

Observando o comportamento de exames médicos de COVID-19 no desempenho de classificação da ResNet50

Márcus V. L. Costa¹, Ricardo M. Marcacini¹

¹Instituto de Ciências Matemáticas e de Computação (ICMC)
Universidade de São Paulo (USP)
CEP: 13566-590 – São Carlos – SP – Brasil

marcusvllc@usp.br, ricardo.marcacini@usp.br

Resumo. Com a rápida disseminação da doença de COVID-19 em todo o mundo, a radiografia de tórax (CXR) e a tomografia computadorizada (CT) se tornaram meios para detectar, avaliar ou monitorar a doença. Ainda é um desafio a detecção e diferenciação da COVID-19 de outras doenças (pneumonias). Desta forma, este trabalho visa utilizar uma técnica de transfer learning, a ResNet50. O intuito é observar o poder de classificação da ResNet50 nas modalidades de imagens de exames médicos de CT e CXR. Em nossos resultados, o CXR acabou performando melhor, com uma acurácia de 85% ($AUC = 0,87$).

1. Introdução

Com a rápida disseminação da doença proveniente do SARS-CoV-2, ou mais sucintamente, COVID-19 em todo o mundo, a radiografia de tórax (CXR) tem sido utilizada como um dos métodos para detectar COVID-19 e avaliar sua gravidade ou monitorar seu prognóstico em hospitais devido ao baixo custo, baixa dose de radiação e ampla acessibilidade [Litmanovich et al. 2020]. A tomografia computadorizada (CT) também é comumente utilizada para este meio, pois tem maior sensibilidade em detectar os achados da COVID-19 [Larici et al. 2020, Poortahmasebi et al. 2020].

Ainda é um desafio à precisa detecção e diferenciação da pneumonia por COVID-19 dos outros tipos de pneumonias [Heidari et al. 2021]. No atual estado da arte, numerosos estudos têm apresentado métodos baseados em *machine learning* e *deep learning* aplicados na detecção e diagnóstico da doença de modo a criar um sistema de auxílio ao diagnóstico. Esses sistemas podem incluir um pré-processamento para a melhoria na qualidade da imagem [Rahman et al. 2021], além da diferenciação das síndromes respiratórias agudas utilizando apenas com técnicas de aprendizado de máquina, redes neurais convolucionais [Heidari et al. 2021, Marques et al. 2020], ou modelos que juntam ambas as abordagens, denominados sistemas híbridos [Shankar and Perumal 2021, Turkoglu 2021].

[Farooq and Hafeez 2020] apresentaram uma abordagem baseada em *deep learning* para distinguir os casos de COVID-19 dos CXRs. Eles relataram uma precisão de 96,23% com 41 épocas usando ResNet50, onde 8 casos positivos COVID-19 foram usados. [Narin et al. 2021] utilizou arquiteturas de *deep learning* para detectar COVID-19. Com 50 casos positivos de COVID-19, e relatou precisões de 97%, 98% e 87% de três arquiteturas: InceptionV3, ResNet50, e Inception-ResNetV2, respectivamente.

O objetivo do projeto é utilizar uma técnica de *transfer learning*, a ResNet50, para classificação de exames médicos de imagens de COVID-19 e não COVID-19, a

fim de observar o comportamento de ambas as modalidades de imagens, CXR e CT. O presente projeto tem como *background* o trabalho de [Mukherjee et al. 2021], no qual será realizado uma abordagem simplista ao realizado pelo autor.

2. Metodologia

2.1. Coleta de dados

As bases de dados de CT e CXR utilizados neste projeto são provenientes de três fontes. Isto significa que, para evitar enviesamentos que podem acontecer na utilização da ResNet50, em nossos experimentos, foram criados dois conjuntos de dados de diferentes fontes.

- (a) **Coleta dos CXR:** Duas fontes de *datasets* foram utilizados para essa coleta. A primeira coleta foi realizada no *dataset covid-chestxray-dataset* [Cohen et al. 2020], é composta por 196 casos positivos de COVID-19, e 38 casos de Non-COVID-19. Non-COVID-19 incluem casos de outras doenças pulmonares como MERS, SARS e ARDS. Para balancear o número de instâncias nos casos de Non-COVID-19, foram adicionadas 70 casos normais de pacientes saudáveis e outros 76 casos de pacientes com pneumonia, coletados do *dataset Chest X-Ray Images (Pneumonia)* [Kermany et al. 2018]. Ao todo, 196 casos de COVID-19 e 196 casos de Non-COVID-19.
- (b) **Coleta das CT:** Esta coleta de dados foi realizada somente no *dataset COVID-CT* [Zhao et al. 2020b]. Foram coletados 196 casos de COVID-19 e 196 casos de Non-COVID-19.

