

Apoio para Triagem da Infecção por COVID-19 Através de Aprendizagem Profunda Sobre Imagens Segmentadas de Radiografias do Tórax

1st Cecília de Moura Costa

Universidade Federal do Rio de Janeiro - UFRJ
Programa de Pós-Graduação em Medicina (Radiologia)
Rio de Janeiro, Brasil
ceciliamcost@gmail.com

2nd José Manoel de Seixas

Universidade Federal do Rio de Janeiro - UFRJ
Laboratório de Processamento de Sinais, COPPE/POLI
Rio de Janeiro, Brasil
seixas@lps.ufrj.br

3rd Bianca Gutfilen

Universidade Federal do Rio de Janeiro - UFRJ
Departamento de Radiologia, Faculdade de Medicina
Rio de Janeiro, Brasil
bianca.gutfilen@gmail.com

4th Sérgio Augusto Lopes de Souza

Universidade Federal do Rio de Janeiro - UFRJ
Departamento de Radiologia, Faculdade de Medicina
Rio de Janeiro, Brasil
sergioalsouza@gmail.com

5th Susie Medeiros Oliveira

Universidade Federal do Rio de Janeiro - UFRJ
Hospital Universitário Clementino Fraga Filho, Departamento de Radiologia
Rio de Janeiro, Brasil
medeiros.susie@gmail.com

6th Carlos Danilo Miranda Regis

Instituto Federal da Paraíba - IFPB
Departamento de Engenharia Elétrica
João Pessoa, Brasil
danilo.regis@ifpb.edu.br

7th Ricardo de Sousa Farias

Instituto Federal da Paraíba - IFPB
Programa de Pós-Graduação em Engenharia Elétrica
João Pessoa, Brasil
ricardo.farias@academico.ifpb.edu.br

8th Kaique Rijkaard de Sousa Oliveira

Instituto Federal da Paraíba - IFPB
Programa de Pós-Graduação em Engenharia Elétrica
João Pessoa, Brasil
kaique.rijkaard@academico.ifpb.edu.br

Resumo—A busca por eficiência, na triagem de pacientes com indicativo de infecção por COVID-19, tem impulsionado pesquisadores de diversas áreas no desenvolvimento de sistemas de apoio, com particular ênfase no aprendizado profundo. Neste trabalho, consideramos as imagens de Radiografias do tórax, que são segmentadas para identificar os pulmões como região de interesse. Para tal, o modelo MultiResUnet é utilizado, obtendo um Índice de Jaccard de 93,34% e Coeficiente Dice de 96,53%. As imagens segmentadas alimentaram dois modelos profundos, com diferentes níveis de complexidade e um baseado em uma Máquina de Vetores de Suporte (SVM) do tipo Classe Única. Os modelos de aprendizado profundo apresentaram desempenho superior, quando comparados ao SVM, obtendo uma sensibilidade de 97,03%(±1,24) para a identificação do COVID-19. Porém, observou-se que o aumento da complexidade não propiciou ganhos expressivos no desempenho.

Palavras-chave—COVID-19, Radiografias do tórax, Segmentação, Aprendizado Profundo, Classificação.

I. INTRODUÇÃO

Declarada pandemia em março de 2020, pela Organização Mundial da Saúde (OMS) [1], a infecção por COVID-19 tem se espalhado rapidamente e afetado a vida de bilhões de pessoas ao redor do mundo. Com um alto número de infectados

e cerca de 4,5 milhões de mortes [2], a infecção causada pelo vírus pode causar danos significativos nos pulmões e insuficiência respiratória progressiva [3].

O teste para detecção do COVID-19 é baseado na Reação em Cadeia da Polimerase de Transcrição Reversa (RT-PCR) [4]. Embora amplamente utilizado, o teste possui limitações como disponibilidade do material necessário à realização do exame, e sensibilidade relativamente baixa. Além de atestar o COVID-19 com o RT-PCR, exames de diagnóstico por imagens como Tomografia Computadorizada, Radiografias e Ultrassonografias têm favorecido a compreensão do processo inflamatório da doença. Considerando o forte impacto da doença no sistema respiratório, imagens de Radiografias do tórax têm sido bastante utilizadas para a triagem de pacientes e avaliação da progressão da doença, em casos de internação hospitalar [5], [6].

A necessidade de apoiar, a triagem e ao diagnóstico de doenças contagiosas de forte impacto em saúde coletiva, como o COVID-19, tem impulsionado o desenvolvimento de sistemas que utilizam técnicas de aprendizado de máquina, notadamente o aprendizado profundo [7]. Alguns trabalhos

propõem ainda modelos de segmentação das imagens, buscando identificar os pulmões como região de interesse [8]. No caso, as imagens foram segmentadas por um modelo baseado em uma topologia U-Net [9], e aplicadas à diferentes modelos de classificação. O modelo que obteve melhor desempenho foi o InceptionV3 [8], com um valor de F1 de 0,9841, na distinção de COVID-19 em relação à pneumonia adquirida na comunidade (CAP) e ausência de achados radiológicos (Normal).

Uma alternativa *Ensemble* também foi explorada a partir da segmentação pulmonar, com uma adaptação do modelo U-Net baseado em DenseNet103 [10] como *backend* [3]. Na identificação das classes COVID-19, pneumonia e normal, o modelo obteve uma acurácia de 99,2%, enquanto que a segmentação atingiu 92% para o coeficiente Dice, medida comumente utilizada para avaliar desempenho de modelos de segmentação [11].

