

Detecção de pneumonia por redes neurais Convolucionais: AlexNet, ResNet e VGG16.

Giovanni Torres Chaves 198105 g198105@dac.unicamp.br
Leticia Domingos Duarte 178797 1178797@dac.unicamp.br
Marcelo da Silva Lourenço 221418 m221418@dac.unicamp.br
Nathan Brusamarello de Souto 222854 n222854@dac.unicamp.br

I. RESUMO

A pneumonia é uma doença com risco de vida, que ocorre nos pulmões causada por qualquer bactéria ou infecção viral. Pode ser fatal se não for tratada no momento certo e, portanto, o início o diagnóstico de pneumonia é vital. O objetivo do trabalho é detectar automaticamente pneumonias usando imagens de raios-x digitais. Três diferentes pré-treinados Rede Neural Convolucional profunda (CNN): AlexNet, ResNet50 e VGG16 foram usado para transferência de aprendizagem. Um total de 6.006 imagens de raios-X de tórax consistindo de pneumonia e normais. As imagens de raios-x de tórax foram pré-processadas e treinadas para a tarefa de classificação baseada em aprendizagem de transferência. A precisão da classificação máxima de imagens normais e de pneumonia, foi de 90% proveniente pela VGG16, pré-treinada pelo Keras Applications, com a adição de layers adicionais, que foram treinados.

II. INTRODUÇÃO

Pneumonia é uma categoria de infecção respiratória que pode atingir pessoas de todas as idades, especialmente crianças e idosos [Respira 2019]. Por ser uma das mais comuns infecções que geram necessidade de internação, casos dessa doença são considerados problemáticos [Rozenbauma et al. 2015], uma vez que está se enquadra na categoria de doenças que mais matam no mundo [Super Interessante, 2018]. Só no Brasil, em 2019 foram registrados 99.117 óbitos causados por pneumonia [aos fatos 2020], o que mostra o quão perigosa é a doença.

O diagnóstico geralmente é baseado nos sintomas e no exame físico, tendo o auxílio de radiografias do tórax para aumentar a assertividade do laudo. Desse modo, parte significativa do diagnóstico depende diretamente do treinamento médico dedicado a aprender a identificar as informações relevantes nesse momento e assim, minimizar complicações biológicas provenientes de um diagnóstico incorreto.

Já a opacidade em vidro fosco visto na radiografia do tórax, na Figura 1, se traduz pelo aumento do coeficiente de atenuação dos pulmões, mas sem apagar as marcas bronco vasculares é um dos principais indícios de pneumonia, bem como a diminuição da radio transparência pulmonar na região peri-hilar, de forma simétrica, poupando a periferia e configurando o aspecto em “asa de borboleta”, visto pela Figura 2.

