

# Aplicação de Deep Learning no pré-diagnóstico da COVID-19 através de imagens de raio-x

Luan Silva

Graduando em Sistemas de Informação  
Instituto de Geociências e Engenharias  
Pesquisador do Laboratório de Computação Científica

Victor Ferreira

Graduando em Sistemas de Informação  
Instituto de Geociências e Engenharias  
Pesquisador do Laboratório de Computação Científica

Leandro Araújo

Graduando em Sistemas de Informação  
Instituto de Geociências e Engenharias  
Pesquisador do Laboratório de Computação Científica

Adam Santos

Doutor em Engenharia Elétrica  
Instituto de Geociências e Engenharias  
Coordenador do Laboratório de Computação Científica

13 de maio de 2020

## Resumo

Atualmente o mundo está lidando com uma pandemia relacionada com um novo coronavírus, *severe acute respiratory syndrome coronavirus 2* ou SARS-CoV-2, que passou a ser denominada como COVID-19. Uma das alternativas para auxiliar na identificação do COVID-19 é a utilização da radiografia do tórax, a qual mostra características semelhantes a outras pneumonias causadas por outros coronavírus. Entretanto, uma interpretação radiológica rápida de imagens nem sempre está disponível. A aplicação de técnicas de *deep learning* (aprendizagem profunda) para classificação de imagens de raio-x teve um crescimento considerável nos últimos anos. Neste trabalho foram utilizadas duas arquiteturas de redes neurais convolucionais, InceptionResNetV2 e ResNetX50, com o objetivo de resolver o problema de classificação de imagens de raio-x de pessoas com pneumonia, a fim de ajudar no pré-diagnóstico, principalmente, da COVID-19, podendo se tornar um possível método de triagem de pacientes. O modelo gerado a partir da arquitetura ResNetX50 obteve os melhores resultados para todas as métricas definidas.

**Palavras-chave:** Covid-19. Deep learning. Reconhecimento de padrões. Transfer learning.

## 1 Introdução

Atualmente o mundo está lidando com uma pandemia de um novo coronavírus, *severe acute respiratory syndrome coronavirus 2* ou SARS-CoV-2, que passou a ser denominada como COVID-19 (SHOJI et al., 2020), tendo seu primeiro caso notificado em Wuhan, China, no final do ano de 2019. Devido sua rápida disseminação, tornou-se um grave problema de saúde pública no mundo.

A fim de diagnosticar a COVID-19, o teste que vem sendo utilizado é o RT-PCR (NISHIOKA, 2020), mas devido sua falta em algumas localidades e a demora para obter os resultados, se torna necessário a identificação e desenvolvimento de ferramentas que possam auxiliar os profissionais de saúde.

Uma das alternativas para auxiliar na identificação da COVID-19 é a utilização da radiografia do tórax, que mostra características semelhantes a outras pneumonias causadas por outros coronavírus (HOSSEINY et al., 2020). Entretanto, uma interpretação radiológica rápida de imagens nem sempre está disponível, particularmente nos locais com poucos recursos em que a pneumonia (causada por vírus e bactérias) tem a maior incidência e maiores taxas de mortalidade.

O processo de interpretação de agentes (tumores cerebrais ou anomalias nos pulmões, por exemplo) é uma atividade complexa e, por isso, faz-se necessário o uso de técnicas de processamento de imagem, frequentemente combinadas à técnicas de aprendizado de máquina, para identificá-los (NETO; HIDALGA, 2017).

A aplicação de técnicas de *deep learning* (aprendizagem profunda) para classificação de imagens de raio-x teve um crescimento considerável nos últimos anos. Várias pesquisas abordam esse tema, como realizar a classificação de imagens

para ajudar no diagnóstico precoce da tuberculose, e classificar lesões através de radiografia torácica (KARNKAWINPONG; LIMPIYAKORN, 2018), além de serem usadas para a classificação de imagens de raio-x de pessoas com pneumonia e outras lesões (KERMANY et al., 2018a; WANG et al., 2017).