Uma outra abordagem treina modelos de aprendizado profundo para duas tarefas: segmentação pulmonar e detecção da área de acometimento [12]. Ao utilizar um modelo modificado da U-Net com a EfficientNet [13] como *backbone*, as imagens segmentadas propiciaram ao modelo uma acurácia de 90,7% e com concordância $\kappa = 0,71$ no diagnóstico radiológico.

Portanto, utilizar a segmentação pulmonar na identificação do COVID-19 tem favorecido os modelos de classificação. Em [14], é proposta uma modificação do modelo U-Net, que obteve um índice Jaccard de 94,3% [15] e coeficiente de Dice de 96,94%. Os melhores resultados foram alcançados pela ChexNet com 96,29% para acurácia e sensibilidade e 97,27% de especificidade [14].

Neste trabalho, propomos a segmentação pulmonar com o modelo MultiResUnet e classificação binária com três modelos com um mais complexo e sofisticado, proposto por [16] e adaptado a um conjunto de dados de imagens de Radiografias do tórax, uma vez que a proposta original do modelo utilizou imagens de Tomografia Computadorizada (TC). Embora as imagens de TC também sejam utilizadas para a identificação da COVID-19 [17], [18], a grande disponibilidade de imagens de Radiografias do tórax têm favorecido à sua utilização [19]. O modelo de classificação é comparado com dois modelos de referência: uma rede neural convolucional (CNN) [20] e um SVM (*Support Vector Machine*) do tipo classe única [21], visando a triagem de pacientes com indicação de infecção por COVID-19.

O artigo está organizado da seguinte maneira, na Seção 2 é apresentada uma breve descrição dos conceitos utilizados nos modelos desenvolvidos. Na Seção 3, os passos necessários para elaboração deste trabalho são expostos. Os resultados obtidos são apresentados na Seção 4 e, por fim, na Seção 5 estão as conclusões e propostas futuras.

II. APRENDIZADO DE MÁQUINA

Por muitos anos, a humanidade buscou construir máquinas inteligentes. Com inspiração na maneira como o cérebro humano é capaz de aprender, diversos pesquisadores tentaram simular seu funcionamento nas mais variadas tarefas como

classificação, reconhecimento de padrões, processamento de imagens e processamento de linguagem natural. O conceito de aprendizado de máquina (*Machine learning* - ML) fornece aos sistemas a capacidade de aprender e/ou melhorar automaticamente com a experiência, sem ser explicitamente programado [22].

Dentre as muitas soluções existentes para problemas de classificação as Máquinas de Vetor Suporte (SVM) têm mostrado desempenho robusto em diversas aplicações [21], [23]. Como uma extensão natural das SVMs, as Máquinas de Vetor Suporte de classe única (SVM *single class*) estimam uma distribuição sob observações de uma determinada classe e rotulam classes diferentes como *outliers* ou novidades. Em outras palavras, estimam uma função de distribuição de probabilidade que torna a maioria dos dados observados mais prováveis do que o restante, e uma regra de decisão que separa essas observações pela maior margem possível [24].

A Fig. 1 ilustra a ideia do algoritmo do SVM (classe única), cujo objetivo consiste em encontrar eventos incomuns, produzindo uma função de predição ϕ treinada para “capturar” a maioria das características dos dados de treinamento. Para esse propósito, uma função *kernel* é usada para mapear os dados em um espaço de recurso, em que o SVM é empregado para encontrar o hiperplano com margem máxima da origem do espaço de recurso. Neste uso, a margem a ser maximizada entre as duas classes (em duas classes SVM) torna-se a distância entre a origem e os vetores de suporte que definem os limites do círculo circundante, (ou hipersfera em um espaço de alta dimensão) que inclui a única classe [24].

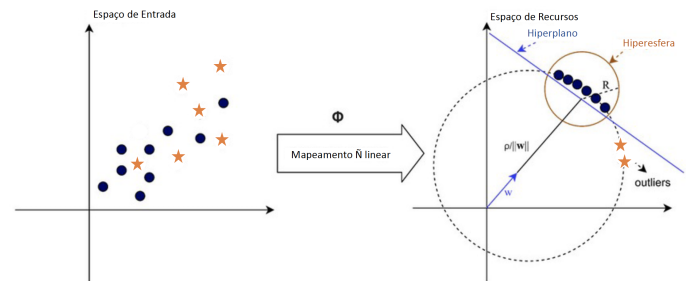


Figura 1. SVM de classe única com kernel RBF.

Na formulação do SVM (classe única), inicialmente o espaço de entrada x é transformado em espaço de recursos $\phi(x)$, como apresentado na Eq. 1. Para computar a margem, é preciso maximizar a distância do hiperplano, segundo a Eq. 2, em que F representa o espaço de recursos e w um vetor de parâmetros.

$$x = (x_1, \dots, x_m) \mapsto \phi(x) = (\phi_1(x), \dots, \phi_m(x)) \quad (1)$$

$$\min_{w \in F} \frac{1}{2} \|w\|^2 \text{ com } \langle w \cdot \phi(x_i) \rangle \geq \rho \quad \forall i \quad (2)$$

Para que a maioria dos dados possam ser separados por uma grande margem aplicam-se limites de erros $v \in (0, 1]$, resolvendo uma determinada restrição com uma variável de

folga ξ_i . A Eq. 3 calcula a distância entre a margem e os *outliers*, resolvendo o problema com a restrição descrita na Eq. 4.

