# Automatic Detection of COVID-19 in X-Ray Images Using Fully-Connected Neural Networks

Élisson Carlos de Carvalho\*, Raian Campos Malta \*, Alessandra Martins Coelho \*,

Matheus de Freitas Oliveira Baffa<sup>†</sup>

\*Federal Institute of Education, Science and Technology of Southeast of Minas Gerais (IF SudesteMG)

Rio Pomba, MG, Brazil

<sup>†</sup>University of São Paulo (USP)

Ribeirão Preto, Brazil

Email: elissoncarvalho314@gmail.com, raianmalta71@gmail.com, alessandra.coelho@ifsudestemg.edu.br,

mfreitas826@gmail.com

Abstract—The coronavirus pandemic remains a problem of worldwide interest. The diagnosis of COVID-19 is difficult due to its high rate of occurrence and the limited number of test kits. Medical imaging is already widespread and has been used to quickly provide lung visualization. It's needed some expertise from the radiologist to detect elements in the image that allow differentiating the sick and healthy patterns. Therefore, our goal with this paper is to provide a computer-aided diagnosis tool to help radiologists to accurately diagnose the COVID-19 using X-Ray images. For that, a model based on Fully-Connected Neural Networks was proposed for the detection of patients infected with coronavirus, through the analysis of texture characteristics, such as Haralick and Threshold Adjacency Statistics (TAS) descriptors, extracted from chest X-Ray images. Using 10-Fold Cross-Validation, the proposed method achieved an accuracy of 98.39%, showing itself as an option to aid the disease diagnosis.

Index Terms—computer vision, COVID-19, deep learning, xray classification

## I. INTRODUÇÃO

Nos últimos meses o mundo se viu enfrentando uma de suas maiores crises. Em 31 de dezembro de 2019, 27 casos inexplicáveis de pneumonia foram identificados em *Wuhan*, província de *Hubei*, China e associados aos chamados "mercados úmidos", que vendem frutos do mar e carne fresca de uma variedade de animais, incluindo morcegos e pangolins. Desde então, países da América, África, Ásia, Europa e Oceania vêm sendo afetados pelo vírus [6]. A pneumonia encontrada é causada por um vírus identificado como Síndrome Respiratória Aguda Grave CoronaVírus-2 (SARSCoV-2) [18], com a doença associada denominada de coronavírus-2019 ou simplesmente COVID-19. O vírus é transmitido por gotículas de saliva, espirros, acessos de tosse, contato próximo e superfícies contaminadas, podendo sobreviver por até cinco dias em algumas superfícies [3].

No dia 11 de março de 2020, a Organização Mundial da Saúde (OMS) [4] declarou a pandemia do novo coronavírus (COVID-19). Esse vírus, que tem como principal grupo de risco os idosos [7], infectou mais de 10 milhões de pessoas e tirou mais de 500 mil vidas no primeiro semestre de 2020, além de deixar vários países ao redor do mundo em estado de *Lockdown*. Países de primeiro mundo viram seus sistemas

de saúde em total colapso, como o caso da Itália, que, em certo ponto da pandemia, priorizou dentre os pacientes, os que teriam mais chances de sobreviver, pois não havia como atender a todos.

Os relatos iniciais da infecção caracterizaram o quadro como uma pneumonia de origem desconhecida, sendo muitos pacientes tratados para pneumonia, porém sem sucesso com a implementação da terapia antibiótica usual. Os primeiros casos apresentaram opacificação mal definida na radiografia de tórax, bilateral e periférica na maioria das vezes e na Tomografia Computadorizada apresentou-se com um padrão em "vidro fosco" e zonas de mosaico. Hoje, sabemos que cerca de 59% dos pacientes apresentam alterações no exame de imagem, no entanto, uma pessoa que possui seu exame de imagem sem alterações não pode ser diagnosticada como negativo para a COVID-19, pois somente o exame de imagem sem alteração não é fator de exclusão [17].

As imagens médicas têm sido utilizadas como forma inicial de diagnóstico do coronavírus. Isso se deve ao fato do alto custo de outras formas de diagnóstico, e à a alta taxa de ocorrência da doença ao redor do planeta. Atualmente, o protocolo padrão para casos de suspeitas da doença e urgência de diagnóstico é a utilização de imagens de Tomografia Computadorizada (*Computed Tomography* – CT) ou de Radiografia do Tórax (Raio-X) [13]. A principal vantagem na utilização deste exame é o seu fácil acesso. Como são métodos utilizados no diagnóstico de grande quantidade de doenças, muitos hospitais e clínicas já possuem aporte para realizá-lo.

