

Classificação de Radiografias Torácicas

Relatório Final - NoIA

Informações Gerais

Data: 12/12/2025
Membros: Gustavo Nascimento Pavanelli - 242028065
Gabriel Campello Dalbuquerque Lima - 241008110
Marcel Capistrano Almeida Rodrigues - 242035391
Matéria: Noções de Inteligência Artificial
Grupo: Grupo 2

1 Introdução

1.1 Contextualização

A análise de imagens médicas, especificamente radiografias de tórax, é uma etapa crítica no diagnóstico de diversas doenças respiratórias. Neste contexto, sistemas de auxílio ao diagnóstico baseados em *Deep Learning* surgem como ferramentas poderosas para oferecer uma "segunda opinião" automatizada, priorizando casos críticos e aumentando a eficiência clínica. O objetivo deste trabalho é desenvolver um classificador capaz de distinguir entre pulmões saudáveis, com pneumonia e com tuberculose.

1.2 Fonte de Dados

O conjunto de dados utilizado provém de uma coleção pública disponível no Kaggle Dataset de Imagens de Raio-X do Tórax, contendo milhares de imagens organizadas em subpastas de treino, validação e teste. A distribuição das imagens apresenta um desbalanceamento natural, comum em datasets médicos, com variações significativas na quantidade de exemplos para cada classe (Normal, Pneumonia e Tuberculose).

1.2.1 Estrutura do Dataset

Classe	Treinamento	Validação	Teste
Normal	7,263	900	925
Pneumonia	4,674	570	580
Tuberculose	8,513	1,064	1,064
Total	20,450	2,534	2,569

Tabela 1: Quantidade de imagens por etapa.

2 Metodologia e Implementação

A implementação foi realizada utilizando a biblioteca PyTorch, seguindo um pipeline de ingestão, processamento, modelagem e avaliação.

2.1 Pré-processamento e Aumento de Dados

Para adequar as imagens à rede neural e aumentar a robustez do modelo, foram aplicadas as seguintes transformações:

- **Padronização:** Redimensionamento (*Resize*) de todas as imagens para 512×512 pixels.
- **Conversão de Canal:** Conversão para escala de cinza (*Grayscale*), visto que a informação de cor é irrelevante em Raio-X, reduzindo a complexidade computacional.

- **Data Augmentation:** Durante o treino, aplicou-se equalização de histograma, rotações leves (5 graus), transformações afins e variações de brilho/contraste para evitar *overfitting*.

2.2 Arquitetura do Modelo

Optou-se pela arquitetura **DenseNet169** utilizando a técnica de *Transfer Learning*. O modelo foi pré-treinado no dataset ImageNet, o que permitiu o aproveitamento de extratores de características já convergidos.

As modificações realizadas na arquitetura original foram:

1. Adaptação da primeira camada convolucional para aceitar 1 canal de entrada (Grayscale) em vez de 3 (RGB).
2. Substituição da camada de classificação final (*Fully Connected*) por uma nova camada linear com 3 saídas, correspondentes às classes do problema.

2.3 Estratégia de Treinamento

O treinamento utilizou a função de custo *CrossEntropyLoss* e o otimizador Adam. Para otimizar o tempo e evitar o *overfitting*, implementou-se um mecanismo de **Early Stopping** com paciência de 5 épocas, monitorando a perda no conjunto de validação.

3 Experimentos e Tentativas Preliminares

Antes de chegar à arquitetura final descrita acima, foram realizados experimentos exploratórios para definir os melhores hiperparâmetros e estratégias.

3.1 Tentativa 1: Arquitetura ResNet18

Inicialmente, testou-se uma arquitetura mais leve, a ResNet18. Embora o treinamento fosse mais rápido, o modelo apresentou dificuldade em distinguir características sutis entre Pneumonia e Tuberculose, resultando em uma acurácia de validação estagnada em torno de 65%. A troca para a DenseNet169, com suas conexões densas, permitiu uma melhor reutilização de *features*, elevando o desempenho.

3.2 Tentativa 2: Treinamento sem Data Augmentation

Em uma segunda iteração, treinou-se a DenseNet169 sem as técnicas de aumento de dados (*Data Augmentation*). O modelo convergiu rapidamente no conjunto de treino (acurácia próxima a 99%), mas falhou em generalizar para o conjunto de validação (alto *overfitting*). A introdução de rotações e variações de contraste foi crucial para obrigar a rede a aprender a estrutura pulmonar e não apenas memorizar os pixels das imagens de treino.

3.3 Tentativa 3: Estratégias de Congelamento de Camadas

Experimentou-se o congelamento e descongelamento progressivo de camadas da rede (*Freezing/Unfreezing*), focando inicialmente apenas no classificador e liberando gradualmente os blocos densos. O objetivo era preservar as *features* do ImageNet, mas a adaptação não superou o treinamento completo da rede.

3.4 Tentativa 4: Ajuste de Hiperparâmetros (Fine-tuning)

Realizou-se uma varredura nos hiperparâmetros, reduzindo a taxa de aprendizado (*Learning Rate*) e testando diferentes tamanhos de lote (*Batch Size*) para buscar um mínimo global mais estável. Embora tenha reduzido a oscilação da perda, a acurácia de validação não apresentou ganhos substanciais.

3.5 Tentativa 5: Pré-processamento Avançado

Visando superar a estagnação da acurácia em 77%, implementaram-se técnicas de pré-processamento mais agressivas e normalizações distintas para realçar contrastes pulmonares. Apesar de aumentar a variabilidade dos dados de entrada, o modelo manteve a dificuldade em distinguir casos normais com maior precisão.

