

Otimização de Hiperparâmetros de Redes Neurais Profundas para Detecção de Cardiomegalia em radiografias do tórax.

Saulo Enock Rodrigues Fernandes, Ricardo Costa da Silva Marques,
João Dallyson Sousa de Almeida, Anselmo Cardoso de Paiva, Geraldo Braz Junior

¹Núcleo de Computação Aplicada – Vision and Image Processing Lab
Universidade Federal do Maranhão (UFMA) Campus do Bacanga – São Luís – MA

{saulorodrigues, ricardo.marques, jdallyson, paiva, geraldo}@nca.ufma.br

Abstract. *Chest X-Ray is one of the most commonly used exams for diagnosing thoracic diseases. However, due to the complexity of chest diseases, there is a demand for skilled and experienced physicians to mitigate the error chances in diagnosing such pathology. Multiple approaches using Neural Networks have been devised to aid specialists in detecting chest diseases such as lung opacity and pneumonia. This paper proposes a methodology striving to help health professionals diagnose thoracic diseases. Therefore, we present an image quantity balancing and hyperparameter optimization approach. The tests were focused on cardiomegaly classification and the results obtained proved to be promising with values of 0.919 (AUC), 0.873 (Accuracy), 0.842 (Precision), 0.876 (F1-Score), 0.913 (Sensitivity), 0.790 (Specificity), approaching the best metrics present in the literature.*

Keywords: *Deep Neural Networks; Diagnosis of Cardiomegaly; Hyperparameters Optimization.*

Resumo. *A radiografia do tórax é um dos exames mais comumente utilizados para o diagnóstico de doenças torácicas. Devido à complexidade das doenças torácicas, existe a demanda de médicos experientes e habilidosos na área para mitigar as chances de erros no diagnóstico de tais patologias. Diversas pesquisas utilizando Redes Neurais foram elaboradas para auxiliar especialistas na detecção de doenças da região do tórax como, por exemplo, a opacidade pulmonar e a pneumonia. Este trabalho apresenta uma metodologia que pretende fornecer auxílio aos profissionais da saúde no diagnóstico de doenças da região do tórax. É proposta uma abordagem baseada no balanceamento de quantidades de imagens e otimização de hiperparâmetros do modelo. Os testes foram focados na classificação da cardiomegalia e os resultados obtidos se provaram promissores alcançando 0,919 de AUC, 0,873 de acurácia, 0,842 de precisão, 0,876 de F1-Score, 0,913 de sensibilidade e 0,790 de especificidade, aproximando-se das melhores métricas presentes na literatura.*

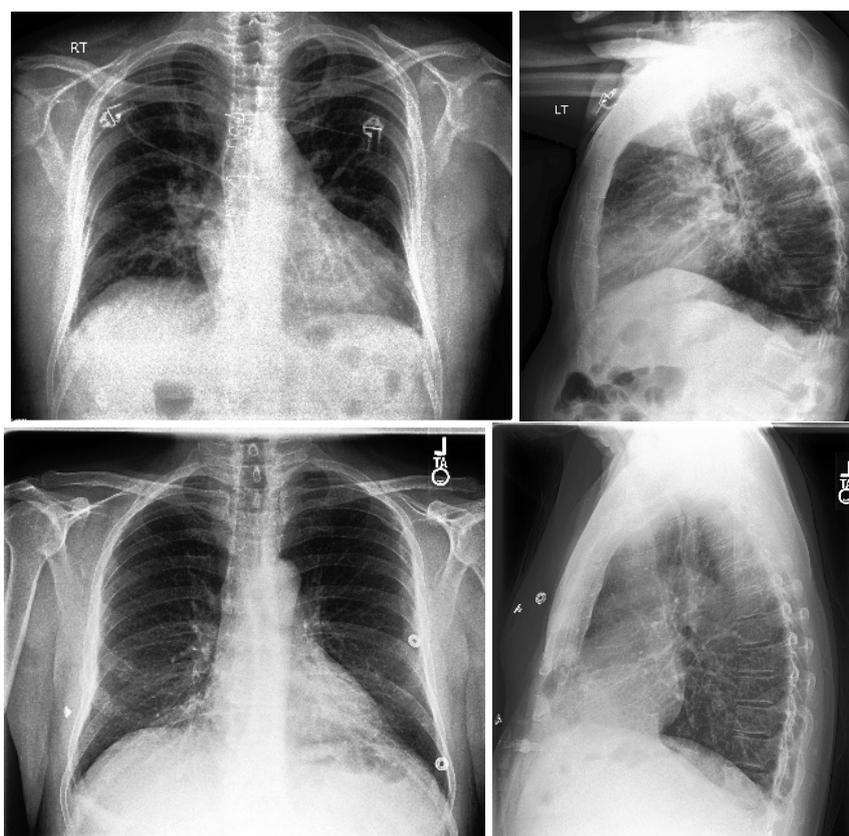
Palavras-chave: *Redes Neurais Profundas; Diagnóstico de Cardiomegalia; Otimização de Hiperparâmetros.*

