

Classifying COVID-19 Positive X-Ray using Deep Learning Models

I. Rodrigues, G. L. Santos, D. F. H. Sadok, P. T. Endo

Abstract—COVID-19 is a pandemic characterized by uncertainty not only in transmission and pathogenicity, but also in disease-specific control options. Despite many governmental measures, the disease is spreading and in many countries, the public health system is close to be collapsed. Alternative techniques should be taken in order to minimize the COVID-19 negative impacts on the society. This work presents preliminary results of deep learning models to classify COVID-19 positive based on X-ray images. We provide binary classification (COVID-19 vs healthy, and COVID-19 vs pneumonia) and also multiclass (COVID-19 vs pneumonia vs healthy) regarding five metrics: accuracy, precision, sensibility, specificity and F1-score. Results show that VGG models present the best results, achieving 98.81% of precision in binary classification, and 91.68% in multiclass classification.

Index Terms—COVID-19, X-ray, pneumonia, deep learning, binary classification, multiclass classification.

I. INTRODUÇÃO

De acordo com a *World Health Organization* (WHO), os coronavírus são uma grande família de vírus que podem causar doenças em animais ou humanos [1]. Em humanos, sabe-se que vários coronavírus causam infecções respiratórias, variando do resfriado comum a doenças mais graves, como a *Middle East Respiratory Syndrome* (MERS) e a *Severe Acute Respiratory Syndrome* (SARS) [2], [3].

O coronavírus descoberto recentemente causa a doença denominada COVID-19 [4], [5]. Esse novo vírus, que foi nomeado SARS-CoV-2, era desconhecido até antes do início do surto em Wuhan, na China, em dezembro de 2019; e atualmente é uma pandemia que afeta muitos países do mundo. Até o dia 13 de maio de 2020, o mundo contabilizava 4.170.424 casos confirmados de COVID-19 e 287.399 óbitos pela doença.

De acordo com Grasselli [6], a maioria das pessoas infectadas com o COVID-19 experimenta doença respiratória leve a moderada, enquanto alguns desenvolvem pneumonia severa que ocasiona dificuldade respiratória, que demandam utilização das unidades de terapia intensiva (UTI). Neste cenário de rápida disseminação do vírus e grande número de mortos devido à complicações por COVID-19, duas situações são alarmantes: (a) a grande probabilidade de colapso dos sistemas de saúde e (b) a escassez de kits de teste disponíveis.

Para minimizar os impactos nos sistemas de saúde, os agentes responsáveis por formulação de políticas públicas de saúde, nas esferas federais, estaduais e locais, precisam medir, monitorar e responder a fatores que afetam a vida social e econômica de suas comunidades. Porém, devido a novidade da doença (como a incerteza na forma de transmissão, patogenicidade e falta de medidas de controle específico), a tomada de decisão é complicada e a eficácia das intervenções é incerta em diversos níveis. Além disso, as normas e os comportamentos comunitários não são homogêneos, nem as intervenções públicas.

Com relação a possível indisponibilidade de testes de COVID-19, outras opções rápidas de diagnóstico alternativo devem ser pensadas e implementadas para evitar o mascaramento do número de pessoas infectadas pela doença. Modelos de *deep learning* têm sido aplicados com sucesso nos mais diversos exames de imagem, como detecção de Alzheimer através de exames de ressonância magnética (MRI) [7], predição de câncer de mama maligno em exames de MRI [8], classificação de exames de raio-x do tórax para detecção de tuberculose [9], dentre outras aplicações.

Baseado neste contexto, o presente trabalho apresenta análises preliminares de modelos de *deep learning* e *transfer learning* para classificação e diagnóstico automático de COVID-19 baseado em raio-x do tórax, como uma alternativa rápida e de baixo custo. Uma vez que a COVID-19 pode ser confundida com outros tipos de pneumonia viral, foi proposta uma classificação que não somente diferencia pacientes com COVID-19 de pacientes saudáveis, mas também leva-se em consideração pacientes com pneumonia viral. Os modelos de *deep learning* e *transfer learning* propostos neste trabalho têm uso prático na rotina dos profissionais de saúde para auxílio no processo de diagnóstico, especialmente no atendimento a pacientes durante a atual pandemia da COVID-19. Os mesmos podem ser aplicados para auxiliar a triagem dos pacientes, tomar decisões sobre internação, acompanhar pacientes durante a internação, minimizar a necessidade de realização de tomografia computadorizada, e também complementar a análise de outros exames de imagens.

O restante deste trabalho está organizado da seguinte forma: A Seção II apresenta uma fundamentação teórica, além de apresentar os trabalhos relacionados para fins de comparação. A Seção III apresenta os métodos, base de dados e configurações experimentais dos modelos propostos. A Seção IV apresenta os resultados e discussão a respeito dos experimentos realizados. E por fim, a Seção V apresenta as principais contribuições, desafios e trabalhos futuros para este trabalho.

I. Rodrigues, Universidade Federal de Pernambuco (UFPE), Pernambuco, Brasil, irrs@cin.ufpe.br

G. L. Santos, Universidade Federal de Pernambuco (UFPE), Pernambuco, Brasil, gls4@cin.ufpe.br

D F. H. Sadok, Universidade Federal de Pernambuco (UFPE), Pernambuco, Brasil, jamel@cin.ufpe.br

P. T. Endo, Universidade de Pernambuco (UPE), Pernambuco, Brasil, patricia.endo@upe.br

II. CONCEITOS BÁSICOS

A. Deep Learning

Há um crescimento significativo de estudos na literatura usando aprendizado de máquina em vários contextos, entre eles análise de mineração de dados [10], previsão do mercado de ações [11]; produtos de consumo [12] e veículos autônomos [13]. Diversas técnicas podem ser usadas no desenvolvimento dessas soluções, incluindo *support vector machine* (SVM), *Random Forest*, Naive-Bayes, entre outros [11]. Cada técnica tem suas particularidades e pode ser usada para resolver diferentes categorias de problemas.

O processamento de dados brutos de séries temporais (por exemplo, texto, *streaming* de vídeo, dados do sensor etc.) é um desafio para as técnicas convencionais de aprendizado de máquina, pois pressupõe que todas as entradas são independentes umas das outras [14]. Como tal, uma representação interna e abstrata, geralmente chamada de *feature*, precisa ser extraída manualmente com base no conhecimento do domínio [15]. Algumas técnicas de aprendizado de máquina surgiram para resolver essa necessidade de extração manual a partir de dados brutos, como as redes neurais.