4 Resultados e Discussão

O modelo final teve seu treinamento interrompido precocemente na época 18 pelo mecanismo de *Early Stopping*, garantindo o uso dos pesos que geraram a menor perda de validação.

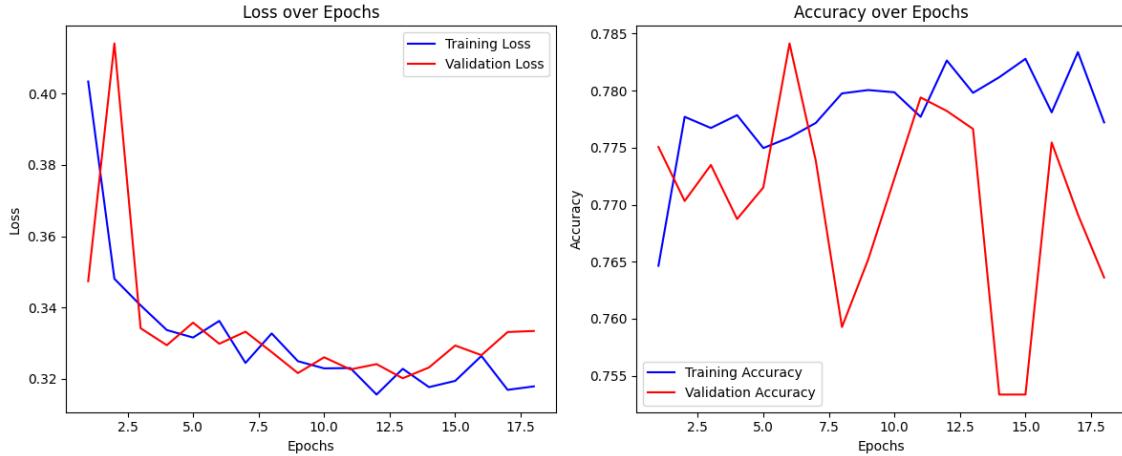


Figura 1: Análises por Épocas.

4.1 Métricas de Desempenho

No conjunto de teste, o modelo alcançou uma **Acurácia Global de 77%**. A análise detalhada por classe revela o comportamento do classificador:

Classe	Precision	Recall (Sensibilidade)	F1-Score
Normal	0.92	0.38	0.54
Pneumonia	0.79	0.95	0.86
Tuberculose	0.71	1.00	0.83

Tabela 2: Resultados de classificação no conjunto de teste.

4.2 Análise da Matriz de Confusão

A matriz de confusão demonstra um comportamento claro do modelo: ele prioriza a detecção de doenças.

- **Alto Recall para Doenças:** O modelo identificou 100% dos casos de Tuberculose e 95% dos casos de Pneumonia. Falsos negativos para doenças são extremamente raros.
- **Falsos Positivos:** O baixo recall da classe Normal (0.38) indica que muitos pacientes saudáveis foram classificados incorretamente como doentes (principalmente como Tuberculose, com 424 casos).

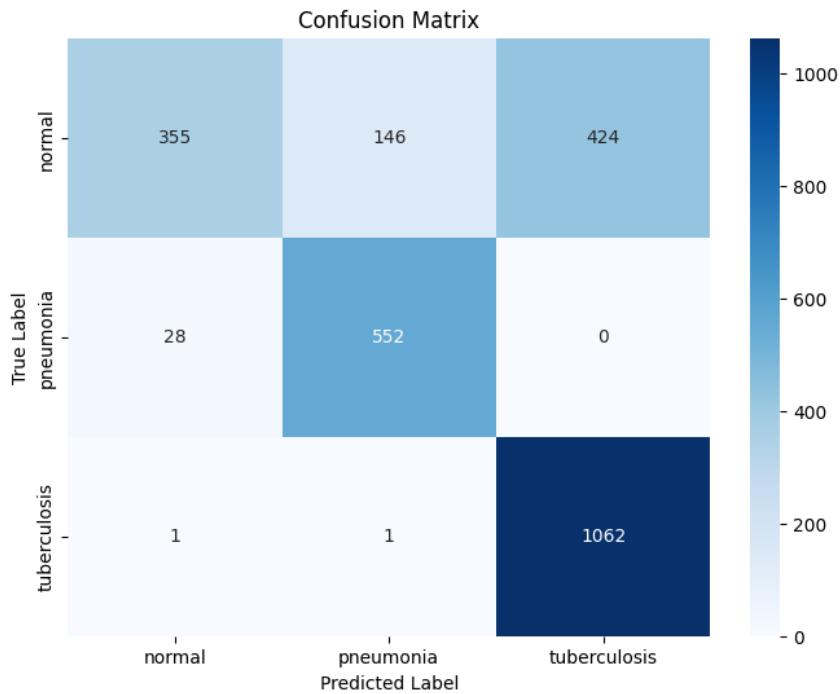


Figura 2: Matriz de Confusão.

5 Conclusão

O projeto demonstrou a eficácia do uso de *Transfer Learning* com a DenseNet169 para a extração de características em imagens de Raio-X. O classificador resultante atua como uma ferramenta de **alta sensibilidade**, ideal para cenários de triagem onde o custo de não detectar uma doença (Falso Negativo) é muito maior do que o custo de um alarme falso (Falso Positivo).

Para trabalhos futuros, sugere-se focar no equilíbrio da classe "Normal" através de técnicas de balançamento de dados (*Oversampling*) ou funções de perda ponderada (*Weighted Loss*), além da aplicação de técnicas de explicabilidade (Grad-CAM) para validar as regiões de interesse observadas pela rede.