

Detecção de COVID-19 e Avaliação de Nível de Severidade: Uma abordagem com BPPC e Redes Neurais Artificiais Rasas

Afonso U. Fonseca*, Juliana P. Felix*, Gabriel S. Vieira**
Deborah S. A. Fernandes* Fabrizzio Soares*

* Instituto de Informática, Universidade Federal de Goiás, GO, Brasil

** Instituto Federal Goiano - Urutaí, GO, Brasil

Abstract: Identifying COVID-19 on chest radiographic images (CXR) remains an essential task in patient tracking and follow-up. Several initiatives have worked on this task, and although the results presented so far are encouraging, few are reflected in applications adopted in clinical environments. Thus, understanding that there is still a field of research in this area, this work presents a method based on feature extraction by using Binary Patterns of Phase Congruency (BPPC) on segmented CXR images. The objective of this work is to be a mild and fast alternative for the automatic detection of COVID-19 and its severity levels while requiring few computational resources. To do this, radiomic features are extracted from CXR images, a selection process based on SVM is used, and two models of shallow *Feedforward* networks are built. The results presented far surpasses previous works, with an average accuracy for COVID-19 detection of 98.24%, and 94.74% in identifying PCR+ images taken from people without any infectious marks and Normal diagnosis. A second model is also presented, in which the task is to classify CXR images of COVID-19 in different severity levels, and the presented AUC is 98.05%. The solution's high performance makes it a viable option as a computer-aided diagnostic tool, which can represent a significant gain in the speed and accuracy of COVID-19 diagnosis and in identifying the severity of the disease.

Resumo:

A identificação de COVID-19 em imagens radiográficas de tórax (RXT) continua sendo uma importante tarefa no rastreamento e acompanhamento de pacientes. Diversas iniciativas têm trabalhado essa tarefa e embora os resultados apresentados até o momento sejam animadores, poucos se reverberam em aplicações adotadas em ambientes clínicos. Assim, entendendo que ainda há campo de pesquisa nesta área, este trabalho apresenta um método baseado em extração de características por Padrões Binários de Fase de Congruência (BPPC do inglês *Binary Pattern of Phase Congruency*) em imagens RXT segmentadas. O objetivo do trabalho é ser uma alternativa leve e rápida para a detecção automática de COVID-19 e seus níveis de severidade, ao passo que exija poucos recursos computacionais. Para isso, são extraídas características radiômicas de imagens RXT, um processo de seleção baseado em SVM é utilizado e dois modelos de redes rasas *Feedforward* são construídas. Os resultados apresentados superam com grande margem trabalhos anteriores, com acurácia média para detecção de COVID-19 de 98,24%, e 94,74% em imagens sem quaisquer marcas de infecção com diagnóstico Normal e exame PCR+. Um segundo modelo também é apresentado, nele a tarefa é classificar imagens RXT de COVID-19 em diferentes níveis de severidade, e a AUC apresentada é de 98,05%. O alto desempenho da solução a torna uma opção viável como ferramenta de diagnóstico auxiliado por computador, o que pode representar um ganho significativo na velocidade e precisão do diagnóstico da COVID-19 e na identificação do grau de severidade da doença.

Keywords: COVID-19 Detection; Severity Level; Feature Extraction; Binary Pattern of Phase Congruency; Feedforward Neural Network; Shallow Artificial Neural Network.

Palavras-chaves: Detecção de COVID-19; Nível de Severidade; Extração de Características; Rede Neural Feedforward; Rede Neural Artificial Rasa.

1. INTRODUÇÃO

Nos últimos dois anos, o mundo tem passado por uma das piores catástrofes sanitárias de todos os tempos, a pandemia da SARS-CoV-2 (COVID-19). Embora o avanço da cobertura vacinal em vários países venha nos trazendo de volta à 'normalidade', não podemos ainda dizer que o mundo está livre dessa doença. O Brasil, ainda em julho de 2022, registra média diária de cerca de 200 mortes, acumulando mais de 670 mil vidas perdidas, enquanto no mundo o número diário fica acima de 6 mil mortes (Han-

nah Ritchie and Roser, 2020). Esses números demonstram que decorridos mais de dois anos de pandemia a batalha ainda não foi vencida e que ainda pode persistir por algum tempo (Kissler et al., 2020).

Portanto, para frear a disseminação e reduzir óbitos a identificação precoce da doença e o acompanhamento de sua evolução continuam sendo as melhores opções. Nesse sentido, as radiografias de tórax (RXT) vêm mostrando ser uma escolha de primeira ordem para atender essas tarefas. Pesquisadores de diferentes nacionalidades têm investigado técnicas de inteligência artificial associadas às RXTs para auxiliar em diagnósticos mais precisos e

* Afonso U. Fonseca: Corresponding author: (afonsoueslei@ufg.br)

rápidos, e os resultados, embora animadores, pouco se traduzem em aplicações para uso clínico [Kallianos et al. \(2019\)](#). Logo, pesquisas e investigações devem avançar até que seus resultados forneçam níveis de confiabilidade e segurança adequados para adoção em ambientes clínicos.

Atualmente, grande parte das pesquisas que aplicam inteligência artificial em imagens RXT concentram-se no diagnóstico através da classificação de casos normais e de infecção por COVID-19. Geralmente, utilizando-se de imagens RXT públicas coletadas de diferentes fontes e sem qualquer definição de coorte ou protocolo conhecido. Por outro lado, em menor número, outras utilizam dados privados com coorte e protocolo específicos, mas de difícil checagem por terceiros ([Wynants et al., 2020](#); [Ismael and Şengür, 2021](#); [Jain et al., 2021](#); [Khan et al., 2020](#); [Basu et al., 2020](#); [López-Cabrera et al., 2021](#); [Roberts et al., 2021](#)). Logo, embora importantes, estas pesquisas podem ter vieses de seleção e amostragem, colocando em suspeição e com alto risco de previsão seus modelos construídos.

Assim, para melhorar a eficiência, segurança e confiabilidade destes sistemas temos entre outras alternativas:

- a construção e compartilhamento de conjuntos de dados bem formatados, seguindo protocolos bem definidos e com metadados associados;
- a combinação de dados demográficos entre coortes;
- a avaliação de vieses de seleção e
- as análises de robustez e sensibilidade dos modelos construídos.

