

**UNIVERSIDADE ESTADUAL PAULISTA "JÚLIO DE MESQUITA FILHO"**

**FACULDADE DE CIÊNCIAS - CAMPUS BAURU**

**DEPARTAMENTO DE COMPUTAÇÃO**

**BACHARELADO EM CIÊNCIA DA COMPUTAÇÃO**

**KAIO HENRIQUE AVELINO DOMENI DE PAULO**

**INTELIGÊNCIA ARTIFICIAL PARA AUXÍLIO NO DIAGNÓSTICO DE  
INFECÇÕES FÚNGICAS E NEOPLASIAS PULMONARES EM  
RADIOGRAFIAS TORÁCICAS**

**BAURU**

**Novembro/2024**

KAIO HENRIQUE AVELINO DOMENI DE PAULO

**INTELIGÊNCIA ARTIFICIAL PARA AUXÍLIO NO DIAGNÓSTICO DE  
INFECÇÕES FÚNGICAS E NEOPLASIAS PULMONARES EM  
RADIOGRAFIAS TORÁCICAS**

Trabalho de Conclusão de Curso do Curso  
de Bacharelado em Ciência da Computa-  
ção da Universidade Estadual Paulista “Jú-  
lio de Mesquita Filho”, Faculdade de Ciên-  
cias, Campus Bauru.

Orientador: Prof. Dr. João Paulo Papa

Coorientador: Marcel Koenigkam Santos

BAURU

Novembro/2024

P331i

Paulo, Kaio Henrique Avelino Domeni de

Inteligência artificial para auxílio no diagnóstico de infecções fúngicas e neoplasias pulmonares em radiografias torácicas / Kaio Henrique Avelino

Domeni de Paulo. -- Bauru, 2024

42 p. : il., tabs.

Trabalho de conclusão de curso (Bacharelado - Ciência da Computação) - Universidade Estadual Paulista (UNESP), Faculdade de Ciências, Bauru

Orientador: João Paulo Papa

Coorientador: Marcel Koenigkam Santos

1. Redes neurais. 2. Inteligência artificial. 3. Infecções fúngicas. 4. Tumores. 5. Tórax Doenças. I. Título.

Kaio Henrique Avelino Domeni de Paulo

# **Inteligência Artificial para Auxílio no Diagnóstico de Infecções Fúngicas e Neoplasias Pulmonares em Radiografias Torácicas**

Trabalho de Conclusão de Curso do Curso de Bacharelado em Ciência da Computação da Universidade Estadual Paulista "Júlio de Mesquita Filho", Faculdade de Ciências, Campus Bauru.

Banca Examinadora

**Prof. Dr. João Paulo Papa**

Orientador

Universidade Estadual Paulista "Júlio de Mesquita Filho"

Faculdade de Ciências

Departamento de Computação

**Profa. Dra. Simone das Graças Domingues Prado**

Universidade Estadual Paulista "Júlio de Mesquita Filho"

Faculdade de Ciências

Departamento de Computação

**Prof. Dr. Leandro Aparecido Passsos Junior**

Universidade Estadual Paulista "Júlio de Mesquita Filho"

Faculdade de Ciências

Departamento de Computação

Bauru, 12 de Novembro de 2024.

*Dedico esse trabalho à todos os familiares e pessoas queridas que conheci na jornada  
até aqui.*

# Agradecimentos

A estrada até aqui foi longa e conheci pessoas maravilhosas durante todos esses anos. Deixo aqui, um agradecimento aos meus familiares, minha mãe, meu pai, meus queridos avós, minha querida irmã Aghata, minha madrastra, minha prima Maria Eduarda, meu primo Raphael, meus tios Thiago e Anderson e meus doguinhos Jiló e Pluto que sempre me recebem com alegria em casa, além de Bob e Bucky, peludos que já não estão mais entre nós. Com vocês, aprendi sobre valores e aprendi a evoluir como pessoa. Tenho um carinho imenso por cada um de vocês e sou grato por todos os ensinamentos a apoio durante a minha vida.

Nos últimos quatro anos, fiz amigos para a vida de diversas classes e turmas. Sou grato por todas essas amizades e por influenciarem positivamente minha jornada na Universidade. Alguns nomes para lembrar são: Mavi, Junks, Leonel, Rhay, Dororo, Herminio e Raul, do Caosnesp; Alex, Garcia, Yuki, Ian e Franco, do *Bulking*; meus bixos, Henrique, Luana, Daniel, Chris, Vini e Sofia; Luan e Laura de SP; Natan, Léo e Jovem, de BSI; meus veteranos, Gu, Nih, Nic, Marry, Toki, Cassiano, Mateus e Xilsu; Júlia, Mimi e Isadora, de Bio; Paiola, Yasmin, Danilo e Roder, do Recogna; e Anselmo, Vitas, Rayane, Cunha, Koda e Amandinha, amigos que me acompanham desde antes da graduação. Obrigado a todos vocês por tornarem meus dias maravilhosos e participarem de momentos importantes ao meu lado.

Sou grato aos grupos que participei, em especial, o Laboratório Recogna, que alavancou minha vida acadêmica e formou bases fortes para os conhecimentos que tenho hoje e ao time do Protiva que proporcionou desafios emocionantes e onde pude compartilhar conhecimento para ajudar outras pessoas.

Agradeço imensamente ao João Paulo Papa, Marcel Koenigkam Santos e João Renato que me orientaram para a realização deste trabalho e foram de grande importância para meu desenvolvimento acadêmico. Sou grato à FAPESP (Fundação de Amparo a Pesquisa do Estado de São Paulo) processo #2023/12569-0.

E por fim, sou grato aos membros da banca e professores que tive o prazer de ser aluno na UNESP-Bauru: Clayton, Nilceu, Borges, Roberta, Cristiane, Hercules, Andrea, Adriana, Rene, Perea, Sonia, Wallace, Wilson, Crepaldi, Sacoman, Simone, Sementille, Graef, Rodolfo, Tatiana, Márcia, Kelton, Leandro, Juliana, Zanata, Higor e Napolitano.

*"We will make our own destiny."*  
(SANTA MONICA STUDIO, 2022)

# Resumo

A radiografia simples do tórax representa na maioria das situações o primeiro exame radiológico na avaliação das doenças torácicas e pulmonares. Grande parte das doenças broncopulmonares se apresentam na radiografia como opacidades, como é o caso da tuberculose, infecções fúngicas, doenças intersticiais pulmonares, câncer de pulmão e as pneumonias virais. Técnicas de aprendizado de máquina vêm sendo aplicadas para auxiliar na detecção de alterações nos exames de imagem, mas a falta de exames de radiografia anotados de pacientes com infecção fúngica em bases de dados torna difícil a sua classificação. O objetivo desta pesquisa é testar e aprimorar Redes Neurais Convolucionais para auxílio ao diagnóstico de pacientes com infecções fúngicas e neoplasia, utilizando uma base de dados anotada de exames de radiografia fornecidos pelo Hospital das Clínicas de Ribeirão Preto. Para tal, foram testados diferentes técnicas de processamento de imagem e *Ensemble* com o intuito de maximizar a acurácia e evitar o *overfitting* do modelo, concluindo com o desenvolvimento de uma interface que simule uma ferramenta de auxílio computadorizado ao diagnóstico.

**Palavras-chave:** Pneumonia Fúngica; Neoplasia; Redes Neurais Convolucionais; Radiografias do Tórax; CAD.



# Abstract

Chest X-ray is, in most cases, the first radiological exam used to evaluate thoracic and pulmonary diseases. A large number of bronchopulmonary diseases present in chest X-rays as opacities, such as tuberculosis, fungal infections, interstitial lung diseases, lung cancer, and viral pneumonias. Machine learning techniques have been applied to aid in the detection of abnormalities in imaging exams, but the lack of annotated Chest X-ray exams from patients with fungal infections in available databases makes classification challenging. The objective of this research is to test and improve Convolutional Neural Networks to assist in the diagnosis of patients with fungal infections and neoplasia, using an annotated Chest X-ray dataset provided by the Hospital das Clínicas of Ribeirão Preto. For this purpose, different image processing techniques and Ensemble methods were tested to maximize accuracy and prevent model overfitting, concluding with the development of an interface simulating a computer-aided diagnosis tool.

**Keywords:** Fungal Pneumonia; Neoplasia; Convolutional Neural Networks; Chest X-rays; CAD.

# Lista de figuras

Figura 1 – Arquitetura de uma CNN simples com camadas de convolução e camadas totalmente conectadas. . . . .	18
Figura 2 – Estrutura de um bloco residual. . . . .	20
Figura 3 – Exemplos de imagens presentes na base de dados. . . . .	23
Figura 4 – Distribuição de imagens por classe. . . . .	24
Figura 5 – Visão geral do modelo com duplo <i>ensemble</i> proposto. . . . .	26
Figura 6 – Evolução da Acurácia, <i>Loss</i> e <i>F1 Score</i> ao longo de 60 épocas nos conjuntos de treino e validação. . . . .	28
Figura 7 – Matriz de confusão do modelo com <i>ensemble</i> no conjunto de teste. . . . .	28
Figura 8 – Evolução da Acurácia, <i>Loss</i> e <i>F1 Score</i> ao longo de 20 épocas nos conjuntos de treino e validação da ResNet-101. . . . .	29
Figura 9 – Evolução da Acurácia, <i>Loss</i> e <i>F1 Score</i> ao longo de 100 épocas nos conjuntos de treino e validação da ResNet-152. Apesar da tentativa de treinar a rede por mais épocas, não há ganho de desempenho após 15 épocas. . . . .	30
Figura 10 – Evolução da Acurácia, <i>Loss</i> e <i>F1 Score</i> ao longo de 80 épocas nos conjunto de treinos e validação do <i>Ensemble Model 1</i> . . . . .	31
Figura 11 – Matriz de confusão do modelo com <i>ensemble</i> da ResNet-101 e ResNet-152 no conjunto de teste. . . . .	31
Figura 12 – Evolução da Acurácia, <i>Loss</i> e <i>F1 Score</i> ao longo de 80 épocas nos conjuntos de treino e validação do modelo final com duplo <i>ensemble</i> . . . . .	32
Figura 13 – Matriz de confusão do modelo final com duplo <i>ensemble</i> . . . . .	33
Figura 14 – Diagrama do funcionamento da Aplicação CAD. . . . .	34
Figura 15 – Tela do Menu Principal. À esquerda, a interface no modo escuro e à direita, a interface no modo claro. . . . .	36
Figura 16 – Tela de Diagnóstico sem uma imagem carregada. . . . .	36
Figura 17 – Tela de Diagnóstico com uma imagem carregada. . . . .	37
Figura 18 – Tela de Diagnóstico com uma imagem carregada e estatísticas que denotam incerteza. . . . .	38
Figura 19 – Tela de Créditos. . . . .	38

# Lista de tabelas

Tabela 1	– Métricas obtidas ao avaliar os modelos utilizando o conjunto de teste. Nota-se que o <i>Recall</i> foi extremamente baixo no modelo com <i>ensemble</i> .	28
Tabela 2	– Métricas obtidas ao avaliar os modelos utilizando o conjunto de teste. Todas as redes, com exceção da LeNet vieram com os pesos pré-treinados na base de dados Imagenet.	29
Tabela 3	– Métricas obtidas ao avaliar os modelos utilizando o conjunto de teste. O duplo <i>ensemble</i> obteve as melhores métricas no <i>dataset</i> avaliado.	32
Tabela 4	– Métricas obtidas ao avaliar o modelo final nos conjuntos de treino, validação e teste. Além da alta taxa de acertos no treino, o modelo final obteve uma acurácia consideravelmente mais alta no conjunto de validação.	33

# Lista de abreviaturas e siglas

AM	Aprendizado de Máquina
CAD	<i>Computed Aided Diagnosis</i>
CNN	<i>Convolutional Neural Network</i>
IA	Inteligência Artificial
ResNet	<i>Residual Network</i>
RX	Radiografia
TC	Tomografia Computadorizada

