

Segmentação e Data Augmentation em Imagens de Raio-X de Pacientes com Tuberculose ou COVID-19.

Alexandre Ferreira Oliveira
Programa de Engenharia Elétrica
COPPE – UFRJ
Rio de Janeiro, Brasil
email:alexandre.oliveira@coppe.ufrj.br

Marcos Vinícius G. Rocha
Programa de Engenharia Elétrica
COPPE – UFRJ
Rio de Janeiro, Brasil
email : : marcosv.jf@gmail.com

José Manoel Seixas
Programa de Engenharia Elétrica
COPPE – UFRJ
Rio de Janeiro, Brasil
email:seixas@lps.ufrj.br

Abstract—O emprego da Inteligência Artificial (IA) tem ganhado importância nas mais diversas áreas da ciência. A tuberculose é uma doença transmissível que está entre uma das 10 principais causas de morte em todo o mundo. O SARS Cov-2 é um tipo de coronavírus que pode atingir diferentes partes do corpo humano, que fez com que a OMS decretasse o estágio de pandemia, com severos impactos em escala global. Este trabalho tem por objetivo empregar redes neurais profundas a fim de avaliar o impacto da segmentação de pulmões e *data augmentation* na classificação de pacientes com tuberculose ou COVID-19, em base de dados de tamanho reduzido. Desenvolvemos modelos que foram capazes de identificar pacientes com tuberculose (acurácia de 97,7%), ou COVID (acurácia de 99,4%), a partir de *data augmentation* sugerindo a viabilidade do emprego no diagnóstico assistido por computador.

Keywords—Tuberculose, COVID-19, CNN, Inception, MLP, Data Augmentation, Segmentação, Triagem, Diagnóstico.

I. INTRODUÇÃO

O emprego de redes neurais na predição de resultados em medicina e suporte a diagnósticos encontra grande amparo científico e vem se ampliando continuamente [1] [2], e os avanços recentes no aprendizado profundo levaram a um desempenho promissor em muitas tarefas de análise de imagens médicas [3]. Neste contexto, o presente trabalho busca tratar de aspectos relacionados a segmentação e *data augmentation*, em aplicações de aprendizado profundo, na área de saúde, mais especificamente relacionados a análise de imagens de raio-x da região do pulmão de pacientes com tuberculose ou COVID-19.

A tuberculose (TB) é uma doença transmissível que era a principal causa de morte por um único agente infeccioso (classificação acima de HIV/AIDS), segundo a Organização Mundial de Saúde (OMS), até a atual pandemia de coronavírus se estabelecer, respondendo por 1,4 milhão de mortes em 2019. TB é causada pelo bacilo *Mycobacterium tuberculosis*, que é propagado quando as pessoas que estão doentes com tuberculose expõem bactérias no ar, afetando principalmente os pulmões [4].

A rápida disseminação do novo coronavírus SARS-CoV2 [5] deu origem a uma doença designada por COVID-19 pela OMS, que apresenta um espectro clínico variando de infecções assintomáticas a quadros graves e elevada, em particular no Brasil, onde aproximadamente 20% dos casos detectados requerem atendimento hospitalar por apresentarem os pacientes dificuldade respiratória, dos quais aproximadamente 5% podem necessitar de suporte ventilatório [6].

O real impacto da COVID-19 é incerto ainda, entretanto, a doença representou uma grande demanda nos sistemas de saúde, comprometendo a oferta de recursos humanos e outros insumos essenciais para o funcionamento adequado dos serviços de saúde [7].

Neste estudo, usando técnicas de aprendizado de máquina (ML, do inglês, *machine learning*), aplicamos topologias de aprendizado profundo e redes neurais não profundas com os seguintes propósitos:

- Avaliar o impacto da segmentação de pulmões e *data augmentation* na classificação de pacientes com TB, em base de dados de tamanho reduzido ($<10^3$ imagens);
- Avaliar o impacto de *data augmentation* na classificação de pacientes com COVID-19, a partir de *transfer learning* de modelos treinados para TB, em base de dados de tamanho reduzido ($<10^3$ imagens).

Para a elaboração do trabalho utilizamos bases de dados públicas distintas de imagens de raio-X do tórax, sendo a primeira com imagens de pacientes sem TB e com TB [8] e a segunda com pacientes não infectados e infectados por COVID-19 [9].

