



Algorithms to Aid in the Diagnosis of Lung Diseases Using Artificial Intelligence

Gustavo Chichanoski and Maria Bernadete França

EasyChair preprints are intended for rapid dissemination of research results and are integrated with the rest of EasyChair.

April 28, 2022

ALGORITMOS PARA AUXÍLIO NO DIAGNÓSTICO DE DOENÇAS PULMONARES UTILIZANDO INTELIGÊNCIA ARTIFICIAL

Gustavo Chichanoski * Maria Bernadete de Moraes França **

* Pós Graduação em Engenharia Elétrica - Mestrado, Universidade Estadual de Londrina, PR, (e-mail: gustavo.chichanoski@uel.br).

** Departamento de Engenharia Elétrica, Universidade Estadual de Londrina, PR, (e-mail: mbmorais@uel.br)

Abstract: The world has been hit by a pandemic caused by the coronavirus that has impacted all of humanity. The disease caused by this virus affects several systems of the human body, among them the respiratory system is the one that suffers the most, leading to the formation of consolidations and ground-glass opacity in the lungs. These symptoms can be detected in radiology exams. To assist the medical community in diagnosing Covid-19, several works were developed with the aim of generating new tools and methods. In this work, a Covid-19 detection system and probabilistic Grad-CAM generation was developed, using the neural network models ResNet50V2, DenseNet121, InceptionResnetV2 and VGG-19. The results of the two models were compared using the precision and specificity of each one for Covid-19, where the DenseNet121 network obtained precision values of 99.28% and the ResNet50V2 a specificity of 99.72%, which were higher than those obtained in the reference literature.

Resumo: O mundo foi atingido por uma pandemia causada pelo coronavírus que impactou toda a humanidade. A doença ocasionada por esse vírus afeta diversos sistemas do corpo humano, entre eles o sistema respiratório é o que mais sofre, levando a formação de consolidações e opacidade em vidro fosco nos pulmões. Estes sintomas podem ser detectados em exames de radiologia. Para auxiliar a comunidade médica no diagnóstico da Covid-19, diversos trabalhos foram desenvolvidos com o objetivo de gerar novas ferramentas e métodos. Nesse trabalho foi desenvolvido um sistema de detecção de Covid-19 e geração do Grad-CAM probabilístico, utilizando os modelos de redes neurais ResNet50V2, DenseNet121, InceptionResnetV2 e VGG-19. Os resultados dos modelos foram comparados usando a precisão e especificidade de cada um para Covid-19, onde a rede DenseNet121 obteve valores de precisão de 99,28% e a ResNet50V2 uma especificidade de 99,72%, que foram superiores aos obtidos na literatura de referência.

Keywords: Covid-19; Artificial Intelligence; Machine Learning; Deep Learning; Segmentation.

Palavras-chaves: Covid-19; Inteligência Artificial; Aprendizado de máquina; Aprendizado profundo; Segmentação.

1. INTRODUÇÃO

Em dezembro de 2019, o mundo foi surpreendido pela pandemia causada pelo vírus SARS-CoV2, que poderia levar a uma pneumonia severa dos infectados, levando-os à morte. Até abril de 2022, foram registrados mais de 491 milhões de casos e 6,15 milhões de mortes no mundo, chegando a 30 milhões de casos e 660 mil mortos no Brasil.

A Covid-19 pode ser classificada nos quatro estágios seguintes: Leve, Moderado, Grave e Crítico, Shi et al. (2021). No estágio Leve, a pessoa apresenta sintomas leves e sem sinais de pneumonia na radiologia do tórax. No Moderado, apresenta febre e sintomas respiratórios, e com sinais de pneumonia na radiologia do tórax. No Grave, o paciente tem dificuldades respiratórias, saturação de oxigênio e pressão arterial parcial do oxigênio. Por fim, o estágio Crítico apresenta falha respiratória, requerendo ventilação mecânica.

Os sintomas mais presentes nos pacientes com Covid-19, segundo Zhao et al. (2020), são GGO (opacidade de vidro fosco), ocorrendo em 86,1% dos pacientes não emergenciais, e chegando a 100% dos casos emergenciais; GGO com consolidação aparece em 64,4% dos paciente não emergenciais, e Consolidação, se apresentando em 41,4% dos casos não emergenciais e 64,4% dos casos emergenciais 64,4%. Todos esses sintomas podem ser visualizados em radiologias.