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>13</b>
<b>1.1</b>	<b>Problemática</b>	<b>13</b>
<b>1.2</b>	<b>Justificativa</b>	<b>14</b>
<b>1.3</b>	<b>Objetivos</b>	<b>14</b>
1.3.1	Objetivo geral	14
1.3.2	Objetivos Específicos	15
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>16</b>
<b>2.1</b>	<b>Aprendizado de Máquina</b>	<b>16</b>
<b>2.2</b>	<b>Redes Neurais Convolucionais</b>	<b>17</b>
2.2.1	Estrutura de uma CNN	17
<b>2.3</b>	<b>ResNet</b>	<b>19</b>
<b>2.4</b>	<b><i>Ensemble</i></b>	<b>20</b>
<b>3</b>	<b>MATERIAIS E MÉTODOS</b>	<b>22</b>
<b>3.1</b>	<b>Configuração Experimental</b>	<b>22</b>
<b>3.2</b>	<b>Pré-Processamento</b>	<b>22</b>
<b>3.3</b>	<b>Base de Dados</b>	<b>23</b>
<b>3.4</b>	<b>Métricas</b>	<b>23</b>
<b>3.5</b>	<b>Modelo Proposto</b>	<b>25</b>
<b>4</b>	<b>RESULTADOS E DISCUSSÃO</b>	<b>27</b>
<b>4.1</b>	<b>Treinamento dos modelos com duas classes</b>	<b>27</b>
<b>4.2</b>	<b>Treinamento dos modelos com três classes</b>	<b>29</b>
<b>4.3</b>	<b>Primeiro <i>ensemble</i></b>	<b>30</b>
<b>4.4</b>	<b>Modelo final com <i>ensemble</i></b>	<b>31</b>
<b>5</b>	<b>APLICAÇÃO DE AUXÍLIO AO DIAGNÓSTICO</b>	<b>34</b>
<b>5.1</b>	<b>Processo de Desenvolvimento</b>	<b>35</b>
<b>5.2</b>	<b>Interfaces da Aplicação</b>	<b>35</b>
<b>6</b>	<b>CONSIDERAÇÕES FINAIS</b>	<b>39</b>
	<b>REFERÊNCIAS</b>	<b>40</b>

# 1 Introdução

O exame de radiografia (RX) do tórax representa na maioria das situações o primeiro exame radiológico na avaliação das doenças torácicas e pulmonares devido à sua disponibilidade, baixo custo e por ser simples do ponto de vista tecnológico (WADA; RODRIGUES; SANTOS, 2019). Doenças pulmonares como tumores e pneumonias virais, bacterianas ou fúngicas apresentam alta morbimortalidade e necessitam de investigação complementar, com tomografia computadorizada (TC) e/ou exame laboratoriais.

Mesmo sendo considerado um exame simples, é necessário um profissional experiente para interpretar corretamente um RX de tórax. Além disso, há uma falta de profissionais especialistas no local da realização do exame ou estes estão sobrecarregados pelo grande volume de exames e por essa razão, nos últimos anos, ferramentas de auxílio ao diagnóstico computadorizadas (CAD, do inglês "*computed aided diagnosis*") têm sido criadas para ajudar na detecção das alterações nos exames de imagem, na interpretação dos achados e no fluxo de trabalho dos serviços de saúde, melhorando a acurácia do diagnóstico feito pelo radiologista que enfrenta dificuldades devido a fatores como baixa qualidade de imagem, fadiga visual, distração e a própria natureza minuciosa das análises de RX (AZEVEDO-MARQUES, 2001). Para tanto, tem-se destacado o desenvolvimento de ferramentas de inteligência artificial (IA).

As pneumonias fúngicas são incomuns, porém representam causa importante de morbimortalidade principalmente na população de indivíduos imunossuprimidos (SILVA, 2010). Já o câncer de pulmão representa a principal causa de morte relacionada a câncer no mundo (SCHABATH; COTE, 2019) e, apesar de o rastreamento não dever ser realizado com RX, muitas vezes é o exame radiográfico o primeiro a detectar a presença de um nódulo ou massa pulmonar suspeita. O pulmão também é um dos sítios mais comuns de metástases de outros tumores.

Este trabalho terá como foco as infecções fúngicas, que podem se apresentar de diversas maneiras e serem fatais para pacientes que tenham o sistema imunológico debilitado, e a neoplasia, que é uma lesão relacionada ao crescimento anormal das células (EAGEN; LEWIS, 1977), também podendo ser fatal.

## 1.1 Problemática

Apesar da existência de alguns trabalhos que utilizam Redes Neurais Convolucionais (CNN, do inglês "*convolutional neural network*") para classificação de tuberculose

(LAKHANI; SUNDARAM, 2017) e pneumonias virais ou bacterianas, há uma escassez de trabalhos do tipo que tratem sobre infecções fúngicas e neoplasia pulmonar na radiografia.

Trabalhos como os de (WANG et al., 2023), (ANGELINI; SHAH, 2021) e (ZHANG et al., 2021) falam sobre o uso de IA para a detecção de pneumonias fúngicas utilizando exames de TC, que são consideravelmente mais complexos e mais caros de serem realizados apesar de alcançarem uma acurácia alta. Exames de RX também possuem consideravelmente menor resolução do que exames de TC, tornando a detecção de anomalias difícil até para radiologistas experientes (HASAN; MUHAMMAD; ISLAM, 2024).

## 1.2 Justificativa

Em casos de infecção de pacientes imunossuprimidos, é desejável que um tratamento menos específico se inicie imediatamente para diminuir a chance de fatalidade. No entanto, diagnósticos mais específicos podem ser requeridos devido à toxicidade de tratamentos comuns em doenças pulmonares e à alta taxa de mortalidade que está relacionada com um diagnóstico errado. Em casos desse tipo, amostras de tecido são ideias para um diagnóstico acurado, mas exames invasivos nem sempre podem ser realizados em pacientes imunossuprimidos (GUILLAMET et al., 2018), tornando métodos mais simples como o exame de RX a forma mais segura de obter um diagnóstico.

Estudos realizados nos últimos anos mostram que abordagens que fazem uso de Aprendizado de Máquina (AM) têm se mostrado eficientes para o desenvolvimento de CADs para diagnóstico de doenças pulmonares como o de Barbosa e Canuto (BARBOSA; CANUTO, 2021) e o de Lakhani e Sundaram (LAKHANI; SUNDARAM, 2017).

## 1.3 Objetivos

### 1.3.1 Objetivo geral

O objetivo do estudo é desenvolver e verificar a aplicabilidade de um modelo de CNN aplicado ao auxílio diagnóstico em imagens de RX do tórax, com foco na detecção e classificação das infecções fúngicas e neoplasia. Este projeto busca ser uma contribuição para estudos que envolvam outras doenças que causam opacidades em radiografias do tórax e necessitam de investigação complementar, como as doenças intersticiais pulmonar, tuberculose e pneumonia da Covid-19.

### 1.3.2 Objetivos Específicos

- Estruturar e pré processar a base de dados contendo as imagens de RX do tórax.
- Testar e comparar técnicas de pré-processamento de imagens para facilitar a extração de características relevantes para a classificação.
- Estruturar e implementar modelos de AM para análise de exames de RX do tórax.
- Projetar, implementar e avaliar um modelo que classifique pacientes saudáveis, com infecção fúngica e neoplasia pulmonar através de exames de RX do tórax;
- Estruturar e implementar uma interface que simule uma aplicação CAD que possa inferir o diagnóstico das doenças pulmonares avaliadas nesse estudo utilizando técnicas de AM.



## 2 Fundamentação Teórica

Neste capítulo, serão apresentados alguns dos conceitos importantes para o desenvolvimento do modelo e formar a base do método proposto discutido no capítulo 3. Inicialmente será dada uma introdução aos conceitos de AM e CNN e detalhar a funcionalidade de uma ResNet e *Ensemble*, que serão importantes para desenvolver o modelo de CNN proposto nos objetivos do capítulo 1.

### 2.1 Aprendizado de Máquina

O aprendizado de máquina é um campo da Inteligência Artificial (IA) que se concentra em criar algoritmos capazes de aprender e melhorar automaticamente a partir de dados, sem a necessidade de serem explicitamente programados para cada tarefa específica. Isso é alcançado por meio de técnicas que permitem aos computadores identificar padrões e fazer previsões com base em grandes volumes de dados, sendo amplamente utilizado em diversas áreas, como reconhecimento de imagens, diagnóstico médico, processamento de linguagem natural e até sistemas autônomos.

O conceito de AM é fundamentalmente orientado a dados, o que significa que o desempenho dos algoritmos melhora conforme a quantidade e a qualidade dos dados disponíveis para treinamento aumentam. Historicamente, o campo começou a ganhar forma com as primeiras tentativas de criar programas que imitassem o comportamento humano em situações específicas, como o jogo de xadrez, desenvolvido por Alan Turing e Dietrich Prinz na década de 1950 (SHINDE; SHAH, 2018). O grande salto, no entanto, veio com a combinação de redes neurais artificiais e técnicas mais sofisticadas de treinamento, como o perceptron de Frank Rosenblatt e o algoritmo de retropropagação de erros desenvolvido por Rumelhart e McClelland na década de 1980.

Existem três tipos principais de aprendizado de máquina: supervisionado, não supervisionado e por reforço (LUDERMIR, 2021). Focaremos no aprendizado supervisionado que foi o utilizado durante a pesquisa. No aprendizado supervisionado, os algoritmos são treinados com dados rotulados, ou seja, para cada exemplo de entrada, o sistema também recebe a resposta correta, permitindo que ele aprenda a generalizar para novos dados. Esse tipo de aprendizado é amplamente usado em tarefas de classificação, como diferenciar imagens ou prever preços de imóveis. No contexto da presente pesquisa, será utilizado para classificação diferenciar imagens e classificação de doenças.

Dentro do aprendizado supervisionado, as CNNs ganharam destaque em áreas como a visão computacional, enquanto Redes Neurais Recorrentes se mostraram eficazes para tarefas que envolvem dados sequenciais, como a previsão de séries temporais e processamento de linguagem natural. Com o avanço tecnológico, as redes neurais profundas passaram a utilizar várias camadas ocultas, aumentando a complexidade e a capacidade de processamento dos modelos.

A evolução do aprendizado de máquina foi acelerada por alguns fatores chave: a disponibilidade de grandes quantidades de dados (Big Data), o aumento da capacidade computacional (em especial, o uso de GPUs para treinamento de redes neurais), e a melhoria contínua de algoritmos. Essas inovações permitiram que tarefas anteriormente impossíveis para máquinas, como o reconhecimento de voz, a tradução automática e a geração de imagens, se tornassem realidade.

No contexto de diagnósticos médicos, por exemplo, algoritmos de aprendizado de máquina conseguem analisar imagens de exames como radiografias e ressonâncias magnéticas, detectando anomalias com precisão comparável a de especialistas humanos.

Os desafios atuais para o aprendizado de máquina incluem a necessidade de algoritmos mais explicáveis, que permitam aos humanos entender como as decisões são tomadas, e a capacidade de aprender com menos dados ou a partir de interações mais complexas, simulando de forma mais fiel o aprendizado humano. Esses aspectos são fundamentais para que a IA se torne mais confiável e ética, principalmente em contextos críticos, como a tomada de decisões médicas.

## 2.2 Redes Neurais Convolucionais

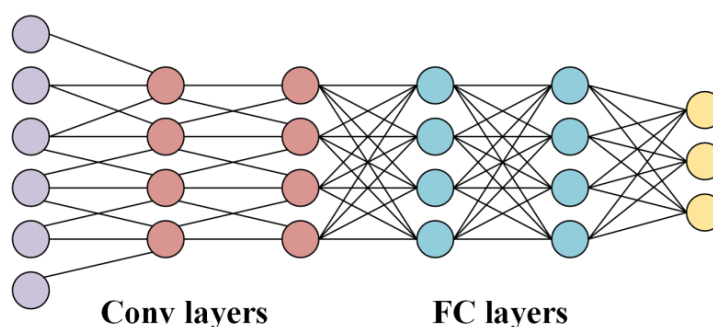
CNNs são uma classe de redes neurais artificiais amplamente utilizadas em tarefas de aprendizado de máquina, especialmente no campo da visão computacional. Inspiradas pela estrutura do córtex humano, as CNNs são projetadas para processar dados com estrutura de matriz, como imagens (HE et al., 2015a). Elas utilizam convoluções, que são operações matemáticas que permitem a extração automática de características diretamente dos dados brutos, eliminando a necessidade de extração manual de características, como ocorre em métodos tradicionais.