1. Introdução

O exame de Raio-X (Figura 1), também identificado como radiografia do tórax (*Chest X-ray* ou *CXR*), é um dos exames radiológicos mais costumeiros empregados no diagnóstico de patologias torácicas, em particular para doenças relacionadas ao coração e

aos pulmões [Anderson et al. 2015]. Entre essas doenças, existem as com consequências extenuantes, por exemplo, a atelectasia [Randtke et al. 2015], a qual consiste no colapso de um dos pulmões. E há também aquelas essencialmente mortais como o pneumotórax [Bellaviti et al. 2016] e a cardiomegalia [Agostoni et al. 2000].

Figura 1. Exames de raios X frontal e lateral do *CheXpert Dataset*. Caso de cardiomegalia (acima), caso saudável (abaixo).



Fonte: *CheXpert Dataset*, [Irvin et al. 2019].

Cardiomegalia é uma denominação abrangente para uma variedade de condições causadoras do aumento do coração, as quais geralmente permanecem sem diagnóstico até que os sintomas se manifestem. Segundo a Sociedade Brasileira de Cardiologia, essa patologia pode ser caracterizada pelo índice cardiorádico acima de 0,5 (obtido em exames laterais e pósterio-anterior ou PA), podendo ser descoberta em pacientes que apresentam tipos graves de cardiomiopatia, como, por exemplo, na cardiopatia chagásica crônica e na cardiomiopatia dilatada [Silva 2011]. É uma patologia grave, em que a dilatação do coração pode levar a um quadro letal de insuficiência cardíaca [Bui et al. 2011, Cardenas et al. 2020], assim, tanto a prevenção quanto o tratamento clínico são essenciais.

Na literatura existem diversas pesquisas baseadas em Redes Neurais Convolucionais (CNNs) empregadas à *CheXpert Dataset*. Esse estudo retrata a metodologia elaborada e compara com os resultados presentes em outras pesquisas. Por exemplo, Irvin et al. (2019) apresentou o *CheXpert Dataset* desenvolvendo sobre a construção da base de dados do início ao fim, desde o recolhimento de exames e a extração de rótulos, até o tratamento de imagens e a configuração de treino aplicada.

Allaouzi e Ahmed (2019) apresentam uma combinação da eficácia das Redes Neurais Convolucionais para a extração de características das imagens e os classificadores supervisionados de múltiplos rótulos na tarefa de detecção de patologias torácicas nos exames de Raio-X. Esse método atingiu o melhor resultado no conjunto de dados *ChestX-ray14 Dataset*.

Pham et al. (2020) construíram uma abordagem inovadora na qual os rótulos incertos do *CheXpert Dataset* (Tabela 1) e as dependências hierárquicas existentes entre doenças da base de dados são explorados de forma eficiente, superando estudos anteriores com bons resultados.

Yang et al. (2019) assumiram os rótulos incertos como uma classe distinta durante o processo de treinamento, permitindo ao modelo produzir menos previsões de confiança em casos ambíguos, oposto aos modelos treinados com rótulos binários (*U-ones*, etc). Os resultados de todas as abordagens mencionadas são retratados na Tabela 6.

Diego et al. (2020) validaram cinco modelos distintos (EfficientNetB2, DenseNet121, Xception, InceptionV3, and MobileNet) na tarefa de detecção automática de cardiomegalia, treinando-os com um *perceptron* multicamadas customizado, onde se substituiu as camadas totalmente conectadas originais por um *perceptron* de 3 camadas, padronizando as camadas de classificação para todas as CNNs. Os experimentos foram realizados nos datasets PadChest e OpenI. Obtiveram com a métrica *Area Under a Curve (AUC)* de 0,91 no modelo com EfficientNet e 0,88 com a *Inception*.

Nesse contexto, este trabalho demonstra uma metodologia que visa estudar e aplicar Redes Neurais Profundas empregadas ao diagnóstico de Cardiomegalia em imagens de exames radiográficos do tórax, do *CheXpert Dataset*. Para cumprir esse objetivo, utiliza-se uma abordagem pautada no balanceamento de quantidades de imagens utilizadas no treino e executada concomitantemente com a otimização de hiperparâmetros para a entrega do modelo otimizado para o problema.

Como principais contribuições temos, uma proposta de balanceamento de quantidades utilizada no treino para igualar as proporções de casos positivos e negativos, evidenciando a relevância de um conjunto de dados balanceado. E a seleção dentre os diferentes conjuntos de imagens, utilizando exames de pacientes apenas frontais, laterais ou ambos juntos, demonstrando sua significância na aprendizagem do modelo. Por último, diferente da maioria dos trabalhos presentes na literatura, os quais se limitam à observação de apenas algumas métricas, principalmente a *AUC*, os resultados obtidos nesta pesquisa são observados com um total de 6 métricas distintas.