Diferenciar entre pacientes doentes ou saudáveis a partir de exames de imagem é um problema comumente abordado na literatura de Visão Computacional aplicada à Saúde [5]. A utilização de inteligência artificial em Sistemas de Auxílio ao Diagnóstico tem se mostrado eficaz, tornando-se uma aliada no diagnóstico preciso de doenças [14].

Diante do exposto, apresentamos neste artigo um método que utiliza uma rede de Aprendizado profundo, treinada a partir de imagens de Raio-X, para detecção da COVID-19.

## II. TRABALHOS RELACIONADOS

Atualmente o mundo vem passando por uma crise sanitária devido à pandemia causada pela COVID-19. O seu diagnóstico está tipicamente associado aos sintomas de pneumonia, que podem ser revelados por testes de imagem de Raio-X. Visando mais uma ferramenta para auxiliar o diagnóstico médico, vários trabalhos têm sido realizados com o objetivo de desenvolver sistemas de classificação automática, usando principalmente *Deep Learning* [22].

Hemdan et al. [10], por exemplo, desenvolveram o COVIDx-Net, uma estrutura composta por sete arquiteturas diferentes de Redes Neurais Convolutivas (CNN). Cada uma das arquiteturas foram utilizadas para classificar imagens de Raio-X, mostrando um desempenho significativo dos modelos de *Deep Learning*, onde a VGG19 e a DenseNet201 foram as arquiteturas que apresentaram os melhores resultados, com uma acurácia de 90%.

No trabalho desenvolvido por Afshar et al. [2], foi proposta a COVID-CAPS, uma estrutura baseada em *Capsule Network* para identificação do COVID-19 usando imagens de Raio-X, alcançando uma acurácia de 95,7%.

Em Wang et al. [20] foi criado um conjunto de dados composto por 13.975 imagens de Raio-X, de 13.870 casos de pacientes de cinco repositórios de dados distintos. Esses dados foram utilizados para treinar o COVID-Net, um projeto de Rede Neural Convolucional profunda para a detecção de casos COVID-19, alcançando uma acurácia de 93%.

Utilizando a arquitetura ResNet, Abbas et al. [1] realizaram um estudo que se mostrou eficiente, tendo 95,12% de acurácia. Apresentaram soluções robustas para a classificação de casos COVID-19, com a capacidade de lidar com irregularidade de dados e com um número limitado de imagens de treinamento.

Três diferentes modelos baseados em Redes Neurais Convolucionais (ResNet50, InceptionV3 e Inception-ResNetV2) foram propostas para a detecção de pacientes infectados por pneumonia por coronavírus usando Raio-X, em Narin et al. [12]. InceptionV3 e Inception-ResNetV2 alcançaram, respectivamente, 97% e 87% de acurácia. ResNet50 apresentou melhor o desempenho, obtendo 98% de acurácia.

Já em Kasssini et al. [11] foi realizada a comparação das estruturas populares de extração de características, baseada em *Deep Learning*, para a classificação automática do COVID-19. Como componentes do aprendizado, MobileNet, DenseNet, Xception, ResNet, InceptionV3, InceptionResNetV2, VG-GNet, NASNet foram escolhidos entre um conjunto de subarquiteturas de Redes Neurais Convolucionais profundas. Os recursos extraídos foram alimentados em vários classificadores de Aprendizado de Máquina, para classificar os assuntos como um caso de COVID-19 ou um caso de controle.

Essa abordagem evitou dados específicos da tarefa, métodos de pré-processamento, para suportar uma melhor capacidade de generalização de dados não vistos.

O extrator de recursos DenseNet121 com classificador de árvore de ensacamento, alcançou o melhor desempenho, com precisão de classificação de 99%. O segundo melhor resultado foi um método híbrido ResNet50, treinado pela LightGBM, com uma precisão de 98%.

