



Trabalho de Conclusão de Curso

Melhorando o Desempenho na Classificação de Imagens Médicas: Uma Análise Comparativa de Técnicas de Aumento de Dados

Phyllipe Matheus B. Alves
pmba@ic.ufal.br

Orientador:
Prof. Dr. Marcelo Costa de Oliveira

Maceió, Junho de 2023

Phyllipe Matheus B. Alves

Melhorando o Desempenho na Classificação de Imagens Médicas: Uma Análise Comparativa de Técnicas de Aumento de Dados

Monografia apresentada como requisito parcial para
obtenção do grau de Bacharel em Ciência da Com-
putação do Instituto de Computação da Universidade
Federal de Alagoas.

Orientador:

Prof. Dr. Marcelo Costa de Oliveira

Maceió, Junho de 2023

Catálogo na fonte
Universidade Federal de Alagoas
Biblioteca Central
Divisão de Tratamento Técnico
Bibliotecária: Taciana Sousa dos Santos – CRB-4 – 2062

A474m Alves, Phyllipe Matheus B.
Melhorando o desempenho na classificação de imagens médicas : uma
análise comparativa de técnicas de aumento de dados / Phyllipe Matheus B.
Alves. – 2023.
35 f. : il. color.

Orientador: Marcelo Costa de Oliveira.
Monografia (Trabalho de Conclusão de Curso em Ciência da
Computação) – Universidade Federal de Alagoas. Instituto de Computação.
Maceió, 2023.

Bibliografia: f. 31-35.

1. Imagens médicas – Classificação. 2. Redes Neurais Convolucionais. 3.
Aumento de dados. I. Título.

CDU: 004.6

Agradecimentos

Este trabalho não seria possível sem a ajuda e apoio de muitas pessoas queridas. Primeiramente, gostaria de agradecer à minha mãe e minha avó, cujo amor e carinho incondicional foram fundamentais ao longo de toda a minha vida. Elas sempre acreditaram em mim e me incentivaram a alcançar meus sonhos e metas, sem julgamentos e sempre com muita compreensão.

Ao meu querido avô, que infelizmente não conseguiu me ver concluir esse ciclo da minha vida, mas que durante a sua passagem nessa vida me aconselhou, guiou e me fez tentar sempre visualizar o lado bom das coisas.

Também gostaria de agradecer à minha noiva, cujo apoio e compreensão durante todo o processo de elaboração deste trabalho foram inestimáveis. Seu amor e apoio me deram a força necessária para continuar trabalhando, mesmo nos momentos mais difíceis.

Não posso esquecer de agradecer as amigas que fiz e cultivei durante o curso. Elas tornaram o dia a dia da faculdade muito mais simples e divertido.

Além disso, quero agradecer aos meus professores, que fizeram crescer a constante sede de conhecimento e interesse pela ciência que tenho hoje, eles foram fonte de inspiração e ensinamentos preciosos que serão levados para o resto da minha vida. Especialmente o meu orientador Prof. Dr. Marcelo Costa de Oliveira que me auxiliou durante toda a minha jornada de construção desse trabalho.

Por fim, gostaria de agradecer a todas as pessoas que, de alguma forma, contribuíram para a realização deste trabalho. O apoio de vocês foi muito valioso e será sempre lembrado por mim.

A todos vocês, muito obrigado!

Resumo

Os métodos de classificação e segmentação de imagens médicas baseados em redes neurais convolucionais (CNNs) possuem o estado da arte quando se trata de acurácia. No entanto, a escassez de dados médicos rotulados disponíveis para o treinamento supervisionado das CNNs pode comprometer sua performance. Para contornar essa limitação, os métodos de aumento de dados têm sido empregados com sucesso. Esses métodos consistem na aplicação de transformações nos dados de entrada, como rotação, translação, escala e espelhamento, a fim de gerar novas imagens sintéticas que possam enriquecer o conjunto de treinamento. Além disso, podem ser aplicadas a diferentes tipos de dados, incluindo aplicações com o uso de diferentes tipos de imagens médicas.

O objetivo desse trabalho é analisar a eficácia de quatro métodos automáticos de aumento de dados: Método base, AugMix, RandAugment e TrivialAugment. Propõe-se assim, a análise da eficácia de um modelo de CNN na categorização de imagens de raio-x de pneumonia por COVID-19, em que o conjunto de dados está categorizado em quatro diferentes classes e será aumentado a partir de quatro cenários de aumento de dados diferentes, ao final dos treinamentos e validações os resultados foram comparados.

Os resultados mostraram que o uso do método AugMix obteve o melhor resultado dentre os métodos avaliados, com F1-Score de 0.910368, cerca de 5% maior quando comparado ao método base. Seguido dos métodos RandAugment e TrivialAugment com valores de F1-Score de 0.896375 e 0.893653 respectivamente.

Além disso, notou-se que em comparação com o método base, os métodos avaliados foram capazes de reduzir o *overfitting* do modelo, melhorando sua capacidade de generalização para novos casos. Espera-se que o resultado desse estudo possam contribuir para a melhor compreensão da importância do aumento de dados, auxiliar na escolha do melhor método de aumento de dados para o domínio de imagens de raio-x de pneumonia por COVID-19 e analisar o impacto de diferentes métodos automáticos de aumento de dados no desempenho de modelos supervisionados de classificação de imagens de pneumonia por COVID-19.

Palavras-chave: processamento gráfico; imagens médicas; visão computacional; aprendizagem profunda; aumento de dados;

Abstract

Classification and segmentation methods for medical image analysis based on convolutional neural networks (CNNs) have achieved state-of-the-art accuracy. However, the scarcity of labeled medical data available for supervised training of CNNs can compromise their performance. To overcome this limitation, data augmentation methods have been successfully employed. These methods involve applying transformations to the input data, such as rotation, translation, scaling, and mirroring, to generate synthetic images that enrich the training set. Additionally, they can be applied to different types of data, including applications involving various types of medical images.

The objective of this study is to evaluate the effectiveness of four automatic data augmentation methods: Baseline, AugMix, RandAugment, and TrivialAugment. We aim to assess the performance of a CNN model in classifying X-ray images of COVID-19 pneumonia, utilizing a dataset divided into four distinct classes. The dataset will be augmented using four different data augmentation scenarios. Subsequently, the results obtained from the training and validation processes will be compared and analyzed.

The results showed that the AugMix method achieved the best performance among the evaluated methods, with an F1-Score of 0.910368, approximately 5% higher than the baseline method. The RandAugment and TrivialAugment methods followed, with F1-Scores of 0.896375 and 0.893653, respectively.

Furthermore, it was noticed that compared to the baseline method, the evaluated methods were able to reduce model overfitting, improving its generalization capability for new cases. It is expected that the results of this study will contribute to a better understanding of the importance of data augmentation, assist in selecting the best data augmentation method for the domain of COVID-19 pneumonia X-ray images, and analyze the impact of different automatic data augmentation methods on the performance of supervised classification models for COVID-19 pneumonia image analysis.

Key-words: graphics processing; medical images; computer vision; deep learning; data augmentation;

Conteúdo

1	Introdução	6
1.1	Objetivos	9
2	Fundamentação Teórica	10
2.1	Trabalhos Relacionados	10
2.1.1	Aumento de dados	10
2.1.2	Aplicação em áreas médicas	10
2.2	Arquitetura do modelo	11
2.3	Aumento de Dados e Técnicas Utilizadas no Trabalho	13
2.3.1	Método 1: Método base	14
2.3.2	Método 2: AugMix	14
2.3.3	Método 3: RandAugment	15
2.3.4	Método 4: TrivialAugment	17
2.4	Métricas de avaliação	18
2.5	Treinamento e validação	19
3	Métodos e Materiais	20
3.1	Dataset de imagens médicas de COVID-19	20
3.2	Pré-processamento das Imagens	21
3.3	Arquitetura do modelo utilizado	22
3.4	Aumento de dados	23
3.4.1	Método 1: Método Base	24
3.4.2	Método 2: AugMix	24
3.4.3	Método 3: RandAugment	24
3.4.4	Método 4: TrivialAugment	25
3.5	Métricas de avaliação	25
4	Resultados Experimentais	26
4.1	Aumento de dados	26
4.1.1	Parâmetros do RandAugment	27
4.2	Análise dos resultados	28
5	Conclusões Finais	30
	Referências bibliográficas	31

1

Introdução

A pandemia de COVID-19, causada pelo vírus SARS-CoV-2, desencadeou uma crise global sem precedentes, afetando milhões de pessoas em todo o mundo. A rápida disseminação do vírus e a gravidade das complicações pulmonares associadas à doença destacaram a necessidade urgente de diagnóstico preciso e eficiente. A radiografia de tórax, uma técnica de imagem amplamente utilizada, tem desempenhado um papel crucial na detecção e no monitoramento da pneumonia por COVID-19 [23, 46]. No entanto, a correta identificação e classificação das imagens de raio-x tornam-se um desafio, especialmente quando se trata de diferenciar a pneumonia por COVID-19 de outras doenças pulmonares com sintomas semelhantes.

A distinção precisa entre a pneumonia por COVID-19 e outras condições pulmonares é fundamental para orientar o tratamento adequado e ajudar na contenção da propagação do vírus. No entanto, a tarefa de classificação computacional de imagens de raio-x apresenta uma série de desafios, principalmente devido à natureza desbalanceada dos conjuntos de dados. A disponibilidade limitada de imagens de raio-x de pneumonia por COVID-19 em comparação com outras doenças pulmonares pode levar a um viés de classificação, afetando a acurácia dos algoritmos de aprendizado de máquina [3].

Redes neurais (RN) têm transformado o campo da análise de imagens médicas nos últimos anos, sendo inspiradas no cérebro humano e capazes de aprender a reconhecer padrões em dados, tornando-as adequadas para tarefas como classificação, segmentação e registro de imagens [32, 19, 44]. Um tipo comum de rede neural utilizado nessa área é a rede neural convolucional (CNN), especialmente indicada para análise de imagens devido à sua capacidade de identificar padrões em dados espaciais, o que a torna ideal para classificação de imagens médicas.

As CNNs têm sido utilizadas para obter resultados de ponta em diversas tarefas de imagem médica. Por exemplo, elas têm sido empregadas para classificar tumores em mamografias [1], diagnosticar retinopatia diabética em imagens de retina [40] e segmentar tumores cerebrais em exames de ressonância magnética (MRI) [35]. Além de sua precisão, as CNNs também são conhecidas pela alta capacidade de generalizar a partir dos dados de treinamento e aplicar o

conhecimento aprendido em exemplos nunca antes vistos. Tornando-as uma escolha favorável para tarefas de classificação em tempo real.