$$\min_{w, \xi_i, \rho} \frac{1}{2} \|w\|^2 + \sum_{i=1}^m (\xi_i - v\rho) \quad (3)$$

$$\langle w \cdot \phi(x_i) \rangle \geq \rho - \xi_i, \quad \xi_i \geq 0, \quad i = 1, \dots, m. \quad (4)$$

Por fim, seleciona-se um método de *kernel* apropriado (linear, RBF, polinomial, entre outros) e resolve-se a Eq. 5, em que α_i e α_j representam multiplicadores de Lagrange sendo a função de *kernel*, $K(x, y) = \phi(x) \cdot \phi(y)$.

$$\min \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j K(x_i, x_j) \quad (5)$$

A. Aprendizado Profundo

O aprendizado profundo é um subcampo do aprendizado de máquina e tem sua estrutura baseada em redes neurais artificiais. O princípio da sua aplicação consiste na aprendizagem de representação criando camadas sucessivas de representações significativas, em que redes neurais artificiais (ANNs) são empilhadas umas sobre as outras [25].

A base das ANNs profundas são as redes chamadas de *feedforwards*, em que a informação flui por meio de uma função que está sendo avaliada, para determinar uma saída. O aprendizado profundo abrange diversas técnicas como as redes neurais convolucionais (CNN), que são compostas por diversas camadas que utilizam a operação de convolução para realizar a extração de características, por isso o termo convolucional. A Figura 2 baseada em [26], ilustra a operação de convolução que a partir de uma janela de dados deslizante, chamada de filtro convolutivo, percorre toda a imagem de entrada operando de forma análoga à sobreposição dos campos visuais. Os valores dos filtros podem ser ajustados de acordo com as características distintas da imagem de entrada [27].

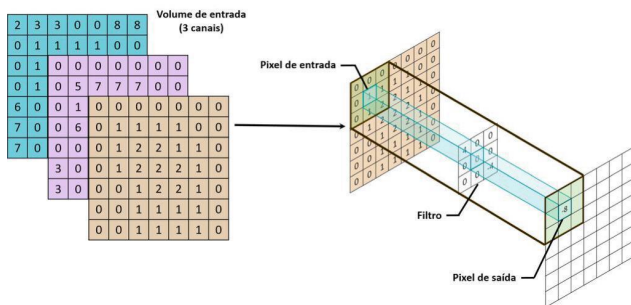


Figura 2. Operação de convolução

Sua arquitetura, normalmente é composta por camadas convolucionais, e entre elas operações como *pooling* (agrupamento), normalização, *padding* (para manter a dimensionalidade na imagem resultante durante a operação) e ao final uma camada densa ou totalmente conectada, como as ANNs tradicionais. As transformações sobre os dados de entrada,

ao longo dessas camadas, formam mapas de recursos que carregam informações específicas como bordas, intensidade, cor, contorno e formas que permitirão aos modelos realizarem determinadas tarefas [27].

III. APOIO À TRIAGEM PARA COVID-19

O modelo desenvolvido é constituído de duas etapas: segmentação e classificação. Na primeira etapa um modelo de aprendizado profundo, MultiResUnet é utilizado para segmentar as imagens de Radiografias do tórax, extraindo os pulmões como regiões de interesse. Na etapa seguinte, um modelo de aprendizado profundo mais complexo, COVNet é implementado visando a comparação com dois modelos de referência (mais simples): um baseado em CNN e outro em SVM de classe única. Os modelos de classificação são avaliados prioritariamente pela sensibilidade, visando o apoio na triagem de pacientes com indicação de infecção por COVID-19. Além disso, outras medidas de desempenho, como especificidade, acurácia, pontuação F1, Recall e área sob a curva ROC (*Receiver Operating Characteristics*), AUC, também são utilizadas.

A. Banco de dados

O COVID-19 *Radiography Database* é um banco de dados de domínio público hospedado no Kaggle, de imagens de Radiografias do tórax composto por quatro classes: COVID-19, Normal, Opacidades Pulmonares e Pneumonia Viral. Este banco está em constante evolução e é elaborado em parceria pela Universidade do Qatar e Universidade de Dhaka, Bangladesh juntamente com colaboradores especialistas do Paquistão e da Malásia [28].

O banco conta atualmente com 3616 imagens na classe COVID-19 e cerca de 10000 imagens na classe normal. As imagens disponibilizadas estão em formato PNG e com dimensões 299×299 acompanhadas de um arquivo de metadados com informações sobre as referências da imagem [14].

B. Segmentação

A arquitetura U-Net tem sido proeminente nas aplicações que visam a segmentação de imagens médicas, sendo amplamente utilizada, assim como suas variantes [9]. Embora esta estrutura tenha boas pontuações em competições de segmentação, algumas modificações em sua arquitetura clássica têm apresentado ganhos de desempenho significativos. Um exemplo desses ganhos está atrelado ao modelo MultiResUnet, que propõe melhorias em dois aspectos da U-Net: variação de escala em imagens médicas e na lacuna semântica entre os níveis correspondentes do codificador-decodificador [29].

