

Classificação de Pneumonia com CNN

Gabriel Luciano Gomes
265673
g265673@dac.unicamp.br

Gabriel Oliveira Campos
265146
g265146@dac.unicamp.br

Paulo Junio Reis Rodrigues
265674
p265674@dac.unicamp.br

Resumo—A pneumonia é uma doença que provoca inflamação nos sacos de ar dos pulmões e pode colocar em risco a vida humana se não for diagnosticada logo nos primeiros sintomas. Além disso, ela pode ser causada por vírus (pneumonia viral), por bactéria (pneumonia bacteriana) e fungos (pneumonia fúngica). Sendo assim, o objetivo deste trabalho é detectar se um paciente possui essa enfermidade, e caso haja, identificar qual é o causador (vírus ou bactéria), a fim de auxiliar no tratamento e, potencialmente, reduzir o quadro de óbitos. Para tal, neste projeto serão utilizadas imagens de radiografias pulmonares de uma base de dados com contribuições de diferentes fontes. Por fim, é apresentado algumas das técnicas já exploradas para resolução deste problema e realizado um novo estudo, com os modelos *DenseNet201* e *InceptionV3*, a fim de avaliar o desempenho em relação aos outros já existentes. Consequentemente, será feito o levantamentos sobre o material estudando, apontando quais foram os resultados obtidos e quais as limitações identificadas.

I. INTRODUÇÃO

No ano de 2018, entre Janeiro e Agosto, foram registrados cerca de 417 mil pacientes hospitalizados por pneumonia no Brasil. Por este motivo, mais de R\$ 378 milhões foram gastos em serviços médicos [1]. Pouco depois, no início 2019, cerca de 800 mil crianças, menores de 5 anos, morreram pela mesma enfermidade. Além disso, cerca de 9 milhões de menores poderão morrer em uma década, caso ações não sejam adotadas, diz a Unicef [2].

Diante do atual cenário de pandemia, os problemas respiratórios se tornaram foco em diversas áreas de pesquisa, para pacientes de qualquer faixa etária. Isso porque a COVID-19 tem características semelhante a pneumonia, pois o alojamento final da doença é no pulmão e pode possibilitar a enfermidade em questão. Tendo isso em vista, qualquer diferença presente neste órgão, comparado a um em estado saudável, é passivo de apresentar a doença. Ademais, a pneumonia pode ser causada por vírus ou bactéria, o que dificulta ainda mais sua identificação.

O objetivo deste trabalho é identificar, por meio de análise de raio x pulmonar, a presença de pneumonia nos pacientes. Junto a isso, caso positivo, identificar qual o motivo desta doença, seja ela causada por vírus ou por bactéria. Com isso, este trabalho está dividido nas seguintes seções: i) descrição da base que será utilizada no projeto; ii) modelo baseline utilizado e resultados obtidos por meio dele; iii) proposta de um novo modelo para melhoria e seu respectivo resultado; Por fim, é apresentado os resultados e limitações encontradas neste projeto.

II. REVISÃO BIBLIOGRÁFICA

Mediante o agravamento de doenças crônicas, diversos autores do estado da arte tem apresentado soluções para facilitar a identificação e até mesmo auxiliar no tratamento do paciente. Junto a isso, é realizado uma revisão bibliográfica para identificar o que já foi feito em relação ao tema apresentado.

Em [3], os autores realizam um estudo para identificação de pneumonia em pacientes, por meio da aplicação de uma *Convolutional Neural Network* (CNN) utilizando radiografias pulmonares. Junto a isso, foi explorado a classificação entre pneumonia viral e bacteriana no projeto, utilizando as redes AlexNet, ResNet18, DenseNet201 e SqueezeNet. Por fim, foi identificado que a rede DenseNet201 apresentou melhores resultados e obteve 93% na acurácia para predição de pneumonia bacteriana ou viral. Seguindo esta ideia, em [4] os autores exploram o desempenho de outras redes aplicado a este problema. Nele são analisados os modelos *ResNet50*, *VGG-16* e *InceptionV3*, onde o *ResNet50* apresentou os melhores resultados, com o *f1-score* de 92.82%.