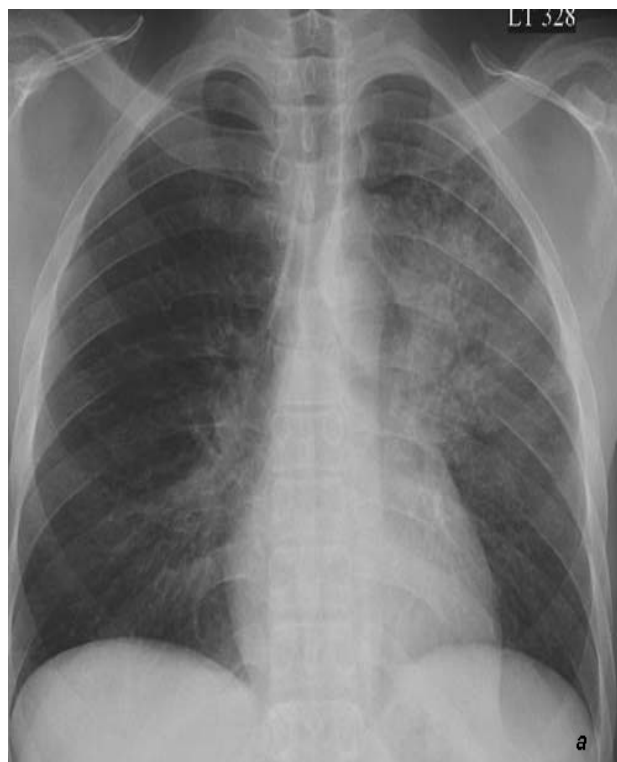


Figura 1. Exemplos de imagens raio-x

Tendo em vista o contexto supracitado, a problemática a ser investigada se refere à identificação de pneumonia a partir de imagens de raios-x dos pulmões de diferentes pacientes. Atualmente, sistemas computacionais de suporte à decisão clínica (SAD) vêm sendo amplamente utilizados para auxiliar os profissionais de saúde nessa análise.

De forma complementar, em meio à pandemia de COVID-19 sabe-se que a pneumonia é o quadro mais comum apresentado pelos pacientes infectados [Hospital Osvaldo Cruz, 2018]. Tendo em vista a grande quantidade de infectados seria inviável a avaliação individual do raio-x, por um médico, tendo em vista o aumento da demanda com o quadro pandêmico do COVID-19.



Figura 2. Exemplos de imagens raio-x

Desse modo, a utilização de Machine Learning pode ser uma alternativa à agilidade desse processo, bem como a garantia de diminuição de diagnósticos errôneos.

III. TRABALHOS RELACIONADOS

O referido estudo tem como insumo trabalhos anteriores, incorporando técnicas selecionadas para incorporação no algoritmo. Dessa maneira, a critério de enriquecer e maximizar o potencial preditivo, selecionou-se três artigos disponibilizados pela Applied Science e kaggle, respectivamente.

Em ambos, o conjunto inicialmente foi treinado com um CNN AlexNet e aprimorado com rede neurais mais estruturadas, além de adição de layers para um melhor ajuste.

Segundo o artigo publicado na Applied Science, utilizou-se quatro diferentes técnicas de pré-treinamento de Rede Neural Convolutiva profunda (CNN): AlexNet, ResNet18, DenseNet201 e SqueezeNet foram usado para transferência de aprendizagem. Diferente da problemática sugerida no referido trabalho, o estudo incorporava a classificação de pneumonia distinguida pela causa: bacteriana ou viral. Dessa forma, a precisão da classificação de imagens normais e de pneumonia, imagens de pneumonia bacteriana e viral e normal, bacteriana, e pneumonia viral foram 98%, 95% e 93.3%, respectivamente.

De forma complementar, os posteriores artigos serviram como insumo à implementação, uma vez que seus algoritmos eram disponibilizados, assim, auxiliando a incorporação desses a composição do estudo. Não obstante, o artigo inicial

elaborado por Salman Ibne Eunus, utilizou da uma única técnica de Rede Neural Convolutiva, AlexNet com três layers auxiliares, resultando em uma acurácia de 84.8%, o respectivo serviu como insumo para a estrutura do baseline.

Por fim, o trabalho elaborado e disponibilizado por Gaurav Rajpal, no Kaggle, fundamenta sua implementação em uma Rede Neural Convolutiva (VGG16) com a incorporação de layers que atingiu uma acurácia de 99.52%, sendo esta a maior predição atingida entre os três trabalhos relacionados e, portanto, exclusivamente utilizada para avaliação final do referido estudo.

IV. METODOLOGIA

A. Conjunto de dados

O conjunto de dados original consiste em duas pastas principais (isto é, pastas de treinamento e teste) e duas subpastas contendo imagens de raios-X de tórax de pneumonia (P) e normal (N), respectivamente. Um total de 6.006 imagens de raios-X de tórax ântero-posterior foram cuidadosamente escolhidas de pacientes. Para equilibrar a proporção de dados atribuídos ao conjunto de treinamento e validação, a categoria de dados original foi modificada. Reorganizou-se todos os dados em um conjunto de treinamento, validação e teste. Um total de 5.232 imagens foram atribuídas ao conjunto de treinamento, 624 ao conjunto de validação e 523 ao conjunto de teste.

