aproximado à arquitetura *COVIDNet-CXR4-A* (WANG; LIN; WONG, 2021), modelo mais atual da COVIDNet para a tarefa de classificação multiclasse, como pode ser visto na Tabela 5.

Tabela 5 – Comparativo com o Estado da Arte.

| | Normal | Pneumonia | COVID |
|------------------------|--------|-----------|--------|
| COVIDNet-CXR4-A | 94,00% | 94,00% | 95,00% |
| Dense | 91,00% | 88,57% | 97,44% |
| Efficient | 90,00% | 93,33% | 94,89% |

Fonte: Acervo do Autor.

Para a classificação binária, o modelo de rede CNN *Efficient* ultrapassa o *recall* da melhor abordagem de classificação binária, *COVIDNet-CXR-2* (WANG; LIN; WONG, 2021), como pode ser visto na Tabela 6.

Tabela 6 – Estado da Arte em Classificação Binária.

| | Saudável | Doente |
|-----------------------|----------|--------|
| COVIDNet-CXR-2 | 97,00% | 95,50% |
| Dense | 91,00% | 97,88% |
| Efficient | 90,00% | 98,94% |

Fonte: Acervo do Autor.

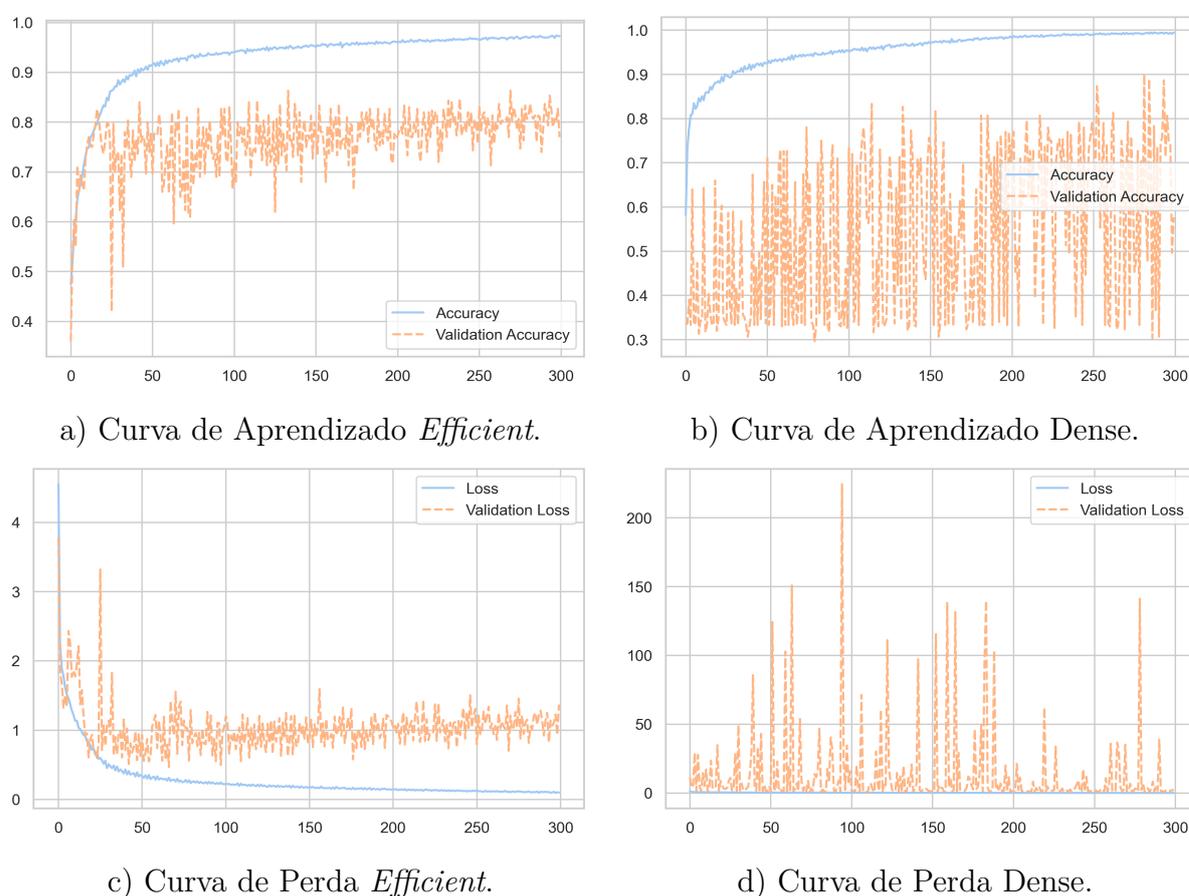
Também, estes resultados apresentam novos referenciais para a distribuição de dados da base COVIDx-CXR 8, visto que durante nossa revisão bibliográfica não foram encontrados experimentos desempenhados na mesma distribuição, com mais amostras de treino e teste.

