



Técnicas de Aprendizado Profundo para auxílio na detecção do COVID-19 em imagens de raios-X do tórax

Deep Learning techniques to aid in the detection of COVID-19 in chest X-ray images

Sthefanie Monica Premebida (orientado) *,

Marcella Scoczynski Ribeiro Martins (orientador)[†], Thiago Fellipe Ortiz de Camargo[‡],
Vinicios Roberto Soares[§], Cristhiane Gonçalves[¶], Virgínia Helena Varotto^{||}

RESUMO

Dado o grande número de casos COVID-19 em todo o mundo, uma solução prática para diminuir e aliviar a fila de pacientes nos hospitais e nos sistemas de saúde é bem-vinda. Diagnóstico rápido e confiável com base em ferramentas tecnológicas pode apoiar os profissionais médicos para gerenciar esta situação de gargalo, como o diagnóstico baseado em técnicas de imagem, que permite procedimentos não intrusivos. Neste artigo, propomos uma metodologia prática utilizando aprendizado profundo para detectar e classificar os pulmões afetados pelo COVID-19 por meio da radiografia de tórax. A arquitetura RetinaNet é considerada aqui. Esta arquitetura é uma detecção de objeto de um estágio usando perda focal frequentemente aplicada com objetos densos, pequenos e desequilibrados. Consideramos um conjunto de dados com 2.500 imagens para treinamento do modelo e 1.000 imagens para validar o modelo. Além disso, um conjunto de 1000 imagens de dois conjuntos de dados diferentes é aplicado para testar a abordagem do pipeline. Os resultados obtidos mostram uma pontuação de especificidade de 0,54, precisão de 0,68, evocação de 0,994 e mAP de 0,913. A alta pontuação de evocação explica que um paciente com COVID-19 será classificado corretamente.

Palavras-chave: COVID-19 Raio-X de pulmão Aprendizado profundo

ABSTRACT

Given the large number of COVID-19 cases around the world, a practical solution to decrease and relieve the queue of patients in the hospitals and in the health care systems is welcome. Fast and reliable diagnosis based on technological tools can support medical professionals to manage this bottleneck situation, such as the diagnostic based on image techniques, which allows non-intrusive procedures. In this paper, we propose a practical methodology using deep learning to detect and classify lung affected by COVID-19 using Chest X-ray radiography. RetinaNet architecture is considered here. This architecture is a one-stage object detection using *focal loss* often applied with dense, small and imbalance objects. We consider a dataset with 2500 images for model training and 1000 images to validate the model. Besides, a set of 1000 images from two different datasets are applied to test the pipeline approach. The obtained results show a specificity score of 0.54, precision of 0.68, recall of 0.994, and mAP of 0.913. The high recall score explains that a patient with COVID-19 will be classified correctly.

Keywords: COVID-19 detection. Chest-X ray images. DeepLearning

* Engenharia Elétrica; ✉ sthefanie@alunos.utfpr.edu.br.

† Departamento de Engenharia Elétrica; ✉ marcella@utfpr.edu.br.

‡ Engenharia Elétrica; ✉ tcamargo@alunos.utfpr.edu.br.

§ Engenharia Elétrica; ✉ viniciossoares@alunos.utfpr.edu.br.

¶ Departamento de Engenharia Elétrica; ✉ cristhianeg@utfpr.edu.br.

|| Departamento de Engenharia Elétrica; ✉ virginia@utfpr.edu.br.



1 INTRODUÇÃO

Doenças infecciosas acompanham a humanidade por milhares de anos e normalmente a causa de epidemias são desconhecidas, porém, com o progresso de várias áreas da ciência, métodos de prevenção desses tipos de doenças foram aprendidas e repassadas em todo o mundo (MARTINS et al., 1997, p.59-80). Em 2019 uma doença infecciosa surgiu e começou a assolar vários países, e ao invés de ser apenas uma epidemia, tomou a proporção de uma pandemia (WHO, 2021). A doença em questão é a COVID-19, causada pelo betacoronavírus SARS-CoV-2 (ISMAEL; ŞENGÜR, 2021).