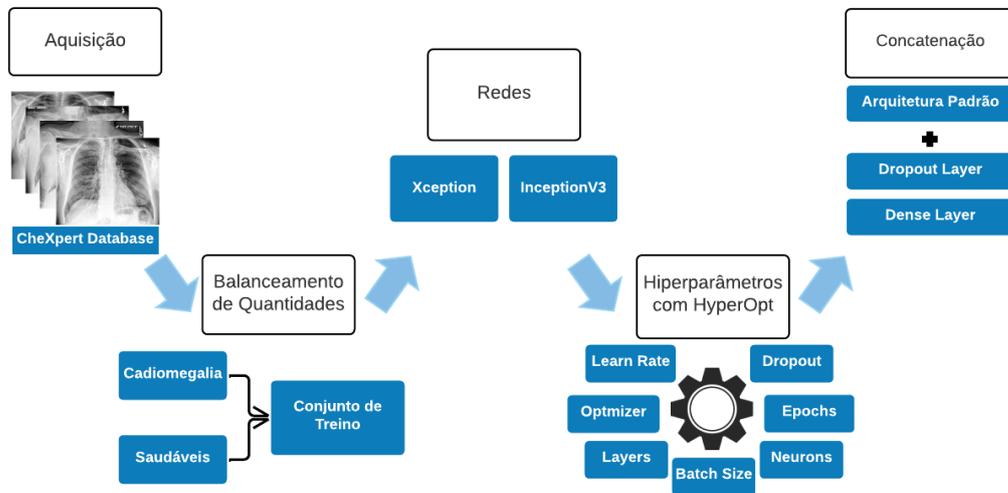
2. Metodologia

A metodologia proposta divide-se em cinco etapas principais, sendo estas: (1) Aquisição e Preparação de Dados, uma etapa adicional de (2) Balanceamento de Quantidade de Imagens; (3) Escolha de Arquiteturas, subdividida em observações sobre cada rede aplicada ao problema; (4) Otimização de Hiperparâmetros, e por fim uma etapa de (5) Concatenação de Camadas. As etapas estão ilustradas na Figura 2.

2.1. Aquisição e Preparação dos Dados

O *CheXpert dataset* [Irvin et al. 2019], selecionado para esta pesquisa, é um conjunto de dados disponibilizado ao público para o estudo de exames radiográficos do tórax, con-

Figura 2. Etapas da metodologia proposta.



Fonte: Autor.

tendo 224.316 exames de raio-X do tórax referentes a 65.240 pacientes.

Os exames, podendo ser frontais ou laterais, foram realizados no *Stanford Hospital* entre o período de outubro de 2002 e julho de 2017, feitos tanto em ambulatórios como em centros de internamento. Cada uma das imagens de raio-X é classificada dentre catorze descobertas clínicas distintas: Atelectasia, Cadiomegalia, Cadiomediastino Alargado, Consolidação, Pneumonia, Pneumotórax, Edema Pulmonar, Opacidade Pulmonar, Lesão Pulmonar, Derrame Pleural, Pleural Outros, Fratura, Dispositivos de Suporte e Sem Descoberta. Essa distribuição de achados clínicos está disposta na Tabela 1.

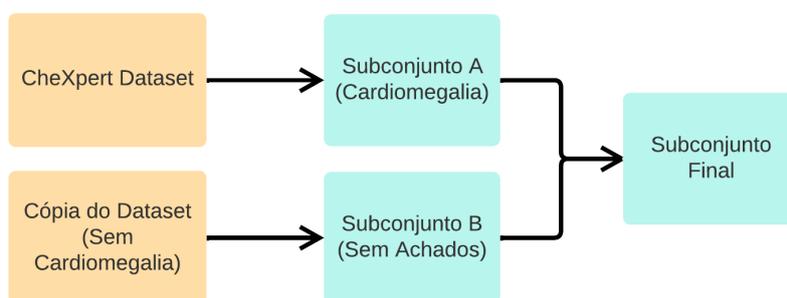
A preparação dos dados ocorreu na forma de alterações pontuais nas instâncias da base de dados. As imagens foram redimensionadas em dois tamanhos diferentes se adequando às entradas das respectivas redes, sendo $224 \times 224 \text{ pixels}$ para a VGG16, ResNet50 e DenseNet121, já para a Inceptionv3 e Xception foram $299 \times 299 \text{ pixels}$ de dimensão. Os rótulos incertos foram tratados através da abordagem de *U-ignore* (ignorar as classificações incertas), para que apenas os rótulos positivos e negativos sejam utilizados no treino.

