

# Classificação e Recuperação de Imagens Baseada em Conteúdo Aplicada a Imagens de Tomografia do Tórax

Helio R. V. Couto Júnior<sup>1</sup>, Romuere R. V. e Silva<sup>1,2</sup>

<sup>1</sup> Sistemas de Informação (UFPI) Picos, Brasil.

<sup>2</sup>Programa de Pós-Graduação em Engenharia Elétrica (UFPI) Picos/Teresina, Brasil.

junior0helio15@gmail.com, romuere@ufpi.edu.br

**Abstract.** *Early diagnosis and treatment of pulmonary nodules can significantly improve the patient's survival rate. For a lung cancer examination, patients are submitted to X-ray, CT or MRI scans to differentiate the development of lung abnormalities, enabling the use of computers for autonomous classification. In this work we performed performance tests between methods based on Convolutional Neural Networks previously trained with the ImageNet image base for the extraction of pulmonary nodule attributes. Features were classified using content-based image retrieval models.*

**Resumo.** *O diagnóstico e o tratamento precoce de nódulos pulmonares podem melhorar significativamente a taxa de sobrevivência do paciente. Para um exame do câncer de pulmão, os pacientes são submetidos a exames de raio X, tomografia computadorizada ou ressonância magnética para diferenciar os desenvolvimentos de anomalias no pulmão, possibilitando o uso de computadores para classificação autônoma. Neste trabalho realizamos testes de performance entre métodos baseados em Redes Neurais Convolucionais previamente treinados com a base de imagens ImageNet para a extração de atributos dos nódulos pulmonares. As características foram classificadas através de modelos de recuperação de imagens baseada em conteúdo.*

## 1. Introdução

O câncer de pulmão, de acordo com o Instituto Nacional de Câncer [INCA 2019], é o segundo mais comum no Brasil e o primeiro no mundo desde 1985, tanto na incidência quanto na mortalidade. Na estimativa mundial de 2012 apresentou cerca de 1,8 milhões de novos casos onde 68% homens e 32% mulheres. No Brasil há uma estimativa de 31.270, sendo 60% homens e 40% mulheres [INCA 2019], em 2015 foram 26.498 de mortes pelo câncer de pulmão.

Entre todos os casos de câncer de pulmão, mais da metade é diagnosticada em um estágio avançado, diminuindo a taxa de sobrevivência para 4%. A detecção precoce e a localização de nódulos podem melhorar muito a taxa de sobrevivência para o câncer de pulmão, o que requer a identificação de nódulos de malignos a inocentes como parte essencial para o diagnóstico.

Com o avanço da tecnologia em Tomografia Computadorizada (TC), o espaço entre as camadas digitalizadas está cada vez menor. Com esse crescimento de informações,

os médicos precisam de muito mais tempo para detectar e diagnosticar. A tarefa de encontrar nódulos e classificá-los entre benigno e maligno é árdua ao olho humano, trazendo assim uma necessidade da utilização de computadores.

Dessa forma, o objetivo deste trabalho é avaliar a viabilidade de das principais arquiteturas de redes neurais de convolução (CNNs) na classificação e recuperação de nódulos pulmonares pequenos e não pequenos.

## 2. Trabalhos relacionados

No trabalho desenvolvido por [Yuan et al. 2018] foi utilizado as bases de imagens pública LIDC-IDRI (*Lung Image Database Consortium image collection*) e ELCAP (*Early Lung Cancer Action Program*), eles propuseram um novo método de classificação para nódulos pulmonares utilizando imagens de TC baseadas em características híbridas. Nele utiliza descritores híbridos que consistem em CNNs com múltiplas escalas e características geométricas de codificações de vetores Fisher (FV) baseados em transformações de características invariantes em escala (SIFT). Primeiro eles aproximaram o raio do nódulo baseado na amostragem do icosaedro e análise de intensidade. Em seguida, aplicaram a análise da medida de conteúdo de alta frequência para obter visualizações de amostragem com informações mais abundantes. Depois disso, com base em visualizações re-amostradas, treinam as CNNs para extrair as características e calcular as codificações FV como características geométricas. Por fim conseguiram mesclar características estatísticas e geométricas com base no aprendizado múltiplo do kernel e classificando os tipos de nódulos por meio de uma máquina de vetores de suporte de várias classes.