### 2.2.1 Estrutura de uma CNN

A arquitetura básica de uma CNN é composta por camadas de convolução, *pooling*, e camadas totalmente conectadas como descrito abaixo e podendo ser visualizado na figura 1:

1. **Camadas de Convolução:** Realizam a extração de características. Filtros (ou *kernels*) são aplicados nas imagens de entrada, e esses filtros capturam padrões como bordas, texturas e, em camadas mais profundas, estruturas mais complexas.
2. **Camadas de *Pooling*:** Reduzem a dimensionalidade dos dados, mantendo as características mais importantes. O *pooling* é realizado, tipicamente, com técnicas como o *max pooling*, que seleciona o valor máximo de uma região específica da imagem, reduzindo o número de parâmetros e o custo computacional.
3. **Camadas Totalmente Conectadas:** Na fase final da rede, os neurônios são totalmente conectados entre si, assim como nas redes neurais tradicionais. Elas são responsáveis pela tomada de decisão final, como a classificação das imagens processadas.

Figura 1 – Arquitetura de uma CNN simples com camadas de convolução e camadas totalmente conectadas.



Fonte: (LI et al., 2022).

As CNNs apresentam diversas vantagens em relação a outros métodos de aprendizado de máquina. Uma de suas principais vantagens é o uso de pesos compartilhados. Isso significa que o mesmo conjunto de filtros é aplicado em toda a imagem, o que reduz significativamente o número de parâmetros a serem treinados. Esse fator facilita o treinamento da rede, diminui o risco de sobreajuste (*overfitting*) e permite que o modelo generalize melhor em novos dados. Outra vantagem fundamental é a invariância espacial, que confere à CNN a capacidade de detectar características em uma imagem, independentemente de sua posição. Isso torna as CNNs extremamente eficazes para o reconhecimento de padrões, objetos e formas visuais em diferentes contextos, já que conseguem identificar esses elementos em qualquer parte da imagem.

As CNNs são amplamente utilizadas em uma variedade de aplicações. No campo da visão computacional, elas revolucionaram tarefas como a classificação de imagens, a detecção de objetos e a segmentação semântica, com impacto significativo em áreas como reconhecimento facial, sistemas de vigilância e veículos autônomos. No diagnóstico médico por imagem, as CNNs têm se mostrado essenciais para detectar

anomalias em exames como radiografias e tomografias (TANG et al., 2020). Elas ajudam a identificar doenças como tumores e condições pulmonares, aumentando a precisão diagnóstica, mesmo para patologias de difícil detecção, como infecções fúngicas.

## 2.3 ResNet

A ResNet (*Residual Network*) é um tipo de rede neural profunda que utiliza conexões residuais (ou "atalhos") para facilitar o treinamento de redes com muitas camadas. Introduzida em 2015 (HE et al., 2015a), a ResNet tem como principal objetivo resolver o problema da degradação da performance de redes muito profundas, onde a adição de mais camadas pode acabar piorando os resultados de treinamento, em vez de melhorá-los.

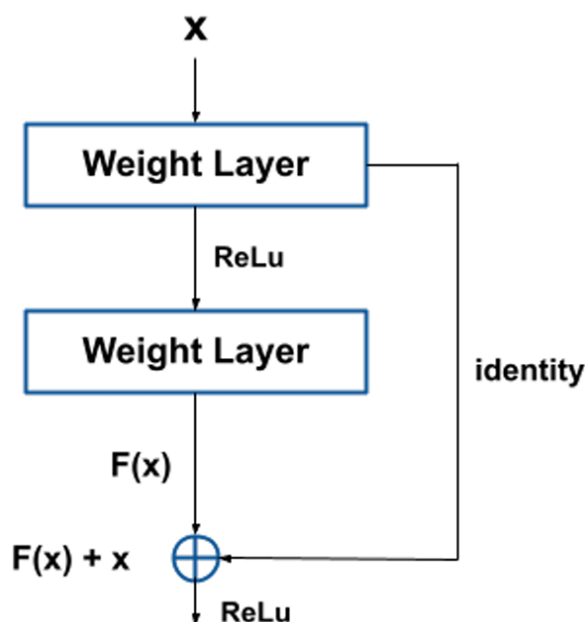
A estrutura da ResNet é baseada em blocos residuais. Cada bloco contém uma ou mais camadas de redes neurais, mas o diferencial está nas conexões diretas entre a entrada e a saída do bloco. Essas conexões diretas, ou atalhos, permitem que a entrada do bloco seja somada à sua saída. A ideia central é que, em vez de aprender uma função mapeadora completa, cada bloco residual aprenda a ajustar o resíduo, ou seja, a diferença entre a entrada e o resultado esperado. Na prática, isso facilita o aprendizado de funções mais complexas ao reduzir a carga de trabalho de cada camada.

A estrutura matemática de um ResNet é descrita da seguinte forma: em vez de calcular diretamente a saída  $y = F(x)$  onde  $F(x)$  representa a função de mapeamento aprendida, o bloco residual calcula  $y = F(x) + x$ , onde  $x$  é a entrada original. A estrutura do bloco residual pode ser observada na figura 2. Esse simples ajuste tem um grande impacto na eficiência de aprendizado de redes muito profundas. A adição desse "atalho" evita que a rede caia em problemas de desaparecimento de gradiente, o que é comum em redes profundas. O uso de conexões residuais permite que o fluxo de gradientes durante o treinamento seja mais estável, tornando possível a utilização de redes com centenas ou até milhares de camadas.

Existem diferentes variantes da ResNet, sendo que uma das principais é o uso de "*bottlenecks*", onde as camadas intermediárias do bloco residual têm dimensões reduzidas, otimizando o custo computacional sem perder a capacidade de aprendizado. Além disso, variantes como "*Wide ResNet*" e "*ResNeXt*" surgiram para explorar maneiras de ampliar ou empilhar blocos residuais.

A principal contribuição da ResNet é a facilidade de treinamento de redes muito profundas. Antes da ResNet, redes com dezenas de camadas já eram difíceis de treinar, mas os ResNets permitiram a construção de redes com até 1001 camadas (LI

Figura 2 – Estrutura de um bloco residual.



Fonte: (SAJWAN; MISHRA, 2024).

et al., 2016). Graças às conexões residuais, a ResNet se mostrou eficaz em várias tarefas de visão computacional, alcançando resultados de ponta em benchmarks como o ImageNet (HE et al., 2015b).

## 2.4 Ensemble

*Ensemble learning* é um método de aprendizado de máquina que combina múltiplos modelos de aprendizado para formar um modelo mais robusto e preciso. Ao contrário de métodos que dependem de um único algoritmo, o ensemble learning visa utilizar a diversidade e complementaridade dos modelos para obter previsões melhores. Essa abordagem é amplamente utilizada para melhorar a performance em tarefas complexas, como classificação e previsão, onde um único modelo pode não capturar todas as nuances dos dados.

O conceito central do *ensemble learning* é a combinação de vários classificadores ou modelos de aprendizado para criar um sistema mais forte. A base do *ensemble learning* está em utilizar algoritmos simples, chamados de classificadores fracos, que individualmente podem não ter um desempenho notável, mas, quando combinados, formam um classificador forte. A fusão desses resultados pode ser feita por diversos métodos de votação, como a média ponderada ou a maioria dos votos, dependendo da natureza dos modelos e da tarefa específica.

Uma vantagem clara do *ensemble learning* é a melhoria da acurácia, robustez e generalização do modelo, especialmente em cenários onde há ruído nos dados ou

os dados são altamente complexos. A combinação de múltiplos modelos, cada um capturando diferentes padrões dos dados, gera um sistema capaz de realizar previsões mais precisas.

A aplicação do *ensemble learning* em combinação com técnicas modernas como *deep learning* e aprendizado por reforço tem mostrado resultados promissores em áreas como processamento de imagem, análise de sentimentos e sistemas de recomendação (DONG et al., 2019). Por exemplo, combinações de redes neurais profundas com *ensemble learning* permitem a captura de representações semânticas mais ricas dos dados, enquanto em aprendizado por reforço, técnicas de *ensemble* podem reduzir a complexidade da amostragem e melhorar a estabilidade do aprendizado.

No entanto, alguns desafios permanecem, como a determinação do número ideal de modelos no *ensemble*, o balanceamento entre a diversidade dos classificadores e a complexidade do sistema, além da necessidade de mais estudos teóricos para compreender melhor as relações entre essas variáveis. Por essa razão, foram testadas várias formas para realização do *ensemble* com as redes utilizada na base de dados.

Em suma, o *ensemble learning* se destaca por sua versatilidade e capacidade de melhorar significativamente a performance de modelos de aprendizado de máquina ao combinar os pontos fortes de múltiplos algoritmos, proporcionando resultados que, em muitos casos, são superiores aos de um único modelo isolado.

## 3 Materiais e Métodos

Esta seção apresenta a metodologia de execução dos experimentos, as métricas de avaliação, a configuração experimental utilizada para condução dos testes e o, ao final do capítulo, o modelo proposto.

### 3.1 Configuração Experimental

Os experimentos foram conduzidos em um *notebook* com um processador Intel(R) Core(TM) i7-10750H, 16GB de RAM DDR4, uma GPU Nvidia GTX 1660 Ti com 6GB de VRAM e Sistema Operacional Windows 10.

Durante o treinamento, um *batch* de 32 elementos foi utilizado, em conjunto com uma taxa de aprendizado de  $1 \times 10^{-3}$  associada ao otimizador Adam. O monitoramento do modelo foi feito sobre a *loss* no conjunto de validação, com uma paciência de 20 épocas. A *seed* de seleção das imagens para os conjuntos de treino, validação e teste foi 123.

Os códigos foram desenvolvidos na linguagem de programação Python e as bibliotecas Tensorflow e Scikit-learn foram utilizadas para construção dos modelos de CNN utilizados nesta pesquisa.

### 3.2 Pré-Processamento

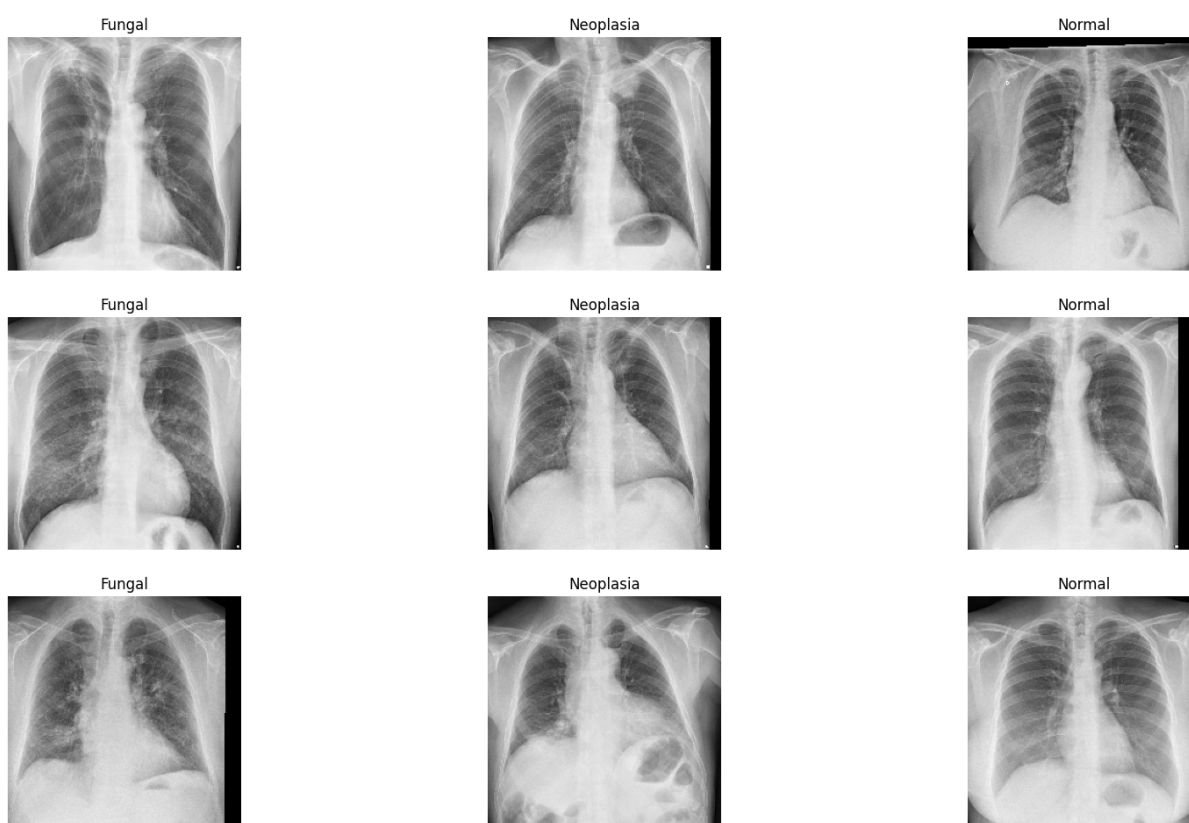
As imagens do *dataset* foram recebidas na extensão DICOM e foram convertidas para a extensão PNG. Para realização do treinamento, as imagens foram redimensionadas para uma altura e largura de 224x224 pixels e profundidade de cores RGB. Além disso, técnicas simples de *data augmentation* foram aplicadas como um giro horizontal e um pequeno *zoom* e, no caso da ResNet-101, um filtro de normalização específico para essa rede foi aplicado para melhorar a extração de características. Trabalhos como o de (OH et al., 2022) mostram que o uso correto de filtros pode aumentar a acurácia de redes residuais profundas. Entretanto, não foram utilizados filtros na ResNet-152 e nem nos modelos que realizam *ensemble* com média ponderada, pois não foi observado uma melhora desses modelos ao usar filtros.