# III. MATERIAIS E MÉTODOS

Nas Figuras 1 e 2, temos, respectivamente, o Raio-X de uma pessoa com a COVID-19 e de uma pessoa saudável. Ao observá-las, podemos perceber as diferenças entre a imagens. As opacidades na Figura 1 (marcas brancas circulares), associadas a um padrão de vidro fosco, são indicativos da infecção causada pela COVID-19 [19].

Com a utilização de algoritmos de aprendizagem de máquina, busca-se encontrar padrões em uma base de dados contendo imagens de pacientes com e sem a COVID-19. Espera-se, a partir da base de dados usada e de melhorias no algoritmo, obter um diagnóstico cada vez mais preciso.

A linguagem de programação Python foi usada para treinar os modelos de aprendizado de transferência profunda propostos. Todos os experimentos foram realizados no ambiente de programação Anaconda (versão 1.7.2) utilizando o sistema operacional Windows 10 Pro de 64 bits.

O Spyder é uma IDE Python dedicada à computação matemática, integrada com o interpretador IPython, e com pacotes Numpy (álgebra linear), Scipy (processamento de imagens) e Matplotlib (plotagem 2D e 3D). Para o préprocessamento da base de dados utilizou-se a biblioteca OpenCV (versão 4.2.0.34). Já o modelo de Inteligência Artificial foi criado a partir das bibliotecas Tensorflow (versão 2.1.0) e Keras (versão 2.3.1).

Os testes foram feitos em um computador com placa gráfica GTX 960 4GB Windforce e processador Intel core I5 4460 3.4 GHz, com 16 GB de memória RAM.

#### A. Base de Dados

Neste trabalho, usamos um banco de dados de imagens de Raio-X do tórax para casos positivos de COVID-19, obtidas em Tawsifur et al. [15] e composto por 2905 imagens de tamanho 1024x1024 pixels, divididas em três classes (219 imagens de pacientes com COVID-19, 1.341 imagens de pacientes saudáveis e 1.345 imagens de pacientes com pneumonia viral). Para este trabalho, foram selecionadas apenas as duas primeiras classes, que estão representadas, respectivamente nas Figuras 1 e 2.



Fig. 1. Representação de imagens de Raio-X de pacientes com COVID-19. Fonte: Tawsifur et al. [15].



Fig. 2. Representação de imagens de Raio-X de pacientes saudáveis. Fonte: Tawsifur et al. [15].

# B. Pré-Processamento da Base de Dados

No pré-processamento da base de dados, as imagens foram carregadas na escala cinza, seguidas da aplicação da técnica de processamento de imagem de equalização de histograma, com o objetivo de melhorar o aspecto da imagem como um todo, reforçando as características texturais importantes para este trabalho.



Fig. 3. Representa a mesma imagem da Figura 1, porém agora com o préprocessamento da base de dados. Fonte autor.

A equalização de histogramas é uma operação que melhora o contraste, uniformizando o histograma da imagem de forma automática, redistribuindo os níveis de cinza existentes e mapeando-os para novos níveis. Embora os picos e vales do histograma sejam mantidos, eles são deslocados após a equalização. Esse procedimento (1) faz com que o número de intensidades na imagem resultante seja igual ou menor que na imagem original.

$$Histograma: (rk), k \in [0, L - 1h(rk), k \in [0, L - 1]$$
(1)

#### C. Extração de Características

Após a etapa de pré-processamento, as imagens foram submetidas à extração de características texturais, que possibilita descrever uma imagem analisando sua textura. Existem várias técnicas para a extração dessas características. Neste trabalho foi utilizada a combinação de dois descritores de textura: o descritor de *Haralick* e o *Threshold Adjacency Statistics* (TAS), gerando um vetor com 67 características. Haralick descreve uma imagem por meio da técnica de Matriz de Coocorrência de Níveis de Cinza (*Grey-Level Co*occurrence Matrix - GLCM) [9]. São definidos um conjunto de 14 medidas de características, sendo algumas relacionadas com as características texturais específicas da imagem, como homogeneidade, contraste e a presença de estrutura organizada dentro da imagem, e outras que caracterizam a complexidade e a natureza das transições de tons de cinza que ocorrem na imagem.

