

Uso de imagens de pulmão para identificação de Covid-19

Brenda Uemura Silva

ICT – Instituto de Ciência e Tecnologia
Universidade Federal do Estado de São
Paulo (UNIFESP)

São José dos Campos – São Paulo
brenda.uemura@unifesp.br

Conrado Gobato

ICT – Instituto de Ciência e Tecnologia
Universidade Federal do Estado de São
Paulo (UNIFESP)

São José dos Campos – São Paulo
conrado.gobato@unifesp.br

Abstract— The Covid-19 pandemic was declared in 2020 by the World Health Organization. In face of the pandemic scenario confronted in Brazil, this paper have the objective of providing assistance to the health sector through the developing of an artificial intelligence, using neuron networks and machine learning, capable of detecting in X-ray images characteristics that may be caused by the SARS-cov-2 virus, considering that the illness have as its primary characteristic it's effects on the respiratory system.

Key-words—Covid-19, Machine learning, Artificial Intelligence, Convolutional Neuron Network, Pulmonary X-ray Images .

Resumo— A pandemia de Covid-19 foi declarada em 2020 pela Organização Mundial da Saúde. Frente ao Quadro pandêmico vivido no Brasil e no mundo, esse trabalho tem por objetivo auxiliar à área da saúde através do desenvolvimento com redes neurais e treinamento machine learn uma inteligência artificial capaz de detectar em imagens Raio X pulmonares características que possam ser advindas do vírus SARS-cov-2, visto que a doença possui a característica de atacar principalmente o trato respiratório.

Palavras-Chave—Covid-19, Machine learn, Inteligência Artificial, Rede Neural Convolucional, Imagnes de Raio-X, Pulmão.

I. INTRODUÇÃO E MOTIVAÇÃO

Tendo seu primeiro caso registrado em 17 de novembro de 2019, o coronavírus ocasionou mais de 6.874.084 mortes mundiais. Causando ainda óbitos no território nacional brasileiro em decorrência da doença todos os dias. Visto os danos causados pelo vírus, suas graves consequências caso o diagnóstico e o tratamento não sejam realizados de forma rápida e correta, nosso trabalho visa corroborar com o diagnóstico precoce, através da identificação de irregularidades pulmonares em imagens radiográficas. Isso devido ao fato da doença atacar principalmente a área respiratória, como citado no trecho “Os coronavírus (CoV) são uma grande família de vírus que causam doenças que variam do resfriado comum a enfermidades mais graves, como Síndrome Respiratória do Oriente Médio (MERS-CoV) e Síndrome Respiratória Aguda Grave (SARS-CoV)”, do artigo “COVID-19 e a infecção por SARS-Cov-2 em um panorama geral” [1].

II. CONCEITOS FUNDAMENTAIS

Com o avanço rápido de mortes na pandemia, as radiografias de pulmão tiveram um papel essencial como ferramenta para o diagnóstico e acompanhamento da doença nos pacientes. As radiografias são úteis na identificação de características típicas da infecção pelo coronavírus, como a presença de opacidades pulmonares, distribuídas perifericamente e nas regiões subpleurais. O envolvimento de múltiplos lobos, em particular os inferiores, foi documentado na maioria dos pacientes, podendo as áreas de consolidação focal acompanhar estas imagens, assim como reticulações intralobulares. Com a evolução da doença, surgem sinais de consolidação que denotam pneumonia em organização vista através do sinal do halo invertido. Tendo assim a radiografia sua importância relacionada com a rapidez e precisão diagnóstica, sendo um exame não invasivo e de baixo custo.

Essa importância também é descrita no trecho ‘O diagnóstico precoce da Covid-19 é essencial para a estabilização do paciente e contenção da pandemia, assim sendo, os casos suspeitos podem ser investigados e posteriormente confirmados, por meio da associação dentre teste laboratoriais, exames de imagem e investigação clínico-epidemiológica’, presente no artigo ‘Diagnóstico de COVID-19: importância dos testes laboratoriais e dos exames de imagem’ [2].

