

**UNIVERSIDADE FEDERAL DE ALAGOAS
CAMPUS ARAPIRACA
CIÊNCIA DA COMPUTAÇÃO – BACHARELADO**

JOSÉ YTALO RAMON DE ALMEIDA GONÇALVES

**TÉCNICAS MULTI-ESCALA PARA IDENTIFICAÇÃO DE NÓDULOS
PULMONARES BASEADAS EM REDES CONVOLUCIONAIS**

ARAPIRACA

2023

José Ytalo Ramon de Almeida Gonçalves

Técnicas multi-escala para identificação de nódulos pulmonares baseadas em redes
convolucionais

Monografia apresentada ao Curso de Ciência da
Computação da Universidade Federal de Alagoas –
Campus Arapiraca, como requisito parcial à
obtenção do título de Bacharelado em Ciência da
Computação.

Orientador: Prof. Dr. Rodolfo Carneiro Cavalcante.

Arapiraca

2023



Universidade Federal de Alagoas – UFAL
Campus Arapiraca
Biblioteca Setorial *Campus Arapiraca* - BSCA

G635t Gonçalves, José Ytalo Ramon de Almeida
 Técnicas multi-escala para identificação de nódulos pulmonares baseadas em
 redes convolucionais [recurso eletrônico] / José Ytalo Ramon de Almeida Gonçalves. –
 Arapiraca, 2023.
 23 f.: il.

 Orientador: Prof. Dr. Rodolfo Carneiro Cavalcante.
 Trabalho de Conclusão de Curso (Bacharelado de Ciência da Computação) -
 Universidade Federal de Alagoas, *Campus Arapiraca*, Arapiraca, 2023.
 Disponível em: Universidade Digital (UD) – UFAL (*Campus Arapiraca*).
 Referências: f. 23.

 1. Aprendizado profundo. 2. Pulmões - Câncer. 3. Multi-escala (Imagem).
 4. Filtragem espacial. I. Cavalcante, Rodolfo Carneiro. II. Título.

CDU 004

Bibliotecário responsável: Nestor Antonio Alves Junior
CRB - 4 / 1557

José Ytalo Ramon de Almeida Gonçalves

Técnicas multi-escala para identificação de nódulos pulmonares baseadas em redes
convolucionais

Monografia apresentada ao Curso de Ciência da
Computação da Universidade Federal de Alagoas –
Campus Arapiraca, como requisito parcial à
obtenção do título de Bacharelado em Ciência da
Computação.

Data de Aprovação: 12/05/2023.

Banca Examinadora

Documento assinado digitalmente
 RODOLFO CARNEIRO CAVALCANTE
Data: 17/05/2023 15:20:31-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Rodolfo Carneiro Cavalcante
Universidade Federal de Alagoas – UFAL
Campus Arapiraca
(Orientador)

Documento assinado digitalmente
 ELTHON ALLEX DA SILVA OLIVEIRA
Data: 16/05/2023 09:11:54-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Elthon Allex da Silva Oliveira
Universidade Federal de Alagoas – UFAL
Campus Arapiraca
(Examinador)

Documento assinado digitalmente
 TACITO TRINDADE DE ARAUJO TIBURTINO
Data: 16/05/2023 15:48:55-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Tácito Trindade de Araújo Tiburtino Neves
Universidade Federal de Alagoas – UFAL
Campus Arapiraca
(Examinador)

AGRADECIMENTOS

Dedico este trabalho a todas aquelas pessoas que contribuíram e me acompanharam durante essa jornada, em especial, para aqueles que sempre estiveram ao meu lado incentivando e ajudando da forma que podiam: minha mãe, Maria José de Almeida, meu pai, José Maria Gonçalves Filho, meus irmãos, Isadora, Isaque, Maria Aparecida (Cidinha) e Odiley, e meu tio Claudivanio.

Além deles, gostaria de agradecer também ao Wendell Rafael, Valmir Júnior, Daniel Ferreira e a Noemy Roberta por serem grandíssimos amigos.

Ao Alan Duda, a Crislaine Costa e o Ivillys Gomes por terem sido os meus principais parceiros nos incontáveis trabalhos acadêmicos ocorridos durante essa caminhada. E ao Eduardo Vieira, pelos dois ciclos de pesquisa que estivemos juntos.

Agradeço a todos os professores por terem – nas suas áreas de especialidade (às vezes não) – e do seu jeito – partilhado um pouco de seu conhecimento. E ao Tácito Tiburtino e ao Thiago Sales por além de serem excelentes professores terem sido meus orientadores de monitoria.