Durante a construção de uma rede neural, na maioria dos casos, é necessário que o treinamento seja realizado com um conjunto de dados extenso e diversificado para que o modelo seja capaz de aprender padrões complexos e extrair características relevantes [3]. No entanto, na área de imagens médicas, existe uma problemática significativa: a escassez de dados rotulados. A obtenção de uma quantidade suficiente de dados rotulados nessa área é desafiadora e pode ser um processo demorado, exigindo tempo e expertise de profissionais da área médica [32, 15]. A falta de dados para treinar modelos de RN em imagens médicas compromete a performance e a capacidade de generalização dos modelos, resultando em limitações na aplicabilidade clínica dessas técnicas [3].

Para resolver o problema de conjuntos de dados desbalanceados no treinamento de redes neurais, diversas técnicas de aumento de dados (do inglês *data augmentation*, DA) foram desenvolvidas para gerar novos dados de treinamento. DA foi inicialmente popularizado por Martin A. Tanner e Wing Hung Wong [37] como uma forma de tornar simulações mais viáveis e simples. Em suma, o aumento de dados consiste em aplicar diversas transformações e operações em um dado original, no intuito de modificá-lo suficientemente, adicionando ruído, variações ou criando novas amostras por meio de modelos generativos. O aumento de dados pode ser usado para reduzir o desbalanceamento de classe, evitar o *overfitting* e, em geral, aumentar a variabilidade dos dados. Para o caso de imagens, por exemplo, transformações como rotação, translação, espelhamento, escala, ajuste de brilho, ajuste de saturação, dentre outras, podem ser aplicadas com o objetivo de gerar imagens sintéticas. (figura 1.1)

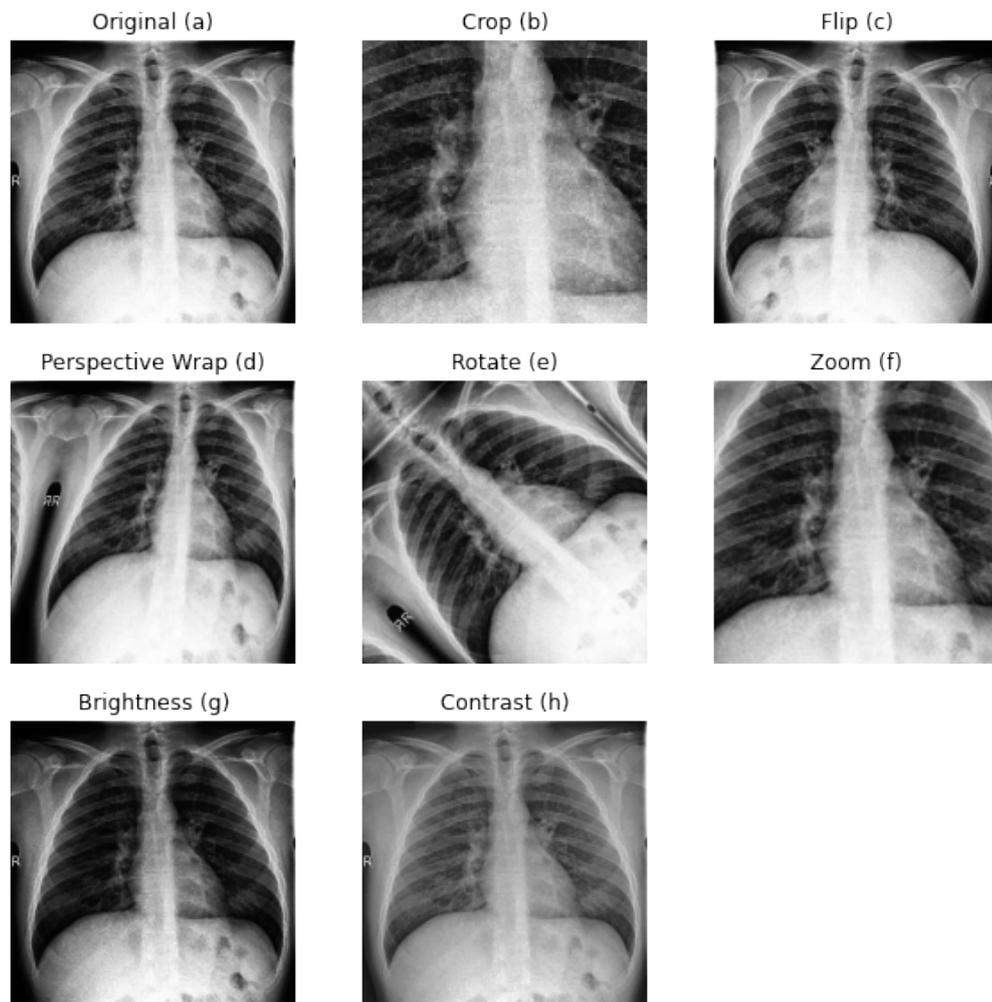


Figura 1.1: Exemplos de imagens geradas (b-h) utilizando diferentes técnicas de aumento de base, a partir imagem original (a).

Embora o aumento de dados venha sendo utilizado amplamente para ampliar o volume de dados de treinamento, a necessidade de configuração manual das transformações e suas intensidades, é relativamente complexa e pode levar a vieses ou escolhas subjetivas no processo. Como consequência, métodos automáticos de aumento de dados baseados em algoritmos que aplicam automaticamente diversas transformações aos dados, surgiram como uma abordagem para resolver esse problema. Em [5, 18] por exemplo, foram apresentados métodos de aumento de dados que encontram automaticamente as melhores políticas de aumento, isto é, transformações e forças das transformações a serem aplicadas no dado original. Contudo, o custo de encontrar políticas ideais está diretamente relacionado ao tamanho e qualidade do conjunto de dados.

1.1 Objetivos

O objetivo deste trabalho foi analisar a eficácia de diferentes métodos de DA aplicados a um conjunto de dados desbalanceado de imagens médicas. Para tal, será utilizado um conjunto de dados contendo imagens de pneumonia por COVID-19.

Este trabalho foi norteado pelas seguintes perguntas de pesquisas:

- PQ1: Qual o melhor método de aumento de dados para o domínio de imagens de raio-x COVID?
- PQ2: Qual é a porcentagem de dados sintéticos que traz o maior benefício para o desempenho de um modelo supervisionado de classificação de imagens de pneumonia de COVID-19?

2

Fundamentação Teórica

2.1 Trabalhos Relacionados

2.1.1 Aumento de dados

Devido ao custo de busca de políticas ideais estar diretamente relacionado com o tamanho do conjunto de dados, métodos de aumento de dados, que reduzem o custo de busca das políticas, surgiram como proposta para otimizar essa etapa.

Cubuk et al. [6], apresentaram o RandAugment, um novo método para aumento de dados que contém um espaço de busca reduzido em comparação a outros métodos e elimina a necessidade do processo separado de busca já mencionado. Seu funcionamento consiste em aplicar transformações escolhidas aleatoriamente a partir de um conjunto de dados pré-definido de operações e aplicar essas transformações no dado original com uma determinada intensidade pré-definida.

Posteriormente, outros métodos de aumento de dados que eliminam a necessidade da etapa de busca foram introduzidos. Em [11] é apresentado o método AugMix, seu método consiste em gerar uma nova imagem a partir da combinação de outras imagens aumentadas, com foco na melhoria do desempenho do classificador e na otimização da perda de consistência do dado. [22] Apresentaram o TrivialAugment, método que elimina a necessidade de ajustes de parâmetros e aplica somente uma transformação a cada imagem de entrada.

2.1.2 Aplicação em áreas médicas

Diversas abordagens inovadoras têm sido propostas para aprimorar a classificação e segmentação de diferentes condições médicas. Em um estudo conduzido por Tang et al. [36], uma abordagem foi apresentada para a classificação de nódulos pulmonares, utilizando técnicas de

umento de dados como rotação e espelhamento aplicadas em um conjunto de dados de exames de tomografia computadorizada de nódulos pulmonares. Os resultados demonstraram uma notável eficácia, com uma área sob a curva (AUC) de 0.971, especificidade de 0.956, sensibilidade de 0.919 e acurácia de 0.936, superando diversas abordagens de classificação de nódulos avançadas atualmente disponíveis.

No estudo conduzido por Müller et al. [21], uma abordagem foi proposta para a segmentação pulmonar e segmentação de lesões de COVID-19. Essa abordagem empregou uma combinação de transformações de imagem, como escala, espelhamento, deformações, rotações, mudança de contraste e adição de ruído Gaussiano. Ao ser avaliada em um conjunto de dados de exames de tomografia computadorizada de pulmões e lesões de COVID-19, a abordagem demonstrou resultados promissores, com um score Dice de 0.804, sobreposição de IoU de 0.672, sensibilidade de 0.778 e especificidade de 0.999, superando várias abordagens estado da arte para segmentação pulmonar e segmentação de lesões de COVID-19.

Adicionalmente, Soulami et al. [33] propuseram uma abordagem para a classificação de lesões de mama, utilizando técnicas de aumento de rotação e espelhamento aplicadas em um conjunto de dados de lesões de mama. Ao ser avaliada obteve resultados significativos, com uma área sob a curva (AUC) de 0.9, F1-Score de 77.45%, acurácia de 77.78%, sensibilidade de 83.15%, recall de 84.54%, especificidade de 83.15% e precisão de 83.15%, superando as 18 abordagens baseadas em aprendizado profundo mais populares na classificação de lesões de mama.

Yadav e Yadav [43] desenvolveram uma abordagem para o diagnóstico de catarata, que emprega técnicas de aumento de dados, como rotação, espelhamento, recorte e translação aplicadas a um conjunto de dados de imagens oculares. Os resultados obtidos foram promissores, com uma sensibilidade de 96.25%, especificidade de 98.74%, escore F1 de 97.47%, precisão de 96.24%, AUC de 96.25% e acurácia de 97.96%. De acordo com os autores, essa abordagem superou os modelos existentes aplicados ao diagnóstico de catarata.

