

Redes Neurais e Detecção de COVID-19 a partir de Imagens de raios-X do tórax

Palavras-Chave: Redes neurais profundas, aprendizado de máquina, COVID-19

Autores/as:

João Paulo Christofolletti - FEEC

Prof. Dr. Romis Ribeiro de Faissol Attux (orientador) - FEEC

INTRODUÇÃO:

A doença respiratória COVID-19, reconhecida em 2019 a partir de suspeitas iniciais de pneumonia de etiologia desconhecida, tornou-se rapidamente uma pandemia e uma grande ameaça à saúde mundial (*World Health Organization* (2020)) [1]. Dada a severidade, a rápida disseminação e o surgimento de novas variantes, bem como a enorme importância de meios para diagnóstico, é natural que haja interesse em estratégias, em diferentes níveis, para identificação da contaminação pelo vírus. Diante disso, a detecção por meio da análise de imagens de raios-X ou de tomografia do tórax tem sido utilizada por radiologistas para a análise de indícios da infecção viral pelo SARS-CoV-2 (COVID-19) [2].

As redes neurais profundas (*deep neural networks*, DNNs), que são estruturas de processamento da informação inspiradas pelo funcionamento do sistema nervoso humano [3], apresentam-se como uma estratégia interessante para a rápida detecção de infecção viral em imagens. O desenvolvimento de redes neurais profundas já foi testado por pesquisadores na identificação da pneumonia, os quais obtiveram sucesso na identificação da pneumonia a partir de imagens de raios-X, apresentando resultados até melhores que os de radiologistas [2].

Portanto, o presente estudo pretende focar sua pesquisa em redes neurais profundas convolucionais e sequenciais para a detecção da enfermidade, uma vez que o uso de redes neurais já foi previamente estudado, e foram obtidos resultados promissores na identificação de doenças por meio de imagens de Raio-X com uso de *Dense Convolutional Networks* (*DenseNets*) [4]. Assim, pode-se avaliar a extensão do treinamento a outras redes para obtenção de resultados para o caso específico do coronavírus [4].

METODOLOGIA:

Primeiramente, houve uma revisão dos fundamentos de aprendizado de máquina, incluindo a temática de redes neurais, a fim de rever os conceitos mais importantes sobre o tema. O trabalho teve início com um estudo da literatura sobre as arquiteturas de redes neurais, a fim de se definir a que melhor se aplica a nosso problema. Segundo os pesquisadores, o uso de DNNs para ajudar na detecção da

doença através da análise de imagens de raio-X está obtendo resultados positivos, chegando a apresentar 96,6% de acurácia em teste e 87% de precisão com as imagens de COVID-19 [2].

Diante disso, em nosso trabalho optamos por iniciar estudando uma arquitetura chamada VGG19, a qual é uma estrutura profunda com 19 camadas de peso formatadas na configuração E, ou seja, são 16 camadas de convolução, 3 camadas *fully connected*, 5 camadas *MaxPool* e 1 camada *SoftMax* [5]. A escolha dessa rede se deu por ela apresentar um alto desempenho em processamento de imagens, conforme demonstrado previamente [6]; modularidade, uma vez que é possível escolher a quantidade de parâmetros a serem treinados e, assim, não precisar treinar todos, e por se tratar de uma rede já conhecida pelo grupo.

Foi estudado também o conjunto de datasets, o qual contém dados de imagem de Raio-X de tórax com classificação em três categorias – ‘Normal’, ‘Pneumonia viral’ e ‘COVID-19’. Esses dados foram extraídos do ‘COVID-19 Radiography Database’, Dataset com 3616 imagens de positivo COVID-19, 10200 imagens normais e 1345 de pneumonia viral, disponível através do Kaggle¹.