No trabalho de [Tajbakhsh and Suzuki 2017], foram realizados testes de performances entre Massive Training Artificial Neural Network (MTANNs) e CNNs. Foram considerados dois tópicos bem estudados no campo da análise de imagens médicas: a detecção de nódulos pulmonares e a distinção dos nódulos pulmonares entre benignos e malignos na TC. A análise consiste em duas arquiteturas MTANN otimizadas e quatro arquiteturas CNN distintas que possuem diferentes profundidades. Nestas experiências demonstraram que o desempenho de MTANNs era substancialmente mais alto que o da CNNs ao usar apenas dados limitados de treinamento. Com um conjunto de dados de treinamento maior, a diferença de desempenho tornou-se menos evidente, embora a margem ainda fosse significativa.

[Nishio et al. 2018], Desenvolveram um método de Diagnóstico Auxiliado por Computador (CADx) para a classificação entre nódulo benigno, câncer de pulmão primário e câncer de pulmão metastático. Entre 1240 pacientes de banco de dados previamente construído, imagens de tomografia computadorizada e informações clínicas de 1236 pacientes foram incluídos. Para o método convencional, o CADx foi realizado usando padrão binário local de padrão uniforme com rotação invariante em três planos ortogonais com uma máquina de vetor de suporte. Para o método Deep Convolutional Neural Network (DCNN), o CADx foi avaliado utilizando-se a CNN VGG-16 com e sem aprendizado de transferência, e a otimização hiperparamétrica do método DCNN foi realizada por busca aleatória.

### 3. Referencial Teórico

Para dar um entendimento maior sobre o assunto a ser tratado neste trabalho é necessário conhecer alguns conceitos sobre o câncer de pulmão e métodos de processamento de imagem. Nesta sessão serão abordados os assuntos: *Transfer Learning*, *ImageNet*, *VGG16*, *VGG19*, *Dense 121*, *Inception Resnet v2*, *Xception* e *Inception 3*.

#### 3.1. Transfer Learning

*Transfer Learning* é uma técnica de aprendizado de máquina na qual um modelo é treinado e desenvolvido para uma tarefa e é então utilizado em uma segunda tarefa relacionada [Hussain et al. 2018]. Neste trabalho analisaremos vários modelos que foram treinados com a base de imagens e serão reaproveitados para aprender a categorizar os nódulos pulmonares entre benigno e maligno.

#### 3.2. VGG16 e VGG19

Introduzida por [LeCun et al. 1995], uma CNN é um tipo especial de redes neurais multicamada, projetadas para reconhecer padrões visuais diretamente de imagens de pixel com pré-processamento mínimo. Esta rede é caracterizada pela sua simplicidade, usando apenas camadas convolucionais de tamanho  $3 \times 3$  empilhadas umas sobre as outras em profundidade crescente [Simonyan and Zisserman 2014]. A redução do tamanho do volume é tratada pelo pool máximo. Existem ainda duas camadas totalmente conectadas, cada uma com 4096 nós. Os “16” e “19” representam o número de camadas de peso na rede.

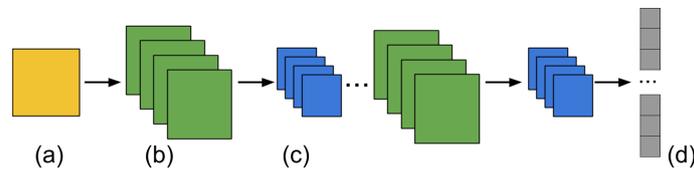


Figura 1. Arquitetura da VGG.

Na Figura 1 é mostrada a arquitetura da VGG. (a) imagem de entrada, (b) Camada de convolução, (c) Camada de pool, (d) Saída da CNN - Atributos da imagem.

#### 3.3. Dense 121

No trabalho desenvolvido por [Huang et al. 2016] é proposta a arquitetura Dense 121, esta CNN tem como principal característica conectar as camadas com todas as suas camadas a frente. A DenseNet tem várias vantagens convincentes: eles aliviam o problema do gradiente de desaparecimento, reforçam a propagação de recursos, estimulam a reutilização de recursos e reduzem substancialmente o número de parâmetros.

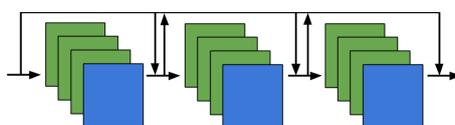


Figura 2. Parte da arquitetura da DenseNet.

