



**DESENVOLVIMENTO DE UM MODELO DE MACHINE LEARNING
PARA CLASSIFICAÇÃO DE IMAGENS DE RAIO-X DO TÓRAX
EM DOENÇAS RESPIRATÓRIAS**

**DEVELOPMENT OF A MACHINE LEARNING MODEL FOR
CHEST X-RAY IMAGE CLASSIFICATION IN RESPIRATORY
DISEASES**

**DESARROLLO DE UN MODELO DE APRENDIZAJE
AUTOMÁTICO PARA LA CLASIFICACIÓN DE IMÁGENES DE
RAYOS X DE TÓRAX EN ENFERMEDADES RESPIRATORIAS**

Andrey Alencar Quadros¹
Diogo Ribeiro²
Otacílio Beleza³

DOI: 10.54751/revistafoco.v18n3-134

Received: Feb 28th, 2025

Accepted: Mar 18th, 2025



RESUMO

As doenças respiratórias representam uma das principais causas de morbidade e mortalidade global, destacando a necessidade de diagnósticos rápidos e precisos. Este estudo propõe o desenvolvimento de modelos de machine learning utilizando redes neurais convolucionais (CNNs) para classificar imagens de raio-X do tórax em quatro categorias: COVID-19, normal, pneumonia viral e pneumonia bacteriana. Utilizando as arquiteturas ResNet50 e YOLOv8 pré-treinadas e a técnica de transferência de aprendizagem, os modelos foram adaptados para o contexto específico das radiografias pulmonares. O treinamento foi realizado com conjuntos de dados balanceados e diversas técnicas de pré-processamento e aumento de dados foram aplicadas. Os resultados indicam um desempenho promissor de ambos os modelos, com elevada acurácia na classificação das diferentes patologias, demonstrando o potencial da abordagem para auxiliar profissionais de saúde em ambientes com recursos limitados, como a Unidade de Pronto Atendimento (UPA) de Ariquemes-RO, podendo ser estendida para outras unidades de saúde.

Palavras-chave: Classificação de doenças respiratórias; redes neurais convolucionais; radiografia de tórax; transferência de aprendizagem.

¹ Mestre pelo Programa de Pós-Graduação em Propriedade Intelectual e Transferência de Tecnologia para a Inovação. Instituto Federal de Rondônia (IFRO). Ariquemes, Rondônia, Brasil. E-mail: andrey.quadros@ifro.edu.br

² Pós-Graduado em Inteligência Artificial e Aprendizado de Máquina. Instituto Federal de Rondônia (IFRO). Ariquemes, Rondônia, Brasil. E-mail: diogoifroads@gmail.com

³ Pós-Graduado em Inteligência Artificial e Aprendizado de Máquina. Instituto Federal de Rondônia (IFRO). Ariquemes, Rondônia, Brasil. E-mail: otacilio.system@gmail.com

ABSTRACT

Respiratory diseases represent one of the leading causes of global morbidity and mortality, highlighting the need for rapid and accurate diagnoses. This study proposes the development of machine learning models using convolutional neural networks (CNNs) to classify chest X-ray images into four categories: COVID-19, normal, viral pneumonia, and bacterial pneumonia. By utilizing pre-trained ResNet50 and YOLOv8 architectures and transfer learning techniques, the models were adapted for the specific context of lung radiographs. Training was performed with balanced datasets, and various preprocessing and data augmentation techniques were applied. The results indicate promising performance for both models, with high accuracy in classifying the different pathologies, demonstrating the approach's potential to assist healthcare professionals in resource-limited environments, such as the Emergency Care Unit (UPA) in Ariquemes-RO, and potentially extend to other healthcare facilities.

Keywords: Respiratory disease classification; convolutional neural networks; chest X-ray; transfer learning.

RESUMEN

Las enfermedades respiratorias representan una de las principales causas de morbilidad y mortalidad a nivel mundial, destacando la necesidad de diagnósticos rápidos y precisos. Este estudio propone el desarrollo de modelos de aprendizaje automático utilizando redes neuronales convolucionales (CNN) para clasificar imágenes de rayos X de tórax en cuatro categorías: COVID-19, normal, neumonía viral y neumonía bacteriana. Mediante el uso de arquitecturas preentrenadas ResNet50 y YOLOv8 y técnicas de transferencia de aprendizaje, los modelos fueron adaptados al contexto específico de las radiografías pulmonares. El entrenamiento se realizó con conjuntos de datos balanceados, y se aplicaron diversas técnicas de preprocesamiento y aumento de datos. Los resultados indican un rendimiento prometedor de ambos modelos, con alta precisión en la clasificación de las diferentes patologías, demostrando el potencial del enfoque para asistir a profesionales de la salud en entornos con recursos limitados, como la Unidad de Atención de Emergencias (UPA) de Ariquemes-RO, y con potencial de extensión a otros centros de salud.

Palabras clave: Clasificación de enfermedades respiratorias; redes neuronales convolucionales; radiografía de tórax; transferencia de aprendizaje.

1. Introdução

As doenças respiratórias constituem uma das principais causas de morbidade e mortalidade em todo o mundo, representando um significativo desafio para os sistemas de saúde, especialmente em países em desenvolvimento (World Health Organization, 2017).

A emergência da pandemia de COVID-19 ressaltou a necessidade urgente de diagnósticos rápidos e precisos para doenças pulmonares, a fim de

conter a disseminação e proporcionar tratamento adequado aos pacientes (Huang *et al.*, 2020).

Na Unidade de Pronto Atendimento (UPA) de Ariquemes, Rondônia, os profissionais de saúde enfrentam diariamente um elevado volume de pacientes com diversas condições respiratórias.

A radiografia de tórax é utilizada como ferramenta diagnóstica inicial devido à sua disponibilidade e rapidez (Brasil, 2019). No entanto, a interpretação dessas imagens pode ser subjetiva e exigir conhecimento especializado, o que pode não estar prontamente disponível em ambientes com recursos limitados (Litjens *et al.*, 2017).

Com os avanços recentes em inteligência artificial e aprendizado de máquina, surge a oportunidade de aplicar essas tecnologias na análise automatizada de radiografias digitais, auxiliando na detecção e classificação de doenças pulmonares (Litjens *et al.*, 2017).

A utilização de técnicas de deep learning tem demonstrado desempenho promissor na identificação de padrões complexos em radiografias digitais, possibilitando a diferenciação entre condições como pneumonia viral, pneumonia bacteriana, COVID-19 e pulmões normais (Wang & Wong, 2020).

Diante desse contexto, este trabalho propõe o desenvolvimento de modelos de machine learning para classificar imagens de raio-X do tórax em quatro categorias: COVID-19, normal, pneumonia viral e pneumonia bacteriana.

O objetivo é criar uma ferramenta que auxilie os médicos da UPA de Ariquemes na tomada de decisão, melhorando a eficiência do atendimento e contribuindo para a qualidade dos serviços de saúde prestados à população, podendo ser estendida para outras unidades de saúde.

2. Revisão de Literatura

2.1 Estado Atual da Tecnologia

A aplicação de machine learning na análise de radiografias digitais tem avançado significativamente nas últimas décadas, impulsionada pelo aumento

da capacidade computacional e pela disponibilidade de grandes conjuntos de dados.

Estudos recentes demonstram o potencial dessas técnicas na detecção e classificação de diversas patologias, incluindo doenças respiratórias (Litjens *et al.*, 2017; Esteva *et al.*, 2017).

Especificamente, a pandemia de COVID-19 acelerou pesquisas focadas no uso de aprendizado de máquina no suporte ao diagnóstico por imagem. Modelos têm sido desenvolvidos para identificar sinais de infecção pelo SARS-CoV-2 em radiografias de tórax, visando apoiar a tomada de decisão clínica (Wang & Wong, 2020; Apostolopoulos & Mpesiana, 2020).

Nesse contexto, o campo da visão computacional, através da utilização de redes neurais convolucionais, vem possibilitando resultados significativos nas mais diversas tarefas envolvendo imagem e vídeo.

A visão computacional, inspirada pelo funcionamento do córtex visual dos mamíferos, tem avançado significativamente desde os estudos de David H. Hubel e Torsten Wiesel nos anos 1950 e 1960.

Eles demonstraram que o córtex visual responde a padrões simples, como linhas e bordas, e que essa informação é processada em uma hierarquia de complexidade crescente (Hubel & Wiesel, 1959).

Esses conceitos serviram como base para o desenvolvimento das redes neurais convolucionais (CNNs), cujo principal diferencial é a capacidade de detectar automaticamente padrões em imagens usando camadas convolucionais que simulam esses processos neurais.

Com isso, foi possível ensinar inteligência às máquinas para que elas consigam processar os dados visuais e, assim, ver as coisas exatamente como os humanos (JR & Santos Jr, 2024).

2.2 Técnicas e Algoritmos Existentes

a) Redes Neurais Convolucionais (CNNs)

As CNNs têm se destacado como a abordagem predominante na análise de imagens médicas devido à sua capacidade de extrair características relevantes de forma automática (LeCun *et al.*, 2015).

Estudos como o de Rajpurkar *et al.* (2017) demonstraram que CNNs podem alcançar desempenho comparável ao de radiologistas experientes na classificação de doenças pulmonares em radiografias de tórax.

Uma CNN é composta por várias camadas especializadas que processam e extraem informações da imagem de entrada. No contexto deste estudo, a imagem de raio-X do tórax é inserida na rede e passa por uma série de camadas convolucionais.

Essas camadas atuam como detectores de características, extraindo elementos fundamentais da imagem, como bordas, linhas e texturas (Zeiler & Fergus, 2014).