As técnicas de imagem mais comuns que conseguimos identificar o COVID-19 são: tomografia computadorizada (TC), ultrassom dos pulmões (USL) e o raio-X de tórax (RXT), sendo esse último (RXT) o mais acessível, por ser mais barato, de ser feito pelo SUS (Sistema Único de Saúde) e ainda provém uma análise rápida, ou seja, é um exame que sai em pouco tempo depois de ter sido feito (ROY et al., 2020, p.2676-2687)(WANG, S. et al., 2020). Sabendo disso, os exames de RXT foram escolhidos como primeira técnica de imagem para auxiliar no diagnóstico de doenças respiratórias, bem como definir a severidade da doença para guiar na priorização de tratamento.

Com o auxílio das imagens para identificar os casos que precisam ser priorizados, houve uma minimização de saturação dos sistemas de saúde durante a pandemia da COVID-19 (OH; PARK; YE, 2020, p.2688-2700), porém existe uma falta de acurácia dos diagnósticos por conta de que muitos profissionais da saúde que foram para a linha de frente, não tinham experiência clínica suficiente para que a interpretação desses exames não tivesse uma acurácia elevada (CHEN et al., 2018, p.124-149).

Técnicas de visão computacionais, como redes neurais convolucionais, podem ser usadas no caso de processamento de imagens. As redes neurais convolucionais são redes que tem um conjunto de camadas de convolução e *maxpooling*, seguidas por n camadas densas de neurônios. Modelos de detecção de um estágio são empregados em tarefas de localização e classificação de objetos em imagens. A abordagem consiste na geração de um número fixo de propostas de localização, e cada uma tem uma probabilidade de realmente conter o objeto e a classe, então as propostas passam por um filtro de probabilidade de conter o objeto. Porém, esse tipo de abordagem tem problemas, como classificar o fundo da imagem como uma classe, e desse modo chegamos ao desbalanceamento de classes, pois em grande parte das imagens o fundo tem uma fração grande da imagem, o que leva o aprendizado não chegar no seu ótimo para detecção e classificação. Esse problema é chamado de *foreground-background imbalance* e pode ser amenizado utilizando algumas técnicas ou até outras arquiteturas de rede, como é o caso da RetinaNet (LIN; DOLLÁR et al., 2017, p.2117-2125).

2 MÉTODO

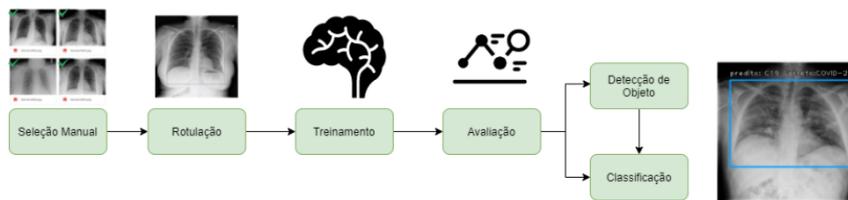
2.1 Redes Neurais e técnicas de Machine Learning aplicadas à Raio-X

Neste artigo, foram aplicadas redes neurais com camadas profundas (*deep learning*) ao invés dos métodos clássicos de aprendizado de máquina (ML), devido ao volume de dados e custo computacional. Os métodos clássicos têm um ponto de saturação em relação à quantidade de dados, algo que não acontece ao usar *Deep Learning* (DL), já que é feito para funcionar com Big Data. Além disso, o DL tem um desenvolvimento mais rápido de alto desempenho em sistemas computacionais paralelos, como *clusters* de GPU (ZHAO et al., 2019, p.3212-3232).