Em uma tentativa de atender a esses critérios, [Tabik et al. \(2020\)](#) construíram um conjunto de dados (COVIDGR-1.0) equilibrado e homogêneo para desmitificar os níveis de sensibilidade e robustez dos modelos de classificação de COVID-19. Esse conjunto foi criado em parceria com o Hospital Universitário Clínico San Cecilio (Granada, Espanha) e contém imagens com todos os graus de severidade da doença, seguindo o *score* de Avaliação Radiográfica de Edema Pulmonar (RALE) ([Warren et al., 2018](#)). Além disso, os autores, como prova de conceito, criaram uma rede, que chamaram COVID-SDnet, para melhorar a capacidade de generalização de modelos de classificação de COVID-19, e alcançaram resultados satisfatórios para diferentes níveis de severidade da doença. A Figura 1 ilustra a escala em que o grau de severidade é definido pela quantidade de regiões da radiografia com características de consolidação e opacidades em vidro fosco.

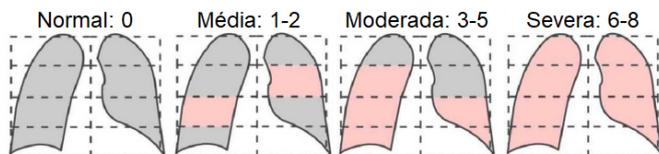


Figura 1. Escala da Severidade: A estratificação da gravidade radiológica do COVID-19 ([Tabik et al., 2020](#))

[Panetta et al. \(2021\)](#), um pouco depois, propuseram um novo descritor de características baseado em padrões Fibonacci-p dependente de forma, que usaram para discriminar casos de COVID-19. Os autores, utilizando diferentes classificadores de aprendizagem de máquina para o mesmo conjunto usado em ([Tabik et al., 2020](#)), relataram resultados similares na classificação entre casos COVID-19 e normais, porém não avaliaram a sensibilidade para os diferentes níveis de severidade da doença. Outras estratégias de extração de características, combinadas a aprendizagem de máquina, também foram usadas na detecção de COVID-19 e apresentaram resultados consistentes [Elaziz et al. \(2020\)](#); [Gil et al. \(2020\)](#); [Fonseca et al. \(2021\)](#).

Em [Gazda et al. \(2021\)](#) uma rede neural convolucional profunda auto-supervisionada baseada no modelo pré-treinado ResNet-50 foi usada para extrair características, depois essas características foram usadas para alimentar um classificador de regressão logística otimizado na tarefa de classificação de doenças respiratórias. Utilizando uma estratégia similar, [Öksüz et al. \(2021\)](#) também lidaram com a extração de características profundas de diversos modelos de redes pré-treinadas, acrescentando um método de seleção de características *wrapper* baseado em pontuações laplacianas com uma Máquina de Vetores de Suporte (SVM) como avaliador. Em comparação ao trabalho de [Tabik et al. \(2020\)](#), alcançaram desempenho superior tanto na detecção da COVID-19 quanto na discriminação dos níveis de severidade.

Uma abordagem de rede de atenção adaptativa foi apresentada por [Lin et al. \(2021\)](#). Essa rede é formada por dois componentes principais: uma ResNet deformável adaptável e um codificador baseado em atenção. A primeira ajusta de forma adaptativa os campos receptivos para aprender representações das características conforme a forma e a escala das regiões infectadas, e o último modela interações não-locais por mecanismo de auto-atenção que aprende informações de contexto ricas para detectar as regiões da lesão com formas complexas. O modelo sofisticado adotado pelos autores é completo com uma rede *Feed-Forward* (FFN) na saída que detecta os casos de COVID-19 e seus níveis de severidade.

Neste trabalho é proposta uma solução para auxiliar na detecção de COVID-19 e, adicionalmente, também definir o grau de severidade da doença. A proposta combina segmentação dos pulmões de imagens RXT, extração de características por Padrões Binários de Fase de Congruência (BPPC) e ranqueamento e seleção das características mais relevantes por SVM para alimentar redes de classificação FFN (*Feed-Forward*). São criados modelos para dois esquemas de classificação: a) pacientes saudáveis contra doentes (Normal x COVID-19) e, b) nível de severidade da COVID-19 (Normal com PCR+, Leve, Moderado e Severo). A proposta apresentada é uma solução de alta sensibilidade, viável para ambientes e equipamentos com recursos computacionais limitados e apresenta desempenho superior quando comparado aos trabalhos relacionados.

As principais contribuições de pesquisa deste artigo podem ser resumidas da seguinte forma:

- Extração de características radiômicas para detecção de COVID-19 por meio de BPPC;
- Segmentação dos pulmões para minimizar ruídos ou influência de regiões não relacionadas à patologia;
- Ranqueamento e seleção de características com alto poder discriminante para detecção e classificação da severidade da COVID-19;
- Construção de um modelo de rede *Feed-Forward* para classificação de COVID-19 com alta sensibilidade;
- Construção de um modelo de rede *Feed-Forward* capaz de distinguir os níveis de severidade da COVID-19.

O artigo está organizado da seguinte forma. Os materiais e métodos usados na elaboração do método proposto são detalhados na Seção 2. Os resultados e a discussão são explicados na Seção 3. Por fim, o artigo é concluído na Seção 4 com apresentação de possíveis trabalhos futuros.

2. MATERIAIS E MÉTODOS

Nesta seção são apresentados algoritmos, técnicas e materiais usados nas etapas do método proposto. Na Figura 2

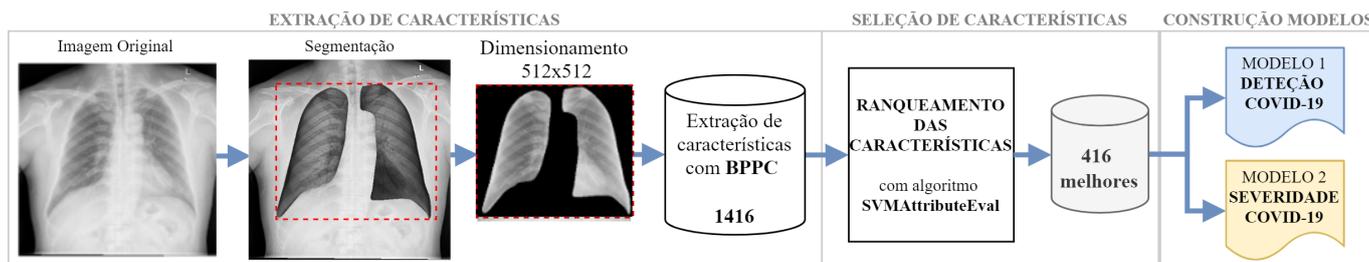


Figura 2. Fluxograma do método proposto para os dois modelos de classificação.