Portanto, esses estudos evidenciam a relevância e o potencial das abordagens inovadoras propostas para aprimorar a classificação e segmentação de diferentes condições médicas, como nódulos pulmonares, lesões de COVID-19, lesões de mama e diagnóstico de catarata. As técnicas de aumento de dados, como rotação e espelhamento, têm demonstrado eficácia na melhoria dos resultados e superação de abordagens estado da arte em seus respectivos domínios. Esses avanços oferecem perspectivas promissoras para o desenvolvimento de métodos mais precisos e eficientes na prática clínica.

2.2 Arquitetura do modelo

O modelo baseado para esse estudo (DarkCovidNet [28]) possui 17 camadas convolucionais. Essas camadas são constituídas por blocos chamados DarkNet e 3xConv. O bloco DarkNet é

constituído por uma sequência de uma camada convolucional, uma camada de normalização por *batch* e uma camada LeakyReLU. O bloco 3xConv é constituído de uma sequência de três blocos DarkNet.

A operação de normalização por *batch* (*batch normalization*), além de ser um auxílio na redução do tempo de treinamento e na estabilidade do modelo, serve para padronizar os *inputs* do modelo. Já LeakyReLU é uma variação da função de ativação ReLU [Eq. 2.1] que é utilizada na prevenção da eliminação de nós. Em comparação com ReLU, LeakyReLU possui um coeficiente para ser utilizados em casos negativos, prevenindo assim a eliminação de nós por valores zerados nos casos negativos [Eq. 2.2].

$$f(x) = \max(0, x) \quad (2.1)$$

$$f(\alpha, x) = \max(\alpha \cdot x, x) \quad (2.2)$$

Observando a figura 2.1 torna-se mais evidente a diferença entre ambos. A função LeakyReLU também varia a depender do valor de α [Eq. 2.2].

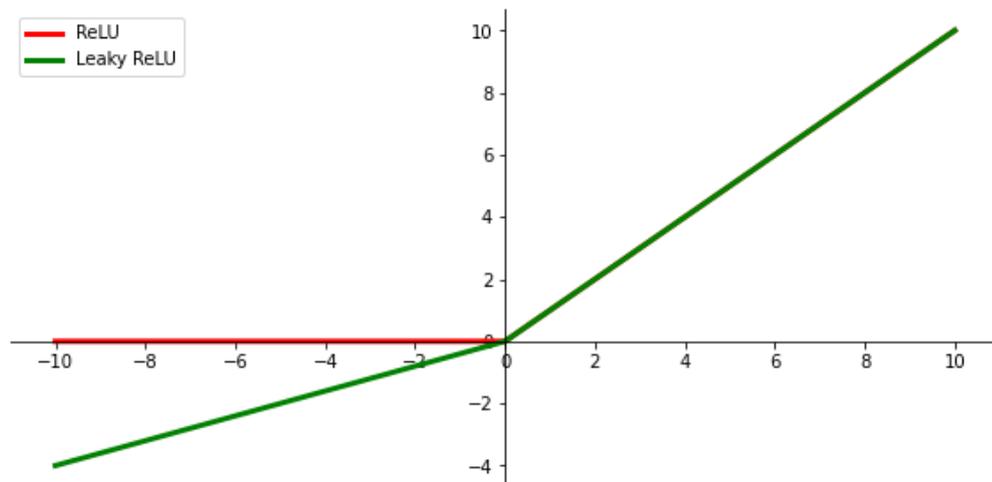


Figura 2.1: Comparação das funções de ativação, Leaky ReLU vs ReLU

O método MaxPooling também é utilizado no modelo em todas as operações de *pooling*, sua funcionalidade consiste em reduzir o tamanho da entrada calculando os valores máximos de uma determinada área definida por um filtro. Na figura 2.2 é possível observar a matriz inicial (*Input*) de tamanho 4x4 ao ser aplicada ao MaxPool com filtro 2x2 uma varredura é feita pela matriz calculando o valor máximo de cada submatriz de tamanho 2x2, resultando assim na matriz de saída (*output*) com os valores máximos.

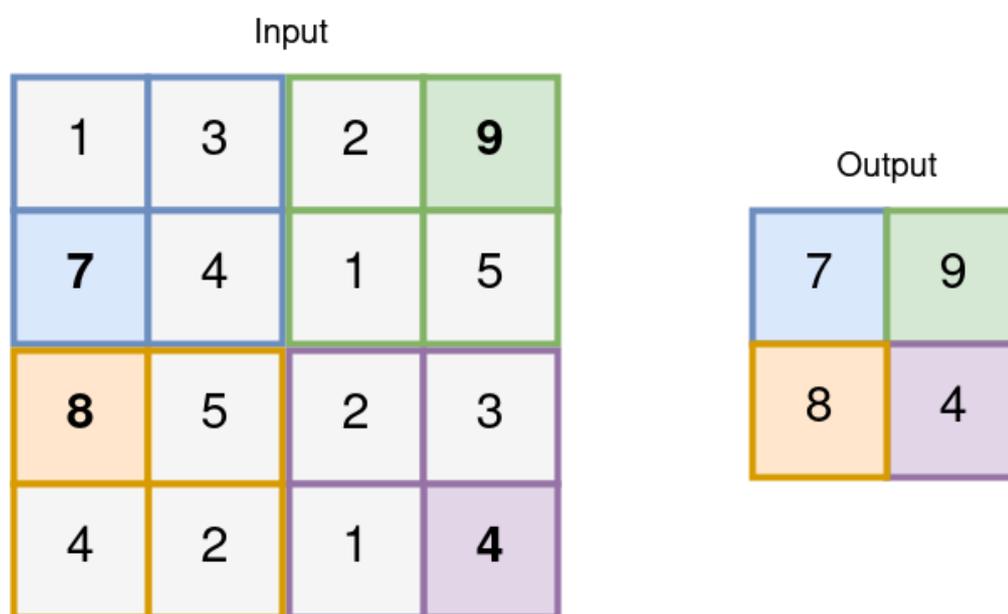


Figura 2.2: Exemplo de funcionamento do método MaxPool

O modelo proposto possui 1.171.076 parâmetros e foi utilizado *cross entropy loss function* como função de custo, no qual compara a probabilidade de classe para cada predição com a classe desejada e a perda é calculada de modo a penalizar baseado na distância da predição para a classe esperada.

2.3 Aumento de Dados e Técnicas Utilizadas no Trabalho

Neste trabalho foram analisados quatro métodos de aumento automático de dados no contexto das imagens de raio-x de COVID-19. O primeiro método (seção 2.3.1) é baseado na biblioteca *fastai* (v1.0.61) e foi selecionado como uma linha de base por adotar uma abordagem aleatória e simplificada. Os métodos 2 e 3 (seções 2.3.2, 2.3.3 e 2.3.4) foram implementados usando a biblioteca *PyTorch* (v0.15) e o método 4 (2.3.4) foi retirado do repositório do Github listado em [22]. Os 3 últimos métodos representam o estado da arte no campo de aumento automático de dados [6, 11, 22].

O principal problema de trabalhar com dados desbalanceados consiste basicamente em como garantir uma acurácia consistente do modelo, tanto para a maioria quanto para a minoria dos dados, tendo em vista que caso a ocorrência de uma classe seja consideravelmente maior do que as outras, o modelo pode vir a se tornar enviesado em suas predições [3].

Neste contexto, é importante mencionar que a determinação da porcentagem ideal de dados aumentados para obter o melhor desempenho no conjunto de dados é uma questão desafiadora. Não existe uma resposta única ou conhecida para qual porcentagem de dados sintéticos é a mais adequada em todos os cenários.

Por esse motivo, o experimento foi realizado com diferentes porcentagens de dados aumentados (10%, 20%, 30%, 40% e 50%) a fim de investigar o impacto dessas variações no desempenho do modelo de classificação. O método base foi utilizado como referência para estabelecer uma base comparativa. Na formula 2.3 é possível calcular a quantidade de dados aumentados (A_s) a partir da quantidade de dados originais (O) e da porcentagem de dados aumentados (p) desejados no conjunto.

$$A_s = \frac{O}{(1-p)} - O \quad (2.3)$$

$$p \in \mathbb{R} \mid 0 < p < 1$$

$$O \in \mathbb{Z} \mid O > 0$$

2.3.1 Método 1: Método base

O primeiro método utiliza a biblioteca *fastai*. Através da função de utilidade "get_transforms"[9] é possível criar facilmente uma lista de transformações para processamento de imagens. Essas transformações incluem operações como espelhamento, rotação, zoom, distorção e ajuste de iluminação. Essa lista de transformações é criada com base nos parâmetros fornecidos, como a probabilidade de aplicação de cada transformação e os intervalos de variação para os parâmetros específicos de cada transformação.

Uma vez criada a lista de transformações, é possível aplicá-la às imagens usando o método "apply_tfms"[8]. Esse método é usado para aplicar as transformações especificadas a um conjunto de imagens ou a uma única imagem. No contexto do *fastai* v1, esses métodos são amplamente utilizados para manipular e processar imagens durante tarefas de visão computacional.

2.3.2 Método 2: AugMix

Seu método consiste em aplicar transformações aleatórias na imagem original, esse processo é repetido várias vezes com novas transformações a cada iteração, onde cada uma das ramificações gera uma nova imagem aumentada única. Depois que todas as ramificações são criadas, as imagens aumentadas são combinadas através de uma mistura ponderada (eq. 2.4, sendo w a força da imagem na mistura) e o resultado da combinação é então misturado com a imagem original utilizando mistura ponderada. A cada nova iteração, o AugMix gera aleatoriamente as transformações a serem aplicadas, as quantidades e intensidades dessas transformações e os pesos das imagens na combinação. Exemplo do método de aumento de dados utilizado no Augmix pode ser visualizado na figura 2.3.

$$\begin{aligned}
 \text{imagem}_{\text{combinada}} &= w * \text{imagem}_1 + (1 - w) * \text{imagem}_2 \\
 w &\in \mathbb{R} \mid 0 < w < 1
 \end{aligned}
 \tag{2.4}$$

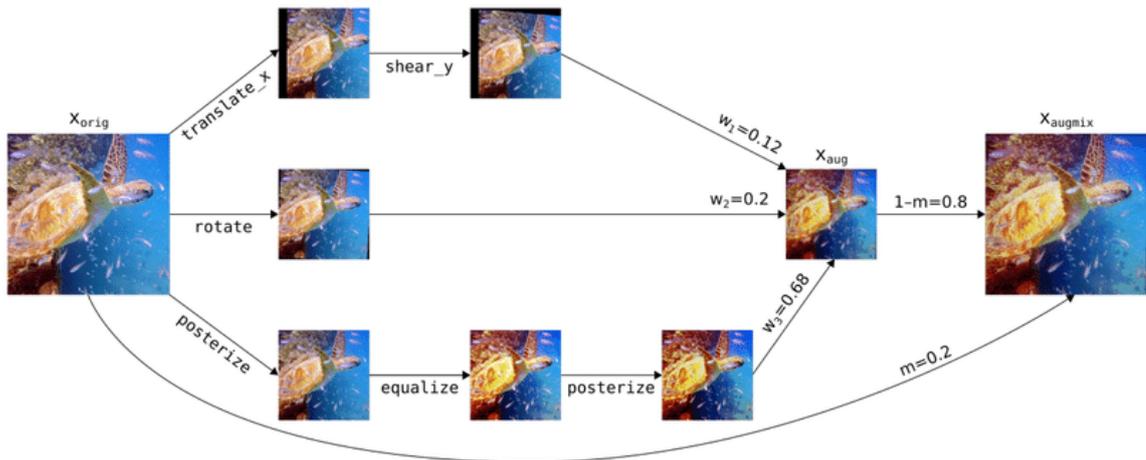


Figura 2.3: Procedimento de aumento de dados utilizando o Augmix, múltiplas transformações são aplicadas de maneira aleatória em imagens de diferentes ramificações, ao final do procedimento as imagens são combinadas com pesos aleatórios.