As redes neurais são compostas por módulos simples, mas não lineares, conectados, chamados neurônios. Em geral, esses neurônios são organizados em camadas sequenciais e a saída de uma camada é a entrada da próxima camada (denominado de *feedforward*), conforme ilustrado na Figura 1.

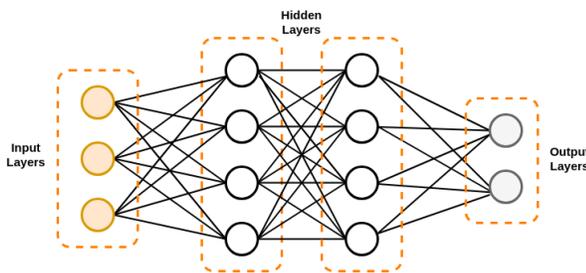


Fig. 1. Exemplo de uma rede neural genérica [15]

A primeira camada (camada de entrada) recebe os dados de entrada brutos, enquanto a última camada (camada de saída) é responsável pela classificação dos dados. As camadas intermediárias, também chamadas de camadas ocultas, são compostas por neurônios, capazes de transformar os dados de entrada em uma representação mais abstrata [12]. Cada camada armazena uma matriz de pesos, que são atualizados de forma iterativa durante o processo de aprendizado. Após o treinamento, as matrizes dos pesos representam as *features* extraídas dos dados, ou seja, as informações relevantes extraídas para fazer novas inferências com base em novos dados.

Deep learning é uma subárea das redes neurais que se caracteriza principalmente por conter várias camadas ocultas [12]. Sendo assim, esse tipo de rede consegue extrair representações mais abstratas a partir dos dados brutos, além de fazer transformações necessárias para se processar diversos tipos de dados tais como imagens, áudio, texto, vídeo, etc

[16]. Entre os diversos tipos de técnicas, as redes neurais convolucionais têm recebido atenção na última década devido ao sucesso obtido em campos com processamento de imagem e visão computacional.

B. Redes Neurais Convolucionais

As redes neurais convolucionais, do inglês *convolutional neural networks* (CNN) são projetadas para processar dados armazenados como matrizes de múltiplas dimensões: 1D para dados sequenciais (vetores), como dados de sensores, sinais e textos; matrizes 2D para imagens ou áudio; e 3D para imagens volumétricas ou vídeos. O objetivo por trás da CNN é o aproveitamento das propriedades do formato dos dados brutos [12].

A Figura 2 apresenta a arquitetura genérica de uma CNN. As duas principais camadas da CNN são: camadas convolucionais e de *pool*. As camadas convolucionais geram um conjunto de mapas de recursos, dependendo do número de filtros pré-definidos. Cada filtro representa um conjunto de pesos que define a operação de convolução discreta que vai ser aplicada nos dados de entrada [17]. Depois que os mapas de recursos são calculados, uma função de ativação não linear é aplicada nos dados proveniente da convolução.

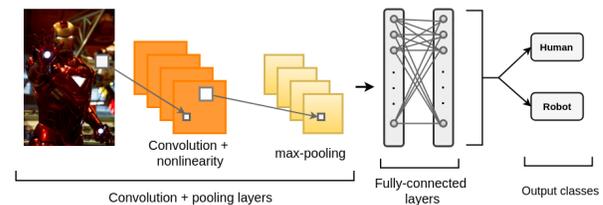


Fig. 2. Exemplo de uma CNN para classificação de imagens (Adaptado de [18])

A saída da camada convolucional geralmente é a entrada para uma camada de *pool*. O papel dessa camada é agregar recursos semanticamente semelhantes (da camada anterior) em uma representação única [12]. A técnica de *pooling* mais usada comumente é o valor máximo, onde o valor máximo de ativação na região de subamostragem é selecionado como saída para a próxima camada. No entanto, outras técnicas podem ser aplicadas, como a média da região de subamostragem. Essa técnica de agrupamento reduz a dimensão dos dados e cria uma invariância a pequenas mudanças e distorções presentes nos dados de entrada [12], [17].

Após sequências de camadas de convolucionais e de *pool*, geralmente camadas totalmente conectadas são adicionadas para classificar ou fazer previsões [19], [20].

C. Transfer Learning

Para treinar modelos *deep learning*, é imprescindível que se tenha um grande quantidade de dados e máquinas com grande capacidade computacional. Considerando CNNs, se um modelo for treinado para classificar imagens de forma binária, uma grande quantidade de imagens de cada categoria deve estar disponível para o treinamento. Além disso, CNNs podem

ter grande quantidade de parâmetros, que são atualizados com base em operações matriciais durante o treinamento [12]. Assim, máquinas com hardware dedicado (tais como GPU) podem ser utilizadas para o treinamento de tais modelos.

Contudo, tais recursos podem não estar disponíveis para a criação dos modelos. Se a quantidade de dados for pequena, os modelos podem não aprender a generalizar para novos exemplos; por outro lado, o treinamento pode demorar muito tempo caso o poder computacional disponível não for suficiente, tornando a aprendizagem inviável. Assim, o conceito de *transfer learning* foi proposto para lidar com esses problemas.

Em *transfer learning*, o conhecimento aprendido para executar uma tarefa pode ser utilizado para executar outra tarefa [21]. Por exemplo, se um modelo CNN foi treinado para classificar cachorros e gatos a partir de imagens, o conhecimento desse modelo pode ser utilizado para classificar, por exemplo, cavalos e tigres.

Existem várias maneiras de se fazer *transfer learning*, mas a mais popular para lidar com imagens foi proposta por Yosinski [22]. Nesse processo, as *features* que foram aprendidas por um modelo para desempenhar uma certa tarefa podem ser transferidas para um outro modelo, para desempenhar outra tarefa diferente. Um pequeno ajuste (denominado *fine tuning*) precisa ser feito nas últimas camadas do modelo (normalmente as camadas totalmente conectadas) [23]. Isso se deve ao fato que essas camadas são responsáveis pela classificação, assim precisam ser atualizadas para lidar com os dados referentes à nova tarefa. Uma vez que os parâmetros de poucas camadas do modelo vão ser ajustadas, o treinamento também é menos custoso que treinar o modelo inteiro novamente. Portanto, o processo de *fine tuning* permite que pesquisadores reutilizem seus modelos (e modelos desenvolvidos por outras pessoas) para novas tarefas [21].