Visando auxiliar a comunidade científica, foram desenvolvidas diversas ferramentas que utilizam inteligência artificial, tanto na previsão dos casos quanto para a classificação da doença, como mostrado em Rustam et al. (2020). Entre essas ferramentas, foram publicados os artigos de Oh et al. (2020) e Tabik et al. (2020) que desenvolveram uma nova forma de realizar a análise do pulmão.

Oh et al. (2020) demonstra uma forma de como interpretar os resultados obtidos pelo modelo de inteligência artificial

através do Grad-CAM Probabilístico, que é mais preciso do que a modalidade atual, Grad-CAM. Tabik et al. (2020) por outro lado, utiliza esse novo método como forma de fornecer imagens para análise sobre o estado do pulmão através do índice RALE (*Radiographic Assessment of Lung Edema*).

O principal objetivo desse trabalho é desenvolver um sistema capaz de auxiliar no diagnóstico de doenças pulmonares através de exames de raio-x. Realizando a segmentação do pulmão para a classificação dos estados de saúde que será usada pelo modelo de classificação para gerar o mapa de calor da doença, caso seja classificado como coronavírus.

Para isso foi utilizado os *datasets* de Jaeger et al. (2014) para a segmentação do pulmão, e Cohen et al. (2020) para a classificação.

Utilizando os trabalhos de Oh et al. (2020) como base, alterando o modelo base para ResNet50V2, InceptionResnetV2, DenseNet121 e VGG-19, de modo que esses novos modelos fornecessem mais precisão aos resultados obtidos e gerando as imagens necessárias para as análises dos médicos, produzidas por Tabik et al. (2020).

Para isso foi realizada a segmentação da imagem do pulmão, e variando os modelos bases da classificar a doença do pulmão e gerar o Grad-CAM Probabilístico, para análise futura.

2. FUNDAMENTAÇÃO

Nessa seção serão discutidas as diferentes arquiteturas usadas na segmentação dos pulmões nas imagens de raio-x do tórax e classificação entre Covid-19, Pneumonia e Sem Doença.

Baseando-se no mesmo princípio de funcionamento da rede neural de Oh et al. (2020), foi alterado o modelo de rede utilizado na classificação, tendo como objetivo diminuir o número de parâmetros sem alterar a sua performance.

Todo o trabalho foi desenvolvido em Python com o auxílio do *framework* Keras e disponibilizado em:

<https://github.com/GustavoChichanoski/COVID>

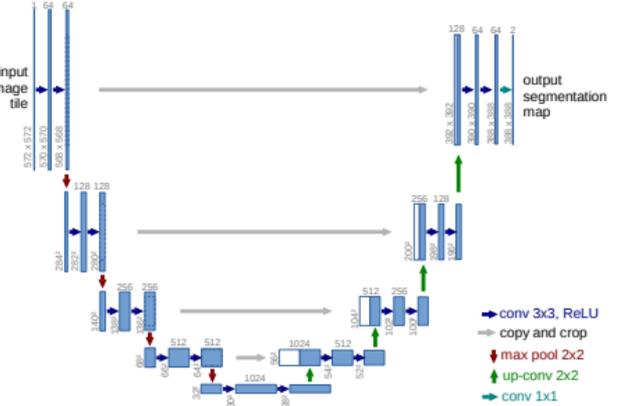
2.1 Rede de Segmentação

O objetivo da rede de segmentação é extrair os pulmões das imagens de raio-x. Para isso foi usado a rede U-Net, sugerido em Ronneberger et al. (2015).

A rede U-Net consiste em caminhos de compressão e expansão. No caminho de compressão, a imagem passa por 2 camadas de convoluções com filtros de tamanhos 3 x 3 e normalização de pacotes, seguidos por uma camada de ativação e uma camada de *max pooling*. Após esse processo, o número de filtros na camada de convolução é dobrado. Essas etapas definem um passo e a rede utilizada neste trabalho possui 5 passos. Sua arquitetura é mostrada na Figura 1.

Quando o último passo for atingindo, começa o caminho da expansão, que consiste em uma camada de *up sampling*, seguida por uma concatenação com a última camada de convolução do caminho de compressão no mesmo nível do

Figura 1. Arquitetura da rede U-Net.



Fonte: Ronneberger et al. (2015).

passo atual, seguido por duas camadas de convoluções, até o primeiro passo novamente.

Finalmente, a última camada de convolução apresenta apenas um filtro de tamanho 1 x 1, seguida de uma camada de ativação (Sigmoide).

Erro Para o cálculo do erro da rede foi utilizado o coeficiente Sorensen-Dice, que visa determinar a correlação entre dois dados, descrito em:

$$D = \frac{2 \cdot TP}{(2 \cdot TP) + FP + FN}, \quad (1)$$

onde TP são os valores verdadeiros positivos, FP os falsos positivos, FN os falsos negativos e D o coeficiente de Dice.