Ressalta-se que estes dados são disponibilizados em repositórios gratuitos. Além disso, as imagens possuem tamanhos diferentes e ruídos, dado o fato acima. A Tabela 1 apresenta a distribuição das imagens, e na Figura 1 é apresentado alguns exemplos dos conjuntos de dados.

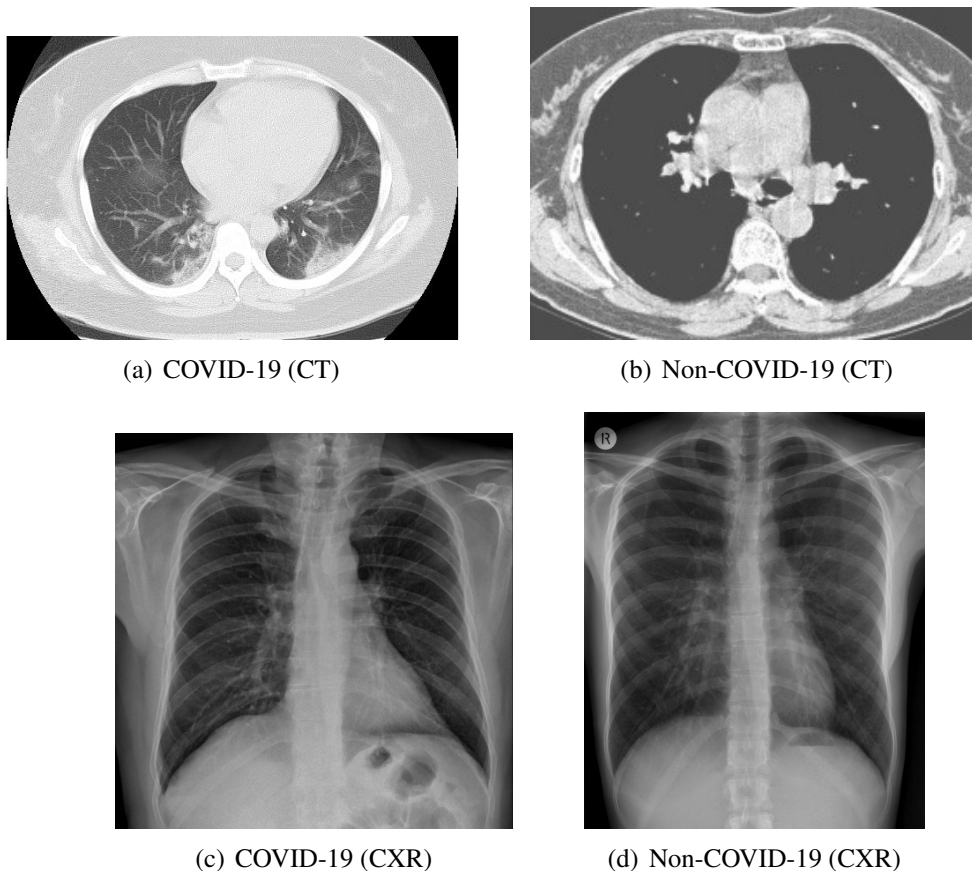
Tabela 1. Coleção dos *datasets*.

| Tipo | Casos COVID-19 | Casos Non-COVID-19 | Total |
|--------|----------------|--------------------|-------|
| CT | 196 | 196 | 392 |
| CXR | 196 | 196 | 392 |
| CT+CXR | 392 | 392 | 784 |

2.2. Data augmentation

O *data augmentation* ajuda a criar exemplos mais recentes através de aplicações de transformações diferentes aleatoriamente disponíveis nas imagens de treinamento. Neste trabalho foram acrescidos as seguintes transformações nas imagens, rotação (50 graus), deslocamento da altura e largura (rotação em 20 graus), *zoom* (10%) e rotações aleatórias na vertical e horizontal. O *data augmentation* aumenta os dados de treinamento de entrada e melhora a generalização do modelo de treinamento.

