RECONHECIMENTO DE PADRÕES EM IMAGENS DE RAIO-X DE TÓRAX PARA CLASSIFICAÇÃO DE PNEUMONIAS BACTERIANA E VIRAL UTILIZANDO TÉCNICAS DE APRENDIZADO DE MÁQUINA

Leonardo de Oliveira Godoy (PIBIC/CNPq/FA/UEM), Yandre Maldonado e Gomes da Costa (Orientador). E-mail: yandre@din.uem.br

Universidade Estadual de Maringá, Centro de Ciências da Tecnologia, Maringá, PR.

Ciências Exatas e da Terra /Ciência da Computação

Palavras-chave: Classificação de Imagens; Redes Neurais Convolucionais; Radiografia.

RESUMO

O presente trabalho teve como objetivo a avaliação do emprego de técnicas de aprendizado de máquina na identificação e classificação de imagens de radiografia de tórax de pulmões saudáveis, com pneumonia bacteriana e com pneumonia viral. Para a realização do projeto, foi criado um código em Python utilizando a biblioteca PyTorch e foi empregado o uso de modelos de Redes Neurais Convolucionais pré-treinados. Uma base de dados disponível na plataforma Kaggle, com imagens de radiografia de pulmão, foi utilizada durante o estudo. As imagens passaram por uma etapa de pré-treinamento, seguido pelas etapas de treinamento e teste dos modelos. Foram escolhidas como métricas a acurácia, f1-score e matrizes de confusão. Posteriormente, foi aplicada uma técnica de *late-fusion*, a fim de combinar os resultados dos diferentes modelos em busca de resultados melhores. Por fim, foi executada uma técnica de *Explainable Artificial Inteligence*, responsável por gerar mapas de calor referentes aos pontos das imagens que os modelos mais se basearam para tomar as decisões. As métricas mostram que os modelos apresentaram bom desempenho, com cerca de 90% de acurácia e de f1-score. Além disso, o emprego da técnica de *late-fusion* contribuiu para melhora nos resultados, chegando a elevar as métricas para aproximadamente 94%. Por fim, os mapas de calor ajudaram no entendimento melhor do funcionamento dos modelos e mostraram que eles podem gerar resultados errôneos. Desse modo, o uso dos modelos pode auxiliar no trabalho dos médicos, mas não deve substituí-los.

INTRODUÇÃO

A pneumonia é uma infecção pulmonar que afeta os alvéolos pulmonares, fazendo com que se encham de fluido ou de pus. Ela pode ser causada por diferentes agentes, como vírus e bactérias e possui uma variedade de sintomas, podendo levar à internação em casos mais sérios. Dessa forma, é de suma importância realizar o diagnóstico nos estágios iniciais da doença, de forma rápida e eficaz. O diagnóstico se dá principalmente por meio da análise de imagens de radiografia do tórax, devido

ao seu baixo custo. Isso requer um treinamento adequado e o resultado pode variar de acordo com a interpretação do médico.
Diante disso, o presente trabalho busca avaliar o emprego de técnicas de aprendizado de máquina que possam auxiliar os profissionais no diagnóstico de pneumonia viral e bacteriana a partir de imagens de raio-x do tórax, de forma semelhante ao que foi feito pelo Lucas Teixeira no trabalho “Impact of Lung Segmentation on the Diagnosis of COVID-19 in Chest X-ray Images”.

MATERIAIS E MÉTODOS

Inicialmente, foi definida uma base de dados que seria a fonte das imagens utilizadas no decorrer do projeto. A base escolhida foi a “ChestX6: Multi-Class X-ray Dataset”, disponível gratuitamente na plataforma Kaggle. Ela contém 18036 imagens de raio-x de tórax em escala de cinza. As imagens escolhidas estão divididas em 3 categorias, sendo elas *Normal*, *Pneumonia-Bacterial* e *Pneumonia-Viral*. Além disso, estão organizadas em conjuntos de treino e de teste.
Para a realização dos experimentos, foi criado um código na linguagem Python, utilizando a biblioteca PyTorch, amplamente utilizada em trabalhos de machine learning. As imagens passaram por um pré-processamento, sendo normalizadas e redimensionadas para o tamanho 224x244.
A fim de avaliar o desempenho de diferentes redes pré-treinadas, foram escolhidos 7 modelos de Redes Neurais Convolucionais (CNNs): DenseNet121, EfficientNet, Inception-V3, MobileNet-V2, ResNet-18, ResNet-50 e VGG-16. Para os modelos, foi utilizado *transfer learning* com modelos pré-treinados da ImageNet, tendo a última camada ajustada para classificação de imagens de raio-x. Cada modelo passou por 10 épocas de treinamento, número escolhido considerando que os modelos já foram pré-treinados. O tempo decorrido no treinamento foi anotado. Além disso, foram adotadas como métricas de avaliação a acurácia, f1-score e matriz de confusão.
Posteriormente, foi adotada uma técnica de *late fusion*, a qual combina as saídas de diferentes redes para definir a classificação, a fim de buscar melhores resultados. A estratégia de combinação utilizada foi a regra da soma, que calcula a soma das predições feitas pelos modelos e seleciona a classe com maior predição acumulada.
Por fim, foi aplicada uma abordagem de Explainable Artificial Intelligence (XAI) a fim de compreender melhor a forma como os modelos tomaram as decisões. A técnica empregada foi o Grad-CAM, o qual gera mapas de calor referentes às regiões da imagem que foram mais relevantes para o trabalho do classificador.