III. TRABALHOS RELACIONADOS

Dentre os vários estudos realizados, dentro e fora de sala de aula, trazemos aqui outros trabalhos presentes na literatura que contenham temas relacionados e correlacionados com nossa atual pesquisa.

Para incentivo e aprendizado do nosso projeto, encontramos alguns estudos já feitos sobre a IA inserida na área da saúde, e sobre redes convolucionais.

1. O artigo "Artificial intelligence and forecasting of death by COVID-19 in Brazil: A comparative analysis of the algorithms Logistic Regression, Decision Tree, and Random Forest"[3] escrito por Risomario Silva e Darcy Ramos da Silva Neto, demonstrou uma pesquisa feita por eles, que contribui com evidências que auxiliam na previsão da morte por Covid-19. Para esse estudo, foram utilizados 3 modelos diferentes sendo eles: Regressão Logística, Árvore de decisão e Floresta aleatória, que consiste em uma combinação de decision trees. Como resultado, foi obtido que pacientes que apresentam quadro clínico severo, associado a presença de

comorbidades, estão mais propensos a fazer o uso de suporte ventilatório, e, portanto, eles têm o maior risco de morte, já que os dados apontaram que 80% dos pacientes que foram intubados vieram a óbito.

2. Em "Segmentação de Nódulos Pulmonares Baseada em Redes Neurais Convolucionais (CNN)"[4], é tratado o uso de redes convolucionais para a detecção de nódulos pulmonares orientados a patches e múltiplas orientações. Os resultados obtiveram altas taxas de acurácia quando comparadas com especialistas, e tornou possível reduzir o número de falsos negativos.

3. A utilização de redes convolucionais para diagnóstico de Alzheimer "Utilização de Redes Convolucionais Para Classificação e Diagnóstico da Doença de Alzheimer"[5] utilizou de bases de dados MIRIAD, para realizar a classificação de indivíduos. Os experimentos apresentaram valores mais elevados de acurácia (acima de 82%) tanto com poucas épocas (5 e 15) como também quando foi utilizada uma quantidade maior de épocas (35, 45 e 50).

4. O artigo "Análise da seleção de hiperparâmetros de Data Argumentation na detecção de Covid-19 em imagens de raio-x com Deep Learning"[6] aborda sobre a escolha de hiperparâmetros (zoom, rotação, intensidade do

brilho e espelhamento horizontal) da imagem, para obter um melhor resultado, na predição de diagnóstico de covid-19. Na pesquisa, é apresentado a acurácia obtida com cada conjunto de parâmetros selecionados.

5. A detecção do vírus da Covid-19, por meio de redes convolucionais, é uma área que foi muito pesquisada durante a pandemia. Em "Detecção de Covid-19 em Imagens de Raio-x Utilizando Redes Convolucionais"[7] foi-se utilizada uma Rede Neural Convolucional, em que a base de dados passou por todos os processos de pré-processamento, e a rede foi treinada do zero, levando aos seguintes resultados: 85,37% de acurácia, 83,60% de precisão, 86,44% de sensibilidade e 85% de F1-Score.

6. Na pesquisa "Rede Neural Convolucional para o diagnóstico de Leucemia"[8], redigida por Luis H. S. Vogado, Rodrigo M. S. Veras, Flavio H. D. Araujo, Romuere R. V. Silva, Kelson R. T. Aires, uma rede convolucional pré treinada acompanhada da técnica de transferência de aprendizagem (utilização de um modelo base, com ajustes) foi capaz de prever imagens de células com Leucemia e atingiu 98,84% de acurácia. No artigo, ainda foi destacado as vantagens da utilização de redes pré treinadas, como a diminuição do custo computacional.

