

Antonio Oseas de Carvalho Filho

# **Diagnóstico precoce de câncer pulmonar usando 3D Deep Features**

Relatório de Execução do Objeto - PRO-  
PESQI/PRPG/UFPI Nº 07/2020 – Bolsas  
de Produtividade em Pesquisa

Universidade Federal do Piauí - UFPI

Laboratório de Pesquisas Avançadas em Visão e Inteligência Computacional -  
PAVIC/UFPI

Programa de Pós-Graduação em Engenharia Elétrica - PPGEE/UFPI

Departamento de Sistemas de Informação/Engenharia Elétrica

Picos, PI - Brasil

2021

# Sumário

<b>1</b>	<b>IDENTIFICAÇÃO</b> .....	<b>3</b>
<b>1.1</b>	<b>Resumo</b> .....	<b>3</b>
<b>1.2</b>	<b>Abstract</b> .....	<b>4</b>
<b>2</b>	<b>IMPACTOS DO PROJETO PARA AVANÇO DO ESTADO DA ARTE NA ÁREA DO CONHECIMENTO</b> .....	<b>5</b>
<b>3</b>	<b>CONTRIBUIÇÃO DO PROJETO PARA INOVAÇÃO DE PRODUTOS, PROCESSOS OU POLÍTICAS PÚBLICAS</b> .....	<b>5</b>
<b>4</b>	<b>CONTRIBUIÇÃO DO PROJETO PARA FORMAÇÃO DE RECURSOS HUMANOS ESPECIALIZADOS PARA A ACADEMIA, EDUCAÇÃO BÁSICA E SUPERIOR, INDÚSTRIA, SETOR DE SERVIÇOS E SETOR PÚBLICO</b> .....	<b>6</b>
<b>5</b>	<b>CONTRIBUIÇÃO DO PROJETO PARA DIFUSÃO E TRANSFERÊNCIA DE CONHECIMENTO</b> .....	<b>6</b>

# 1 Identificação da Proposta

- Título

Português: Diagnóstico precoce de câncer pulmonar usando 3D Deep Features

Inglês: Early diagnosis of lung cancer using 3D Deep Features

- Identificação Proponente

Antonio Oseas de Carvalho Filho - antoniooseas@ufpi.edu.br - (030.355.953-56)

Prof. Dr., Adjunto III do Departamento de Sistemas de Informação - SI/UFPI, docente do Programa de Pós-Graduação em Engenharia Elétrica - PPGEE/UFPI, Assessor de Ações de Pesquisa do Campus Senador Helvídio Nunes de Barros - UFPI, e pesquisador do laboratório de Pesquisas Avançadas em Visão e Inteligência Computacional - PAVIC/UFPI

## 1.1 Resumo

O câncer de pulmão é apontado como a principal causa de morte entre os pacientes com câncer. As altas taxas de mortes e registros de ocorrências desse câncer em todo o mundo demonstram a importância do desenvolvimento e investigação, a fim de produzir meios para a detecção e o diagnóstico precoce dessa doença. O tempo dispendido para trabalhar com exames por imagem, a subjetividade dos atributos extraídos e a necessidade contínua de investigação para o progresso na área, tem feito surgir novas técnicas de processamento e análise das imagens médicas que melhoram a qualidade do diagnóstico. Dessa forma, com intuito de aumentar a precisão no diagnóstico de lesões, auxiliando o especialista com uma segunda opinião e em processos de triagem, ferramentas de cunho computacional denominadas Computer-Aided Diagnosis (CADx), tem sido amplamente exploradas. Além disso, o cenário conturbado em que vivenciamos, isto é, a pandemia causada pelo SARS-CoV-2 (COVID-19), mostra o quão importante é, a investigação de doenças que acometem o aparelho respiratório, em especial o pulmão. Contribuir para o diagnóstico precoce e preciso do câncer de pulmão é o objetivo principal deste projeto. Normalmente, os processos envolvidos nos sistemas CADx são compostos por quatro etapas: 1) a aquisição de imagem; 2) extração/segmentação dos nódulos para análise; 3) extração de características; e 4) classificação. Nesse contexto, este projeto pretende desenvolver métodos capazes de mensurar as diferenças entre as classes de lesões pulmonares e classificar entre: i) tecido saudável; ii) lesão benigna; iii) lesão maligna; e iv) tipos de pneumonias.

Para isso, serão adaptados e propostas novas técnicas através das análises de forma e textura em conjunto com técnicas baseadas em deep learning.