Neste trabalho ambas as redes utilizadas apresentam um número consideravelmente reduzido de parâmetros quando em contraste à COVIDNet-CXR, que pode ser considerada o estado da arte. A rede Dense proposta, por exemplo, apresenta 1.42 milhões parâmetros treináveis. A rede *Efficient*, 4.87 milhões. Enquanto isso, a rede *COVIDNet-CXR* é composta por 11.75 milhões parâmetros (WANG; LIN; WONG, 2020). O número reduzido de parâmetros pode indicar que o modelo está menos sujeito à *overfit*, que tende a crescer de acordo com o número de parâmetros. Além disso, subsequentes treinos são mais leves dado que algoritmos baseados em gradiente aplicam a regra da cadeia em menos filtros convolucionais.

Nas Figura 16 podem ser observadas as curvas de aprendizado e perda para as redes com melhor desempenho, *Efficient* 1:2:4 e Dense 1:2:2. A partir destas Figuras, pode-se concluir que a utilização do *gradiente descendente estocástico* não afeta de forma negativa os resultados obtidos pela rede *Efficient*, considerando suavidade da curva de aprendizado e da curva de perda.

Para a rede Dense 1:2:2, temos durante o processo de treino, diversos picos na perda. Ainda que o classificador alcance bons resultados, estas perdas aplicam um gradiente elevado e conseqüentemente há perda de informação. Considerando esta ocorrência, há indícios de que ajustes à função e perda e à taxa de aprendizado possam melhorar o processo de treino, em desempenho e agilidade.

Figura 16 – Curvas de Aprendizado e Perda.



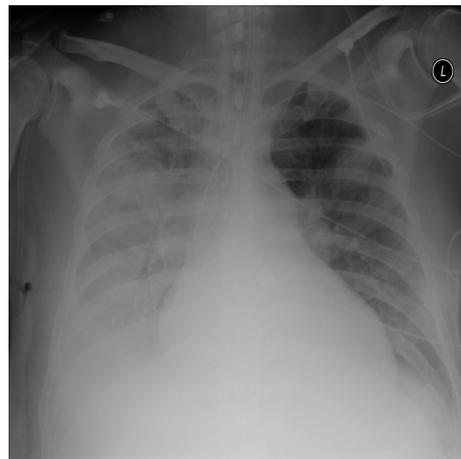
Fonte: Acervo do Autor.

Na Figura 17 estão presentes duas amostras da base de teste. A a) é um falso negativo para pneumonia enquanto a b) é um caso de verdadeiro positivo. A radiografia do falso negativo é um caso difícil de ser classificado, pois toda a estrutura do Pulmão está visível, assim como o coração, indicando que o paciente está em um estado moderado da infecção. Em contraste, vemos que o caso de verdadeiro positivo está opaco ao ponto em que o coração não é visível e sua silhueta se expande à regiões onde não há obstruções.

Figura 17 – Exemplos da classe Pneumonia.



a) Falso Negativo.



b) Positivo Verdadeiro.

Fonte: Acervo do Autor.