II. TRABALHOS RELACIONADOS

Em trabalho recente, Stirenko et al. [10], a partir de uma base de dados disponibilizada pelo Hospital de Shenzhen, com 662 imagens da região do tórax de pacientes sem e com TB, propôs o uso de duas técnicas de *data augmentation*, designadas como

LossLess e Lossy, assim como a aplicação de segmentação dos pulmões, para o diagnóstico assistido por computador, especificamente para pequenas bases de dados ($<10^3$ imagens). Os autores apontam que uma melhor segmentação, *data augmentation*, estratificação de *dataset* e exclusão de outliers não evidentes poderiam ser o caminho para obter melhores resultados em conjuntos de dados pequenos e não balanceados.

Jaeger et al. [11] efetuaram a segmentação dos pulmões de pacientes sem TB e com TB, para posteriormente capturar características que permitiram que um classificador binário efetuasse a classificação entre normal ou TB. Eles utilizaram duas bases de dados, sendo a primeira de Montgomery Count e a segunda do Hospital de Shenzhen, obtendo respectivamente 78,3% e 84% de acurácia. Para a primeira base um radiologista obteve 82% de acurácia, indicando a viabilidade da solução proposta pelos autores.

Tartaglione et al. em [12] ressaltam que o raio-x de tórax (CXR) pode ser uma ferramenta útil no cenário de COVID-19, especialmente em ambientes de emergência, onde pode ajudar em uma primeira avaliação aproximada da extensão do envolvimento pulmonar, podendo ser obtido na cama do paciente usando dispositivos portáteis. Assim eles estudaram o uso de algumas topologias de redes profundas (ResNet-18, ResNet-50, COVID-Net, Dense-Net-121) na classificação de pacientes, a partir do pré-processamento de imagens composto por equalização, segmentação dos pulmões e normalização das imagens.

Estudo realizado por Tsiknakis et al. [13] focou na classificação de COVID-19 em exames CXR de rotina sem qualquer etapa de pré-processamento quanto a segmentação, utilizando a arquitetura Inceptionv3. Foram gerados mapas de atenção, que foram avaliados por especialistas e classificados entre 5 tipos (0-homogêneo em toda a imagem/1- focando em áreas totalmente irrelevantes fora do pulmão/2- focando nas áreas pulmonares, mas também em outras estruturas extrapulmonares/3-focando principalmente nas áreas pulmonares/4-focando exclusivamente na doença pulmonar), além de avaliados em termos de indicadores quantitativos. A principal novidade do modelo proposto encontra-se na classificação de COVID-19 contra casos comuns de pneumonia e não contra normais (saudáveis). A discriminação entre COVID-19 e as pneumonias, especialmente de origem viral, é intuitivamente mais complexa, dado os sinais e sintomas clínicos inespecíficos [14].

Os métodos de aprendizado de máquina oferecem uma grande promessa para a detecção e prognóstico rápidos e precisos da COVID-19, através de CXR e imagens de tomografia computadorizada (TC) de tórax [15], entretanto, os autores identificaram 2.212 estudos sobre COVID-19, dos quais 415 foram analisados e após revisão concluíram que nenhum dos modelos identificados são de uso clínico potencial devido a falhas metodológicas e/ou vieses subjacentes. Identificando esse como o grande ponto fraco de estudos de ML, dada a urgência com que modelos validados para COVID-19 são necessários.

Em [16] Ke et al. compararam o desempenho de 16 CNN populares (DenseNet121/169/201, EfficientNetB0/B1/B2/B3, InceptionV3/V4, MNASNet, MobileNetV2/V3, ResNet101/18/34/50) em um grande conjunto de dados de CXR

(CheXpert) e não encontraram relação entre modelos sem pré-treinamento e modelos com pré-treinamento na ImageNet quando aplicados a CheXpert. Identificaram também que modelos pré-treinados na ImageNet seriam demasiadamente grandes para a aplicação na CheXpert e que o truncamento de blocos finais de redes poderia melhorar a eficiência de parâmetros em 3,25 vezes. Os autores apontaram também que a Inceptionv3 foi a arquitetura sem pré-treinamento com o melhor resultado 0,866 (0,851-0,880) AUC.

Segundo Zhuang et al. [17] o *transfer learning* visa melhorar o treinamento de uma nova rede a partir de domínios correlatos, transferindo o conhecimento contido em domínios de origem diferentes, mas relacionados. Adicionalmente Yosinski et al. [18] identificaram que a inicialização de modelos com recursos transferidos pode melhorar o desempenho de generalização fino substancialmente em uma nova tarefa, o que pode ser uma técnica útil para melhorar o desempenho da rede neural profunda.