A arquitetura do modelo MultiResUnet é apresentada na Figura 3, com base em [29], em que os blocos "MultiRes" são compostos por três camadas convolucionais de *kernels* 3×3 , 5×5 e 7×7 conectadas em paralelo, com o intuito de aumentar a capacidade de análise multiresolução em uma imagem. Os caminhos "Res" constituem uma cadeia de camadas convolucionais com conexões residuais, de modo a

diminuir a discrepância entre os níveis dos mapas de recursos do codificador e decodificador, que são mesclados, durante o processo de segmentação [29].

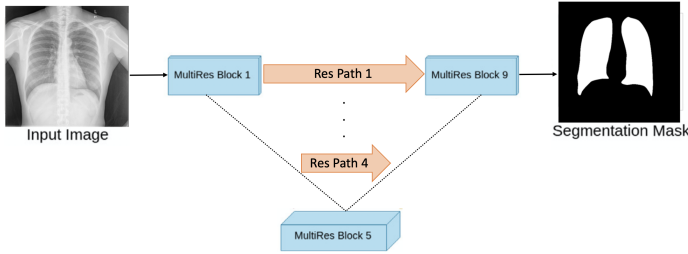


Figura 3. Arquitetura MultiResUnet utilizada para segmentar imagens de Radiografias do tórax.

O modelo MultiResUnet, em sua proposta original, foi treinado em um conjunto de dados heterogêneo e que reúne cinco conjuntos de dados de referência em imagens médicas, afim de favorecer à sua arquitetura com uma variedade de modalidades de imagens. Para realizar a segmentação pulmonar, nas imagens de Radiografias, foi aplicado o ajuste fino no MultiResUnet treinando a última camada, em um conjunto de dados contendo 399 imagens de Raio-x de pulmões normais e suas respectivas máscaras de segmentação, disponibilizados por [30]. Devido a pequena quantidade de dados do conjunto de treinamento, a última camada da rede foi ajustada para uma saída binária, mantendo todos os parâmetros aprendidos pela rede original e realizando o treinamento apenas na camada ajustada.

Para diversificar o conjunto de dados, antes do processo de treinamento, realizou-se um aumento de dados (*Data Augmentation*) no qual, para cada imagem, 15 novas imagens são geradas [31], resultantes de operações como, rotações verticais e horizontais, acréscimo de 15% de ruído branco e aumento (zoom) de 30%. Após o treinamento, o conhecimento adquirido pelo modelo foi transferido para fazer a segmentação das imagens do banco de dados de interesse, COVID-19 *Radiography Database*. Essa transferência foi realizada aplicando, como conjunto de teste, o conjunto de dados de interesse para que o modelo realizasse a predição das respectivas máscaras, nas novas amostras.

Foram utilizadas 1201 imagens da classe COVID-19 e 1232 imagens Normais, que representam a quantidade de dados disponibilizados no banco de dados, antes de sua recente atualização. O modelo recebe, como entrada, as imagens a serem segmentadas e gera as respectivas máscaras de segmentação. A Figura 4 ilustra esse processo, em que ao final, a imagem original Figura 4(a) é multiplicada pela máscara gerada pelo modelo Figura 4(b) resultando em uma imagem que contém apenas a região pulmonar Figura 4(c). As imagens segmentadas são salvas no mesmo formato e dimensões das imagens originais.

Na segmentação semântica, geralmente os pontos de interesse compreendem em uma pequena parte da imagem inteira, assim o desempenho no treinamento foi avaliado conforme medidas de similaridade, como Índice de Jaccard e Coeficiente

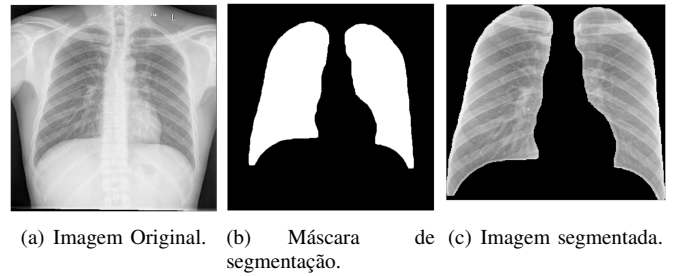


Figura 4. Segmentação do pulmão utilizando as máscaras geradas pelo modelo MultiResUnet.

Dice [11], [15]. O Índice de Jaccard é uma medida estatística que avalia a sobreposição entre duas imagens binárias e pode ser calculado conforme a Eq. 6, em que A representa a máscara binária verdadeira e B a máscara predita pelo modelo [15].

$$\text{Índice de Jaccard} = \frac{A \cap B}{A \cup B} \quad (6)$$

De modo similar, o Coeficiente Dice avalia sobreposição espacial entre imagens e pode ser obtido por meio da Eq. 7, em que A e B representam as imagens das máscaras real e predita, respectivamente [11].

$$\text{Dice}(A,B) = 2 * \frac{|A \cap B|}{|A| + |B|} \quad (7)$$

C. Classificação

No contexto de triagem de pacientes, uma classificação binária foi realizada por meio dos modelos CNN, SVM de classe única e pelo modelo proposto por [16], denominado COVNet para uma classificação mais sofisticada. Esse modelo tem como *backbone* uma Resnet50, que extrai características que são formuladas em um mapa de recursos, analisado em uma camada totalmente conectada e ativada com a função *softmax*. A arquitetura do modelo é apresentado na Figura 5 baseada em [16], que originalmente realizou a detecção da COVID-19 em imagens de Tomografia Computadorizada.