B. Pré-processamento

Inicialmente dividimos as imagens classificando-as como normais (N) e com pneumonia (P), em conjunto de treino, validação e teste. Precedente a aplicação do nosso método proposto, realizamos uma tratativa nas imagens, mediante a incorporação de um filtro que distinguiu os padrões por escala de cor, facilitando a detecção de padrões pelo modelo. É importante notar que este filtro utiliza na sua escala as cores RGB, um exemplo do pré-processamento pode ser visto na Figura 3.

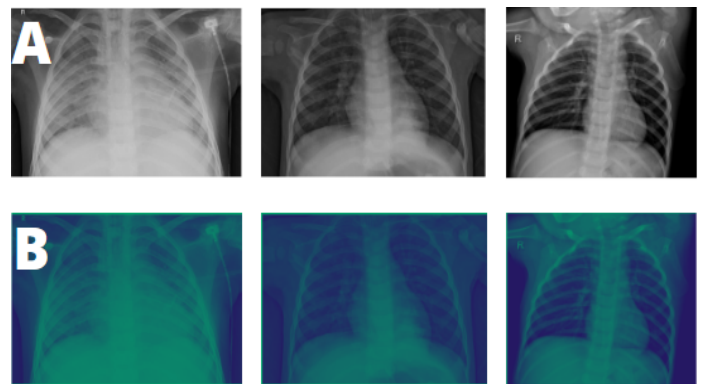


Figura 3. Imagem original (A), Imagem com o filtro RGB (B)

C. Matriz de desempenho para classificação

Considerando a estrutura supracitada, implementou-se uma rede CNN, com camadas de Convolução (utilizando Relu), seguidas por Max Pooling, terminando com uma camada Flatten e duas camadas Dense. A ilustração da estrutura de Rede Neural Convolucionadas, podem ser vistas pela Figura 4.

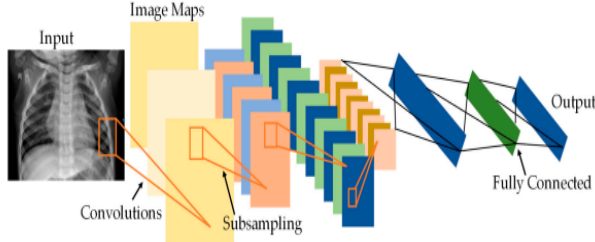


Figura 4. Arquitetura da Rede Neural Convolucionada (CNN)

Posteriormente, como forma de maximizar as medidas de predição, adotou-se redes neurais mais estruturadas e retirando o viés do desbalanceamento dos dados por meio da métrica AUC, que incorpora essa proporcionalidade. Dessa maneira, foi implementado dois modelos distintos e comparando seu poder preditivo, ResNet e VGG16, sua estrutura pode ser visualizada pela Figura 5.

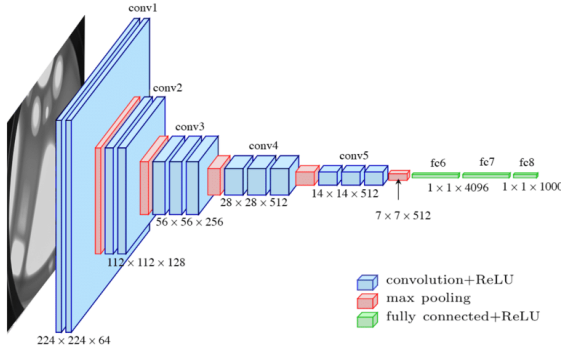


Figura 5. Arquitetura da Rede Neural Convolucionada (VGG16)

Em ambos casos, a predição foi inflacionada à favor de casos de pneumonia decorrente do desbalanceamento, mesmo com a métrica AUC como ponderação. Considerando esse viés e incorporando estudos pré-existentes, adicionou-se layers, resultando em um ganho na predição. A listagem abaixo mostram as seis métricas de desempenho:

$$\text{Acurácia} = \frac{(TP+TN)}{(TP+FN)+(FP+TN)} \quad (1)$$

$$\text{Sensitividade} = \frac{(TP)}{(TP+FN)} \quad (2)$$

$$\text{Especificidade} = \frac{(TN)}{(FP+TN)} \quad (3)$$

$$\text{Precisão} = \frac{(TP)}{(TN+FP)} \quad (4)$$

Neste estudo, a linguagem Python foi utilizada para treinar, avaliar e testar diferentes algoritmos. A Figura 6 ilustra a visão geral da metodologia deste estudo.