D. Trabalhos Relacionados

Devido a novidade do tema, poucos trabalhos foram publicados em periódicos e/ou conferências. De fato, a grande maioria dos trabalhos estão disponíveis em formato de *preprint* em plataformas como o arXiv e o medRxiv, onde os mesmos ainda não foram revisados e certificados por pares (como [24]–[26]).

O único trabalho relacionado encontrado foi publicado em [27]. Esse trabalho realizou uma comparação entre várias arquiteturas de modelos *deep learning* para classificar imagens de raio-x com o intuito de realizar diagnóstico automático de COVID-19. Os autores analisaram cinco modelos diferentes de CNNs, e o treinamento foi realizado utilizando a técnica *transfer learning*, para classificação binária e multi-classes. Na classificação binária, as classes consideradas foram (a) pessoas saudáveis e (b) pessoas com COVID-19 ou pneumonia; já na classificação multi-classes, foram consideradas as seguintes classes: (a) pessoas saudáveis, (b) pessoas com pneumonia e (c) pessoas com COVID-19.

Os resultados mostraram que a arquitetura VGG-19 [28] obteve os melhores resultados de acurácia, tanto para classificação binária (98.75%) e multi-classes (93.48%). A

arquitetura MobileNet v2 [29] obteve o melhor resultado de sensibilidade (99.10%); e a arquitetura Xception [30] obteve o melhor resultado de especificidade (99.99%).

Embora bastante próximos, este trabalho possui diferenças significativas ao apresentado pelos autores em [27]. A primeira diferença é com relação a base de dados utilizada. Neste trabalho, a base de dados apresenta um desbalanceamento maior, o que torna o processo de treinamento dos modelos CNNs mais difícil. A segunda diferença é com relação aos modelos avaliados. Este trabalho avaliou arquiteturas que não foram contempladas em [27] na classificação automática da COVID-19. A última diferença é com relação a forma de classificação. Enquanto os autores em [27] consideraram classificação binária como pessoas saudáveis e doentes (independente do tipo de pneumonia), este trabalho realiza classificação binária de pessoas com COVID-19 vs saudáveis; pessoas com COVID-19 vs pneumonia; e classificação multiclasse de pessoas com COVID-19 vs saudáveis vs pneumonia.

III. MATERIAIS E MÉTODOS

A. Base de Dados e Classificação

Este trabalho utiliza uma base de dados disponibilizada por uma equipe de pesquisadores da Universidade do Qatar, Doha, Qatar e da Universidade de Dhaka, Bangladesh, juntamente com colaboradores do Paquistão e da Malásia em colaboração com médicos, sendo disponibilizada para fins de pesquisa na plataforma Kaggle [31]. Esta base de dados é composta por 219 imagens de raio-x de pacientes diagnosticados com COVID-19, 1341 imagens de pacientes saudáveis e 1345 imagens de outros tipos de pneumonia viral. A Figura 3 apresenta exemplos de imagens referentes às classes citadas que podem ser encontrados na base de dados.

Para equilíbrio da base de dados em termos de quantidade, aplicou-se o processo de *data augmentation*. Este processo cria imagens artificiais da classe com menor número de instâncias e as mesmas são utilizadas apenas no conjunto de treinamento. Os parâmetros levados em consideração para criação de imagens artificiais são *flips* horizontais de imagens e variação de *zoom*. O processo de balanceamento de dados pode garantir um melhor aprendizado das classes com menores quantidades de dados [32].

Neste trabalho, analisamos três diferentes classificações utilizando a base de dados apresentada: (a) pacientes com COVID-19 vs pacientes saudáveis, (b) pacientes com COVID-19 vs pacientes com diversos tipos de pneumonia viral, e (c) classificação para diferenciação de todas as três classes (COVID-19, saudável e pneumonia).

B. Modelos de Deep Learning para Classificação

Foram utilizados modelos de *deep learning*, mais especificamente CNN para classificação COVID-19 baseado em imagens de raio-x. A escolha destes modelos se dá pelo seu sucesso na classificação de imagens, ou mais especificamente imagens de raio-x [33] [34]. Com o objetivo de garantir que esta proposta seja efetivamente transferida para toda

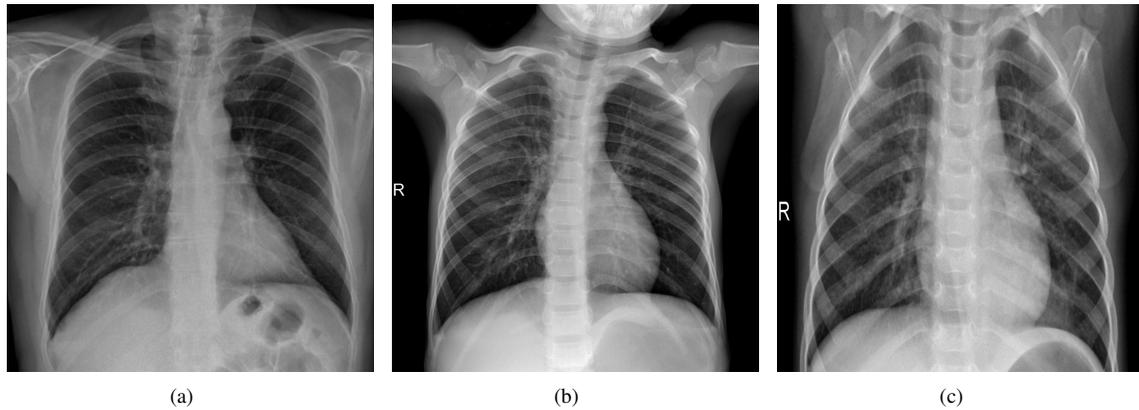


Fig. 3. Exemplos de imagens de raio-x de pacientes pertencentes as classes (a) COVID-19, (b) outros tipos de pneumonia viral e (c) saudável.

comunidade acadêmica, os modelos estão disponíveis e podem ser publicamente acessados através do link [35].

Neste trabalho utilizamos arquiteturas de *transfer learning* devido a quantidade de dados existentes na base disponível. Com isso, os modelos de *deep learning* utilizados neste trabalho constituem-se em dois sub-modelos: (a) modelo de CNN do estado da arte com pesos pré-treinados em outra grande base de dados [36] e (b) camadas densas, onde é realizado o *fine tuning* dos modelos, ou novo aprendizado na base de dados de raio-x. Os modelos são treinados separadamente em cada um dos conjuntos de dados definidos pela etapa de classificação (Seção III-A). A Figura 4 apresenta como consistem os modelos de *deep learning* propostos neste trabalho.