7. No artigo "Detecção de pneumonia usando redes neurais convolucionais treinadas com destilação do conhecimento obscuro"[9], os autores: José Vitor Santos Silva, Leonardo Nogueira Matos criaram um modelo para diagnosticar pulmões com pneumonia. Nessa pesquisa, foi adotado a técnica de 'Destilação do Conhecimento', que visa melhorar a capacidade de generalização de uma rede neural. Basicamente, cria-se dois modelos, um modelo professor (modelo profundo), e um modelo estudante (modelo de arquitetura mais simples), e o objetivo é transferir o conhecimento do modelo professor para o estudante.

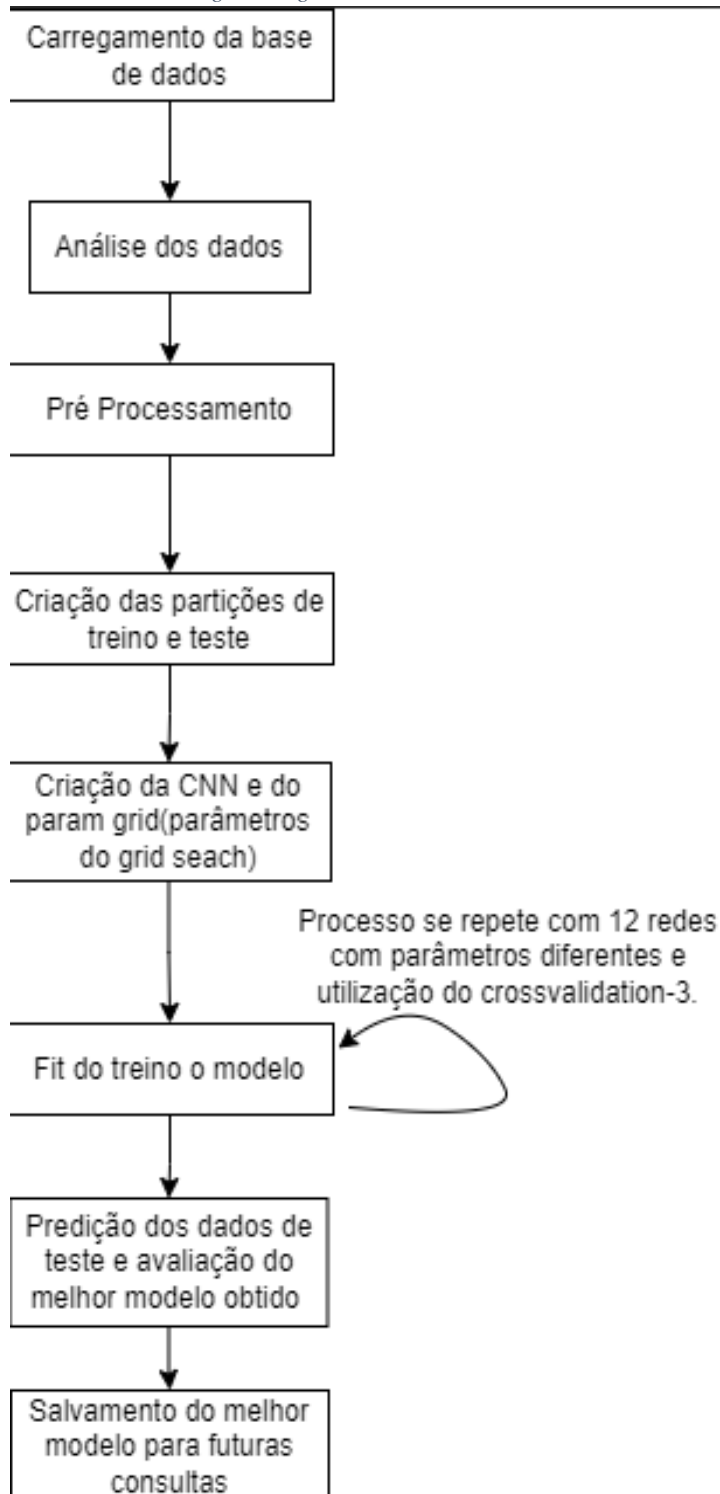
IV. OBJETIVOS

Como Machine Learn é capaz de prever com maior precisão através de treinamentos diagnósticos por imagens, nosso objetivo consiste em contribuir para a comunidade médica, acadêmica e populacional através da aprendizagem de máquina. Sabendo que um diagnóstico preciso de imagens é essencial para o monitoramento da evolução da doença e avaliação da eficácia do tratamento aplicado pelo médico, pois possibilitam avaliar a progressão ou regressão dos sintomas, nosso projeto se propõe a otimizar a identificação de casos suspeitos, podendo assim ser iniciado rapidamente o tratamento adequado juntamente com o parecer médico.

V. METODOLOGIA EXPERIMENTAL

Para a realização do projeto, iremos utilizar principalmente de duas bases de dados, a primeira contendo imagens de pulmões diagnosticados com covid-19 e a segunda com imagens de pulmões saudáveis registrados em imagens de raio-x. Para a manipulação das imagens será utilizado a biblioteca Pillow existente na linguagem de programação Python, possibilitando trabalhar com o tamanho e filtros de imagens. Utilizaremos além disso a biblioteca Sklearn e seus métodos para auxiliar na manipulação dos dados, transformá-los em categóricos e avaliação do desempenho dos classificadores. Os modelos serão baseados em 12 setups para uma Rede Neural convolucional (um para cada combinação de parâmetros