Neste trabalho foram utilizadas duas arquiteturas de Redes Neurais Convolucionais (RNC), InceptionResNetV2 (SZEGEDY; IOFFE; VANHOUCKE, 2016) e ResNetX50 (XIE et al., 2016), com o objetivo de resolver o problema de classificação de imagens de raio-x de pessoas com pneumonia, a fim de ajudar no diagnóstico, principalmente, da COVID-19, podendo se tornar um método de triagem de pacientes. O restante deste estudo está organizado da seguinte forma. Na seção 2 serão abordadas mais informações sobre *deep learning* e RNC. Na seção 3, a metodologia utilizada, ferramentas, técnicas e base de dados serão discutidos. Na seção 4, os resultados obtidos com os modelos testados serão apresentados. Por fim, na seção 5, as considerações finais deste trabalho serão sumarizadas, destacando possíveis trabalhos futuros.

## 2 Deep Learning

O paradigma de aprendizagem profunda (do inglês *deep learning*) é uma sub-área da aprendizagem de máquina que aborda o aprendizado automatizado sendo realizado através da utilização de sucessivas camadas profundas dentro de uma arquitetura de rede neural (FRANCOIS, 2017). Com isso, cada camada passa o resultado do seu aprendizado para a camada seguinte. Quanto maior a quantidade de camadas, mais profunda a rede neural se torna.

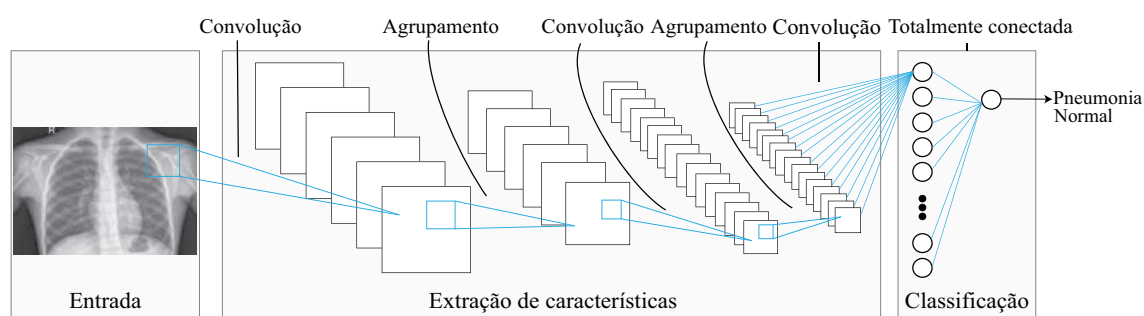
A principal característica da aprendizagem profunda é a utilização de camadas para realizar a extração de características e classificação dos dados (RIGHETTO, 2016). Com isso, algoritmos de *deep learning* se tornam opções viáveis para tarefas de classificação de imagens, por exemplo.

### 2.1 Redes Neurais Convolucionais

RNC são arquiteturas biologicamente inspiradas capazes de serem treinadas e aprenderem/generalizarem representações invariantes a escala, translação, rotação e transformações afins (LIU; ROSENBERG; ROWLEY, 2007; JURASZEK et al., 2014). As RNCs compõem um dos tipos de algoritmos da área conhecida como aprendizagem profunda e são projetadas para uso com dados de múltiplas dimensões, tornando-as boas candidatas para a solução de problemas envolvendo reconhecimento de imagens (AREL et al., 2010).

De maneira semelhante aos processos tradicionais de Visão Computacional, uma RNC é capaz de aplicar filtros em dados não-estruturados, mantendo a relação de vizinhança entre os *pixels* da imagem ao longo do processamento da rede. A Figura 1 demonstra a estrutura de uma RNC.

Figura 1 – Estrutura de uma RNC.