Já em [5], é realizado um estudo em larga escala que foca no desempenho de generalização dos modelos atuais. Para isto, é analisado os resultados obtidos quando utilizado fontes de instituições distintas — neste trabalho foram utilizadas radiografias de 7 bases de dados diferentes — que foram rotuladas por médicos ou outras ferramentas de classificação. Para os modelos de predição, foi utilizado a DenseNet (não especificado qual foi o modelo específico) e, aliado a isso, foi realizado o *data augmentation* para obter maiores generalizações para o modelo. Por fim, foi analisado que as melhores generalizações foram identificadas para trabalhos de análise de cardiomegalia, edema e efusão. Ademais, o trabalho também apresenta resultados satisfatórios para reconhecimento de atelectasia, consolidação, enfisema, hérnia e opacidade pulmonar.

III. METODOLOGIA

Para este projeto, será utilizado a base [6] dos autores *Joseph Paul Cohen, Paul Morrison and Lan Dao*, encontrada na plataforma **Kaggle**. A base é constituída por 5.910 imagens (JPEG) radiográficas da parte torácica dos pacientes. O conteúdo está categorizadas em dois grupos, pulmões saudáveis (1576 imagens) e pulmões com presença de pneumonia (4334 imagens). Este grupo está subdividido em outros grupos menores, representados na tabela I.

A base está dividida em duas partes. A primeira delas, é a seção de treinamento, onde se encontra 5.286 imagens,

Causa	Imagens
Fumo por estresse	2
Vírus (sem categoria)	1493
Vírus (COVID-19)	58
Vírus (SARS)	4
Bactéria (sem categoria)	2772
Bactéria (Streptococcus)	5

Tabela I: Divisão da base de dados.

partilhadas em dois grupos: 1.342 figuras de pulmões sem nenhuma pneumonia e 3.944 figuras de pulmões com pneumonia. Já a segunda parte é utilizada para a fase de testes, que é constituída por 624 imagens. Sendo assim, a base de treino será parcialmente dividida para formar um novo conjunto, o de validação. Consequentemente, 80% destes dados formarão o conjunto de treino e os 20% restantes serão utilizados para a fase de validação.

Todas as imagens foram retiradas de fontes públicas, bem como coleta indireta de hospitais e médicos.

A. Pré-processamento

Antes do processamento da base, será necessário verificar como identificar a classe de cada um dos elementos nela contidos. Para isso, um arquivo de metadados (CSV), que acompanha os arquivos descarregados, será analisado para auxiliar nesta etapa. Este documento é responsável por descrever os seguintes campos:

- **X_ray_image_name**: nome da imagem
- **Label**: identificação de presença ou não de Pneumonia ('Normal' ou 'Pneumonia')
- **Dataset_type**: informação de qual conjunto a imagem pertence ('TRAIN' ou 'TEST')
- **Label_2_Virus_category**: informação do vírus detectado ('NaN', 'Streptococcus', 'COVID-19', 'ARDS' ou 'SARS')
- **Label_1_Virus_category**: informação de como a doença foi contraída ('NaN', 'Virus', 'bacteria', 'Stress-Smoking')

Após esta análise, foi observado a presença de campos nulos nesta tabela, que podem impactar no resultados obtidos. Sendo assim, todos estes valores foram convertidos em entradas desconhecidas ('unknown'). A tabela II ilustra o procedimento descrito.

Coluna	Valores nulos
X_ray_image_name	0
Label	0
Dataset_type	0
Label_2_Virus_category	5841
Label_1_Virus_category	1576

Tabela II: Número de ocorrência de valores nulos.

Uma outra medida adotada foi o balanceamento da base de dados. Isto porque a quantidade de pneumonia bacteriana no conjunto de treino (2027) é maior do que a quantidade de pneumonia (1126) viral. Sendo assim, 911 radiologias da doença provocada por bactéria foram descartadas.