### 3.3 Base de Dados

A avaliação foi feita sobre os dados coletados de pacientes do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto (HCFMRP-USP). A base de dados contém 714 imagens de radiografias de pacientes brasileiros, pertencentes aos grupos de doenças analisadas. Entretanto, imagens de perfil não foram consideradas para a análise, sobrando assim 376 exames de RX de pacientes com infecção fúngica e neoplasia e 333 exames de RX de pacientes saudáveis.

Para o modelo final, 496 imagens foram utilizadas para treinamento, 85 imagens foram utilizadas para validação e 128 imagens foram utilizadas para teste do modelo. A figura 3 mostra três exemplos das três classes avaliadas na presente pesquisa. Nos exemplos, nenhum filtro foi aplicado, mas as imagens estão no formato correto para entrada do modelo. Já a figura 4 mostra a quantidade de exames de RX por classe.

Figura 3 – Exemplos de imagens presentes na base de dados.



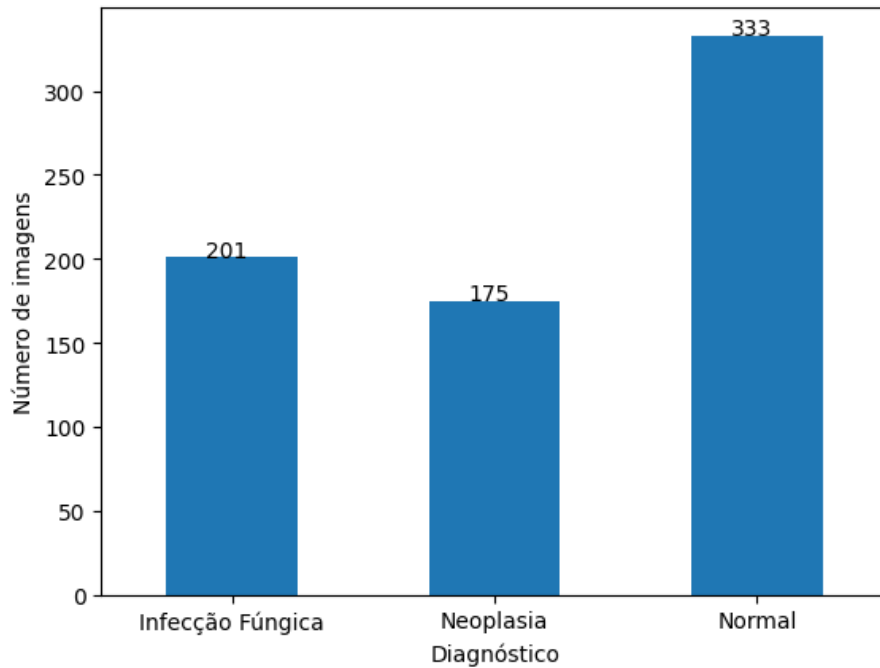
Elaborada pelo Autor.

### 3.4 Métricas

Os métodos propostos e desenvolvidos neste trabalho foram avaliados por meio das seguintes métricas utilizadas como padrão em outros trabalhos encontrados na literatura:



Figura 4 – Distribuição de imagens por classe.



Elaborada pelo Autor.

**Categorical Cross-Entropy Loss:** Também chamada de *Softmax Loss*, calcula o erro entre a probabilidade da classe prevista pelo modelo e a classe real da imagem. Sua fórmula pode ser vista na equação 3.1:

$$-\sum_i^C t_i \log(f(s)_i), \quad (3.1)$$

onde  $C$  é o número de classes,  $t_i$  representa a probabilidade real da classe,  $s$  representa um vetor de probabilidade para cada classe e  $f(s)_i$  é a probabilidade prevista pelo modelo para uma classe  $i$  após passar por uma função de ativação *softmax* usando a fórmula 3.2:

$$f(s)_i = \frac{e^{s_i}}{\sum_j^C e^{s_j}}, \quad (3.2)$$

onde  $s_i$  é a probabilidade da classe prevista,  $s_j$  é a probabilidade de cada uma das classes e  $e$  é a base de um logaritmo natural;

**Acurácia:** Diz respeito à quantidade de *labels* que o modelo acertou em relação à todas as *labels* e é calculado como na equação 3.3:

$$Accuracy = \frac{\sum TP + TN}{\sum TP + FP + FN + TN}, \quad (3.3)$$

onde  $TP$  representa os verdadeiros positivos (*True Positives*),  $TN$  os verdadeiros negativos (*True Negatives*),  $FP$  os falsos positivos (*False Positives*) e  $FN$  os falsos negativos (*False Negatives*);

**Precisão:** Diz respeito à quantidade de verdadeiros positivos em relação aos verdadeiros positivos e falsos positivos e sua fórmula é dada pela equação 3.4:

$$Precision = \frac{\sum TP}{\sum TP + FP}; \quad (3.4)$$

**Recall:** Diz respeito à quantidade de verdadeiros positivos em relação aos verdadeiros positivos e falsos negativos como na fórmula em 3.5:

$$Recall = \frac{\sum TP}{\sum TP + FN}; \quad (3.5)$$

**F1 Score:** Média harmônica entre a precisão e o *recall* e inclui as duas métricas anteriores com a mesma importância usando a fórmula 3.6:

$$F1\ Score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (3.6)$$

**Matriz de confusão:** Uma tabela que permite a visualização do desempenho de um modelo classificatório (LIANG, 2022). Ela apresenta de forma detalhada o resultado da classificação, comparando as previsões do modelo com os valores reais dos dados. Através dela, podemos obter as variáveis para o cálculo da acurácia, precisão e *Recall*.

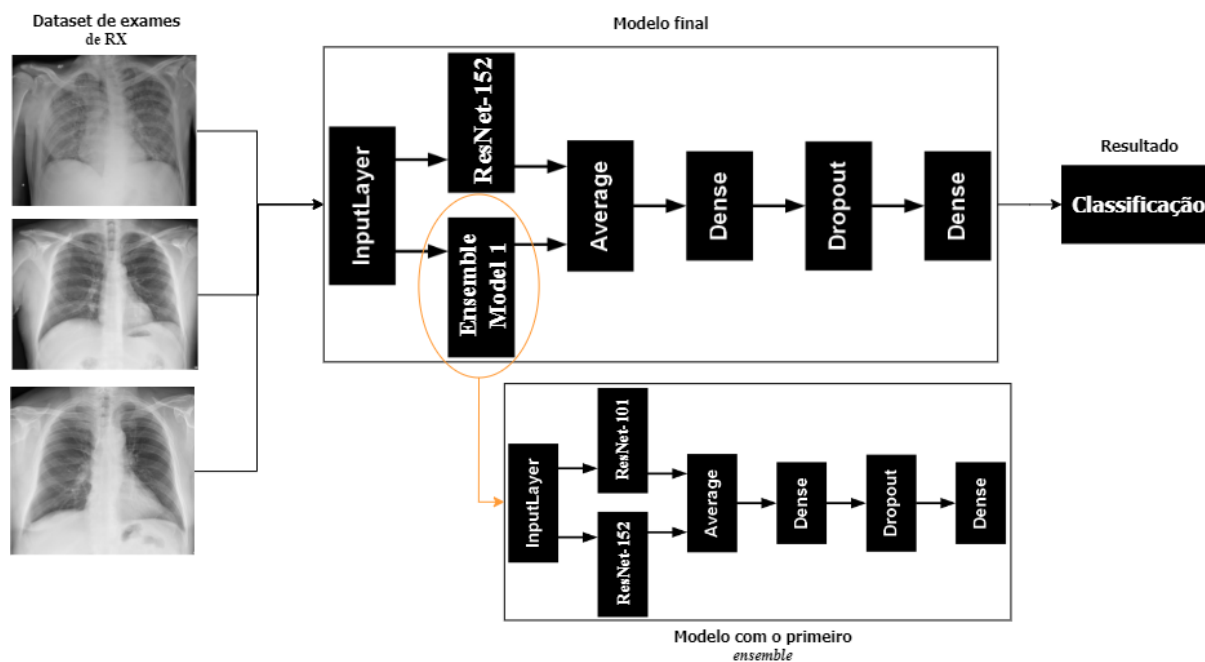
### 3.5 Modelo Proposto

Para lidar com o problema proposto e levando em consideração a limitação do tamanho da base de dados, neste trabalho, propomos testar diferentes Redes Neurais como ResNet, VGG, DenseNet e LeNet semelhante à trabalhos que buscavam aplicar aprendizado profundo para detecção de Covid-19 (KHAN et al., 2022) mesclado com diferentes pré-processamento das imagens de RX do tórax e a técnica de *Ensemble*, que se mostrou eficaz na classificação de múltiplas doenças em exames de RX (TRAN et al., 2021).

Como as ResNets demonstraram excelente desempenho em trabalhos semelhantes e são capazes de extrair detalhes mais minuciosos na área de processamento de imagens, será realizado o *ensemble* que faça a média ponderada de diferentes tipos de ResNet, como a ResNet-50, ResNet-101 e ResNet-152. O método de *ensemble* será utilizado para extrair a média de cada um dos modelos. Após o treinamento desse modelo, será realizado outro *ensemble* com uma outra rede.

Espera-se que utilizando o *ensemble* em uma rede que já passou pelo processo de *ensemble learning*, o modelo poderá extrair dados mais relevantes das imagens de RX e alcançar um desempenho melhor no momento de classificar os exames. A figura 5 dá uma visão geral do modelo proposto.

Figura 5 – Visão geral do modelo com duplo *ensemble* proposto.



Fonte: Elaborado pelo Autor.

Como pode ser observado na figura 5, uma imagem de exame de RX do tórax entra na rede e passa por duas redes com as camadas congeladas, ou seja, seus pesos não mudam durante o treinamento. Uma das redes é uma ResNet-152, enquanto a outra foi chamada de *Ensemble Model 1*, que, analisando mais profundamente, trata-se de um *ensemble* de uma ResNet-152 com uma ResNet-101 (também congeladas) onde são extraídas suas médias. O modelo final utiliza as médias do *Ensamble Model 1* e das ResNet-152 para realizar a classificação da imagem na forma de uma probabilidade.

## 4 Resultados e Discussão

Nesta seção, os resultados obtidos pelo método proposto serão expostos. Para a confecção do modelo final exposto no capítulo 3, é necessário treinar as redes individualmente com o melhor pré-processamento de imagem para cada uma. As redes que obtiverem os melhores resultados, farão parte do *Ensemble Model 1* e esse último estará com o melhor modelo individual no modelo final.