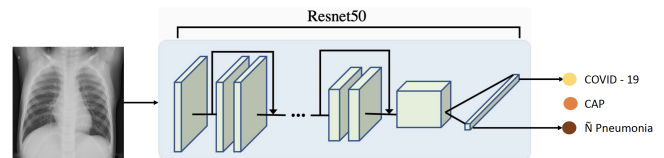


Figura 5. Arquitetura COVNet.

O modelo COVNet foi adaptado para realizar a classificação COVID-19 \times Normal em imagens de Radiografias do tórax segmentadas. As imagens segmentadas são inicialmente redimensionadas para 224×224 , visto que, são as dimensões normalmente utilizadas por modelos convolucionais pré-treinados [32]. Essas mesmas dimensões também foram utilizadas, nos demais modelos implementados. Os modelos convolucionais recebem essas imagens como tensores, na forma (amostras \times dimensões \times canais), em que o número de canais

corresponde ao código de cores RGB, portanto 3. Os conjuntos de treino, validação e teste foram definidos, na proporção 70%, 20% e 10% respectivamente, em uma validação cruzada de Kfold com $k = 10$ [33]. Foram configuradas 50 épocas de treinamento em *batchsize* de 32.

Enquanto modelo de referência, a arquitetura da CNN conta com três camadas, sendo uma convolucional de *kernel* 3×3 e 64 neurônios, uma maxpooling 2×2 e uma totalmente conectada, de saída binária e ativada com *softmax*. Como o SVM de classe única aplica outra abordagem sobre os dados, elevando os dados de entrada para um espaço de recursos de alta dimensão ao invés utilizar sucessivas camadas de representação para esses dados, sua configuração foi ligeiramente diferente.

Dois configurações foram realizadas, aprendendo na classe COVID-19, ou seja, incluindo a classe em um espaço de alta densidade (hiperesfera) e detectando a classe normal como *outliers* e inversamente, aprendendo na classe normal e classificando a outra como novidade. A motivação no uso desse modelo consiste na detecção de anomalias, ou seja em entender como o algoritmo pode aprender com uma determinada classe e como identificar se um novo dado deve ou não pertencer ao grupo.

As imagens foram fornecidas ao modelo no formato *amstras* \times *features*, no qual, cada *features* corresponde a um valor de pixel, resultante da concatenação de cada linha da imagem. Na implementação foi utilizado *kernel* 'RBF' e uma folga na margem proveniente de um *grid search*, busca pelo melhor parâmetro, dentro da validação cruzada.

IV. RESULTADOS

Para ambas as classes o modelo de segmentação gera imagens conforme ilustrado na Figura 6, em que é possível observar que a região segmentada se assemelha à região delimitada pelos pulmões na imagem original. Observa-se ainda, que os ruídos provenientes de artefatos de exposição (relacionados a objetos estranhos presentes nas imagens) não interferem diretamente no processo de segmentação, como mostra a imagem pertencente à classe COVID.

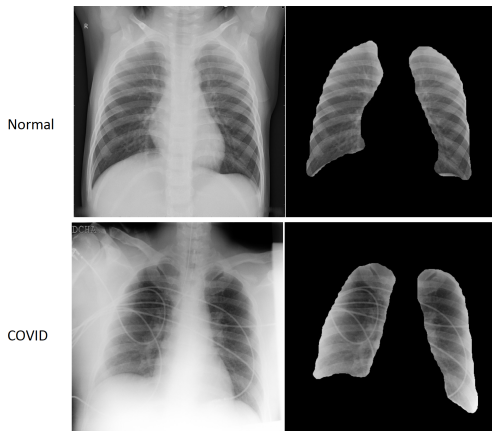


Figura 6. Segmentação nas classes normal e COVID.

Ao utilizar um banco de dados público, cujas imagens são provenientes de diversas fontes, fica evidente que não há um mesmo padrão de aquisição. Na Figura 7, algumas imagens de ambas as classes são apresentadas, nas quais o posicionamento do paciente, o enquadramento da região pulmonar e a qualidade das imagens são divergentes. Essas características interferem no processo de segmentação, uma vez que influenciam o modelo na geração de máscaras que possam ser inconsistentes com a região pulmonar como na Figura 7(a), onde regiões fora da delimitação dos pulmões foram consideradas. A baixa qualidade das imagens pode, ainda, fazer com que o modelo não seja capaz de gerar máscaras, como pode ser visto na Figura 7(b) na classe COVID.

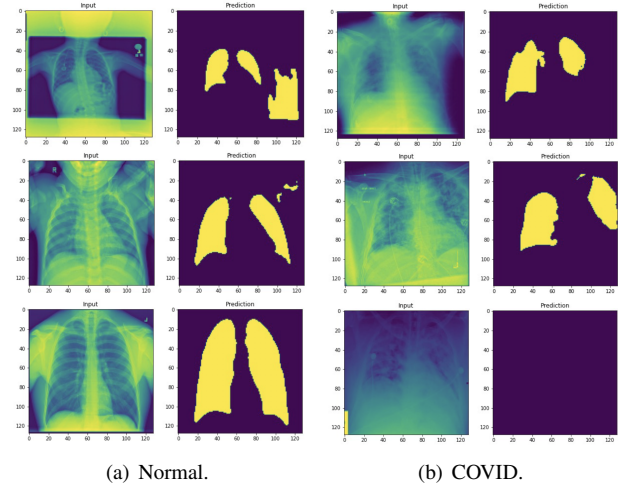


Figura 7. Divergências no padrão de aquisição das imagens em ambas as classes e geração de máscaras inconsistentes com a região pulmonar.