O objetivo do modelo é fazer o valor de D o mais próximo de 1, que acontece quando FP e FN forem igual a 0, logo a função de custo de Dice pode ser calculada como:

$$L = 1 - D, \quad (2)$$

onde L é o valor de custo da correlação entre as imagens da máscara real e a predita pela rede.

Todavia, para melhorar o desempenho da rede foi utilizado uma variação desse erro descrito em Jadon (2020), *Log-Cosh Dice Loss*. Seu erro pode ser calculado por:

$$LD = \log(\cosh(L)), \quad (3)$$

onde LD é o erro para o modelo *Log-Cosh Dice Loss*. A adição do cosseno hiperbólico incrementa o valor do erro quando L for alto e a função logarítmica mantém a equação finita para o intervalo de -1 a 1 .

2.2 Rede de Classificação

Com o objetivo de comparar a performance das redes ResNet50V2, DenseNet121, InceptionResnetV2 e VGG-19, criou-se quatro modelos que utilizaram o mesmo *dataset* de entrada. Os rótulos são divididos em três categorias: Covid-19, Pneumonia e Sem Doença, assim como em Oh et al. (2020). Os pesos da rede de classificação foram pré-carregados para o *dataset ImageNet* a fim de compensar o pequeno número de imagens do *dataset*.

A rede seguirá o procedimento descrito por Oh et al. (2020), dividindo-se em duas abordagens: local e global. Na abordagem local, o modelo prediz os pacotes da imagem, que são pedaços da imagem segmentada de dimensões $m \times n$, que foram cortadas aleatoriamente, evitando as áreas vazias, de tamanhos $u \times v$. Esses pacotes servem de entrada para a abordagem global, que soma seus valores de predição e define a categoria pelo maior valor.

Na parte da inferência da rede, foram escolhidos K pacotes da imagem segmentada, também adquiridos randomicamente evitando as partes vazias, a fim de representar toda a imagem. Para tal é necessário que K seja um valor elevado que cubra toda a imagem. Assim, cada pacote entra na rede gerando a decisão na saída, concedendo um voto para cada doença. O vencedor é selecionado pelo maior número de votos.

2.3 Grad-CAM Probabilístico

Um dos métodos mais utilizadas para visualização dos valores de importância de cada neurônio para uma decisão de interesse é *Gradient-weighted Class Activation Mapping* (Grad-CAM), apresentado em Selvaraju et al. (2020), onde seu cálculo é descrito como:

$$L_{Grad-CAM}^c = ReLU \left(\sum_k \alpha_k^c \cdot A^k \right). \quad (4)$$

A função *ReLU* é utilizada pois os pesos negativos não são interessantes, α_k^c representa a importância dos pesos do neurônio e A^k a entrada do neurônio. O Grad-CAM gera um mapa de calor, *heatmap*, de mesmo tamanho da camada convolucional de interesse, que no caso deste trabalho é a última camada convolucional de tamanho $u \times v$, 7×7 pixels.

O peso α_k^c representa uma linearização parcial da rede profunda através de A , a importância da *feature map* k da classe alvo c , é calculado segundo a equação:

$$\alpha_k^c = \frac{1}{Z} \sum_i^u \sum_j^v \frac{\partial y^c}{\partial A_{i,j}^k}, \quad (5)$$

onde y^c é o *score* da classe c , e Z é o número de pixels do *feature map*, isto é, $Z = u \cdot v$.

Todavia, para conseguir um melhor entendimento de como a rede prevê os pacotes das imagens originais Oh et al. (2020) sugere uma modificação na equação (4) que resulta nas equações (6) e (7), parte local e global respectivamente.

A primeira parte referente a predição do pacote é chamado de aproximação local, nela calcula-se o Grad-CAM do pacote com o objetivo de se obter o mapa de calor, mostrado em:

$$l^c(x) = Up \left(ReLU \left(\sum_k \alpha_k^c \cdot f^k(x) \right) \right) \in \mathbb{R}^{m \times n}, \quad (6)$$

onde $x \in \mathbb{R}^{m \times n}$ é um pacote da imagem de entrada sendo $m \times n$ o valor de 1024×1024 pixels, a função

$f(x) \in \mathbb{R}^{u \times v}$ gera o *feature map* de um pacote de x_k e $Up(\cdot)$ é o operador de *upsampling* do mapa de calor gerado pela *ReLU*, aumentando o de $u \times v$ para $p \times q$ de 224×224 pixels.