Fonte: Autoria Própria (2019).

As RNCs têm sido usadas pela comunidade científica em imagens médicas por causa de seu excelente desempenho demonstrado em Visão Computacional. Usualmente, como mostra a Figura 1, três tipos de camadas são usadas para construir uma RNC (MOUSSER; OUADFEL, 2019): camada convolutiva, camada de agrupamento (*pooling*) e camada totalmente conectada.

### 3 Metodologia

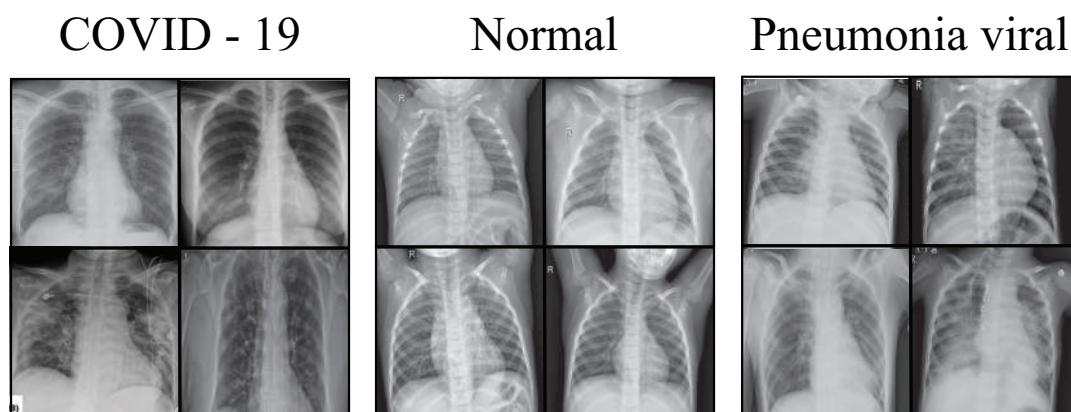
A linguagem de programação Python foi utilizada para a implementação das RNCs, com o auxílio da biblioteca Keras, amplamente utilizada na construção de redes neurais profundas, executando sobre a biblioteca de código aberto TensorFlow (FOUNDATION, 2020; KERAS, 2020; TENSORFLOW, 2020). As arquiteturas empregadas neste estudo são a InceptionResNetV2, que possui 572 camadas; e a ResNeXt50, que possui 50 camadas.

O treinamento das arquiteturas propostas foi executado em uma máquina virtual (VM) fornecida pelo Google Colaboratory (Colab) (COLAB, 2020). O Colab tem o objetivo de disseminar e incentivar a investigação na área de aprendizagem de máquina. O Colab disponibiliza VMs pré-configuradas com as bibliotecas essenciais de aprendizagem de máquina e inteligência artificial, como TensorFlow, Matplotlib e Keras.

#### 3.1 Base de dados

Para o treinamento dos modelos foi utilizada uma base de dados com imagens de radiografia do tórax, amplamente utilizadas no diagnóstico de pneumonia. Seus desenvolvedores (CHOWDHURY et al., 2020) utilizaram quatro bases de dados: SIRM COVID-19 (INTERVENTISTICA, 2020), base de dados pública disponível no GitHub (COHEN; MORRISON; DAO, 2020), coleta de imagens em diferentes artigos e a base de dados *chest x-ray* disponível no Kaggle (KERMANY et al., 2018b).

Figura 2 – Exemplos das radiografias utilizadas.



Fonte: Autoria Própria (2020).

A Figura 2 mostra alguns exemplos de imagens de raio-x presentes na base de dados. A radiografia do tórax de um paciente infectado pela COVID-19 (painel da esquerda) mostra manchas opacas nos pulmões, conhecidas na radiologia como padrão vidro fosco. A radiografia do tórax de um paciente em estado normal (painel do meio) mostra pulmões limpos sem nenhuma área de opacificação anormal na imagem. A radiografia do tórax de um paciente infectado por outros vírus (painel da

direita), que causam a pneumonia, se manifesta com um padrão “intersticial” mais difuso em ambos os pulmões.