RESULTADOS E DISCUSSÃO

A Tabela 1 mostra os resultados obtidos pelos modelos individualmente, bem como de uma combinação feita pelo *late fusion*. Os modelos apresentaram bons resultados tanto na acurácia quanto no f1-score, variando de 80% a 90%, aproximadamente. O tempo de treinamento foi variado entre os modelos. No geral, a combinação gerou resultados melhores, com casos que chegaram a 94% de acurácia e f1-score.

Tabela 1

Modelo	Acurácia	F1-score	Tempo
DenseNet121	0.8282	0.8322	13.41 min
EfficientNet	0.8947	0.8939	6.69 min
Inception-V3	0.9113	0.9110	20.51 min
MobileNet-V2	0.9080	0.9074	5.04 min
ResNet-18	0.8991	0.8981	4.20 min
ResNet-50	0.8780	0.8768	14.31 min
VGG-16	0.8891	0.8905	21.48 min
DenseNet121, EfficientNet, MobileNet-V2, ResNet-18, ResNet-50 e VGG-16	0.9401	0.9400	Soma do tempo dos modelos utilizados

Por fim, a Figura 1 mostra um exemplo de imagem gerada pelo Grad-CAM. A primeira imagem representa a imagem original, seguida pelo mapa de calor de cada modelo e o mapa de calor do ensamble, que representa a combinação dos mapas. Essas imagens auxiliam a observar as peculiaridades de cada modelo, bem como ajudam a compreender melhor o resultado do *late fusion*, já que os modelos que acertam a predição acabam "corrigindo" os que não acertaram. Além disso, é possível observar que, em muitas ocasiões, os modelos tomam suas decisões sem focar nos pulmões, que seriam as partes relevantes da imagem no diagnóstico de pneumonia. Isso corrobora os resultados encontrados por Teixeira, evidenciando que as métricas por si só não devem ser confiáveis e que os modelos podem estar aprendendo a identificar outros elementos da imagem, e não a pneumonia.

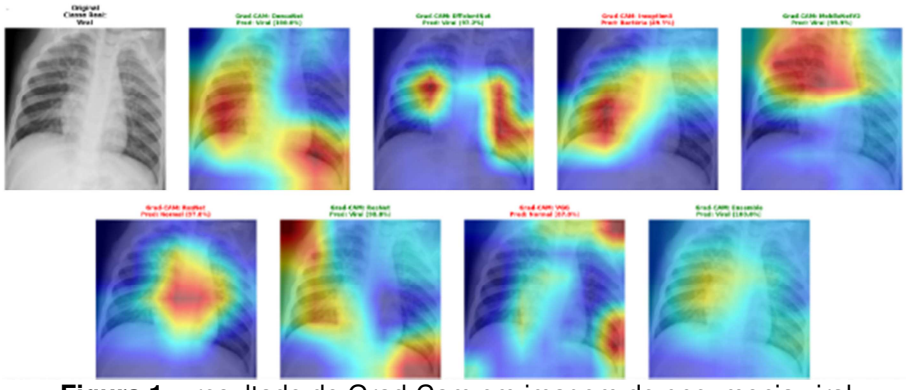


Figura 1 – resultado do Grad-Cam em imagem de pneumonia viral

CONCLUSÕES

Os resultados obtidos no projeto mostram que o uso de CNNs pode auxiliar na classificação de imagens de radiografia de pulmões saudáveis, com pneumonia bacteriana e pneumonia viral, considerando as boas métricas alcançadas. Além disso, a combinação de diferentes modelos se mostrou bem favorável na melhora dos resultados. Entretanto, conforme visto pelos mapas de calor obtidos pelo Grad-CAM, os modelos são suscetíveis a considerar partes das imagens que não deveriam no momento de tomar as decisões. Dessa forma, enquanto o uso dessas técnicas pode auxiliar e facilitar o trabalho dos médicos, elas não devem ser empregadas sozinhas de modo a substituir os especialistas.

AGRADECIMENTOS

Agradeço ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) por todo o apoio financeiro e técnico durante a execução do projeto.

REFERÊNCIAS

PEREIRA, R. M. et al. COVID-19 identification in chest X-ray images on flat and hierarchical classification scenarios. **Computer Methods and Programs in Biomedicine**, v. 194, 2020. Disponível em: https://doi.org/10.1016/j.cmpb.2020.105532. Acesso em: 30 ago. 2026.

ROSSI, R. Tópicos em Inteligência Artificial - Aula 10 - Redes Neurais Convolucionais para a Classificação de Imagens. **YouTube**, [202-]. Playlist. Disponível em: https://www.youtube.com/playlist?list=PLAHmHkSA6KkX7yrpL1GA-joyog9Z16N6V. Acesso em: 30 ago. 2026.

TEIXEIRA, L. O. et al. Impact of lung segmentation on the diagnosis and explanation of COVID-19 in chest X-ray images. **Sensors**, v. 21, n. 21, 7116, 2021. Disponível em: https://doi.org/10.3390/s21217116. Acesso em: 30 ago. 2026.