2.3.3 Método 3: RandAugment

O terceiro método utilizado é o RandAugment [6]. Seu funcionamento consiste em escolher aleatoriamente duas variáveis n e m , onde n é um número inteiro maior que zero e m pertence ao conjunto dos números inteiros, limitado ao intervalo fechado entre 0 e 10. n é definido como a quantidade de transformações para aplicar sequencialmente e m é a força de todas as transformações, sendo 10 o máximo valor possível de força. Para cada imagem a ser aumentada, n transformações são escolhidas de maneira aleatória (t_1, t_2, \dots, t_n) e todas as transformações escolhidas são aplicadas de maneira sequencial a imagem de entrada com intensidade m . Na figura 2.4 é possível observar $n = 2$ transformações aplicadas a imagem original a partir de diferentes intensidades. Todas as transformações possíveis de serem escolhidas são descritas na tabela 2.1.

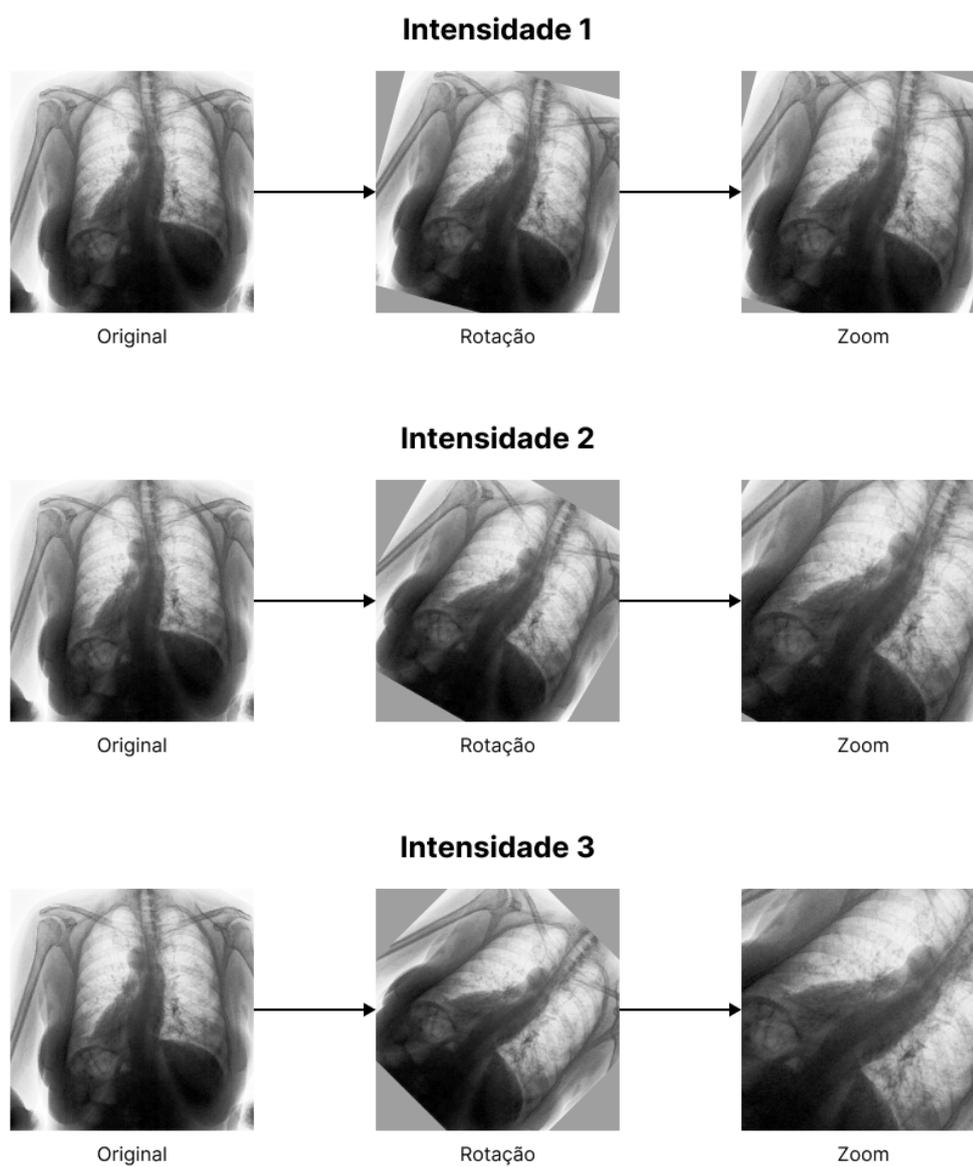


Figura 2.4: Exemplo de $n = 2$ transformações aplicadas com diferentes intensidades (m).

Transformação	Descrição
identify	Não faz nada e mantém a imagem original
autoContrast	Ajusta automaticamente o contraste da imagem, removendo os valores extremos de brilho.
equalize	Equaliza o histograma da imagem, tornando a distribuição de brilho mais uniforme.
rotate	Rotaciona a imagem em um determinado ângulo.
solarize	Inverte as cores da imagem para valores abaixo de um limite de intensidade.
color	Ajusta o equilíbrio de cores da imagem.
posterize	Reduz o número de bits usados para representar cada pixel da imagem.
contrast	Ajusta o contraste da imagem em um valor arbitrário.
brightness	Ajusta o brilho da imagem em um valor arbitrário..
sharpness	Ajusta a nitidez da imagem em um valor arbitrário..
shear-x	Aplica uma distorção na imagem, deslocando-a ao longo do eixo x.
shear-y	Aplica uma distorção na imagem, deslocando-a ao longo do eixo y.
translate-x	Deslocam a imagem ao longo do eixo x.
translate-y	Deslocam a imagem ao longo do eixo y.

Tabela 2.1: Transformações utilizadas pelo método RandAugment e suas descrições.

2.3.4 Método 4: TrivialAugment

O método a ser analisado é o TrivialAugment [22]. Sua abordagem é simples: primeiramente, uma transformação é escolhida aleatoriamente de um conjunto de operações possíveis (tabela 2.2). Em seguida, a força da transformação, representada pela variável m e com m pertencendo aos números inteiros ($m \in \mathbb{Z}$), é selecionada aleatoriamente de um intervalo que varia de 0 a 30. Por fim, essa transformação escolhida é aplicada à imagem de entrada. O pseudocódigo detalhado do método é apresentado no algoritmo 1.

Transformação	Descrição
*	Inclui todas as transformações listadas na tabela 2.1
cutout	Remove aleatoriamente uma seção retangular da imagem.
flip_lr	Inverte horizontalmente a imagem.
flip_ud	Inverte verticalmente a imagem.
sample_pairing	Mistura duas imagens de treinamento em um único exemplo.
smooth	Suaviza a imagem, reduzindo ruídos e irregularidades.
invert	Inverte as cores da imagem.
blur	Suaviza a imagem, borrando as bordas e reduzindo transições bruscas de intensidade.

Tabela 2.2: Transformações utilizadas pelo método TrivialAugment e suas descrições.

Algoritmo 1 Abordagem TrivialAugment

procedure TRIVIALAUGMENT(*img* : imagem)

Sorteia uma transformação α

Sorteia uma força s de $\{0, \dots, 30\}$

Retorna $\alpha(\textit{img}, s)$

▷ Retorna a imagem aumentada

end procedure

2.4 Métricas de avaliação

Neste trabalho foram utilizadas métricas de acurácia, sensibilidade, especificidade e F1-score.

Acurácia é uma medida comum de desempenho em tarefas de classificação, que mede a proporção de predições corretas em relação ao número total de predições. No contexto deste trabalho, a acurácia mede a proporção de imagens de raio-x classificadas corretamente pelo modelo.

Sensibilidade é uma métrica que mede a proporção de exemplos positivos que foram corretamente identificados pelo modelo. No contexto deste trabalho por exemplo, a sensibilidade da classe "Aparência Típica" é definida como todos os casos da classe que foram corretamente classificados como "Aparência Típica".

Especificidade é uma métrica que mede a proporção de exemplos negativos que foram corretamente identificados pelo modelo. No contexto deste trabalho, a especificidade da classe "Aparência Típica" é definida como todos os casos negativos que foram corretamente classificados como não sendo "Aparência Típica".

F1-score é uma média harmônica que combina a especificidade e sensibilidade em uma única medida, que varia de 0 a 1. O F1-score é especialmente útil em tarefas de classificação desequilibradas, em que as classes de interesse têm proporções muito diferentes e foi escolhida

devido a desproporção de dados sintéticos após a etapa de aumento de dados e a baixa, porém existente, probabilidade de que imagens sintéticas iguais possam ser geradas a partir dos eventos aleatórios.

2.5 Treinamento e validação

Para realizar o treinamento do modelo foi utilizado o k-fold. *k-fold cross-validation* é uma técnica comum de validação cruzada que divide o conjunto de dados em k subconjuntos de tamanhos iguais, onde um dos subconjuntos é usado como conjunto de validação e os outros $k - 1$ subconjuntos são usados como conjunto de treinamento.