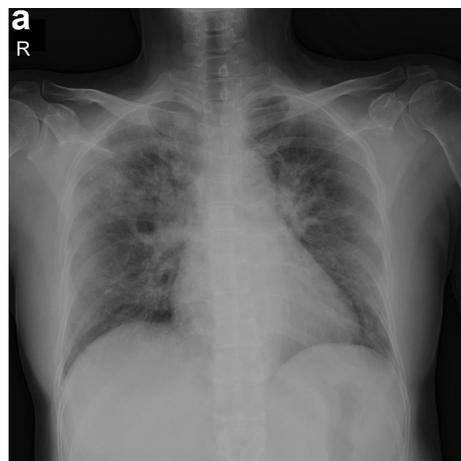
Na Figura 18 estão presentes duas amostras da base de teste. A a) é um falso negativo para Sars-CoV-2 enquanto a b) é um caso de verdadeiro positivo. Com o alto grau de transparência dos alvéolos, não há distinção clara desta amostra para uma saudável. Em casos positivos a opacidade idealmente é visível, como na radiografia b) da Figura 18 ou b) da Figura 17.

Além disso, nosso exemplo de Falso Negativo é uma das imagens da base de dados que não apresenta exposição consistente, que pode ser verificado pela taxa de proporção diferenciada e desalinhamento vertical. Ainda que este não seja o principal fator, deve ser levado em consideração que a exposição diferenciada atrapalhe a localização das regiões onde as obstruções estão presentes.

Figura 18 – Exemplos da classe COVID.



a) Falso Negativo.

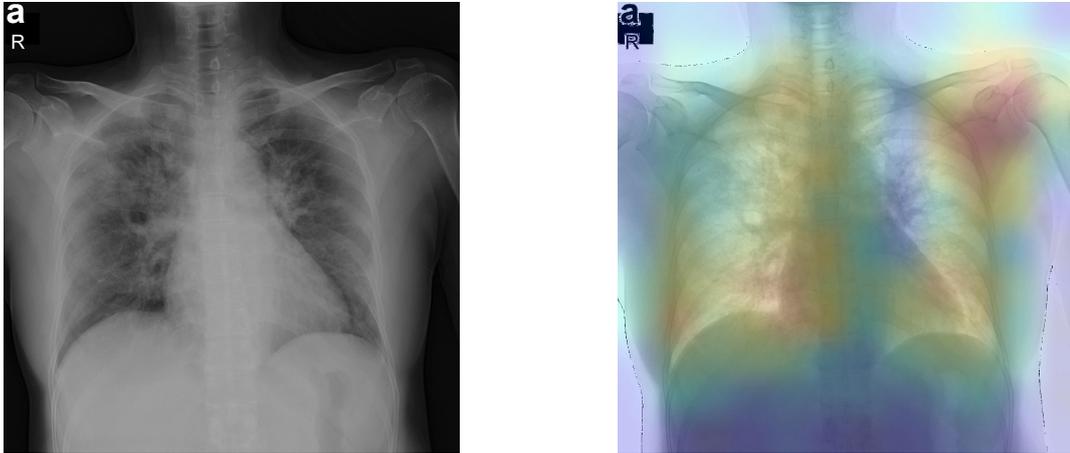


b) Positivo Verdadeiro.

Fonte: Acervo do Autor.

Na Figura 19 pode-se observar o mapa de ativação da rede *Efficient* em um caso de Sars-CoV-2. Na Figura 20, o mapa de ativação da rede em um caso de Pneumonia.

Figura 19 – Mapa de Ativação em Caso de COVID.