Coluna Label_1_Virus_category		Coluna Label_2_Virus_category	
Campos	Nº de entradas	Campos	Nº de entradas
bacteria	2777	unknown	5841
unknown	1576	COVID-19	58
Virus	1555	Streptococcus	5
Stress-Smoking	2	SARS	4
		ARDS	2

Tabela III: Dados após remoção de campos nulos.

B. DenseNet201

A princípio, o modelo RestNet50 seria utilizado para processamento dos dados. Entretanto, os resultados obtidos não foram muito convincentes; uma grande frequência de *overfitting* foi identificada, bem como acurácia não superior a 80%. Sendo assim, como recomendado pelos autores da base de dados, o modelo DenseNet201 foi escolhido para dar continuidade neste trabalho, uma vez que apresentou resultados satisfatórios, que serão descritos posteriormente.

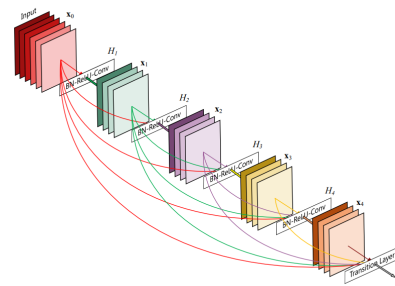


Figura 1: Bloco denso com 5 camadas e taxa de crescimento $k = 4$. Cada camada considera os mapas de características anteriores como entrada.

A DenseNet201 é uma *Dense Convolutional Network* (DenseNet) [7] de 201 camadas. Ela melhora o fluxo de informação e gradientes pela rede, fazendo com que a otimização de redes muito profundas seja possível mais facilmente. Ainda assim, para este projeto, a camada *fully connected* final desta rede foi substituída por uma outra que gera apenas uma saída, na qual foi utilizada uma função de ativação sigmoide, visto que se trata de uma classificação binária. Por fim, os pesos originais foram mantidos e incluída uma nova camada de *Max-Pooling* global, resultando no modelo presente na figura 2.

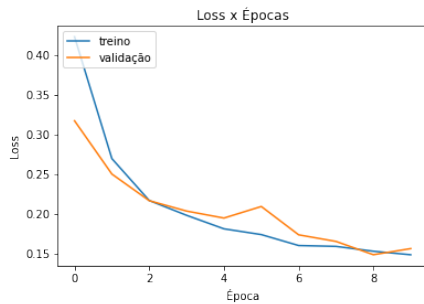
Layer (type)	Output Shape	Param #
densenet121 (Functional)	(None, 8, 8, 1024)	7037504
global_average_pooling2d_2 ((None, 1024)		0
dense_7 (Dense)	(None, 1)	1025
Total params: 7,038,529		
Trainable params: 1,025		
Non-trainable params: 7,037,504		

Figura 2: Modelo baseline resultante.

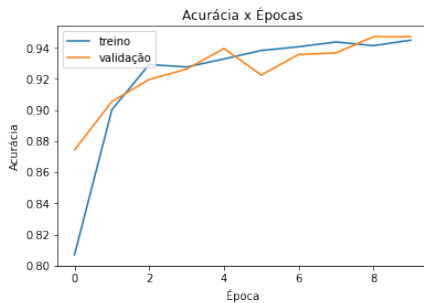
C. Considerações Iniciais

Foram analisadas duas execuções para o modelo resultante. A primeira delas diz respeito a identificação de pneumonia. Já a segunda, estuda como a enfermidade foi contraída (vírus ou bactérias). Ambas estão descritas nas seções abaixo. Ainda assim, foi utilizado o otimizador Adam, com todos os parâmetros padrões, pesos da ImageNet, *batch size* de 64 e analisado a métrica de acurácia.