III. MÉTODO DE PESQUISA

A. Base de Dados

Para o treinamento dos modelos foram utilizadas imagens de CXR de duas bases de dados distintas. A primeira contendo 800 imagens e 704 máscaras do pulmão de pacientes com tuberculose ou normais, que foram compiladas de dados primários do Hospital de Shenzhen (SH) (662 imagens/ 566 máscaras) e Montgomery County (MC) (138 imagens/ máscaras) e disponibilizadas na plataforma GitHub [8]. A segunda contendo 800 imagens do pulmão de pacientes com COVID-19 ou normais, de dados primários que foram compiladas por Qatar University, Tampere University e Hamad Medical, e selecionadas da base QaTa-COV19 disponibilizadas na plataforma Kaggle [9]. Com o objetivo de captar imagens com as maiores e menores variações de pixels, os arquivos da base QaTa-COV19 foram ordenadas por tamanho selecionando as 200 maiores imagens e as 200 menores imagens de pacientes normais e com COVID-19 respectivamente de forma a perfazer também 800 imagens, para que se possa avaliar o impacto de *data augmentation* em base de dados de tamanho reduzido ($<10^3$ imagens), tal como Stirenko et al. [10].

B. Pré-Processamento

Foi utilizada a plataforma Google Colab Pro com Python 3.7.10, Pytorch 1.8.1, Keras 2.4.0, TensorFlow 2.4.1 e TensorBoard 2.4.0. Antes de iniciar o treinamento das redes foi realizado um pré-processamento dos bancos de dados, gerando 3 bases derivadas de imagens com *data augmentation* seguindo a abordagem proposta por [10], quais sejam Normal, LossLess e Lossy para as quais foram aplicadas as transformações apresentadas na Tabela 1. As máscaras correspondentes a cada imagem também sofreram as mesmas transformações apontadas.

Propósito BD	Data Augmentation	Imagens
Normal	N.A.	800
LossLess	Flip LR, Flip UP, Rotação 90°, 180° e 270°	4.800
Lossy	Rotação 5°, 10° e 355° e 350°	4.000

Tabela 1 – Propósitos das Bases de Dados.

C. Treinamento e Teste

Para o processo de treinamento e escolha da configuração dos modelos, cada uma das bases foi dividida em 3 subconjuntos, de Treinamento, Validação e Teste, respeitando a proporção de 60%, 20% e 20%, respectivamente.

O modelo U-net utilizado na segmentação do pulmão [19], foi treinado por 75 épocas. Os modelos utilizados para a classificação das imagens foram treinados por 300 épocas, com a gravação dos parâmetros do melhor modelo e parada prematura (*early stopping*) em um intervalo de 100 épocas, durante o processo de avaliação dos modelos ou com 200 épocas quando do processo de avaliação da variabilidade dos modelos em relação a inicialização dos hiperparâmetros.

D. Indicadores de Avaliação

Durante o processo de treinamento da segmentação das imagens dos pulmões foram avaliados o Índice de Jaccard [20] e o Coeficiente de Dice [21].

Para avaliação da classificação dos modelos simulados foram elaboradas matrizes de confusão, para que fossem extraídos os indicadores de Sensibilidade (Recall), Acurácia, Especificidade, Precisão e F1. Considerando que o contexto de aplicação desse estudo é o de triagem de pacientes, temos a Sensibilidade como o índice mais relevante.

E. Faseamento

O processo de pesquisa do presente trabalho pode ser subdividido em 2 fases com os seguintes propósitos:

- Fase 1 – Avaliação de Segmentação de pulmão e *Data Augmentation* em pacientes com TB e sem TB (Base SH/MC);
- Fase 2 – Avaliação de *Data Augmentation* em pacientes com COVID e normais (Base QaTa-COV19) com *transfer learning* de modelos treinados da Base SH/MC.

A seguir serão apresentadas as etapas seguidas em cada uma das fases, sendo seus resultados apresentados e analisados na Seção de Resultados e Discussões.

1) Fase 1

Na fase 1 avaliou-se na base SH/MC o impacto da segmentação do pulmão e o *data augmentation*, segundo proposto por [8] [19], na classificação de pacientes sem TB e com TB. Para isso foram feitas simulações para a segmentação do pulmão utilizando o modelo U-Net [19] nas bases auxiliares Normal, LossLess e Lossy.

Os classificadores avaliados foram: MLP, como um *baseline* não profundo; CNN, como um modelo profundo típico utilizado na classificação de imagens; Inceptionv3, como um modelo profundo que, segundo Ke et al. [16], apresentou os melhores resultados entre diversas CNN.