Por último, mas não menos importante, não mesmo, agradeço ao Rodolfo Carneiro Cavalcante por ter sido um excelente professor, orientador de monitoria, pesquisa e de TCC. Prof. Rodolfo, “OBRIGADO” pelos direcionamentos e, principalmente, pelos “puxões de orelha”.

RESUMO

Os tecidos e órgãos do corpo humano são compostos por células que normalmente crescem, se dividem e morrem o tempo todo de forma ordenada e controlada. Entretanto, devido a diversos fatores, nem sempre o corpo humano consegue manter esta dinâmica, ocasionando tumores que em seu pior estado podem evoluir para um câncer. Segundo a Organização Mundial da Saúde (OMS), o câncer de pulmão é o tipo de câncer que mais causa mortes no mundo. Somente em 2020 foram mais de 2,2 milhões de casos que resultaram na morte de cerca de 1,8 milhão de pessoas. Devido à alta taxa de letalidade advinda deste tipo de câncer, é muito importante que ele seja diagnosticado o mais rápido possível. Entretanto, a detecção precoce e a sobrecarga de trabalho nos especialistas da área – devido ao grande volume de exames – são alguns dos fatores que dificultam e contribuem para este cenário catastrófico. Diante desta realidade e a relevância do tema para sociedade, nos últimos anos vem se intensificando o estudo de técnicas baseadas em Visão Computacional e Aprendizado de Máquina que possam de alguma forma auxiliar especialistas na tarefa de identificação de nódulos em imagens de Tomografia Computadorizada (TC). Sendo assim, este estudo propõe-se a investigar uma abordagem baseada em Redes Neurais Convolucionais com múltiplas entradas, denominada de Multi-Escala, em busca de obter melhores e mais precisas predições das classes dos objetos presentes nas imagens. Além disso, o estudo também se propõe a investigar a aplicação de métodos de pré-processamento sobre as imagens de TC com base em Filtragem Espacial, visando aprimorá-las e realçar características. Os resultados obtidos a partir dos experimentos realizados neste estudo são promissores e demonstram, a princípio, que o uso de Redes Neurais Convolucionais Multi-Escala podem ser de fato uma solução para predição de imagens médicas que apresentam uma grande variabilidade de tamanho dos objetos, ao alcançar a marca de 86% de acurácia. Ademais, o emprego de técnicas de Processamento Digital de Imagens se mostraram um grande aliado no realce das informações presentes nas imagens de TC, conseguindo, em alguns casos, aumentar a acurácia da rede em até 4%.

Palavras-chave: aprendizado profundo; detecção de nódulo; multi-escala; filtragem espacial.

ABSTRACT

The tissues and organs of the human body are composed of cells that normally grow, divide and die all the time in an orderly and controlled manner. However, due to several factors, the human body is not always able to maintain this dynamic, causing tumors that in their worst state can evolve into cancer. According to the World Health Organization (WHO), lung cancer is the type of cancer that causes the most deaths in the world. In 2020 alone, there were more than 2.2 million cases that resulted in the death of about 1.8 million people. Due to the high lethality rate resulting from this type of cancer, it is very important that it be diagnosed as soon as possible. However, early detection and the workload of specialists in the field – due to the large volume of exams – are some of the factors that hinder and contribute to this catastrophic scenario. Faced with this reality and the relevance of the theme to society, in recent years the study of techniques based on Computer Vision and Machine Learning that can somehow help specialists in the task of identifying nodules in Computed Tomography (CT) images has been intensified. Therefore, this study proposes to investigate an approach based on Convolutional Neural Networks with multiple inputs, called Multi-Scale, in order to obtain better and more accurate predictions of the classes of objects present in the images. In addition, the study also proposes to investigate the application of pre-processing methods on CT images based on Spatial Filtering, aiming to improve it and highlight features. The results obtained from the experiments carried out in this study are promising and demonstrate, in principle, that the use of Multi-Scale Convolutional Neural Networks can in fact be a solution for predicting medical images that present a great variability in the size of objects, while reach the 86% accuracy mark. In addition, the use of Digital Image Processing techniques proved to be a great ally in enhancing the information present in the CT images, managing, in some cases, to increase the accuracy of the network by up to 4%.

Keywords: deep learning; nodule detection; multi-scale; spatial filtering.