Após esses estudos, foi feito um teste preliminar com a rede citada anteriormente, para poder medir seu desempenho (utilizando um notebook com processador I5-5200u e 8GB RAM), e foi observado que o tempo de treinamento foi relativamente grande, chegando a durar dias, visto que há uma grande quantidade de imagens no dataset e uma limitação de desempenho do dispositivo utilizado, o qual não possui altas configurações de desempenho. Assim, esse modelo foi descartado e passamos a estudar outros modelos sequenciais e convolucionais.

Layer (type)	Output Shape	Param #
rescaling (Rescaling)	(None, 150, 150, 3)	0
conv2d (Conv2D)	(None, 148, 148, 32)	896
max_pooling2d (MaxPooling2D)	(None, 74, 74, 32)	0
conv2d_1 (Conv2D)	(None, 72, 72, 32)	9248
max_pooling2d_1 (MaxPooling2D)	(None, 36, 36, 32)	0
conv2d_2 (Conv2D)	(None, 34, 34, 32)	9248
max_pooling2d_2 (MaxPooling2D)	(None, 17, 17, 32)	0
flatten (Flatten)	(None, 9248)	0
dense (Dense)	(None, 98)	906402
dense_1 (Dense)	(None, 3)	297
Total params: 926,091		
Trainable params: 926,091		
Non-trainable params: 0		

Figura 1 - Modelo sequencial com 3 camadas de convolução, 3 camadas *MaxPool* e 1 camada *Dense* com função de ativação *relu*.

O primeiro modelo a ser estudado foi um sequencial com 3 camadas de convolução, 3 camadas *MaxPool* e 1 camada *Dense* com função de ativação *relu*, como mostra a Fig. 1 ao lado. Para treinar esta rede foi utilizado 20% do *dataset* para teste e o restante para treino, além de usar um total de dez épocas, porém, a baixa acurácia apresentada por este modelo nos levou a melhorá-lo continuamente até que os resultados fossem satisfatórios. Portanto, após diversas mudanças chegamos em um modelo final, o qual apresenta 5 camadas de convolução, 5 camadas *MaxPool* e 2 camadas *Dense* uma com função de ativação *relu* e a outra *SoftMax*, como mostra a Fig. 2. Já para esta

rede foi utilizado 30% do *dataset* para teste e o restante para treino, além de usar um total de 30 épocas.

¹ Disponível em: <https://www.kaggle.com/tawsifurrahman/covid19-radiography-database>

Layer (type)	Output Shape	Param #
rescaling (Rescaling)	(None, 150, 150, 3)	0
conv2d (Conv2D)	(None, 148, 148, 32)	896
max_pooling2d (MaxPooling2D)	(None, 74, 74, 32)	0
conv2d_1 (Conv2D)	(None, 72, 72, 32)	9248
max_pooling2d_1 (MaxPooling2D)	(None, 36, 36, 32)	0
conv2d_2 (Conv2D)	(None, 34, 34, 32)	9248
max_pooling2d_2 (MaxPooling2D)	(None, 17, 17, 32)	0
conv2d_3 (Conv2D)	(None, 15, 15, 32)	9248
max_pooling2d_3 (MaxPooling2D)	(None, 7, 7, 32)	0
conv2d_4 (Conv2D)	(None, 5, 5, 32)	9248
max_pooling2d_4 (MaxPooling2D)	(None, 2, 2, 32)	0
flatten (Flatten)	(None, 128)	0
dense (Dense)	(None, 128)	16512
dropout (Dropout)	(None, 128)	0
dense_1 (Dense)	(None, 3)	387

Total params: 54,787
 Trainable params: 54,787
 Non-trainable params: 0

Figura 2 - Modelo sequencial com 5 camadas de convolução, 5 camadas MaxPool e 2 camadas Dense uma com função de ativação relu e outra SoftMax

RESULTADOS E DISCUSSÃO:

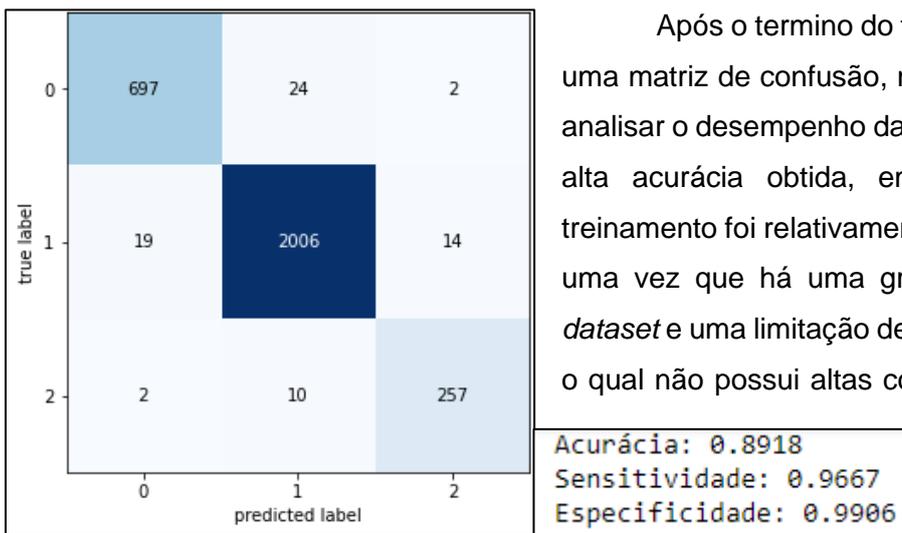


Figura 3 - Matriz de confusão e estatísticas obtidas através do treinamento da rede VGG19

Após o término do treino da rede VGG19, foi montada uma matriz de confusão, mostrada na Fig. 3, a fim de poder analisar o desempenho da rede. Foi concluído que apesar da alta acurácia obtida, em torno de 89%, o tempo de treinamento foi relativamente grande, chegando a durar dias, uma vez que há uma grande quantidade de imagens no *dataset* e uma limitação de desempenho dispositivo utilizado, o qual não possui altas configurações de desempenho. Por este motivo, descartamos esta rede e iniciamos os estudos de outros modelos.

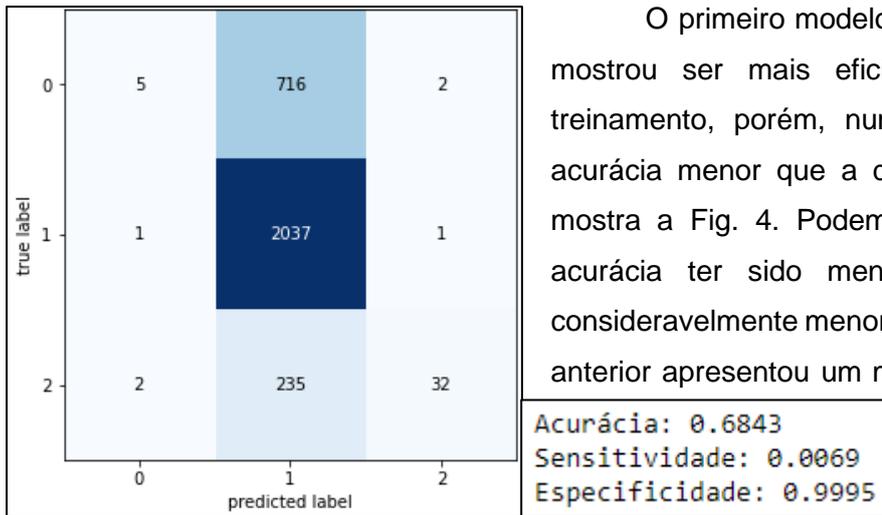


Figura 4 - Matriz de confusão e estatísticas obtidas através do treinamento do primeiro modelo sequencial convolucional

O primeiro modelo, que foi descrito na seção anterior, mostrou ser mais eficiente em questão de tempo de treinamento, porém, num primeiro momento, obteve uma acurácia menor que a do modelo anterior (VGG19), como mostra a Fig. 4. Podemos notar também que, além de a acurácia ter sido menor, o valor da sensibilidade foi consideravelmente menor nesse novo modelo. Logo, o modelo anterior apresentou um melhor desempenho nas estatísticas quando comparado com o novo, no entanto, o tempo para treinamento desse modelo foi menor, levando aproximadamente 1 hora.