Essa base de dados é composta por 2905 imagens divididas em três classes, tal que as classes normal, pneumonia viral e COVID-19 possuem 1341, 1345 e 219 imagens, respectivamente. A fim de diminuir o custo computacional, todas as imagens foram redimensionadas para  $75 \times 75$  pixels. A base de dados foi dividida em 70% para treino e 30% para teste dos modelos.

### 3.2 Técnicas e Hiperparâmetros utilizados

Para o treinamento dos modelos foram adotados alguns hiperparâmetros, que se mantiveram iguais em todas as RNCs. Por se tratar de um problema que envolve a classificação de múltiplas classes, foi selecionada a *Softmax* como função da camada de saída. Para a função de custo foi utilizada a *Categorical Crossentropy*, uma das principais funções utilizadas na literatura para obter o valor de custo em problemas com múltiplas classes.

Por meio de observações experiências anteriores, foi possível definir alguns dos hiperparâmetros, como a utilização do otimizador *Adaptive Moment Estimation* (ADAM); a quantidade de épocas igual a 200; e o *dropout* em 0,5 a fim de evitar *overfitting*. Para o tamanho do lote (*batch*) foi definido o valor de 64 e na taxa de aprendizagem foi realizada a seguinte estratégia: valor inicial de  $1 \times 10^{-3}$ , sendo multiplicada por taxas menores no decorrer das épocas, isto é; a partir da época 81,  $1 \times 10^{-1}$ ; época 121,  $1 \times 10^{-2}$ ; época 161,  $1 \times 10^{-3}$ ; e época 181:  $0,5 \times 10^{-3}$ . Essa estratégia permitiu que as redes em suas épocas finais obtivessem um baixo desvio padrão no que diz respeito a acurácia.

Em todas as arquiteturas serão utilizados pesos estimados por modelos já treinados para identificar se uma radiografia pertence a uma pessoa com ou sem pneumonia. A utilização de pesos pré-treinados nos modelos selecionados trata-se de uma técnica de aprendizado de máquina conhecida como transferência de aprendizagem. Essa técnica possui o objetivo de melhorar a aprendizagem da tarefa atual, aproveitando o conhecimento adquirido em outra tarefa (OLIVAS et al., 2009).

### 3.3 Métricas para avaliação dos resultados

Para avaliação dos resultados, as métricas de Acurácia, Precisão, Revocação, pontuação-F1 e matriz de confusão foram empregadas. A matriz de confusão é geralmente utilizada no aprendizado de máquina, possuindo informações sobre as classificações reais e previstas realizadas por um classificador (DENG et al., 2016). Em uma matriz de confusão as linhas são valores reais em cada classe, enquanto as colunas são as predições realizadas pelo modelo.

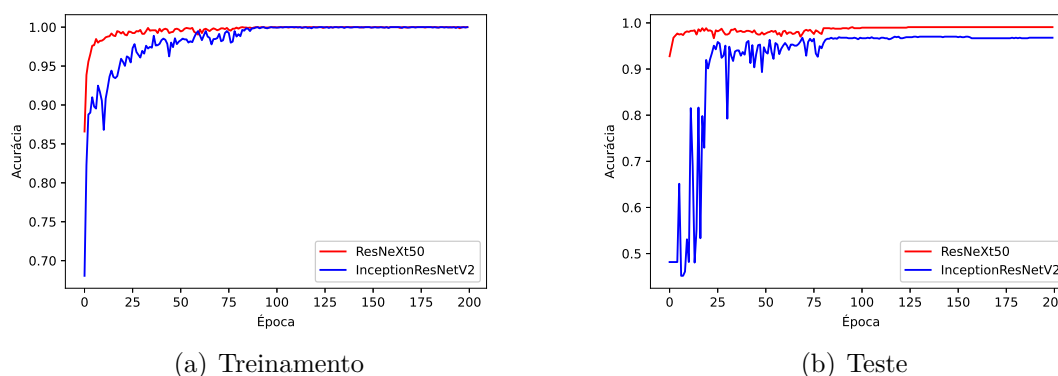
Os valores obtidos a partir da matriz de confusão também são utilizados para gerar algumas métricas de suma importância para a avaliação dos modelos, tais como: Acurácia, Precisão, Revocação e pontuação-F1 (SOARES et al., 2019). Essas métricas são comumente utilizadas na avaliação de modelos de aprendizagem