Esse processo é repetido k vezes, de forma que cada subconjunto seja usado exatamente uma vez como conjunto de validação (fig. 2.5). Dessa forma, é possível avaliar a eficácia do modelo em diferentes conjuntos de dados e evitar problemas de *overfitting* ou *underfitting*. A cada época de treinamento, os pesos do modelo são atualizados com base na função de perda e o modelo é avaliado em um conjunto de validação. O objetivo é encontrar os pesos do modelo que minimizam a função de perda e maximizam a eficácia do modelo em dados não vistos. Ao final do treinamento, as métricas de avaliação são coletadas para cada modelo e usadas para comparar a eficácia dos métodos de aumento de dados em cada cenário.

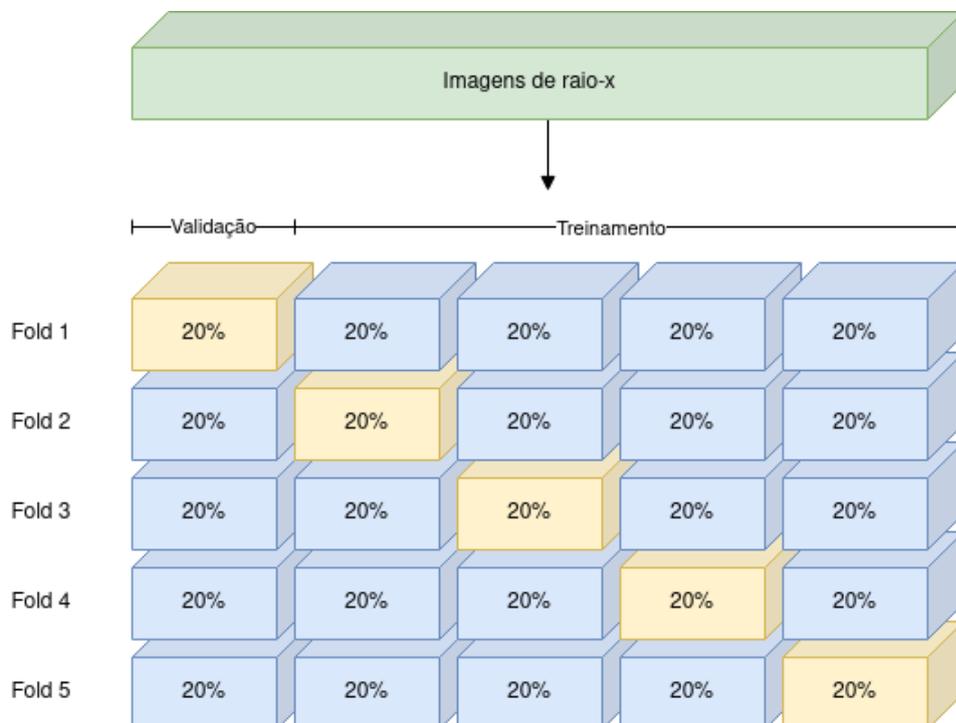


Figura 2.5: Distribuição do treinamento e validação utilizando o procedimento de validação cruzada 5-Fold. Na imagem é possível observar a parcela do conjunto de dados que está sendo utilizada para validação (em amarelo) correspondendo a 20% dos dados do *split* enquanto os outros 80% são utilizados para treinamento (em azul).

3

Métodos e Materiais

3.1 Dataset de imagens médicas de COVID-19

Neste trabalho foram utilizadas imagens de raio-x de COVID-19 disponibilizadas pela *Society for Imaging Informatics in Medicine* (SIIM), onde consta dados de duas fontes de dados públicas, MIDRC-RICORD e BIMCV [41, 39].

A base de imagens possui 6.334 imagens de raio-x do tórax no padrão DICOM, todos os exames foram classificados por radiologistas experientes. Posteriormente as imagens foram anonimizadas e publicadas como parte de um desafio para identificação e localização de anormalidades de COVID-19 [10], exemplos das imagens utilizadas podem ser vistos na Figura 3.1. O conjunto é classificado entre quatro diferentes grupos; Aparência típica, Aparência indeterminada, Aparência atípica e Negativo para pneumonia. As descrições e distribuições das classes contidas no conjunto são apresentados nas tabelas 3.1 e 3.2.

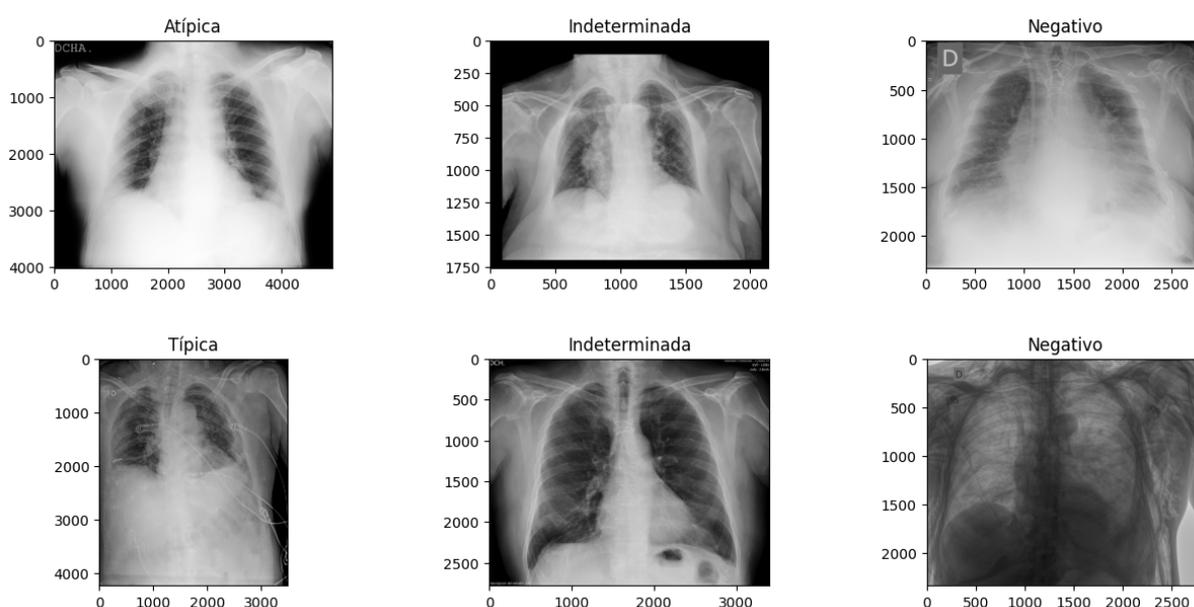


Figura 3.1: Exemplos de imagens contidas no conjunto de dados

Classificação	Descrição
Aparência típica	Aparência típica de pneumonia por COVID-19
Aparência indeterminada	Características não específicas de pneumonia por COVID-19.
Aparência atípica	Características incomuns ou não relatadas de pneumonia por COVID-19.
Negativo para pneumonia	Nenhuma característica de pneumonia

Tabela 3.1: Definição das classes encontradas no conjunto de dados

Classificação	Quantidade de dados	Porcentagem do ocorrência
Aparência típica	3.007	47.5%
Aparência indeterminada	1.736	27.4%
Aparência atípica	1.108	17.5%
Negativo para pneumonia	483	7.6%

Tabela 3.2: Distribuição das classes do conjunto de dados

3.2 Pré-processamento das Imagens

Inicialmente foi observado que os dados dispostos no conjunto de dados não estavam necessariamente padronizados, sendo assim, alguns tratamentos de imagem foram utilizados na padronização. A princípio todas as imagens foram redimensionadas para um tamanho padrão de 255x255 pixels; em seguida, uma técnica de equalização de histograma foi aplicada com o

intuito de balancear os níveis de cinza das imagens; por fim, a resolução de contraste foi normalizada entre o intervalo fechado de $[0, 255]$. Na figura 3.2 é possível ver a comparação visual pré e pós processamento das imagens.

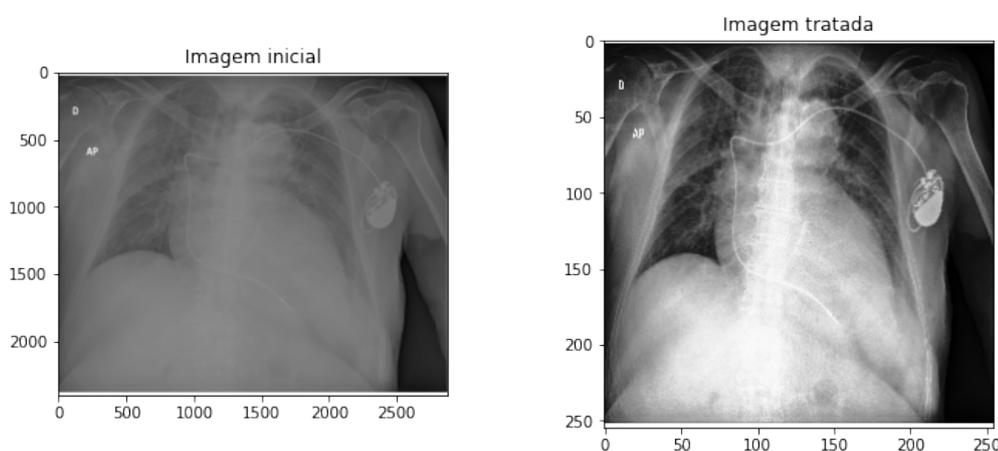


Figura 3.2: Comparação do dado inicial com o dado processado.

Além da padronização e normalização das imagens, o processamento das imagens foi muito útil para melhorar a eficiência do modelo, visto que é computacionalmente menos custoso processar uma imagem de 255×255 pixels em comparação com as resoluções espaciais das imagens de Raio-X (ex.: 2.400×2.880). Por último, uma busca linear foi realizada nos dados para verificar possíveis casos de dados listados nas tabelas que não possuíam uma imagem correspondente, ao total 255 linhas das tabelas foram removidas.

