

Análise de técnicas de classificação de imagens de radiografias na detecção de pneumonia e COVID-19

Eurico Delmondes Do Nascimento Junior *
Jeanderson de Sousa Gomes * Lucas Silva Lopes *

* Universidade Federal do Piauí - PI, (e-mail: {euricodelmondes, jeandsgomes10, lucaslopes092020}@ufpi.edu.br).

Abstract: Amid the COVID-19 pandemic, a highly contagious disease that has spread throughout the world, the development of solutions in the fields of medicine, biology, data science, and deep learning to detect and control the disease has become essential. In this context, a study proposes a computational methodology for classifying chest X-ray images into three categories: normal, pneumonia, and COVID-19. This approach can offer an important contribution to the medical field, as it can assist in the early identification of lung pathologies, with reduced costs compared to other diagnostic techniques, such as RT-PCR. The study obtained significant results, with an accuracy of 95.59%, an f1-score of 95.60%, and a kappa of 93.38%, using the combination of the MobileNet image descriptor with the SVM classifier, without the need for image preprocessing.

Resumo: Em meio à pandemia da COVID-19, uma doença altamente contagiosa que se espalhou pelo mundo, tornou-se essencial o desenvolvimento de soluções nas áreas de medicina, biologia, ciência de dados e aprendizado profundo para detectar e controlar a doença. Nesse contexto, um estudo propôs uma metodologia computacional para classificar imagens de raio-x do tórax em três categorias: normal, pneumonia e COVID-19. Essa abordagem pode oferecer uma importante contribuição para a área médica, já que pode auxiliar na identificação precoce de patologias pulmonares, com custos reduzidos em relação a outras técnicas diagnósticas, como o RT-PCR. O estudo obteve resultados significativos, com uma acurácia de 95,59%, f1-score de 95,60% e kappa de 93,38%, utilizando a combinação do descritor de imagem MobileNet com o classificador SVM, sem a necessidade de pré-processamento das imagens.

Keywords: COVID-19; Pneumonia; Deep Learning; Classification; Chest X-ray images.

Palavras-chaves: COVID-19; Pneumonia; Aprendizado profundo; Classificação; Imagens de raio-x do tórax.

1. INTRODUÇÃO

A COVID-19 é uma doença altamente contagiosa causada pelo novo coronavírus (SARS-COV-2), que se espalhou pelo mundo provocando uma pandemia. Até o momento, mais de 202 milhões de pessoas foram infectadas com COVID-19 e mais de 4 milhões de mortes foram registradas em 223 países e regiões ONU (2021). Os sintomas da COVID-19 podem variar em cada paciente, sendo que alguns apresentam quadros clínicos leves e se recuperam rapidamente, enquanto outros podem ter problemas graves de saúde, como pneumonia, insuficiência respiratória, falência de órgãos e outras complicações Kazimierczuk and Jozwik (1990). Alguns dos sintomas menos comuns incluem dores musculares, dor de garganta, dor de cabeça, congestão nasal, conjuntivite, perda do olfato e do paladar e erupções cutâneas Samir et al. (2021). A detecção precoce da doença é essencial para aumentar a taxa de sobrevivência dos pacientes e evitar a contaminação de mais pessoas.

* Reconhecimento do suporte financeiro deve vir nesta nota de rodapé.

Essa doença apresenta uma ampla variedade de sintomas, muitos dos quais são semelhantes aos de outras doenças respiratórias. Isso torna o diagnóstico da COVID-19 difícil, especialmente em áreas onde a incidência de outras doenças respiratórias é alta. Além disso, a COVID-19 pode se manifestar de formas muito diferentes em cada paciente, o que aumenta a dificuldade no diagnóstico precoce Samir et al. (2021). Por exemplo, alguns pacientes podem apresentar apenas sintomas leves, como febre e tosse, enquanto outros desenvolvem complicações graves, como pneumonia e insuficiência respiratória, que exigem hospitalização imediata ?. Com isso, é fundamental que os profissionais de saúde estejam bem informados sobre os diferentes sintomas que a COVID-19 pode apresentar, e que utilizem ferramentas de diagnóstico precisas para identificar a doença o mais cedo possível.