do Grid Search) contendo três camadas de convolução 2D com parâmetros variados (sendo uma camada de input), duas camadas de Pooling 2D 2x2, uma camada de Flatten para ajustar os pixels à rede, e 2 camadas densas comuns (sendo uma de saída), que irá prever a classe do paciente conforme a imagem dada. Para sua criação iremos usufruir da biblioteca Keras e suas funções. Com o objetivo de alcançar o melhor resultado possível, será também utilizado a técnica de grid search, que consiste em variar parâmetros da rede, visando sempre uma melhora no desempenho, que será realizado através da variação do filtro e tamanho das imagens com o objetivo de verificar se ocorre uma mudança positiva no parâmetro de acertos.

Fig 1 - Diagrama de Blocos



Fonte: de autoria própria

1. BASE DE DADOS

As bases de dados utilizadas, foram encontradas separadamente. A base contendo pacientes diagnosticados com covid-19, foi obtida em um repositório no GitHub [10], possuindo 930 imagens, esses dados, não apresentaram nenhuma inconsistência. Já para referência de pulmões saudáveis, foi necessário particionar uma base presente no Kaggle [11], com 1575 imagens, de pulmões

com pneumonia e saudáveis. Assim, juntamos as informações obtidas, e transformamos em apenas um conjunto, contendo todas as imagens necessárias.

2. PROTOCOLO DE VALIDAÇÃO

Como protocolo de validação, utilizamos as técnicas de Grid Search, e Cross Validation. O Grid Search consiste na variação de parâmetros da rede (batch size, número de épocas, loss, otimizador...) para buscar um melhor resultado. Já o Cross Validation, é uma técnica que particiona os dados em k folds, que utiliza um fold para teste, e o restante para treino, variando os folds utilizados até o fim da base. Para a implementação desses protocolos, foi utilizado a biblioteca Sklearn, e sua função chamada GridSearchCV, que recebe como parâmetros o modelo a ser utilizado, o param grid (parâmetros definidos dentro de um dicionário), e o número de folds desejado.