de máquina, portanto, elas serão utilizadas neste estudo a fim de avaliar a eficácia dos modelos gerados.

## 4 Resultados e discussões

A partir dos resultados obtidos, é possível visualizar na Figura 3 o desempenho dos modelos gerados, nas etapas de treinamento e teste. O parâmetro de análise utilizado foi a acurácia por época, ou seja, taxa de acerto em cada época. A ResNetX50 obteve uma convergência mais rápida do que a InceptionResNetV2, tanto na etapa de treinamento quanto no teste. Vale a pena destacar que houve uma discrepância entre as curvas na etapa de teste, o que evidencia que a ResNetX50 obteve um desempenho melhor.

Figura 3 – Acurácia no treino e teste dos modelos aplicados.



Fonte: Autoria Própria (2020).

Na Tabela 1 é possível visualizar os resultados de precisão, revocação e pontuação-f1 de cada classe para cada arquitetura. Para todas as classes e métricas, a ResNetX50 obteve o melhor resultado. Com isso, o modelo da ResNetX50 é capaz de não confundir as classes (precisão) e consegue encontrar o maior número possível de imagens de cada classe (revocação). A pontuação-f1 é a média ponderada das duas métricas já citadas.

Tabela 1 – Precisão, revocação e pontuação-F1 dos modelos aplicados.

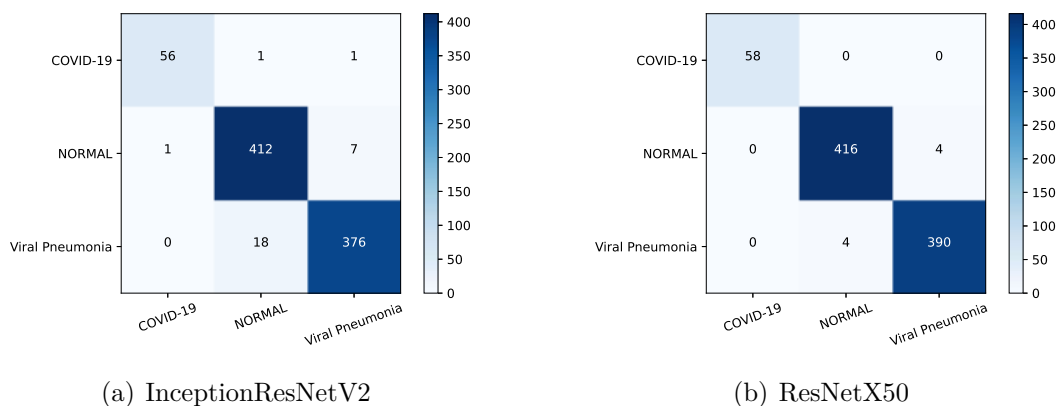
Arquitetura		Precisão	Revocação	Pontuação-F1
ResNetX50	COVID-19	<b>100%</b>	<b>100%</b>	<b>100%</b>
	NORMAL	<b>99%</b>	<b>99%</b>	<b>99%</b>
	Pneumonia Viral	<b>99%</b>	<b>99%</b>	<b>99%</b>
InceptionResNetV2	COVID-19	98%	97%	97%
	NORMAL	96%	98%	97%
	Pneumonia Viral	98%	95%	97%