Os modelos de *deep learning* utilizados para o sub-modelo (i) foram: DenseNet [37], InceptionV3 [38], MobileNet [29], ResNet50 [39] e VGG (arquiteturas VGG-16 e VGG-19 separadamente) [28]. Cada um destes modelos possuem diferentes quantidades de camadas convolucionais, *poolings*, dentre outras configurações. A escolha destes modelos se dá pela possibilidade de alta capacidade de generalização nos mais diversos problemas de classificação de imagens, como por exemplo o desafio da base de dados Imagenet [36].

É aplicado o *pooling* de média global nas *features* extraídas por cada um modelo de *deep learning* do sub-modelo (i). Para o sub-modelo (ii), a definição da quantidade de neurônios, bem como a quantidade de camadas densas foi baseada em testes empíricos. Foi-se definido a quantidade de duas camadas densas, cada uma contendo 256 neurônios. Com isso, é feito neste sub-modelo o *fine tuning* com os dados extraídos por cada uma sub-arquitetura da etapa (i). Isso consiste na classificação dos dados para cada uma das classes definidas neste trabalho.

C. Configuração dos Modelos Propostos

A Tabela I apresenta as configurações dos modelos, com seus atributos e seus respectivos valores para a realização dos testes. Os testes foram realizados em uma máquina com processador Core i7-2600 CPU - 3.4 GHz, tendo 8GB de memória RAM e placa de vídeo dedicada (GPU) GeForce GTX 1060 6GB.



Fig. 4. Modelo geral utilizado para classificação de COVID-19. O primeiro sub-modelo é a arquitetura de *transfer learning* utilizada para extração de características nas imagens. O segundo sub-modelo, o qual consiste em camadas densas, ativação ReLU, e camada totalmente conectada são responsáveis pelo *fine tuning* e classificação.

O *dropout* de 0.25 foi escolhido para o congelamento de 25% dos neurônios das camadas densas a cada época, com o objetivo de evitar o *overfitting*, que ocorre quando os modelos aprendem apenas no conjunto de treinamento, sem conseguir generalizar para o novos dados de teste [40]. O tamanho do *batch* é o número de dados provenientes do conjunto de treinamento a serem utilizados em memória na rede neural.

TABELA I
CONFIGURAÇÃO UTILIZADA PARA OS MODELOS DE REDES NEURAIAS
PROPOSTOS NESTE TRABALHO.

Atributo	Valor
<i>batch</i>	32
<i>dropout</i>	0.25
épocas	50
otimizador de hiper-parâmetros	adam
taxa de aprendizagem	0.001
função <i>loss</i>	<i>binary cross entropy</i> e <i>categorical cross entropy</i>

Foi adotado um tamanho de *batch* igual a 32 devido às especificações do PC utilizado nos experimentos. Foram feitos testes preliminares e *batch* escolhido proveu um treinamento sem travamentos.

Em relação ao número de épocas, selecionamos 50 pois percebemos que após esse número a variação na acurácia do modelo é muito pequena. Isso caracteriza uma convergência das redes neurais no aprendizado, não sendo necessário mais aumentar o número de épocas. O otimizador Adam [41] foi adotado nos experimentos. Ele fornece uma melhor e mais rápida otimização de hiper-parâmetros. Este otimizador é uma combinação de RMSprop e gradiente descendente estocástico. Devido aos baixos requisitos de memória, é uma boa escolha como configuração experimental. Foram escolhidas duas funções *loss*: (i) *binary cross entropy* para o treinamento de classificação binária, e (ii) *categorical cross entropy* para o treinamento multiclasse.

D. Validação dos Modelos

A validação do modelo é uma etapa importante para o processo de aprendizagem, na qual proporciona a separação dos dados para treinamento e avaliação. Neste trabalho, o método de *cross-validation* [42] foi adotado para validação dos modelos de *deep learning* propostos. Esse método consiste em particionar os dados que constituem a base de dados em K conjuntos. Com isso, o modelo é validado K vezes com $K - 1$ conjuntos sendo escolhidos para treinamento e um conjunto para teste a cada iteração. Neste trabalho o número de conjuntos $K = 10$ foi selecionado.

Este método de validação de modelos proporciona uma maior generalização dos dados no conjunto de treinamento e teste, garantindo que ele contenha exemplos suficientes para ambas as classes. Com isso, pode-se ser feita uma análise mais justa dos modelos.

E. Métricas

Será feita uma avaliação experimental de cada um dos modelos de *deep learning*, com o objetivo de verificar se cada um destes oferece uma boa capacidade de generalização para o problema deste trabalho, que é a classificação de pacientes com COVID-19. Para avaliação dos modelos, foram adotadas quatro métricas comumente utilizadas para classificação de dados. As métricas são baseadas nas classificações corretas e incorretas de dados: verdadeiros positivos (VP) e verdadeiros negativos (VN), falso positivo (FP) e no falso negativo (FN).

Esses valores mostram o quanto os modelos de *deep learning* acertaram na predição, com base no valor real.

As métricas selecionadas para avaliação foram as seguintes: acurácia (Equação 1), precisão (Equação 2), sensibilidade (Equação 3), especificidade (Equação 4) e *F1-score* (média harmônica entre precisão e sensibilidade).

$$ACC = \frac{VP + VN}{VP + VN + FP + FN} \quad (1)$$

$$PCS = \frac{VP}{VP + FP} \quad (2)$$

$$SEN = \frac{VP}{VP + FN} \quad (3)$$

$$SPE = \frac{VN}{VN + FP} \quad (4)$$

A respeito destas métricas, a acurácia provê uma análise geral do modelo, ou seja, dentre todos os dados da base de dados a métrica indica a porcentagem de dados que foram classificados corretamente. A métrica precisão provê uma porcentagem de quantas instâncias positivas foram classificadas corretamente, enquanto a métrica sensibilidade provê a porcentagem de quantas instâncias positivas totais da base foram classificadas corretamente. A respeito destas duas últimas métricas, é importante destacar que o cálculo destas na classificação multiclasse foi feita de forma balanceada, ou seja, de forma ponderada em relação a quantidade de dados presentes em cada classe. A métrica *F1-score* indica o comportamento geral da precisão e sensibilidade, e conseqüentemente da classificação como um todo. Se a *F1-score* é baixa, é entendido que uma das duas métricas é baixa.

IV. RESULTADOS E DISCUSSÃO

A. Classificação COVID-19 vs Saudável

A Tabela II apresenta os resultados de classificação de pacientes com COVID-19 vs pacientes saudáveis.