Para a aproximação global,

$$[l_{prob}^c(x)]_i = \frac{1}{K_i} \left[\sum_{k=1}^K r^c(x_k) \cdot \mathcal{Q}_k(l^c(x_k)) \right]_i, \quad (7)$$

tem se como objetivo gerar o *heatmap* de toda a imagem, onde x_k representa o pacote k da imagem de entrada; $\mathcal{Q}_k : \mathbb{R}^{p \times q} \mapsto \mathbb{R}^{m \times n}$ é o operador que copia para a posição adequada na nova matriz de zeros $\mathbb{R}^{m \times n}$, o *heatmap* gerado pela equação (7); $r^c(x_k)$ a probabilidade da classe c calculada antes da camada de *Softmax* para o pacote k e K_i denota a frequência que um determinado pixel i foi utilizado.

3. METODOLOGIA

Para produzir o Grad-CAM Probabilístico é necessário primeiramente realizar a segmentação e classificação das imagens. Através do modelo de segmentação é gerada a imagem segmentada, onde é delimitado a área a ser analisada, separando os pulmões do plano de fundo. Na classificação é realizado o treinamento dos pesos para a predição dos rótulos, gerando os mapas de calor dos pacotes de recortes. Por fim, são gerados os Grad-CAM dos pacotes de acordo com os pesos da etapa de classificação. A seguir serão discutidos os processos necessários em cada uma destas etapas.

Na segmentação foram utilizados dois *datasets* públicos, disponíveis em Jaeger et al. (2014). O primeiro *dataset* foi fornecido pelo Departamento de Saúde e Serviços Humanos de Montgomery County (MC), em Maryland, Estados Unidos, e o outro pelo Hospital das Pessoas de Shenzhen 3 da Universidade Médica Guangdong em Shenzhen (Shenzhen), China.

Estes *datasets* juntos contém um total de oitocentas imagens de raio-X do tórax no formato PNG de 12 bits em escala de cinza, com tamanhos de 4020×4892 ou 4892×4020 pixels. As imagens estão divididas em normal e tuberculose, conforme apresentado na Tabela 1. A Figura 2 exhibe uma das imagens fornecidas pelos *datasets*.

Tabela 1. Classificação dos *datasets* de segmentação.

<i>Dataset</i>	Sem Doença	Tuberculose	Total
MC	80	58	138
Shenzhen	326	336	662
Total	406	394	800

As máscaras utilizadas na segmentação foram disponibilizadas por Stirenko et al. (2018), baseadas nos *datasets* fornecidos em Jaeger et al. (2014). Segundo o próprio autor, as máscaras foram geradas a partir de um *dataset* pequeno, formado por menos de 10^3 imagens, o que pode produzir predições com baixa acurácia. Na Figura 3 é apresentada uma das máscara utilizada na segmentação.

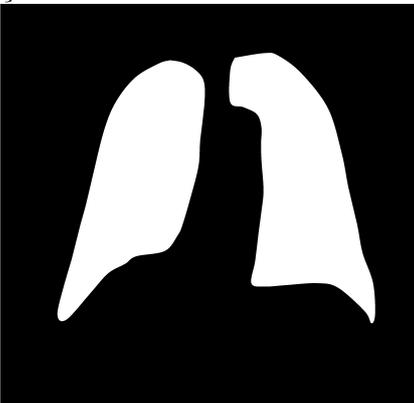
O conjunto formados pelas imagens dos pulmões e das máscaras foi dividido em três partes: treinamento, validação e teste. As imagens foram divididas aleatoriamente em

Figura 2. Exemplo de uma imagem utilizada na segmentação.



Fonte: Jaeger et al. (2014).

Figura 3. Máscara fornecida por um dataset para a segmentação.



Fonte: Stirenko et al. (2018).

uma proporção de 72%, 18% e 10% respectivamente, como mostrado na Tabela 2, tendo como regra a não repetição de uma imagem de uma parte, em qualquer outra parte, visando diminuir o enviesamento da rede.

Tabela 2. Dispersão do *dataset* entre treinamento, validação e teste.

<i>Dataset</i>	Total Imagens	Porcentagem [%]
Treinamento	576	72
Validação	144	18
Teste	80	10
Total	800	100

Fonte: Autor.

A parte do *dataset* destinada ao treinamento será usada como entrada da rede durante o treinamento dos pesos. A validação serve para verificar se as métricas obtidas no treinamento são válidas, buscando evitar o efeito de sobre amostragem da rede. O teste consiste na avaliação da arquitetura proposta, utilizando imagens inéditas para a rede, que não fizeram parte dos processos de treinamento e validação.