2.2 Pipeline

A Fig. 1 apresenta o fluxo de processamento da abordagem proposta. Primeiramente, 6688 imagens foram manualmente selecionadas das classes COVID-19 e Normal. Posteriormente, as regiões de interesse são anotadas nas imagens selecionadas, desenhando-se uma caixa delimitadora e se atribuindo a respectiva classe. Com as imagens e suas anotações, o modelo é treinado e posteriormente avaliado. A avaliação da performance mede a capacidade da rede de detectar e classificar pulmões dentre as classes propostas.

Figura 1 – Fluxo de processamento da abordagem do trabalho.



Fonte: Autoria própria (2021)

A aplicação de *pooling* nas regiões de interesse foi proposto por (HE et al., 2016, p.770-778) em que o autor utiliza *maxpooling* para converter as características dentro da região de interesse em uma para de características, provendo um algoritmo capaz de aprender a localização de um objeto e o classificar. Faz-se importante ressaltar que as caixas que representam as regiões de interesse não possuem tamanho ou formato fixo. As anotações foram feitas através de um programa de código aberto.

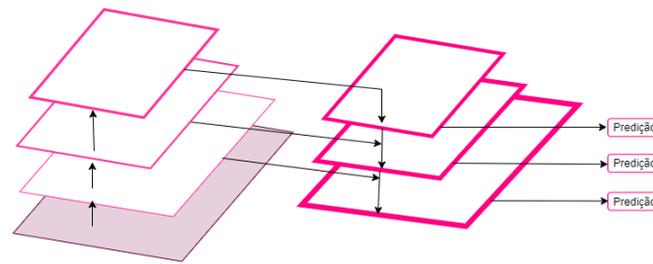
O modelo de aprendizado profundo aplicado é o RetinaNet (LIN; DOLLÁR et al., 2017, p.2117-2125), que é um detector de objetos de um estágio, munido de uma função de perda chamada *focal loss*. Esta função é uma *cross-entropy* dinamicamente escalada. À medida que o coeficiente de confiança aumenta, o fator de escala modula a perda para zero. O fator de escala ameniza o impacto das amostras com alto fator de confiança alto, assim, focando nos exemplos mais difíceis.

A implementação do modelo RetinaNet foi realizada com a ferramenta Keras, uma biblioteca de construção de fluxo computacional por blocos. Para utilizar tal estrutura, os arquivos XML precisam ser convertidos para CSV, em que todas as colunas são descritas pelo caminho para o arquivo de imagem, região de interesse e classe.

A técnica de transfer-learning é aplicada para minimizar o custo computacional para treinar o modelo, melhorar os resultados e evitar treinos desnecessários. A arquitetura RetinaNet adota a *Feature Pyramid Network* (FPN) como a espinha dorsal do modelo. A Fig. 2 apresenta a FPN proposta por (LIN; GOYAL et al., 2017, p.2980-2988), FPN é similar a camadas convolucionais estacadas, em que cada camada extrai características em diferentes escalas e níveis semânticos, um conceito importante na tarefa de reconhecimento de objetos.

Além da FPN, a ResNet50 também compõe a espinha dorsal. ResNet é uma abreviação para *Residual Network* (ISMAEL; ŞENGÜR, 2021), em que as camadas possuem uma conexão de atalho como funções residuais com referência à camada de entrada. As conexões residuais são importantes para estabilizar a rede evitando o problema do desaparecimento do gradiente (KESIM; DOKUR; OLMEZ, 2019, p.1-5). Isso descreve um problema durante o treinamento do modelo de aprendizado profundo: a rede é incapaz de propagar a informação do gradiente para as multicamadas por conta de ser um valor pequeno, o que torna difícil saber para que direção os pesos devem ser ajustados.

Figura 2 – Arquitetura da rede FNP (*Feature Pyramid Network*).



Fonte: Autoria própria (2021)

3 RESULTADOS

O estudo computacional é organizado em fases de treinamento e teste. A fase de treinamento apresenta o banco de imagens anotadas e os métodos de validação utilizados para avaliar os resultados. Na fase de teste o modelo é aplicado em imagens ainda nunca vistas. A parcela de treinamento é composta de 1500 casos de COVID-19 e 1000 casos normais, e a parcela de validação de 500 COVID-19 e 500 Normais. O modelo é treinado por 50 épocas, cada uma com 500 passos.