2.1.1. Balanceamento de Quantidade de Imagens

O balanceamento do número total de imagens utilizadas é proporcionado por meio da criação de um subconjunto **A** formado por casos positivos de cadiomegalia, simultaneamente com um subconjunto **B** compreendido por imagens de pacientes classificados na classe Sem Achados e com quantidade aproximada ao subconjunto **A**. O processo é ilustrado na Figura 3.

O modelo é treinado usando um conjunto de volume correspondente ao número de imagens do rótulo de interesse, visando evitar que a quantidade no aprendizado de tal modelo seja muito grande para ser alienado de exemplos negativos. A desigualdade proporcional entre exemplos positivos e negativos da totalidade do *dataset* pode ser cons-

Figura 3. Processo de balanceamento de quantidades.



Fonte: Autor.

tatada na Tabela 1. Apenas os casos positivos de ambas as classes foram utilizados na criação dos subconjuntos A e B.

Tabela 1. Distribuição da quantidade de imagens do *Chexpert Dataset*.

Patologia	Positivo (%)	Incerto (%)	Negativo (%)
Sem achados	16627 (8,86)	0 (0,0)	171014 (91,14)
Cardiomegalia	23002 (12,26)	6597 (3,52)	158042 (84,23)

Fonte: [Irvin et al. 2019]

2.2. Escolha de Arquiteturas

As arquiteturas VGG16, DenseNet121 e ResNet50 foram selecionadas por sua utilização em pesquisas relacionadas a esse trabalho, por outro lado, as redes InceptionV3 e Xception foram selecionadas por apresentarem melhor valor de especificidade em comparação com as demais, conforme apresentado na Tabela 4. Além disso, a rede utilizada nos testes é inicializada com pesos ImageNet Large Scale Visual Recognition Challenge (ILSVRC) e tem suas camadas descongeladas para aprendizado em imagens de raios-X.

2.2.1. VGGNet

VGGNet, ou VGG-16, é uma arquitetura de Rede Neural Convolutacional concebida por Karen Simonyan e Andrew Zisserman da *University of Oxford* no ILSVRC (*ImageNet Large Scale Visual Recognition Challenge*) 2014 [Simonyan and Zisserman 2014].

É uma rede marcada pela sua simplicidade e arquitetura uniforme, com 5 etapas de camadas convolucionais 3x3 uma após a outra em profundidade progressiva, ela lida com na redução da altura e da largura da arquitetura através de camadas de *MaxPooling*. Sua saída se dá por meio de duas camadas completamente conectadas com 4.096 nós cada, seguido de uma função *Softmax*.

Desta forma, a arquitetura VGG16 dispõe de boa eficácia na extração de características de imagens, superando com larga margem as versões dos modelos apresentados nas competições ILSVRC-2012 e ILSVRC-2013. Sua desvantagem é seu número de parâmetros demasiado de 130 milhões, tornando-a uma rede computacionalmente custosa no caso do treinado ser inicializado com pesos aleatórios.

2.2.2. DenseNet

A rede DenseNet foi proposta no artigo *Densely Convolutional Networks* [Huang et al. 2016]. Esta arquitetura foi elaborada para mitigar a perda de informação decorrente do fato das redes possuírem muitas camadas.

Entre elas, cada camada recebe entradas complementares de todas as camadas anteriores e propaga seu próprio mapa de recursos para todas as camadas subsequentes através de conexões. Cada camada obtém o resultado atual e uma cópia da entrada da camada anterior. Deste modo, as camadas recebem os mapas de características das camadas precedentes, tornando a arquitetura mais robusta na atualização dos mapas de características.

2.2.3. ResNet

A Rede Neural Residual, também conhecida como ResNet [He et al. 2015], venceu a competição ImageNet 2015 com a ideia de acrescentar aprendizado residual ao conceito das Redes Neurais Convolucionais tradicionais.

Resolvendo assim o problema da dispersão de gradientes e degradação de precisão (no conjunto de treino) em Redes Neurais Profundas, de modo que a rede consiga receber cada vez mais profundidade mantendo o controle sobre a precisão e reduzindo o custo de treinamento.

A ResNet apresenta uma conexão de atalho de identidade, pulando uma ou mais camadas [He et al. 2015]. Deste modo, fica mais simples deixar as camadas empilhadas encaixarem num mapeamento residual do que deixá-las se envolver imediatamente no mapeamento subjacente desejado.

2.2.4. InceptionV3

Inception *network*, ou GoogLeNet, possui 42 camadas e a menor taxa de erro da competição, sendo a arquitetura classificada em segundo lugar no ILSVRC (*ImageNet Large Scale Visual Recognition Competition*) 2015 [Szegedy et al. 2015].

Ao contrário da VGGNet, em vez de tornar a rede mais profunda, a Inception faz a arquitetura mais larga, com camadas convolucionais de diferentes tamanhos em um mesmo nível, concatenando a saída. A arquitetura foi atualizada com melhoras iterativas, a segunda versão aperfeiçoou o conceito das camadas da rede com novos princípios, fatorando as camadas em tamanhos menores, aumentando por 2,78 vezes mais a eficiência e diminuindo o custo computacional.