Posteriormente foram feitas simulações com resolução de 128x128, otimizador ADAM e *learning rate* de 0,001, para as seguintes redes neurais:

- U-Net: Implementada via PyTorch com encoder seguindo a VGG11 e batch normalization;
- CNN: Implementada via layers Keras tendo 20 canais convolucionais de 5x5, ativação ReLu e padding valid, seguida de camada MaxPooling 2x2, seguido de novo filtro convolucional 3x3, seguido de nova camada MaxPooling, completamente conectada a uma camada com 30 e 15 neurônios respectivamente;
- MLP: Implementada via layers Keras com função de ativação ReLu, batch normalization e camada densamente conectada com 1.024 neurônios;
- Inceptionv3: Implementada via Applications Kera InceptionV3, com max pooling sem pré-treinamento.

Todos os modelos treinados para as imagens segmentadas também foram treinados com imagens não segmentadas para avaliarmos o impacto da segmentação.

Finalmente, para avaliar a variabilidade dos modelos quanto a inicializações dos hiperparâmetros foram feitas 30 simulações dos modelos CNN, MLP e Inceptionv3. Já para avaliar o impacto quanto a variação dos dados de entrada foram feitas simulações com validação cruzada k-fold com 10 subconjuntos distintos de treinamento, validação e teste desses mesmos modelos.

2) Fase 2

Na fase 2 avaliou-se, em um novo banco de dados (Base QaTa-COV19) com imagens de pacientes normais e COVID-19, o impacto do *data augmentation*. Não foi avaliado o impacto da segmentação dos pulmões, pelo fato dos modelos sem segmentação terem apresentados melhores resultados na Fase 1.

Os melhores modelos de cada uma das topologias treinadas na Fase 1 serviram de base de inicialização, via *transfer learning*, para os treinamentos dos respectivos modelos na base QaTa-COV19.

IV. RESULTADOS E DISCUSSÕES

Nessa seção os resultados serão apresentados para cada uma das fases do processo de pesquisa.

A. Fase 1

Na 1ª Etapa foram efetuadas as segmentações dos pulmões da base SH/MC Normal, LossLess e Lossy, com otimizador Adam e *learning rate* de 0,0005, sendo avaliado o índice de Jaccard e Dice. Observou-se que o processo de *data augmentation* gerou melhores resultados de segmentação para

a base LossLess e Lossy, se comparadas a base Normal (Tabela 2).

	Normal		LossLess		Lossy	
	val jaccard	val dice	val jaccard	val dice	val jaccard	val dice
Médio	85,7%	90,2%	91,1%	94,9%	90,7%	94,2%
Mediana	91,9%	95,7%	92,3%	95,9%	93,3%	96,5%
Máximo	93,6%	96,7%	92,8%	96,2%	96,8%	98,4%
RMS	3,7%	3,4%	0,5%	0,5%	1,5%	1,2%

Tabela 2 - Comparativo de indicadores de segmentação para diferentes data augmentation, onde observa-se que a LossLess e Lossy apresentaram melhores resultados.

A partir da segmentação dos pulmões para cada uma das bases foram efetuadas as classificações de pacientes sem TB e com TB utilizando-se redes CNN, MLP e Inceptionv3.

Os resultados estão apresentados na Tabela 3, como pode-se observar a topologia Inceptionv3 foi a que apresentou um melhor resultado e o baseline de MLP o pior resultado. Em todas as topologias o pré-processamento Lossy apresentou os melhores resultados.

Foram avaliados os modelos em relação às imagens não segmentadas dos pulmões, sendo os resultados apresentados também na Tabela 3. Como pode-se observar os resultados foram melhores que os alcançados com a segmentação do pulmão, tendo também a Inceptionv3 obtido o melhor resultado para o índice de Sensibilidade. Também nessa etapa a base Lossy apresentou um melhor resultado para cada uma das topologias.

O melhor resultado para as imagens não segmentadas do pulmão não está de acordo com o proposto por [19] [8], mas pode estar vinculado a perda de regiões contendo informações importantes durante o processo de segmentação, ou a existência de outros artefatos ou marcadores, que foram capturados durante o processo de treinamento com as redes não segmentadas [10]. Investigação adicional se faz necessária para esclarecer esse ponto.

As topologias foram avaliadas quanto ao impacto na inicialização dos parâmetros, dessa forma foram feitas 30 simulações com os resultados apresentados na Figura 1 onde se observa que a aplicação do data augmentation LossLess e Lossy traduziram-se em um aumento da Acurácia e Recall dos modelos, sendo o Inceptionv3 a topologia com os melhores resultados.