TABELA II
RESULTADOS DE CLASSIFICAÇÃO BINÁRIA PARA AS CLASSES COVID-19
vs SAUDÁVEL.

Modelo	Acurácia	Precisão	Sensibilidade	F1-score	Especificidade
DenseNet	86,03±00,27	75,35±04,58	86,03±00,27	79,62±00,53	85,92±00,40
InceptionV3	85,96±00,20	73,89±00,35	85,96±00,20	79,47±00,29	85,83±00,41
MobileNet	69,10±20,85	82,82±02,58	69,10±20,85	73,27±15,33	70,38±19,41
ResNet50	81,90±12,97	68,56±17,07	79,90±19,29	74,97±14,40	79,82±19,26
VGG-16	99,10±00,87	99,13±00,84	99,10±00,87	99,09±00,90	98,65±01,41
VGG-19	96,41±01,25	98,47±00,67	96,41±01,25	97,32±00,95	96,92±01,11

Os resultados de classificação mostram que, de forma geral, os modelos propostos apresentam boas taxas de classificação em relação as métricas de avaliação.

Destacam-se positivamente os modelos VGG. Estes modelos obtiveram uma acurácia média superior a 96%, em especial o VGG-16, o qual obteve uma acurácia média de 99,10%. Já os modelos DenseNet, InceptionV3 e ResNet50 obtiveram resultados ente 81% e 86%. O MobileNet obteve a menor acurácia média, registrando 69,10% de acertos, mesmo com o intervalo de desvio padrão considerando 10 pontos

percentuais para mais ou para menos, ainda assim não atingiu o intervalo médio dos resultados dos outros modelos, com exceção do ResNet50, o qual obteve uma média de 81,90% de acurácia e 12,97 de desvio padrão.

Os resultados de acurácia poderiam ser considerados dentro da normalidade para tarefas de classificação em aprendizagem de máquina, entretanto por se tratar de um problema de classificação binária e com aplicação de diagnóstico médico, não podem ser considerados bons resultados. Isso se deve ao fato de que um diagnóstico errado pode afetar negativamente a saúde do paciente em análise, requerendo resultados mais confiáveis e mais altos possíveis. A respeito disso, pode-se afirmar que o modelo VGG-16 foi superior em relação ao VGG-19, em termos de acurácia. Isso se deve ao fato de que os intervalos denominados pelo desvio padrão de cada um não se sobrepõem.

É possível inferir que as *features* extraídas pelas arquiteturas convolucionais que não obtiveram bons resultados não foram tão generalistas a ponto de proporcionar evidências para diferenciação de pacientes com COVID-19 e pacientes saudáveis. Tal feito foi alcançado pelos modelos baseados nas arquiteturas VGG. Uma possibilidade para explicação para este fenômeno é o fato da arquitetura VGG ter mais parâmetros que as demais arquiteturas, tonando-a mais complexa e capaz de obter as melhores *features* para a etapa de classificação com *fine tuning*.

A respeito da métrica precisão, pode-se ainda observar a superioridade dos modelos VGG, com resultados superiores a 98%. Ainda ao se analisar esta métrica, é possível confirmar que a dificuldade de classificação do ResNet foi na classificação das imagens positivas (COVID-19), ou seja, das vezes que em o classificador atestou COVID-19, obteve uma média 68,56% de acertos. O mesmo comportamento foi atestado com os outros modelos, com exceção dos que eram baseados em VGG. A mesma lógica segue-se na sensibilidade. Entretanto o MobileNet obteve os menores índices de sensibilidade, com 69,10% de acertos de todas as imagens da classe positiva, e com os modelos VGG demonstrando ainda superioridade. Além de proporcionarem melhores resultados para detecção de pacientes com COVID-19, os modelos VGG também foram superiores na métrica especificidade, ou seja, os falsos negativos foram minimizados, obtendo resultados superiores a 96% na referida métrica.

Apesar da discrepância de diferença entre os modelos nas métricas anteriores, com a métrica *F1-score*, é possível verificar que todos os modelos que obtiveram resultados mais baixos de classificação seguem um mesmo padrão comportamental para classificação de pacientes com COVID-19, diminuindo as taxas de acerto geral dos modelos.

Apesar da base de dados desbalanceada, foi possível obter bons resultados de acurácia com os modelos VGG. Os resultados são superiores aos apresentados em [27], que obteve acurácia média igual 98,75 (porém não há informação sobre o valor do desvio padrão).

B. Classificação COVID vs Pneumonia Viral

A Tabela III apresenta os resultados de classificação de pacientes com COVID-19 vs pacientes com outros tipos de

pneumonia viral.

TABELA III
RESULTADOS DE CLASSIFICAÇÃO BINÁRIA PARA AS CLASSES COVID-19 vs PNEUMONIA VIRAL.

Modelo	Acurácia	Precisão	Sensibilidade	F1-score	Especificidade
DenseNet	86,13±00,31	76,86±06,09	86,13±00,31	79,82±00,65	86,03±01,52
InceptionV3	86,00±00,20	73,96±00,34	86,00±00,20	79,52±00,28	85,30±01,26
MobileNet	86,06±00,25	75,41±04,56	86,06±00,25	79,67±00,51	86,03±00,25
ResNet50	78,80±22,76	66,76±22,77	78,80±22,76	71,92±24,06	79,49±20,92
VGG-16	98,27±01,17	98,36±01,06	98,27±01,17	98,26±01,17	98,40±01,32
VGG-19	98,79±00,77	98,81±00,75	98,79±00,77	98,77±00,80	99,05±00,92

Novamente os modelos VGG-16 e VGG-19 destacam-se na performance da classificação, com acurácia média de 98,27% e 98,79%, respectivamente. Em relação aos demais modelos, a acurácia média ficou entre 78,80% e 86,13%. Novamente o modelo ResNet apresentou-se como um modelo instável, com média de acurácia igual a 78,80%, com alto desvio padrão (22,76), podendo atingir o intervalo médio dos modelos VGG em alguma rodada de execução.

Uma vez que sabe-se que o paciente contém algum tipo de pneumonia, a classificação apresentada nesta seção é importante para determinar se o mesmo encontra-se infectado com COVID-19, ou se seu problema de saúde é ocasionado por outro tipo de pneumonia viral. Com isso, nota-se que é importante a taxa de acertos, bem como foi enfatizado na classificação COVID-19 vs saudável. Neste cenário, os modelos VGG apresentam-se como os melhores para a classificação COVID-19 vs pneumonia viral. Existe uma maior confiabilidade nestes modelos devido a alta acurácia e baixo desvio padrão. Neste caso, não é possível afirmar qual o melhor modelo, pois os desvios padrões dos modelos VGG indicam que os intervalos se sobrepõem. Novamente os modelos VGG proporcionaram melhor capacidade de generalização, diferente dos outros modelos.