Nesse contexto, a pesquisa e o desenvolvimento de soluções nas áreas de medicina, biologia, ciência de dados e aprendizado profundo têm sido essenciais para prevenir e controlar a pandemia Romanov (2020). Uma ferramenta tecnológica que permita um diagnóstico rápido e preciso

da infecção por COVID-19 pode ser extremamente útil para os profissionais de saúde. A ferramenta clínica mais comum para diagnóstico da COVID-19 é a reação da transcriptase reversa seguida pela reação em cadeia da polimerase (RT-PCR), que é cara, menos sensível e requer médicos especializados. Por outro lado, a utilização de inteligência artificial (IA) em ferramentas de classificação de imagens de radiografias pode ser uma ótima alternativa para diagnóstico da COVID-19, sendo uma opção mais barata e precisa Chowdhury et al. (2020).

2. TRABALHOS RELACIONADOS

Nesta seção serão apresentados os trabalhos relacionados ao trabalho apresentado.

O artigo Carvalho et al. (2021) apresenta uma metodologia para detecção de COVID-19 a partir de imagens de tomografia computadorizada (CT), utilizando uma arquitetura de rede neural convolucional para extrair recursos das imagens, otimizando seus hiperparâmetros através de um estimador de árvore de Parzen, aplicando um algoritmo genético para seleção de recursos e quatro classificadores diferentes para a classificação. Os resultados obtidos em dois conjuntos de dados foram satisfatórios, com índice kappa de 0,995 e accuracy e F-score de 0,997 no conjunto de dados SARS-CoV-2 CT-Scan, e índice kappa de 0,975 e accuracy e F-score de 0,987 e 0,986, respectivamente, no conjunto de dados COVID-CT. A metodologia proposta é promissora para a detecção de COVID-19 a partir de imagens de CT.

Marques et al. (2022) propuseram a utilização de uma CNN para realizar descrição das imagens e a aplicação dos classificadores LeNet-5 e VGG-19. A base de dados utilizada para teste foi a intitulada COVIDxCT. Esse método podendo chegar a 0,99 acurácia para classificação.

O artigo Chowdhury et al. (2020) propõe a utilização de aprendizado por transferência em conjunto com oito classificadores diferentes e aumento de imagem para treinar e validar diversas redes neurais convolucionais (CNNs) profundas pré-treinadas. Os autores criaram um banco de dados público com uma mistura de 423 radiografias de COVID-19, 1.485 de pneumonia viral e 1.579 imagens normais de radiografia de tórax, combinando vários bancos de dados e também coletando imagens de artigos recentes. As redes foram treinadas para classificar dois esquemas diferentes: pneumonia normal e COVID-19, e pneumonia normal, viral e COVID-19 com e sem aumento de imagem, obtendo uma accuracy e F-score de 0,994 e 0,995, respectivamente, sem aumento de imagem e com aumento de imagem, uma accuracy e F-score de 0,997 e 0,995, respectivamente. Os resultados mostram grande potencial para a detecção de COVID-19 a partir de radiografias de tórax.

O modelo ULNet proposto por Wu et al. (2021) utiliza uma rede neural convolucional (CNN) para detectar COVID-19 em imagens de raios X de tórax, treinando e testando em um conjunto de dados da COVID-19 Radiography Database com classificação binária (COVID-19 vs Normal) e multiclasse (COVID-19 vs Normal vs Pneumonia Viral). O modelo alcança uma precisão de detecção de COVID-19 de 0,994 de acurácia e 0,988 de F1 na tarefa de classificação

binária, e 0,952 de acurácia e 0,94 de F1 na tarefa de classificação multiclasse.

3. MÉTODO PROPOSTO

Nesta seção, serão descritas as etapas da metodologia utilizada para obter os resultados apresentados nas próximas seções. Um resumo visual da metodologia adotada pode ser encontrado na Figura 1.