Pré-processamento Como as imagens são de diferentes tamanhos e tipos, é necessários primeiramente normalizá-las para que possam servir de entrada para a rede. Desta forma, foi feito o pré-processamento dos *datasets*, realizando as seguintes ações:

- Alteração do tipo de dado de entrada (de uint8 para float32);
- Redimensionamento da imagem (256 x 256 pixels);
- Normalização da imagem (0 a 1);
- Correção Gamma ($\gamma = 0,5$).

Os pesos da rede foram inicializados com uma distribuição de Glorot. Para o treinamento foi utilizado a função erro *Log-Cosh Dice Loss*, junto com o otimizador *Adamax*, cuja a taxa de aprendizado utilizada foi de 10^{-3} , sendo ela reduzida de acordo com o erro de validação e parando o aprendizado caso este não mude por 10 épocas.

3.1 Classificação

Na classificação foi utilizado o *dataset* de imagens disponibilizado de forma pública em Cohen et al. (2020). Ele possui mais de 7.103 imagens, das quais 853 são de pulmões infectados com Covid-19, 1.887 Sem Doenças e 4.363 com Pneumonia. Um exemplo de cada tipo de imagem pode ser visto na Figura 4. Conforme mostra a Tabela 3, o *dataset* foi separado em três partes, sendo 61% destinado para o treinamento, 15% para a validação e 24% para o teste.

Figura 4. Imagens fornecidas pelos *dataset* de classificação.



(a) Covid-19 (b) Sem Doença (c) Pneumonia

Fonte: Cohen et al. (2020).

Tabela 3. Divisão das imagens do *dataset* de classificação.

Imagens	Treinamento	Validação	Teste	<i>Dataset</i>
Covid-19	457	114	282	853
Sem Doença	1.069	267	551	1.887
Pneumonia	2.790	698	875	4.363
<i>Dataset</i>	4.316	1.079	1.708	7.103

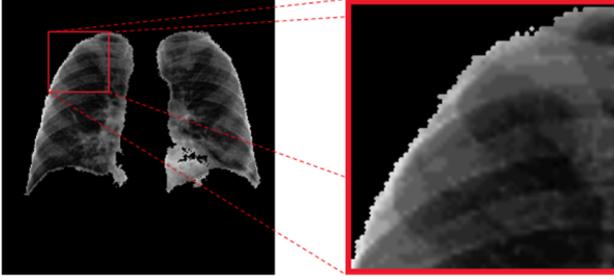
Fonte: Autor.

Primeiramente é feito o pré-processamento das imagens que irão fazer parte da classificação, com o objetivo de corrigir-las para a entrada da rede. Assim, foram feitos os seguintes ajustes nas imagens:

- Normalização da imagem: 0 a 1;
- Redimensionamento da imagem: (1024 x 1024 pixels);
- Correção Gamma ($\gamma = 0,5$);
- Multiplicação pela máscara gerada pela segmentação.

A seguir, a imagem é cortada em diversos pacotes aleatórios de tamanho 224 x 224 pixels, evitando áreas nulas. A finalidade deste processo é produzir as entradas para a rede neural do modelo de classificação, a partir do qual são elaboradas as previsões para cada pacote. Um exemplo do pacote gerado a partir da imagem pré-processada pode ser visto na Figura 5.

Figura 5. Exemplo de um pacote gerado para a rede de classificação.



Fonte: Autor.

O modelo de rede escolhido consiste em uma rede neural composta pelas camadas de entrada, de convolução, camada base e uma camada densa. Esta arquitetura permite a troca entre os modelos a serem testados, sem interferir nos resultados.

A camada de entrada recebe a imagem e converte para tensores do *TensorFlow*. A camada de convolução converte a imagem com apenas um canal para uma imagem de três canais, com três filtros de tamanho 3 x 3, com uma camada de ativação ReLU. A camada base varia conforme o tipo de modelo selecionado: ResNet50V2, DenseNet121, InceptionResnetV2 e VGG-19. A camada densa calcula as previsões da imagem, através da camada de ativação Softmax.

Para o treinamento da rede, os pesos foram inicializados com uma distribuição de Glorot, a função erro utilizada foi *CrossEntropyLoss*, junto com o otimizador *Adamax*, cuja a taxa de aprendizado utilizada foi de 10^{-3} , sendo ela reduzida de acordo com o erro de validação e parando o aprendizado caso este não mude por 10 épocas.