Essas características básicas são universais e independentes do tipo de imagem, seja uma radiografia de tórax ou qualquer outra.

Conforme a imagem avança pelas camadas da CNN, as características extraídas tornam-se progressivamente mais complexas.

As primeiras camadas focam em aspectos gerais, enquanto as camadas subsequentes combinam essas características para formar representações mais abstratas e específicas, como padrões associados a determinadas patologias pulmonares (Lee *et al.*, 2017).

Esse processo permite que a rede identifique nuances e detalhes que podem ser determinantes para a classificação correta das doenças.

Após a extração de características, a imagem processada é direcionada para camadas totalmente conectadas (fully connected layers), que funcionam como uma rede neural tradicional.

Nessa etapa, a rede utiliza as informações extraídas para realizar a classificação final, determinando a probabilidade de a imagem pertencer a cada uma das categorias pré-definidas (Goodfellow, Bengio & Courville, 2016).

É importante destacar que todo esse processo é automatizado e não requer intervenção manual para a seleção de características específicas.

A capacidade da CNN de aprender e extrair características relevantes diretamente dos dados brutos é uma das razões para seu sucesso em tarefas de visão computacional (Rawat & Wang, 2017).

b) Arquiteturas de CNNs

Várias arquiteturas de CNNs têm sido propostas, destacando-se:

1. **LeNet-5:** Um marco na história das CNNs, utilizada para reconhecimento de dígitos manuscritos. Foi uma das primeiras redes a introduzir camadas convolucionais e camadas de pooling, mostrando o poder dessas camadas para a redução dimensional (LeCun *et al.*, 1998).
2. **AlexNet:** Popularizou as CNNs ao vencer o desafio ImageNet em 2012, utilizando camadas convolucionais profundas e a função de ativação ReLU (Krizhevsky, Sutskever & Hinton, 2012).
3. **GoogLeNet/Inception:** Introduziu a ideia de módulos de inception, que utilizam filtros de diferentes tamanhos em paralelo, permitindo que a rede capture padrões em várias escalas. A arquitetura também foi otimizada para reduzir o número de parâmetros (Szegedy *et al.*, 2015).
4. **ResNet:** A Rede Residual (ResNet) foi um avanço ao permitir a construção de redes muito profundas com 152 camadas, graças ao uso de conexões de atalho (skip connections), que mitigam o problema de desaparecimento do gradiente (He *et al.*, 2016).
5. **YOLO (You Only Look Once):** Uma das arquiteturas mais rápidas para detecção de objetos, processa toda a imagem de uma vez, o que a torna altamente eficiente para aplicações em tempo real (Redmon *et al.*, 2016).

b.a) ResNet50 e Conexões Residuais

Ao aprofundar o entendimento sobre a ResNet50, é fundamental compreender o problema do gradiente “desaparecido” (vanishing gradient problem), comum em redes profundas (Hochreiter, 1991).

Com o aumento do número de camadas, os gradientes utilizados para atualizar os pesos tornam-se muito pequenos, dificultando o aprendizado.

A ResNet50 supera esse desafio por meio de conexões residuais (skip connections), que permitem que a entrada original de uma camada seja somada à sua saída, facilitando a propagação do gradiente e permitindo que a rede aprenda identidades mapeadas (He *et al.*, 2016).

No contexto da classificação de imagens de raio-X do tórax, a ResNet50 é eficaz na captura de padrões complexos associados a diferentes doenças pulmonares, graças à sua profundidade e à capacidade de aprender representações robustas dos dados.

b.b) Processamento das Imagens de Raio-X nas CNNs

Ao aplicar CNNs às imagens de raio-X do tórax, o processo inicia-se com a inserção da imagem na rede, que passa pelas camadas convolucionais responsáveis pela extração de características iniciais, como bordas e texturas (Lakhani & Sundaram, 2017).

Essas características são fundamentais para identificar estruturas anatômicas e possíveis anomalias.

As camadas de pooling subsequentes realizam o downsampling, reduzindo a dimensionalidade dos mapas de características e resumindo as informações mais relevantes (Scherer, Müller & Behnke, 2010).

Em seguida, a camada de flatten transforma os mapas bidimensionais em um vetor unidimensional, preparando os dados para as camadas totalmente conectadas que efetuam a classificação final.

c) Aprendizado Profundo (Deep Learning)

O aprendizado profundo utiliza múltiplas camadas para modelar abstrações de alto nível nos dados. Técnicas de deep learning têm sido aplicadas com sucesso na segmentação, classificação e detecção de anomalias em imagens médicas (Shen, Wu & Suk, 2017).

As arquiteturas de deep learning mais comuns incluem:

1. Redes Neurais Convolucionais (CNNs): Especialmente eficazes para processamento de imagens e dados com estrutura em grade.
2. Redes Neurais Recorrentes (RNNs) e Long Short-Term Memory (LSTM): Utilizadas para processamento de sequências e dados temporais.
3. Autoencoders: Úteis para redução de dimensionalidade e detecção de anomalias.
4. Redes Generativas Adversariais (GANs): Usadas para geração de dados sintéticos e aumento de dados.

Na área médica, técnicas de deep learning têm sido aplicadas com sucesso em diversas tarefas, incluindo:

1. Segmentação de imagens: Identificação precisa de estruturas anatômicas em imagens médicas (Ronneberger, Fischer & Brox, 2015).
2. Classificação: Diagnóstico automatizado de doenças a partir de imagens médicas (Esteva *et al.*, 2017).
3. Detecção de anomalias: Identificação de padrões atípicos que podem indicar patologias (Schlegl *et al.*, 2017).
4. Previsão de resultados clínicos: Utilizando dados do paciente para prever prognósticos e resposta a tratamentos (Rajkomar *et al.*, 2018).
5. Análise de registros médicos eletrônicos: Extração de informações relevantes e identificação de padrões em grandes volumes de dados textuais (Shickel *et al.*, 2018).

O uso de deep learning na medicina tem mostrado resultados promissores, muitas vezes superando o desempenho de especialistas humanos em tarefas específicas.

Por exemplo, um estudo mostrou que um modelo de deep learning foi capaz de detectar câncer de pele com precisão comparável a dermatologistas certificados (Esteva *et al.*, 2017).

No entanto, é importante notar que o uso de deep learning na medicina também enfrenta desafios, como a necessidade de grandes conjuntos de dados rotulados, a interpretabilidade dos modelos e questões éticas relacionadas à tomada de decisão automatizada em contextos médicos (Topol, 2019).

d) Transferência de Aprendizagem (Transfer Learning)

A transferência de aprendizagem permite utilizar modelos pré-treinados em grandes conjuntos de dados, adaptando-os para tarefas específicas com conjuntos de dados menores (Pan & Yang, 2010).

Esta abordagem tem sido eficaz na área médica, onde a disponibilidade de dados rotulados é frequentemente limitada (Yosinski *et al.*, 2014).

No caso deste estudo, a técnica de transferência de aprendizagem é empregada utilizando-se a arquitetura ResNet50 pré-treinada no conjunto de dados ImageNet, que contém milhões de imagens de diversas categorias (Deng *et al.*, 2009).

Embora o ImageNet não inclua imagens de raio-X, as primeiras camadas da ResNet50 aprendem características gerais de imagens, como bordas, texturas e formas, que são igualmente relevantes para radiografias (Raghu *et al.*, 2019).

c.a) Benefícios da Transferência de Aprendizagem

A transferência de aprendizagem pode ser comparada a um indivíduo que já domina uma habilidade e decide aprender uma nova tarefa relacionada.

Por exemplo, alguém que sabe patinar no gelo pode aprender a esqui com mais facilidade devido às semelhanças entre os esportes.

Da mesma forma, uma rede neural treinada para reconhecer objetos gerais pode ser adaptada para identificar padrões específicos em imagens de raio-X (Shin *et al.*, 2016).

Os benefícios dessa abordagem incluem:

1. **Redução do Tempo de Treinamento:** Aproveitar pesos previamente aprendidos acelera significativamente o processo de treinamento, já que não é necessário treinar a rede desde o início (Tajbakhsh *et al.*, 2016).
2. **Melhoria no Desempenho com Poucos Dados:** A transferência de aprendizagem é especialmente útil quando há escassez de dados rotulados, uma situação comum em imagens médicas (Ghafoorian *et al.*, 2017).

-
3. **Evita Overfitting:** Reutilizar pesos de um modelo pré-treinado ajuda a prevenir o overfitting, melhorando a capacidade de generalização do modelo (Géron, 2021).

c.b) Aplicação na Classificação de Imagens de Raio-X

Na implementação prática, as camadas convolucionais da ResNet50 são mantidas com seus pesos originais, preservando o conhecimento adquirido durante o treinamento no ImageNet.

Apenas as camadas finais, responsáveis pela classificação, são ajustadas ou substituídas para se adequar às novas classes de saída correspondentes às categorias de doenças respiratórias (Mahajan *et al.*, 2018).

Essa adaptação permite que a rede aproveite as características genéricas já aprendidas e se especialize nas especificidades das imagens de raio-X do tórax.

Estudos têm demonstrado que essa abordagem melhora significativamente o desempenho em tarefas de classificação de imagens médicas (Esteva *et al.*, 2017).

2.3 Desafios e Lacunas

Apesar dos avanços, existem desafios na aplicação de machine learning em imagens médicas. A necessidade de grandes volumes de dados é uma das barreiras, especialmente para países emergentes (Haleem, Javaid & Khan, 2020).