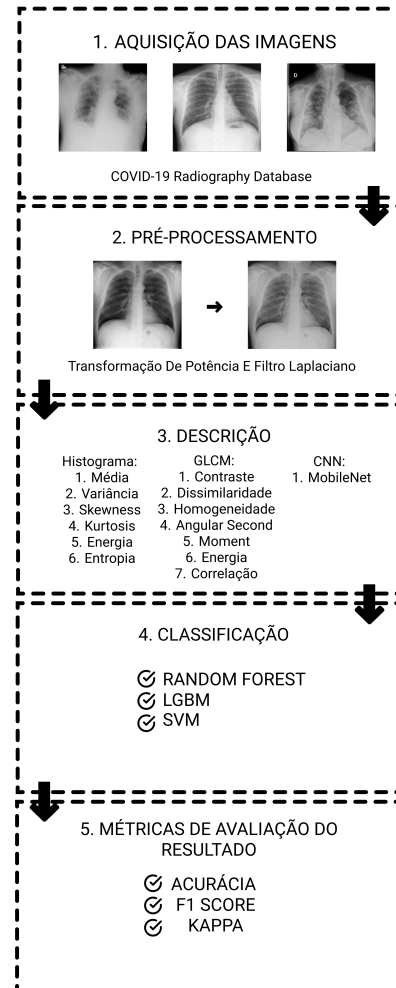


Figura 1. Metodologia proposta.

Para os testes realizados por esse trabalho foi utilizada uma amostragem de 3900 imagens da base de dados intitulada COVID-19 Radiography Database, composta por 21.165 imagens. As Figuras 2,3,4 apresentam exemplos de imagens presentes na base. As 3900 imagens da amostragem são compostas por 1300 imagens de cada classe, sendo elas COVID-19, pneumonia e normal. A amostragem foi realizada por dois motivos principais: reduzir o custo computacional dos testes e balancear as classes do conjunto de imagens utilizadas nos testes.

Após a obtenção das imagens, foi aplicada inicialmente as técnicas de pré-processamento de transformação (com gama igual a 0.5) de potência e realce de bordas usando o filtro Laplaciano. A transformação de potência é uma transformação de intensidade que aplica uma potência aos níveis de cinza da imagem Gonzalez et al. (2010). O filtro

Tabela 1. Resumo dos trabalhos relacionados

Trabalho	Dataset	Descritor	Classificador	Métricas
Carvalho et al. (2021)	SARS-COV-2 CT-Scan Dataset	CNN	MLP	accuracy (A) = 0.997 F-score (F) = 0.997 kappa (K) = 0.995
	COVID-CT Dataset	CNN	MLP	accuracy (A) = 0.987 F-score (F) = 0.986 kappa (K) = 0.975
Marques et al. (2022)	2*COVIDxCT	2*CNN	LeNet-5	accuracy (A) = 0.99 F-score (F) = 0.99 kappa (K) = 0.99
			VGG-19	accuracy (A) = 1 F-score (F) = 1 kappa (K) = 1
Chowdhury et al. (2020)	Dataset criado sem augmentation	CNN	ResNet18 ou ChexNet	accuracy (A) = 0,996 F-score (F) = 0,923
	Dataset criado com augmentation	CNN	DenseNet201	accuracy (A) = 0,997 F-score (F) = 0,995
Wu et al. (2021)	COVID-19 Radiography Database COVID-19 VS Normal	CNN	ULNet	accuracy (A) = 0,994 F-score (F) = 0,988
	COVID-19 Radiography Database COVID-19 VS Normal VS Viral Pneumonia	CNN	ULNet	accuracy (A) = 0,952 F-score (F) = 0,94

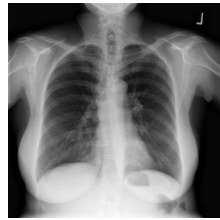


Figura 2. COVID-19. Figura 3. Normal.