Após o treinamento é possível gerar as previsões para cada pacote, que são somadas, formando um vetor global. A classificação da imagem é o maior valor do vetor global.

A performance da rede é avaliada por meio dos parâmetros acurácia, precisão, revocação e *F1 Score*. Estes parâmetros são descritos respectivamente por:

$$A = \frac{TN + TP}{TN + TP + FN + FP}, \quad (8)$$

$$P = \frac{TP}{TP + FP}, \quad (9)$$

$$R = \frac{TP}{TP + FN}, \quad (10)$$

$$F1 = 2 \cdot \left(\frac{P \cdot R}{P + R} \right), \quad (11)$$

onde *TP* representa os valores verdadeiros positivos, *TN* os verdadeiros negativos, *TN* os falsos positivos e *FN* falsos negativos.

3.2 Grad-CAM Probabilístico

Para a geração das imagens do Grad-CAM Probabilístico, foram utilizados os mapas de calor gerados pelo pacote no

modelo de classificação, removendo os valores negativos. Adicionou-se os mapas de calor a uma nova imagem, com 1024 x 1024 pixels, na mesma posição original do pacote. Por fim, dividiu o valor de cada pixel pelo número de pacotes que continham sua posição.

O número de pacotes pode variar conforme a necessidade, gerando imagens mais nítidas de acordo com o número maior de pacotes. Para a análise dos resultados, os Grad-CAM Probabilísticos desse trabalho foram gerados a partir de 400 pacotes.

4. RESULTADOS

Nesse seção, serão apresentado os resultados obtidos a partir do modelo de segmentação e classificação da doença.

4.1 Segmentação

Para avaliar o desempenho da rede de segmentação foi usado o parâmetro do coeficiente de Jaccard, onde os resultados são mostrados na Tabela 4, apresentando os 5 maiores e menores valores obtidos da predição. O modelo obteve uma média de 94,93% com um desvio padrão de 3,48%, valor superior a literatura de referência Oh et al. (2020), que obteve uma similaridade de 94,30%. A rede de segmentação desenvolvida possui 4,7 milhões de parâmetros, contra os 9,4 milhões de parâmetros de Oh et al. (2020).

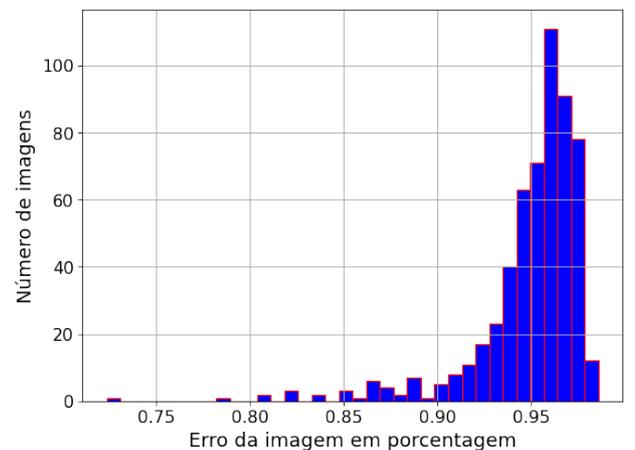
Tabela 4. Comparação entre os 5 melhores e piores resultados obtido pela similaridade de Jaccard.

Valores	1	2	3	4	5
Mínimos	71,62	77,13	78,41	79,13	79,87
Máximos	98,42	98,24	98,10	98,06	98,05

Fonte: Autor.

O histograma dos coeficientes de Jaccard para todas as imagens do *dataset* treino é apresentado na Figura 6. Nela é possível notar que das 576 imagens, 538 das previsões se encontram acima dos 94,3% de similaridade, enquanto 38 estavam abaixo, sendo o pior resultado apresentando uma diferença de 26,48% em comparação com o melhor.

Figura 6. Histograma dos coeficientes de Jaccard das previsões realizadas pelo modelo de segmentação.



Fonte: Autor.

A Figura 7 apresenta exemplo de imagem segmentada pelo modelo desenvolvido, a Figura 7(a) é a imagem do pulmão pré-processada, a Figura 7(b) mostra a imagem da máscara de segmentação de treino e por fim a Figura 7(c) apresenta a imagem dos pulmões segmentado.

Figura 7. Exemplos de imagens segmentadas pelo modelo desenvolvido. A Figura 7(a) é a imagem original fornecida pelo *dataset*, a Figura 7(b) é a máscara original fornecida também pelo *dataset*, e por último a Figura 7(c) é a imagem segmentada.



(a) Original (b) Máscara (c) Predição

Fonte: Autor.