Figura 1. Alguns exemplos das imagens utilizadas.



2.3. Arquitetura da rede

Ao invés de propor uma arquitetura, optou-se por utilizar uma arquitetura de rede neural convolucional (CNN) que apresenta excelentes resultados em modelos de *deep learning*, em especial nas tarefas de classificação. A arquitetura utilizada foi a *residual neural network* com total de 50 camadas, chamada ResNet50 [He et al. 2016]. A ResNet50 prover uma boa combinação de desempenho, número de parâmetros e um rápido treinamento.

Os pesos utilizados na ResNet50 são pré-treinados com o conjunto de dados da ImageNet [Deng et al. 2009]. Para o aprendizado por transferência, a parte não treinada é alterada para a sequência de *average pooling 2d*, com camadas de 1024, 1024, 512, 256 e 2 neurônios acrescido da função *softmax*, respectivamente. A Tabela 2 apresenta como ficou estruturada a parte não treinada da rede.

2.4. Treinamento da rede

Na fase de treinamento as imagens foram redimensionadas para 224x224 *pixels* com 3 canais de cor, acrescidos do *data augmentation*. O modelo foi treinado por 100 épocas, com *batch size* de tamanho 16 e com a função de ativação *Adam*. Estes parâmetros foram escolhidos empiricamente.

Tabela 2. Tipos de camadas utilizadas no modelo ResNet50.

| Camada (Tipo) |
|------------------------------|
| ResNet50 (Modelo) |
| global_average_pooling2d |
| dense (Dense) |
| dense_1 (Dense) |
| dense_2 (Dense) |
| dense_3 (Dense) |
| dense_4 (Dense) |
| Total params: 27,392,130 |
| Trainable params: 27,339,010 |
| Non-trainable params: 53,120 |

2.5. Protocolo de avaliação

Para validar o modelo optou-se pela utilização da técnica *holdout*, 80:20. 80% dos dados foram utilizados para treinamento e 20% para teste. Para avaliar o desempenho da rede, foram utilizadas as seguintes métricas: Acurácia e a Área sob a curva ROC (AUC). Eles foram computados como:

$$Acurácia = \frac{VP + VN}{VP + VN + FP + FN}, \quad (1)$$

onde VP, VN, FP, e FN se referem a verdadeiro positivo, verdadeiro negativo, falso positivo, e falso negativo, respectivamente.

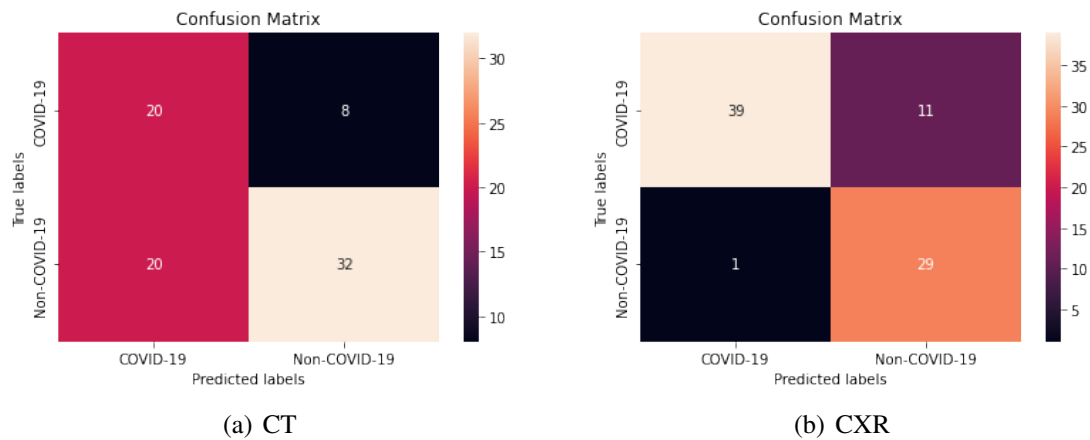
3. Resultados e discussões

Os resultados obtidos são apresentados nesta seção. Para cada modalidade imagem apresentamos a acurácia, matriz de confusão, curva ROC e as curvas de aprendizagem. Na Tabela 3 é possível perceber que o modelo CXR performou melhor, ao contrário do CT.