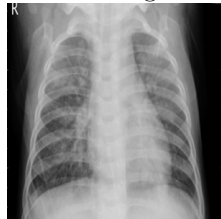


Figura 4. Pneumonia.

Laplaciano enfatiza as regiões de descontinuidade e suaviza as regiões de variação lenta de níveis de intensidade. Após a aplicação do filtro, foi criado um fundo preservando as descontinuidades, somando a imagem Laplaciana à imagem original. O uso do filtro Laplaciano permitiu um realce das bordas da imagem.

Posteriormente, foi realizada a etapa de descrição da imagem, para a qual foram utilizados três métodos distintos. O primeiro método é composto por descritores de 1° ordem, que incluem as medidas de média, variância, skewness, kurtosis, energia e entropia. O segundo método é composto por descritores de 2° ordem, baseados na matriz de co-ocorrência de níveis de cinza (GLCM). E, por fim, o terceiro método é o descritor MobileNet. A seguir, serão apresentadas breves descrições de cada um desses descritores.

Média: A média de um histograma, é a frequência média dos píxeis na imagem i . É calculado pela somatória das probabilidades da frequência de intensidades dividido pelo número total de frequências, n ?.

Variância: A variância de um histograma é uma medida da dispersão de uma distribuição. A variância é a média dos desvios ao quadrado da média ?.

Skewness: O skewness, ou valor de assimetria do histograma, é utilizada como uma medida de simetria, onde o valor 0, indicando simetria em relação à média, positivo, indicando uma inclinação para a esquerda ou negativo, indicando uma inclinação para a direita[3].

Kurtosis: A kurtosis é uma medida de "achatamento" do histograma em relação à média. A kurtosis é útil para entender como os níveis de cinza na imagem se concentram em relação à média[3].

Energia: A medida de uma medida da uniformidade do histograma. Ou seja, essa medida pode ser usada como uma medida da "complexidade" da imagem. Por exemplo, uma imagem com muitos detalhes ou textura pode ter uma energia mais alta do que uma imagem mais lisa ou homogênea Gonzalez et al. (2010).

Entropia: A entropia é uma medida da incerteza ou "aleatoriedade" dos níveis de cinza na imagem. Ela pode ser usada para comparar a aleatoriedade dos níveis de cinza entre duas imagens ou para detectar anomalias em uma imagem Gonzalez et al. (2010).

Contraste: O contraste é uma medida da diferença entre os valores dos pixels adjacentes na imagem. Quanto maior o valor do contraste, mais distintas são as diferenças entre os pixels, o que significa que a imagem tem uma maior variedade de tons e é mais nítida.

Dissimilaridade: A dissimilaridade é uma medida da soma das diferenças absolutas dos valores de cinza dos pixels vizinhos. Quanto maior o valor da dissimilaridade, mais distintas são as diferenças entre os pixels, o que significa que a imagem tem áreas com grandes diferenças de brilho.

Homogeneidade: A homogeneidade mede a proximidade dos valores dos pixels vizinhos. Quanto maior o valor da homogeneidade, mais próximos os valores dos pixels estão uns dos outros, o que significa que a imagem tem uma aparência mais uniforme.

Angular Second Moment (ASM): O ASM é uma medida da uniformidade da distribuição de cinza dos pixels da imagem. Quanto mais uniforme a distribuição, maior o valor do ASM, o que significa que a imagem é mais uniforme em termos de seus tons de cinza.

Energia: A energia é a raiz quadrada do ASM e mede a quantidade de energia em cada direção da imagem. Quanto maior o valor da energia, mais concentrada é a energia em uma determinada direção, o que pode indicar a presença de bordas ou outros padrões na imagem.

Correlação: A correlação mede a relação entre os valores dos pixels vizinhos em relação à sua posição. Uma alta correlação indica que os valores dos pixels vizinhos têm uma forte relação em termos de sua posição, o que significa que a imagem tem áreas que são altamente correlacionadas e outras que são menos correlacionadas.