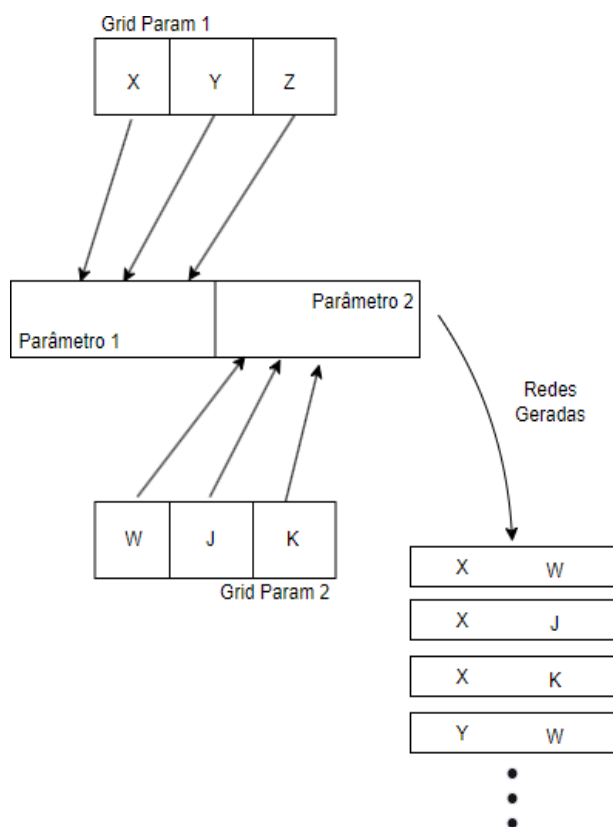
VI. CONCEITOS IMPORTANTES

As estratégias adotadas para a realização do projeto, e para um resultado consistente foram o estudo sobre técnicas utilizadas por outros pesquisadores, além de protocolos de validação eficientes. A utilização do Grid Search, foi escolhida pois realiza uma comparação de uma grande quantidade de setups com diferentes parâmetros que podem melhorar muito a acurácia do modelo. Em alguns testes realizados com o otimizador Adagrad (otimizador que utiliza diferentes taxas de aprendizado para cada parâmetro com base na iteração), o modelo não conseguia passar de 63% de acurácia, isso porque o otimizador não era o melhor dado o problema que estamos tratando. Dessa forma, é possível perceber a importância da utilização do Grid Search, onde será validado qual o conjunto de parâmetros combinados que nos trará os melhores resultados. O funcionamento desse método, pode ser observado na *Figura 2*.

Além do Grid Search, nós utilizamos o protocolo Cross Validation, explicado na subseção II da seção 'Metodologia experimental'.

A rede utilizada consiste em uma CNN densa, com 9 camadas (incluindo camada de entrada e saída). Considerando que nosso problema é de classificação (classifica entre duas classes), as métricas de avaliação escolhidas foram: Acurácia (Verdadeiros positivos + Verdadeiros negativos/ total), Acurácia Balanceada (que se assemelha a acurácia, porém é mais sensível a modelos que erram ou acertam muito apenas uma classe), Precision, F1 Score, Recall e Especificidade. Além disso, também foi utilizado uma matriz de confusão para a melhor visualização das previsões do modelo.