A terceira versão manteve a diminuição de custo da arquitetura, seguida da integração de *Label Smoothing*, do otimizador RMSprop e da normalização de *Batch* nos classificadores auxiliares da primeira versão. Deste modo, a arquitetura Inception é eficiente com baixo custo computacional.

2.2.5. Xception

A Xception foi apresentada por Francois Chollet sendo uma versão alterada da rede Inception [Szegedy et al. 2015], substituindo módulos padrão da arquitetura Inception por uma convolução modificada separável em profundidade, criou uma arquitetura de desempenho superior a Inception-v3 em ambos os conjuntos de dados ImageNet ILSVRC e JFT [Chollet 2017].

A disposição da arquitetura Xception é constituída de camadas de convolução separáveis tratadas como os blocos da Inception, localizadas ao longo de toda a estruturação da rede. Além do que, existem conexões residuais (tal qual proposto na ResNet [He et al. 2015]) por todos os fluxos, contribuindo para a acurácia do modelo e sendo uma parte crucial da rede Xception.

2.3. Otimização de Hiperparâmetros

Esta seção descreve a etapa de otimização de hiperparâmetros do modelo, realizada com apoio da biblioteca HyperOpt [Bergstra et al. 2015] baseado no *Sequential Model-Based Optimization* (SMBO), assim como os ajustes empíricos efetuados durante a fase preliminar deste trabalho.

Os testes iniciais foram aplicados nas cinco redes mencionadas nas seções anteriores, treinadas por 100 épocas cada, com divisão de 80%, 10% e 10% para treino, validação e teste, nessa ordem. Os resultados indicam uma performance superior das arquiteturas InceptionV3 e Xception comparativamente as demais, portanto essas duas redes foram selecionadas como foco dos experimentos descritos a seguir.

Os hiperparâmetros escolhidos para a otimização foram o *Dropout*, *Learn Rate*, *Batch Size*, *Optimizer* e *Epoch*. Adicionalmente, também é adquirida a configuração da estrutura a ser concatenada na saída das arquiteturas, sendo apurados o número de camadas e seu referente número de neurônios. Cada um desses parâmetros possui seu respectivo espaço de busca utilizado na computação da configuração ideal, como pode ser constatado na Tabela 2.

Tabela 2. Espaço de Busca de Hiperparâmetros utilizado na Otimização.

Hiperparâmetro	Espaço de Busca
<i>Batch Size</i>	16, 32, 64, 128, 256
<i>Dropout</i>	Entre 0.1 e 0.5
<i>Learn Rate</i>	0.1, 0.01, 0.001, 0.0001, 0.00001
<i>Optimizer</i>	'Adam', 'SGD', 'Nadam', 'RMSprop'
<i>Epoch</i>	100, 200, 300
<i>Layers</i>	1, 2, 3
<i>Neurons</i>	128, 256, 512, 1024

Fonte: Autor

A última aplicação da otimização visou maximizar a métrica de especificidade e minimizar a *Loss* do modelo. A configuração obtida é exposta na Tabela 3.

Tabela 3. Configuração Final de Hiperparâmetros

<i>Batch Size</i>	<i>Dropout</i>	<i>Learn Rate</i>	<i>Optimizer</i>	<i>Epoch</i>	<i>Layers</i>	<i>Neurons</i>
16	0.452	0.001	Adam	100	2	128

Fonte: Autor

2.4. Concatenação de Camadas

Nesta etapa a camada de classificação do modelo foi alterada, adicionando-se uma camada *Dropout* e uma combinação de camadas densas antes da camada final, baseando-se no trabalho proposto por Fernandes et al. (2021). A configuração dessa concatenação é obtida na fase de Otimização (Seção 2.3) para assim definir o melhor conjunto final de camadas e então gerar a melhor topologia possível para o modelo. Esta metodologia se baseia no fato de que as arquiteturas foram previamente treinadas na ImageNet, logo, nem todas as camadas são adequadas para o problema tratado. Por isso, estima-se, na fase de otimização, qual é a melhor composição das camadas de saída para a detecção de cardiomegalia. A concatenação pode ser observada na Figura 4.

Figura 4. Estrutura da Concatenação.



Fonte: Autor.

3. Resultados e Discussão

Inicialmente foram realizados experimentos com ajustes manuais dos hiperparâmetros e o balanceamento da quantidade de imagens para verificar o melhor *backbone* a ser utilizado. Todas as redes foram testadas com os mesmos parâmetros. Os resultados obtidos estão disponíveis na Tabela 4.