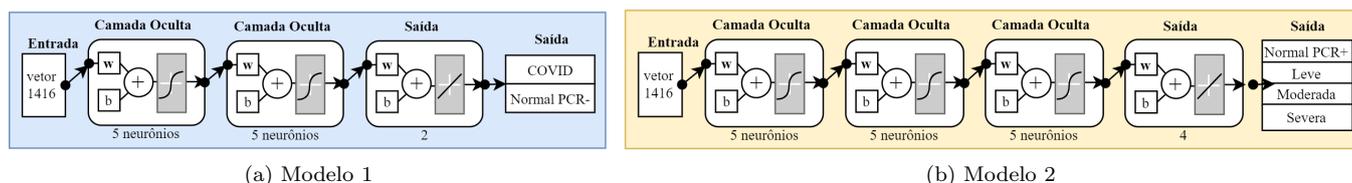


Figura 3. Redes *Feed-Forward* em detalhes. (a) Detecção de COVID e (b) Classificação do Nível de Severidade.

é mostrada uma visão geral dessas etapas, e nas Figuras 3-(a) e 3-(b), detalhes de cada um dos modelos.

2.1 Conjuntos de Dados

Este estudo utiliza o conjunto COVIDGR-1.0, criado e disponibilizado pelo trabalho de [Tabik et al. \(2020\)](#). O conjunto foi escolhido por permitir uma melhor avaliação, em diferentes níveis de severidade, do comportamento e da real sensibilidade dos modelos de inteligência artificial de detecção de COVID-19, enquanto possibilita a construção de modelos menos enviesados, dado que é mais homogêneo e equilibrado.

Neste conjunto, todas as imagens foram obtidas por um mesmo equipamento RXT e usando as mesmas configurações. O conjunto contém 852 imagens RXT organizadas em duas classes, distribuídas em 426 casos positivos para COVID-19, confirmados por exame padrão ouro de Reação em cadeia da polimerase – transcriptase reversa (RT-PCR+), e 426 casos negativos. As imagens foram tomadas em projeção posteroanterior (PA) e os casos positivos divididos segundo a gravidade da infecção em 76 normais, 100 leves, 171 moderadas e 79 severas, seguindo a pontuação RALE conforme apresentado na Figura 1. Na Tabela 1 é apresentada uma descrição sumarizada do conjunto.

Tabela 1
Descrição sumarizada do conjunto COVIDGR-1.0

Classe	COVID-19	NORMAL
Projeção das imagens	PA	PA
Dimensão das imagens	[514-3840] × [466-300]	[492-3840] × [464-2877]
Número imagens (%)	426 (1,00)	426 (1,00)
Homens (%)	236 (0,55)	187 (0,44)
Normal RTPCR+ (%)	76 (0,18)	n/a
Leve infecção (%)	100 (0,23)	n/a
Moderada infecção (%)	171 (0,40)	n/a
Severa infecção (%)	79 (0,19)	n/a

n/a: Não se aplica.

2.2 Segmentação dos pulmões

O problema da segmentação pulmonar já foi enfrentado e abordado com sucesso em [Candemir and Antani \(2019\)](#); [Mittal et al. \(2017\)](#); [Pietka \(1994\)](#). Segmentar apenas os pulmões, descartando o resto da RXT, potencialmente elimina possíveis fontes de viés, como, por exemplo, a

presença de dispositivos médicos (normalmente correlacionados a pacientes doentes) ou anotações incorporadas na varredura. Outra questão é o tamanho e posicionamento do paciente que pode implicar na inclusão de mais partes do corpo, como cabeça, braços, pescoço e estômago.

Como todas as imagens do conjunto COVIDGR-1.0 são RXT frontais e incluem regiões externas aos pulmões, que não estão relacionadas à detecção da COVID-19, realizou-se a segmentação dos pulmões para reduzir a influência destas regiões externas presentes nas imagens, mas não relacionadas à patologia da COVID-19, e que poderiam, portanto, distorcer o resultado da sua detecção.

A abordagem utilizada para realização desta tarefa é um algoritmo de segmentação baseado em atlas proposto por [Candemir et al. \(2013\)](#). O atlas é um conjunto de RXTs de vários pacientes com os limites pulmonares delimitados por especialistas. O algoritmo executa primeiro uma busca entre os modelos para encontrar aqueles com a forma do pulmão mais semelhante ao RXT do paciente, em seguida uma deformação e realizada nos modelos selecionados usando o algoritmo de registro de fluxo SIFT ([Liu et al., 2010](#)). Então, o modelo do pulmão específico do paciente é definido como a média dos modelos registrados. Por fim, o sistema combina os valores de intensidade da RXT e o modelo definido com uma função objetivo para definir a segmentação final.

Executado o algoritmo para segmentar as RXTs obtemos como uma das saídas uma máscara que corresponde a região delimitada dos pulmões. Então, para garantir que partes dos pulmões não sejam perdidas neste processo de segmentação, uma operação de dilatação desta máscara é realizada com elemento estruturante em forma de disco e tamanho 25. Ao final a máscara dilatada é então usada para definir a região pulmonar de interesse (ROI, *region of interest*) que será redimensionada para 512 × 512 pixels e na sequência usada como entrada para etapa de extração de características. A Figura 4 mostra os passos desta etapa.

2.3 Congruência de Fase

Os modelos de congruência de fase (PC do inglês *Phase Congruency*) são importantes descritores de padrão, pois fornecem uma medida que é independente da magnitude geral do sinal, tornando-o invariável às variações na iluminação e contraste da imagem, isso porque, ao contrário de detectores de borda focados em detectar mudanças

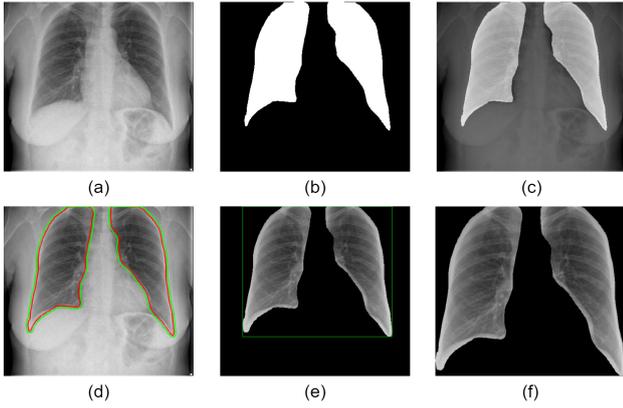


Figura 4. Segmentação etapas. (a) RXT do paciente; (b) máscara gerada algoritmo baseado em atlas; (c) sobreposição da máscara na RXT; (d) Em **vermelho** fronteiras da máscara da segmentação pelo algoritmo baseado em atlas, em **verde** fronteiras da máscara dilatada com elemento estruturante; (e) pulmões segmentados e ROI limitada e (f) Resultado do processo de segmentação a ser usado na etapa de extração de características.

bruscas de intensidade, eles captam pontos de ordem no espectro de fase. Esse atributo dá a esses modelos, alta capacidade de descrever e distinguir imagens, desejável em modelos cujo objetivo é separar amostras entre classes (Kovesi et al., 1999).