LISTA DE FIGURAS

Figura 1 - Anatomia de uma imagem de TC.....	12
Figura 2 - Passos fundamentais do Projeto (<i>Pipeline</i>).....	13
Figura 3 - Exemplos de distorções/ruídos apresentados nos <i>patches</i>	14
Figura 4 - Comparativo de <i>patches</i> com e sem a aplicação do Filtro Bilateral.....	16
Figura 5 - <i>Patches</i> – região de interesse e segmentação em <i>grid</i>	17
Figura 6 - Agrupamento dos <i>patches</i> nas entradas da rede.....	18
Figura 7 - Arquitetura (resumida) da CNN Multi-Escala.....	19
Figura 8 - Matriz de Confusão.....	20

LISTA DE TABELAS

Tabela 1 - Resultados dos experimentos com a CNN.....	21
---	----

LISTA DE ABREVIATURAS E SIGLAS

2D	Bidimensional
3D	Tridimensional
API	Application Programming Interface
CNN	Convolutional Neural Network
FN	Falso Negativo
FP	Falso Positivo
IDRI	Image Database Resource Initiative
LIDC	Lung Image Database Consortium
LUNA16	LUNg Nodule Analysis 2016
ML	Machine Learning
OMS	Organização Mundial da Saúde
PX	Píxel
ROI	Region Of Interest
TC	Tomografia Computadorizada
VC	Visão Computacional
VN	Verdadeiro Negativo
VP	Verdadeiro Positivo

SUMÁRIO

1	INTRODUÇÃO.....	10
2	METODOLOGIA.....	12
2.1	Conjunto de Dados.....	13
2.2	Extração das Regiões de Interesse.....	14
2.2.1	Filtragem Espacial.....	14
2.3	Rede Neural Convolucional.....	16
3	EXPERIMENTOS E RESULTADOS.....	20
3.1	Experimentos.....	20
3.2	Resultados.....	21
4	CONSIDERAÇÕES FINAIS.....	22
	REFERÊNCIAS.....	23

1 INTRODUÇÃO

Os tecidos e órgãos do corpo humano são compostos por células que normalmente crescem, se dividem e morrem o tempo todo de forma ordenada e controlada. Entretanto, devido a diversos fatores, nem sempre o corpo humano consegue manter esta dinâmica, ocasionando tumores que em seu pior estado podem evoluir para um câncer. O Câncer é um termo genérico dado a um grande grupo de doenças que podem afetar qualquer parte do corpo, sendo ele caracterizado por uma disfunção celular que ocasiona um crescimento desordenado das células e produzem excesso de tecidos (PFIZER, [ca. 2020]).

Segundo a Organização Mundial da Saúde (OMS), o câncer de pulmão é o tipo de câncer que mais causa mortes no mundo. Somente em 2020 foram mais de 2,2 milhões de casos que resultaram na morte de cerca de 1,8 milhão de pessoas, representando cerca de 18% das quase 10 milhões de mortes por câncer que ocorreram em todo mundo no referido ano (WHO, 2022).

Devido à alta taxa de letalidade advinda deste tipo de câncer é muito importante que ele seja diagnosticado o mais rápido possível, porém, existem diversos aspectos que dificultam esta tarefa. A primeira delas está na detecção precoce do tumor, já a segunda é que a triagem de câncer de pulmão por exames de Tomografia Computadorizada (TC) ainda é um processo mecânico realizado por radiologistas – o que além torná-los um agente propenso a falhas traz um fardo enorme quando há um grande volume de exames. Uma falha durante este estágio comprometerá o resto do processo, pois, a predição da não existência de um nódulo em um exame (falso negativo) dispensará a fase de classificação posterior, o que consequentemente resultará em um diagnóstico falho do paciente.

Diante desta realidade, relevância do tema para a sociedade e os avanços tecnológicos e computacionais ocorrido nos últimos anos, cada vez mais estão sendo estudadas e aplicadas técnicas e algoritmos baseados em Visão Computacional (VC) e Aprendizagem de Máquina (em inglês, *Machine Learning*, ou simplesmente ML) em mecanismos que possam auxiliar especialistas na tarefa de identificação de nódulos em imagens de TC. Atualmente, algumas técnicas de aprendizado de máquina, como o Aprendizado Profundo (em inglês, *Deep Learning*), têm sido aplicadas com relativo sucesso em tarefas de análise de imagens médicas e auxílio ao diagnóstico baseado em imagens, porém, ainda assim nos deparamos com diversos desafios na tarefa de automatização desse processo, principalmente relacionados com a escassez de conjuntos de dados que possam ser utilizados para estudo e as altas taxas de ruídos presentes nas imagens médicas. Diante disso, o presente estudo tem como objetivos:

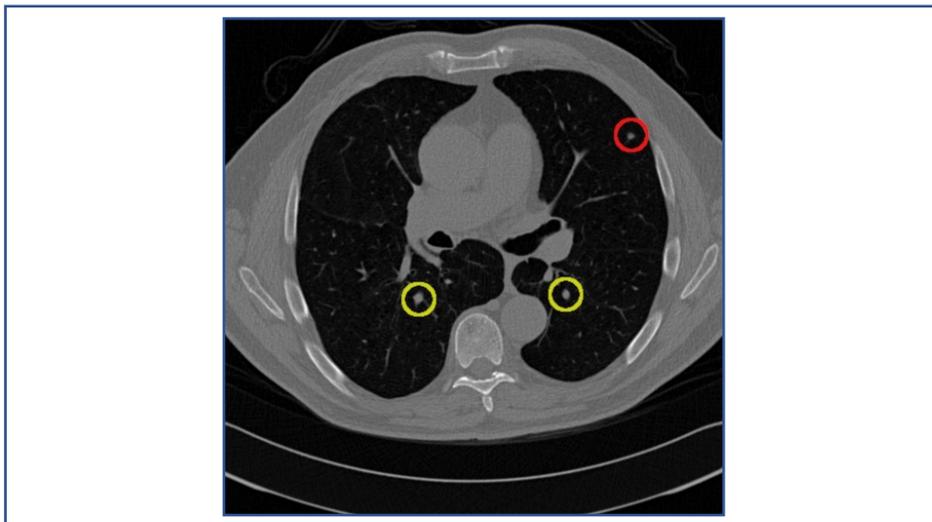
1. Construção de um sistema automático de identificação de nódulos pulmonares em imagens de TC.
2. Estudo dos métodos de Aprendizado Profundo, com foco em Redes Neurais Convolucionais.
3. Estudo e aplicação de métodos de pré-processamento sobre as imagens de TC para aprimorá-las e realçar características e estruturas da imagem.
4. Implementar, treinar e validar o modelo.

O restante deste trabalho está estruturado como se segue. No Capítulo 2, será apresentado o método proposto, onde abordaremos sobre o conjunto de dados utilizado, processamento das imagens e rede neural. No Capítulo 3, são descritos os detalhes sobre a condução dos experimentos e os resultados obtidos com o método. E por fim, no Capítulo 4, traremos as conclusões e as perspectivas futuras sobre o tema.

2 METODOLOGIA

A utilização de técnicas baseadas em Aprendizado Profundo para classificação de imagens médicas comumente enfrenta diversos desafios, estando os principais relacionados com: (1) a escassez de imagens que possam ser utilizadas para treinamento; (2) a qualidade das imagens que, geralmente, apresentam um alto índice de ruídos; (3) as regiões de interesse (em inglês, *Region Of Interest*, ou simplesmente ROI) que são relativamente pequenas, “confusas” e apresentam tamanho variável, dificultando a automatização do processo de predição. Na Figura 1 é demonstrada uma típica imagem de TC obtida a partir do conjunto de dados do *LUNG Nodule Analysis 2016* (LUNA16). Nela, é apresentada uma imagem de tamanho 512x512 (píxeis ou px), circulado em vermelho um nódulo e em amarelo regiões visualmente semelhantes, porém que não são nódulos.

Figura 1 - Anatomia de uma imagem de TC.

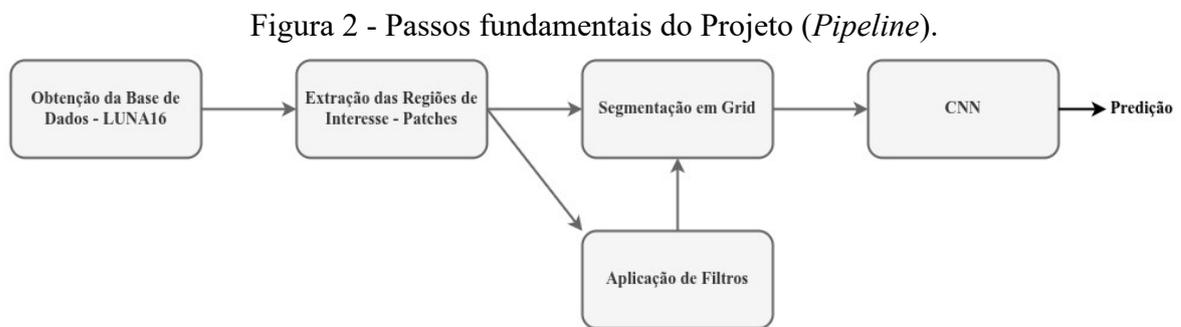


Fonte: O autor (2023).