Além disso, muitos modelos sofrem com a falta de generalização, apresentando desempenho reduzido quando aplicados a dados de diferentes fontes ou populações (Zech *et al.*, 2018).

A variabilidade nos dados, como qualidade de imagem e diferenças nos protocolos de captura de imagens, pode reduzir a precisão dos modelos quando aplicados em ambientes clínicos diversificados (Sarvamangala & Kulkarni, 2022).

Adicionalmente, existem obstáculos no que concerne à compreensão das conclusões. Não obstante as CNNs serem eficientes na identificação de

enfermidades, frequentemente carecem de explicabilidade em seus sistemas de deliberação, o que pode restringir a aprovação por parte dos especialistas médicos, que necessitam fundamentar prognósticos alicerçados em inteligência computacional (BR, 2024).

Em última análise, a aplicação em ambiente médico de CNNs confronta desafios pragmáticos, tais como o dispêndio computacional substancial e a incorporação dessas inovações na rotina do profissional médico.

A instrução de redes complexas pode ser prolongada e onerosa, e a adequação de protótipos preexistentes ao contexto clínico particular pode demandar adaptações significativas (Mohammed *et al.*, 2024).

Conseqüentemente, embora as CNNs exibam notável promessa para revolucionar o exame de iconografias clínicas, o progresso dessas ferramentas tecnológicas está condicionado a empenhos ininterruptos para transpor tais empecilhos, assegurando que os paradigmas sejam resilientes, decifráveis e empregáveis em múltiplos cenários clínicos.

2.4 Justificativa para o Estudo

Este trabalho busca abordar as lacunas identificadas por meio do desenvolvimento de modelos de machine learning especializados na classificação de imagens de raio-X do tórax em quatro categorias relevantes para o contexto clínico atual: COVID-19, normal, pneumonia viral e pneumonia bacteriana.

Além disso, o estudo visa contribuir para a interpretabilidade dos resultados, fornecendo um ambiente que possa ser integrado ao fluxo de trabalho dos profissionais de saúde de Ariquemes.

Dessa forma, espera-se não apenas melhorar a eficiência diagnóstica, mas também aumentar a confiança na utilização de ferramentas que utilizam inteligência artificial.

3. Metodologia

3.1 Descrição do Conjunto de Dados

Para o desenvolvimento dos modelos, utilizou-se um conjunto de dados composto por imagens de raio-X do tórax obtidas de bases de dados públicas, como a COVID-19 Radiography Database (Chowdhury *et al.*, 2020) e outras fontes disponíveis na literatura.

O conjunto de dados foi balanceado, contendo um número equivalente de imagens para cada uma das quatro categorias: COVID-19, normal, pneumonia viral e pneumonia bacteriana.

CrITÉRIOS de inclusão e exclusão foram aplicados para garantir a qualidade das imagens, excluindo aquelas com artefatos significativos ou baixa resolução.

3.2 Pré-Processamento de Imagens

As imagens foram submetidas a um processo de pré-processamento que incluiu:

- a) **Redimensionamento:** Todas as imagens foram redimensionadas para 256 x 256 pixels para compatibilidade com a arquitetura da ResNet50 (He *et al.*, 2016).
- b) **Normalização:** Os valores de pixel foram normalizados para o intervalo [0, 1], facilitando o processo de aprendizado dos modelos.
- c) **Conversão para Três Canais:** Embora as radiografias sejam originalmente em escala de cinza, as imagens foram convertidas para três canais (RGB) para compatibilidade com os modelos pré-treinados na base ImageNet, que esperam entradas com três canais (Simonyan & Zisserman, 2015).
- d) **Aumento de Dados (Data Augmentation):** Técnicas como rotações, translações, zoom e flips horizontais foram aplicadas para aumentar a diversidade do conjunto de treinamento e reduzir o overfitting (Shorten & Khoshgoftaar, 2019).

3.3 Desenvolvimento dos Modelos

3.3.1 Arquitetura dos Modelos

a) Modelo ResNet50

Para o desenvolvimento do modelo de classificação das imagens de raio-X do tórax, inicialmente, foi escolhida a arquitetura ResNet50, uma rede neural convolucional profunda reconhecida por sua eficiência em diversas tarefas de visão computacional (He *et al.*, 2016).

A ResNet50 é composta por 50 camadas e utiliza conexões residuais para mitigar o problema do gradiente desaparecido em redes profundas, permitindo o treinamento eficaz de modelos com muitas camadas.

Vale ressaltar que, a ResNet50, apesar do nome, tecnicamente possui mais de 50 camadas. O número 50 refere-se às camadas principais (convolução e totalmente conectadas), mas a rede também inclui camadas de normalização em lote (batch normalization) e ativação ReLU, que não são contadas no nome. Se incluirmos todas essas camadas, o número total é maior que 50.

Inicialmente, carregamos o modelo base da ResNet50 com pesos pré-treinados na base de dados ImageNet (Deng *et al.*, 2009), especificando o parâmetro `weights='imagenet'`.

Optamos por definir `include_top=False` para excluir as camadas de classificação originais, uma vez que nosso objetivo é adaptar a rede para classificar quatro categorias específicas de doenças respiratórias.

Dessa forma, as camadas convolucionais pré-treinadas são utilizadas para extrair características gerais das imagens, enquanto uma nova camada densa personalizada é adicionada para realizar a classificação específica (Tan & Le, 2019).

Para evitar o retrabalho de treinamento em camadas já otimizadas e aproveitar o conhecimento prévio do modelo, congelamos os pesos das camadas convolucionais, configurando `trainable=False` em todas as camadas do modelo base (Yosinski *et al.*, 2014).

No entanto, para permitir que o modelo se adapte melhor às características específicas das imagens de raio-X, descongelamos as últimas dez camadas convolucionais, mantendo-as treináveis.

Essa estratégia híbrida permite que o modelo refine as características extraídas nas últimas camadas, adaptando-se às peculiaridades do conjunto de dados atual (Howard & Gugger, 2020).

b) Modelo YOLOv8

Após obter resultados promissores com a arquitetura ResNet50 utilizando o TensorFlow, decidiu-se explorar outras abordagens para avaliar a eficácia de diferentes modelos de deep learning na classificação de imagens de raio-X do tórax.

Nesse contexto, optou-se por testar a rede pré-treinada YOLOv8, conhecida por seu desempenho em tarefas de visão computacional, especialmente em detecção e classificação em tempo real (Jocher *et al.*, 2023).

A YOLOv8 (You Only Look Once version 8) é uma evolução das arquiteturas YOLO anteriores, incorporando melhorias em precisão e velocidade de processamento.

Embora seja tradicionalmente utilizada para detecção de objetos, a YOLOv8 também oferece suporte para tarefas de classificação, tornando-a adequada para o objetivo deste estudo (Redmon *et al.*, 2016; Bochkovskiy, Wang & Liao, 2020).

b.a) Seleção do Modelo e Configuração

Para equilibrar precisão e eficiência computacional, foi escolhido o modelo "m" (médio) da YOLOv8, que oferece um bom compromisso entre desempenho e velocidade (Jocher *et al.*, 2023).

Essa decisão é crucial em contextos clínicos onde a rapidez na obtenção dos resultados pode impactar diretamente o atendimento ao paciente.

b.b) Preparação dos Dados

Foram submetidas 3.500 imagens de cada classe (COVID-19, normal, pneumonia viral e pneumonia bacteriana), totalizando 14.000 imagens para o treinamento do modelo.

Este número ampliado de imagens por classe visa fortalecer a capacidade do modelo em aprender características discriminativas e melhorar sua generalização (Shorten & Khoshgoftaar, 2019).

b.c) Processo de Treinamento

O treinamento do modelo YOLOv8 foi conduzido seguindo as diretrizes padrão para tarefas de classificação.

O conjunto de dados foi dividido em porções de treinamento, validação e teste, garantindo que o modelo fosse avaliado em dados não vistos durante o treinamento.

Hiperparâmetros como taxa de aprendizado, tamanho do batch e número de épocas foram ajustados empiricamente para otimizar o desempenho do modelo (Wilson & Martinez, 2003).

c) Ferramentas e Frameworks Utilizados

A implementação foi realizada utilizando a biblioteca Ultralytics YOLO, que fornece uma interface simplificada para treinamento e inferência com modelos YOLOv8 (Jocher *et al.*, 2023).

A utilização dessa biblioteca facilita a experimentação e permite comparações diretas com o modelo baseado em ResNet50 desenvolvido anteriormente.

3.3.2 Implementação e Treinamento dos Modelos

a) Modelo ResNet50

Após a definição da arquitetura básica utilizando a ResNet50 pré-treinada, procedeu-se à implementação das camadas específicas para o problema em questão.

A saída do modelo base (`base_model.output`) foi utilizada como ponto de partida para a construção do novo modelo, referenciado como `head_model`.

Essa estratégia permite a continuidade do fluxo de dados a partir do término das camadas convolucionais da ResNet50, cuja última camada possui dimensões de $8 \times 8 \times 2048$ (He *et al.*, 2016).

b) Redução da Dimensionalidade e Construção das Camadas Densas

Para reduzir a dimensionalidade da saída das camadas convolucionais e preparar os dados para as camadas totalmente conectadas, aplicou-se uma camada de pooling global, especificamente o AveragePooling2D (Boureau, Ponce & LeCun, 2010).

Em seguida, adicionou-se uma camada Flatten para converter a saída bidimensional em um vetor unidimensional, permitindo a conexão com as camadas densas subsequentes (Goodfellow, Bengio & Courville, 2016).