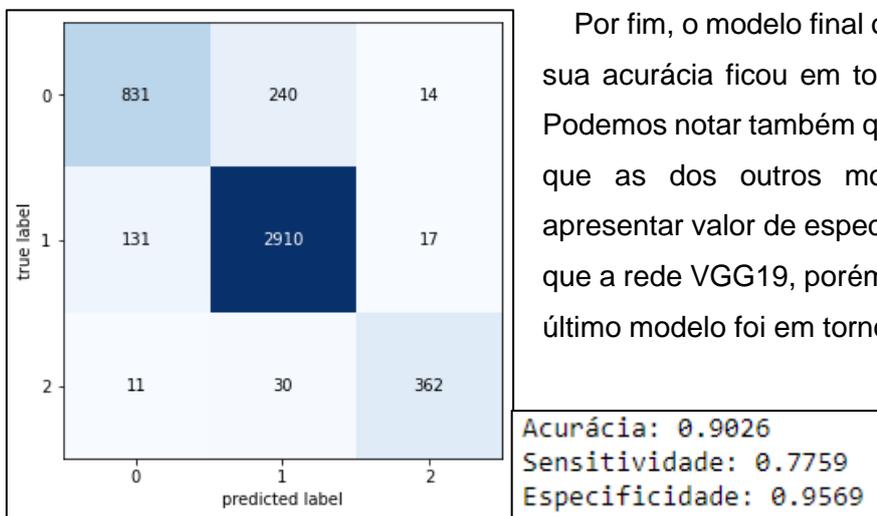


Figura 5 - Matriz de confusão e estatísticas obtidas através do treinamento do segundo modelo sequencial convolucional

Por fim, o modelo final obteve bons resultados, uma vez que sua acurácia ficou em torno de 90%, como mostra a Fig. 5. Podemos notar também que, além de a acurácia ter sido maior que as dos outros modelos anteriores, este chegou a apresentar valor de especificidade e sensibilidade menores do que a rede VGG19, porém o tempo que levou para treinar este último modelo foi em torno de duas horas e meia.

A fim de comparar os dois modelos sequenciais, foram traçados os gráficos das Figs. 6 e 7, os quais nos mostram também quais foram os efeitos das mudanças aplicadas no último modelo, que além de deixa-lo com uma acurácia maior, melhoraram também a estabilidade da rede, uma vez que ela apresentou resultados mais contínuos e com poucas variações e consequentemente deixando-a mais confiáveis para possíveis aplicações na vida real.