4.2 Classificação

A Tabela 5 apresenta os resultados obtidos pelas diversas redes, cujos os valores foram obtidos utilizando 100 recortes de cada imagem do *dataset* de teste. Os valores em negritos são os melhores resultados obtidos pelo parâmetro da linha.

A rede ResNet50V2 se mostra a melhor nas métricas de VP, acurácia, revocação e especificidade para a categoria de Covid-19. A rede InceptionResnetV2 obteve bons resultados em quase todas as métricas, mas não se mostrou melhor em nenhuma e ainda exigia o maior número de parâmetros. Enquanto, a rede VGG-19 não se mostrou a melhor em nenhuma métrica e ainda obteve o pior resultado na maioria delas. Por fim, a rede DenseNet121 que obteve a melhor precisão entre todas as redes,.

Todavia, todas as redes poderiam ser utilizadas como auxílio no diagnóstico de Covid-19, visto que apresentaram um valor de precisão superior a 69%, obtida pela classificação dos médicos dos exames de raio-x em Oh et al. (2020).

Tabela 5. Comparação das métricas de avaliação para Covid-19 usando 100 recortes.

Métricas	ResNet50V2	DenseNet121	InceptionResnetV2	VGG-19
VP	278	274	276	260
VN	1420	1423	1420	1422
FP	5	2	4	3
FN	4	8	6	22
Precisão	98,23	99,28	98,57	98,86
Acurácia	99,47	99,41	99,41	98,54
Revocação	98,58	97,16	97,87	92,20
Especificidade	99,72	99,44	99,58	98,48
Parâmetros	23.570.980	7.040.613	54.341.381	20.025.957

Fonte: Autor.

Comparando ainda com o resultados obtidos em Oh et al. (2020) na Tabela 6, apresentando os melhores valores de cada coluna em negritos. A tabela mostra que apesar de possuir uma menor especificidade para Covid-19 se comparado ao trabalho de Oh et al. (2020), todos os outros parâmetros se mostraram superiores, mostrando que a alteração dos modelos de redes utilizados pode gerar um ganho na precisão e especificidade do sistema.

Tabela 6. Comparação entre a precisão e a revogação do atual trabalho com a literatura.

	Precisão		
	Covid-19	Sem Doença	Pneumonia
Oh et al. (2020)	90,3	95,7	90,3
ResNet50V2	98,23	94,32	95,90
DenseNet121	98,91	95,51	94,65
InceptionResnetV2	98,57	94,98	95,05
VGG-19	98,86	94,55	94,10

Fonte: Autor.

Tabela 7. Comparação entre a precisão e a revogação do atual trabalho com a literatura.

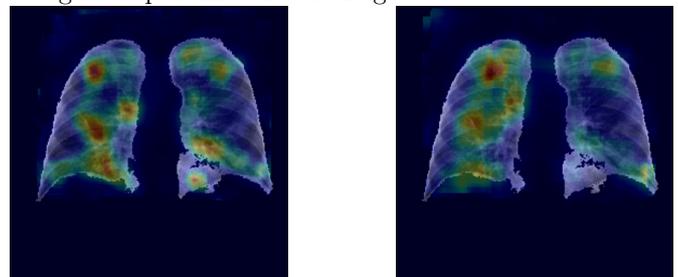
	Especificidade		
	Covid-19	Sem Doença	Pneumonia
Oh et al. (2020)	100	90	93
ResNet50V2	99,72	96,98	96,02
DenseNet121	99,30	96,59	96,91
InceptionResnetV2	99,58	96,58	96,45
VGG-19	98,48	97,32	96,19

Fonte: Autor.

4.3 Grad-CAM Probabilístico

A Figura 8 apresenta a comparação entre os diversos modelos de redes utilizados na geração dos Grad-CAM Probabilístico utilizando 400 pacotes gerados pelos modelos ResNet50V2, DenseNet121, InceptionResnetV2 e VGG-19, mostrados respectivamente nas figuras 8(a), 8(b), 8(c) e 8(d).

Figura 8. Exemplos de imagens Grad-CAM probabilístico gerado pelo modelo do artigo.



(a) ResNet50V2

(b) DenseNet121



(c) InceptionResnetV2



(d) VGG-19

Fonte: Autor.

As regiões em vermelho em cada Grad-CAM Probabilístico representam os locais que mais acionaram cada rede. Nota-se que as regiões em vermelho apresentam-se em maior parte no pulmão à esquerda da imagem, tendo a parte superior e inferior ressaltadas, e em locais semelhantes porém com diferentes intensidades.