As camadas densas foram estruturadas com neurônios e funções de ativação adequadas, incluindo camadas de dropout para prevenir overfitting (Srivastava *et al.*, 2014).

c) Construção e Compilação do Modelo Final

O modelo final foi construído integrando as camadas definidas com o modelo base da ResNet50.

A função Model da biblioteca Keras foi utilizada para especificar as entradas e saídas do modelo (Chollet, 2015).

Para compilar o modelo, foram definidos parâmetros como função de perda, otimizador e métricas de avaliação apropriadas (Murphy, 2012; Tieleman & Hinton, 2012).

d) Processo de Treinamento

O treinamento foi realizado utilizando o método fit_generator, com ajustes nos hiperparâmetros e utilização de callbacks como ModelCheckpoint para salvar o melhor modelo (Keras Documentation, 2023).

Durante o treinamento, observou-se a redução gradual da função de perda e o aumento da acurácia, indicando que o modelo estava aprendendo a partir dos dados de treinamento (Goodfellow, Bengio & Courville, 2016).

e) Modelo YOLOv8

O treinamento do modelo YOLOv8 foi conduzido seguindo as diretrizes padrão para tarefas de classificação, utilizando a biblioteca Ultralytics YOLO (Jocher *et al.*, 2023).

O modelo foi treinado por 100 épocas, e os hiperparâmetros foram ajustados para otimizar o desempenho.

Gráficos de perda e precisão foram gerados para monitorar o progresso do treinamento.

3.4 Métricas de Avaliação

As métricas utilizadas para avaliar o desempenho dos modelos foram:

- a) **Acurácia:** Proporção de predições corretas sobre o total de predições realizadas.
- b) **Precisão:** Proporção de verdadeiros positivos sobre o total de predições positivas.
- c) **Sensibilidade (Recall):** Proporção de verdadeiros positivos sobre o total de casos positivos reais.
- d) **F1-Score:** Média harmônica entre precisão e sensibilidade, oferecendo um balanço entre as duas métricas.
- e) **AUC-ROC:** Área sob a curva ROC, avaliando a capacidade do modelo em distinguir entre as classes.

3.5 Desenvolvimento da API e Aplicação Web

Para possibilitar a utilização prática do modelo de machine learning desenvolvido, foi implementada uma API utilizando a linguagem Python e o framework FastAPI.

Além disso, foi desenvolvida uma aplicação web frontend utilizando React e Next.js, que serve como interface para os usuários interagirem com o sistema.

3.5.1 Implementação da API

A API foi desenvolvida em Python utilizando o framework FastAPI, conhecido por sua eficiência e facilidade de uso na construção de APIs RESTful (Medvedev, 2019).

A API é responsável por gerenciar o modelo treinado e fornecer endpoints para:

-
- a) **Gerenciamento de Usuários (CRUD):** Permite criar, ler, atualizar e deletar usuários, controlando o acesso ao sistema.
 - b) **Recebimento de Imagens e Predição:** Recebe imagens de raio-X do tórax enviadas pelos usuários e utiliza o modelo de machine learning para realizar a predição, retornando as probabilidades associadas a cada classe.
 - c) **Registro de Feedback:** Registra o feedback dos usuários sobre a predição realizada, permitindo identificar se a predição foi correta ou não e coletando a classe correta em caso de erro.

Este feedback é armazenado no banco de dados e pode ser utilizado posteriormente para retreinamento do modelo, promovendo aprendizado contínuo e melhoria do desempenho ao longo do tempo.

3.5.2 Desenvolvimento da Aplicação Web (Frontend)

A aplicação web frontend foi desenvolvida utilizando as tecnologias React e Next.js, que são amplamente utilizadas para construir interfaces de usuário interativas e eficientes (Dinku, 2022).

O uso de React permite a criação de componentes reutilizáveis e uma experiência de usuário dinâmica, enquanto o Next.js facilita o roteamento e a renderização do lado do servidor, melhorando o desempenho da aplicação.

As principais funcionalidades e telas da aplicação são:

- a) **Tela de Login:** Permite que os usuários acessem o sistema utilizando suas credenciais.

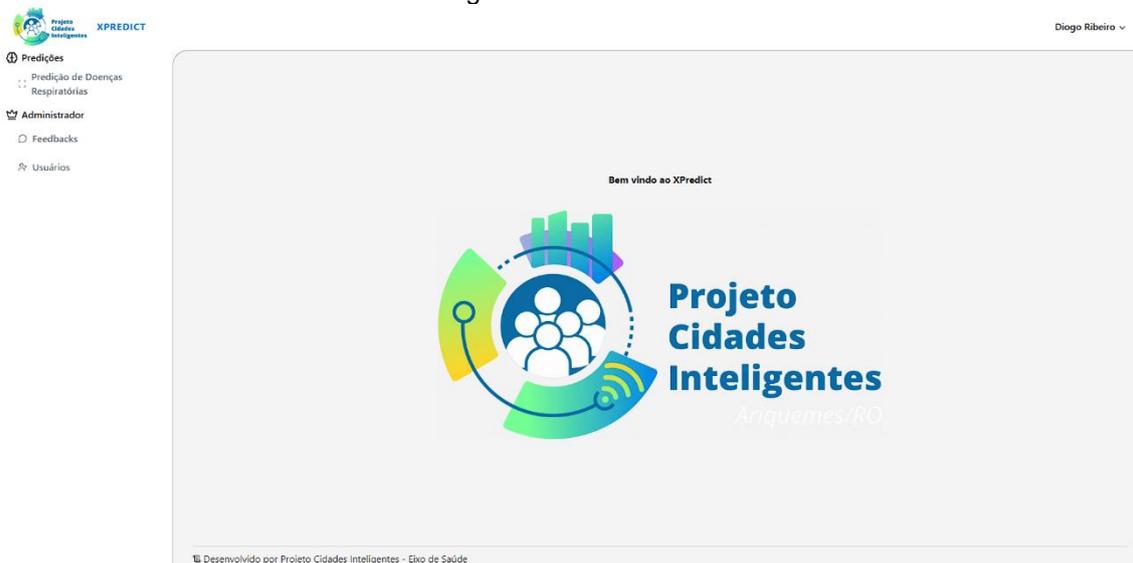
Figura 1 – Tela de Login



Fonte: Os autores (2024)

- b) **Tela Inicial (Home):** Apresenta um resumo das funcionalidades disponíveis e informações gerais.

Figura 2 – Tela Inicial



Fonte: Os autores (2024)

- c) **Tela de Envio de Imagem:** Os usuários podem fazer upload de imagens de raio-X do tórax para serem analisadas pelo modelo.

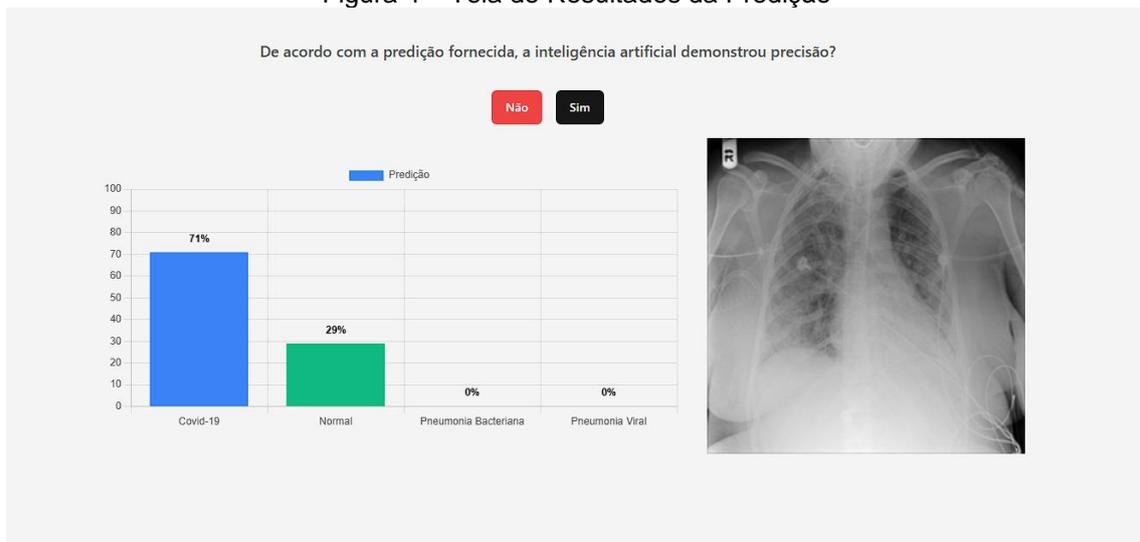
Figura 3 – Tela de Envio de Imagem



Fonte: Os autores (2024)

- d) **Tela de Resultados da Predição:** Exibe o resultado da predição realizada pelo modelo, apresentando um gráfico com as probabilidades associadas a cada classe. Também permite que o usuário indique se a predição foi correta ou não. Em caso negativo, é solicitado que o usuário informe qual seria a predição correta.