### 3.4. Inception 3

A microarquitetura Inception foi introduzida por [Szegedy et al. 2015]. O objetivo desse modelo de criação é atuar como um "extrator de vários níveis", computando convoluções de  $1 \times 1$ ,  $3 \times 3$  e  $5 \times 5$  dentro do mesmo módulo da rede - a saída desses filtros é então empilhada junto a dimensão do canal e antes de ser alimentado na próxima camada da rede.

### 3.5. Inception Resnet v2

Na Inception Resnet desenvolvida por [Szegedy et al. 2016] foi usada blocos Inception menores que o Inception original. Cada bloco de iniciação é seguido por uma camada de expansão de filtro que é usada para escalar a dimensionalidade do banco de filtros antes da adição para coincidir com a profundidade da entrada. Isso é necessário para compensar a redução de dimensionalidade induzida pelo bloco Inception.

### 3.6. Xception

O Xception, desenvolvida por [Chollet 2016], inspirada no Inception, os módulos de Inception foram substituídos por convoluções separáveis em profundidade. A Xception possui o mesmo número de parâmetros que o Inception V3 pois tem um uso mais eficiente dos parâmetros do modelo.

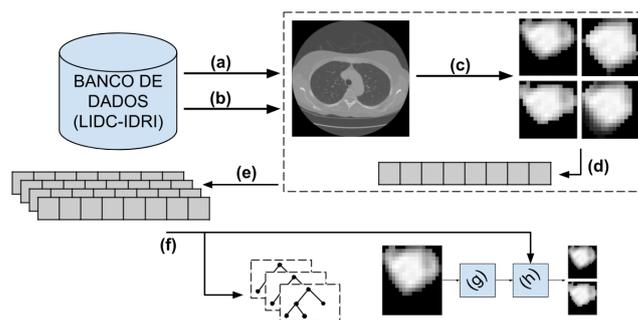
### 3.7. Base de dados

O conjunto de dados que usamos foi o LIDC-IDRI, dela utilizamos 516 casos, cada caso consiste em imagens de Tomografia Computadorizada (TC) do tórax onde cada TC tem anotações fornecidas por quatro radiologistas. A região de interesse foi extraída de acordo com as informações feitas pelos especialistas. Em seguida, colocamos cada faixa do nódulo em uma imagem de  $224 \times 224$  (outros com resolução  $299 \times 299$  dependendo da CNN). Ao final, nosso banco de dados consiste em 23263 imagens.

## 4. Metodologia

Para a realização deste trabalho seguimos o fluxograma apresentado na Figura 3. Estes passos serão realizados para todas as CNNs. A partir da base de dados LIDC-IDRI, obtemos todas as regiões de nódulos utilizando as imagens de TC e a marcação dos especialistas. Após isso, extraímos o atributos das imagens de cada região de nódulo utilizando seis diferentes arquiteturas de CNNs. O passo seguinte é treinar uma *Random Forest* [Breiman 2001a] para separar os nódulos em benignos e malignos. A performance do método é obtida em termos de Acurácia, Especificidade e Sensibilidade para o sistema de classificação, e através do MAP (*Mean Average Precision*) nos experimentos CBIR [Gudivada and Raghavan 1995].

A Figura 3 apresenta o fluxograma da metodologia utilizada neste trabalho nos experimentos de classificação. (a) Imagens TC do banco de dados LIDC-IDRI. (b) Informações dos especialistas sobre o diagnóstico e marcação de cada nódulo. (c) Combinar a informações de cada nódulo e extraí-los das imagens TC. (d) Extrair os atributos de cada nódulos utilizando uma CNN. (e) Agrupar as características para o treinamento/teste do classificador. (f) Treinar e testar o classificador Random Forest (RF) para diferenciar os nódulos em malignos e benignos. Além disso, foram feitos experimentos CBIR substituindo a etapa (f) pela recuperação de imagens. (g) Extração dos atributos. (h) Correspondência de similaridade.



**Figura 3. Fluxograma da metodologia.**

## 5. Resultados e discussão

Neste trabalho foram utilizadas faixas CT dos nódulos na base LIDC-IDRI com um total de 23.263 imagens. Para obtenção dos resultados, foram utilizadas seis arquiteturas de CNNs: VGG16, VGG19, Dense 121, Inception Resnet v2, Xception e Inception 3.

Para o treinamento de máquina foi usado o classificador RF [Breiman 2001b]. Utilizamos 100 estimadores da RF. Foram utilizadas 80% das imagens para treinamento e as outras 20% para os testes. Nos experimentos CBIR utilizamos a distância euclidiana como métrica de similaridade. Na base LIDC-IDRI o grau de malignidade dos nódulos está na faixa de 1 a 5, consideramos que entre 1 a 3 seja benigna e entre 4 e 5 maligna.