No cálculo da PC de uma imagem, é necessária a informação da sua frequência localizada espacialmente, normalmente realizada pela transformada *wavelet* (Kovesi et al., 1999).

Seja I a imagem RXT e M_n^e e M_n^o as *wavelets* simétricas par e ímpar na escala n respectivamente. Então a resposta de cada par de filtros de quadratura forma um vetor de resposta dado por:

$$[e_n(x), o_n(x)] = [I(x) * M_n^e, I(x) * M_n^o]. \quad (1)$$

As informações de amplitude e fase da transformada em uma determinada escala *wavelet* são formuladas como:

$$A_n(x) = \sqrt{e_n(x)^2 + o_n(x)^2}. \quad (2)$$

$$\varphi_n(x) = \text{atan2}(e_n(x), o_n(x)). \quad (3)$$

Então a PC é expressa como a energia local normalizada pela soma dos componentes de amplitude de *Fourier* como:

$$PC(x) = \frac{E(x)}{\sum_n A_n(x)}. \quad (4)$$

onde $E(x)$ é a energia local definida como:

$$E(x) = \sqrt{F(x)^2 + H(x)^2}. \quad (5)$$

$$F(x) = \sum_n e_n(x). \quad (6)$$

$$H(x) = \sum_n o_n(x). \quad (7)$$

Levando em consideração o ruído, característico do equipamento usado na geração da RXT, a expressão de PC pode ser modificada da seguinte forma:

$$PC(x) = \frac{\lfloor E(x) - T \rfloor}{\sum_n A_n(x) + \varepsilon}. \quad (8)$$

onde o símbolo $\lfloor \rfloor$ denota que a quantidade incluída não pode ser negativa. T é um limiar de ruído que é determinado a partir das estatísticas da resposta do filtro

aos dados calculados conforme descrito em Kovesi et al. (1999) e ε é um pequeno deslocamento para evitar a divisão por zero.

Para detectar as características da imagem RXT em várias orientações, um banco de filtros orientados pode ser projetado. Neste artigo, foram utilizados 8 direções e assim como realizado por (Kovesi et al., 1999) as *wavelets* logarítmicas de *Gabor*.

2.4 Padrões Binários de Fase de Congruência

Padrões Binários de Fase de Congruência (BPPC do inglês *Binary Pattern of Phase Congruency*) são uma adaptação aos Padrões Binários Locais (LBP do inglês *Local Binary Pattern*), onde o valor de intensidade necessário para calcular o LBP tradicional é substituído pelo valor PC no *pixel* correspondente. Como a imagem do PC descreve diferentes categorias de características, o código LBP que descreve cada categoria de característica é projetado para ser diferente.

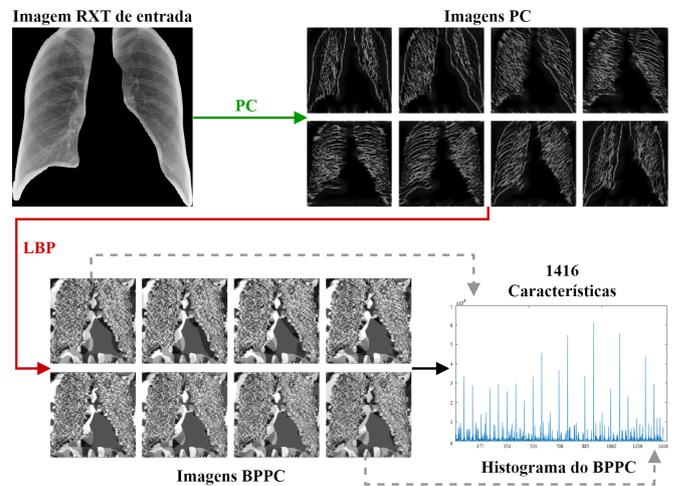


Figura 5. Visão geral do processo de extração de características por BPPC.

O primeiro passo no BPPC é encontrar o PC da imagem em diferentes orientações. Considerando que temos 8 direções dividindo o plano de frequência uniformemente, o conjunto resultante de características da imagem de entrada é então obtido pela concatenação dos histogramas de cada imagem BPPC numa direção. Para cada histograma, o mesmo número adotado em (Turan and Lam, 2018) de 177 faixas foi utilizado, de modo que, o vetor final tenha 1416 características. O processo geral do método BPPC proposto é mostrado na Figura 5.

2.5 Ranqueamento e Seleção de Características

O principal objetivo do processo de seleção e classificação de características é reduzir o número de variáveis de entrada em um modelo de aprendizado. Com a redução de variáveis busca-se otimizar o modelo eliminando características irrelevantes que podem gerar ruído, o que resulta em maior precisão e menor tempo para desenvolvimento e treinamento do modelo e, por consequente, o uso mais eficiente dos recursos computacionais. Outro benefício desse processo é que ele ajuda a evitar o problema de sobreajuste do modelo (*overfitting*), pois a manutenção de características irrelevantes pode resultar em modelos com fraco poder de generalização Ng (1998).

Neste trabalho, utilizou-se como solução para esta tarefa o algoritmo *SVMAAttributeEval* Guyon et al. (2002), que

avalia o valor de uma característica usando um classificador SVM, que então a classifica pelo quadrado do peso atribuído pelo SVM. Se temos um modelo de múltiplas classes, então o algoritmo ranqueia as características para cada classes separadamente usando a estratégia de um contra todos (OvA do inglês *One-vs-All*) e, em seguida “negociando” do topo de cada pilha para obter um ranqueamento final. Os valores dos parâmetros usados pelo algoritmo neste trabalho são dados a seguir.

- Limiar de eliminação constante = 0,
- Filtragem usada pelo SVM = normalização dos dados de treinamento,
- Tolerância $T = 1, 0^{-10}$,
- Complexidade $C = 1, 0$,
- Epsilon $P = 1, 0^{-25}$,
- Taxa constante eliminação de característica = 0,
- Taxa percentual eliminação de característica = 0 e
- Checagem do avaliador = Sim.