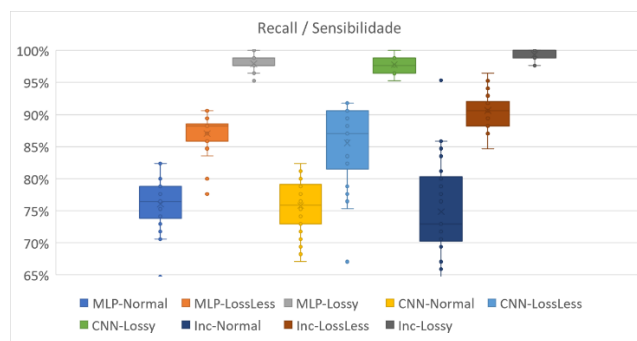
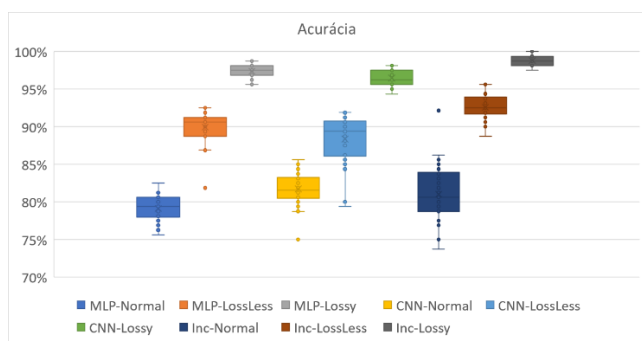


Figura 1- Resultados das simulações dos indicadores de Acurácia e Recall das diversas topologias, quanto a inicialização de parâmetros, para imagens não segmentadas da base SH/MC, onde observa-se que o Lossy apresenta melhores resultados.

	Acurácia		Especificidade	
	Segmentado	Não Segmentado	Segmentado	Não Segmentado
MLP-Normal	78,1% ± 2,7%	79,1% ± 0,6%	79,6% ± 5,8%	74,5% ± 3,1%
MLP-LossLess	85,7% ± 6,5%	90% ± 0,8%	90,5% ± 2,7%	87,8% ± 1,1%
MLP-Lossy	94,8% ± 1,1%	97,4% ± 0,3%	94,5% ± 1,8%	97,9% ± 1,1%
CNN-Normal	82,1% ± 1,2%	81,7% ± 0,8%	82,8% ± 1,7%	76,9% ± 2,0%
CNN-LossLess	89,4% ± 3,1%	88,4% ± 1,2%	92,1% ± 1,5%	86,7% ± 5,8%
CNN-Lossy	94,2% ± 0,9%	96,5% ± 0,3%	93,1% ± 2,0%	97,5% ± 0,7%
Inc-Normal	80,5% ± 3,5%	81,0% ± 1,4%	89,1% ± 4,7%	75,2% ± 4,4%
Inc-LossLess	83,7% ± 2,7%	92,6% ± 0,6%	85,0% ± 6,5%	89,6% ± 1,6%
Inc-Lossy	93,1% ± 1,4%	98,8% ± 0,2%	93,8% ± 2,6%	99,2% ± 0,6%

	Sensibilidade / Recall	
	Segmentado	Não Segmentado
MLP-Normal	77,1% ± 5,1%	76% ± 1,4%
MLP-LossLess	81,9% ± 9,0%	87,1% ± 1,1%
MLP-Lossy	95,1% ± 2,0%	97,9% ± 0,4%
CNN-Normal	81,5% ± 3,0%	75,8% ± 1,5%
CNN-LossLess	86,6% ± 5,5%	85,6% ± 2,3%
CNN-Lossy	95,3% ± 1,7%	97,8% ± 0,4%
Inc-Normal	72,7% ± 8,6%	74,9% ± 3,0%
Inc-LossLess	82,9% ± 6,2%	90,6% ± 1,0%
Inc-Lossy	92,5% ± 3,5%	99,4% ± 0,3%

Piores resultados
Melhores resultados

Tabela 3 - Resultados das simulações para a base SH/MC de pacientes com TB onde observa-se que os melhores resultados foram obtidos com a rede Inceptionv3 com o data augmentation Lossy para as imagens não segmentadas.