**Tabela 1. Análise das performances de classificação.**

	VGG16	VGG19	Dense 121	Inception Resnet v2	Xception	Inception 3
Sensibilidade	0,87090	0,86581	0,84536	0,84259	0,84294	0,84479
Especificidade	0,63739	0,62400	0,63109	0,64463	0,60469	0,62293
Acurácia	0,78709	0,78022	0,76564	0,76871	0,75390	0,76494

A Tabela 1 fornece a análise comparativa das extrações pelos modelos CNNs, sendo três métricas de validação: Acurácia, Especificidade e Sensibilidade. Observa-se que os resultados obtidos foram similares nas diferentes arquiteturas, não tendo nenhuma se sobressaindo das demais.

**Tabela 2. Análise das performances de CBIR.**

	VGG16	VGG19	Dense 121	Inception Resnet v2	Xception	Inception 3
MAP Benigno	0,7302	0,7322	0,6774	0,6803	0,6881	0,6787
MAP Maligno	0,4568	0,4380	0,4221	0,3738	0,4136	0,4550

A Tabela 2 mostra os experimentos CBIR realizados. Para melhorar analisar os resultados dividimos as métricas por classe, ou seja, obtivemos um MAP para a classe de nódulos malignos e outro para a classe de benignos. Observamos que a arquitetura Inception 3 obteve uma ligeira queda de desempenho quando comparada às demais. Observa-se também uma performance menor na recuperação dos nódulos malignos, provavelmente pela falta de um padrão nesses tipos de nódulos.

## 6. Conclusão

Nós fizemos um teste de performances de CNNs, pré-treinados com *ImageNet*, para extração de características de nódulos pulmonar. O objetivo da análise apresentada é mostrar a precisão da extração dos nódulos divididas em faixas de imagem de TC. Em nosso trabalho, analisamos seis extratores, todos eles apresentaram uma performances semelhantes, uma das possíveis causas desta precisão é que os nódulos pulmonares 3D foram divididas em faixas 2D, com isso algumas imagens podem ter uma pequena parte do nódulo completo, outra possibilidade é o tamanho dos nódulos pois as CNNs estudadas variam o tamanho das imagens de entrada entre  $224 \times 224$  e  $299 \times 299$ .

## Referências

- Breiman, L. (2001a). Random forests. *Machine learning*, 45(1):5–32.
- Breiman, L. (2001b). Random forests. *Mach. Learn.*, 45(1):5–32.
- Chollet, F. (2016). Xception: Deep learning with depthwise separable convolutions. *CoRR*, abs/1610.02357.
- Gudivada, V. N. and Raghavan, V. V. (1995). Content based image retrieval systems. *Computer*, 28(9):18–22.
- Huang, G., Liu, Z., and Weinberger, K. Q. (2016). Densely connected convolutional networks. *CoRR*, abs/1608.06993.
- Hussain, M., Bird, J. J., and Faria, D. R. (2018). A study on cnn transfer learning for image classification. In *UK Workshop on Computational Intelligence*, pages 191–202. Springer.
- INCA (2019). Câncer de pulmão.
- LeCun, Y., Bengio, Y., et al. (1995). Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks*, 3361(10):1995.
- Nishio, M., Sugiyama, O., Yakami, M., Ueno, S., Kubo, T., Kuroda, T., and Togashi, K. (2018). Computer-aided diagnosis of lung nodule classification between benign nodule, primary lung cancer, and metastatic lung cancer at different image size using deep convolutional neural network with transfer learning. *PloS one*, 13(7):e0200721.
- Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Szegedy, C., Ioffe, S., and Vanhoucke, V. (2016). Inception-v4, inception-resnet and the impact of residual connections on learning. *CoRR*, abs/1602.07261.
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. (2015). Rethinking the inception architecture for computer vision. *CoRR*, abs/1512.00567.
- Tajbakhsh, N. and Suzuki, K. (2017). Comparing two classes of end-to-end machine-learning models in lung nodule detection and classification: Mtanns vs. cnns. *Pattern recognition*, 63:476–486.
- Yuan, J., Liu, X., Hou, F., Qin, H., and Hao, A. (2018). Hybrid-feature-guided lung nodule type classification on ct images. *Computers & Graphics*, 70:288–299.