Tabela 4. Resultados iniciais com ajustes empíricos e balanceamento de quantidades. ¹Acurácia, ²Precisão, ³F1-Score, ⁴Sensibilidade, ⁵Especificidade

Arquitetura	Épocas	AUC	Acc. ¹	Prec. ²	F1. ³	Sens. ⁴	Spec. ⁵
Xception	100	0,879	0,803	0,728	0,815	0,925	0,600
InceptionV3	100	0,877	0,809	0,725	0,818	0,939	0,596
VGG16	100	0,893	0,813	0,601	0,747	0,986	0,387
ResNet50	100	0,852	0,773	0,561	0,713	0,980	0,250
DenseNet121	100	0,903	0,825	0,650	0,782	0,981	0,469
Xception	200	0,867	0,803	0,752	0,812	0,881	0,701
InceptionV3	200	0,857	0,796	0,740	0,799	0,857	0,708
VGG16	200	0,904	0,817	0,565	0,720	0,991	0,270
ResNet50	200	0,873	0,802	0,537	0,697	0,993	0,190
DenseNet121	200	0,886	0,805	0,613	0,754	0,980	0,391

Fonte: Autor

Deste modo, tomando estes resultados como embasamento, foram selecionadas as arquiteturas Xception e InceptionV3 para serem utilizadas em novos experimentos realizados incluindo a otimização de hiperparâmetros e a etapa de concatenação.

Os conjuntos de treino, validação e teste foram separados de maneira aleatória a partir do conjunto obtido por meio do método apresentado na Seção 2.1.1, respectivamente em 70%, 15% e 15%. Após isso, com a configuração dos parâmetros foram obtida no conjunto de treino com o método descrito na Seção 2.3.

Os resultados demonstrados são de treinamentos efetuados com três distribuições distintas de imagens: apenas frontais, apenas laterais e apenas pacientes com ambos os

tipos de exame. A quantidade de imagens de cada uma dessas distribuições são: 23.103 de exames frontais, 7.646 de exames laterais, e 19.437 frontais e laterais. A Tabela 5 mostra os resultados dos experimentos.

Tabela 5. Resultados finais da abordagem de balanceamento de quantidades e otimização de hiperparâmetros aplicada à classificação de cardiomegalia.¹Acurácia, ²Precisão, ³F1-Score, ⁴Sensibilidade, ⁵Especificidade

Arquitetura	AUC	Acc. ¹	Prec. ²	F1. ³	Sens. ⁴	Spec. ⁵
Imagens frontais						
Xception	0,914	0,871	0,859	0,886	0,915	0,797
InceptionV3	0,905	0,863	0,859	0,888	0,918	0,788
Imagens laterais						
Xception	0,882	0,823	0,772	0,811	0,854	0,778
InceptionV3	0,879	0,834	0,802	0,835	0,870	0,798
Imagens frontais e laterais						
Xception	0,919	0,873	0,842	0,876	0,913	0,790
InceptionV3	0,915	0,879	0,853	0,887	0,922	0,795

Fonte: Autor

A partir dos resultados apresentados, notou-se que o experimento com a rede Xception (Frontal e Lateral) atingiu o melhor resultado de *AUC* de **0,919**, a InceptionV3 (Frontal e Lateral) obteve melhor sensibilidade de **0,922** e Xception (Frontal) obteve melhor acurácia de **0,871** e precisão de **0,859**. A comparação das abordagens testadas para a classificação binária (normal/cardiomegalia) com os resultados presentes na literatura podem ser observados na Tabela 6.

Tabela 6. Comparação dos resultados com trabalhos da literatura.

Métodos	AUC
U-Ignore [Irvin et al. (2019)]	0,818
U-Zeros [Irvin et al. (2019)]	0,811
U-Ones [Irvin et al. (2019)]	0,858
U-MultiClass [Irvin et al. (2019)]	0,821
U-SelfTrained [Irvin et al. (2019)]	0,833
U-Ignore+LP [Allaouzi et al. (2019)]	0,720
U-Ignore+BR [Allaouzi et al. (2019)]	0,720
U-Ignore+CC [Allaouzi et al. (2019)]	0,700
U-Ignore+CT [Pham et al. (2020)]	0,780
U-Zeros+CT+LSR [Pham et al. (2020)]	0,806
U-Ones+CT+LSR [Pham et al. (2020)]	0,825
Ensemble with TTA [Pham et al. (2020)]	0,909
Nossa Abordagem (Preliminar)	0,879
Nossa Abordagem (Final)	0,919