O modelo de segmentação obteve valores de Índice de Jaccard de 93,34% (+0,11) e Coeficiente Dice de 96,53% (+0,06). Esses resultados condizem com o estado da arte de modelos de segmentação de imagens médicas, como mostram os trabalhos de [3] que, em trabalho semelhante, apresentou uma acurácia de 99,2% e um Dice de 92% e [8] que apresenta Índice Jaccard de 94,3% e coeficiente de Dice de 96,94%.

Embora tenha apresentado desempenho significativo, o modelo de segmentação também gerou máscaras inconsistentes com a região pulmonar, que propiciaram à uma segmentação mal sucedida, como mostra a Figura 8. Na imagem pode-se notar, que na classe normal, parte da região pulmonar é perdida na imagem segmentada e na classe COVID há um recorte falho do que representaria à respectiva região pulmonar. Essas imagens, para as quais a segmentação foi falha, impactaram diretamente nos modelos de classificação, uma vez que, permitiram a esses modelos aprender com informações que não caracterizam as classes de interesse.

Durante o treinamento do SVM de classe única, na classe COVID, os parâmetros resultantes do *grid search* foram, uma variável de folga $\nu = 0,1$, coeficiente de kernel $\gamma = 1/\text{número de recursos}$ e kernel RBF. Na etapa de teste foram gerados 102 vetores de suporte. Para o SVM com treinamento

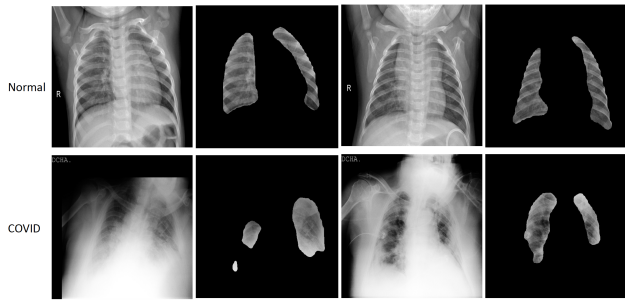


Figura 8. Amostras cuja segmentação foi falha.

na classe normal, foi encontrado o parâmetro $nu = 0,5$, os demais parâmetros semelhantes ao da configuração anterior e foram gerados 494 vetores de suporte.

A Tabela I apresenta o desempenho dos modelos de classificação implementados, na qual pode-se observar diferenças significativas entre os modelos que utilizam técnicas de aprendizado profundo (COVNet e CNN) e o modelo de aprendizado de máquina SVM de classe única. Ao observar a sensibilidade, por exemplo, é possível notar que os modelos COVNet e CNN têm desempenho semelhante, pois apresentaram resultados com valores próximos, $95,91\%(\pm 1,49)$ e $97,03\%(\pm 1,24)$ respectivamente, considerando a barra de erro. O fato do conjunto de dados possuir um número restrito de imagens não permite, à amostra, uma avaliação definitiva quanto a possíveis ganhos para um modelo mais complexo.

Entretanto, o SVM não obteve um valor expressivo para essa medida, quando aprendido na classe COVID-19, pois os $48,9\%$ teve uma variação muito grande, em torno de $\pm 5\%$ expondo falhas na caracterização das classes por parte do modelo. Entretanto, o SVM aprendido na classe normal obteve um desempenho inferior ao dos modelos de aprendizado profundo, uma vez que a sensibilidade foi de $89,72(\pm 2,92)$. É importante destacar, que o SVM de classe única visando à aplicação em COVID-19, deve mapear os pacientes infectados como novidade, portanto o resultado anterior era esperado.

Em termos de especificidade, que consiste na probabilidade de um indivíduo avaliado e normal ter seu teste normal (em outras palavras, negativo) os valores obtidos pelos modelos foram mais próximos, $99,32\%(\pm 0,50)$ para o COVNet, $99,66\%(\pm 0,41)$ na CNN e $98,66\%(\pm 0,61)$ para o SVM (COVID), exceto para o SVM aprendido na classe normal, cujo valor foi de $81,99\%(\pm 2,74)$. Esses resultados mostram a capacidade dos modelos em caracterizar a classe normal.

Quanto as demais medidas, F1 score, Recall e AUC score o melhor desempenho foi obtido com a CNN, cujos valores correspondem respectivamente a $98,40\%(\pm 0,48)$, $99,63\%(\pm 0,48)$ e $98,35\%(\pm 0,51)$.

V. CONCLUSÃO

A ideia deste trabalho consistiu em mostrar um comparativo entre o desempenho de um modelo complexo, de aprendizado profundo, e modelos mais simples a partir da análise de imagens segmentadas de Radiografias do tórax. O propósito

desta análise, é buscar técnicas que possam nortear o desenvolvimento de ferramentas de apoio ao processo de triagem de pacientes com COVID-19. Na segmentação foi utilizado o modelo MultiResUnet, que gera as máscaras de segmentação das respectivas imagens de entrada. A identificação da COVID-19 é realizada pelos modelos COVNet, CNN e SVM (classe única), na tarefa de classificação binária e no contexto de triagem de pacientes.