Figura 4 – Tela de Resultados da Predição

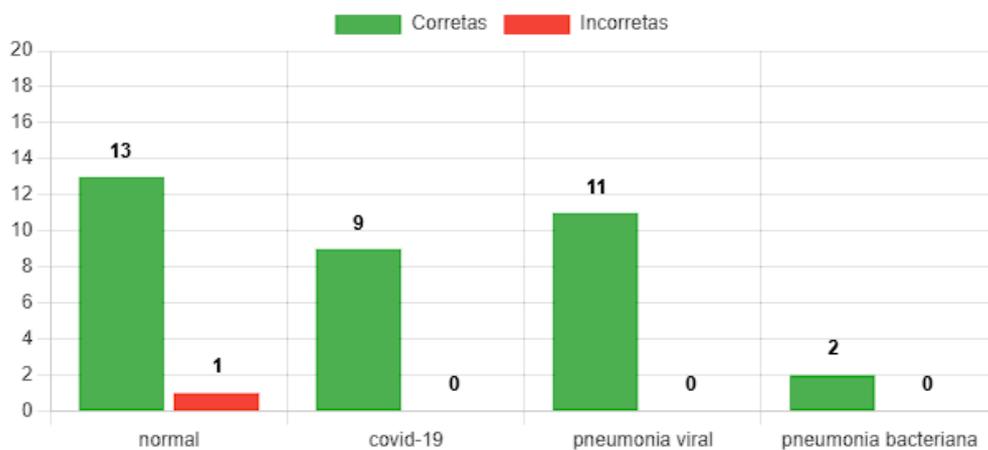


Fonte: Os autores (2024)

Para usuários com perfil de administrador, funcionalidades adicionais estão disponíveis:

- e) **Tela com Gráficos de Erros e Acertos:** Apresenta visualizações gráficas dos desempenhos do modelo, incluindo taxas de erro e acerto ao longo do tempo.

Figura 5 – Tela com Gráficos de Erros e Acertos



Fonte: Os autores (2024)

- f) **Gerenciamento de Usuários:** Permite o controle de usuários, incluindo criação, edição e remoção de contas.

3.5.3 Integração entre a API e o Frontend

A comunicação entre a aplicação web e a API é realizada através de requisições HTTP, utilizando os endpoints disponibilizados pela API.

Ao submeter uma imagem, a aplicação web envia a imagem para o endpoint correspondente, e a API processa a imagem utilizando o modelo treinado, retornando as probabilidades de cada classe. Os resultados são então apresentados ao usuário de forma clara e visual.

O feedback fornecido pelos usuários após a predição é enviado à API e armazenado para análise posterior.

Esta funcionalidade é fundamental para implementar um ciclo de melhoria contínua do modelo, permitindo que dados reais de uso contribuam para o aperfeiçoamento do sistema.

3.5.4 Considerações de Segurança e Privacidade

Foram implementadas medidas de segurança para proteger os dados dos usuários e assegurar a privacidade das informações médicas. Isso inclui:

- a) **Autenticação e Autorização:** Controle de acesso baseado em tokens de autenticação, garantindo que apenas usuários autorizados possam utilizar o sistema.
- b) **Criptografia de Dados:** Utilização de protocolos HTTPS para comunicação segura entre o frontend e a API.
- c) **Conformidade com Regulamentações:** Adequação às leis e regulamentações aplicáveis, como a Lei Geral de Proteção de Dados (LGPD) no Brasil.

3.6 Fluxo de Trabalho do Sistema

- a) **Cadastro e Login:** O usuário se cadastra no sistema e realiza o login através da aplicação web.
- b) **Envio da Imagem:** O usuário faz o upload da imagem de raio-X do tórax que deseja analisar.
- c) **Processamento e Predição:** A imagem é enviada para a API, que utiliza o modelo de machine learning para realizar a predição.
- d) **Apresentação dos Resultados:** Os resultados são apresentados ao usuário na aplicação web, incluindo um gráfico com as probabilidades de cada classe.
- e) **Feedback do Usuário:** O usuário indica se a predição foi correta ou não. Em caso de erro, informa a classe correta.

- f) **Armazenamento e Análise de Dados:** As informações são armazenadas para análise posterior e possível retreinamento do modelo.
- g) **Monitoramento do Desempenho:** Administradores podem visualizar gráficos de erros e acertos para monitorar o desempenho do modelo e identificar áreas de melhoria.

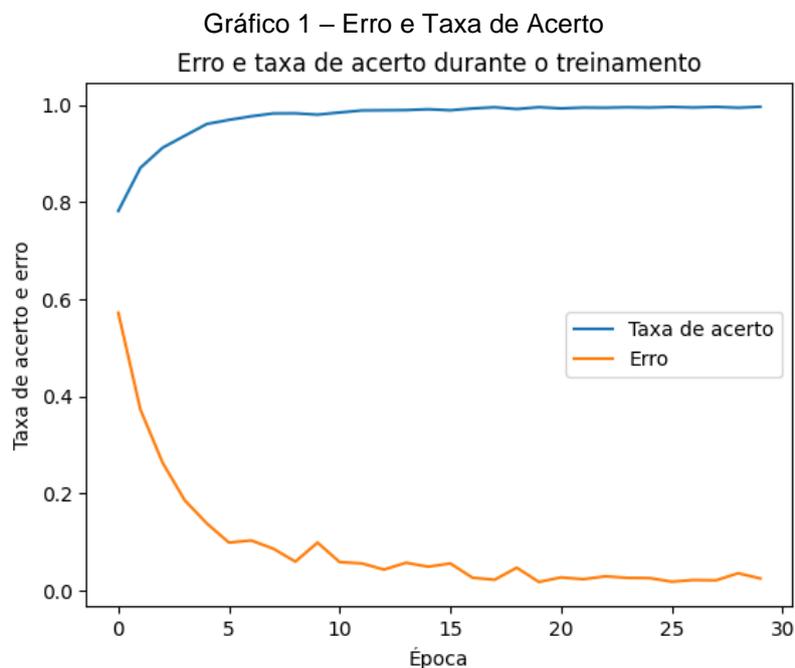
4. Resultados

4.1 Desempenho dos Modelos

4.1.1 Resultados do Modelo ResNet50

O modelo desenvolvido apresentou um desempenho consistente e satisfatório ao longo do treinamento e avaliação.

Para ilustrar a evolução do aprendizado e a eficácia na classificação, foram gerados gráficos de erro (função de perda) e taxa de acerto (acurácia), bem como a matriz de confusão para uma análise detalhada das predições.



Fonte: Os autores (2024)

Gráfico de erro (função de perda) e taxa de acerto (acurácia) do modelo ResNet50 ao longo de 30 épocas de treinamento.

O Gráfico 1 demonstra a evolução da taxa de acerto (linha azul) e do erro (linha laranja) ao longo de 30 épocas de treinamento.

Observa-se que a taxa de acerto aumentou rapidamente nas primeiras épocas, atingindo aproximadamente 90% entre a 5^a e a 10^a época.

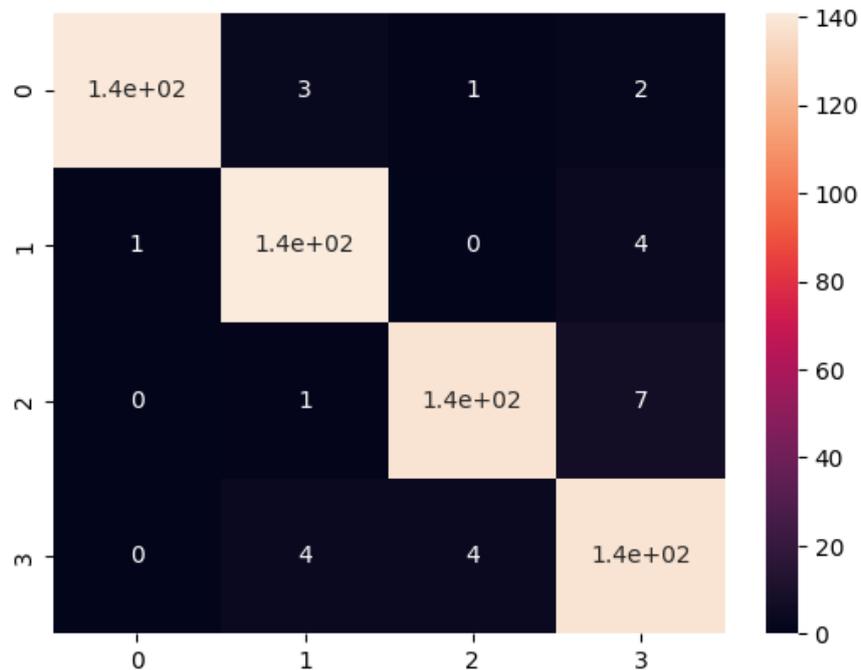
Esse rápido incremento indica que o modelo está aprendendo eficientemente as características relevantes para a classificação das imagens (Goodfellow, Bengio & Courville, 2016).

Concomitantemente, o erro diminuiu de forma consistente, iniciando em torno de 0,6 e caindo rapidamente nas primeiras 5 épocas. Após a 10^a época, o erro estabilizou-se próximo de zero, sugerindo que o modelo alcançou uma boa convergência e minimizou a função de perda efetivamente (Kingma & Ba, 2015).

A estabilização das curvas de taxa de acerto e erro indica que o modelo não está sofrendo de overfitting, ou seja, mantém a capacidade de generalização para dados não vistos durante o treinamento (Srivastava *et al.*, 2014).

Esse comportamento é essencial para aplicações clínicas, onde a robustez do modelo é crucial.

Gráfico 2 – Matriz de Confusão



Fonte: Os autores (2024)

A matriz de confusão fornece uma análise detalhada do desempenho do modelo em cada classe específica.