O TAS é obtido por meio da aplicação de um limiar à imagem para criar uma imagem binária. Esse limiar é escolhido a partir de um intervalo selecionado para maximizar a diferença visual das imagens [8]. Então, para cada pixel branco, o número de pixels brancos adjacentes é contado. A primeira estatística é então o número de pixels brancos sem vizinhos brancos; a segunda é o número com um vizinho branco e assim sucessivamente até o máximo de oito. Esse procedimento resulta em um histograma do número de pixels brancos adjacentes. A partir disso, o histograma é normalizado, dividindo cada compartimento pelo número total de pixels brancos. Os valores numéricos de cada um dos nove compartimentos do histograma consiste nas estatísticas. Essas podem ser usadas como características em testes de classificação.

#### D. Metodologia de Classificação

Para classificar as imagens de Raio-X de pacientes infectados ou não com COVID-19, a partir das características descritas anteriormente, foi utilizada uma rede neural totalmente conectada (*Fully-Connected Neural Networks* - FCNN).

A FCNN consiste em uma série de camadas totalmente conectadas, em que cada camada é uma função do  $\mathbb{R}^m$  para  $\mathbb{R}^n$ , onde  $\mathbb{R}^{m \times n}$  é o espaço real de uma matriz  $m \times n$ . Cada dimensão de saída depende de cada dimensão de entrada [16]. Uma camada totalmente conectada é representada conforme apresentada na Figura 4.

A principal vantagem das redes totalmente conectadas é que elas são independentes da estrutura, ou seja, nenhuma suposição especial precisa ser feita sobre a entrada (por exemplo, que a entrada consiste em imagens ou vídeos). [16].

O modelo de FCNN proposto neste trabalho é constituído por seis camadas totalmente conectadas, com dezesseis neurônios cada. Foi definido um *dropout* de 20% após a função de ativação de cada camada. Além disso, a rede é composta por uma camada de saída com um neurônio.

A rede foi treinada por 100 épocas, tendo como parâmetros um *batch size* igual 32 e *optimizer adam*. A função de ativação usada na camada de saída foi o *Sigmoid*. Nas demais camadas, foi utilizada a Unidade Linear Retificada (*Rectified Linear Unit* - ReLU) como função de ativação. Na Figura 5 temos uma ilustração representativa dessa rede.

# IV. EXPERIMENTOS E RESULTADOS

A validação cruzada por *K-Fold* é um procedimento popular para se estimar o desempenho de um algoritmo de classificação ou comparação do desempenho entre dois algoritmos de



Fig. 4. Representa uma rede neural onde cada dimensão de saída depende de cada dimensão de entrada. Fonte: Ramsundar et al. [16].



Fig. 5. Representação visual da arquitetura final da rede neural. Fonte: autor.

classificação em um conjunto de dados [21]. Esse procedimento divide aleatoriamente um conjunto de dados em k partes (*folds*), com aproximadamente o mesmo tamanho. Cada parte é usada para testar o modelo induzido das outras k - 1 *folds*, por um algoritmo de classificação. O desempenho da classificação do algoritmo é avaliado pela média das k precisões resultantes da validação cruzada.

Utilizando a validação cruzada *k-fold*, com k = 10, este trabalho avalia o desempenho do modelo proposto, a partir da seguintes métricas de performance: acurácia, sensibilidade e especificidade. Essas métricas podem ser determinadas a partir da contagem correta de exemplares de cada classe, em que Verdadeiro Positivo (VP) é o número de exemplos da classe corretamente reconhecidos; Verdadeiro Negativo (VN) é o número exemplos corretamente reconhecidos que

não pertencem à classe; Falso positivo (FP) é o número de exemplos que incorretamente foram atribuídos à classe; e Falso Negativo (FN) é a quantidade de exemplos que não foram reconhecidos como pertencentes à classe.

A Acurácia (2) é a métrica mais importante para os resultados de classificadores de aprendizado profundo. Ela retorna o acerto médio por classe de um classificador.

$$Acuracia = \frac{VP + VN}{VP + FP + FN + VN}$$
(2)

A Sensibilidade (3) mede a eficácia do classificador para identificar os verdadeiros positivos, ou seja, ela mede a capacidade do classificador em reconhecer corretamente os volumes.

$$Sensibilidade = \frac{VP}{VP + FN} \tag{3}$$

A Especificidade (4) mede a eficácia do classificador em identificar os verdadeiros negativos.

$$Especificidade = \frac{VN}{VN + FP} \tag{4}$$

Utilizando imagens de Raio-X para treinar uma FCNN e classificar os pacientes em saudáveis ou infectados com COVID-19, o modelo proposto alcançou 98,39% de acurácia, 99,18% de sensibilidade e 93,47% de especificidade. A sensibilidade e especificidade alta mostram que o desbalanceamento na base de dados não impactou na qualidade do método proposto uma vez que não houve uma classe dominante. Na Tabela I são encontrados os valores de todos *k-fold* da validação cruzada. O resultado final é a media desses valores, calculada para cada métrica avaliada.