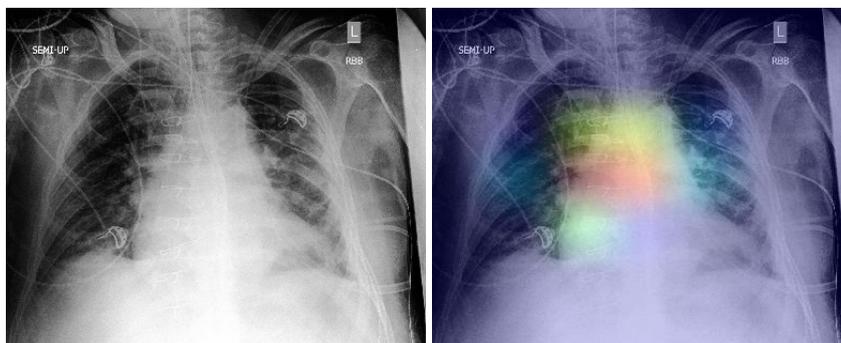
Fonte: Autor

Ao contrário dos demais trabalhos utilizando o *Chexpert Dataset* presentes na literatura (Tabela 6), nossa pesquisa usou um menor volume de imagens na etapa de treino. Enquanto Irvin et al. (2019) usaram a totalidade do banco de dados e Allaouzi et al. (2019) utilizaram 134.327 imagens após ignorar rótulos incertos, neste trabalho com a seleção descrita na Seção 2.1.1 foram utilizados pouco mais de 20.000 exames na distribuição mais numerosa. Outra diferença está na classificação binária utilizada, oposta à identificação de múltiplos rótulos aplicados por Irvin et al. (2019), Pham et al. (2020) e Allaouzi et al. (2019), além da observação de seis métricas diferentes, e não apenas a *Area Under the Curve (AUC)*.

Nas Figuras 5 e 6 há exemplos de *Heat Maps* de detecção positiva de Cardiome-

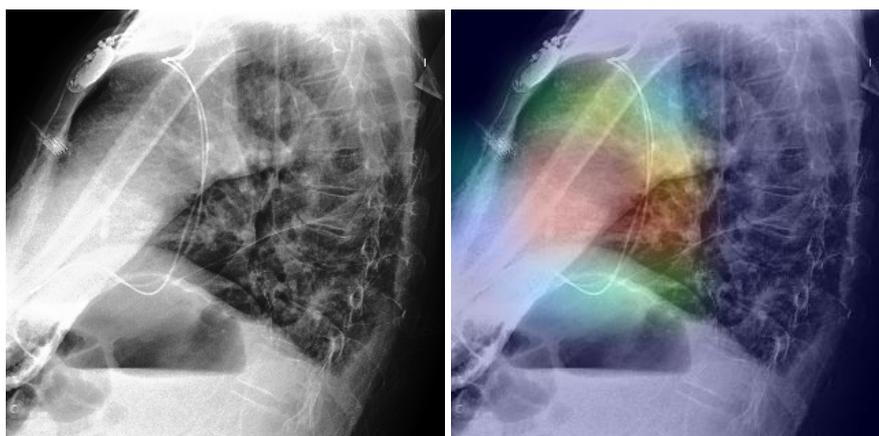
galia do modelo InceptionV3 (Frontal and Lateral) posteriormente ao processo de treino. A detecção ocorreu em ambos os tipos, Frontal e Lateral, resultando em uma ativação concisa na região de interesse do órgão afetado pela patologia.

Figura 5. Exame frontal com detecção do modelo Xception frontal e lateral. *Original e Heat Map*.



Fonte: Autor.

Figura 6. Exame lateral com detecção do modelo Xception frontal e lateral. *Original e Heat Map*.



Fonte: Autor.

Dentre todos os resultados, é pertinente apontar as limitações da abordagem apresentada. Não foram feitos testes com imagens de outras bases de dados tal como [Cardenas et al. 2020]. Apesar das boas métricas de *AUC*, precisão e sensibilidade, não apresentamos a mesma qualidade na especificidade. Por último, no meio das 14 distintas classificações existentes no *Chexpert Dataset*, este trabalho limitou-se à detecção de Cardiomegalia.

4. Conclusão e Trabalhos Futuros

A abordagem demonstrada neste trabalho utiliza-se do balanceamento das quantidades de imagens para equalizar o volume de imagens utilizadas nos conjuntos de treino, validação e teste, evitando que o modelo fique equivocadamente inclinado para uma classe devido à discrepância anteriormente existente em seus números. Além disso, com a otimização de hiperparâmetros, foi identificada uma configuração aproxima ao ideal para o problema, dessa forma, potencializando o uso dos conjuntos de imagens na tarefa de detecção de doenças torácicas em imagens de exames de Raios-X.

Os experimentos executados alcançaram bons resultados na métrica *AUC*, utilizada como principal parâmetro de validação pelos trabalhos presentes na literatura. Os valores obtidos foram de 0,915 com a rede InceptionV3 e 0,919 para Xception. Portanto, vale evidenciar que os resultados obtidos nessa pesquisa se provaram robustos e competitivos ao superarem a métrica do melhor resultado atualmente mostrado na literatura referente à classificação de Cardiomegalia.