MobileNet: Consiste em uma rede neural convolucional (CNN) projetada para processar imagens em dispositivos móveis, com alta eficiência computacional e baixo consumo de energia. Ele utiliza filtros de convolução de tamanho reduzido (3x3) em vez dos convencionais (5x5 ou 7x7), além de uma técnica chamada de "depthwise separable convolution", que divide a convolução em duas etapas separadas para reduzir o número de parâmetros e operações. Essas otimizações permitem que o modelo MobileNet atinja uma precisão comparável a outras CNNs mais complexas, enquanto consome menos recursos computacionais.

Durante a etapa de classificação, utilizamos dois classificadores: o Random Forest, LightGBM (LGBM) e Support Vector Machine (SVM). O Random Forest é um algoritmo de aprendizado de máquina baseado em árvores de decisão, que usa a combinação de múltiplas árvores para obter um resultado final mais preciso e robusto.

O LightGBM (LGBM) é uma implementação de algoritmo de boosting baseado em árvores de decisão, que usa uma técnica chamada de "gradient-based one-side sampling" para selecionar as melhores características e otimizar o processo de treinamento. Isso torna o LGBM mais rápido e preciso em comparação com outros algoritmos de boosting.

O Support Vector Machine (SVM) é um algoritmo de classificação que usa uma técnica de separação de margem máxima para encontrar um hiperplano que separe os dados em duas classes distintas. O SVM pode ser usado tanto para problemas de classificação linear quanto não linear e é conhecido por sua eficácia em conjuntos de dados com alta dimensionalidade.

Para avaliar a performance dos modelos treinados, selecionamos três métricas: acurácia, f1 score e coeficiente de kappa. A acurácia é calculada como a proporção de acertos do modelo em relação ao total de exemplos avaliados. Já o f1 score é uma métrica que combina a precisão e o recall do modelo, fornecendo uma medida da eficácia geral do modelo na classificação das amostras. O kappa de Cohen é uma ferramenta estatística que possibilita avaliar a concordância entre duas classificações, sejam elas realizadas por diferentes avaliadores ou pelo mesmo avaliador em momentos distintos.

4. RESULTADOS E DISCUSSÃO

As Tabelas 2 e 3 contêm os resultados dos classificadores Random Forest, LightGBM e SVM utilizando os descritores de primeira e segunda ordem, bem como a MobileNet. A Tabela 2 apresenta os resultados sem a aplicação de pré-processamento na base de dados, enquanto a Tabela 3 exibe os resultados após a aplicação do pré-processamento descrito na seção de metodologia.

Tabela 2. Resultados sem pré-processamento

descritor	classificador	Accuracy	f1-score	kappa
Primeira Ordem	Random Forest	63,74 ± 4,54	63,69 ± 4,59	45,62 ± 6,80
	SVM	44,92 ± 3,49	40,08 ± 4,74	17,38 ± 5,24
	LightGBM	64,92 ± 4,77	64,85 ± 4,80	47,38 ± 7,16
Segunda Ordem	Random Forest	69,85 ± 4,24	69,78 ± 4,15	54,77 ± 6,37
	SVM	50,59 ± 2,66	44,46 ± 1,19	25,88 ± 3,99
	LightGBM	69,51 ± 3,78	69,49 ± 3,70	54,27 ± 5,67
MobileNet	Random Forest	92,51 ± 0,19	91,50 ± 0,19	87,27 ± 0,28
	SVM	95,59 ± 0,73	95,60 ± 0,73	93,38 ± 1,10
	LightGBM	94,35 ± 0,90	94,36 ± 1,01	91,53 ± 1,08