TABELA I Resultado do desempenho obtido pela FCNN no método 10-fold de validação cruzada. Fonte: autor.

#	Acurácia (%)	Sensibilidade (%)	Especificidade (%)
0	97,60	97,34	100,00
1	97,60	99,12	81,81
2	98,40	100,00	90,00
3	97,60	98,14	94,11
4	100,00	100,00	100,00
5	97,60	98,14	94,11
6	100,00	100,00	100,00
7	98,38	99,08	93,33
8	97,58	100,00	86,95
9	99,19	100,00	94,44
Média	98,39	99,18	93,47

Na Tabela II, temos um comparativo com outros métodos de classificação propostos na literatura. Como pode ser observado, o modelo proposto neste artigo, foi o que alcançou maior acurácia. Além disso, a sensibilidade alcançada, apesar de não ser a maior, mostra que o método é eficiente na detecção de verdadeiros positivos.

## V. CONCLUSÃO

A detecção eficiente de pacientes com COVID-19 é vital para evitar a propagação da doença a outras pessoas e salvar vidas. Neste estudo, propusemos uma abordagem baseada em

TABELA II Comparação dos resultados entre este e outros trabalhos. As abreviações são: Acurácia (Ac), Sensibilidade (Se), Especificidade (Es). Fonte: autor.

Autor	Classificador	Ac (%)	Se (%)	Es (%)
Abbas et al.	ResNet	95.12	97.91	91.87
Afshar et al.	CapsNet	95,7	90	95,8
Hemdan et al.	VGG19	90	-	-
Hemdan et al.	DenseNet201	90	-	-
Hemdan et al.	ResNetV2	70	-	-
Hemdan et al.	InceptionV3	50	-	-
Hemdan et al.	InceptionResNetV2	80	-	-
Hemdan et al.	Xception	80	-	-
Hemdan et al.	MobileNetV2	60	-	-
Narin et al.	InceptionV3	97	100	-
Narin et al.	ResNet50	98	100	-
Narin et al.	InceptionResNetV2	87	90	-
Wang et al.	COVID-Net	93	91	-
Método	ECNN	08.20	00.19	02.47
Proposto	FUNIN	98,39	99,10	95,47

redes neurais totalmente conectadas e recursos de textura, usando imagens de Raio-X do tórax, obtidas de pacientes com COVID-19 e de pacientes saudáveis, para predizer automaticamente pacientes com COVID-19. Os resultados de desempenho mostram que o modelo de Rede Neural Totalmente Conectada produziu um método de alta eficácia. À luz de nossas descobertas, acredita-se que isto ajudará os médicos a tomar decisões na prática clínica devido ao alto desempenho e ajudar a salvar cada vez mais vidas.

Em trabalhos futuros, este método poderá ser utilizado para diferenciar pacientes com pneumonia viral, de pacientes infectados com COVID-19.