1) *Identificação de Pneumonia*: Os resultados para esta seção estão representados na figura 3. Junto a isso, foi observado que os valores obtidos foram bastante satisfatórios, com taxa de acerto próximo a 94% e ausência de *overfitting*. Porém, o problema em questão não é complicado de ser resolvido, que por sua vez, não agrega resultados chave para o projeto.



(a) Loss por época.



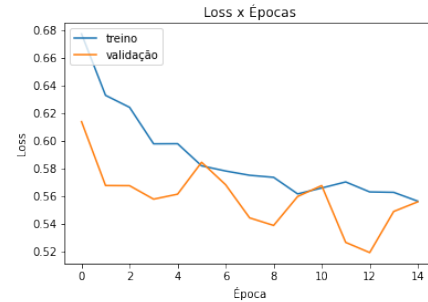
(b) Acurácia por época.

Figura 3: Resultados encontrados para identificação da doença.

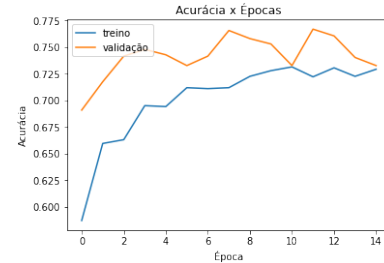
2) *Contração da Enfermidade*: Os resultados para esta seção estão representados na figura 4. Pode-se observar que a acurácia tem taxa de acerto próxima a 73%. Ademais, uma tentativa realizada a fim de otimizar esse resultado, foi aplicar o *fine-tuning* na rede, que permite a re-treinagem de todos os pesos da DenseNet, considerando a base de dados informada nesse projeto. Todavia, após este procedimento, pôde-se observar na figura 5, que os resultados pioraram em relação ao modelo anterior, além de apresentar *overfitting*.

D. Solução proposta

A fim de aumentar os resultados obtidos no modelo *DenseNet201*, um outro modelo será avaliado para o mesmo problema. Após o estudo de outros métodos, o modelo *InceptionV3* será utilizado para investigação. Outro fato que influenciou nesta escolha, foi o fato de estar presente em

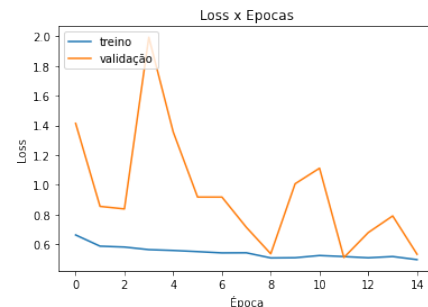


(a) Loss por época.

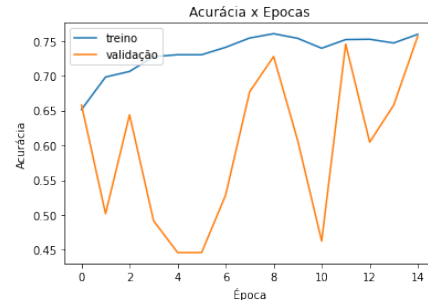


(b) Acurácia por época.

Figura 4: Resultados encontrados para identificação do tipo da doença sem *fine-tuning*.



(a) Loss por época.



(b) Acurácia por época.

Figura 5: Resultados encontrados para identificação do tipo da doença com *fine-tuning*.

diversos trabalhos encontrados no local de publicação da base de dados.

1) *InceptionV3*: O *InceptionV3* [8] é terceira versão uma CNN que é utilizada para identificação de objetos. A princípio, era parte de um módulo da *GoogLeNet*, que foi difundido e originou a rede *Inception*. Sua melhor característica é fazer a extração de características com a menor quantidade de parâmetros possível. Além disso, apresenta acurácia superior a 78% no conjunto de imagens da *ImageNet*.

Ele é composto por módulos chamados *Inception* que apresentam diferentes tipos camadas em cada um deles, sendo assim componentes simétricos e assimétricos. Por fim, a normalização é fortemente utilizada em todo o modelo e também aplicada às entradas de ativação e, calcula a perda por meio da função *softmax*. A figura 6 ilustra o modelo.