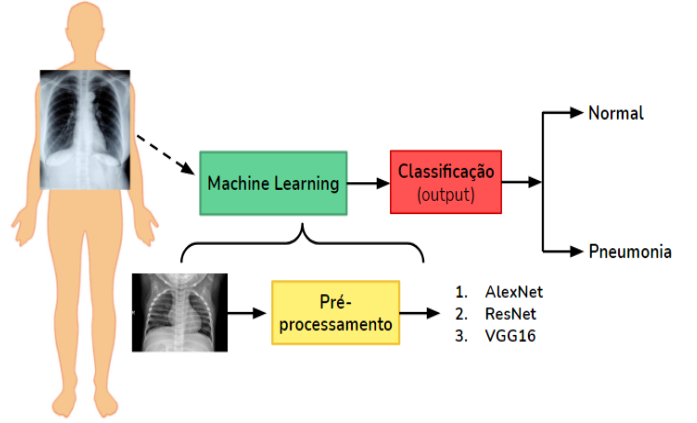


Figura 6. Visão geral da metodologia

V. EXPERIMENTOS

Primeiramente, é importante ressaltar que, nesse projeto, utilizamos como base a implementação de um trabalho sobre classificação de imagens (com pneumonia ou não) de raio-x, supracitado. Inicialmente dividimos as imagens classificadas como normais e com pneumonia em conjunto de treino e conjunto de validação. Antes de realizar a aplicação do método, utilizou-se um filtro nas imagens que altera a escala de cor, tornando as imagens mais legíveis pelo modelo. Vale ressaltar que o filtro utiliza a escala de cores RGB.

Alguns exemplos de como as imagens ficaram podem ser visualizados na Figura 7.

Após isso, normalizou-se os dados dividindo os valores de cada observação do banco por 255 (para que estes se encontrassem em um intervalo de 0 a 1, já que os valores de 1 a 255 representam uma escala de “cinza” para o pixel da imagem). Além disso, alterou-se de forma aleatória a altura e largura das imagens para o banco de dados original sofrer mudanças e assim, o modelo ter uma taxa de aprendizagem melhor. Dessa maneira, chegou-se a 5.232 imagens no banco de treino, 624 no banco de validação e 523 no banco de teste, sendo duas classes consideradas: normal e pneumonia. Portanto, decidiu-se por utilizar uma rede neural com convolução e funções de ativação para elaborar uma classificação inicial binária, entre normal e pneumonia.

O modelo implementado foi uma rede CNN, com camadas de Convolução seguidas por Max Pooling, terminando com uma camada Flatten e duas camadas Dense. O layer da