#### REFERENCES

- Asmaa Abbas, Mohammed M Abdelsamea, and Mohamed Medhat Gaber. Classification of covid-19 in chest x-ray images using detrac deep convolutional neural network. *arXiv preprint arXiv:2003.13815*, 2020.
- [2] Parnian Afshar, Shahin Heidarian, Farnoosh Naderkhani, Anastasia Oikonomou, Konstantinos N Plataniotis, and Arash Mohammadi. Covidcaps: A capsule network-based framework for identification of covid-19 cases from x-ray images. arXiv preprint arXiv:2004.02696, 2020.
- [3] Ministério da Saúde. Quanto tempo o vírus sobrevive nas superfícies. Disponível em: https://coronavirus.saude.gov.br/index.php/ perguntas-e-respostas. Acessado em: 30 jun. 2020.
- [4] Organização Mundial das Nações Unidas (ONU). Organização mundial da saúde declara novo coronavírus uma pandemia. Disponível em: https: //news.un.org/pt/story/2020/03/1706881. Acessado em: 25 jun. 2020.
- [5] M. de Freitas Oliveira Baffa and L. Grassano Lattari. Convolutional neural networks for static and dynamic breast infrared imaging classification. In 2018 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), pages 174–181, 2018.
- [6] Site de notícias G1. países da Ásia, oceania, américa do norte, europa, oriente médio, américa do sul e África são afetados pela covid-19. Disponível em: https://g1.globo.com/ciencia-e-saude/noticia/2020/01/23/numero-de-paises-com-casos-confirmados-de-coronavirus.ghtml. Acessado em: 29 jun. 2020.
- [7] Folha de São Paulo. Grupo de risco do novo coronavírus. Disponível em: https://www1.folha.uol.com.br/equilibrioesaude/2020/05/ homens-e-idosos-sao-quem-mais-morre-de-covid-19-no-estado-de-sp. shtml. Acessado em: 29 jun. 2020.
- [8] Nicholas A Hamilton, Radosav S Pantelic, Kelly Hanson, and Rohan D Teasdale. Fast automated cell phenotype image classification. *BMC bioinformatics*, 8(1):110, 2007.

- [9] Robert M Haralick, Karthikeyan Shanmugam, and Its' Hak Dinstein. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, pages 610–621, 1973.
- [10] Ezz El-Din Hemdan, Marwa A Shouman, and Mohamed Esmail Karar. Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images. arXiv preprint arXiv:2003.11055, 2020.
- [11] Sara Hosseinzadeh Kassani, Peyman Hosseinzadeh Kassasni, Michal J Wesolowski, Kevin A Schneider, and Ralph Deters. Automatic detection of coronavirus disease (covid-19) in x-ray and ct images: A machine learning-based approach. arXiv preprint arXiv:2004.10641, 2020.
- [12] Ali Narin, Ceren Kaya, and Ziynet Pamuk. Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. arXiv preprint arXiv:2003.10849, 2020.
- [13] Caio Vinicius de Oliveira. Coronavírus: uso de tomografia computadorizada na detecção. Disponível em: https://brasilrad.com.br/artigos/ coronavirus-uso-de-tomografia-computadorizada-na-deteccao/. Acessado em: 30 jun. 2020.
- [14] Omir Antunes Paiva and Luciano M Prevedello. O potencial impacto da inteligência artificial na radiologia. *Radiologia Brasileira*, 50(5):V–VI, 2017.
- [15] Tawsifur Rahman, Dr. Muhammad Chowdhury, and Amith Khandakar. Covid-19 radiography database. Disponível em: https://www.kaggle. com/tawsifurrahman/covid19-radiography-database. Acessado em: 25 jun. 2020.
- [16] B. Ramsundar and R.B. Zadeh. TensorFlow for Deep Learning: From Linear Regression to Reinforcement Learning, chapter 4. O'Reilly Media, 2018.
- [17] sanarMed. Diagnóstico do coronavírus. Disponível em: https://www.sanarmed.com/ coronavirus-origem-sinais-sintomas-achados-tratamentos. Acessado em: 2 julho. 2020.
- [18] saude do viajante. Covid-19 associado a sars-cov-2 mundial. Disponível em: http://www.saudedoviajante.pr.gov.br/2020/04/94/ COVID-19-associado-a-SARS-CoV-2-Mundial.html. Acessado em: 25 jun. 2020.
- [19] Fengxiang Song, Nannan Shi, Fei Shan, Zhiyong Zhang, Jie Shen, Hongzhou Lu, Yun Ling, Yebin Jiang, and Yuxin Shi. Emerging 2019 novel coronavirus (2019-ncov) pneumonia. *Radiology*, 295(1):210–217, 2020.
- [20] Linda Wang and Alexander Wong. Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. arXiv preprint arXiv:2003.09871, 2020.
- [21] Tzu-Tsung Wong. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recognition*, 48(9):2839–2846, 2015.
- [22] Rikiya Yamashita, Mizuho Nishio, Richard Do, and Kaori Togashi. Convolutional neural networks: an overview and application in radiology. *Insights into Imaging*, 9, 06 2018.