3.3 Arquitetura do modelo utilizado

Neste projeto foi utilizado o DarkCovidNet [28] como modelo de classificação. O modelo é uma variação do modelo DarkNet-19 [31]. O DarkNet-19 é um modelo de classificação fundamental para o sistema de detecção de objetos em tempo real chamado YOLO (*You Only Look Once*) [31], que possui o estado da arte quando se trata de detecção de objetos. [38]

O modelo DarkCovidNet 3.3 possui 17 camadas convolucionais. Dentre essas camadas, existem camadas DarkNet (DN) e 3xConv. A camada DarkNet contém uma camada convolucional (*Convolutional Layer*), seguida de uma normalização por *batch* (*Batch Normalization*) e uma operação LeakyReLU, assim como a camada 3xConv é formada por três camadas DarkNet.

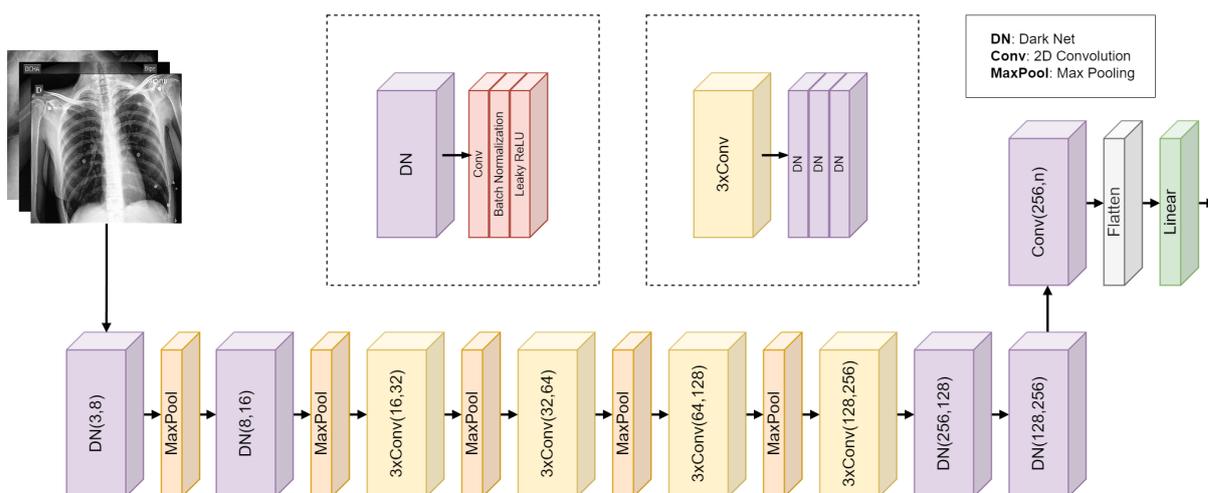


Figura 3.3: Arquitetura do modelo DarkCovidNet

O modelo proposto possui 1.171.076 parâmetros e foi utilizado *cross entropy loss function* como função de custo, no qual compara a probabilidade de classe para cada predição com a classe desejada e a perda é calculada de modo em que uma penalização seja aplicada baseada na distância da predição para a classe esperada.

3.4 Aumento de dados

Ao analisar a distribuição das classes no conjunto de dados fica clara a discrepância na distribuição das classes, na figura 3.4 fica mais claro observar, que da classe com mais ocorrências (Aparência típica) para a classe com menos ocorrências (Negativo para pneumonia) existe uma diferença de 2.524 casos.

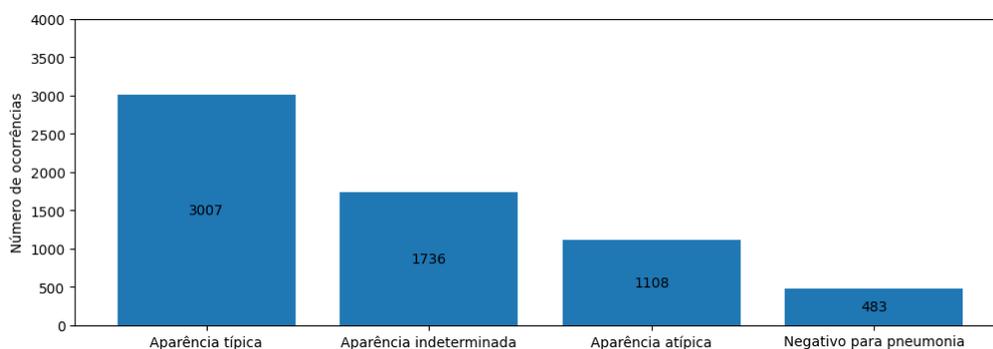


Figura 3.4: Distribuição de classes no conjunto de dados

Para determinar a melhor porcentagem de dados sintéticos, um experimento foi realizado utilizando validação cruzada K-fold, com $k = 10$, durante 100 épocas. 80% dos dados do conjunto de dados foram utilizados para treinamento e 20% para validação, sendo os dados sintéticos utilizados somente no treinamento.

3.4.1 Método 1: Método Base

O método base utiliza a função de utilidade "get_transforms"[9] contida na biblioteca *fastai* (v1.0.61). Ao total 2.603 imagens sintéticas foram geradas, sendo aproximadamente 46% dessas imagens da classe "Negativo para pneumonia", 34% da classe "Aparência Atípica" e 20% da classe "Aparência Indeterminada". Os dados sintéticos foram adicionados ao conjunto de dados original.

Para o treinamento, 80% dos dados foram utilizados como dados de treinamento, contendo dados originais e sintéticos, e 20% dos dados foram utilizados como dados de validação, contendo somente dados originais. O modelo foi treinado durante 100 épocas utilizando validação cruzada k-fold, com $k = 10$, o treinamento foi repetido 3 vezes e a média aritmética foi calculada para apresentação dos resultados.

3.4.2 Método 2: AugMix

O método AugMix utilizado pertence a biblioteca *PyTorch* (v0.15), nenhum dos parâmetros padrões da biblioteca foram alterados para manter sua consistência com método inicialmente apresentado em [11]. Ao total 2.603 imagens sintéticas foram geradas, sendo aproximadamente 46% dessas imagens da classe "Negativo para pneumonia", 34% da classe "Aparência Atípica" e 20% da classe "Aparência Indeterminada". Os dados sintéticos foram adicionados ao conjunto de dados original.

Para o treinamento, 80% dos dados foram utilizados como dados de treinamento, contendo dados originais e sintéticos, e 20% dos dados foram utilizados como dados de validação, contendo somente dados originais. O modelo foi treinado durante 100 épocas utilizando validação cruzada k-fold, com $k = 10$, o treinamento foi repetido 3 vezes e a média aritmética foi calculada para apresentação dos resultados.

3.4.3 Método 3: RandAugment

O método RandAugment utilizado pertence a biblioteca *PyTorch* (v0.15). Durante a etapa de aumento de dados com o método RandAugment foram realizadas 16 rodadas com diferentes parâmetros n e m , quantidade de transformações e intensidade das transformações respectivamente. O intuito de testar com diferentes parâmetros é justamente encontrar o melhor resultado possível para o nosso conjunto de dados. Os parâmetros testados começaram em $n = 2$ e $m = 2$ para divergir do método base que utiliza apenas uma transformação e foram testados até que o resultado obtido fosse aproximado ao resultado do método base $n = 5$ e $m = 5$ (Eq. 3.1).

$$\forall (n, m) \mid n \in [2, 5] \wedge m \in [2, 5] \quad (3.1)$$

Tratando-se de um conjunto de dados de imagens de raio-x, é esperado que muitas trans-

formações com altas intensidades resultem em uma maior dificuldade do modelo em encontrar semelhanças e categorizar as imagens, e conseqüentemente, uma menor acurácia final no treinamento.

Ao total 2.603 imagens sintéticas foram geradas, sendo aproximadamente 46% dessas imagens da classe "Negativo para pneumonia", 34% da classe "Aparência Atípica" e 20% da classe "Aparência Indeterminada". Os dados sintéticos foram adicionados ao conjunto de dados original.

Para o treinamento, 80% dos dados foram utilizados como dados de treinamento, contendo dados originais e sintéticos, e 20% dos dados foram utilizados como dados de validação, contendo somente dados originais. O modelo foi treinado durante 100 épocas utilizando validação cruzada k-fold, com $k = 10$, o treinamento foi repetido 3 vezes e a média aritmética foi calculada para apresentação dos resultados.

3.4.4 Método 4: TrivialAugment

O método TrivialAugment utilizado foi retirado do repositório do Github listado em [22]. Ao total 2.603 imagens sintéticas foram geradas, sendo aproximadamente 46% dessas imagens da classe "Negativo para pneumonia", 34% da classe "Aparência Atípica" e 20% da classe "Aparência Indeterminada". Os dados sintéticos foram adicionados ao conjunto de dados original.

Para o treinamento, 80% dos dados foram utilizados como dados de treinamento, contendo dados originais e sintéticos, e 20% dos dados foram utilizados como dados de validação, contendo somente dados originais. O modelo foi treinado durante 100 épocas utilizando validação cruzada k-fold, com $k = 10$, o treinamento foi repetido 3 vezes e a média aritmética foi calculada para apresentação dos resultados.

3.5 Métricas de avaliação

Para avaliar a eficácia de cada um dos quatro métodos de aumento de dados, as métricas de acurácia, sensibilidade, especificidade e F1-score foram utilizadas para calcular os resultados e ao final as médias aritméticas simples dos valores de acurácia e F1-Score foram apresentadas.



Resultados Experimentais

4.1 Aumento de dados

A partir dos experimentos realizados para determinar a melhor porcentagem de dados sintéticos para o conjunto de dados, foram obtidos valores médios de F1-Score e as quantidades de dados utilizadas nas etapas de treinamento e validação listados na tabela 4.1.

Porcentagem	F1-Score	Nº de dados de treinamento	Nº de dados de validação
10%	0.868660	5402	1351
20%	0.859752	6077	1520
30%	0.862551	6945	1737
40%	0.833725	8103	2026
50%	0.815674	9724	2432

Tabela 4.1: Resultados do experimento com diferentes porcentagens de dados aumentados

Os resultados médios obtidos mostraram que a porcentagem de 10% apresentou o melhor desempenho em relação aos outros experimentos. No entanto, por uma diferença mínima, o experimento com 30% ficou logo atrás. Considerando a maior diversidade de dados presente nesse cenário, a porcentagem de 30% foi escolhida como a mais adequada para o estudo.

Sendo assim, nesse estudo foram geradas 2.603 imagens sintéticas com cada um dos quatro métodos de aumento de dados, resultando em um tamanho final do conjunto de dados de 8.682 casos e reduzindo a diferença de ocorrência entre as classes "Aparência típica" e "Negativo para pneumonia" em aproximadamente 52%. A distribuição de classes após o processo de aumento de dados e a proporção de dados sintéticos para cada dado original são apresentados na tabela 4.2.