Realizado o ranqueamento das características, um limiar para o número de características a ser usado nos dois modelos propostos é definido por um processo incremental de remoção de características em blocos de 100 (partindo das características menos relevantes) seguido pela avaliação da acurácia do modelo. Inicia-se usando o número total de características e mede-se a acurácia do modelo, continuando pela remoção de características enquanto for mantida ou aumentada esta acurácia inicial. O processo termina quando a remoção de um bloco resulta em uma menor acurácia. Assim, as características finais são definidas com o conjunto restante do último passo de remoção que obteve a maior acurácia para o modelo. Neste trabalho, este número de característica foi de 416.

2.6 Divisão dos Dados e Avaliação do Modelo

Validação cruzada é uma técnica para avaliar a generalização de um modelo, com base em um conjunto de dados (Kohavi et al., 1995). Esta técnica é amplamente utilizada em problemas onde o objetivo da modelagem é a previsão. Neste trabalho, foi utilizado esse processo com divisão igual a 10 para cada um dos modelos propostos. É fundamental entender que essa validação cruzada é um *loop* externo independente. Já para avaliação interna dos modelos, em cada uma das 10 rodadas de validação, outro processo é adotado. Nesta avaliação interna, o conjunto é dividido na seguinte proporção: 70%, 15% e 15% para treino, validação e teste, respectivamente.

2.7 Redes Feed-Forward

Redes neurais *Feed-Forward* (FFN) foram as primeiras e mais simples categorias de redes neurais artificiais desenvolvidas (Schmidhuber, 2015). Nestas redes, as conexões entre nós não formam um ciclo, de modo que as informações se movem apenas em uma direção – para a frente – partindo dos nós da camada de entrada, e passando através dos nós das camadas ocultas até os nós da camada de saída (Zell, 1994).

Neste estudo, optou-se por estas redes por sua grande capacidade de previsão, independente das informações de distribuição de probabilidade dos dados. Além disso, estas redes têm sido aplicadas com sucesso em uma ampla gama de aplicações médicas, incluindo a doença de Alzheimer (Jha et al., 2017), doença renal crônica (Almansour et al., 2019), detecção de câncer de pulmão (Arulmurugan and Anandakumar, 2018) e previsão de surtos de COVID-19 (Aljaaf et al., 2021). Sua estrutura paralela e a capacidade de melhorar o desempenho pela experiência explicam a eficiência destas redes.

Como mostrado nas Figuras 3-(a) e 3-(b), duas redes foram consideradas neste trabalho. A primeira, construída com duas camadas ocultas, cada qual com 5 neurônios, é usada para diagnosticar casos de COVID-19 entre casos normais (saudáveis/controle). A segunda rede, usada para classificar os níveis de severidade dos casos COVID-19, é construída com uma estrutura similar, agora com três camadas ocultas e igual número de neurônios. Essa estrutura reduzida garante às redes um rápido treinamento, baixo consumo de recursos computacionais e alto desempenho. Para ambas as redes, com exceção do número de camadas ocultas e a camada de saída, os demais parâmetros foram definidos igualmente – com função de treinamento *Levenberg-Marquardt backpropagation* (Sapna et al., 2012) para atualização dos pesos (W) e viés (B); número máximo de épocas igual a 50, e máximo de falhas de validação igual a 5 para evitar sobreajuste.

3. ANÁLISE DOS RESULTADOS E DISCUSSÃO

Nesta seção, o método proposto é avaliado em um conjunto de métricas de desempenho que incluem acurácia (ACC), análise da curva de características operacionais do receptor (curva ROC) e análise de regressão, entre outras. Uma análise dos resultados é trazida, e com ela uma discussão do potencial do método para cada um dos modelos. Além disso, os resultados alcançados pelo Modelo 1 são comparados com trabalhos da literatura, que realizaram a mesma tarefa no mesmo conjunto COVIDGR-1.0, usando diversas métricas de desempenho, incluindo estatística *Kappa*. Foi também realizada uma estimativa para recuperar os valores ausentes dos trabalhos, para que a comparação fosse o mais justa possível, (veja Tabela 2).

Na avaliação de acurácia do Modelo 1, referente ao nível de severidade da infecção COVID-19, seguiu-se o mesmo cálculo usado em (Tabik et al., 2020) para permitir comparações com outros trabalhos conforme mostra Tabela 3.

Adicionalmente, o Modelo 2 foi criado no sentido de avaliar o poder do método em classificar imagens RXT entre os diferentes níveis de severidade da infecção. Dessa forma, uma avaliação diferente da realizada no Modelo 1 é realizada em que se observava apenas a sensibilidade num esquema binário de classificação.

3.1 Desempenho das Redes Feed-Forward

As imagens da Figura 7-(a) e Figura 7-(b) trazem as curvas de Erro Médio Absoluto no processo de treinamento (validação interna) do Modelo 1 e Modelo 2, respectivamente. A Figura 8 mostra a representação desses erros na distribuição do histograma em 20 faixas. Essas imagens comprovam a eficiência dos modelos no sentido da rápida convergência do treinamento e na baixa taxa de erro mostrados na primeira figura, e na concentração desses erros nas faixas próximas ao erro zero da segunda figura.

Ainda, outra informação que ajuda a entender o desempenho dos modelos é a evolução nos valores dos parâmetros de treinamento, conforme mostrado na Figura 6-(a) e Figura 6-(b), que também justificam a adoção do reduzido número de pontos de verificação, para garantir melhor generalização dos modelos.

3.2 Performance do Modelo 1: Detecção de COVID-19

Conforme observado na Figura 9-(a), o método apresenta boa capacidade tanto na identificação dos casos normais (Especificidade/TNR) com apenas 6 erros, quanto para o diagnóstico correto dos casos COVID-19 (Sensibilidade/TPR) com 9 erros, dos quais apenas 2 são de

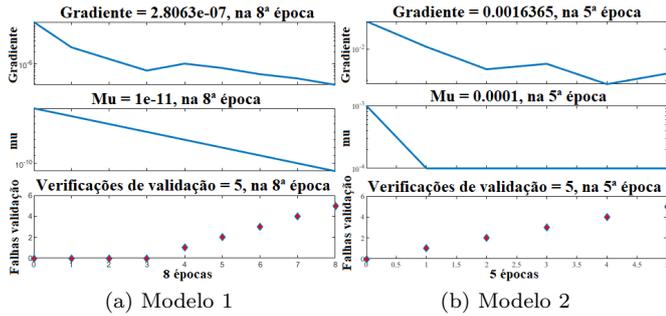


Figura 6. Evolução dos parâmetros de treinamento dos modelos na validação interna, dados do último passo da validação cruzada.

