

INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E  
TECNOLOGIA DO CEARÁ  
CÂMPUS MARACANAÚ

ALAN FÉLIX DA SILVA

MÉTODO AUTOMÁTICO DE  
CLASSIFICAÇÃO DE NÓDULOS  
PULMONARES EM IMAGENS DE TC DO  
TÓRAX

FORTALEZA

2020

ALAN FÉLIX DA SILVA

# MÉTODO AUTOMÁTICO DE CLASSIFICAÇÃO DE NÓDULOS PULMONARES EM IMAGENS DE TC DO TÓRAX

Trabalho de Conclusão de Curso  
apresentado ao Instituto Federal de  
Educação, Ciência e Tecnologia de  
Ceará, como parte dos requisitos para  
a obtenção do grau de Bacharelado.

Área de Concentração: Visão  
Computacional

Orientador: Prof. Dr. Igor Rafael Silva  
Valente

FORTALEZA

2020

# 1 INTRODUÇÃO

O câncer pulmonar se tornou um dos maiores problemas de saúde mundial, situando-se entre as principais causas de mortes em vários países. A mortalidade por câncer está sempre crescendo e os motivos são complexos. O crescimento populacional, envelhecimento e desenvolvimento socioeconômico são fatores que contribuem para este aumento (BRAY et al., 2018).

Nos últimos anos houve um grande aumento no desenvolvimento de sistemas de detecção e diagnóstico, conhecidos como CADs, sendo cada vez mais utilizados para o auxílio ao diagnóstico médico. Estes tipos de sistemas realizam marcações em imagens, mostrando possíveis anomalias, servindo como um suporte aos médicos na decisão de diagnósticos (PETRICK et al., 2013).

O sistema CAD podem ser divididos em dois subsistemas: CADe, subsistema capaz de analisar as imagens e detectar regiões suspeitas, revelando as anomalias e facilitando então a observação dos médicos; CADx, subsistema responsável por classificar as regiões previamente identificadas de acordo com seu tipo, gravidade e estágio (VALENTE et al., 2016).

Xie et al. (2018) testaram uma nova técnica para a classificação de nódulos pulmonares, utilizando a união da textura, forma e informações profundas do modelo aprendido (Fuse-TSD), o algoritmo é realizado de duas formas, sendo um treinamento online e offline, na primeira fase offline são extraídas as regiões de interesse de cada tomografias computadorizada (TC), na sequência é realizado o treinamento através da rede Deep convolutional neural network (DCNN) para extrair a textura, formas e usar os recursos extraídos para realizar o cálculo dos pesos, então usando Back-Propagation Neural Networks (BPNN) para classificar. Na fase online, ocorre a extração de nódulos em fatia retirada da TC, aplica-se ao DCNN testado anteriormente para obtendo novas características profundas, textura e forma, na sequência são aplicadas ao classificador com suas respectivas características ao conjunto já treinado, então classificar o nódulo com base na etiqueta de cada fatia de nódulo. **A técnica foi avaliada utilizando a base de dados The Lung Image Database Consortium (LIDC-IDRI) contendo 1018 exames e foi usado o esquema proposto por Han et al. (2015) e Shen et al. (2015) onde a base de dados é subdividida de acordo com o grau de malignidade, um e dois é considerado nódulos benignos, quatro e cinco é considerado maligno e grau três não há certeza sobre benigno ou maligno. A subdivisão da base foi realizada em três grupos, grupo um contendo nódulos benignos indicados pelo grau um e dois e malignos quatro e cinco; O segundo grupo contendo como**

nódulos benignos, nodos de grau um, dois e três, e malignos de grau quatro e cinco; O terceiro grupo contendo como nodos benignos nodos de grau um e dois, e malignos de grau três, quatro e cinco. O resultado obtido para os grupos foram respectivamente AUC 96.65%, 94.45% e 81.24%.