A Figura 4 evidencia as matrizes de confusão das duas arquiteturas. A diagonal principal de cada matriz demonstra os acertos na etapa de teste. O modelo



da InceptionResNetV2 acertou 56 imagens de pacientes diagnosticados com COVID-19; 412 imagens de pacientes em condições normais; e 376 imagens de pacientes com pneumonia causados por outros vírus. O modelo da ResNetX50 acertou 58 imagens de pacientes diagnosticados com COVID-19; 416 imagens de pacientes em condições normais; e 390 imagens de pacientes com pneumonia causados por outros vírus.

Figura 4 – Matrizes de confusão dos modelos aplicados.



Fonte: Autoria Própria (2020).

## 5 Considerações finais e trabalhos futuros

Com base nos resultados alcançados, podemos observar que o modelo gerado através da arquitetura da ResNetX50 obteve o melhor desempenho no diagnóstico da COVID-19, foco deste estudo. Em todas as métricas analisadas a ResNetX50 obteve um resultado superior à InceptionResNetV2. Por conseguir uma taxa de acerto (acurácia) maior do que a InceptionResNetV2, a ResNetX50 torna-se uma escolha factível de arquitetura para ser mais explorada em trabalhos futuros. É importante ressaltar que este estudo tem como premissa servir futuramente como uma forma alternativa de triagem de pacientes.

A partir deste estudo, novas possibilidades podem ser exploradas. Um sistema WEB poderá ser desenvolvido com o intuito de servir como ambiente de teste do modelo que foi treinado, ou seja, o usuário poderá fazer o *upload* de uma imagem de raio-x do tórax de um paciente com suspeita de COVID-19 e o modelo irá gerar um possível pré-diagnóstico da imagem. Como forma de otimizar e melhorar a proposta anterior, será necessário utilizar mais bases de dados que contenham imagens de pacientes diagnosticados com COVID-19 ou outras doenças que possam ser diagnosticadas através de imagens de raio-x, para potencializar e generalizar o modelo.



## Referências

- AREL, I. et al. Deep machine learning-a new frontier in artificial intelligence research. *IEEE computational intelligence magazine*, The University of Tennessee, v. 5, n. 4, p. 13–18, 2010. Citado na página 4.
- CHOWDHURY, M. E. H. et al. *Can AI help in screening Viral and COVID-19 pneumonia?* 2020. Citado na página 5.
- COHEN, J. P.; MORRISON, P.; DAO, L. Covid-19 image data collection. *arXiv 2003.11597*, 2020. Disponível em: <<https://github.com/ieee8023/covid-chestxray-dataset>>. Citado na página 5.
- COLAB, G. *Google Colab*. 2020. Disponível em: <<https://colab.research.google.com/notebooks/intro.ipynb>>. Citado na página 5.
- DENG, X. et al. An improved method to construct basic probability assignment based on the confusion matrix for classification problem. *Information Sciences*, v. 340-341, p. 250 – 261, 2016. ISSN 0020-0255. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S002002551600044X>>. Citado na página 6.
- FOUNDATION, P. S. *Python*. 2020. Disponível em: <<https://www.python.org/>>. Citado na página 5.
- FRANCOIS, C. *Deep learning with Python*. [S.l.]: Manning Publications Company, 2017. Citado na página 4.
- HOSSEINY, M. et al. Radiology perspective of coronavirus disease 2019 (covid-19): Lessons from severe acute respiratory syndrome and middle east respiratory syndrome. *American Journal of Roentgenology*, American Roentgen Ray Society, v. 214, n. 5, p. 1078–1082, maio 2020. ISSN 0361-803X. Disponível em: <<https://doi.org/10.2214/AJR.20.22969>>. Citado na página 2.
- INTERVENTISTICA, S. I. di Radiologia Medica e. *COVID-19 DATABASE*. 2020. Acessado: 05-05-2020. Disponível em: <<http://www.sirm.org/category/senza-categoria/covid-19/>>. Citado na página 5.
- JURASZEK, G. D. et al. Reconhecimento de produtos por imagem utilizando palavras visuais e redes neurais convolucionais. Universidade do Estado de Santa Catarina, 2014. Citado na página 4.
- KARNKAWINPONG, T.; LIMPIYAKORN, Y. Chest x-ray analysis of tuberculosis by convolutional neural networks with affine transforms. In: ACM. *Proceedings of the 2018 2nd International Conference on Computer Science and Artificial Intelligence*. [S.l.], 2018. p. 90–93. Citado na página 3.
- KERAS. *Keras*. 2020. Disponível em: <<https://keras.io/>>. Citado na página 5.