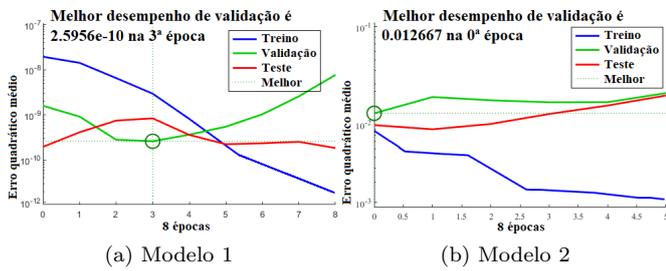


Figura 7. Curvas de Erro dos modelos na validação interna, dados do último passo da validação cruzada.

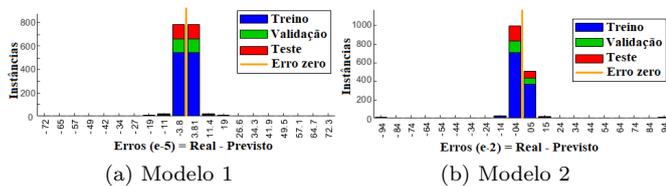


Figura 8. Histograma de Erro na validação interna, dados do último passo da validação cruzada.

RXT com níveis de infecção Moderada e Severa, confirma ainda essa análise, a leitura da Figura 9-(b), curvas ROC do modelo e os valores de AUC superiores à 0,998%.

Ainda sobre o do Modelo 1, a Figura 9-(c) apresenta sua análise de regressão, que demonstra o bom ajuste da rede na resposta adequada dos seus alvos, com saídas sobrepondo perfeitamente suas metas e com coeficiente de correlação igual a 1 ($R = 1$).

Finalizando a análise do Modelo 1, uma comparação com estudos da literatura, que abordaram a mesma tarefa sobre o mesmo conjunto de dados é trazida, primeiro na Tabela 2, onde observa-se que a proposta supera as abordagens anteriores, com grande margem em todas as métricas avaliadas, com a vantagem de usar um modelo simples, que requer pouco poder de processamento e memória. Segundo, na Tabela 3 onde é mostrada a resposta de sensibilidade aos diferentes níveis de severidade, novamente o método apresenta extraordinário desempenho, com especial destaque para o nível Normal PCR+ em que o método atinge 94,74%. Importante dizer que aqui no Modelo 1 é executada uma classificação binária, portanto a sensibilidade ao nível de severidade se dá pela resposta à detecção da COVID-19 e ao posterior cruzamento da informação do nível de severidade que cada imagem COVID-19 pertence.

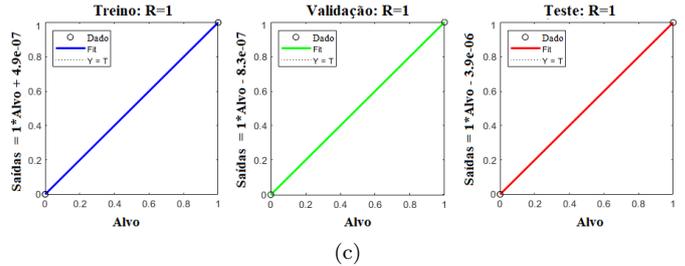
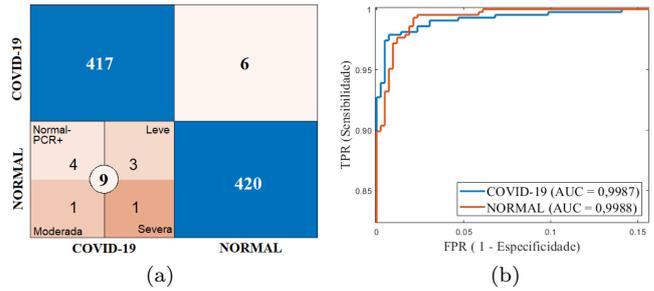


Figura 9. Modelo 1: Detecção de COVID-19, (a) Matriz de Confusão média estratificada por níveis de severidade, (b) Curvas ROC em zoom, e (c) Regressão do modelo.

3.3 Performance do Modelo 2: Classificação do Nível de Severidade da COVID-19

O Modelo 2 aborda uma nova tarefa, a capacidade da rede distinguir imagens COVID-19 em seus diferentes níveis de severidade. Nesta tarefa de multi-classe, portanto, a rede é alimentada apenas com imagens da classe COVID-19 rotuladas de acordo com seus níveis de severidade, com uma avaliação no estilo todos contra todos.

A Figura 10 apresenta os resultados para esta tarefa, respectivamente, com a matriz de confusão, as curvas ROC e a análise de regressão.

Iniciando pela matriz de confusão na Figura 10-(a), dois aspectos chamam a atenção, o primeiro relativo à precisão do método observada pelos valores da diagonal principal (em azul), o segundo pelos 6 erros entre os níveis Leve e Severo, e 1 erro entre os níveis Leve e Normal-PCR+, onde o esperado era encontrar menores erros nos extremos de severidade (bordas da matriz) e maiores entre severidades mais próximas (paralelas da diagonal principal). A hipótese é que as razões estejam em: i) desbalanceamento das classes, podendo o modelo estar favorecendo as classes majoritárias e ii) processo de segmentação que em alguns casos pode excluir partes dos pulmões, principalmente as mais comprometidas pela COVID-19, vez que o processo de segmentação não é totalmente preciso.

Em relação à Figura 10-(b), as linhas das curvas ROC mostram que o modelo tem bom desempenho com independência à distribuição de classes e apresenta valores de AUC superiores as 97% para todas as classes.

Sobre a análise de regressão deste modelo, os resultados mostram a boa correlação entre as metas e os resultados, com valores médios de $R=0,97423$, inclinação = 0,95 e interceptação = 0,013.

Completando a análise do Modelo 2, na última linha da Tabela 2 estão presentes os valores médios alcançados nas diferentes métricas de avaliação, com destaque para o valor de especificidade média (TNR) de 96,79%, que está em consonância ao valor pelo Modelo 1 na Tabela 3.

Tabela 2
Comparativo dos Resultados com Estudos de Referência para o Esquema de Detecção de COVID-19.