B. Fase 2

Nesta etapa os modelos treinados para pacientes com tuberculose serviram de base, tendo seus pesos pré-carregados, para o treinamento dos modelos de avaliação de pacientes com COVID-19 da base QaTa-COV19. Como os resultados para imagens não segmentadas apresentavam os melhores resultados os modelos foram treinados com imagens não segmentadas, sendo seus resultados apresentados na Tabela 4. Como pode-se observar também temos aqui resultados comparativamente melhores com data augmentation que os dados normais.

Todos os modelos treinados para COVID-19 alcançaram acurácias maiores que os correspondentes treinados para tuberculose, tendo o modelo Inceptionv3 Lossy convergido

mais rapidamente que o modelo análogo treinado para TB.

Foram feitas análises da variabilidade na inicialização dos parâmetros, para pacientes com COVID-19 e os resultados obtidos sugerem que o *data augmentation* reduz a variabilidade dos mesmos, tendo a base Lossy apresentado os melhores resultados quanto a acurácia e recall (Figura 2). Observa-se também que para a base Normal os modelos CNN e Inceptionv3 apresentam resultados equivalentes.

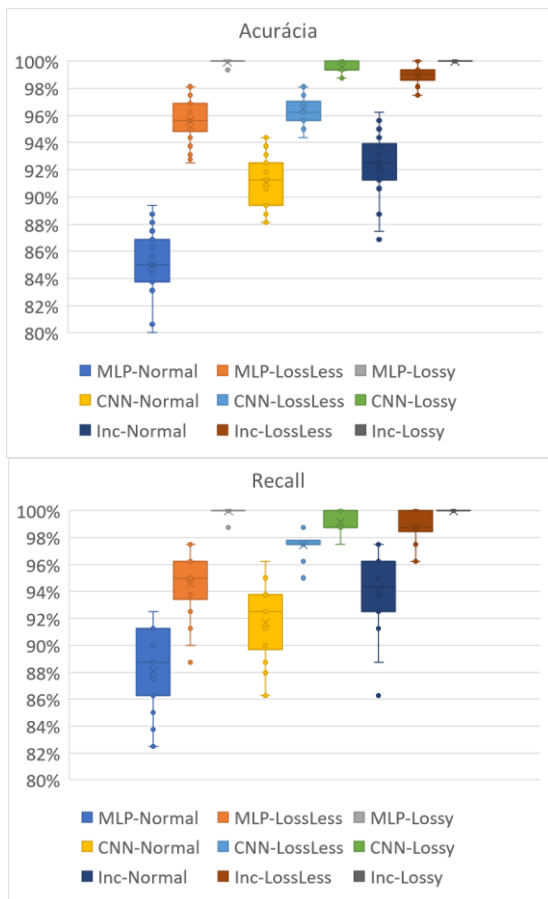


Figura 2 - Resultados das simulações das diversas topologias, quanto a inicialização de parâmetros, para imagens não segmentadas do base QaTa-COVID19, onde observa-se que o Lossy apresenta melhores resultados.

	Acurácia Não Segmentado	Sensibilidade / Recall Não Segmentado	Especificidade Não Segmentado
MLP-Normal	84,5% ± 1,7%	88,8% ± 2,5%	80,4% ± 3,3%
MLP-LossLess	94,1% ± 0,8%	94,6% ± 1,5%	93,5% ± 1,6%
MLP-Lossy	99,1% ± 0,8%	99,1% ± 0,9%	99,1% ± 0,9%
CNN-Normal	88,0% ± 2,0%	90,2% ± 3,7%	85,7% ± 3,8%
CNN-LossLess	97,0% ± 1,2%	97,3% ± 1,4%	96,7% ± 1,5%
CNN-Lossy	99,4% ± 0,4%	99,3% ± 0,6%	99,5% ± 0,5%
Inc-Normal	88,8% ± 1,7%	90,9% ± 2,5%	86,7% ± 2,7%
Inc-LossLess	94,7% ± 5,3%	97,2% ± 1,1%	93,8% ± 6,2%
Inc-Lossy	99,4% ± 0,6%	99,3% ± 0,7%	99,6% ± 0,4%

Piores resultados Melhores resultados

Tabela 4 - Resultados das simulações para a base QaTa-COVID19 de pacientes com COVID-19 onde observa-se que os melhores resultados foram obtidos com a rede Inceptionv3 com o *data augmentation* Lossy para as imagens não segmentadas.

V. CONCLUSÃO E TRABALHOS FUTUROS

O presente trabalho aplicou redes profundas e não profundas em dados reais de pacientes com TB e COVID-19, obtendo resultados promissores, o que o torna relevante no atual contexto do sistema de saúde nacional.