Classificação	Dados originais	Dados sintéticos	Proporção média
Aparência típica	2.888	0	0:1
Aparência indeterminada	1.666	517	0.3:1
Aparência atípica	1.058	877	0.8:1
Negativo para pneumonia	467	1209	2.5:1

Tabela 4.2: Detalhes da quantidade de dados originais, aumentados e as respectivas proporções após a etapa de aumento de dados.

4.1.1 Parâmetros do RandAugment

Para o experimento realizado com o intuito de encontrar os melhores parâmetros do método RandAugment, os valores de F1-Score do modelo se comportaram como previsto anteriormente, quanto maior a intensidade e a quantidade de transformações menor o valor de F1-Score do modelo na classificação das imagens. Para os parâmetros (n, m) sendo $(2, 2)$ e $(3, 2)$ foram obtidas os maiores valores, 0.894868 e 0.895066 respectivamente, já para $(5, 5)$ foi obtido o menor valor de todos os cenários, 0.829693. Os valores arredondados de F1-Score obtidos para cada par de parâmetro utilizado podem ser vistos na imagem 4.1.

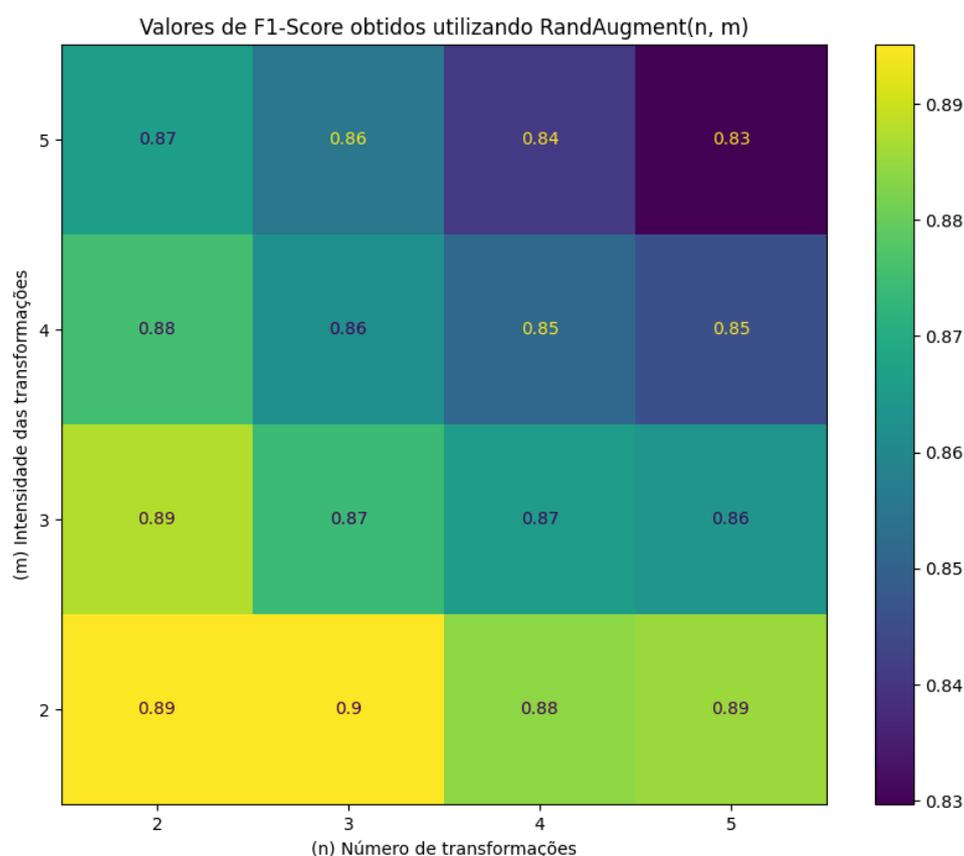


Figura 4.1: F1-Score do modelo ao ser treinado com dados aumentados a partir de cada par de parâmetros (n, m) descritos na equação 3.1 (Valores arredondados).

Com base nos resultados, os valores dos parâmetros n e m para o método RandAugment adotados nesse estudo foram $n = 3$ e $m = 2$.

4.2 Análise dos resultados

A análise dos resultados mostra que o método AugMix obteve o melhor valor de F1-Score (0.910368), seguido pelo método RandAugment (0.896375), e por fim o método TrivialAugment obteve o pior resultado (0.893653). Os valores médios de F1-Score e acurácia para cada método analisado estão disponíveis na tabela 4.2.

Método de DA	Acurácia	F1-Score
Método Base	0.865170	0.862135
TrivialAugment	0.893371	0.893653
RandAugment	0.896314	0.896375
AugMix	0.910225	0.910368

Tabela 4.3: Valores médios de acurácia e F1-Score para cada método de aumento de dados analisado.

No entanto algumas características importantes de cada método devem ser levadas em consideração. Para o método RandAugment, devido aos seus dois parâmetros: n (quantidade de transformações) e m (intensidade das transformações), o ajuste manual para obter as melhores estratégias possíveis para o seus dados de entrada pode se tornar uma etapa consideravelmente custosa e podendo obter tanto resultados eficazes (0.895066 de F1-Score), mas também resultados bem abaixo do método base (0.829693 de F1-Score) (fig. 4.1). Tendo isso em vista, recomendo que uma busca pelos parâmetros seja realizada a partir de uma amostra do conjunto de dados original, visando reduzir o custo de tempo para essa etapa.

O mesmo problema pode também ocorrer ao utilizar o AugMix, porém devido a etapa final de combinação das diferentes imagens transformadas com a imagem original (Figura 2.3) a consistência da imagem acaba permanecendo maior do que em casos onde o RandAugment aplica uma política com intensidade alta (Figura 4.2).

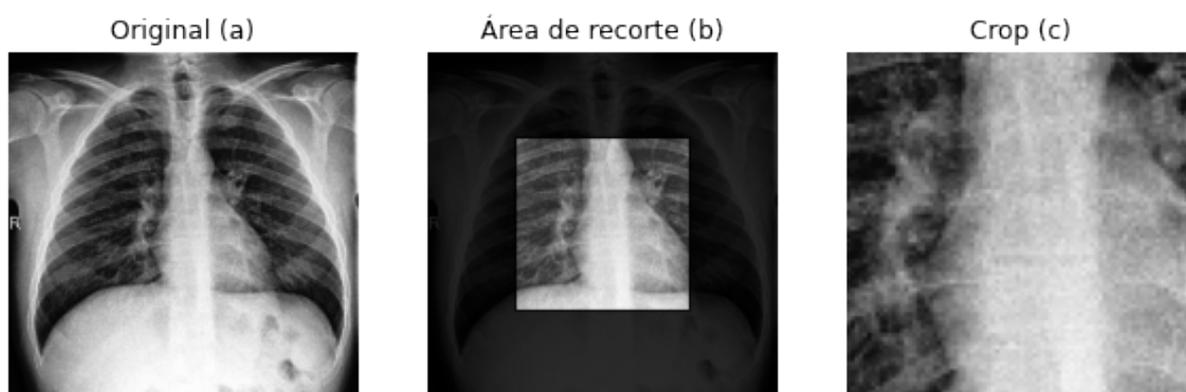


Figura 4.2: Transformação de recorte (*crop*) aplicada a imagem original (a). Na imagem (b) é possível analisar a quantidade de informação perdida ao aplicar o recorte com essa intensidade, aproximadamente 75% da imagem original foi perdida.

Para o TrivialAugment alguns pontos devem ser analisados. A ausência de parâmetros o torna o mais prático dos métodos. Contudo, justamente por não possuir ajuste, o mesmo problema de perda de informação citado anteriormente pode acontecer. Transformações com intensidades altas podem acabar reduzindo consideravelmente a quantidade de informação presente no dado original, resultando em uma perda maior do modelo na etapa de treinamento.

5

Conclusões Finais

Em resumo, o estudo atual examinou a eficácia de várias técnicas de aumento de dados para conjuntos de dados de imagens médicas. Experimentos foram realizados em uma base de dados de imagens de raio-x torácicas, classificadas em quatro categorias distintas e em quatro cenários diferentes. As técnicas de aumento de dados avaliadas foram o método base, AugMix, RandAugment e TrivialAugment.

Os resultados demonstraram que as técnicas de aumento de dados AugMix e RandAugment tiveram melhor desempenho, alcançando valores médios de F1-Score de 0.910368 e 0.896375, respectivamente. Seguido do método TrivialAugment que obteve um F1-Score médio de 0.893653. Todos os três métodos melhoraram a precisão do modelo de reconhecimento de imagens, reduziram o *overfitting* e aumentaram a capacidade de generalização para novos casos quando comparados ao método base.

Tendo isso em vista, recomenda-se o uso das técnicas AugMix e RandAugment para conjuntos de dados de imagens médicas, particularmente em cenários com escassez de dados. No entanto, sugiro que pesquisas futuras explorem outras técnicas de aumento de dados e avaliem a eficácia dessas técnicas em diferentes tipos de conjuntos de dados de imagens médicas.

Em conclusão, este estudo forneceu evidências sólidas sobre a eficácia do aumento de dados como técnica para melhorar o desempenho do reconhecimento de imagens médicas, avaliou a eficácia de três diferentes métodos de aumento automático de dados e destacou a importância de seu uso em futuros estudos nessa área.

Referências bibliográficas

- [1] Dina Abdelhafiz, Clifford Yang, Reda Ammar, and Sheida Nabavi. Deep convolutional neural networks for mammography: advances, challenges and applications. *BMC bioinformatics*, 20:1–20, 2019.
- [2] Nithya C Achaiah, Sindhu B Subbarajasetty, and Rajesh M Shetty. R0 and re of covid-19: can we predict when the pandemic outbreak will be contained? *Indian journal of critical care medicine: peer-reviewed, official publication of Indian Society of Critical Care Medicine*, 24(11):1125, 2020.
- [3] AR Ajiboye, Ruzaini Abdullah-Arshah, and Q Hongwu. Evaluating the effect of dataset size on predictive model using supervised learning technique. 2015.
- [4] Jayme Garcia Arnal Barbedo. Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification. *Computers and electronics in agriculture*, 153:46–53, 2018.
- [5] Ekin D Cubuk, Barret Zoph, Dandelion Mane, Vijay Vasudevan, and Quoc V Le. Autoaugment: Learning augmentation strategies from data. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 113–123, 2019.
- [6] Ekin D. Cubuk, Barret Zoph, Jonathon Shlens, and Quoc V. Le. Randaugment: Practical automated data augmentation with a reduced search space, 2019. URL <https://arxiv.org/abs/1909.13719>.
- [7] Zach Eaton-Rosen, Felix Bragman, Sebastien Ourselin, and M Jorge Cardoso. Improving data augmentation for medical image segmentation. 2018.
- [8] fast.ai. vision.image - fastai, . URL https://fastai1.fast.ai/vision.image.html#Image.apply_tfms.
- [9] fast.ai. vision.transform - fastai, . URL https://fastai1.fast.ai/vision.transform.html#get_transforms.