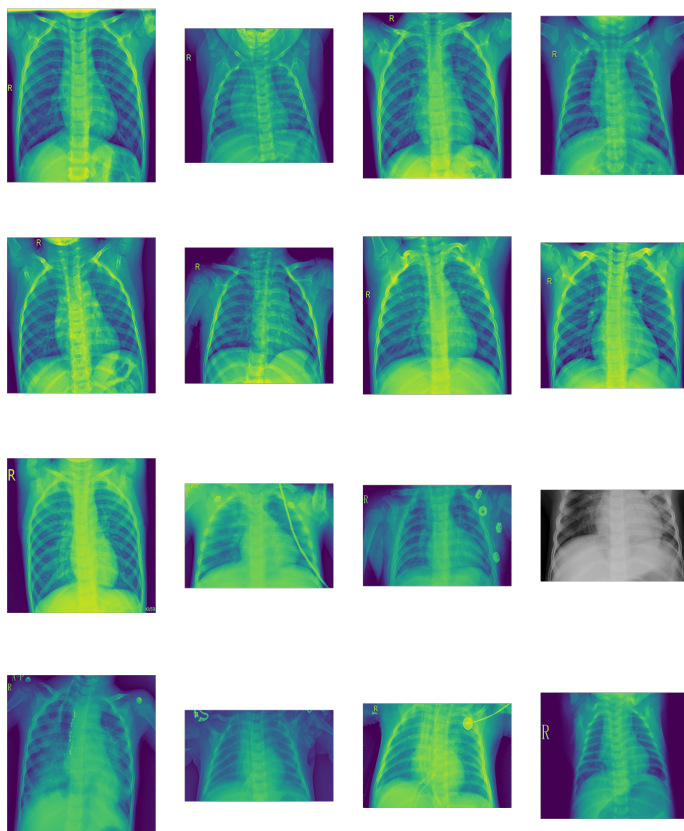


Figura 7. Exemplos de imagens raio-x com o filtro

Convolução aplica a matriz de filtro, enquanto o MaxPooling seleciona o valor máximo de cada matriz de valores. Já o Flatten realiza a transformação em um vetor de valores e o Dense é um layer comum.

Tendo isso feito os resultados obtidos não foram satisfatórios, pois a acurácia obtida ainda podia ser aumentada, considerando os trabalhos relacionados. Com isso, escolhemos as redes neurais ResNet50 e VGG16 para uma maior robustez do algoritmo de classificação.

Considerando a ResNet50 como modelo promissor devido a complexidade de suas camadas, observou-se um forte tendência à classificação de pneumonia, mesmo em casos errôneos. Tal problemática é semelhante ao conflito do desbalanceamento do conjunto de dados e portanto, o modelo não possui uma estrutura adequada ao estudo.

Mediante a problemática referenciada ao ajuste da ResNet50, considerou-se a rede VGG16, que obteve resultados melhores. Porém ainda foi localizada uma tendência em classificar as imagens em "pneumonia". Portanto, foram implementados layers adicionais, com base no trabalho de Gaurav Rajpal.