## 1.2 Abstract

Lung cancer is considered the leading cause of death among cancer patients. The high rates of deaths and records of occurrences of this cancer worldwide demonstrate the importance of development and investigation to produce means for the detection and early diagnosis of this disease. The time spent working with imaging exams, the subjectivity of the extracted attributes, and the continuous need for research for progress in the area, have given rise to new techniques for processing and analyzing medical images that improve the quality of diagnosis. Thus, to increase the accuracy in the diagnosis of injuries, helping the specialist with a second opinion, and in screening processes, computational tools called Computer-Aided Diagnosis (CADx) have been widely explored. In addition, the troubled scenario in which we live, that is, the pandemic caused by SARS-CoV-2 (COVID-19), shows how important it is to investigate diseases that affect the respiratory system, especially the lung. Contributing to the early and accurate diagnosis of lung cancer is the main objective of this project. Typically, the processes involved in CADx systems are composed of four steps: 1) image acquisition; 2) extraction/segmentation of nodules for analysis; 3) feature extraction, and 4) classification. In this context, this project intends to develop methods capable of measuring the differences between the classes of lung lesions and classifying between i) healthy tissue; ii) benign lesion; iii) malignant lesion, and iv) types of pneumonia. For this, new techniques will be adapted and proposed by analyzing shape and texture in conjunction with techniques based on deep learning.

## 2 Impactos do projeto para avanço do estado da arte na área do conhecimento

As novas funcionalidades promovidas pelo desenvolvimento das metodologias deste projeto, proporcionaram mecanismos importantes na classificação de lesões pulmonares e de outras doenças. Nestes meses de execução do projeto, foi possível desenvolver métodos eficientes que concentram contribuições importantes para o estado da arte das áreas de Ciência da Computação com impacto no âmbito das áreas médicas, especificamente, no auxílio ao diagnóstico médico, tão quanto, para triagem de pacientes. Especificamente, dentre outros, pode-se listar os impactos do projeto:

- Construção de técnicas capazes de descrever a textura das lesões pulmonares;
- Adaptação de técnicas de outras áreas do conhecimento capazes de mensurar sobre propriedades de forma e de textura das imagens analisadas;
- Fomentar a construção de modelos híbridos mesclando o estado da arte com as técnicas desenvolvidas no projeto;
- Aplicação e adaptação de técnicas tradicionais, especificamente, baseadas em *deep learning* para caracterização das lesões estudadas; e,
- Possibilidade de expansão das técnicas propostas para análises de outros tipos de imagens.

## 3 Contribuição do projeto para inovação de produtos, processos ou políticas públicas

Por trabalhar com o desenvolvimento de novas técnicas capazes de mensurar sobre propriedades de texturas, sejam elas, em imagens de duas dimensões ou de três dimensões, abre-se um leque de possibilidades de aplicabilidade em diversos problemas no contexto de análises de imagens. Baseado nisso, acredita-se que a possibilidade de inovação seja algo concreto na medida em que os protocolos sejam atendidos, uma vez que, até o presente momento de escrita deste relatório, o presente preponente alcançou concessão do registro de software intitulado **GLCAD**, sob o número de registro: BR512021001795-1. Este produto de software contempla um dos métodos desenvolvidos ao longo deste projeto e adaptados

ao contexto de análises de imagens da retina. Além deste, é importante destacar que outros métodos frutos deste projeto estão em processo de desenvolvimento para novas solicitações de pedidos de registro.

## 4 Contribuição do projeto para formação de recursos humanos especializados para a academia, educação básica e superior, indústria, setor de serviços e setor público

Como resposta direta à sociedade, por meio deste projeto contemplou-se a formação técnico/científica dos pesquisadores e alunos que trabalharam no desenvolvimento do tema proposto, bem como a possibilidade de disparar outros processos que, agregados à linha de pesquisa delineada, estão propiciando o aperfeiçoamento de outros alunos, a saber:

- Desenvolvimento de capital intelectual de qualidade direcionado para o processamento de imagens médicas;
- Desenvolvimento de capital intelectual de qualidade no contexto da inteligência computacional;
- Construção de trabalhos de Iniciação Científica, Iniciação Tecnológica, Graduação e Mestrado;
- Por fim, é importante constatar que, a partir da vivência nos projetos, muitos alunos participantes do projeto, conseguiram oportunidades no mercado de trabalho, graças as expertises adquiridas ao longo das ações do projeto.