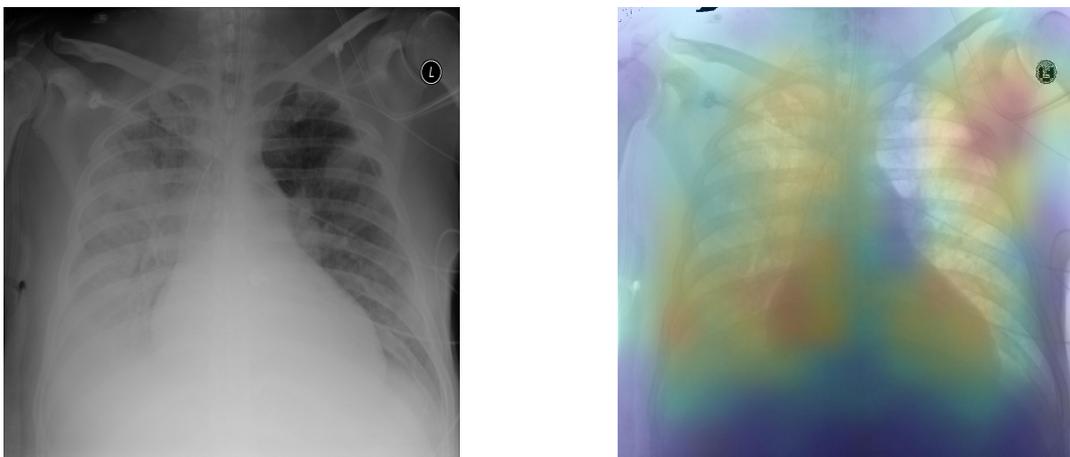


a) Amostra com Sars-CoV-2.

b) Mapa Sobreposto.

Fonte: Acervo do Autor.

Figura 20 – Mapa de Ativação em Caso de Pneumonia.



a) Amostra com Pneumonia.

b) Mapa Sobreposto.

Fonte: Acervo do Autor.

Em ambas as amostras os mapas de ativação se fazem presentes nas regiões onde há sinal de infecção respiratória. Ainda assim, regiões do mapa estão presentes fora do Pulmão, como na clavícula direita, indicando um pequeno grau de superajuste que naturalmente pode ocorrer durante o treino de um classificador.

6 Conclusão

Diagnosticar doenças através de radiografias torácicas é uma prática capaz de acelerar a triagem de pacientes para proporcionar o amplo e ágil atendimento. Devido as circunstancias proporcionadas pela epidemia de Sars-CoV-2, faz-se necessário formas de rápido diagnóstico, através de métodos como RT-PCR. Este trabalho então propõe novas formas de diagnóstico através de modelos de aprendizado de máquina e processamento de imagens.

O modelo proposto foi capaz de alcançar *recall* para COVID elevado em alguns contextos diferentes: multi-classe e de classificação binária. Para o contexto multi-classe, foi possível alcançar *recall* para COVID-19 de 94,89% e 97,44%, como mostra a Tabela 5.

Para a classificação binária, o modelo de rede CNN *Efficient* ultrapassa o *recall* do Estado da Arte para classificação binária, *COVIDNet-CXR-2* (WANG; LIN; WONG, 2021), como pode ser visto na Tabela 6, mesmo estando em seu contexto multi-classe, pois falsos positivos com patologia se confundem entre si e não como casos saudáveis.

Neste trabalho ambas as redes utilizadas apresentam um número consideravelmente reduzido de parâmetros quando em contraste à *COVIDNet-CXR*, trabalho relacionado mais semelhante. Para cada rede temos: *Dense* - 1,42 milhões de parâmetros; *Efficient* - 4,87 milhões de parâmetros; *COVIDNet-CXR* - 11,75 milhões de parâmetros.

O número reduzido de parâmetros indica que o modelo está menos sujeito à *overfit*. Também, que subseqüentes treinos são mais leves dado que algoritmos baseados em gradiente aplicam a regra da cadeia em menos filtros convolucionais. São mais leves e rápidos para a inferência, se situando melhor em ambientes de produção, onde requisitos de *hardware* nem sempre podem ser atendidos.

As propostas de redes CNN apresentadas alcançam resultados que estão aproximados aos modelos propostos de desenvolvidos por Wang, Lin e Wong (2020), oferecendo como diferencial *recall* elevado e modelos até 8.2X mais compactos.