KERMANY, D. S. et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*, Elsevier, v. 172, n. 5, p. 1122–1131, 2018. Citado na página 3.

KERMANY, D. S. et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*, Elsevier, v. 172, n. 5, p. 1122–1131, 2018. Citado na página 5.

LIU, T.; ROSENBERG, C.; ROWLEY, H. A. Clustering billions of images with large scale nearest neighbor search. In: IEEE. *2007 IEEE Workshop on Applications of Computer Vision (WACV'07)*. [S.l.], 2007. p. 28–28. Citado na página 4.

MOUSSER, W.; OUADFEL, S. Deep feature extraction for pap-smear image classification: A comparative study. In: ACM. *Proceedings of the 2019 5th International Conference on Computer and Technology Applications*. [S.l.], 2019. p. 6–10. Citado na página 4.

NETO, C.; HIDALGA, S. A. de L. Reconhecimento de tumores cerebrais utilizando redes neurais convolucionais. Universidade Federal do Pampa, 2017. Citado na página 2.

NISHIOKA, S. de A. *A utilidade da tomografia computadorizada de tórax no diagnóstico de COVID-19*. 2020. Acessado: 09-05-2020. Disponível em: <<https://www.unasus.gov.br/especial/covid19/markdown/149>>. Citado na página 2.

OLIVAS, E. S. et al. *Handbook Of Research On Machine Learning Applications and Trends: Algorithms, Methods and Techniques - 2 Volumes*. Hershey, PA: Information Science Reference - Imprint of: IGI Publishing, 2009. ISBN 1605667668, 9781605667669. Citado na página 6.

RIGHETTO, G. *O uso da rede neural convolucional como extrator de características aplicado ao problema de identificação de escritores*. Dissertação (B.S. thesis) — Universidade Tecnológica Federal do Paraná, 2016. Citado na página 4.

SHOJI, H. et al. Relatório estruturado de tomografia computadorizada de tórax para a pandemia do covid-19. *einstein (São Paulo)*, v. 18, p. eED5720, abril 2020. ISSN 1679-4508. Disponível em: <<https://journal.einstein.br/article/structured-thoracic-computed-tomography-report-for-covid-19-pandemic/>>. Citado na página 2.

SOARES, G. A. d. S. et al. Reconhecimento de estados dos olhos utilizando máquinas de aprendizado profundo a partir de ondas cerebrais. *Revista de Sistemas e Computação-RSC*, v. 9, n. 1, 2019. Citado na página 6.

SZEGEDY, C.; IOFFE, S.; VANHOUCKE, V. Inception-v4, inception-resnet and the impact of residual connections on learning. *CoRR*, abs/1602.07261, 2016. Disponível em: <<http://arxiv.org/abs/1602.07261>>. Citado na página 3.

TENSORFLOW. *TensorFlow*. 2020. Disponível em: <<https://www.tensorflow.org/>>. Citado na página 5.

WANG, X. 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. Citado na página 3.

XIE, S. et al. Aggregated residual transformations for deep neural networks. *CoRR*, abs/1611.05431, 2016. Disponível em: <<http://arxiv.org/abs/1611.05431>>. Citado na página 3.