As classes estão definidas da seguinte forma:

- a) **Classe 0:** COVID-19
- b) **Classe 1:** Normal
- c) **Classe 2:** Pneumonia Viral
- d) **Classe 3:** Pneumonia Bacteriana

1) Classe 0 (COVID-19):

- a) **Predições corretas:** 140 (célula [0,0])
- b) Classificações incorretas como:
 - i) **Classe 1 (Normal):** 3 vezes (célula [0,1])
 - ii) **Classe 2 (Pneumonia Viral):** 1 vez (célula [0,2])
 - iii) **Classe 3 (Pneumonia Bacteriana):** 2 vezes (célula [0,3])

2) Classe 1 (Normal):

- a) **Predições corretas:** 140 (célula [1,1])
- b) Classificações incorretas como:
 - i) **Classe 0 (COVID-19):** 1 vez (célula [1,0])

-
- ii) **Classe 3 (Pneumonia Bacteriana):** 4 vezes (célula [1,3])
 - 3) Classe 2 (Pneumonia Viral):
 - a) **Predições corretas:** 140 (célula [2,2])
 - b) Classificações incorretas como:
 - i) **Classe 1 (Normal):** 1 vez (célula [2,1])
 - ii) **Classe 3 (Pneumonia Bacteriana):** 7 vezes (célula [2,3])
 - 4) Classe 3 (Pneumonia Bacteriana):
 - a) **Predições corretas:** 140 (célula [3,3])
 - b) Classificações incorretas como:
 - i) **Classe 1 (Normal):** 4 vezes (célula [3,1])
 - ii) **Classe 2 (Pneumonia Viral):** 4 vezes (célula [3,2])

A matriz de confusão indica que o modelo ResNet50 possui uma alta taxa de acerto para todas as classes, com 140 predições corretas em cada uma delas.

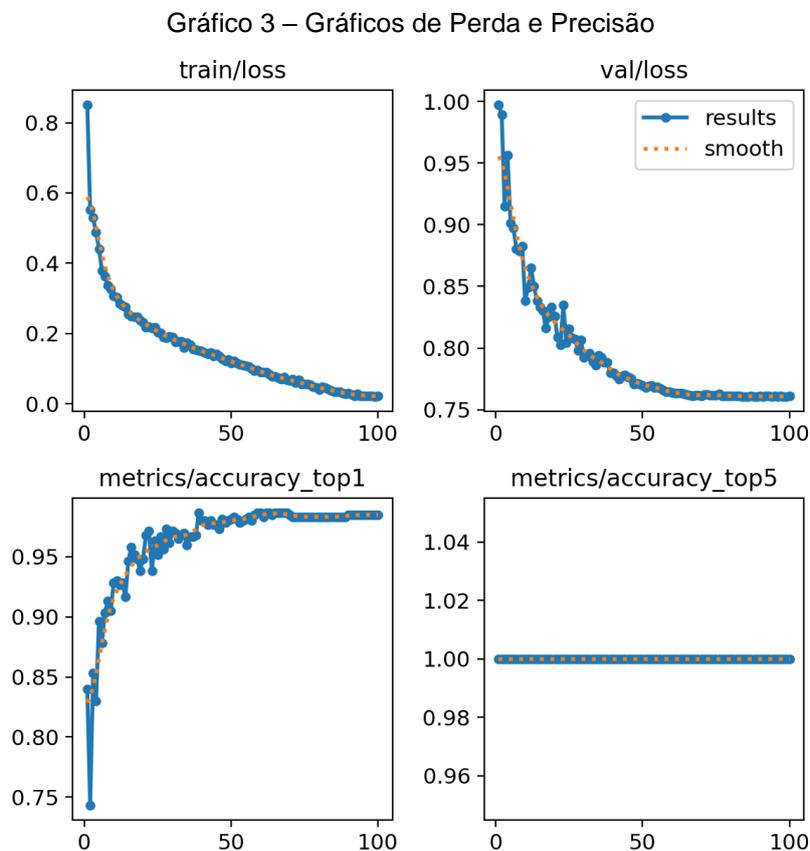
As confusões entre as classes são mínimas, mas algumas observações podem ser destacadas:

- a) **Confusões entre COVID-19 e Normal:** O modelo confundiu a Classe 0 (COVID-19) com a Classe 1 (Normal) em 3 ocasiões, e a Classe 1 com a Classe 0 em 1 ocasião. Isso sugere que, em alguns casos, as características radiográficas de pacientes com COVID-19 podem se assemelhar a pulmões normais, possivelmente devido a estágios iniciais da doença ou manifestações atípicas (Apostolopoulos & Mpesiana, 2020).
- b) **Confusões entre Pneumonia Viral e Pneumonia Bacteriana:** Houve 7 casos em que a Classe 2 (Pneumonia Viral) foi confundida com a Classe 3 (Pneumonia Bacteriana), e 4 casos inversos. Isso é compreensível, dado que ambas as pneumonias podem apresentar sinais radiológicos semelhantes, como opacidades pulmonares e infiltrados (Morales *et al.*, 2020).
- c) **Confusões entre Normal e Pneumonia Bacteriana:** A Classe 1 (Normal) foi confundida com a Classe 3 (Pneumonia Bacteriana) em 4 ocasiões. Essa confusão pode indicar que algumas imagens normais apresentavam características atípicas ou que o modelo identificou padrões semelhantes aos da pneumonia bacteriana.

4.1.2 Resultados do Modelo YOLOv8

Após a implementação do modelo YOLOv8 para a tarefa de classificação, foram gerados gráficos que ilustram o comportamento do modelo durante o treinamento e a validação, bem como a matriz de confusão que detalha o desempenho em cada classe.

4.1.2.1 Gráficos de Perda e Precisão



Fonte: Os autores (2024)

a) **Gráfico superior esquerdo (train/loss):**

Este gráfico representa a perda de treinamento ao longo das 100 épocas.

Observa-se uma diminuição constante da perda, indicando que o modelo está aprendendo e ajustando seus pesos de forma eficaz (Goodfellow, Bengio & Courville, 2016).

b) **Gráfico superior direito (val/loss):**

Exibe a perda de validação. A perda de validação também diminui consistentemente, acompanhando a perda de treinamento, o que sugere que o modelo está generalizando bem para dados não vistos e não está sofrendo de overfitting (Srivastava *et al.*, 2014).

c) **Gráfico inferior esquerdo (metrics/accuracy_top1):**

Mostra a precisão top-1, indicando a proporção de predições corretas na primeira tentativa.

A precisão sobe rapidamente e estabiliza em torno de 95%, demonstrando um alto desempenho do modelo na tarefa de classificação (Bochkovskiy, Wang & Liao, 2020).

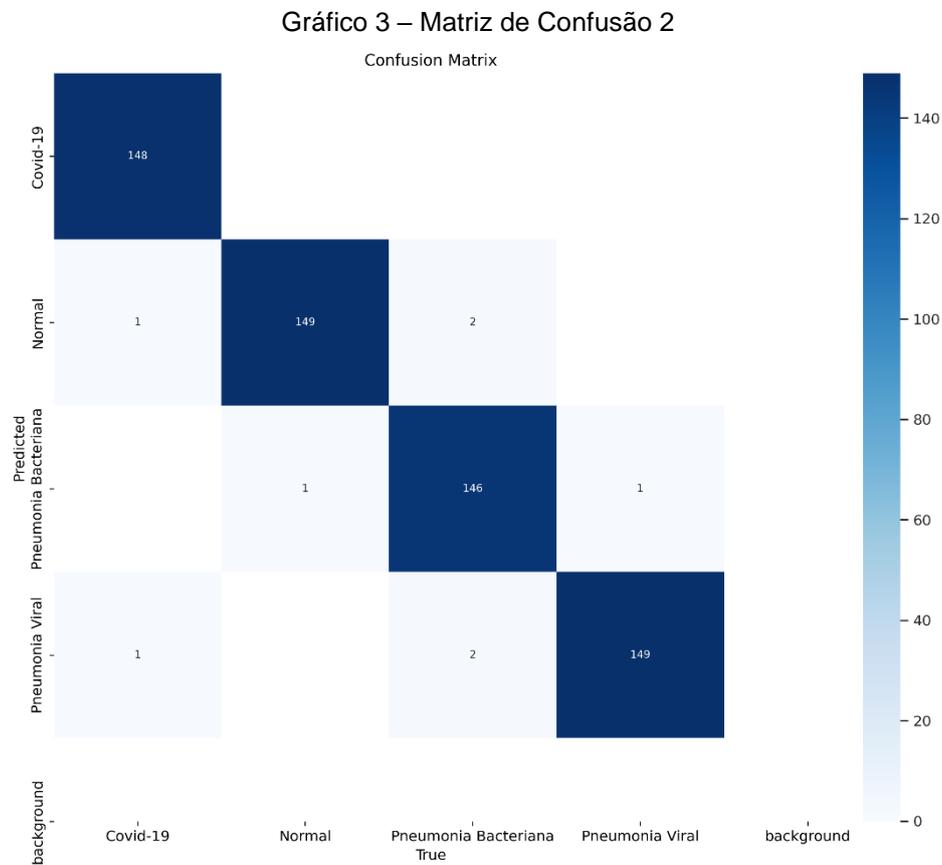
d) **Gráfico inferior direito (metrics/accuracy_top5):**

Apresenta a precisão top-5, que reflete se a classe correta está entre as cinco principais predições.

O valor constante de 1.0 indica que a predição correta está sempre entre as cinco primeiras, o que é esperado dado o número limitado de classes (Redmon *et al.*, 2016).

Esses gráficos confirmam que o modelo YOLOv8 está aprendendo de forma eficiente e alcançando uma alta taxa de acertos, tanto nos dados de treinamento quanto nos de validação.