Em relação a métrica de precisão, os modelos VGG-16 e VGG-19 obtiveram médias iguais a 98,27% e 98,79%, respectivamente, e com desvios padrões que se sobrepõem entre si. O modelo ResNet obteve novamente dificuldades para classificação dos pacientes com COVID-19, demonstrando que este fenômeno não foi uma particularidade da classificação COVID-19 vs saudável. A precisão média dos modelos DenseNet, InceptionV3 e MobileNet foram iguais a 76,86%, 73,96% e 75,41%, respectivamente. A sensibilidade dos modelos VGG também se sobressaíram, enquanto os demais modelos variam entre 78,80% e 86,13%. As métricas de precisão e sensibilidade demonstram que a maioria dos modelos apresentaram dificuldades para a extração de *features* da classe de pacientes com COVID-19. A métrica de *F1-score* demonstra que os modelos VGG possuem um equilíbrio na precisão e sensibilidade. Com relação a métrica especificidade, os modelos VGG superaram em mais de 10% o melhor dos demais modelos, o MobileNet, alcançando cerca de 99% com o modelo VGG-19. Os demais modelos seguem um mesmo padrão comportamental de resultados, com resultados mais baixos.

O trabalho [27] publicado na literatura a respeito de classificação de COVID-19 em exames de raio-x usando aprendizagem de máquina não reportou este tipo de classificação.

C. Classificação Multiclasse

A Tabela IV apresenta os resultados de classificação entre todas as classes anteriormente analisadas (COVID-19 *vs* saudável *vs* pneumonia viral).

TABELA IV
RESULTADOS DE CLASSIFICAÇÃO MULTICLASSE PARA AS CLASSES
COVID-19 *vs* PACIENTES SAUDÁVEIS *vs* PNEUMONIA VIRAL

Modelo	Acurácia	Precisão	Sensibilidade	F1-score	Especificidade
DenseNet	62,75±13,41	64,79±15,65	59,74±12,45	52,29±17,01	75,02±07,71
InceptionV3	69,82±11,89	60,96±09,45	60,96±09,45	56,61±13,28	73,23±09,38
MobileNet	75,61±09,67	73,18±04,31	74,62±09,09	70,96±10,63	84,71±05,86
ResNet50	46,37±00,13	31,39±12,75	46,35±00,11	29,37±00,14	65,35±02,77
VGG-16	91,07±02,36	91,68±01,85	90,91±02,65	90,39±03,05	90,25±02,61
VGG-19	91,33±01,43	91,81±01,36	90,82±01,70	90,55±01,73	91,63±01,10

Os resultados de classificação apresentaram uma queda, se comparado com os métodos de classificação anteriores. Dentre os modelos que se destacaram, VGG-16 e VGG-19 obtiveram acurácia média de 91,07% e 91,33%, respectivamente. Enquanto os outros modelos apresentaram acurácia média entre 46,37% e 75,61%. O modelo ResNet foi o destaque negativo dentre os modelos avaliados, com baixa acurácia média de 46,37%.

Esta abordagem de classificação é importante e sendo uma alternativa para o diagnóstico final de um paciente supostamente infectado pela COVID-19. As acurácias obtidas pelos modelos VGG demonstram suas altas capacidades de generalização na fase de extração de *features*. Isso demonstra que estes modelos podem ser inicialmente utilizados para o suporte na decisão final do epidemiologista, ao diagnosticar o paciente. Devido as baixas acurácias reportadas, os outros modelos não garantem confiabilidade para utilização em um cenário real.

As métricas de precisão e sensibilidade confirmam o que foi discutido anteriormente. Os modelos VGG apresentaram maiores taxas nestas duas métricas citadas, com pouca variação. Este fato indica que os modelos foram capazes de aprender cada uma das classes envolvidas neste problema de classificação multiclasse. Destaque negativo para o modelo ResNet, que obteve precisão 31,39% e sensibilidade de 46,35%. A métrica F1-score demonstra que os modelos VGG seguiram com resultados elevados, acima de 90%, demonstrando o balanceamento entre precisão e sensibilidade. Enquanto os demais modelos apresentaram o mesmo padrão comportamental, com exceção do modelo ResNet, obtendo F1-score de 29,37%. Não foi diferente na métrica especificidade, os resultados superiores a 90% confirmam a melhor aplicabilidade dos modelos VGG para detecção de classes negativas, assim como nos demais tipos de classificação binária anteriormente abordados.

De acordo com o trabalho relacionado reportado em [27], o melhor resultado de classificação reportado foi com o VGG-19 para classificação multiclasse, obtendo uma taxa de 93,48% de acurácia. Apesar das dificuldades apresentadas com a base de dados, os modelos VGG propostos no presente trabalho apresentaram resultados competitivos, e de acordo com os valores reportados para os respectivos desvios padrões, os resultados podem atingir o intervalo de acurácia reportado pelo trabalho relacionado.

A respeito do número de épocas utilizado em todos os modelos dos três tipos de classificação, notou-se que a convergência dos modelos VGG foi, de modo geral, alcançada em torno da época 20. Foi observado também que os demais modelos apresentaram poucas variações no resultado de acurácia de validação durante o treinamento dos dados. Sendo assim, o número de épocas não é um fator relevante para estes modelos, indicando que estes alcançaram uma convergência precoce, porém sem melhoria alguma no aprendizado.

V. CONCLUSÕES E TRABALHOS FUTUROS

Este trabalho apresentou uma comparação entre arquiteturas de modelos CNN presentes na literatura para classificação de COVID-19 baseado em imagens de raio-x. Foi realizado o processo de *transfer learning* com o intuito de aproveitar o treinamento realizado previamente por esses modelos para extração de *features*. Foram comparadas as arquiteturas DenseNet, InceptionV3, MobileNet, ResNet50, VGG-16 e VGG-19. Além disso, foram realizadas diferentes classificações para verificar o desempenho dos modelos. Foram feitas classificações entre pessoas com COVID-19 *vs* saudáveis, pessoas com COVID-19 *vs* com pneumonia viral, e a classificação de todas as classes anteriores.