Como trabalhos futuros pretendemos expandir o processo de meta-otimização à rede *Efficient* e *Dense* com a base de dados mais atual, COVIDx-CXR 8, adicionando mais parâmetros em seu espaço de busca. Também, a aplicação de *transfer learning* para a detecção de pneumonia em radiografias torácicas pediátricas, faixa demográfica com altíssima mortalidade para pneumonia (WHO, 2015).

Referências

- ANDRIOLE, K. P. et al. *RSNA Pneumonia Detection Challenge*. Kaggle, 2018. Disponível em: <<https://www.kaggle.com/c/rsna-pneumonia-detection-challenge/overview/description>>. Citado 2 vezes nas páginas 12 e 23.
- BERGSTRA, J. et al. Algorithms for hyper-parameter optimization. *Advances in neural information processing systems*, v. 24, 2011. Citado na página 25.
- BERGSTRA, J.; YAMINS, D.; COX, D. Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures. In: PMLR. *International conference on machine learning*. [S.l.], 2013. p. 115–123. Citado na página 25.
- CHEN, X. et al. Automatic feature learning for glaucoma detection based on deep learning. In: SPRINGER. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. [S.l.], 2015. p. 669–677. Citado na página 10.
- CHUNG, A. et al. *Actualmed COVID Chest X-ray Dataset*. Github, 2021. Disponível em: <<https://github.com/agchung/Actualmed-COVID-chestxray-dataset>>. Acesso em: 2021-17-08. Citado na página 23.
- CHUNG, A. et al. *Figure-1 COVID Chest X-ray Dataset*. Github, 2021. Disponível em: <<https://github.com/agchung/Figure1-COVID-chestxray-dataset>>. Acesso em: 2021-17-08. Citado na página 23.
- COHEN, J. P. et al. Covid-19 image data collection: Prospective predictions are the future. *arXiv preprint arXiv:2006.11988*, 2020. Citado na página 23.
- DENG, J. et al. Imagenet: A large-scale hierarchical image database. In: IEEE. *2009 IEEE conference on computer vision and pattern recognition*. [S.l.], 2009. p. 248–255. Citado 2 vezes nas páginas 16 e 25.
- DONG, E.; DU, H.; GARDNER, L. An interactive web-based dashboard to track covid-19 in real time. *The Lancet infectious diseases*, Elsevier, v. 20, n. 5, p. 533–534, 2020. Citado na página 5.
- DOSOVITSKIY, A. et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020. Citado na página 16.
- DYK, D. A. V.; MENG, X.-L. The art of data augmentation. *Journal of Computational and Graphical Statistics*, Taylor & Francis, v. 10, n. 1, p. 1–50, 2001. Citado na página 19.
- ESTEVA, A. et al. A guide to deep learning in healthcare. *Nature medicine*, Nature Publishing Group, v. 25, n. 1, p. 24–29, 2019. Citado na página 10.
- GUNRAJ, H.; WANG, L.; WONG, A. Covidnet-ct: A tailored deep convolutional neural network design for detection of covid-19 cases from chest ct images. *Frontiers in medicine*, Frontiers Media SA, v. 7, 2020. Citado 2 vezes nas páginas 12 e 13.
- HINTON, G. E. et al. Improving neural networks by preventing co-adaptation of feature detectors. *arXiv preprint arXiv:1207.0580*, 2012. Citado na página 19.