Tabela 3. Resultado dos dois modelos.

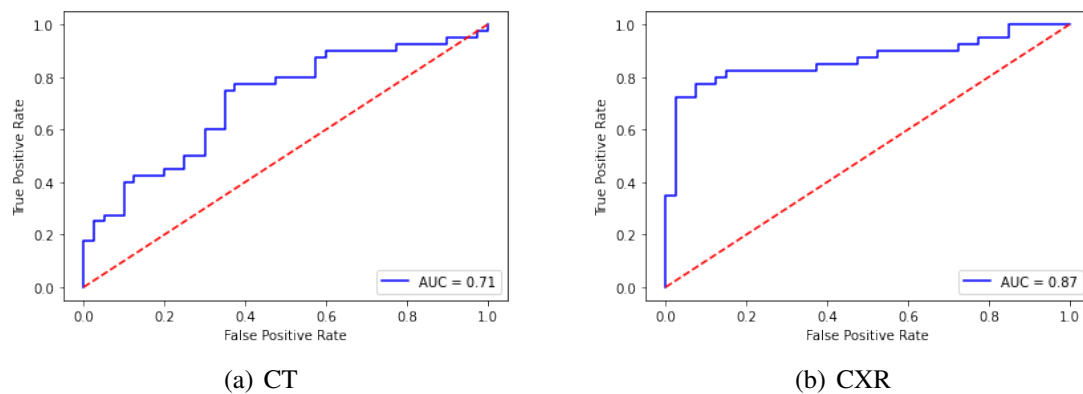
| Modalidade | Acurácia | Desvio Padrão | Acurácia | AUC |
|------------|-----------------|---------------------|-----------|------|
| | Treinamento (%) | | Teste (%) | |
| CT | 57,41 | 0.0658894659858155 | 65 | 0,71 |
| CXR | 69,84 | 0.11238403757194775 | 85 | 0,87 |

Figura 2. Matriz de confusão dos dois modelos.



A Figura 3(a) e Figura 3(b) apresentam, respectivamente, a curva ROC do modelo de CT e CXR. As curvas dos modelos apresentam resultados satisfatórios, o modelo CXR performou melhor com uma $AUC = 0,87$, que de certa forma encontra-se mais próximo do valor 1, proveniente de uma elevada sensibilidade em acertar os dados.

Figura 3. Curva ROC dos dois modelos.



Os gráficos nas Figuras 4 e 5 representam as curvas de aprendizagem do modelo, onde o primeiro gráfico apresenta a acurácia do treino e teste pelas épocas e o segundo gráfico apresenta a perda ou erro dos mesmos, respectivamente. Ambos os gráficos possuem picos desordenados na acurácia do modelo, mas que conseguem convergir para uma resolução.

Figura 4. Curvas de aprendizagem CT.