Os modelos profundos de segmentação dos pulmões alcançaram valores satisfatórios com o índice de Jaccard de $85,6\% \pm 0,7\%$ (Normal), $91,1\% \pm 0,5\%$ (LossLess) e $90,7\% \pm 0,8\%$ (Lossy), mostrando a pertinência do *data augmentation* corroborando o encontrado por [10] para pacientes com TB, entretanto, no presente trabalho a classificação de pacientes sem segmentação dos pulmões apresentou resultados gerais melhores. Ressalta-se também que a inclusão de mais um conjunto de rotações de ± 5 graus no *data augmentation* Lossy gerou melhores resultados que [10] para esse tipo de *data augmentation*.

O emprego do *data augmentation* mostrou um maior impacto na melhoria dos indicadores de classificação de pacientes com TB, que o emprego de arquiteturas profundas, visto que a MLP Normal teve uma acurácia de $78,1\% \pm 2,7\%$ já a MLP Lossy obteve uma acurácia de $94,8 \pm 1,1\%$, enquanto a Inceptionv3 Normal obteve $79,1\% \pm 0,6\%$ e a Inceptionv3 Lossy $98,8\% \pm 0,2\%$. Portanto, a combinação de arquitetura profunda Inceptionv3 com o *data augmentation* Lossy foi a que apresentou os melhores resultados globais para TB, mostrando a pertinência do uso de arquitetura profunda combinada com *data augmentation* na análise de imagens de CXR de pequenos conjuntos de dados ($< 10^3$ imagens).

O *transfer learning* de modelos pré-treinados em TB, mostrou-se adequado para pacientes com COVID-19, tendo também os *data augmentation* LossLess e Lossy apresentados melhorias nos indicadores gerais, tendo também a topologia Inceptionv3 Lossy alcançado os melhores resultados, com por exemplo $99,4\% \pm 0,6\%$ para acurácia em pacientes com COVID-19.

Como trabalhos futuros sugerimos a aplicação de segmentação dos pulmões também para os casos de pacientes com COVID-19. Sugere-se também a plotagem dos mapas de atenção especialmente para a topologia Inceptionv3, tanto para os pacientes segmentados quanto não segmentados, afim de

avaliar se os modelos não segmentados alcançaram melhores resultados por identificar aspectos relevantes que foram perdidos no processo de segmentação, ou outros elementos externos ao pulmão que podem estar correlacionados com TB ou COVID-19. Adicionalmente, sugerimos que sejam testadas configurações de redes Inception com menos camadas, a fim de avaliar se podemos obter resultados equivalentes aos alcançados, mas com um menor custo computacional.

AGRADECIMENTOS

Os autores gostariam de agradecer ao CNPq e à FAPERJ pelo apoio a este trabalho. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior Brasil (CAPES).