Fig 2 - Funcionamento do método Grid Search



Fonte: de autoria própria

VII. RESULTADOS E DISCUSSÕES

Nos primeiros experimentos realizados, utilizamos dois otimizadores que não obtiveram resultados tão convincentes. A princípio, optamos pela utilização do Adagrad que, apesar de obter 63% de acurácia no treino, obteve apenas 50% de acurácia balanceada nos dados de teste, devido a rede ter classificado todos os exemplos como saudáveis. Isso pode ter ocorrido devido a esse otimizador ajustar o learning rate automaticamente, o que pode ter levado ele a se estagnar em um platô quando atingiu a acurácia de 63%. Ao nos depararmos com essa situação, decidimos testar outro otimizador, e nesse caso usamos o RMSprop, com esse parâmetro nossa rede foi capaz de atingir 70% de acurácia nos dados de treino, e 72% nos dados de teste, o que não é um resultado espetacular, porém é uma melhora comparada ao primeiro experimento. Ainda não satisfeitos, e seguindo a estratégia adotada no começo do código, utilizamos a técnica do Grid search, que a princípio geraria 64 setups com parâmetros diferentes. Após alguns testes e análises, percebemos que a utilização de tantos parâmetros, e da base completa no grid, era um custo computacional muito caro, e que não nos traria uma acurácia melhor que utilizar a base reduzida, e menos parâmetros, já que havíamos acertado na escolha deles. Dessa maneira, reduzimos nosso param grid, com os params que estavam nos gerando a melhor acurácia por época, e acabamos utilizando 12 setups para uma rede neural, com três otimizadores diferentes e dois valores para o número de épocas e o batch size. Ao final da comparação, dentre os 12 setups gerados (apresentadas na Fig 4) escolhemos os três melhores para realizarmos o treino com a base completa, e como resultado, obtivemos incríveis 98.6% de acurácia, com a melhor rede do grid search, além de 98.1% de f1 score, 98.6% de acurácia balanceada, 96.8% de

especificidade, 97.3% de precision e 98.9% de recall. Ao gerarmos a matriz de confusão do modelo (apresentada na Fig 3), observamos que dos 500 exemplares, a nossa rede obteve apenas 2 valores para falsos negativos, e 5 valores para falsos positivos, o que é um resultado muito satisfatório, ainda mais levando em conta o peso de apontar diagnósticos falsos, para o caso em que o paciente está doente. Dessa maneira, foi possível concluir a importância da variação de parâmetros, e da análise de cada um deles, para obter um resultado convincente e que pode ser utilizado. Por fim, utilizamos a biblioteca Joblib, para salvar nosso melhor modelo e seus pesos, para uso futuro, em outras aplicações como ajuda em diagnósticos. Os parâmetros que variamos, alteraram nos resultados de forma que: o batch size, é o número de interações que o código faz, antes de atualizar os pesos. Como pudemos ver, 334 foi o melhor resultado talvez por atualizar o peso mais vezes que 668, mas vale ressaltar que em testes feitos anteriormente, o batch size muito pequeno não obteve uma melhora considerável em relação à 334, além de os outros dois melhores resultados virem de um batch size alto que é 668. O número de épocas, de modo semelhante ao batch size, não apresentou uma melhora muito intensa com valores muito grandes. Em testes realizados com valores maiores, percebemos que o custo computacional estava sendo muito elevado, e que os melhores resultados estavam vindo de 5 e 10 épocas. Já os otimizadores escolhidos, foram o SGD, o Adam, e o Adamax. O SGD, que foram utilizados por meio da biblioteca Keras, e vinham configuradas. Um valor interessante e importante padronizados pelos otimizadores da biblioteca é o Learning rate, que para o SGD foi 0.01, para o Adamax (Segundo melhor otimizador) foi 0.001 e o adam (Melhor otimizador utilizado) de 0.001 também. Os otimizadores com o learning rate menor, dão passos menores até chegar no resultado esperado, e isso pode ter influenciado nos que obtiveram melhor acurácia. O resultado obtido, foi muito bom e a utilização dos parâmetros foram certos, além de um custo computacional baixo, comparado ao da estratégia inicial.

Fig 3 – Matriz de confusão do melhor classificador

true label	predicted label	
	0	1
0	312	5
1	2	182

Fonte: de autoria própria

Figura 4 - Tabela de setups gerados

Acurácia	Batch Size	Épocas	Otimizado
77.0%	334	5	adam
62.9%	334	5	SGD
68.8%	334	5	Adamax
83.3%	334	10	adam
63.4%	334	10	SGD
63.4%	334	10	Adamax
69.6%	668	5	adam
62.9%	668	5	SGD
82.0%	668	5	Adamax
79.9%	668	10	adam
62.9%	668	10	SGD
81.5%	668	10	Adamax