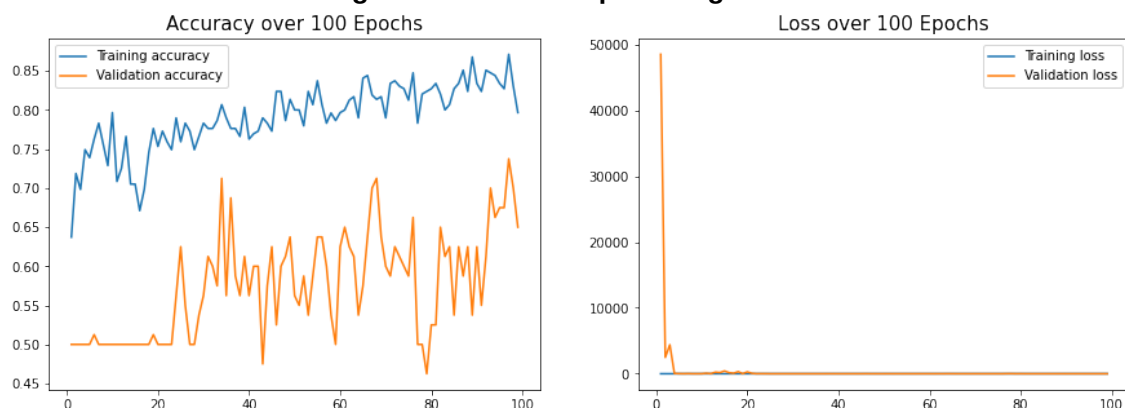
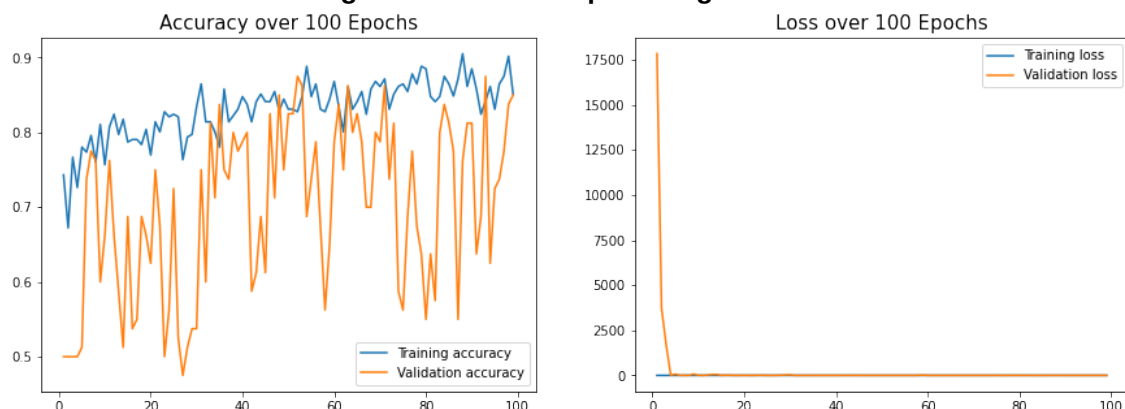


Figura 5. Curvas de aprendizagem CXR.



Portanto, os resultados encontrados são satisfatórios com a ResNet50. Embora os dados sejam coletados de diversas fontes *online*s, não possuem um padrão de qualidade igualitária entre os centros de coleta que as disponibilizam. Isto está ligado ao fato de que os equipamentos utilizados não possuem um paradigma padrão no resultado dos exames [Zhao et al. 2020a]. Além de várias imagens possuírem marcações, resoluções diferentes, contraste, e brilho diferentes, como apresentado na Figura 6.

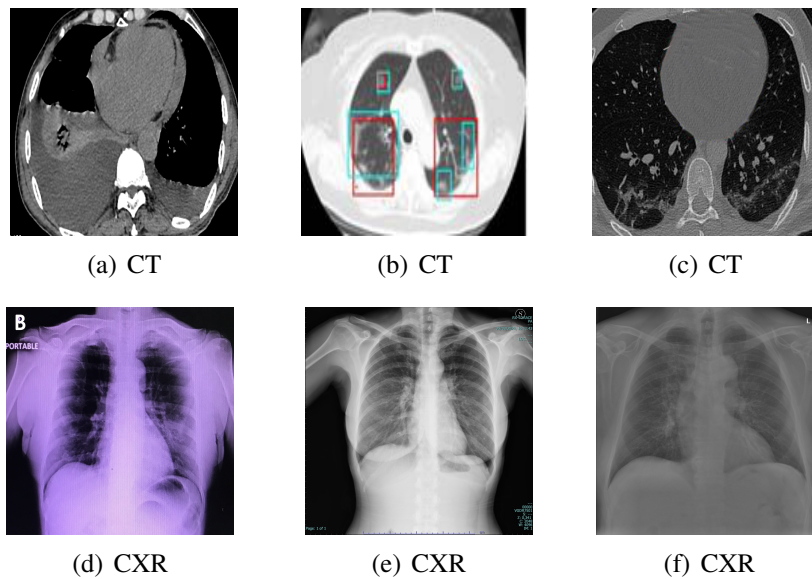
4. Conclusões

Neste projeto, utilizamos a arquitetura ResNet50 para diferentes modalidades de dados para detectar casos positivos COVID-19. Treinamos e testamos ambos os tipos de dados; para as imagens CT alcançamos uma acurácia de 65% ($AUC = 0,71$), e para imagens de CXR alcançamos uma acurácia de 85% ($AUC = 0,87$). É necessário em testes futuros abordar uma forma de normalizar os dados/imagens, além de realizar teste com outras arquiteturas de *transfer learning*.

Apêndices

Os experimentos de todo o projeto estão disponíveis em meu repositório do GitHub.