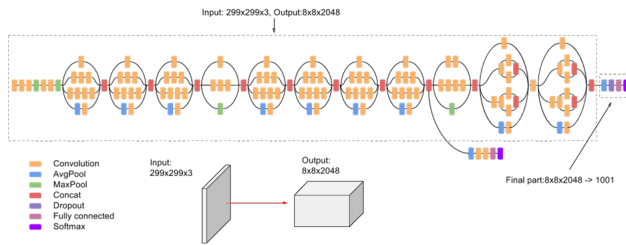


Figura 6: Diagrama geral do modelo *InceptionV3*.

2) *Early Stopping*: O *Early Stopping* é uma técnica de aprendizado de máquina que realiza a parada do treinamento antes que ocorra um *overfitting*, ilustrado na figura 7. Ele é utilizado para evitar que o modelo adote ajustes excessivos e prejudique o aprendizado já realizado. Ele será aplicado no *fine-tuning* do modelo *InceptionV3*, para que não grande perda no conhecimento da rede original. Ainda assim, ele observa durante um número x de iterações (será utilizado $x = 3$), para visualizar se realmente é devera fazer a parada, chamado *patience*.

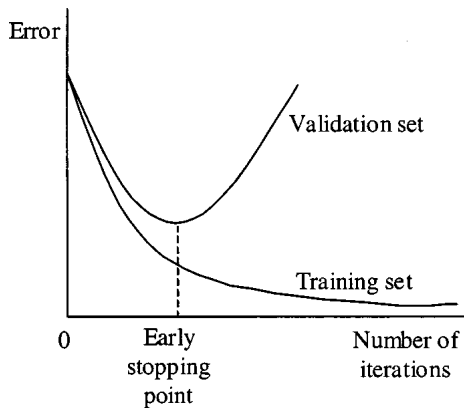


Figura 7: Ponto de *early stopping* [9].

3) *Redução do Learning Rate*: Por fim, é utilizado a redução da taxa de aprendizado (*Learning Rate Reduction*). Esta técnica diz respeito à quantidade de pesos que serão atu-

alizados durante o processo de treinamento de rede. Ela é responsável por controlar o ajuste desnecessário dos parâmetros, que impede que o modelo assuma predições que prejudicam o resultado final do treino. Consequentemente, também será aplicada no *fine-tuning* para que o mínimo possível de erros venham a acontecer. A imagem 8 ilustra o impacto dos diferentes *learn-rates* aplicado no modelo.

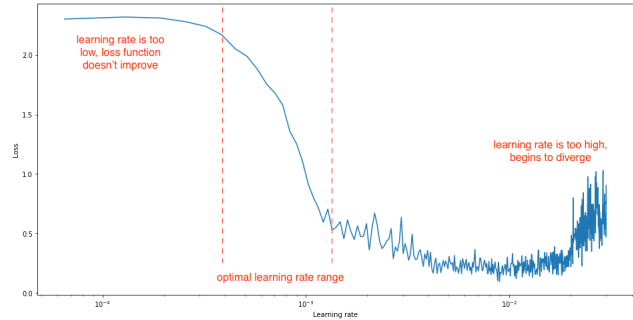


Figura 8: Impactos de diferentes *learn-rates* na rede [10].

IV. EXPERIMENTOS E DISCUSSÕES

Os experimentos deste trabalho são divididos em duas partes, considerando a rede *InceptionV3*; Primeiro será avaliado o desempenho para identificação de Pneumonia; Em seguida será avaliado o seu desempenho para classificação da pneumonia, sem e com o *fine-tuning*. Os resultados estão disponíveis nas subseções seguintes.