As arquiteturas VGG apresentaram melhor desempenho em todos os tipos de classificação. Para a classificação entre pessoas com COVID-19 *vs* saudáveis, a arquitetura VGG-16 obteve os melhores resultados (com acurácia média de 99,10% e F1-score médio de 99,09%), enquanto que para os demais tipos de classificação (COVID *vs* pneumonia viral e multi classes), as arquiteturas VGG-16 e VGG-19 obtiveram resultados muito próximos para todas as métricas.

Para a tarefa de classificação entre pessoas com COVID-19 *vs* pneumonia viral, a arquitetura VGG19 apresentou o melhor desempenho, com acurácia média de 98,79% e F1-score médio de 98,77%. Com relação a classificação multi classes, todas as arquiteturas obtiveram desempenhos inferiores, o que é esperado dado o aumento da complexidade dessa tarefa. Nesse caso, as arquiteturas VGG também obtiveram resultados superiores que as demais arquiteturas, onde a arquitetura VGG-19 apresentou o melhor desempenho com acurácia e F1-score médios de 91,33% e 90,55%, respectivamente.

Ainda que os resultados para classificações binárias tenham sido satisfatórios, ainda há bastante espaço para melhoria do desempenho da classificação multi classes. Como trabalhos futuros, pretende-se analisar outras arquiteturas de CNN, assim como outros classificadores. *Extreme Machine Learning (ELM)*, *Support Vector Machine (SVM)* e *Random Forest* podem ser candidatos a substituir as camadas totalmente conectadas para a classificação de COVID-19. Além disso, pretende-se conseguir bases de dados maiores para que os modelos consigam generalizar e assim melhorar o desempenho da classificação.

AGRADECIMENTOS

Este trabalho foi realizado com apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

REFERÊNCIAS

- [1] World Health Organization and others, "Coronavirus disease 2019 (covid-19)," *World Health Organization*, 2020.
- [2] De Wit, Emmie and Van Doremalen, Neeltje and Falzarano, Darryl and Munster, Vincent J, "Sars and mers: recent insights into emerging coronaviruses," *Nature Reviews Microbiology*, vol. 14, no. 8, p. 523, 2016.
- [3] Yin, Yudong and Wunderink, Richard G, "Mers, sars and other coronaviruses as causes of pneumonia," *Respirology*, vol. 23, no. 2, pp. 130–137, 2018.
- [4] Zheng, Ying-Ying and Ma, Yi-Tong and Zhang, Jin-Ying and Xie, Xiang, "Covid-19 and the cardiovascular system," *Nature Reviews Cardiology*, vol. 17, no. 5, pp. 259–260, 2020.
- [5] Borba, Mayla Gabriela Silva and Val, Fernando Fonseca Almeida and Sampaio, Vanderson Souza and Alexandre, Marcia Almeida Araújo and Melo, Gisely Cardoso and Brito, Marcelo and Mourão, Maria Paula Gomes and Brito-Sousa, José Diego and Baía-da-Silva, Djane and Guerra, Marcus Vinitius Farias and others, "Effect of high vs low doses of chloroquine diphosphate as adjunctive therapy for patients hospitalized with severe acute respiratory syndrome coronavirus 2 (sars-cov-2) infection: a randomized clinical trial," *JAMA network open*, vol. 3, no. 4, pp. e208857–e208857, 2020.
- [6] Grasselli, Giacomo and Zangrillo, Alberto and Zanella, Alberto and Antonelli, Massimo and Cabrini, Luca and Castelli, Antonio and Cereda, Danilo and Coluccello, Antonio and Foti, Giuseppe and Fumagalli, Roberto and others, "Baseline characteristics and outcomes of 1591 patients infected with sars-cov-2 admitted to icus of the lombardy region, italy," *Jama*, 2020.
- [7] Silva, Iago R. R. and Silva, Gabriela S. L. and de Souza, Rodrigo G. and Santos, Wellington P. and Fagundes, Roberta A A, "Model based on deep feature extraction for diagnosis of alzheimer's disease," in *2019 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2019, pp. 1–7.
- [8] Antropova, N and Huynh, B and Giger, M, "Su-d-207b-06: Predicting breast cancer malignancy on dce-mri data using pre-trained convolutional neural networks," *Medical physics*, vol. 43, no. 6Part4, pp. 3349–3350, 2016.
- [9] Lakhani, Paras and Sundaram, Baskaran, "Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks," *Radiology*, vol. 284, no. 2, pp. 574–582, 2017.
- [10] Witten, Ian H and Frank, Eibe and Hall, Mark A and Pal, Christopher J, *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, 2016.
- [11] Patel, Jigar and Shah, Sahil and Thakkar, Priyank and Kotecha, K, "Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques," *Expert Systems with Applications*, vol. 42, no. 1, pp. 259–268, 2015.
- [12] LeCun, Yann and Bengio, Yoshua and Hinton, Geoffrey, "Deep learning," *nature*, vol. 521, no. 7553, p. 436, 2015.
- [13] Zhang, Tianhao and Kahn, Gregory and Levine, Sergey and Abbeel, Pieter, "Learning deep control policies for autonomous aerial vehicles with mpc-guided policy search," in *2016 IEEE international conference on robotics and automation (ICRA)*. IEEE, 2016, pp. 528–535.
- [14] Gupta, Lav and Samaka, M and Jain, Raj and Erbad, Aiman and Bhamare, Deval and Metz, Chris, "Colap: A predictive framework for service function chain placement in a multi-cloud environment," in *2017 IEEE 7th Annual Computing and Communication Workshop and Conference (CCWC)*. IEEE, 2017, pp. 1–9.
- [15] Shickel, Benjamin and Tighe, Patrick James and Bihorac, Azra and Rashidi, Parisa, "Deep ehr: A survey of recent advances in deep learning techniques for electronic health record (ehr) analysis," *IEEE journal of biomedical and health informatics*, vol. 22, no. 5, pp. 1589–1604, 2018.
- [16] Goodfellow, Ian and Bengio, Yoshua and Courville, Aaron and Bengio, Yoshua, *Deep learning*. MIT press Cambridge, 2016, vol. 1.
- [17] Akula, Aparna and Shah, Anuj K and Ghosh, Ripul, "Deep learning approach for human action recognition in infrared images," *Cognitive Systems Research*, vol. 50, pp. 146–154, 2018.
- [18] Adit Deshpande, "A beginner's guide to understanding convolutional neural networks," <https://bit.ly/2JBTPL>, 2016, accessed: August, 2018.
- [19] Li, Xiaogang and Pang, Tiantian and Liu, Weixiang and Wang, Tianfu, "Fall detection for elderly person care using convolutional neural networks," in *Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), 2017 10th International Congress on*. IEEE, 2017, pp. 1–6.
- [20] Lu, Na and Ren, Xiaodong and Song, Jinbo and Wu, Yidan, "Visual guided deep learning scheme for fall detection," in *Automation Science and Engineering (CASE), 2017 13th IEEE Conference on*. IEEE, 2017, pp. 801–806.
- [21] Afridi, Muhammad Jamal and Ross, Arun and Shapiro, Erik M, "On automated source selection for transfer learning in convolutional neural networks," *Pattern recognition*, vol. 73, pp. 65–75, 2018.
- [22] Yosinski, Jason and Clune, Jeff and Bengio, Yoshua and Lipson, Hod, "How transferable are features in deep neural networks?" in *Advances in neural information processing systems*, 2014, pp. 3320–3328.
- [23] Han, Dongmei and Liu, Qigang and Fan, Weiguo, "A new image classification method using cnn transfer learning and web data augmentation," *Expert Systems with Applications*, vol. 95, pp. 43–56, 2018.
- [24] Hemdan, Ezz El-Din and Shouman, Marwa A and Karar, Mohamed Esmail, "Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images," *arXiv preprint arXiv:2003.11055*, 2020.
- [25] Narin, Ali and Kaya, Ceren and Pamuk, Ziyet, "Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks," *arXiv preprint arXiv:2003.10849*, 2020.
- [26] Zhang, Jianpeng and Xie, Yutong and Li, Yi and Shen, Chunhua and Xia, Yong, "Covid-19 screening on chest x-ray images using deep learning based anomaly detection," *arXiv preprint arXiv:2003.12338*, 2020.
- [27] Apostolopoulos, Ioannis D and Mpesiana, Tzani A, "Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks," *Physical and Engineering Sciences in Medicine*, p. 1, 2020.
- [28] Karen Simonyan and Andrew Zisserman, "Very deep convolutional networks for large-scale image recognition," in *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, 2015.
- [29] Howard, Andrew G and Zhu, Menglong and Chen, Bo and Kalenichenko, Dmitry and Wang, Weijun and Weyand, Tobias and Andreetto, Marco and Adam, Hartwig, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *arXiv preprint arXiv:1704.04861*, 2017.
- [30] Chollet, François, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258.
- [31] "Covid-19 radiography database (winner of the covid-19 dataset award)," *Kaggle*. 2020., <https://www.kaggle.com/tawsiurrahman/covid19-radiography-database>. Acesso em 25 Abr. 2020.
- [32] Chawla, Nitesh V and Bowyer, Kevin W and Hall, Lawrence O and Kegelmeyer, W Philip, "Smote: synthetic minority over-sampling technique," *Journal of artificial intelligence research*, vol. 16, pp. 321–357, 2002.
- [33] Spampinato, Concetto and Palazzo, Simone and Giordano, Daniela and Aldinucci, Marco and Leonardi, Rosalia, "Deep learning for automated skeletal bone age assessment in x-ray images," *Medical image analysis*, vol. 36, pp. 41–51, 2017.
- [34] Bar, Yaniv and Diamant, Idit and Wolf, Lior and Greenspan, Hayit, "Deep learning with non-medical training used for chest pathology identification," in *Medical Imaging 2015: Computer-Aided Diagnosis*, vol. 9414. International Society for Optics and Photonics, 2015, p. 94140V.
- [35] "Covid-19-x-ray-ieee-latam repository," *Github*. 2020., <https://github.com/iagorichard/COVID-19-X-ray-IEEE-LATAM>. Acesso em 13 Mai. 2020.
- [36] Deng, Jia and Dong, Wei and Socher, Richard and Li, Li-Jia and Li, Kai and Fei-Fei, Li, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*. IEEE, 2009, pp. 248–255.
- [37] Huang, Gao and Liu, Zhuang and Van Der Maaten, Laurens and Weinberger, Kilian Q, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.
- [38] Szegedy, Christian and Vanhoucke, Vincent and Ioffe, Sergey and Shlens, Jon and Wojna, Zbigniew, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818–2826.
- [39] He, Kaiming and Zhang, Xiangyu and Ren, Shaoqing and Sun, Jian, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