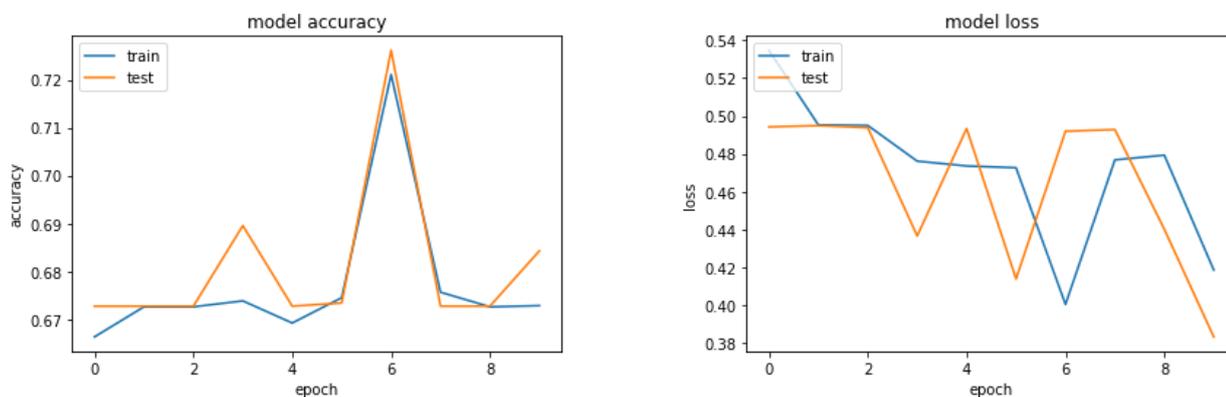


Figura 6 - Resultados obtidos do treinamento do primeiro modelo de rede sequencial

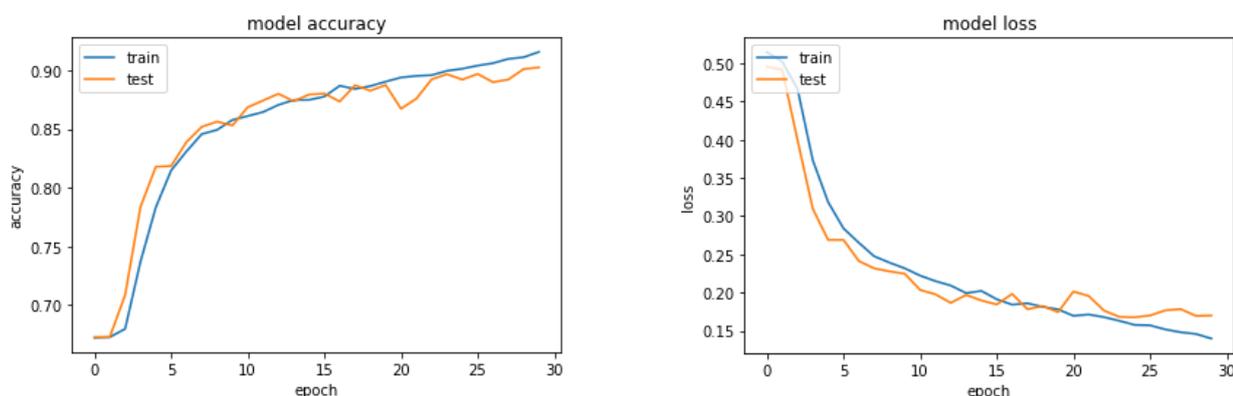


Figura 7 - Resultados obtidos do treinamento do modelo final de rede sequencial

CONCLUSÕES:

Neste trabalho, foi apresentado o estudo teórico realizado, e, apesar de a rede VGG19 não ter sido a escolha final, ela foi de grande aprendizado, pois através dela foram estudadas várias lógicas de programação, como os conceitos e termos mais utilizados no ramo de *Deep Learning*.

A segunda parte abrangeu a continuação do estudo das novas redes, seguida pela implementação e análise da qual obtive o melhor desempenho requerido pelo projeto, chegando a obter uma acurácia em torno de 90% e uma boa estabilidade durante seu treinamento.

BIBLIOGRAFIA

- [1] World Health Organization, <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>, Acesso em: 5 mai. 2021, 2021.
- [2] Bassi P., Attux R., "A Deep Convolutional Neural Network for COVID-19 Detection Using Chest X-Rays", *Research on Biomedical Engineering*, Special Issue: Emerging Technologies for Fighting COVID-19, 2021.
- [3] S. Haykin, *Neural Networks and Learning Machines*, Pearson, 2008.
- [4] Apostolopoulos, I., Mpesiana, T., "Covid-19: Automatic Detection from X-Ray Images Utilizing Transfer Learning with Convolutional Neural Networks", *Physical and Engineering Sciences in Medicine*, Vol. 43, pp. 635 – 640, 2020.