A. Identificação de Pneumonia

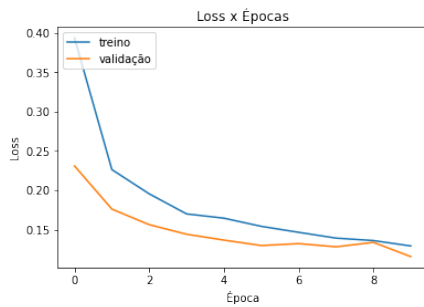
Ao realizar o treinamento da rede para identificação de Pneumonia, foi observado que apresentou resultados pouco melhores aos adquiridos pela *DenseNet* — acurácia de 95.29% vs 94.51% no treino e 95.18% vs 93.86% na validação. Isto demonstra um grande potencial para as próximas etapas, uma vez que já na identificação da doença obteve resultados superiores aos já obtidos pelo outro modelo. Sendo assim, é possível prosseguir com os outros processos a fim de adquirir resultados superiores aqueles previamente obtidos. A figura 9 ilustra quais foram os resultados obtidos para este processo.

B. Classificação de Pneumonia

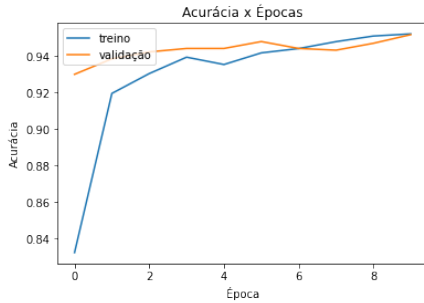
A fim de avaliar corretamente a classificação feita pela rede, a partir destes testes também será incluso a acurácia balanceada. Ela é responsável por informar o resultado final, considerando tanto os acertos quanto os erros. Sendo assim, foi possível adquirir os seguintes resultados da rede.

1) *Sem fine-tuning*: Como primeira tentativa, foi utilizado o *transfer learning* considerando o treinamento já existente da rede. Para isto, foi removido a camada superior para inclusão de uma camada densa com o número de classes do problema (como se trata de um problema binário, apenas 1) e congelada todas as demais. Com isso, é possível avaliar o desempenho da técnica para o problema em questão.

Como foi aplicado o *Early stopping* junto ao *Learn Rate Reduction*, é possível identificar na imagem 10 que o treino para



(a) Loss por época.



(b) Acurácia por época.

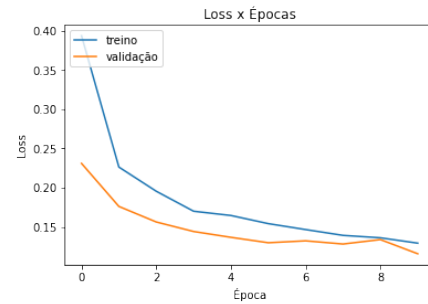
Figura 9: Resultados encontrados para identificação da doença.

antes da quantidade de épocas estabelecida (originalmente 30). Isto porque os pesos que estavam sendo ajustados começaram a ser configurados erroneamente, o que levou a parada do treinamento. Todavia, apresentou resultados levemente melhor ao da *DenseNet* — acurácia de 72.87% vs 73.76% no treino e 73.38% vs 71.36% na validação — mas, ao analisar a acurácia balanceada, os resultados não foram tão satisfatórios (50.44% treino e 50.81% 0.5081 validação). Ao que indica que, o modelo está acertando classificações de apenas uma classe.

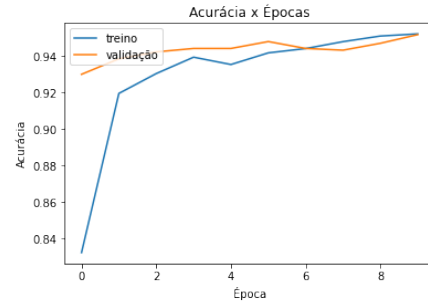
2) Com *fine-tuning*: A fim de melhorar os resultados obtidos, foi realizado o *fine-tuning* da rede. Para isto, após o treinamento anterior, todas as camadas foram descongeladas e submetidas ao mesmo treino. Esta técnica possibilita a extração de mais conhecimentos do conjunto de dados e possivelmente um melhor resultado. Apesar disso, ela se torna mais vulnerável a atualização de pesos erroneamente, consequentemente a taxa de aprendizado foi reduzida de 0.001 para 0.00001 e a quantidade de épocas de 30 para 15. Por tal motivo, não ocorreu o *Early stopping* no treinamento e os resultados aumentaram, que pode se observado na figura 11 ;obteve-se aumento na acurácia de 7.33% no treino e de 3.17% na validação, o que não foi muito relevante. Entretanto a acurácia balanceada reduziu ainda mais, resultando em 49.38% para o treino e 48.99% para validação.