### 4.1 Treinamento dos modelos com duas classes

Inicialmente, houve uma tentativa de treinar os modelos permutando as classes dois à dois, ou seja, normais e infecções fúngicas, normais e neoplasia e infecções fúngicas e neoplasia. A ideia consiste em obter três redes que fossem capazes de diferenciar bem duas classes para que no final fosse realizado um *ensemble* com as três redes para o treinamento com três classes.

A ResNet-50 se mostrou ideal para ser utilizada em duas classes, visto que ela têm menos parâmetros para serem ajustados em relação à ResNet-101 e a ResNet-152. Vale ressaltar, que quando os testes com duas classes foram realizados, apenas 103 das 201 imagens de infecções fúngicas estavam disponíveis.

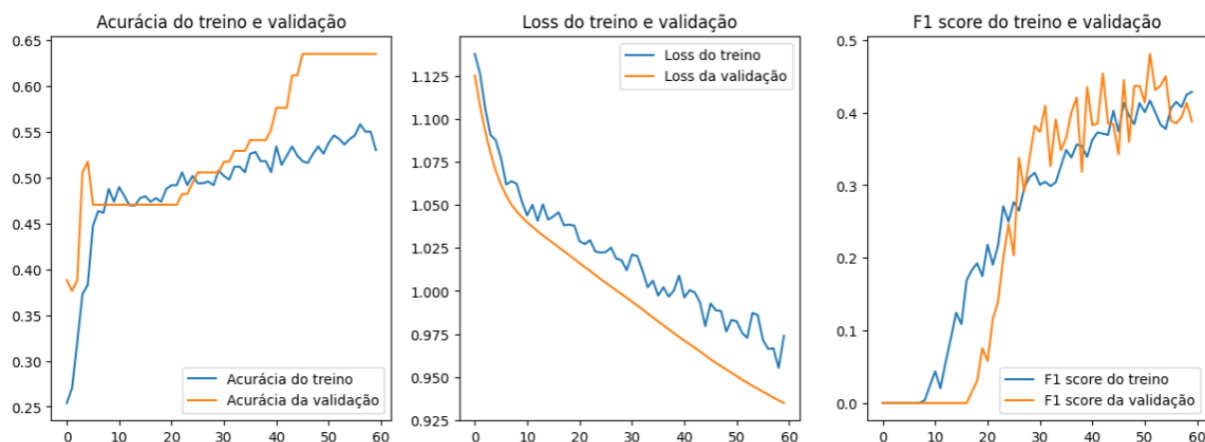
Após serem treinados por 40 épocas, os modelos de ResNet-50 responsáveis por distinguir pessoas saudáveis de uma das doenças analisadas atingiram acurácia superior a 97% no conjunto de treino e superiores a 86% no conjunto de teste. Entretanto, após treinar o modelo responsável pela classificação das infecções fúngicas e neoplasia por 500 épocas, ele foi capaz de atingir apenas 71% de acurácia no conjunto de teste.

Esperava-se que ao realizar *ensemble* dos três modelos citados anteriormente para classificação de todas as classes juntas, seriam obtidas métricas melhores do que os modelos treinados separadamente, mas após treinar o modelo por 50 épocas, ele começou a estagnar e alcançou uma acurácia de 60% no conjunto de teste.

A figura 6 mostra a evolução das métricas durante o treinamento e validação do modelo de *ensemble* das três ResNets e a tabela 1 compara as métricas obtidas no conjunto de teste dos três modelos treinados com duas classes e o treinado com 3 classes.

Analisando a matriz de confusão da figura 7, é possível observar que o modelo não foi capaz de prever corretamente nenhuma classe com a *label* neoplasia, justificando o *recall* tão baixo. Apesar da boa precisão para distinguir normais e infecções

Figura 6 – Evolução da Acurácia, *Loss* e *F1 Score* ao longo de 60 épocas nos conjuntos de treino e validação.



Elaborada pelo Autor.

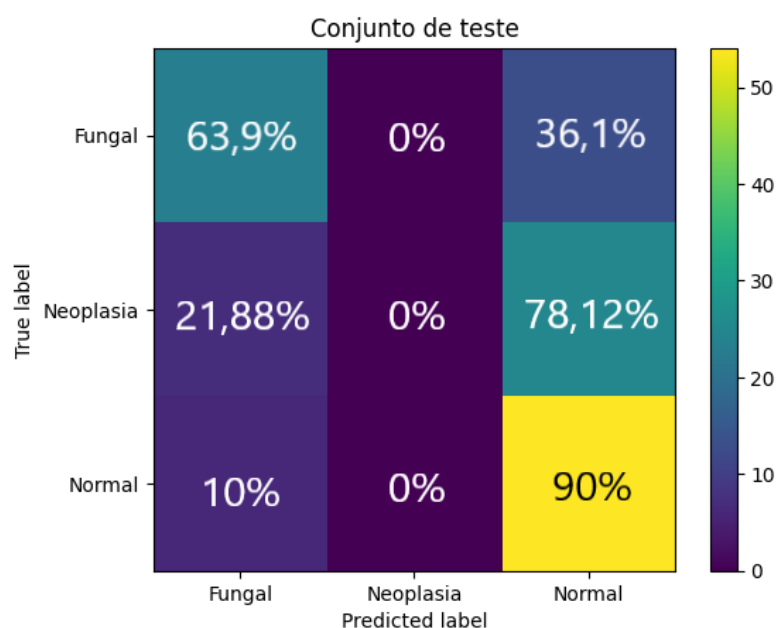
Tabela 1 – Métricas obtidas ao avaliar os modelos utilizando o conjunto de teste. Nota-se que o *Recall* foi extremamente baixo no modelo com *ensemble*.

Modelo/Desempenho no Teste	Acurácia	Loss	Precisão	Recall	F1 Score
ResNet (normais e fungos)	88,64%	0,3883	88,64%	88,64%	88,64%
ResNet (normais e neoplasia)	86,27%	0,8627	86,27%	86,27%	86,27%
ResNet (fungos e neoplasia)	71,43%	1,2308	71,43%	71,43%	71,43%
<b>Ensemble dos modelos</b>	60,16%	0,9433	69,64%	30,47%	42,39%

Fonte: Elaborada pelo Autor.

fúngicas, concluiu-se que essa estratégia não era viável e era necessário que todas redes fossem previamente treinadas com as três classes.

Figura 7 – Matriz de confusão do modelo com *ensemble* no conjunto de teste.



Elaborada pelo Autor.

## 4.2 Treinamento dos modelos com três classes

Ao todo, 6 redes foram treinadas combinando diferentes pré-processamentos das imagens. A tabela 2 mostra detalhadamente as melhores métricas obtidas por cada rede ao serem treinadas com as três classes e o respectivo pré-processamento utilizado na base de dados. Os modelos de ResNet tiveram os melhores desempenhos, excepcionalmente a ResNet-101 e ResNet-152, que foram os escolhidos para a realização do primeiro *ensemble* que compõe o modelo final.

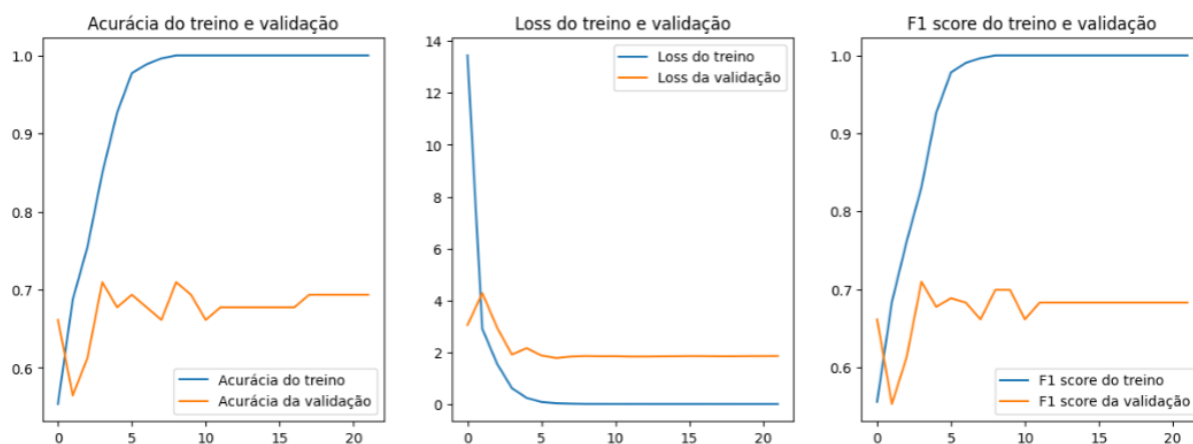
Tabela 2 – Métricas obtidas ao avaliar os modelos utilizando o conjunto de teste. Todas as redes, com exceção da LeNet vieram com os pesos pré-treinados na base de dados Imagenet.

Modelo/Desempenho no Teste	Acurácia	Loss	Precisão	Recall	F1 Score	Pré-processamento
DenseNet	56,00%	3,2343	56,00%	56,00%	56,00%	Data Augmentation + normalização
LeNet	53,00%	0,9537	66,67%	26,00%	37,41%	Data Augmentation
ResNet-50	66,00%	2,5387	66,00%	66,00%	66,00%	Data Augmentation
ResNet-101	74,14%	1,1069	74,78%	74,14%	74,46%	Normalização
ResNet-152	83,10%	1,6195	83,10%	83,10%	83,10%	Data Augmentation
VGG-19	67,00%	0,7442	72,09%	62,00%	66,67%	Data Augmentation

Fonte: Elaborada pelo Autor.

O treinamento das redes residuais foi consideravelmente rápido, visto que ocorreu uma estagnação da *loss* e da acurácia após 15 épocas como pode ser observado nas figuras 8 e 9. Em testes realizados com apenas duas classes, a VGG-19 havia demonstrado um desempenho bem alto em termos de acurácia, apesar de possuir uma grande *loss*, mas com três classes, essa eficiência não foi mantida e por essa razão, apenas as redes residuais foram selecionadas para fazer parte do *Ensemble Model 1*.

Figura 8 – Evolução da Acurácia, *Loss* e *F1 Score* ao longo de 20 épocas nos conjuntos de treino e validação da ResNet-101.



Elaborada pelo Autor.

Figura 9 – Evolução da Acurácia, *Loss* e *F1 Score* ao longo de 100 épocas nos conjuntos de treino e validação da ResNet-152. Apesar da tentativa de treinar a rede por mais épocas, não há ganho de desempenho após 15 épocas.



Elaborada pelo Autor.

### 4.3 Primeiro *ensemble*

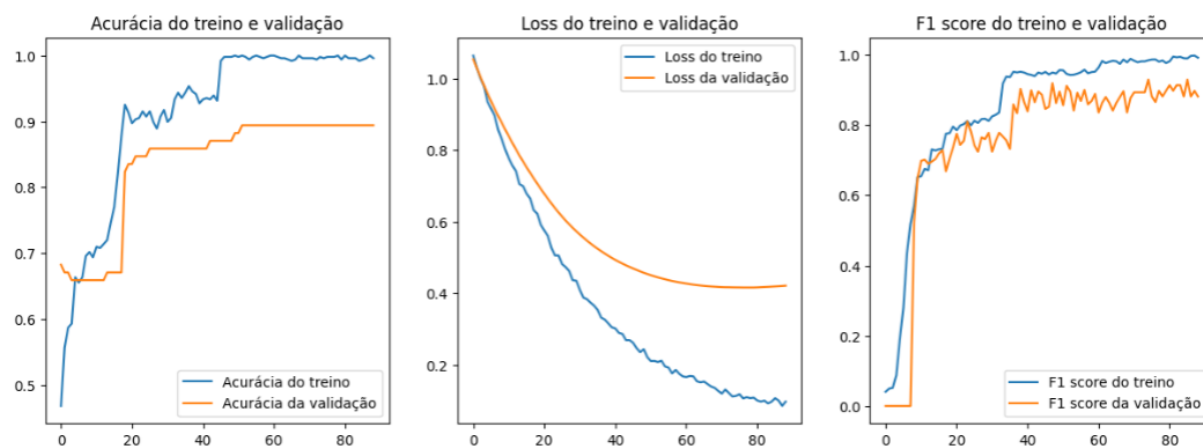
Como já dito anteriormente, as duas melhores redes da secção 4.2 serão utilizadas nesta secção para a realização da classificação em paralelo. O método de *ensemble* utilizado será a média ponderada das duas redes residuais. Serão adicionadas três camadas, sendo elas uma densa, uma de *dropout* e uma última densa responsável pela classificação. Todos os parâmetros das ResNet-101 e ResNet-152 serão congelados e os únicos parâmetros treináveis serão as últimas camadas do modelo.