5. CONCLUSÃO

O presente trabalho propôs um sistema para auxílio no diagnóstico de doenças pulmonares, utilizando dataset públicos para o seu treinamento, com o objetivo de encontrar os modelos para a segmentação, classificação e severidade da doença.

A segmentação e classificação do pulmão, como mostrado obtiveram resultados semelhantes à literatura, com ressalva para a classificação, que mostrou resultados promissores.

A ResNet50V2 se mostrou superior em quase todos os aspectos, porém a pouca diferença em valores dos parâmetros probabilísticos analisados, aliado ao significativo número de parâmetros da ResNet50V2, fazem da DenseNet121 a melhor escolha em quase todas as situações.

As imagens de Grad-CAM Probabilísticos geradas pelos modelos apresentaram regiões de maior interesse em pontos semelhantes em quase todas as redes, exceto na rede VGG-19 que ativou regiões diferentes das demais.

AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. Sinceros agradecimentos à equipe de professores e colegas do Laboratório de Automação e Instrumentação do Departamento de Engenharia Elétrica da Universidade Estadual de Londrina (UEL) e demais setores envolvidos.

REFERÊNCIAS

- Cohen, J.P., Morrison, P., and Dao, L. (2020). Covid-19 image data collection. *arXiv 2003.11597*. URL <https://github.com/ieee8023/covid-chestxray-dataset>.
- Jadon, S. (2020). A survey of loss functions for semantic segmentation. *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology, CIBCB 2020*. doi:10.1109/CIBCB48159.2020.9277638.
- Jaeger, S., Candemir, S., Antani, S., Wang, Y.X.J., Lu, P.X., and Thoma, G. (2014). Two public chest x-ray datasets for computer-aided screening of pulmonary diseases. *Quantitative imaging in medicine and surgery*, 4(6), 475–477. doi:10.3978/j.issn.2223-4292.2014.11.20.
- Oh, Y., Park, S., and Ye, J.C. (2020). Deep learning covid-19 features on cxr using limited training data sets. *IEEE Transactions on Medical Imaging*, 39(8), 2688–2700. doi:10.1109/TMI.2020.2993291.
- Ronneberger, O., Fischer, P., and Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 9351, 234–241. doi:10.1007/978-3-319-24574-4_28. URL <https://arxiv.org/pdf/1505.04597.pdf>.
- Rustam, F., Reshi, A.A., Mehmood, A., Ullah, S., On, B.W., Aslam, W., and Choi, G.S. (2020). COVID-19 Future Forecasting Using Supervised Machine Learning Models. *IEEE Access*, 8, 101489–101499. doi:10.1109/ACCESS.2020.2997311.
- Selvaraju, R.R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D. (2020). Grad-cam: Visual explanations from deep networks via gradient-based localization. *International Journal of Computer Vision*, 128(2), 336–359. doi:10.1007/s11263-019-01228-7. URL <https://arxiv.org/pdf/1610.02391.pdf>.
- Shi, F., Wang, J., Shi, J., Wu, Z., Wang, Q., Tang, Z., He, K., Shi, Y., and Shen, D. (2021). Review of Artificial Intelligence Techniques in Imaging Data Acquisition, Segmentation, and Diagnosis for COVID-19. *IEEE Reviews in Biomedical Engineering*, 14(c), 4–15. doi:10.1109/RBME.2020.2987975.
- Stirenko, S., Kochura, Y., Alienin, O., Rokovyi, O., Gordienko, Y., Gang, P., and Zeng, W. (2018). Chest x-ray analysis of tuberculosis by deep learning with segmentation and augmentation. *arXiv*.
- Tabik, S., Gomez-Rios, A., Martin-Rodriguez, J.L., Sevillano-Garcia, I., Rey-Area, M., Charte, D., Guirado, E., Suarez, J.L., Luengo, J., Valero-Gonzalez, M.A., Garcia-Villanova, P., Olmedo-Sanchez, E., and Herrera, F. (2020). COVIDGR Dataset and COVID-SDNet Methodology for Predicting COVID-19 Based on Chest X-Ray Images. *IEEE Journal of Biomedical and Health Informatics*, 24(12), 3595–3605. doi:10.1109/JBHI.2020.3037127.
- Zhao, W., Zhong, Z., Xie, X., Yu, Q., and Liu, J. (2020). Relation between chest CT findings and clinical conditions of coronavirus disease (covid-19) pneumonia: A multicenter study. *American Journal of Roentgenology*, 214(5), 1072–1077. doi:10.2214/AJR.20.22976.