Estudo	≤NR	Método	ACC ($\mu \pm \sigma$)	TPR	TNR	PPV	F-1	MCC	AUC	Kappa
Tabik et al. (2020)	M	COVID-SDNet	81,00±2,87	76,8	85,20	84,23	80,07	-	-	60,00*
Panetta et al. (2021)	M	FS(Padrões Fibonacci), SVM	75,11±1,76	72,65	77,72	76,41	74,03	-	-	-
Öksüz et al. (2021)	605	FE(Xception), FS(Pontuação Laplaciana), SVM	76,06±2,00	78,64*	73,47*	74,78*	76,66*	-	-	55,71*
Gazda et al. (2021)	128	FS(Self-supervised Resnet-50)	79,50	83,50	75,60	-	-	-	86,60	-
Lin et al. (2021)	M	AAANet (apenas CxN)	84,30*	83,15*	85,54*	86,05*	84,57*	-	-	70,59*
Modelo 1	416	FE(BBPC), FS(SVM), FFN	98,24±3,80	97,89	98,59	98,58	98,23	96,48	99,87	96,53
Modelo 2 (μ)	416	FE(BBPC), FS(SVM), FFN	90,55±15,05	90,81	96,79	89,73	90,26	86,97	98,05	84,37

≤NR Menor número de características usadas pelos modelos avaliados, em que M representa conjunto de milhares de características, * valores estimados a partir dos resultados disponíveis no trabalho citado. FE: Processo de extração de características, FS: Processo de seleção de características, μ : média, σ : Desvio Padrão, ACC: Acurácia, TPR: Sensibilidade, TNR: Especificidade, PPV: Precisão, F1: Medida F-1, MCC: Coeficiente de correlação *Mattheus*, AUC: Área sobre a curva, e Estatística *Kappa*.

Tabela 3
Comparativo de Acurácia (ACC) do Modelo 1 para ao Esquema de Nível de Severidade.

ESTUDO	NÍVEL DE SEVERIDADE			
	Normal PCR+	Leve	Moderada	Severa
Tabik et al. (2020)	-	46,00	85,38	97,22
Öksüz et al. (2021)	36,92	60,51	85,95	96,36
Oh et al. (2020) ^λ	50,00	68,42	91,43	100,00
Lin et al. (2021)	60,00	73,68	91,43	100,00
Modelo 1	94,74	97,00	99,42	98,73

λ: Valores obtidos através do trabalho de Lin et al. (2021), em negrito melhores resultados.

dos pulmões por Padrões Binários de Fase Congruência (BBPC). Um processo de ranqueamento e seleção através de avaliação de características por SVM foi utilizado e então usado para alimentar redes rasas *Feed-Forward* com as características mais relevantes. Dois modelos destas redes foram construídos e avaliados em um esquema de validação cruzada com 10 subconjuntos. Os modelos atingiram desempenho superior a estudos anteriores em todas as métricas avaliadas.

Os resultados mostram que a proposta é promissora e pode representar mais um passo na construção de sistemas de apoio a rastreamento e detecção de COVID-19, com o Modelo 1 apresentando sensibilidade de 94,74 em imagens RXT sem marcas aparentes de infecção (NORMAL-PCR+). A proposta apresentada é eficaz e exige poucos recursos computacionais, número reduzido de características, permitindo sua adoção em equipamentos portáteis e com configuração limitada.

Como trabalhos futuros estão a melhoria do desempenho do Modelo 2, principalmente no que diz respeito à sensibilidade, a exploração de novas formas de extração de características radiômicas, o teste de novos conjuntos à medida que estejam disponíveis e a extensão de nossa proposta para outras doenças, como pneumonia, tuberculose, câncer, etc.

AGRADECIMENTOS

Agradecemos à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CÁPEES Código Financeiro #001) pelo apoio financeiro.

REFERÊNCIAS

Aljaaf, A.J., Mohsin, T.M., Al-Jumeily, D., and Alloghani, M. (2021). A fusion of data science and feed-forward neural network-based modelling of covid-19 outbreak forecasting in iraq. *Journal of Biomedical Informatics*, 118, 103766.

Almansour, N.A., Syed, H.F., Khayat, N.R., Altheeb, R.K., Juri, R.E., Alhiyafi, J., Alrashed, S., and Olatunji, S.O. (2019). Neural network and support vector machine for the prediction of chronic kidney disease: A comparative study. *Computers in biology and medicine*, 109, 101–111.

Arulmurugan, R. and Anandakumar, H. (2018). Early detection of lung cancer using wavelet feature descriptor and feed forward back propagation neural networks classifier. In *Computational vision and bio inspired computing*, 103–110. Springer.

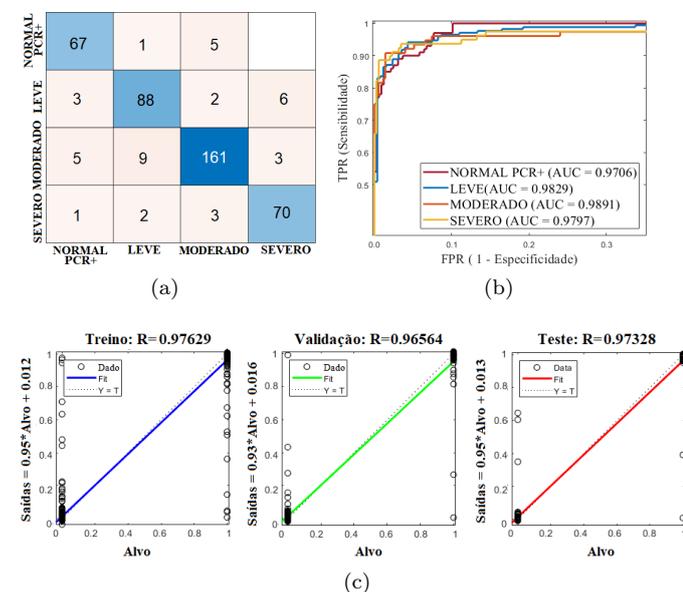


Figura 10. Modelo 2: Classificação Nível de Severidade COVID-19, (a) Matriz de Confusão média (b) Curvas ROC em zoom, e (c) Regressão do modelo.