- HOWARD, A. G. et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*, 2017. Citado 3 vezes nas páginas 17, 18 e 28.
- HUANG, G. et al. Densely connected convolutional networks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. [S.l.: s.n.], 2017. p. 4700–4708. Citado 2 vezes nas páginas 16 e 17.
- IANDOLA, F. et al. Densenet: Implementing efficient convnet descriptor pyramids. *arXiv preprint arXiv:1404.1869*, 2014. Citado na página 25.
- IOFFE, S.; SZEGEDY, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In: PMLR. *International conference on machine learning*. [S.l.], 2015. p. 448–456. Citado na página 20.
- KERMANY, D. S. et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*, Elsevier, v. 172, n. 5, p. 1122–1131, 2018. Citado 2 vezes nas páginas 10 e 12.
- KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, v. 25, p. 1097–1105, 2012. Citado na página 10.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *nature*, Nature Publishing Group, v. 521, n. 7553, p. 436, 2015. Citado 2 vezes nas páginas 15 e 16.
- LECUN, Y.; KAVUKCUOGLU, K.; FARABET, C. Convolutional networks and applications in vision. In: *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*. [S.l.: s.n.], 2010. p. 253–256. Citado 2 vezes nas páginas 15 e 16.
- NARKHEDE, S. Understanding auc-roc curve. *Towards Data Science*, v. 26, p. 220–227, 2018. Citado na página 21.
- O'SHEA, K.; NASH, R. An introduction to convolutional neural networks. *arXiv preprint arXiv:1511.08458*, 2015. Citado na página 15.
- POWERS, D. M. Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation. *arXiv preprint arXiv:2010.16061*, 2020. Citado na página 22.
- RAHMAN, T. et al. Exploring the effect of image enhancement techniques on covid-19 detection using chest x-ray images. *Computers in biology and medicine*, Elsevier, v. 132, p. 104319, 2021. Citado na página 23.
- RAJARAMAN, S. et al. Visualization and interpretation of convolutional neural network predictions in detecting pneumonia in pediatric chest radiographs. *Applied Sciences*, Multidisciplinary Digital Publishing Institute, v. 8, n. 10, p. 1715, 2018. Citado 2 vezes nas páginas 12 e 13.
- RAJARATNAM, J. K. et al. Neonatal, postneonatal, childhood, and under-5 mortality for 187 countries, 1970–2010: a systematic analysis of progress towards millennium development goal 4. *The Lancet*, Elsevier, v. 375, n. 9730, p. 1988–2008, 2010. Citado na página 12.

- RAJPURKAR, P. et al. Deep learning for chest radiograph diagnosis: A retrospective comparison of the cheXnet algorithm to practicing radiologists. *PLoS medicine*, Public Library of Science San Francisco, CA USA, v. 15, n. 11, p. e1002686, 2018. Citado na página 12.
- RAJPURKAR, P. et al. Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. *arXiv preprint arXiv:1711.05225*, 2017. Citado na página 12.
- RUSSAKOVSKY, O. et al. Imagenet large scale visual recognition challenge. *International journal of computer vision*, Springer, v. 115, n. 3, p. 211–252, 2015. Citado na página 10.
- TAN, M.; LE, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In: PMLR. *International Conference on Machine Learning*. [S.l.], 2019. p. 6105–6114. Citado 2 vezes nas páginas 25 e 27.
- TSAI, E. B. et al. The rsna international covid-19 open radiology database (ricord). *Radiology*, Radiological Society of North America, v. 299, n. 1, p. E204–E213, 2021. Citado na página 24.
- TURING, A. M. Computing machinery and intelligence. In: *Parsing the turing test*. [S.l.]: Springer, 2009. p. 23–65. Citado na página 14.
- WANG, L.; LIN, Z. Q.; WONG, A. Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. *Scientific Reports*, Nature Publishing Group, v. 10, n. 1, p. 1–12, 2020. Citado 7 vezes nas páginas 10, 13, 23, 24, 28, 34 e 38.
- WANG, L.; LIN, Z. Q.; WONG, A. *COVIDx-CRX 8*. Github, 2021. Disponível em: <<https://github.com/lindawang/COVID-Net>>. Acesso em: 2021-17-08. Citado 4 vezes nas páginas 13, 33, 34 e 38.
- 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. Citado 2 vezes nas páginas 12 e 13.
- WHO. *Pneumonia*. World Health Organization, 2015. Disponível em: <<https://www.who.int/news-room/fact-sheets/detail/pneumonia>>. Acesso em: 2021-17-08. Citado 2 vezes nas páginas 5 e 38.
- WHO. *COVID Dashboard*. World Health Organization, 2021. Disponível em: <<https://covid19.who.int/>>. Acesso em: 2021-17-08. Citado na página 10.