C. Resultados Finais

Considerando apenas a acurácia, o modelo *InceptionV3* com *Fine-tuning* apresentou melhor resultados em relação à rede *DenseNet201*. Por este motivo, a predição com a base de testes utilizou essa rede e obteve acurácia de 88.21% e

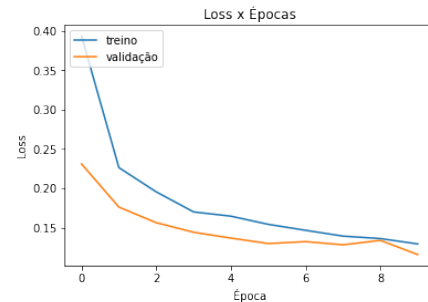


(a) Loss por época.

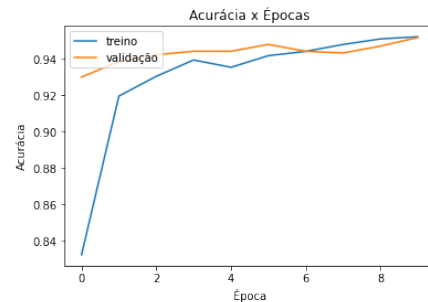


(b) Acurácia por época.

Figura 10: Resultados encontrados para classificação da doença.



(a) Loss por época.



(b) Acurácia por época.

Figura 11: Resultados encontrados para classificação da doença com *fine-tuning*.

acurácia balanceada de 56.84%. A fim de entender melhor o comportamento da acurácia balanceada, uma matriz de confusão foi elaborada, ilustrada na figura 12.

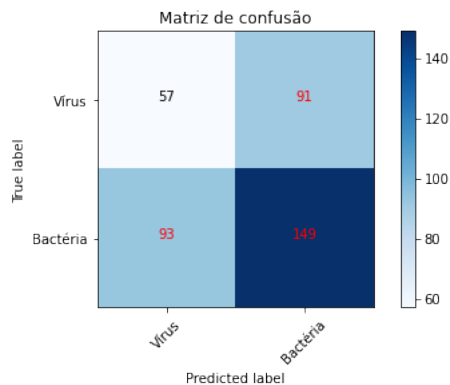


Figura 12: Matriz de confusão para os resultados da base de teste.

Após a construção visual da matriz, pode-se observar que o modelo apresenta grande acerto para a classe de Bactéria. Este fator justifica a acurácia encontrada. Entretanto, é possível visualizar que apresenta dificuldade para a classificação de vírus e a forte tendência a classificar as amostras como sendo da classe bacteriana. Sendo assim, é possível concluir que o *InceptionV3* apresentou desempenho favorável para classificação de pneumonia bacteriana.

V. CONCLUSÕES E TRABALHOS FUTUROS

Pneumonia é uma doença que continua matando milhares de pessoas por ano e quanto antes for identificada, mais fácil será o seu tratamento. Com isso, neste trabalho foi explorado o uso de técnicas de *Deep Learning* a fim de identificar pneumonia por meio de radiografias pulmonares. Por este motivo, o grupo investigou e adquiriu conhecimentos de aplicações que envolvem conhecimento de redes neurais convolucionais. Consequentemente, observou a vasta pesquisa no estado da arte a fim de melhorias acerca do problema abordado, apontando, como exemplo, o forte uso da rede *DenseNet* e suas variantes, como a *CheXNet*.