Rodrigues et al. (2018) Apresentou uma nova abordagem baseada na matriz estrutural co-ocorrência (SCM) para realizar a classificação de nodos pulmonares em benigno e maligno e realizar uma classificação ao nível de malignidade. Inicialmente é realizado uma extração dos nodos pulmonares. A etapa seguinte são aplicadas diferentes configurações ao extrator SCM, sendo filtro Gaussiano, Laplaciano, Media e Sobel, identificando a melhor performance. Na etapa seguinte de classificação os dados que foram gerados são pré-processador e passados aos classificadores perceptron multicamadas (MLP), o vetor de suporte máquina (SVM) e os algoritmos de k-vizinhos mais próximos (k-NN), realizando então o treinamento da rede neural e realizando o teste nos dados. O desempenho foi testado utilizando a base de dados LIDC-IDRI e atingiu 96.7% para as métricas de precisão e F-Score na primeira tarefa, e 74: 5% de precisão e 53.2% F-Score na segunda.

Apesar desse método ter obtido grandes resultados, uma parte das TC, não foi utilizada para realização da classificação, sendo os nodos classificados como  $< 3mm$ . Com isso tem-se a necessidade de classificação de todos os nodos.

Neste contexto este trabalho tem o objetivo de construir um método automático para realizar a classificação dos nodos pulmonares em tomografias computadorizadas TC do tórax, buscando extrair mais informações da base de dados LIDC-IDRI que possa ser utilizado para a melhoria dos classificadores já existente junto com a aplicação de redes neurais convolucionais.

O método proposto será avaliado na base The Lung Image Database Consortium (LIDC-IDRI), tendo um total de 1018 exames de TC do tórax para detecção e diagnóstico de câncer do pulmão com lesões. A elaboração dessa base de dados foi realizada por uma equipe de quatro radiologistas experientes e o processo foi dividido em duas fases, a primeira "vista às cegas", cada radiologista individualmente realizou uma revisão em cada exame identificando as lesões e classificando-os em três categorias: nódulo  $\geq 3$  mm, nódulo  $< 3$  mm e não nódulo  $\geq 3$  mm; Na segunda fase, chamada de "vista não cega" são apresentados os resultados da primeira fase para cada radiologista permitindo então uma nova revisão em nodos classificados como  $\geq 3$  mm e nódulo  $< 3$  mm, sendo os classificados como não nodos opcional para revisão, não foi formado um consenso final, ou seja, existem nodos na base que foram apontados por somente um radiologista (3RD et al., 2011).

# REFERÊNCIAS

3RD, A. S. et al. The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A Completed Reference Database of Lung Nodules on CT Scans. 2011. Disponível em: <<https://wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI>>. Citado na página 4.

BRAY, F. et al. Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. **CA: A Cancer Journal for Clinicians**, 2018. Disponível em: <<https://acsjournals.onlinelibrary.wiley.com/doi/abs/10.3322/caac.21492>>. Citado na página 3.

HAN, F. et al. Texture feature analysis for computer-aided diagnosis on pulmonary nodules. **Journal of Digital Imaging**, 2015. Citado na página 3.

PETRICK, N. et al. Evaluation of computer-aided detection and diagnosis systemsa). **Medical Physics**, 2013. Disponível em: <<https://aapm.onlinelibrary.wiley.com/doi/abs/10.1118/1.4816310>>. Citado na página 3.

Rodrigues, M. B. et al. Health of things algorithms for malignancy level classification of lung nodules. **IEEE Access**, 2018. Citado na página 4.

SHEN, W. et al. Multi-scale convolutional neural networks for lung nodule classification. **Information Processing in Medical Imaging. IPMI 2015.**, 2015. Citado na página 3.

VALENTE, I. R. S. et al. Automatic 3d pulmonary nodule detection in ct images: A survey. **Computer Methods and Programs in Biomedicine**, 2016. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0169260715300298>>. Citado na página 3.

XIE, Y. et al. Fusing texture, shape and deep model-learned information at decision level for automated classification of lung nodules on chest ct. **Information Fusion**, 2018. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1566253516301063>>. Citado na página 3.