4. CONCLUSÃO

Neste trabalho, apresentou-se um estudo para detectar COVID-19 em imagens RXT e adicionalmente determinar seu nível de severidade. Um novo conjunto de dados, homogêneo, equilibrado e com os níveis de severidade estratificados segundo *score* RALE foi utilizado. Investigou-se o uso de características extraídas de regiões segmentadas

- Basu, S., Mitra, S., and Saha, N. (2020). Deep learning for screening covid-19 using chest x-ray images. In *2020 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2521–2527. IEEE.
- Candemir, S. and Antani, S. (2019). A review on lung boundary detection in chest x-rays. *International journal of computer assisted radiology and surgery*, 14(4), 563–576.
- Candemir, S., Jaeger, S., Palaniappan, K., Musco, J.P., Singh, R.K., Xue, Z., Karargyris, A., Antani, S., Thoma, G., and McDonald, C.J. (2013). Lung segmentation in chest radiographs using anatomical atlases with nonrigid registration. *IEEE transactions on medical imaging*, 33(2), 577–590.
- Elaziz, M.A., Hosny, K.M., Salah, A., Darwish, M.M., Lu, S., and Sahlol, A.T. (2020). New machine learning method for image-based diagnosis of covid-19. *Plos one*, 15(6), e0235187.
- Fonseca, A.U., Vieira, G.S., and Soares, F. (2021). Screening of viral pneumonia and covid-19 in chest x-ray using classical machine learning. In *2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC)*, 1936–1941. IEEE.
- Gazda, M., Plavka, J., Gazda, J., and Drotar, P. (2021). Self-supervised deep convolutional neural network for chest x-ray classification. *IEEE Access*, 9, 151972–151982.
- Gil, D., Díaz-Chito, K., Sánchez, C., and Hernández-Sabaté, A. (2020). Early screening of sars-cov-2 by intelligent analysis of x-ray images. *arXiv preprint arXiv:2005.13928*. URL <https://arxiv.org/abs/2005.13928>.
- Guyon, I., Weston, J., Barnhill, S., and Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Machine learning*, 46(1), 389–422.
- Hannah Ritchie, Edouard Mathieu, L.R.G. and Roser, M. (2020). Coronavirus pandemic (covid-19). *Our World in Data*. URL <https://ourworldindata.org/coronavirus>. Acessado em 01/04/2022.
- Ismael, A.M. and Şengür, A. (2021). Deep learning approaches for covid-19 detection based on chest x-ray images. *Expert Systems with Applications*, 164, 114054.
- Jain, R., Gupta, M., Taneja, S., and Hemanth, D.J. (2021). Deep learning based detection and analysis of covid-19 on chest x-ray images. *Applied Intelligence*, 51(3), 1690–1700.
- Jha, D., Kim, J.I., and Kwon, G.R. (2017). Diagnosis of alzheimer’s disease using dual-tree complex wavelet transform, pca, and feed-forward neural network. *Journal of healthcare engineering*, 2017.
- Kallianos, K., Mongan, J., Antani, S., Henry, T., Taylor, A., Abuya, J., and Kohli, M. (2019). How far have we come? artificial intelligence for chest radiograph interpretation. *Clinical radiology*, 74(5), 338–345.
- Khan, A.I., Shah, J.L., and Bhat, M.M. (2020). Coronet: A deep neural network for detection and diagnosis of covid-19 from chest x-ray images. *Computer methods and programs in biomedicine*, 196, 105581.
- Kissler, S.M., Tedijanto, C., Goldstein, E., Grad, Y.H., and Lipsitch, M. (2020). Projecting the transmission dynamics of sars-cov-2 through the postpandemic period. *Science*, 368(6493), 860–868.
- Kohavi, R. et al. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai*, volume 14.2, 1137–1145. Montreal, Canada.
- Kovesi, P. et al. (1999). Image features from phase congruency. *Videre: Journal of computer vision research*, 1(3), 1–26.
- Lin, Z., He, Z., Xie, S., Wang, X., Tan, J., Lu, J., and Tan, B. (2021). Aaenet: Adaptive attention network for covid-19 detection from chest x-ray images. *IEEE Transactions on Neural Networks and Learning Systems*, 32(11), 4781–4792. doi:10.1109/TNNLS.2021.3114747.
- Liu, C., Yuen, J., and Torralba, A. (2010). Sift flow: Dense correspondence across scenes and its applications. *IEEE transactions on pattern analysis and machine intelligence*, 33(5), 978–994.
- López-Cabrera, J.D., Orozco-Morales, R., Portal-Díaz, J.A., Lovelle-Enríquez, O., and Pérez-Díaz, M. (2021). Current limitations to identify covid-19 using artificial intelligence with chest x-ray imaging. *Health and Technology*, 11(2), 411–424.
- Mittal, A., Hooda, R., and Sofat, S. (2017). Lung field segmentation in chest radiographs: a historical review, current status, and expectations from deep learning. *IET Image Processing*, 11(11), 937–952.
- Ng, A.Y. (1998). *On feature selection: learning with exponentially many irreverent features as training examples*. Ph.D. thesis, Massachusetts Institute of Technology.
- 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.
- Öksüz, C., Urhan, O., and Güllü, M.K. (2021). Covid-19 detection with severity level analysis using the deep features, and wrapper-based selection of ranked features. *Concurrency and Computation: Practice and Experience*, e6802.
- Panetta, K., Sanghavi, F., Agaian, S., and Madan, N. (2021). Automated detection of covid-19 cases on radiographs using shape-dependent fibonacci-p patterns. *IEEE Journal of Biomedical and Health Informatics*, 25(6), 1852–1863.
- Pietka, E. (1994). Lung segmentation in digital radiographs. *Journal of digital imaging*, 7(2), 79–84.
- Roberts, M., Driggs, D., Thorpe, M., Gilbey, J., Yeung, M., Ursprung, S., Aviles-Rivero, A.I., Etmann, C., McCague, C., Beer, L., et al. (2021). Common pitfalls and recommendations for using machine learning to detect and prognosticate for covid-19 using chest radiographs and ct scans. *Nature Machine Intelligence*, 3(3), 199–217.
- Sapna, S., Tamilarasi, A., Kumar, M.P., et al. (2012). Backpropagation learning algorithm based on levenberg marquardt algorithm. *Comp Sci Inform Technol (CS and IT)*, 2, 393–398.
- Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, 61, 85–117.
- Tabik, S., Gómez-Ríos, A., Martín-Rodríguez, J.L., Sevillano-García, I., Rey-Area, M., Charte, D., Guirado, E., Suárez, J.L., Luengo, J., Valero-González, M., et al. (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.
- Turan, C. and Lam, K.M. (2018). Histogram-based local descriptors for facial expression recognition (fer): A comprehensive study. *Journal of visual communication and image representation*, 55, 331–341.
- Warren, M.A., Zhao, Z., Koyama, T., Bastarache, J.A., Shaver, C.M., Semler, M.W., Rice, T.W., Matthay, M.A., Calfee, C.S., and Ware, L.B. (2018). Severity scoring of lung oedema on the chest radiograph is associated with clinical outcomes in ards. *Thorax*, 73(9), 840–846.
- Wynants, L., Van Calster, B., Collins, G.S., Riley, R.D., Heinze, G., Schuit, E., Bonten, M.M., Dahly, D.L., Damen, J.A., Debray, T.P., et al. (2020). Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. *bmj*, 369.
- Zell, A. (1994). *Simulation neuronaler netze*. 5.3. Addison-Wesley Bonn.