Em vista dos resultados abordados no trabalho, pode-se concluir que o primeiro problema apresentado, identificação de pneumonia ou não por meio de raio-x, se tornou um problema "trivial" e acurácias acima de 90% são facilmente adquiridas com poucos ajustes em redes já existentes. Entretanto, a identificação do microrganismo que levou à doença ainda é um desafio que apresenta resultados muito variantes na literatura. A rede *DenseNet* garante bons resultados para isto, porém não foi possível obter resultados acima de 75% com esta técnica, além de apresentar tendência ou até mesmo sinais de *overfitting*. Por este motivo, foi explorado outra rede para obtenção de resultados, a rede *InceptionV3*. Todavia, os parâmetros ajustados ou até mesmo as camadas utilizadas para estudar o problema não foram adequados o suficiente para adquirir resultados superiores àqueles obtidos por outros

modelos. Isto porque, apesar de adquirir acurácia de 88% na base de teste, apresentou uma classificação com fortes erros para a classe de vírus, que pode ser observado com a matriz de confusão.

Com isso, como trabalhos futuros pode-se destacar a exploração de camadas mais adequadas a serem incluídas no modelo *InceptionV3*, para adquirir uma maior extração de conhecimentos para melhor classificar as amostras. Um outro fator importante a ser explorado, é a concatenação com outras bases de dados a fim de obter mais amostras para o treinamento, visto que na base utilizada neste projeto, existe grande desbalanceamento entre as classes. Consequentemente, subclassificações poderão ser possíveis, identificando, por exemplo, pneumonia provocada por síndrome respiratória aguda grave (SARS) ou pela COVID-19, em caso de pneumonia viral, trazendo maiores contribuições para a comunidade acadêmica.

REFERÊNCIAS

- [1] B. V. em Saúde, "Dia mundial da pneumonia." <http://bvsmis.saude.gov.br/ultimas-noticias/3085-12-11-dia-mundial-da-pneumonia>, 2019. [Online; acessado em 10-Novembro-2020].
- [2] Unicef, "Nove milhões de crianças podem morrer em uma década, a menos que o mundo aja contra a pneumonia, alertam agências." <http://encurtador.com.br/bloyX>, 2020. [Online; acessado em 10-Outubro-2020].
- [3] T. Rahman, M. E. H. Chowdhury, A. Khandakar, K. R. Islam, K. F. Islam, Z. B. Mahbub, M. A. Kadir, and S. Kashem, "Transfer learning with deep convolutional neural network (cnn) for pneumonia detection using chest x-ray," *Applied Sciences*, vol. 10, p. 3233, May 2020.
- [4] L. Silva, L. Araujo, V. Souza, A. Santos, and R. Neto, "Redes neurais convolucionais aplicadas na detecção de pneumonia através de imagens de raio-x," pp. 1–8, 01 2020.
- [5] J. P. Cohen, M. Hashir, R. Brooks, and H. Bertrand, "On the limits of cross-domain generalization in automated x-ray prediction," *arXiv preprint arXiv:2002.02497*, 2020.
- [6] J. P. Cohen, P. Morrison, and L. Dao, "Covid-19 image data collection," *arXiv 2003.11597*, 2020.
- [7] G. Huang, Z. Liu, and K. Q. Weinberger, "Densely connected convolutional networks," *CoRR*, vol. abs/1608.06993, 2016.
- [8] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," 2015.
- [9] R. Gençay and M. Qi, "Pricing and hedging derivative securities with neural networks: Bayesian regularization, early stopping, and bagging," *Neural Networks, IEEE Transactions on*, vol. 12, pp. 726 – 734, 08 2001.
- [10] J. Jordan, "Setting the learning rate of your neural network.." <https://www.jeremyjordan.me/nn-learning-rate/>, 2018. Acesso: 19-01-2021.