## 5 Contribuição do projeto para difusão e transferência de conhecimento

Neste aspecto, destacam-se as principais publicações em periódicos e congressos internacionais alcançados durante o desenvolvimento do presente projeto, a saber:

1. SALES, N. R.; DE CARVALHO FILHO, ANTONIO OSEAS . Automatic Method for Glaucoma Classification Using Texture Analysis, Xgboost and Grid search. REVISTA DE SISTEMAS E COMPUTAÇÃO - RSC, 2020;
2. BISNETO, T.; DE CARVALHO FILHO, ANTONIO OSEAS; MAGALHAES, D. M. V. . Generative adversarial network and texture features applied to automatic glaucoma detection. APPLIED SOFT COMPUTING, 2020;
3. TORRES, WILLIAM; DE CARVALHO FILHO, ANTONIO OSEAS; LIRA, R. A.; SILVA, R. R. V. E. . Texture analysis of lung nodules in computerized tomography images using functional diversity. COMPUTERS and ELECTRICAL and ENGINEERING, 2020;
4. CARVALHO, E. D.; DE CARVALHO FILHO, ANTONIO OSEAS; DINIZ, J. O. B.; SILVA, R. R. V. E.; ARAUJO, F. H.; SILVA, Aristófanés Corrêa; Cardoso de Paiva, Anselmo; GATTASS, MARCELO . Breast cancer diagnosis from histopathological images using textural features and CBIR. ARTIFICIAL INTELLIGENCE IN MEDICINE, 2020;
5. LIMA, THIAGO JOSE BARBOSA; USHIZIMA, DANIELA; DE CARVALHO FILHO, ANTONIO OSEAS; DE ARAUJO, FLAVIO HENRIQUE DUARTE . Lung CT Screening With 3D Convolutional Neural Network Architectures. In: 2020 IEEE 17th International Symposium on Biomedical Imaging Workshops (ISBI Workshops), 2020;
6. SALES DOS SANTOS, PATRICK RYAN; DE CARVALHO BRITO, VITORIA; DE CARVALHO FILHO, ANTONIO OSEAS; HENRIQUE DUARTE DE ARAUJO, FLAVIO; DE ANDRADE LIRA RABELO, RICARDO; JOSEPH MATHEW, MANO . A Capsule Network-based for identification of Glaucoma in retinal images. In: 2020 IEEE Symposium on Computers and Communications (ISCC), 2020;
7. BARBOSA LIMA, THIAGO JOSE; DUARTE DE ARAUJO, FLAVIO HENRIQUE; DE CARVALHO FILHO, ANTONIO OSEAS; DE ANDRADE LIRA RABELO, RICARDO; DE MELO SOUZA VERAS, RODRIGO; JOSEPH MATHEW, MANO . Evaluation of data balancing techniques in 3D CNNs for the classification of pulmonary nodules in CT images. In: 2020 IEEE Symposium on Computers and Communications (ISCC), 2020;
8. CARVALHO, EDELSON DAMASCENO; CARVALHO, EDSON DAMASCENO; DE CARVALHO FILHO, ANTONIO OSEAS; DE ARAUJO, FLAVIO HENRIQUE DUARTE; ANDRADE LIRA RABELO, RICARDO DE . Diagnosis of COVID-19 in CT image using CNN and XGBoost. In: 2020 IEEE Symposium on Computers and Communications (ISCC), 2020, Rennes;

9. FILHO, JOSE LIMA; DA SILVA BORGES, FABBIO ANDERSON; RABELO, RICARDO DE ANDRADE LIRA; SILVA, IVAN SARAIVA; JUNIOR, RICARDO PINTO TEIXEIRA; DE CARVALHO FILHO, ANTONIO OSEAS . Methods for voltage sag source location by Cluster Algorithm and Decision Rule Labeling with a Comparative Approach of K-means and DBSCAN Clustering Algorithms. In: 2020 5th International Conference on Smart and Sustainable Technologies (SpliTech), 2020;
10. LEAL RODRIGUES JUNIOR, WILSON; MEDEIROS REIS, DYOGO; ANDERSON SILVA BORGES, FABBIO; HENRIQUE DUARTE ARAUJO, FLAVIO; OSEAS DE CARVALHO FILHO, ANTONIO; DE ANDRADE LIRA RABELO, RICARDO . Localization of Voltage Sag Sources Using Convolutional Neural Network in IEEE 34-bus System. In: 2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2020; e,
11. CARVALHO, EDELSON DAMASCENO; CARVALHO, EDSON DAMASCENO; DE CARVALHO FILHO, ANTONIO OSEAS; DE SOUSA, ALCILENE DALILIA; DE ANDRADE LIRA RABULO, RICARDO . COVID-19 diagnosis in CT images using CNN to extract features and multiple classifiers. In: 2020 IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE), 2020.