De um modo geral, os modelos de aprendizado profundo obtiveram resultados mais significativos quando comparados aos resultados obtidos pelo SVM. Porém, entre os dois modelos, a CNN mostrou um desempenho ligeiramente superior ao COVNet mostrando que, ao lidar com os dados de forma similar, o aumento da complexidade não implicou em um desempenho superior na amostra avaliada. Quanto aos resultados obtidos pelo SVM de classe única, ao fazê-lo aprender na classe normal e identificar a classe COVID-19 como *outlier*, ocasionou um desempenho superior ao da configuração inversa, ao observar a acurácia, por exemplo. Ainda que, possam ser melhorados os resultados obtidos pelo SVM, a simplicidade desse modelo deve ser considerada, uma vez que pode ser empregado em tecnologias de menor poder de processamento, como os dispositivos móveis.

Todavia, os valores das medidas para o SVM de classe única aprendido na classe COVID e identificando a classe normal como novidade, possuem uma variação grande, por exemplo a sensibilidade está em torno de 48% , a especificidade de $98,66\%$ e a AUC score próximo de 73% . Essa variação destoante entre as medidas pode estar relacionada à dificuldade do modelo ao interpretar as características que definem a classe COVID, principalmente em imagens cuja a segmentação foi falha.

Como continuidade deste trabalho, melhorias devem ser aplicadas ao modelo segmentação, como um treinamento mais adequado para realizar uma segmentação mais coerente das imagens acometidas por COVID-19. Uma segmentação mais eficaz implicará em uma triagem mais assertiva. Alguns ajustes nos modelos de classificação serão realizadas, o SVM de classe única será revisado no intuito de reduzir as falhas apresentadas durante os testes realizados. Por fim, um número maior de imagens serão utilizadas em novos testes, na tentativa de expor os modelos a uma variabilidade maior da doença e assim poder identificá-la de maneira mais precisa.

AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001. Os autores agradecem ao CNPq, FAPERJ, e ao LPS/UFRJ por todo apoio financeiro e aporte técnico. Agradecem ainda, aos membros e afiliados do grupo de pesquisa do COVID-19 *Radiography Database*, por disponibilizarem dados radiológicos para domínio público. Ao IFPB, pela parceria e troca de conhecimento, gratidão.

Tabela I
DESEMPENHO DOS MODELOS

Medidas	COVNet(%)	CNN (%)	SVM (COVID) (%)	SVM (Normal)(%)
Acurácia	98,51 ($\pm 0,58$)	97,82 ($\pm 1,11$)	73,46 ($\pm 2,71$)	85,72 ($\pm 2,33$)
Sensibilidade	95,91 ($\pm 1,49$)	97,03 ($\pm 1,24$)	48,90 ($\pm 5,26$)	89,35 ($\pm 2,92$)
Especificidade	99,32 ($\pm 0,50$)	99,66 ($\pm 0,41$)	98,66 ($\pm 0,61$)	81,99 ($\pm 2,74$)
F1 Score	97,4 ($\pm 0,66$)	98,40 ($\pm 0,48$)	78,86 ($\pm 1,76$)	85,00 ($\pm 2,42$)
Recall	99,21 ($\pm 0,6$)	99,63 ($\pm 0,48$)	98,66 ($\pm 0,61$)	81,99 (2,74)
AUC score	97,62 ($\pm 0,67$)	98,35 ($\pm 0,51$)	73,78 ($\pm 2,68$)	85,67 ($\pm 2,33$)