REFERÊNCIAS

- [1] J. Trujillano, J. March e A. Sorribas, “Methodological approach to the use of artificial neural networks for predicting results in medicine,” *Medicina Clinica*, vol. Suppl 1, n° 122, pp. 59-67, 2004.
- [2] Q. K. Al-Shayea, “Artificial Neural Networks in Medical Diagnosis,” *International Journal of Computer Science*, vol. 8, n° 2, pp. 150-154, 2011.
- [3] E. Sogancioglu, E. Çalli, B. v. Ginneken, K. G. v. Leeuwen e K. Murphy, “Deep Learning for Chest X-ray Analysis: A Survey,” 15 3 2021. [Online]. Available: <https://arxiv.org/pdf/2103.08700.pdf>.
- [4] W. H. Organization, “Global Tuberculosis report 2020,” World Health Organization, Geneva, 2020.
- [5] A. E. Gorbalenya, A. A. Gulyaeva, C. Lauber, S. I. A. A. M. Leontovich, D. Penzar, S. Baker e R. Baric, “The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2,” *Nature Microbiology*, vol. 5, pp. 536-544, 2020.
- [6] M. d. Saúde, “O que é COVID-19,” Ministério da Saúde . [Online]. Available: <https://coronavirus.saude.gov.br/index.php/sobre-a-doenca>. [Acesso em 18/12/2020].
- [7] K. R. d. Silva, F. G. d. Souza, F. F. Roquete, S. M. d. C. Faria, B. C. F. Peixoto e A. Vieira, “Alocação de recursos para assistência à saúde em tempos da pandemia de COVID-19: revisão integrativa,” *Revista Brasileira de Enfermagem*, vol. 73(Suppl. 2), 2020.
- [8] I. Ovcharenko, “Lung segmentation,”. [Online]. Available: <https://github.com/IlliaOvcharenko/lung-segmentation>. [Acesso em 12/3/2021].
- [9] A. Degerli, “QaTa-COV19 Dataset,”. [Online]. Available: <https://www.kaggle.com/aysendegerli/qatacov19-dataset>. [Acesso em 5/4/2021].
- [10] S. Stirenko, Y. Kochura, O. Alienin e O. Rokovyi, “Chest X-Ray Analysis of Tuberculosis by Deep Learning with Segmentation and Augmentation,” em *EEE 38th International Conference on Electronics and Nanotechnology (ELNANO)*, Kyiv, Ukraine, 2018.
- [11] S. Jaeger, A. Karargyris, S. Candemir, L. Folio, J. Siegelman, F. Callaghan, Z. Xue, K. Palaniappan, R. K. Singh, S. Antani, G. Thoma, Y. Wang, P. Lu e C. J. McDonald, “Automatic Tuberculosis Screening Using Chest Radiographs,” *IEEE Transactions on Medical Imaging*, vol. 33, n° 2, pp. 233-245, 2014.
- [12] E. Tartaglione, C. Barbano, C. Berzovini, M. Calandri e M. Grangetto, “Unveiling COVID-19 from CHEST X-Ray with Deep Learning: A Hurdles Race with Small Data,” *International Journal of Environmental Research and Public Health*, vol. 17, n° <https://doi.org/10.3390/ijerph17186933>, 2020.
- [13] N. Tsiknakis, E. Trivizakis, E. Vassalou, G. Papadakis, D. Spandidos, A. Tsatsakis, J. Sánchez-García, R. López-González, N. Papanikolaou, A. Karantanis e K. Marias, “Interpretable artificial intelligence framework for COVID-19 screening on chest X-rays,” *EXPERIMENTAL AND THERAPEUTIC MEDICINE*, vol. 20, n° DOI: 10.3892/etm.2020.8797, pp. 727-735, 20.
- [14] W. H. Organization, “Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19),” WHO Headquarters (HQ), Geneva, [https://www.who.int/publications/i/item/report-of-the-who-china-joint-mission-on-coronavirus-disease-2019-\(covid-19\)](https://www.who.int/publications/i/item/report-of-the-who-china-joint-mission-on-coronavirus-disease-2019-(covid-19)), 2020.
- [15] M. Roberts, D. Driggs, M. Thorpe, J. Gilbey, M. Yeung, S. Ursprung, A. I. Aviles-Rivero, C. Etmann, C. McCague, L. Beer, J. R. Weir-McCall, Z. Teng e e. al., “Common pitfalls and recommendations for using machine learning to detect and prognosticate for COVID-19 using chest radiographs and CT scans,” *Nature Machine Intelligence*, vol. 3, n° <https://doi.org/10.1038/s42256-021-00307-0>, pp. 199-217, 2021.
- [16] A. Ke, W. Ellsworth, O. Banerjee, A. Ng e P. Rajpurkar, “CheXtransfer: Performance and Parameter Efficiency of ImageNet Models for Chest X-Ray Interpretation,” em *Conference on Health, Inference, and Learning (CHIL)*, New York, NY, 2021.
- [17] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong e Q. He, “A Comprehensive Survey on Transfer Learning,”. Available: <https://arxiv.org/abs/1911.02685>.
- [18] J. Yosinski, J. C. a, Y. Bengio e H. Lipson, “How transferable are features in deep neural networks?,” [Online]. Available: <https://arxiv.org/abs/1411.1792>.
- [19] Yu.Gordienko, P. Gang, J. Hui, W. Zeng, Yu.Kochura, O.Alienin, O. Rokovyi e S. Stirenko, “Deep Learning with Lung Segmentation and Bone Shadow Exclusion Techniques for Chest X-Ray Analysis of Lung Cancer,” *Advances in Computer Science for Engineering and Education*, n° May, p. 638–647, 2018.

[20] R. Shi, K. N. Ngan e S. Li, "Jaccard index compensation for object segmentation evaluation," em *IEEE International Conference on Image Processing (ICIP)*, Paris, 2014.

[21] S. Jha, L. H. Son, R. Kumar, I. Priyadarshini, F. Smarandache e H. V. Long, "Neutrosophic image segmentation with Dice Coefficients," *Measurement*, vol. 134, n° ISSN 0263-2241, pp. 762-772, 2019.