Fonte: de autoria própria

REFERÊNCIAS

- [1] S. Júnior, R. Kairala, A. Pereira, G. Costa, R. Cruz, J. Junior, "COVID-19 e a infecção por SARS-CoV-2 em um panorama geral" 2020 [Online] Avaliable: <https://ojs.brazilianjournals.com.br/ojs/index.php/BJHR/article/view/9101/7732>
- [2] F. Lima, L. Gomes, C. Santos, G. Oliveira, "Diagnóstico de COVID-19: importância dos testes laboratoriais e dos exames de imagem" 2020 [Online] Avaliable: <https://rsdjournal.org/index.php/rsd/article/view/7162/6404>
- [3] R. Silva, D. Neto, "Inteligência artificial e previsão de óbito por covid-19 no Brasil: uma análise comparativa entre os algoritmos Logistic Regression, decision Tree e Random Forest," 2022 [Online] Avaliable: <https://www.scielo.org/article/sdeb/2022.v46nspe8/118-129/pt/#>
- [4] P. Amorim, J. Silva, T. Moraes, H. Pedrini, "Segmentação de Nódulos Pulmonares Baseada em Redes Neurais Convolucionais (CNN)" [Online] 2019. Avaliable: https://www.gov.br/cti/pt-br/publicacoes/producao-cientifica/seminario-pci/ix_seminario_pci-2019/pdf/seminario-2019_paper_23.pdf
- [5] I. Silva, R. Souza, G. Silva, C. Oliveira, L. Cavalcanti, R. Bezerra, "Utilização de Redes Convolucioais Para Classificação e Diagnóstico da Doença de Alzheimer" [Online] Avaliable: https://www.researchgate.net/profile/Wellington-Dos-Santos/publication/327384767_Utilizacao_de_Reddes_Convolucionais_para_Classificacao_e_Diagnostico_da_Doenca_de_Alzheimer/links/5b8b5f62299bf1d5a737f815/Utilizacao-de-Reddes-Convolucionais-para-Classificacao-e-Diagnostico-da-Doenca-de-Alzheimer.pdf
- [6] P. Rici, S. Santos, A. Ottoni, "Análise da seleção de hiperparâmetros de Data Augmentation na detecção de Covid-19 em imagens de raio-x com Deep Learning", 2021. [Online] Avaliable: https://sbic.org.br/wp-content/uploads/2021/09/pdf/CBIC_2021_paper_22.pdf
- [7] P. Leão, N. Freire, R. Pinto, K. Pinto, R. Giusti, E. Santos, " Detecção de Covid-19 em Imagens de Raio-X Utilizando Redes Convolucionais. 2020 [Online] Avaliable: <https://jhi.sbis.org.br/index.php/jhi-sbis/article/view/844/447>
- [8] L. Vogado, R. Veras, F. Araujo, R. Silva, K. Aires, "Rede Neural Convolucional para o Diagnóstico de Leucemia" [Online] Avaliable: <https://sol.sbc.org.br/index.php/sbcas/article/view/6241/6139>
- [9] J. Silva, L. Matos, " Detecção de pneumonia usando redes neurais convolucionais treinadas com destilação do conhecimento obscuro" [Online] Avaliable: <https://sol.sbc.org.br/index.php/erbase/article/view/15460/15303>
- [10] Covid-chestxray-dataset. Github, 2021. [Online] Avaliable: <https://github.com/ieee8023/covid-chestxray-dataset>
- [11] Chest X-Ray Images (Pneumonia). Kaggle, 2018. [Online] Avaliable: <https://www.kaggle.com/datasets/paultimothymooney/chest-xray-pneumonia>
- [12] L. Lobo, "Artificial Intelligence and Medicine" 2017 [Online] Avaliable: <https://www.scielo.br/j/rbem/a/f3kqKJjVQJxB4985fDMVb8b/?lang=pt#>