Tabela 3. Resultados com pré-processamento Laplace e Transformação Potencial

descritor	classificador	Accuracy	f1-score	kappa
Primeira Ordem	Random Forest	65,90 ± 2,96	65,86 ± 2,88	48,85 ± 4,44
	SVM	42,18 ± 3,55	35,22 ± 3,79	13,27 ± 5,32
	LightGBM	65,97 ± 3,20	65,87 ± 3,13	48,96 ± 4,80
Segunda Ordem	Random Forest	74,90 ± 2,79	74,88 ± 2,75	62,35 ± 4,19
	SVM	64,59 ± 4,16	63,11 ± 6,17	46,88 ± 6,24
	LightGBM	73,97 ± 2,92	73,93 ± 2,87	60,96 ± 4,39
MobileNet	Random Forest	87,26 ± 1,74	87,33 ± 1,72	80,88 ± 2,61
	SVM	92,10 ± 1,08	92,12 ± 1,11	88,15 ± 1,62
	LightGBM	91,23 ± 0,94	91,25 ± 0,96	86,84 ± 1,41

O melhor resultado encontrado neste estudo foi obtido ao utilizar a MobileNet como descritor e SVM como

classificador, sem a aplicação de pré-processamento nas imagens. O modelo alcançou uma acurácia de 95,59%, f1-score de 95,60% e kappa de 93,38%, conforme apresentado na Tabela 2. Vale ressaltar que o melhor resultado obtido neste trabalho foi superior ao alcançado por Wu et al. (2021), com acurácia de 95,20% e f1-score de 94%.

5. CONCLUSÃO

O trabalho em questão apresentou uma metodologia computacional para a classificação de imagens de raio-x do tórax em três categorias: normal, pneumonia e covid-19. Essa abordagem pode representar uma importante contribuição para a área médica, uma vez que irá auxiliar na identificação precoce de patologias pulmonares, com custos reduzidos em comparação com outras técnicas diagnósticas, como o RT-PCR.

O melhor resultado atingiu uma acurácia de 95,59%, f1-score de 95,60% e kappa de 93,38%, sem o uso de pré-processamento das imagens e a combinação do descritor de imagem MobileNet com o classificador SVM.

AGRADECIMENTOS

Coloque aqui seus agradecimentos.

REFERÊNCIAS

- Carvalho, E.D., Silva, R.R., Araújo, F.H., de AL Rabelo, R., and de Carvalho Filho, A.O. (2021). An approach to the classification of covid-19 based on ct scans using convolutional features and genetic algorithms. *Computers in biology and medicine*, 136, 104744.
- Chowdhury, M.E., Rahman, T., Khandakar, A., Mazhar, R., Kadir, M.A., Mahbub, Z.B., Islam, K.R., Khan, M.S., Iqbal, A., Al Emadi, N., et al. (2020). Can ai help in screening viral and covid-19 pneumonia? *IEEE Access*, 8, 132665–132676.
- Gonzalez, R.C., Woods, R.E., and Eddins, S.L. (2010). Morphological reconstruction. *Digital image processing using MATLAB, MathWorks*.
- Kazimierczuk, M.K. and Jozwik, J. (1990). Analysis and design of class e zero-current-switching rectifier. *IEEE transactions on circuits and systems*, 37(8), 1000–1009.
- Marques, J.V., Ferreira, J.F., Veras, R.M., and Silva, R.R. (2022). Classificação e segmentação de covid-19 em imagens de tomografia computadorizada usando aprendizado profundo. *Sociedade Brasileira de Computação*.
- ONU (2021). Coronavirus disease (covid-19) pandemic. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>. Acesso em: 31/03/2023.
- Romanov, K. (2020). Coronavirus disease covid-2019. *Safety and Risk of Pharmacotherapy (E-Journal)*, 03–Aug.
- Samir, A., Naguib, N.N., Elnekeidy, A., Baess, A.I., and Shawky, A. (2021). Covid-19 versus h1n1: challenges in radiological diagnosis—comparative study on 130 patients using chest hrct. *Egyptian Journal of Radiology and Nuclear Medicine*, 52, 1–9.
- Wu, T., Tang, C., Xu, M., Hong, N., and Lei, Z. (2021). Ulnet for the detection of coronavirus (covid-19) from chest x-ray images. *Computers in Biology and Medicine*, 137, 104834.

Apêndice A. SUMÁRIO DA GRAMÁTICA SÂNSCRITA

Apêndice B. ALGUM VOCABULÁRIO MAIA