Figura 6. Imagens do *dataset* com várias inconsistências.



Referências

- [Cohen et al. 2020] Cohen, J. P., Morrison, P., Dao, L., Roth, K., Duong, T. Q., and Ghassemi, M. (2020). Covid-19 image data collection: Prospective predictions are the future. *arXiv 2006.11988*.
- [Deng et al. 2009] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition*, pages 248–255. Ieee.
- [Farooq and Hafeez 2020] Farooq, M. and Hafeez, A. (2020). Covid-resnet: A deep learning framework for screening of covid19 from radiographs. *arXiv preprint arXiv:2003.14395*.
- [He et al. 2016] He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778.
- [Heidari et al. 2021] Heidari, M., Mirniaharikandehei, S., Khuzani, A. Z., Danala, G., Qiu, Y., and Zheng, B. (2021). Detecting covid-19 infected pneumonia from x-ray images using a deep learning model with image preprocessing algorithm. 11597:115970V.
- [Kermany et al. 2018] Kermany, D. S., Goldbaum, M., Cai, W., Valentim, C. C., Liang, H., Baxter, S. L., McKeown, A., Yang, G., Wu, X., Yan, F., et al. (2018). Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*, 172(5):1122–1131.
- [Larici et al. 2020] Larici, A. R., Cicchetti, G., Marano, R., Merlino, B., Elia, L., Calandriello, L., Del Ciello, A., Farchione, A., Savino, G., Infante, A., et al. (2020). Multimodality imaging of covid-19 pneumonia: from diagnosis to follow-up. a comprehensive review. *European Journal of Radiology*, page 109217.

- [Litmanovich et al. 2020] Litmanovich, D. E., Chung, M., Kirkbride, R. R., Kicska, G., and Kanne, J. P. (2020). Review of chest radiograph findings of covid-19 pneumonia and suggested reporting language. *Journal of thoracic imaging*, 35(6):354–360.
- [Marques et al. 2020] Marques, G., Agarwal, D., and de la Torre Díez, I. (2020). Automated medical diagnosis of covid-19 through efficientnet convolutional neural network. *Applied soft computing*, 96:106691.
- [Mukherjee et al. 2021] Mukherjee, H., Ghosh, S., Dhar, A., Obaidullah, S. M., Santosh, K., and Roy, K. (2021). Deep neural network to detect covid-19: one architecture for both ct scans and chest x-rays. *Applied Intelligence*, 51(5):2777–2789.
- [Narin et al. 2021] Narin, A., Kaya, C., and Pamuk, Z. (2021). Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. *Pattern Analysis and Applications*, pages 1–14.
- [Poortahmasebi et al. 2020] Poortahmasebi, V., Zandi, M., Soltani, S., and Jazayeri, S. M. (2020). Clinical performance of rt-pcr and chest ct scan for covid-19 diagnosis; a systematic review. *Advanced Journal of Emergency Medicine*, 4(2s):e57–e57.
- [Rahman et al. 2021] Rahman, T., Khandakar, A., Qiblawey, Y., Tahir, A., Kiranyaz, S., Kashem, S. B. A., Islam, M. T., Al Maadeed, S., Zughaier, S. M., Khan, M. S., et al. (2021). Exploring the effect of image enhancement techniques on covid-19 detection using chest x-ray images. *Computers in biology and medicine*, 132:104319.
- [Shankar and Perumal 2021] Shankar, K. and Perumal, E. (2021). A novel hand-crafted with deep learning features based fusion model for covid-19 diagnosis and classification using chest x-ray images. *Complex & Intelligent Systems*, 7(3):1277–1293.
- [Turkoglu 2021] Turkoglu, M. (2021). Covidetectionnet: Covid-19 diagnosis system based on x-ray images using features selected from pre-learned deep features ensemble. *Applied Intelligence*, 51(3):1213–1226.
- [Zhao et al. 2020a] Zhao, G., Bai, J., Wang, P. P., Yang, G., Shi, L., Tao, Y., Lin, Y., and Cheng, J. (2020a). Hs-gs: A method for multicenter mr image standardization. *IEEE Access*, 8:158512–158522.
- [Zhao et al. 2020b] Zhao, J., Zhang, Y., He, X., and Xie, P. (2020b). Covid-ct-dataset: a ct scan dataset about covid-19. *arXiv preprint arXiv:2003.13865*.