O modelo foi treinado por 80 épocas, mas não apresentou melhorias após cerca de 60 épocas como pode ser visto na figura 10. Nota-se também que a curva de aprendizado das métricas foi mais natural do que os modelos citados anteriormente. Os valores obtidos no conjunto de treino ficaram bem próximos do conjunto de validação, mostrando que não ocorreu *overfitting*.

Em relação ao conjunto de teste, o modelo elaborado nesta secção também obteve resultados superiores às redes individuais e ao *ensemble* realizado com as redes residuais treinadas com duas classes. A figura 11 mostra a quantidade de *labels* que o modelo acertou e errou. Dessa vez, todas as classes foram aprendidas. Usando a matriz de confusão, é possível ver que a maioria dos exames de neoplasia preditos de maneira errada, foram preditos como exames sem nenhuma anomalia, enquanto exames que apresentavam sinais de infecção fúngica foram preditos como neoplasia.

Conclui-se que o modelo é capaz de determinar que há uma anomalia nos exames de infecções fúngicas, mas em alguns casos não sabe prever a doença exata. No caso dos exames de neoplasia, ainda há uma dificuldade para extrair a informação

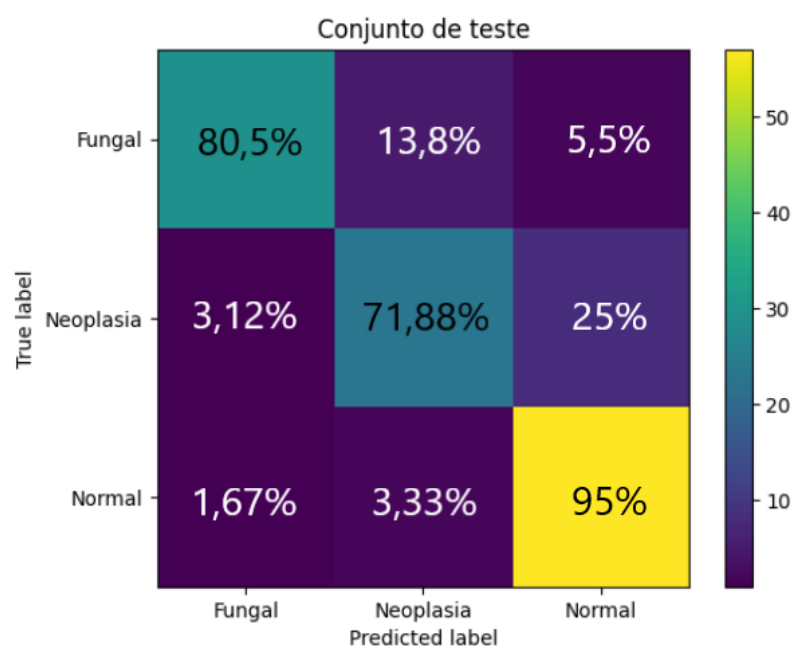
Figura 10 – Evolução da Acurácia, *Loss* e *F1 Score* ao longo de 80 épocas nos conjunto de treinos e validação do *Ensemble Model 1*.



Elaborada pelo Autor.

de que há uma anomalia. Exames de pacientes saudáveis tiveram a maior taxa de acertos.

Figura 11 – Matriz de confusão do modelo com *ensemble* da ResNet-101 e ResNet-152 no conjunto de teste.



Elaborada pelo Autor.

#### 4.4 Modelo final com *ensemble*

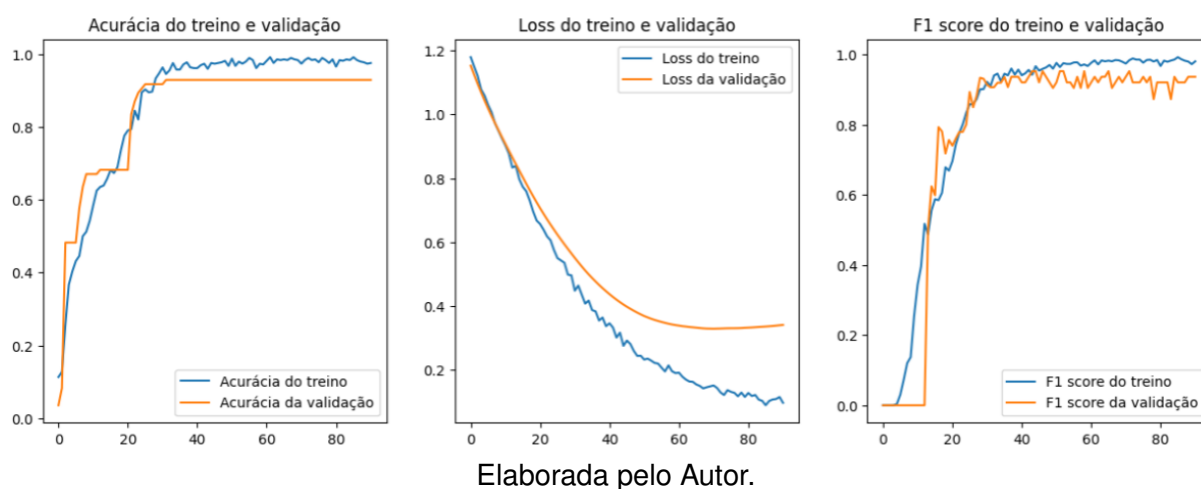
Para a elaboração do modelo proposto no capítulo 3, foi utilizado o modelo da secção 4.3 e uma ResNet-152. Ambos as redes foram congeladas e, novamente, as camadas densas e de *dropout* foram adicionadas após a realização da média



ponderada das probabilidades dos modelos. No total, o modelo conta com mais de 236 milhões de parâmetros congelados e 115 parâmetros para serem treinados.

O modelo também foi treinado por 80 épocas, mas atingiu seu ápice após 60 épocas como pode ser observado na figura 12. Ao analisar a evolução do modelo, percebe-se que a validação e treino ficaram mais próximas do que no modelo da secção anterior. Além disso, esse modelo obteve a menor *loss* e maior acurácia no conjunto de validação

Figura 12 – Evolução da Acurácia, *Loss* e *F1 Score* ao longo de 80 épocas nos conjuntos de treino e validação do modelo final com duplo *ensemble*.



Uma melhora também foi obtida no conjunto de teste, como atestado pela matriz de confusão da figura 13. Houve uma diminuição dos erros de predição em relação aos exames de infecções fúngicas e neoplasia, mostrando que o modelo ficou mais robusto do que qualquer um dos anteriores.

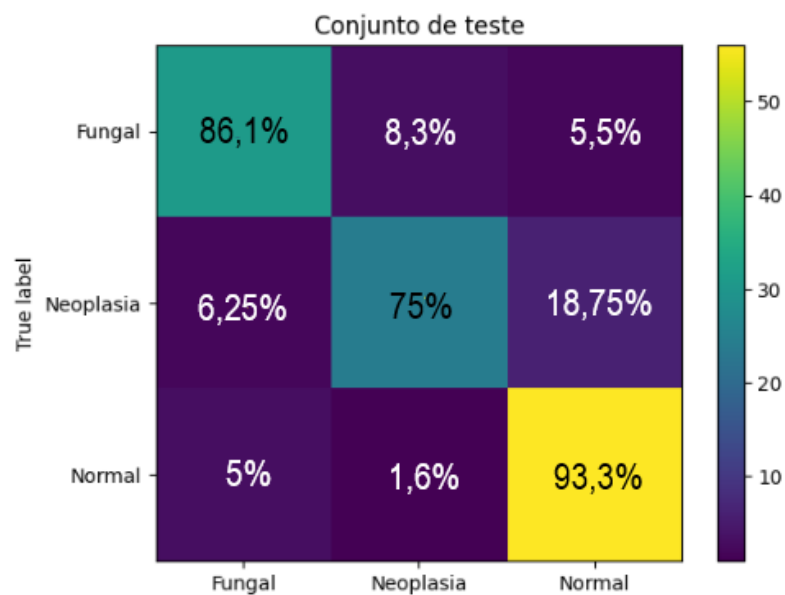
As tabelas a seguir resumem os resultados obtidos nos testes executados. A tabela 3 compara o desempenho das redes residuais e os modelos com *ensemble* no conjunto de treino. Já a tabela 4 diz respeito à todas as métricas obtidas pelo modelo com duplo *ensemble* em todos os conjuntos da base de dados analisada na presente pesquisa.

Tabela 3 – Métricas obtidas ao avaliar os modelos utilizando o conjunto de teste. O duplo *ensemble* obteve as melhores métricas no *dataset* avaliado.

Modelo/Desempenho no Teste	Acurácia	Loss	Precisão	Recall	F1 Score
<b>ResNet-101</b>	74,14%	1,1069	74,78%	74,14%	74,46%
<b>ResNet-152</b>	83,10%	1,6195	83,10%	83,10%	83,10%
<b>Ensemble Model 1</b>	85,16%	0,466	86,40%	84,38%	85,38%
<b>Modelo Final</b>	86,72%	0,4325	86,72%	86,72%	86,72%

Fonte: Elaborada pelo Autor.

Figura 13 – Matriz de confusão do modelo final com duplo *ensemble*.



Elaborada pelo Autor.

Tabela 4 – Métricas obtidas ao avaliar o modelo final nos conjuntos de treino, validação e teste. Além da alta taxa de acertos no treino, o modelo final obteve uma acurácia consideravelmente mais alta no conjunto de validação.

Conjunto/Métricas	Acurácia	Loss	Precisão	Recall	F1 Score
Treinamento	97,51%	0,1439	98,77%	97,37%	98,06%
Validação	92,94%	0,3281	93,71%	93,71%	93,71%
Teste	86,72%	0,4325	86,72%	86,72%	86,72%

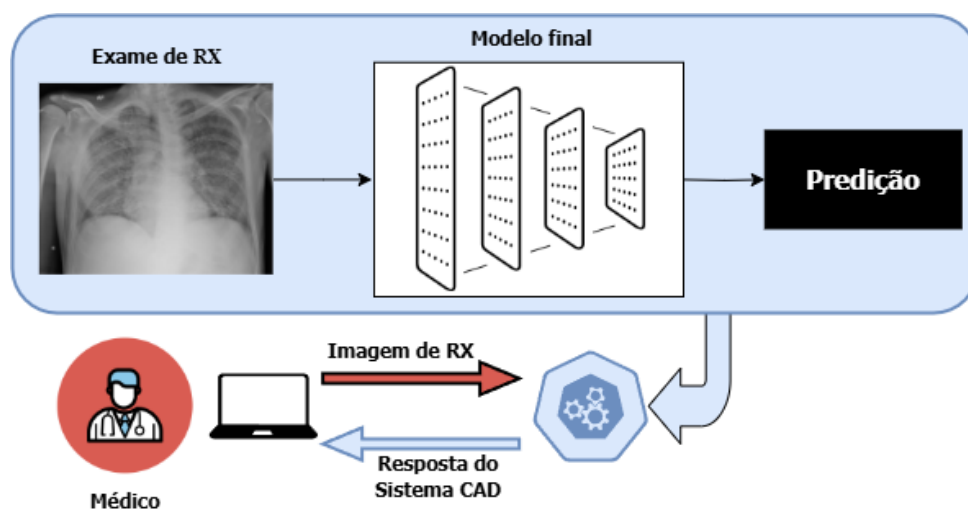
Fonte: Elaborada pelo Autor.

## 5 Aplicação de Auxílio ao Diagnóstico

Concluiremos o presente estudo com o desenvolvimento de uma aplicação prática de auxílio ao diagnóstico utilizando o modelo proposto no capítulo 3 e com os detalhes e resultados esmiuçados no capítulo 4. A intenção do aplicação é simular um possível uso da rede neural construída, mostrando como ela pode ajudar no processo de diagnóstico do profissional.