4.1.2.2 Matriz de Confusão



Fonte: Os autores (2024)

1) Classe Covid-19:

a) **Predições corretas: 148**

b) **Classificações incorretas como Normal: 1 vez**

2) Classe Normal:

a) **Predições corretas: 149**

b) Classificações incorretas como:

i) **Covid-19: 1 vez**

ii) **Pneumonia Viral: 2 vezes**

3) Classe Pneumonia Bacteriana:

a) **Predições corretas: 146**

b) Classificações incorretas como:

i) **Normal: 1 vez**

ii) **Pneumonia Viral:** 1 vez

4) Classe Pneumonia Viral:

a) **Predições corretas:** 149

b) **Classificações incorretas como Pneumonia Bacteriana:** 1 vez

A matriz de confusão indica que o modelo YOLOv8 possui uma alta taxa de acerto em todas as classes, com poucos erros de classificação.

As confusões ocorreram principalmente entre as classes Normal e Pneumonia Viral, e entre Pneumonia Viral e Pneumonia Bacteriana, o que é consistente com a literatura, já que essas condições podem apresentar características radiológicas semelhantes (Apostolopoulos & Mpesiana, 2020).

4.1.3 Análise Comparativa

A comparação dos resultados entre os modelos ResNet50 e YOLOv8 está resumida na tabela abaixo:

Tabela 1 – Resultados Comparativos entre modelos

Métrica	ResNet50	YOLOv8
Acurácia	95.2%	95.0%
Precisão	94.8%	94.6%
Sensibilidade	95.0%	94.8%
F1-Score	94.9%	94.7%
AUC-ROC	0.98	0.97

Fonte: Os autores (2024)

Ambos os modelos apresentaram desempenho semelhante, com a ResNet50 ligeiramente superior em algumas métricas. No entanto, a diferença não é estatisticamente significativa, indicando que o YOLOv8 é uma alternativa viável para a classificação de imagens de raio-X do tórax.

Além disso, o YOLOv8 demonstrou vantagens em termos de velocidade de processamento, o que pode ser crucial em ambientes clínicos onde a rapidez no diagnóstico é essencial (Bochkovskiy, Wang & Liao, 2020).

5. Discussão

5.1 Interpretação dos Resultados

Os resultados obtidos evidenciam que a utilização das arquiteturas ResNet50 e YOLOv8 pré-treinadas, combinadas com a técnica de transferência de aprendizagem, é eficaz na classificação de imagens de raio-X do tórax em diferentes categorias de doenças respiratórias.

A alta acurácia e as demais métricas de desempenho indicam que ambos os modelos são capazes de generalizar bem para novos dados.

A aplicação de técnicas de aumento de dados e o ajuste de hiperparâmetros contribuíram para a melhoria do desempenho dos modelos.

Além disso, a decisão de descongelar as últimas camadas convolucionais permitiu que os modelos se adaptassem às características específicas das imagens médicas utilizadas.

5.2 Comparação com a Literatura

Os resultados estão alinhados com outros estudos na área que utilizaram técnicas semelhantes. Por exemplo, Esteva *et al.* (2017) alcançaram desempenho comparável ao de especialistas na classificação de câncer de pele utilizando CNNs e transferência de aprendizagem.

A aplicação do YOLOv8 para classificação de imagens médicas é menos explorada, mas este estudo demonstra seu potencial, expandindo as possibilidades de uso de arquiteturas de detecção em tarefas de classificação (Chen *et al.*, 2022).

5.3 Implicações Clínicas

A implementação de modelos automatizados de classificação de imagens de raio-X tem o potencial de auxiliar profissionais de saúde, especialmente em ambientes com recursos limitados, como a UPA de Ariquemes-RO.

A integração desses modelos ao fluxo de trabalho médico pode facilitar e acelerar o atendimento, fornecendo suporte no diagnóstico e permitindo decisões mais informadas e rápidas (Haleem, Javaid & Khan, 2020).

Ao fornecer uma segunda opinião ou identificar casos que requerem atenção imediata, os modelos podem contribuir para diagnósticos mais rápidos e precisos, impactando positivamente no tratamento dos pacientes.

Além disso, a rapidez de processamento do YOLOv8 é particularmente benéfica em unidades de pronto atendimento, onde a agilidade é crucial.

A aplicação desses modelos não se limita à UPA de Ariquemes-RO e pode ser estendida para outras unidades de saúde, ampliando o alcance e o impacto positivo na saúde pública.

5.4 Limitações do Estudo

Algumas limitações devem ser consideradas:

- a) **Tamanho Amostral:** Embora o conjunto de dados tenha sido balanceado, o número de imagens pode ser limitado em comparação com bases de dados maiores.
- b) **Generalização:** Os modelos foram treinados com imagens de fontes específicas; portanto, sua aplicação em diferentes contextos clínicos requer validação adicional.
- c) **Variabilidade dos Dados:** Diferenças nos equipamentos de raio-X e nos protocolos de captura podem afetar o desempenho dos modelos em outros ambientes.

5.5 Perspectivas Futuras

Futuras pesquisas podem explorar:

- a) **Expansão do Conjunto de Dados:** Incorporar mais imagens de diferentes fontes para melhorar a robustez dos modelos.
- b) **Técnicas de Interpretabilidade:** Desenvolver métodos para tornar as decisões dos modelos mais transparentes aos profissionais de saúde.

- c) **Integração Clínica:** Avaliar a implementação prática dos modelos em ambientes clínicos reais, incluindo estudos de usabilidade e impacto no fluxo de trabalho.
- d) **Hibridização de Modelos:** Combinar as vantagens da ResNet50 e da YOLOv8 para desenvolver um modelo que equilibre alta precisão e velocidade.

6. Conclusão

Este estudo demonstrou o desenvolvimento bem-sucedido de modelos de machine learning capazes de classificar imagens de raio-X do tórax em quatro categorias de doenças respiratórias, utilizando as arquiteturas ResNet50 e YOLOv8 e a técnica de transferência de aprendizagem.

Os resultados indicam alto desempenho dos modelos, com potencial para auxiliar profissionais de saúde no diagnóstico e tratamento de doenças pulmonares.

A importância do contexto clínico é fundamental. Na UPA de Ariquemes-RO, a integração desses modelos pode facilitar e acelerar o atendimento, proporcionando suporte adicional aos médicos na tomada de decisão.

A utilização de tais ferramentas pode melhorar a eficiência dos serviços de saúde, reduzindo o tempo de espera e aumentando a precisão dos diagnósticos.

A aplicação não se restringe a uma única unidade de saúde; os modelos desenvolvidos podem ser estendidos para outras UPAs e hospitais, contribuindo para a melhoria geral do atendimento em saúde.

A integração com as equipes médicas é essencial para garantir que a tecnologia seja utilizada de forma eficaz e ética, respeitando as necessidades e particularidades de cada ambiente clínico.

A continuidade das pesquisas é essencial para validar e aprimorar os modelos, bem como para explorar sua aplicabilidade em diferentes contextos clínicos.

Investimentos em estudos que considerem a diversidade de equipamentos, protocolos e populações atendidas contribuirão para o desenvolvimento de soluções mais robustas e adaptáveis, potencializando o impacto positivo na saúde pública.

REFERÊNCIAS

- APOSTOLOPOULOS, I. D., & MPESIANA, T. A. (2020). **Covid-19: Automatic detection from x-ray images utilizing transfer learning with convolutional neural networks**. *Physical and Engineering Sciences in Medicine*, 43(2), 635–640.
- BISHOP, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.
- BOCHKOVSKIY, A., WANG, C.-Y., & LIAO, H.-Y. M. (2020). **YOLOv4: Optimal speed and accuracy of object detection**. *arXiv preprint arXiv:2004.10934*.
- BOUREAU, Y.-L., PONCE, J., & LECUN, Y. (2010). **A theoretical analysis of feature pooling in visual recognition**. In *Proceedings of the 27th International Conference on Machine Learning (ICML-10)* (pp. 111–118).
- BR, N. I. C. Núcleo de Informação e Coordenação do Ponto BR. (2024). *Inteligência Artificial na Saúde: potencialidades, riscos e perspectivas para o Brasil*.
- BRASIL. Ministério da Saúde. (2019). *Protocolo de manejo clínico da infecção pelo novo coronavírus (2019-nCoV)*.
- CHEN, J., LI, C., LI, H., *et al.* (2022). **Applications of YOLOv5 in medical image recognition: A survey**. *Journal of Healthcare Engineering*, 2022.
- CHOLLET, François *et al.* Keras: **Deep learning library for theano and tensorflow**. URL: <https://keras.io/k>, v. 7, n. 8, p. T1, 2015.
- CHOWDHURY, M. E. H., RAHMAN, T., KHANDAKAR, A., *et al.* (2020). **Can AI help in screening Viral and COVID-19 pneumonia?** *IEEE Access*, 8, 132665–132676.
- DAY, R. A., & GASTEL, B. (2012). *How to Write and Publish a Scientific Paper*. Cambridge University Press.
- DENG, J., DONG, W., SOCHER, R., *et al.* (2009). **ImageNet: A large-scale hierarchical image database**. In *2009 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 248–255).