Desse modo, com o propósito de aprimorar a metodologia apresentada, como trabalhos futuros, pretende-se adicionar uma etapa de pré processamento de imagens e utilizar estratégias como *Data Augmentation* para ampliar as quantidades de dados relevantes para as classes com poucos indivíduos. Procuramos também continuar com a expansão da pesquisa para compreender mais classes do *CheXpert Dataset*, para fins de validação da metodologia utilizada em um grande volume de dados. Outra possibilidade é expandir este trabalho para outros bancos de dados, visando a detecção de outras patologias pulmonares como, por exemplo, a Covid-19.

5. Agradecimentos

Os autores desejam expressar seus agradecimentos ao apoio financeiro concedido pela CAPES (Finance code 001), CNPQ e FAPEMA para o desenvolvimento desta pesquisa e ao Núcleo de Computação Aplicada (NCA-UFMA) pelo suporte técnico e científico.

Referências

- Agostoni, P., Cattadori, G., Guazzi, M., Palermo, P., Bussotti, M., and Marenzi, G. (2000). Cardiomegaly as a possible cause of lung dysfunction in patients with heart failure. *American Heart Journal*, 140(5):A17–A21.
- Allaouzi, I. and Ahmed, M. B. (2019). A novel approach for multi-label chest x-ray classification of common thorax diseases. *IEEE Access*, 7:64279–64288.
- Anderson, L., Dean, A., Falzon, D., Floyd, K., Baena, I., Gilpin, C., Glaziou, P., Hamada, Y., Hiatt, T., and A. Char, e. a. (2015). *Global tuberculosis report 2015*. World Health Organization, 20th edition.
- Bellaviti, N., Bini, F., Pennacchi, L., Pepe, G., Bodini, B., Ceriani, R., D’Urbano, C., and Vaghi, A. (2016). Increased incidence of spontaneous pneumothorax in very young people: Observations and treatment. *Chest Journal*, 150(4):560A.
- Bergstra, J., Komer, B., Eliasmith, C., Yamins, D., and Cox, D. D. (2015). Hyperopt: a python library for model selection and hyperparameter optimization. *Computational Science & Discovery*, 8(1):014008.
- Bui, A. L., Horwich, T. B., and Fonarow, G. C. (2011). Epidemiology and risk profile of heart failure. *Nature Reviews Cardiology*, 8(1):30–41.
- Cardenas, D., Junior, J. F., Moreno, R., Rebelo, M., Krieger, J., and Gutierrez, M. (2020). Multicenter validation of convolutional neural networks for automated detection of cardiomegaly on chest radiographs. In *Anais do XX Simpósio Brasileiro de Computação Aplicada à Saúde*, pages 179–190, Porto Alegre, RS, Brasil. SBC.
- Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1251–1258.

- Fernandes, V., Junior, G. B., de Paiva, A. C., Silva, A. C., and Gattass, M. (2021). Bayesian convolutional neural network estimation for pediatric pneumonia detection and diagnosis. *Computer Methods and Programs in Biomedicine*, 208:106259.
- He, K., Zhang, X., Ren, S., and Sun, J. (2015). Deep residual learning for image recognition. *CoRR*, abs/1512.03385.
- Huang, G., Liu, Z., and Weinberger, K. Q. (2016). Densely connected convolutional networks. *CoRR*, abs/1608.06993.
- Irvin, J., Rajpurkar, P., Ko, M., Yu, Y., Ciurea-Ilcus, S., Chute, C., Marklund, H., Haggoo, B., Ball, R., Shpanskaya, K., Seekins, J., Mong, D. A., Halabi, S. S., Sandberg, J. K., Jones, R., Larson, D. B., Langlotz, C. P., Patel, B. N., Lungren, M. P., and Ng, A. Y. (2019). Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison. *AAAI-19/IAAI-19/EAAI-20 Proceedings*, 33(1):590–597.
- Pham, H. H., Le, T. T., Tran, D. Q., Ngo, D. T., and Nguyen, H. Q. (2020). Interpreting chest x-rays via cnns that exploit hierarchical disease dependencies and uncertainty labels.
- Randtke, M. A., Andrews, B. P., and Mach, W. J. (2015). Pathophysiology and prevention of intraoperative atelectasis: A review of the literature. *Journal of PeriAnesthesia Nursing*, 30(6):516–527.
- Silva, J. D. D. d. (2011). Repercussões da cardiomegalia na distribuição regional da ventilação na caixa torácica em indivíduos com cardiopatia chagásica crônica. *Programa de Pós-Graduação em Fisioterapia, Universidade Federal de Pernambuco, Dissertação (Mestrado)*.
- Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. (2015). Rethinking the inception architecture for computer vision. *CoRR*, abs/1512.00567.
- Yang, H.-Y., Yang, J., Pan, Y., Cao, K., Song, Q., Gao, F., and Yin, Y. (2019). Learn to be uncertain: Leveraging uncertain labels in chest x-rays with bayesian neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*.