- [40] Srivastava, Nitish and Hinton, Geoffrey and Krizhevsky, Alex and Sutskever, Ilya and Salakhutdinov, Ruslan, “Dropout: a simple way to prevent neural networks from overfitting,” *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.
- [41] Diederik P. Kingma and Jimmy Ba, “Adam: A method for stochastic optimization,” in *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, 2015. [Online]. Available: <http://arxiv.org/abs/1412.6980>
- [42] Shao, Jun, “Linear model selection by cross-validation,” *Journal of the American statistical Association*, vol. 88, no. 422, pp. 486–494, 1993.



Iago Rodrigues é doutorando do Programa de Pós-graduação em Ciência da Computação da Universidade Federal de Pernambuco (UFPE). É mestre em Engenharia da Computação pela Universidade de Pernambuco (UPE). Atualmente é pesquisador do Grupo de Redes e Telecomunicações (GPRT). São suas áreas de interesse atuais: Aprendizado de máquina, processamento de imagens e visão computacional.



Guto Leoni Santos é graduado em sistemas de informação pela Universidade de Pernambuco (UPE), mestre em ciência da computação pela Universidade Federal de Pernambuco (UFPE) e aluno de doutorado em Ciência da Computação na UFPE. Faz parte do Grupo de Pesquisa em Redes e Telecomunicações (GPRT) desde 2016. Seus interesses de pesquisa incluem: computação em nuvem, avaliação de desempenho, Internet das Coisas, *fog computing*, 5G, *machine learning* e *deep learning*.



Djamel Fawzi Hadj Sadok é professor do Centro de Infomática da Universidade Federal de Pernambuco (UFPE) desde 1993. Seus interesses de pesquisa incluem sistemas de comunicação, segurança de redes, computação em nuvem e aprendizagem de máquina. Atualmente, ele lidera o Grupo de Pesquisa em Redes de Computadores e Telecomunicações (GPRT) e coordena uma série de projetos de pesquisa.



Patricia Takako Endo é professora adjunta da Universidade de Pernambuco (UPE), membro permanente do Programa de Pós-Graduação em Engenharia da Computação da UPE e pesquisadora no Grupo de Pesquisa em Redes de Computadores e Telecomunicações (GPRT). Possui doutorado em Ciência da Computação pela Universidade Federal de Pernambuco (UFPE) e suas áreas de interesse são: cloud computing, gerenciamento de recursos, deep learning e análise de dados de saúde.