- [10] Society for Imaging Informatics in Medicine (SIIM). Siim-fisabio-rsna covid-19 detection. URL <https://www.kaggle.com/c/siim-covid19-detection>. Acessado em: 13/08/2022.
- [11] Dan Hendrycks, Norman Mu, Ekin D. Cubuk, Barret Zoph, Justin Gilmer, and Balaji Lakshminarayanan. Augmix: A simple data processing method to improve robustness and uncertainty, 2019. URL <https://arxiv.org/abs/1912.02781>.
- [12] Alex Hernández-García and Peter König. Further advantages of data augmentation on convolutional neural networks. In *Artificial Neural Networks and Machine Learning—ICANN 2018: 27th International Conference on Artificial Neural Networks, Rhodes, Greece, October 4-7, 2018, Proceedings, Part I 27*, pages 95–103. Springer, 2018.
- [13] Zeshan Hussain, Francisco Gimenez, Darvin Yi, and Daniel Rubin. Differential data augmentation techniques for medical imaging classification tasks. In *AMIA annual symposium proceedings*, volume 2017, page 979. American Medical Informatics Association, 2017.
- [14] Our World in Data. Coronavirus (covid-19) vaccinations, 2020. URL https://ourworldindata.org/covid-vaccinations?country=OWID_WRL. Acessado em: 04/09/2022.
- [15] Marc D Kohli, Ronald M Summers, and J Raymond Geis. Medical image data and datasets in the era of machine learning—whitepaper from the 2016 c-mimi meeting dataset session. *Journal of digital imaging*, 30:392–399, 2017.
- [16] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6):84–90, 2017.
- [17] Sungbin Lim, Ildoo Kim, Taesup Kim, Chiheon Kim, and Sungwoong Kim. Fast autoaugment, 2019. URL <https://arxiv.org/abs/1905.00397>.
- [18] Sungbin Lim, Ildoo Kim, Taesup Kim, Chiheon Kim, and Sungwoong Kim. Fast autoaugment. *Advances in Neural Information Processing Systems*, 32, 2019.
- [19] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen Awm Van Der Laak, Bram Van Ginneken, and Clara I Sánchez. A survey on deep learning in medical image analysis. *Medical image analysis*, 42:60–88, 2017.
- [20] Halgurd S Maghdid, Aras T Asaad, Kayhan Zrar Ghafoor, Ali Safaa Sadiq, Seyedali Mirjalili, and Muhammad Khurram Khan. Diagnosing covid-19 pneumonia from x-ray

- and ct images using deep learning and transfer learning algorithms. In *Multimodal image exploitation and learning 2021*, volume 11734, pages 99–110. SPIE, 2021.
- [21] Dominik Müller, Iñaki Soto-Rey, and Frank Kramer. Robust chest ct image segmentation of covid-19 lung infection based on limited data. *Informatics in medicine unlocked*, 25: 100681, 2021.
- [22] Samuel G. Müller and Frank Hutter. Trivialaugument: Tuning-free yet state-of-the-art data augmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 774–782, October 2021.
- [23] Ming-Yen Ng, Elaine YP Lee, Jin Yang, Fangfang Yang, Xia Li, Hongxia Wang, Macy Mei-sze Lui, Christine Shing-Yen Lo, Barry Leung, Pek-Lan Khong, et al. Imaging profile of the covid-19 infection: radiologic findings and literature review. *Radiology: Cardiothoracic Imaging*, 2(1):e200034, 2020.
- [24] World Health Organization. What do we know about sars-cov-2 and covid-19?, 2020. URL https://www.who.int/docs/default-source/coronaviruse/risk-comms-updates/update-28-covid-19-what-we-know-may-2020.pdf?sfvrsn=ed6e286c_2. Acessado em: 03/09/2022.
- [25] World Health Organization. Who coronavirus (covid-19) dashboard, 2022. URL <https://covid19.who.int/>. Acessado em: 29/08/2022.
- [26] World Health Organization. Naming the coronavirus disease (covid-19) and the virus that causes it, 2022. URL [https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-\(covid-2019\)-and-the-virus-that-causes-it](https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it). Acessado em: 26/08/2022.
- [27] World Health Organization. Attacks on health care in the context of covid-19, 2022. URL <https://www.who.int/news-room/feature-stories/detail/attacks-on-health-care-in-the-context-of-covid-19>. Acessado em: 14/08/2022.
- [28] Tulin Ozturk, Muhammed Talo, Eylul Azra Yildirim, Ulas Baran Baloglu, Ozal Yildirim, and U Rajendra Acharya. Automated detection of covid-19 cases using deep neural networks with x-ray images. *Computers in biology and medicine*, 121:103792, 2020.
- [29] Harry Pratt, Frans Coenen, Deborah M Broadbent, Simon P Harding, and Yalin Zheng. Convolutional neural networks for diabetic retinopathy. *Procedia computer science*, 90: 200–205, 2016.

- [30] Roberta Raileanu, Max Goldstein, Denis Yarats, Ilya Kostrikov, and Rob Fergus. Automatic data augmentation for generalization in deep reinforcement learning. *arXiv preprint arXiv:2006.12862*, 2020.
- [31] Joseph Redmon and Ali Farhadi. Yolo9000: better, faster, stronger. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7263–7271, 2017.
- [32] Berkman Sahiner, Aria Pezeshk, Lubomir M Hadjiiski, Xiaosong Wang, Karen Drukker, Kenny H Cha, Ronald M Summers, and Maryellen L Giger. Deep learning in medical imaging and radiation therapy. *Medical physics*, 46(1):e1–e36, 2019.
- [33] Khaoula Belhaj Soulami, Naima Kaabouch, and Mohamed Nabil Saidi. Breast cancer: Classification of suspicious regions in digital mammograms based on capsule network. *Biomedical Signal Processing and Control*, 76:103696, 2022.
- [34] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1):1929–1958, 2014.
- [35] Li Sun, Songtao Zhang, Hang Chen, and Lin Luo. Brain tumor segmentation and survival prediction using multimodal mri scans with deep learning. *Frontiers in neuroscience*, 13: 810, 2019.
- [36] Ning Tang, Rui Zhang, Zeliang Wei, Xicheng Chen, Gaoming Li, Qiuyue Song, Dong Yi, and Yazhou Wu. Improving the performance of lung nodule classification by fusing structured and unstructured data. *Information Fusion*, 88:161–174, 2022.
- [37] Martin A Tanner and Wing Hung Wong. The calculation of posterior distributions by data augmentation. *Journal of the American statistical Association*, 82(398):528–540, 1987.
- [38] Do Thuan. Evolution of yolo algorithm and yolov5: The state-of-the-art object detection algorithm. 2021.
- [39] Emily B Tsai, Scott Simpson, Matthew P Lungren, Michelle Hershman, Leonid Roshkovan, Errol Colak, Bradley J Erickson, George Shih, Anouk Stein, Jayashree Kalpathy-Cramer, et al. The rsna international covid-19 open radiology database (ricord). *Radiology*, 299(1):E204, 2021.
- [40] Carmen Valverde, María García, Roberto Hornero, and Maria I Lopez-Galvez. Automated detection of diabetic retinopathy in retinal images. *Indian journal of ophthalmology*, 64(1):26, 2016.
- [41] Maria De La Iglesia Vayá, Jose Manuel Saborit, Joaquim Angel Montell, Antonio Pertusa, Aurelia Bustos, Miguel Cazorla, Joaquin Galant, Xavier Barber, Domingo

- Orozco-Beltrán, Francisco García-García, et al. Bimcv covid-19+: a large annotated dataset of rx and ct images from covid-19 patients. *arXiv preprint arXiv:2006.01174*, 2020.
- [42] Oliver J Watson, Gregory Barnsley, Jaspreet Toor, Alexandra B Hogan, Peter Winskill, and Azra C Ghani. Global impact of the first year of covid-19 vaccination: a mathematical modelling study. *The Lancet Infectious Diseases*, 22(9):1293–1302, 2022.
- [43] Jay Kant Pratap Singh Yadav and Sunita Yadav. Computer-aided diagnosis of cataract severity using retinal fundus images and deep learning. *Computational Intelligence*, 38(4):1450–1473, 2022.
- [44] Samir S Yadav and Shivajirao M Jadhav. Deep convolutional neural network based medical image classification for disease diagnosis. *Journal of Big data*, 6(1):1–18, 2019.
- [45] Peng Yao, Shuwei Shen, Mengjuan Xu, Peng Liu, Fan Zhang, Jinyu Xing, Pengfei Shao, Benjamin Kaffenberger, and Ronald X Xu. Single model deep learning on imbalanced small datasets for skin lesion classification. *IEEE transactions on medical imaging*, 41(5):1242–1254, 2021.
- [46] Soon Ho Yoon, Kyung Hee Lee, Jin Yong Kim, Young Kyung Lee, Hongseok Ko, Ki Hwan Kim, Chang Min Park, and Yun-Hyeon Kim. Chest radiographic and ct findings of the 2019 novel coronavirus disease (covid-19): analysis of nine patients treated in korea. *Korean journal of radiology*, 21(4):494–500, 2020.
- [47] Hang Yu, Laurence T Yang, Qingchen Zhang, David Armstrong, and M Jamal Deen. Convolutional neural networks for medical image analysis: state-of-the-art, comparisons, improvement and perspectives. *Neurocomputing*, 444:92–110, 2021.
- [48] Amy Zhao, Guha Balakrishnan, Fredo Durand, John V Guttag, and Adrian V Dalca. Data augmentation using learned transformations for one-shot medical image segmentation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 8543–8553, 2019.