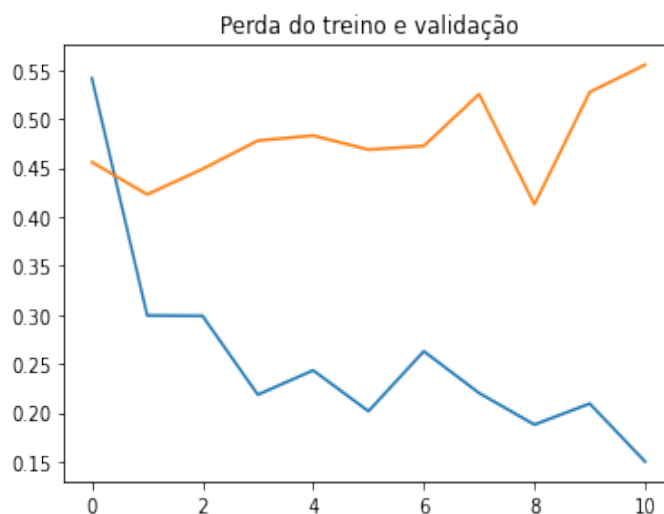


Figura 8. Perdas: treino e validação

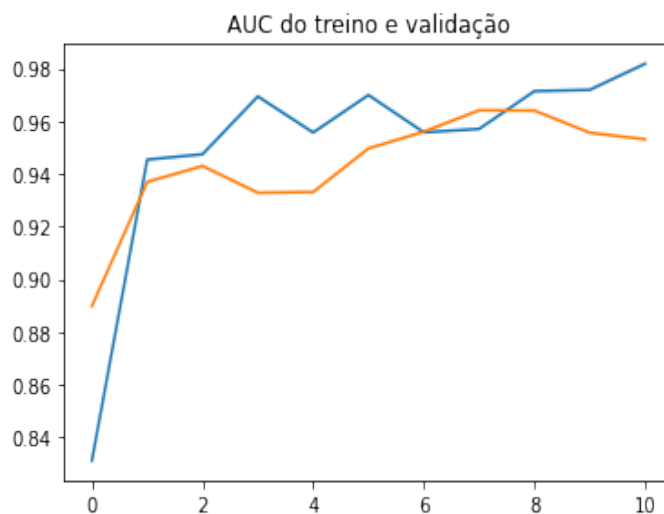


Figura 9. Acurácias balanceadas: treino e validação

Os resultados foram bem mais promissores. Nas figuras 8 e 9 podemos ver a acurácia balanceada e a perda nos conjuntos de dados de treino e validação. Próximo à época 6, de ambos os gráficos, já temos uma estabilização e chegamos ao modelo final.

Com o modelo final, obteve-se acurácias balanceadas (AUC), respectivamente, para treino e validação de 90.4% e 79.7% e enfim para o conjunto de teste, aproximadamente 90%. O resultado dos mesmos podem ser vistos pela Tabela I.

Não obstante, é possível observar uma proporção extremamente baixa de falsos negativos, ou seja, indivíduos que apresentam pneumonia e são classificados como normal. Estes

Tabela I
MATRIZ DE PREDIÇÃO

	Acurácia	AUC
Treino	89.5%	90.4%
Validação	82.7%	79.7%

casos compõem cerca de 1.3% de todas as classificações e os seus resultados podem ser avaliados pela Tabela II.

Tabela II
MATRIZ DE CONFUSÃO

Predição	Normal	Pneumonia
Normal	128	7
Pneumonia	59	329

Nas figuras 10, 11 pode-se observar imagens que foram classificadas corretamente como pneumonia e sem pneumonia, respectivamente.

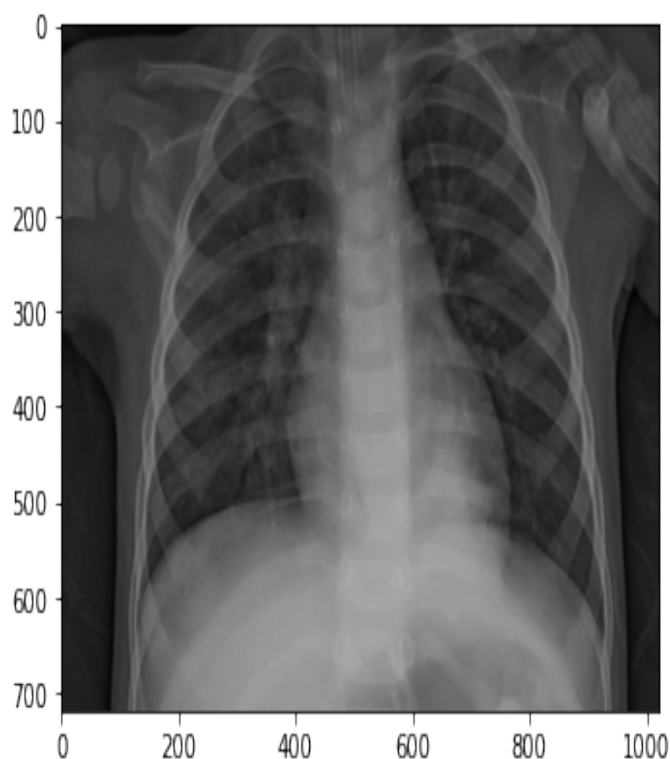


Figura 10. Paciente com pneumonia

Como dito anteriormente a opacidade em vidro fosco é percebida na imagem de um paciente diagnosticado com pneumonia. Se compararmos o paciente contaminado com o saudável podemos perceber nítidas a diferença em suas cavidades torácicas. O paciente com pneumonia tem uma imagem mais fosca com relação ao não contaminado e, provavelmente, a rede neural está captando essa diferença ao classificar os raio-x dos pacientes.