REFERÊNCIAS

- [1] W. H. Organization *et al.*, “Who director-general’s opening remarks at the media briefing on covid-19-11 march 2020,” 2020.
- [2] Worldometers, “Covid-19 coronavirus pandemic.” Disponível em: <https://www.worldometers.info/coronavirus/>. Acesso em: 21 de junho 2021, 2021.
- [3] M. J. Hasan, M. S. Alom, and M. S. Ali, “Deep learning based detection and segmentation of covid-19 & pneumonia on chest x-ray image,” in *2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD)*, pp. 210–214, IEEE, 2021.
- [4] L. Lan, D. Xu, G. Ye, C. Xia, S. Wang, Y. Li, and H. Xu, “Positive rt-pcr test results in patients recovered from covid-19,” *Jama*, vol. 323, no. 15, pp. 1502–1503, 2020.
- [5] X. Xu, X. Jiang, C. Ma, P. Du, X. Li, S. Lv, L. Yu, Q. Ni, Y. Chen, J. Su, *et al.*, “A deep learning system to screen novel coronavirus disease 2019 pneumonia,” *Engineering*, vol. 6, no. 10, pp. 1122–1129, 2020.
- [6] A. Amyar, R. Modzelewski, H. Li, and S. Ruan, “Multi-task deep learning based ct imaging analysis for covid-19 pneumonia: Classification and segmentation,” *Computers in Biology and Medicine*, vol. 126, p. 104037, 2020.
- [7] P. Lakhani and B. Sundaram, “Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks,” *Radiology*, vol. 284, no. 2, pp. 574–582, 2017.
- [8] S. Rajaraman, J. Siegelman, P. O. Alderson, L. S. Folio, L. R. Folio, and S. K. Antani, “Iteratively pruned deep learning ensembles for covid-19 detection in chest x-rays,” *IEEE Access*, vol. 8, pp. 115041–115050, 2020.
- [9] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in *International Conference on Medical image computing and computer-assisted intervention*, pp. 234–241, Springer, 2015.
- [10] S. Jégou, M. Drozdal, D. Vazquez, A. Romero, and Y. Bengio, “The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation,” in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pp. 11–19, 2017.
- [11] L. R. Dice, “Measures of the amount of ecologic association between species,” *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.
- [12] M. Blain, M. T. Kassim, N. Varble, X. Wang, Z. Xu, D. Xu, G. Carrafiello, V. Vespro, E. Stellato, A. M. Ierardi, *et al.*, “Determination of disease severity in covid-19 patients using deep learning in chest x-ray images,” *Diagnostic and interventional radiology*, vol. 27, no. 1, p. 20, 2021.
- [13] M. Tan and Q. Le, “Efficientnet: Rethinking model scaling for convolutional neural networks,” in *International Conference on Machine Learning*, pp. 6105–6114, PMLR, 2019.
- [14] T. Rahman, A. Khandakar, Y. Qiblawey, A. Tahir, S. Kiranyaz, S. B. A. Kashem, M. T. Islam, S. Al Maadeed, S. M. Zughaier, M. S. Khan, *et al.*, “Exploring the effect of image enhancement techniques on covid-19 detection using chest x-ray images,” *Computers in biology and medicine*, vol. 132, p. 104319, 2021.
- [15] P. Jaccard, “The distribution of the flora in the alpine zone. 1,” *New phytologist*, vol. 11, no. 2, pp. 37–50, 1912.
- [16] L. Li, L. Qin, Z. Xu, Y. Yin, X. Wang, B. Kong, J. Bai, Y. Lu, Z. Fang, Q. Song, *et al.*, “Using artificial intelligence to detect covid-19 and community-acquired pneumonia based on pulmonary ct: evaluation of the diagnostic accuracy,” *Radiology*, vol. 296, no. 2, pp. E65–E71, 2020.
- [17] C. Zhao, Y. Xu, Z. He, J. Tang, Y. Zhang, J. Han, Y. Shi, and W. Zhou, “Lung segmentation and automatic detection of covid-19 using radiomic features from chest ct images,” *Pattern Recognition*, p. 108071, 2021.
- [18] N. Dey, V. Rajinikanth, S. J. Fong, M. S. Kaiser, and M. Mahmud, “Social group optimization-assisted kapur’s entropy and morphological segmentation for automated detection of covid-19 infection from computed tomography images,” *Cognitive Computation*, vol. 12, no. 5, pp. 1011–1023, 2020.
- [19] A. Jacobi, M. Chung, A. Bernheim, and C. Eber, “Portable chest x-ray in coronavirus disease-19 (covid-19): A pictorial review,” *Clinical imaging*, 2020.
- [20] S. Albawi, T. A. Mohammed, and S. Al-Zawi, “Understanding of a convolutional neural network,” in *2017 International Conference on Engineering and Technology (ICET)*, pp. 1–6, Ieee, 2017.
- [21] A. J. Smola and B. Schölkopf, *Learning with kernels*, vol. 4. Citeseer, 1998.
- [22] A. L. Samuel, “Some studies in machine learning using the game of checkers,” *IBM Journal of research and development*, vol. 3, no. 3, pp. 210–229, 1959.
- [23] X. Zhu and A. B. Goldberg, “Introduction to semi-supervised learning,” *Synthesis lectures on artificial intelligence and machine learning*, vol. 3, no. 1, pp. 1–130, 2009.
- [24] B. Schölkopf, J. C. Platt, J. Shawe-Taylor, A. J. Smola, and R. C. Williamson, “Estimating the support of a high-dimensional distribution,” *Neural computation*, vol. 13, no. 7, pp. 1443–1471, 2001.
- [25] N. Ketkar and E. Santana, *Deep learning with python*, vol. 1. Springer, 2017.
- [26] F. H. Araújo, A. Carneiro, R. R. Silva, F. N. Medeiros, and D. M. Ushizima, “Redes neurais convolucionais com tensorflow: Teoria e prática,” *SOCIEDADE BRASILEIRA DE COMPUTAÇÃO. III Escola Regional de Informática do Piauí. Livro Anais-Artigos e Minicursos*, vol. 1, pp. 382–406, 2017.
- [27] Y. Li, Z. Hao, and H. Lei, “Survey of convolutional neural network,” *Journal of Computer Applications*, vol. 36, no. 9, pp. 2508–2515, 2016.
- [28] M. E. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M. A. Kadir, Z. B. Mahbub, K. R. Islam, M. S. Khan, A. Iqbal, N. Al Emadi, *et al.*, “Can ai help in screening viral and covid-19 pneumonia?,” *IEEE Access*, vol. 8, pp. 132665–132676, 2020.
- [29] N. Ibtehaz and M. S. Rahman, “Multiresunet: Rethinking the u-net architecture for multimodal biomedical image segmentation,” *Neural Networks*, vol. 121, pp. 74–87, 2020.
- [30] D. Rocha, “Multiresunet-lung-segmentation.” Disponível em: <https://github.com/rochadiego/MultiResUNet-lung-segmentation>. Acesso em: 22 de junho 2021, 2021.
- [31] C. Shorten and T. M. Khoshgoftaar, “A survey on image data augmentation for deep learning,” *Journal of Big Data*, vol. 6, no. 1, pp. 1–48, 2019.
- [32] B. Singh and L. S. Davis, “An analysis of scale invariance in object detection snip,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3578–3587, 2018.
- [33] R. Kohavi *et al.*, “A study of cross-validation and bootstrap for accuracy estimation and model selection,” in *Ijcai*, vol. 14, pp. 1137–1145, Montreal, Canada, 1995.