Na parcela de validação, a rede atinge precisão média (AP) de 0.9028 para a classe Normal e 0.9235 para COVID-19. Além disso, o modelo obteve uma precisão média média (mAP) de 0.9131 e uma mAP ponderada entre classes de 0.9134. Avaliando a performance do modelo através da matriz de confusão, Fig. 3, apenas 16% das imagens com COVID-19 e 15% Normais são classificados erroneamente. Estas percentagens garantem uma precisão de 85%, 84% de sensibilidade, e especificidade de 85%.

Figura 3 – Matriz de confusão dos resultados da rede.

Normal	0.81	0.19
Cl9	0.15	0.89
	Normal	Cl9

Fonte: Autoria própria (2021)

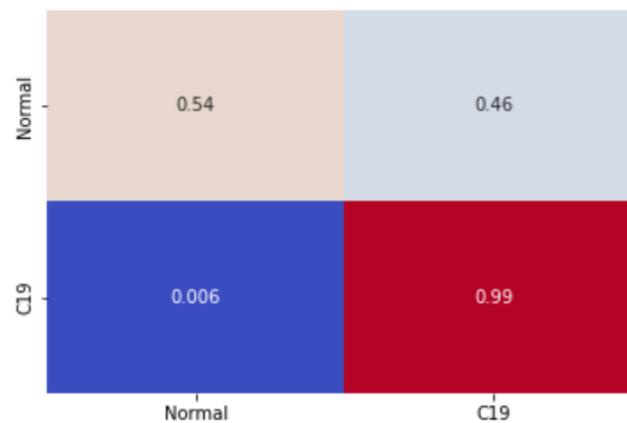
O mAP em paralelo com as métricas: precisão, sensibilidade e especificidades, demonstram que o modelo é factível para detectar e classificar pulmões dentre as classes COVID-19 e Normal em um banco de imagens de raio-X de tórax nunca vistos e com múltiplas características, referindo-se a questões de visão computacional.

A fase de teste nos permite confirmar se a saída é aceitável frente às questões do mundo real. Para tanto, dois bancos de imagens diferentes foram utilizados, A classe normal é composta de imagens aleatórias advindas

do National Institutes of Health (NIH) (WANG, X. et al., 2017, p.2097-2106) com imagens adquiridas antes do começo da pandemia de COVID-19. Para a classe de COVID-19, o banco de imagens do hmHospitales é utilizado de forma randômica.

O teste é realizado utilizando um banco de imagens com 1000 imagens: 500 para COVID-19 do banco hmHospitales, e 500 Normais do banco NIH. É importante ressaltar que as imagens que compõem a parcela de testes não foram expostas ao modelo na fase de treinamento.

Figura 4 – Matriz de confusão da FNP com os dados do banco hmHospitales



Fonte: Autoria própria (2021)

A fim de medir a performance do modelo no mundo real a matriz de confusão se apresenta, Fig. 4. A classe normal é erroneamente classificada 46% das vezes, já a classe COVID-19 apenas 0.006% das vezes. Este fato implica em um escore de especificidade de 54%, precisão de 68%, e sensibilidade de 99,4%.

O alto valor de sensibilidade sugere que os pulmões com positivo para COVID-19 são classificados corretamente, e se o pulmão é classificado como Normal, certamente o paciente não está infectado com COVID-19. Portanto, o modelo é capaz de assistir profissionais da saúde em questões de diagnóstico de COVID-19, oferecendo uma troca custo-computacional-acurácia razoável.