A aplicação foi projetada para rodar em sistemas Windows e roda apenas localmente sem a necessidade de conexão com a *internet*. O funcionamento básico do sistema pode ser observado na figura 14. A escolha de desenvolver o CAD como um aplicativo para o sistema operacional evita possíveis incompatibilidades de navegadores caso a aplicação fosse feita para web. Além disso, o porte para outros sistemas operacionais, como Mac OS e Linux seriam relativamente simples.

Figura 14 – Diagrama do funcionamento da Aplicação CAD.



Fonte: Elaborada pelo autor.

Explicando a figura 14, a ideia é que um profissional da área da saúde envie uma imagem de RX do Tórax de um paciente para o sistema. A aplicação fará uso do modelo desenvolvido no capítulo 4 para realizar a predição de uma possível enfermidade. A predição é, então, exibida para o profissional de saúde juntamente com dados estatísticos sobre o que o modelo conseguiu extrair do exame. Esses dados servirão para que o profissional saiba o quão confiante o sistema está de sua resposta e poderá prosseguir da maneira que julgar adequada com base nisso.

A aplicação foi desenvolvida para ser simples e intuitiva para mostrar o potencial de sistemas CAD para melhorar a detecção de doenças pulmonares através de radiografias do tórax.

## 5.1 Processo de Desenvolvimento

A aplicação foi desenvolvida em Python usando principalmente a biblioteca `customtkinter` para a aparência mais moderna da Interface de Usuário. Visto que o modelo também foi desenvolvido em Python com a biblioteca `Tensorflow`, como detalhado no capítulo 3, a integração entre a interface visual interativa e a rede responsável pela predição dos exames de RX foi relativamente simples.

Ao todo, foram criadas três telas, incluindo um Menu Principal que contém um breve resumo sobre a aplicação e a pesquisa, a tela de diagnóstico, onde o usuário carrega a imagem para ser analisada e recebe o resultado na mesma tela e, por fim, uma tela de créditos com alguns dados sobre os responsáveis pelo projeto. Todas têm opções de um modo escuro ou modo claro, como será exemplificado na próxima seção.

Depois que a interface como um todo foi desenvolvida, o modelo, que foi salvo em duas extensões por questões de segurança (`h5` e `keras`), foi implementado. Quando se salva um modelo treinado em algum desses formatos, todos os parâmetros, pesos e métricas atribuídos ao modelo na hora do seu treinamento são guardados e podem ser recuperados. O modelo é carregado quando o usuário envia a primeira imagem e a primeira vez demora cerca de 5 a 7 segundos para mostrar o resultado, mas todas as outras iterações são praticamente instantâneas.

## 5.2 Interfaces da Aplicação

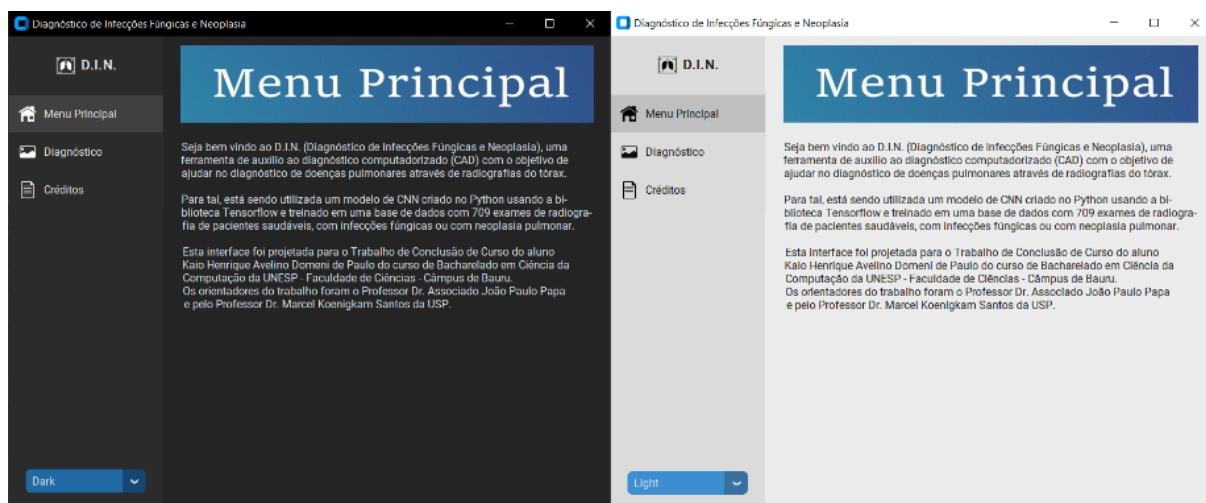
Nesta seção, será abordado com detalhes cada uma das páginas da aplicação, começando pelo menu principal, que pode ser visto na figura 15. Como dito anteriormente, o menu contém um texto que introduz um pouco sobre a aplicação que foi denominada D.I.N. (Diagnóstico de Infecções Fúngicas e Neoplasia).

O menu lateral é fixo e é composto pelo título da aplicação CAD, uma barra de navegação para acessar as páginas do programa e uma caixa de seleção do modo de exibição (escuro, claro ou padrão do sistema). A página selecionada fica sempre à direita, com o título na parte superior e o conteúdo na parte central.

Para melhor visualização, as próximas páginas serão mostradas apenas no modo escuro, mas a figura 15 exemplifica o comportamento da aplicação nos dois modos de exibição.

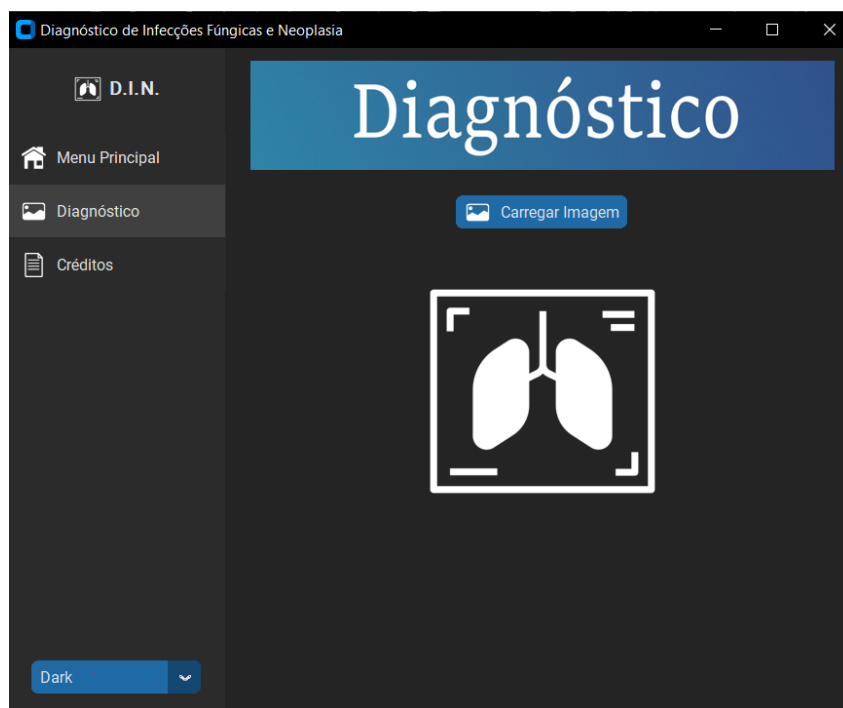
A página de diagnóstico é a de maior interesse. É nessa página que o modelo entra em ação. O *design* da página pode ser visto na figura 16. Quando nenhuma imagem de radiografia está sendo exibida, um template fica no centro da página e, logo acima, um botão para enviar uma imagem de RX do tórax para o sistema.

Figura 15 – Tela do Menu Principal. À esquerda, a interface no modo escuro e à direita, a interface no modo claro.



Fonte: Elaborada pelo autor.

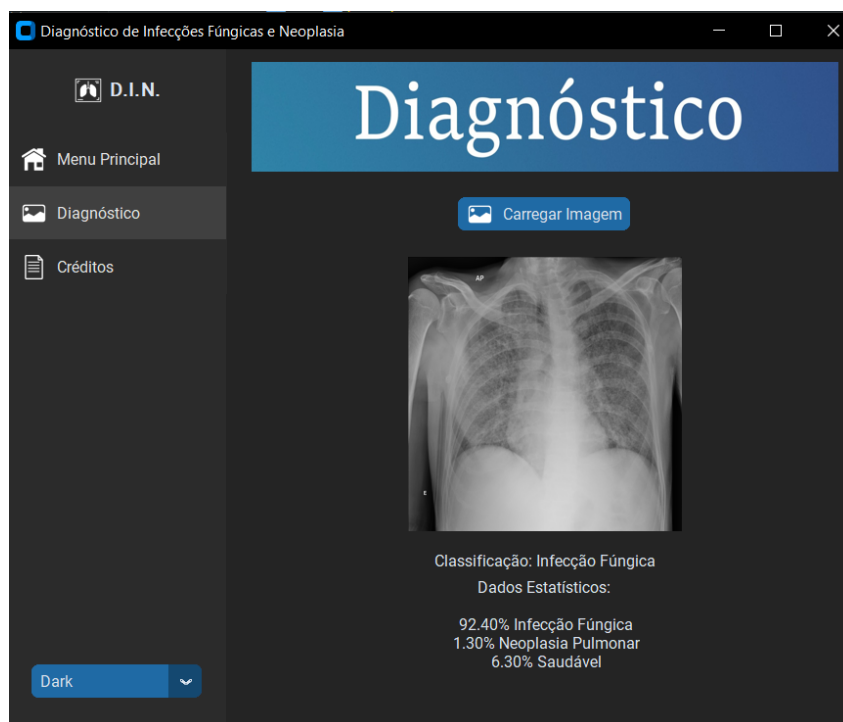
Figura 16 – Tela de Diagnóstico sem uma imagem carregada.



Fonte: Elaborada pelo autor.

Após o carregamento de uma imagem do computador do usuário, o exame é exibido na parte central na área do *placeholder* como pode ser observado na figura 17. Abaixo da imagem, fica a predição do modelo juntamente com informações estatísticas que podem ser interessantes para o médico. Isso porquê a saída da rede neural é um conjunto de probabilidades que aquela imagem tem de ser uma das três classes, no caso, saudável, infecção fúngica e neoplasia.

Figura 17 – Tela de Diagnóstico com uma imagem carregada.



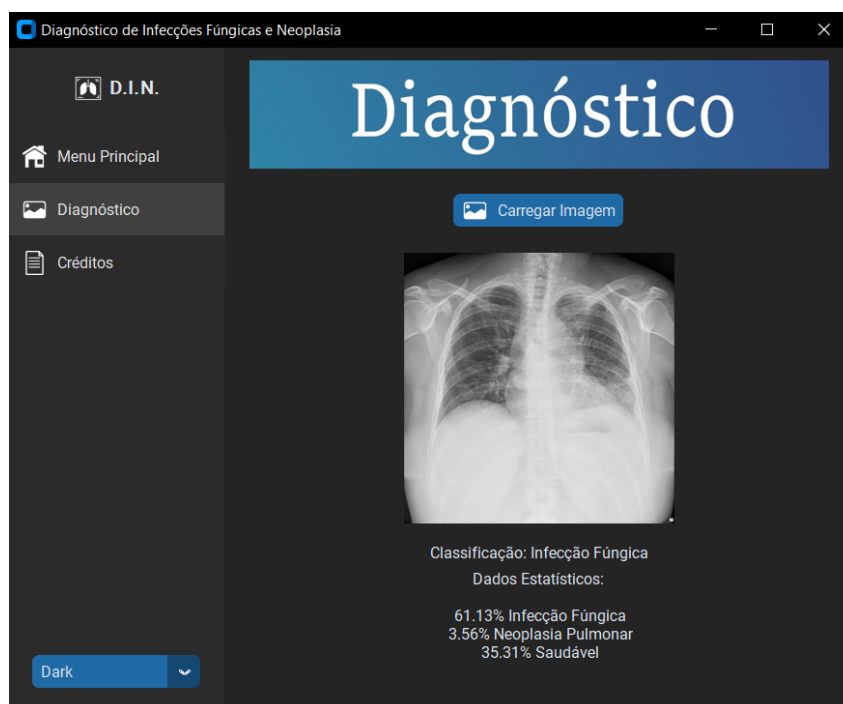
Fonte: Elaborada pelo autor.