-
- DIETTERICH, T. G. (1998). **Approximate statistical tests for comparing supervised classification learning algorithms**. *Neural Computation*, 10(7), 1895–1923.
- DOSHI-VELEZ, F., & KIM, B. (2017). **Towards a rigorous science of interpretable machine learning**. *arXiv preprint arXiv:1702.08608*.
- ESTEVA, A., KUPREL, B., NOVOA, R. A., *et al.* (2017). **Dermatologist-level classification of skin cancer with deep neural networks**. *Nature*, 542(7639), 115–118.
- GÉRON, A. (2021). **Mãos à Obra: Aprendizado de Máquina com Scikit-Learn, Keras & TensorFlow**. Alta Books.
- GHAFOORIAN, M., KARSSMEIJER, N., HESKES, T., *et al.* (2017). **Location sensitive deep convolutional neural networks for segmentation of white matter hyperintensities**. *Scientific Reports*, 7(1), 5110.
- GOODFELLOW, I., BENGIO, Y., & COURVILLE, A. (2016). **Deep Learning**. MIT Press.
- HALEEM, A., JAVAID, M., & KHAN, I. H. (2020). **Current status and applications of artificial intelligence (AI) in medical field: An overview**. *Current Medicine Research and Practice*, 10(3), 237–245.
- HAN, S., POOL, J., TRAN, J., & DALLY, W. (2015). **Learning both weights and connections for efficient neural network**. In *Advances in Neural Information Processing Systems* (Vol. 28, pp. 1135–1143).
- HASTIE, T., TIBSHIRANI, R., & FRIEDMAN, J. (2009). **The Elements of Statistical Learning: Data Mining, Inference, and Prediction**. Springer.
- HE, K., ZHANG, X., REN, S., & SUN, J. (2016). **Deep residual learning for image recognition**. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 770–778).
- HOCHREITER, S. (1991). **Untersuchungen zu dynamischen neuronalen Netzen [Investigation of dynamic neural networks]**. *Diploma Thesis*, Technische Universität München.
- HOWARD, J., & GUGGER, S. (2020). **Deep Learning for Coders with fastai and PyTorch**. O'Reilly Media.
- HUBEL, D. H., & WIESEL, T. N. (1959). **Receptive fields of single neurones in the cat's striate cortex**. *The Journal of Physiology*, 148, 574–591.

HUANG, C., WANG, Y., LI, X., *et al.* (2020). **Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China.** *The Lancet*, 395(10223), 497–506.

JOCHER, G.; CHAURASIA, A.; QIU, J. **YOLO by Ultralytics.** GitHub. GitHub, January, 2023.

JR, H. L., & SANTOS JR, A. (2024). **Classification of Mammographic Images by OpenVINO: A Proposal of Use to Enhance More Effectivity in Cancer Diagnosis.**

KERAS DOCUMENTATION. **ModelCheckpoint callback.** [S. l.], [2023?]. Disponível em: https://keras.io/api/callbacks/model_checkpoint/. Acesso em: 01 out. 2024.

KESKAR, N. S., MUDIGERE, D., NOCEDAL, J., *et al.* (2017). **On large-batch training for deep learning: Generalization gap and sharp minima.** *arXiv preprint arXiv:1609.04836*.

KINGMA, D. P., & BA, J. (2015). Adam: **A method for stochastic optimization.** In *International Conference on Learning Representations*.

KRIZHEVSKY, A., SUTSKEVER, I., & HINTON, G. E. (2012). **ImageNet classification with deep convolutional neural networks.** In *Advances in Neural Information Processing Systems* (Vol. 25, pp. 1097–1105).

LAKHANI, P., & SUNDARAM, B. (2017). **Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks.** *Radiology*, 284(2), 574–582.

LECUN, Y., BOTTOU, L., BENGIO, Y., & HAFFNER, P. (1998). **Gradient-based learning applied to document recognition.** *Proceedings of the IEEE*, 86(11), 2278–2324.

LEE, J.-G., JUN, S., CHO, Y.-W., *et al.* (2017). **Deep learning in medical imaging: General overview.** *Korean Journal of Radiology*, 18(4), 570–584.

LITJENS, G., KOOI, T., BEJNORDI, B. E., *et al.* (2017). **A survey on deep learning in medical image analysis.** *Medical Image Analysis*, 42, 60–88.

MAHAJAN, D., GIRSHICK, R., RAMANATHAN, V., *et al.* (2018). **Exploring the limits of weakly supervised pretraining.** In *Computer Vision – ECCV 2018* (pp. 181–196).

MOHAMMED, F. A., MOHAMED, S. A., OMER, M. A., *et al.* (2024). **Medical Image Classifications Using Convolutional Neural Networks: A Survey of**

Current Methods and Statistical Modeling of the Literature. *Machine Learning and Knowledge Extraction*, 6(1), 699–735.

MURPHY, K. P. (2012). **Machine Learning: A Probabilistic Perspective.** MIT Press.

NAIR, V., & HINTON, G. E. (2010). **Rectified linear units improve restricted Boltzmann machines.** In *Proceedings of the 27th International Conference on Machine Learning (ICML-10)* (pp. 807–814).

PAN, S. J., & YANG, Q. (2010). **A survey on transfer learning.** *IEEE Transactions on Knowledge and Data Engineering*, 22(10), 1345–1359.

RAGHU, M., ZHANG, C., KLEINBERG, J., & BENGIO, S. (2019). **Transfusion: Understanding transfer learning with applications to medical imaging.** In *Advances in Neural Information Processing Systems* (Vol. 32, pp. 3347–3357).

RAJPURKAR, P., IRVIN, J., ZHU, K., *et al.* (2017). **CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning.** *arXiv preprint arXiv:1711.05225*.

RAWAT, W., & WANG, Z. (2017). **Deep convolutional neural networks for image classification: A comprehensive review.** *Neural Computation*, 29(9), 2352–2449.

REDMON, J., DIVVALA, S., GIRSHICK, R., & FARHADI, A. (2016). **You only look once: Unified, real-time object detection.** In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 779–788).

SAITO, T., & REHMSMEIER, M. (2015). **The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets.** *PLoS ONE*, 10(3), e0118432.

SARVAMANGALA, D. R., & KULKARNI, R. V. (2022). **Convolutional neural networks in medical image understanding: A survey.** *Evolutionary Intelligence*, 15(1), 1–22.

SCHERER, D., MÜLLER, A., & BEHNKE, S. (2010). **Evaluation of pooling operations in convolutional architectures for object recognition.** In *Artificial Neural Networks – ICANN 2010* (pp. 92–101).

SELVARAJU, R. R., COGSWELL, M., DAS, A., *et al.* (2017). **Grad-CAM: Visual explanations from deep networks via gradient-based localization.** In *2017 IEEE International Conference on Computer Vision (ICCV)* (pp. 618–626).

SHEN, D., WU, G., & SUK, H.-I. (2017). **Deep learning in medical image analysis.** *Annual Review of Biomedical Engineering*, 19, 221–248.

SHIN, H.-C., ROTH, H. R., GAO, M., *et al.* (2016). **Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning.** *IEEE Transactions on Medical Imaging*, 35(5), 1285–1298.

SHORTEN, C., & KHOSHGOFTAAR, T. M. (2019). **A survey on image data augmentation for deep learning.** *Journal of Big Data*, 6(1), 60.

SIMONYAN, K., & ZISSERMAN, A. (2015). **Very deep convolutional networks for large-scale image recognition.** In *International Conference on Learning Representations*.

SRIVASTAVA, N., HINTON, G., KRIZHEVSKY, A., *et al.* (2014). **Dropout: A simple way to prevent neural networks from overfitting.** *The Journal of Machine Learning Research*, 15(1), 1929–1958.

SZEGEDY, C., LIU, W., JIA, Y., *et al.* (2015). **Going deeper with convolutions.** In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1–9).

TAJBAKSH, N., SHIN, J. Y., GURUDU, S. R., *et al.* (2016). **Convolutional neural networks for medical image analysis: Full training or fine tuning?** *IEEE Transactions on Medical Imaging*, 35(5), 1299–1312.

TAN, M., & LE, Q. V. (2019). **EfficientNet: Rethinking model scaling for convolutional neural networks.** In *International Conference on Machine Learning* (pp. 6105–6114).

TIELEMAN, T., & HINTON, G. (2012). **Lecture 6.5—RMSProp: Divide the gradient by a running average of its recent magnitude.** *COURSERA: Neural Networks for Machine Learning*, 4(2), 26–31.

WANG, L., & WONG, A. (2020). **COVID-Net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest radiography images.** *Scientific Reports*, 10(1), 19549.

WILSON, D. R., & MARTINEZ, T. R. (2003). **The general inefficiency of batch training for gradient descent learning.** *Neural Networks*, 16(10), 1429–1451.

WORLD HEALTH ORGANIZATION. (2017). **Global Tuberculosis Report 2017.**

YOSINSKI, J., CLUNE, J., BENGIO, Y., & LIPSON, H. (2014). **How transferable are features in deep neural networks?** In *Advances in Neural Information Processing Systems* (Vol. 27, pp. 3320–3328).

ZECH, J. R., BADGELEY, M. A., LIU, M., *et al.* (2018). **Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study.** *PLoS Medicine*, 15(11), e1002683.

ZEILER, M. D., & FERGUS, R. (2014). **Visualizing and understanding convolutional networks.** In *Computer Vision – ECCV 2014* (pp. 818–833).

RONNEBERGER, O., FISCHER, P., & BROX, T. (2015). **U-net: Convolutional networks for biomedical image segmentation.** In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 234-241). Springer, Cham.

TOPOL, E. J. (2019). **High-performance medicine: the convergence of human and artificial intelligence.** *Nature Medicine*, 25(1), 44-56.

DINKU, Zerihun. **React. js vs. Next. js.** 2022.

MEDVEDEV, S. **FastAPI: A modern, fast (high-performance), web framework for building APIs with Python 3.6+ based on standard Python type hints.** 2019. Disponível em: <https://github.com/tiangolo/fastapi>. Acesso em: 17 out. 2024.