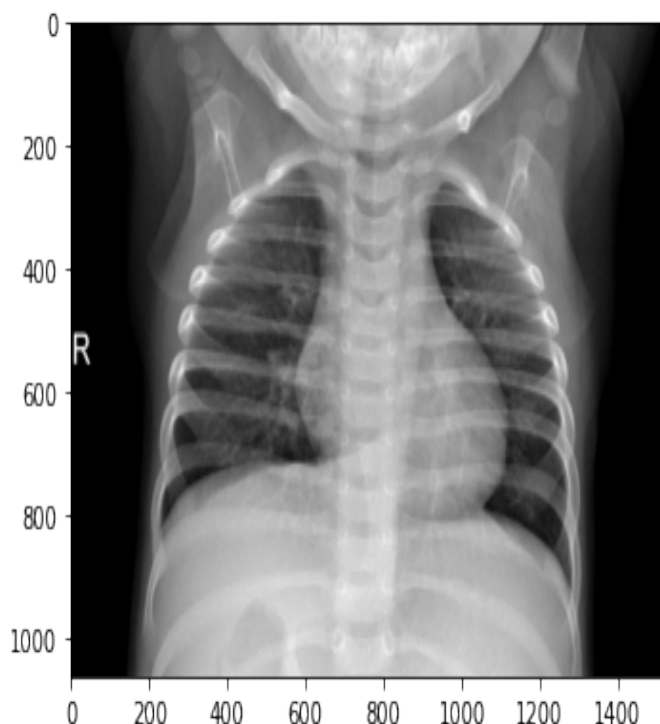


Figura 11. Paciente sem pneumonia

VI. CONCLUSÃO E FUTUROS TRABALHOS

Com os resultados apresentados anteriormente, podemos concluir que, das redes testadas, concluímos que a melhor é a VGG16, pois esta apresentou uma maior acurácia balanceada em todos os estágios (treino, validação e teste).

Como foi visto, na literatura existem diversos trabalhos que abordam classificação de imagens sobre pneumonia. No nosso trabalho, foi abordada uma classificação simples (binária), ou seja, avaliou-se imagens de raio-x e o modelo retornou uma resposta "1" ou "0" (com ou sem pneumonia), por exemplo.

Seria interessante para uma análise futura aumentar o tamanho do banco de dados para ver como o modelo se comportaria, bem como incorporar mais fatores de classificação, por exemplo, identificar dentre os casos classificados como pneumonia qual se este se tratava de bacteriana ou viral.

Entretanto, a segunda abordagem demandaria a seleção de um banco de dados que categorizasse de forma confiável o causador da pneumonia outra abordagem seria aplicar um modelo, para um banco de dados que possuísse dados sobre pacientes com raio-x que apresentam: Pneumonia viral, pneumonia bacteriana e sem pneumonia.

Tais complementos supracitados se destrincham na abordagem de detecção da pneumonia, entretanto, há grupos com maior potencial de risco, mediante doenças pré-existentes ou consumo de drogas lícitas, por exemplo, cigarro. Dessa maneira, seria relevante, clusterizar os pacientes mediante a esses fatores como forma preditiva de definir scores de risco

ou complicações futuras, assim, direcionando os casos que precisariam de pronto atendimento.

REFERÊNCIAS

- [1] Christopher M. Bishop. “Pattern Recognition and Machine Learning”. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.
- [2] Tawsifur Rahman, Muhammad E. H. Chowdhury, Amith Khandakar, Khandaker R. Islam, Khandaker F. Islam, Zaid B. Mahbub, Muhammad A. Kadir, Saad Kashem. “Transfer Learning with Deep Convolutional Neural Network (CNN) for Pneumonia Detection Using Chest X-ray”. Applied Sciences, MDPI, 2020.
- [3] Salman Ibne Eunus. “Pneumonia—X-Ray classification(Accuracy =84.8%)”. Kaggle, 2020.
- [4] Gaurav Rajpal. “Pneumonia Detection - VGG16(99.52%)”. Kaggle, 2020.
- [5] Disponível em: <https://www.aosfatos.org/noticias/registros-de-mortes-por-pneumonia-e-insuficiencia-respiratoria-nao-foram-substituidos-por-covid19-em-2020>. Acesso em 18/01/2021.