4 CONCLUSÕES

Este trabalho apresenta uma metodologia prática para detectar e classificar pulmões por meio de imagens de raio-X de tórax entre a classe COVID-19 e Normal, munindo-se da arquitetura de detecção de um estágio. A metodologia, baseada na RetinaNet, abrange o problema tempo de execução e acurácia, explorando esta troca crucial e o refinando como um problema do mundo real e aplicável. No total, 3500 imagens anotadas por profissionais da medicina foram consideradas: 2500 compuseram a parcela de treinamento e 1000 para validação, além de outras 1000 imagens para a parcela de teste. O resultado alcançado pela metodologia proposta implica em uma especificidade de 54%, precisão de 68%, e sensibilidade de 99,4%. O escore alto de sensibilidade mostra que um paciente com COVID-19 pode ser classificado corretamente, sendo assim, o modelo é capaz de auxiliar os profissionais da saúde com diagnóstico de COVID-19, oferecendo uma ótima troca custo-computacional-acurácia.



AGRADECIMENTOS

Os autores agradecem à Fundação Araucária (FA), pela bolsa de Iniciação Tecnológica e Inovação (Edital 02/2010 - PROPPG - PROGRAMA INSTITUCIONAL DE INICIAÇÃO CIENTÍFICA DA UTFPR) e também à pela participação no Programa de Voluntariado em Iniciação Científica e Tecnológica (Edital PROPPG – PIVICT 2020/2021) concedida aos acadêmicos.

REFERÊNCIAS

CHEN, Jianguo et al. A disease diagnosis and treatment recommendation system based on big data mining and cloud computing. **Information Sciences**, Elsevier, v. 435, p. 124–149, 2018.

HE, Kaiming et al. Deep residual learning for image recognition. In: PROCEEDINGS of the IEEE conference on computer vision and pattern recognition. [S.l.: s.n.], 2016. P. 770–778.

ISMAEL, Aras M; ŞENGÜR, Abdulkadir. Deep learning approaches for COVID-19 detection based on chest X-ray images. **Expert Systems with Applications**, Elsevier, v. 164, p. 114054, 2021.

KESIM, Ege; DOKUR, Zumray; OLMEZ, Tamer. X-ray chest image classification by a small-sized convolutional neural network. In: IEEE. 2019 scientific meeting on electrical-electronics & biomedical engineering and computer science (EBBT). [S.l.: s.n.], 2019. P. 1–5.

LIN, Tsung-Yi; DOLLÁR, Piotr et al. Feature pyramid networks for object detection. In: PROCEEDINGS of the IEEE conference on computer vision and pattern recognition. [S.l.: s.n.], 2017. P. 2117–2125.

LIN, Tsung-Yi; GOYAL, Priya et al. Focal loss for dense object detection. In: PROCEEDINGS of the IEEE international conference on computer vision. [S.l.: s.n.], 2017. P. 2980–2988.

MARTINS, Roberto de Andrade et al. Contágio: história da prevenção das doenças transmissíveis. **São Paulo: Moderna**, p. 59–80, 1997.

OH, Yujin; PARK, Sangjoon; YE, Jong Chul. Deep learning covid-19 features on cxr using limited training data sets. **IEEE transactions on medical imaging**, IEEE, v. 39, n. 8, p. 2688–2700, 2020.

ROY, Subhankar et al. Deep learning for classification and localization of COVID-19 markers in point-of-care lung ultrasound. **IEEE transactions on medical imaging**, IEEE, v. 39, n. 8, p. 2676–2687, 2020.

WANG, Shuo et al. A fully automatic deep learning system for COVID-19 diagnostic and prognostic analysis. **European Respiratory Journal**, Eur Respiratory Soc, v. 56, n. 2, 2020.

WANG, Xiaosong et al. Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In: PROCEEDINGS of the IEEE conference on computer vision and pattern recognition. [S.l.: s.n.], 2017. P. 2097–2106.

WHO, WORLD HEALTH ORGANIZATION -. **Coronavírus disease (COVID-19)**. [S.l.: s.n.], 2021.
Disponível em: [🔗](#).

ZHAO, Zhong-Qiu et al. Object detection with deep learning: A review. **IEEE transactions on neural networks and learning systems**, IEEE, v. 30, n. 11, p. 3212–3232, 2019.