Na maioria dos casos, o modelo realizada a predição com uma alta probabilidade da classe correta, ou seja, mantendo a *loss* baixa. Mas em excessões como ocorreu com o exame de RX exibido na figura 18, o modelo não realiza a predição com tanta precisão. A imagem de radiografia em questão foi anotada como infecção fúngica, mas o modelo previu 35% de chance de ser um exame de um paciente saudável, ou seja, uma *loss* maior do que o ideal. Informações desse tipo podem levar o profissional da saúde a realizar uma análise mais minuciosa para realizar o diagnóstico final.

Por fim, a página de créditos traz algumas informações sobre os responsáveis pela pesquisa. Suas ocupações, nomes e fotos ficam dispostos lado a lado na parte superior da página como pode ser visto na figura 19.

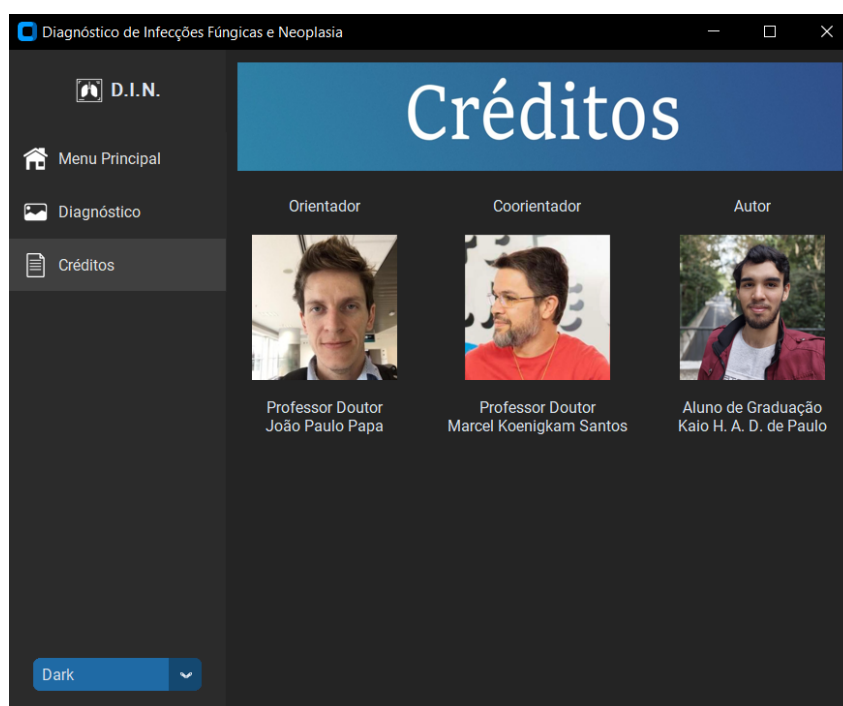
Vale ressaltar que ao utilizar a página de diagnóstico, mesmo que o usuário vá para outras páginas, ao retornar, o exame de RX carregado juntamente com as informações de predição ainda estarão disponíveis e só serão apagadas caso o usuário feche a aplicação.

Figura 18 – Tela de Diagnóstico com uma imagem carregada e estatísticas que denotam incerteza.



Fonte: Elaborada pelo autor.

Figura 19 – Tela de Créditos.



Fonte: Elaborada pelo autor.

## 6 Considerações Finais

Neste trabalho, foi proposto um modelo de CNN com *ensemble* duplo para classificar imagens de RX do tórax. Combinando métodos já estabelecidos no campo do aprendizado de máquina e processamento de imagens, a abordagem se mostrou promissora para lidar com o problema multiclasse apresentado, principalmente considerando as limitações do *dataset* para aplicação de CNNs de aprendizado profundo.

Algumas melhorias poderiam ser realizadas no modelo proposto através de mais testes. Dentre eles, pode-se citar pré-processamentos mais eficazes nas imagens, como aumento de contraste e segmentação para aprimorar a extração de características relevantes para as redes, diferentes tipos de *ensemble* como o por voto, *boosting* e concatenação, além de diferentes combinações de redes individuais para a constituição das redes com *ensemble* aplicado.

A aplicação CAD desenvolvida como uma extensão prática da pesquisa mostrou de forma eficiente como seria uma interface para inferência de diagnósticos de forma automatizada para imagens de RX do tórax. Seu uso poderia ser expandido para abranger outras doenças pulmonares que se apresentam como opacidades em exames do tórax.

Trabalhos futuros podem explorar mais a fundo a proposta de aplicar *ensemble* sobre uma rede que já foi treinada com esse tipo de aprendizado. Além disso, o modelo pode ser testado com mais classes para que seja testada sua capacidade de generalização e técnicas de mapas de calor podem ser utilizadas para entender quais fatores na imagem são essenciais para a tomada de decisão do modelo, além de auxiliar o profissional de saúde a identificar áreas de interesse quando a imagem é carregada na aplicação desenvolvida.



# Referências

- ANGELINI, E.; SHAH, A. *Using Artificial Intelligence in Fungal Lung Disease: CPA CT Imaging as an Example*. Springer Science and Business Media LLC, 2021. 733–737 p. Disponível em: <http://dx.doi.org/10.1007/s11046-021-00546-0>.
- AZEVEDO-MARQUES, P. M. de. Diagnóstico auxiliado por computador na radiologia. *Radiologia Brasileira*, v. 34, n. 5, p. 285–293, 2001.
- BARBOSA, F. M. C.; CANUTO, A. M. de P. Classification of chest x-ray images using machine learning and histogram of oriented gradients. 2021.
- DONG, X.; YU, Z.; CAO, W.; SHI, Y.; MA, Q. *A survey on ensemble learning*. Springer Science and Business Media LLC, 2019. 241–258 p. Disponível em: <http://dx.doi.org/10.1007/s11704-019-8208-z>.
- EAGEN, J. W.; LEWIS, E. J. Glomerulopathies of neoplasia. *Kidney International*, v. 11, n. 5, p. 297–306, 1977. ISSN 0085-2538. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0085253815317397>.
- GUILLAMET, C. V.; HSU, J. L.; DHILLON, G.; GUILLAMET, R. V. Pulmonary infections in immunocompromised hosts. *Journal of Thoracic Imaging*, v. 33, n. 5, p. 295–305, 2018.
- HASAN, R.; MUHAMMAD, S.; ISLAM, R. Recent advancement of deep learning techniques for pneumonia prediction from chest x-ray image. *Medical Reports*, v. 7, p. 100106, 2024. ISSN 2949-9186. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2949918624000718>.
- HE, K.; ZHANG, X.; REN, S.; SUN, J. Deep residual learning for image recognition. arXiv, 2015a. Disponível em: <https://arxiv.org/abs/1512.03385>.
- HE, K.; ZHANG, X.; REN, S.; SUN, J. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. arXiv, 2015b. Disponível em: <https://arxiv.org/abs/1502.01852>.
- KHAN, E.; REHMAN, M. Z. U.; AHMED, F.; ALFOUZAN, F. A.; ALZAHIRANI, N. M.; AHMAD, J. Chest x-ray classification for the detection of covid-19 using deep learning techniques. *Sensors*, v. 22, n. 3, 2022. ISSN 1424-8220. Disponível em: <https://www.mdpi.com/1424-8220/22/3/1211>.
- LAKHANI, P.; SUNDARAM, B. Deep learning at chest radiography: Automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology*, v. 284, n. 2, p. 574–582, 2017.
- LI, S.; JIAO, J.; HAN, Y.; WEISSMAN, T. Demystifying resnet. arXiv, 2016. Disponível em: <https://arxiv.org/abs/1611.01186>.
- LI, Z.; LIU, F.; YANG, W.; PENG, S.; ZHOU, J. A survey of convolutional neural networks: Analysis, applications, and prospects. *IEEE Transactions on Neural*

*Networks and Learning Systems*, v. 33, n. 12, p. 6999–7019, 2022. Disponível em: [pubmed.ncbi.nlm.nih.gov/34111009/](https://pubmed.ncbi.nlm.nih.gov/34111009/).

LIANG, J. Confusion matrix: Machine learning. *POGIL Activity Clearinghouse*, v. 3, n. 4, Dec. 2022. Disponível em: <https://pac.pogil.org/index.php/pac/article/view/304>.

LUDERMIR, T. B. *Inteligência Artificial e Aprendizado de Máquina: estado atual e tendências*. FapUNIFESP (SciELO), 2021. 85–94 p. Disponível em: <http://dx.doi.org/10.1590/s0103-4014.2021.35101.007>.

OH, J.; KIM, H.; BAIK, S.; HONG, C.; LEE, K. M. Batch normalization tells you which filter is important. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*. [S.l.: s.n.], 2022. p. 2645–2654.

SAJWAN, A.; MISHRA, G. Comparative analysis of resnet and densenet for differential cryptanalysis of speck 32/64 lightweight block cipher. In: CHATURVEDI, A.; HASAN, S. U.; ROY, B. K.; TSABAN, B. (Ed.). *Cryptology and Network Security with Machine Learning*. Singapore: Springer Nature Singapore, 2024. p. 495–504. ISBN 978-981-97-0641-9.

SANTA MONICA STUDIO. *God of War Ragnarök*. 2022. [S.l.]. Jogo eletrônico.

SCHABATH, M. B.; COTE, M. L. Cancer progress and priorities: Lung cancer. *Cancer Epidemiology and Prevention Biomarkers*, v. 28, n. 10, p. 1563–1579, 2019.

SHINDE, P. P.; SHAH, S. A review of machine learning and deep learning applications. In: *2018 Fourth International Conference on Computing Communication Control and Automation (ICCCUBEA)*. [S.l.: s.n.], 2018. p. 1–6.

SILVA, R. F. e. Capítulo 8 - infecções fúngicas em imunocomprometidos. *Jornal Brasileiro de Pneumologia*, v. 36, n. 1, p. 142–147, 2010.

TANG, Y.-X.; TANG, Y.-B.; PENG, Y.; YAN, K.; BAGHERI, M.; REDD, B. A.; BRANDON, C. J.; LU, Z.; HAN, M.; XIAO, J.; SUMMERS, R. M. *Automated abnormality classification of chest radiographs using deep convolutional neural networks*. Springer Science and Business Media LLC, 2020. Disponível em: <http://dx.doi.org/10.1038/s41746-020-0273-z>.

TRAN, T. T.; PHAM, H. H.; NGUYEN, T. V.; LE, T. T.; NGUYEN, H. T.; NGUYEN, H. Q. Learning to automatically diagnose multiple diseases in pediatric chest radiographs using deep convolutional neural networks. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*. [S.l.: s.n.], 2021. p. 3314–3323.

WADA, D. T.; RODRIGUES, J. A. H.; SANTOS, M. K. Aspectos técnicos e roteiro de análise da radiografia de tórax. *Medicina (Ribeirão Preto)*, v. 52, n. 1, p. 5–15, 2019.

WANG, F.; LI, X.; WEN, R.; LUO, H.; LIU, D.; QI, S.; JING, Y.; WANG, P.; DENG, G.; HUANG, C.; DU, T.; WANG, L.; LIANG, H.; WANG, J.; LIU, C. *Pneumonia-Plus: a deep learning model for the classification of bacterial, fungal, and viral pneumonia based on CT tomography*. Springer Science and Business Media LLC, 2023. 8869–8878 p. Disponível em: <http://dx.doi.org/10.1007/s00330-023-09833-4>.

ZHANG, Y.-h.; HU, X.-f.; MA, J.-c.; WANG, X.-q.; LUO, H.-r.; WU, Z.-f.; ZHANG, S.; SHI, D.-j.; YU, Y.-z.; QIU, X.-m.; ZENG, W.-b.; CHEN, W.; WANG, J. *Clinical Applicable AI System Based on Deep Learning Algorithm for Differentiation of Pulmonary Infectious Disease*. Frontiers Media SA, 2021. Disponível em: <http://dx.doi.org/10.3389/fmed.2021.753055>.