A detecção por meio de auxílio computadorizado geralmente consiste em dois estágios, sendo eles a detecção de possíveis nódulos e a redução de falsos positivos. Embora os métodos convencionais tenham produzido resultados promissores, eles ainda trazem consigo algumas desvantagens. A primeira delas é que as imagens de TC são inerentemente tridimensionais (3D) e quando reduzimos as buscas a um plano bidimensional (2D) perdemos informações. Outra, é que os sistemas tradicionais comumente detectam candidatos com base em algumas suposições básicas, como, por exemplo, nódulos têm formato esférico. Logo, com variabilidade de formas, tamanhos, texturas e os ruídos, os descritores de baixo nível falham no processo de capturar características discriminatórias dos objetos, resultando assim

em falhas de identificação (DING *et al.*, 2017).

Com o propósito de sanar o problema da identificação, o presente estudo propõe-se a investigar uma abordagem com uso de uma Rede Neural Convolutacional (em inglês, *Convolutional Neural Network* – também conhecida como CNN) com múltiplas entradas, chamada de Multi-Escala¹. A ideia por trás desta abordagem é averiguar se é possível obter resultados satisfatórios de forma que este método possa ser uma alternativa para a resolução do problema da identificação neste contexto que apresenta uma alta variabilidade no tamanho dos objetos. Além disso, para contornar o problema dos ruídos, o estudo investigará também a aplicação de Filtros Espaciais sobre as imagens, visando aprimorá-las e realçar características. A Figura 2 ilustra os passos fundamentais do projeto.



Fonte: O autor (2023).

2.1 Conjunto de Dados

O primeiro passo dado neste trabalho ocorreu pela escolha do conjunto de dados utilizado. Nesse estudo, foi optado pelo conjunto de dados do Desafio *LUNG Nodule Analysis 2016* (LUNA16). O LUNA16 foi um desafio realizado no ano de 2016, aberto ao público, e cujo objetivo foi fornecer uma oportunidade para os participantes testarem seus algoritmos de detecção computadorizada em um conjunto de dados padronizado. O conjunto de dados LUNA16 é um subconjunto LIDC-IDRI (*Lung Image Database Consortium/Image Database Resource Initiative*) que consiste – em sua completude – de um total de 888 exames de TC, exames esses que possuem diferentes quantidades de imagens do pulmão, capturadas em formato transversal, com cada uma possuindo dimensões de 512x512 (píxeis). Além das imagens, o conjunto fornece também anotações de 1186 nódulos realizadas por quatro radiologistas experientes, que incluem, coordenadas tridimensionais e os diâmetros dos nódulos em milímetro (LUNG..., 2016).

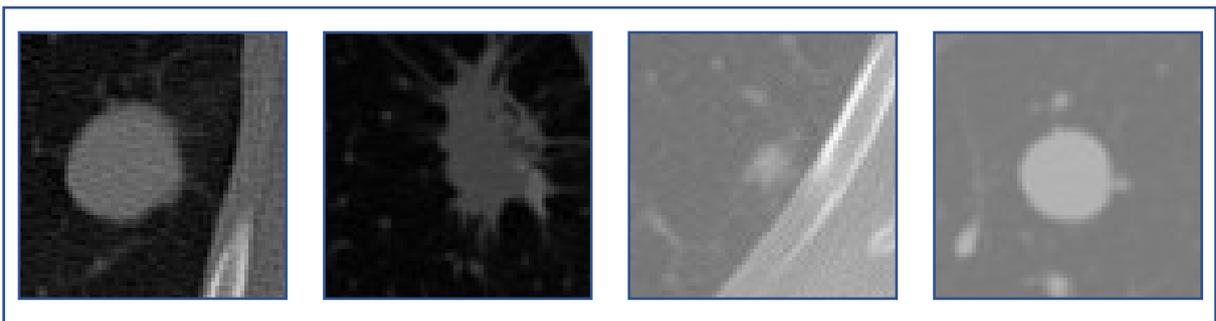
¹ O termo Multi-Escala é uma tradução literal do termo em inglês *Multi-Scale* – nomenclatura amplamente difundida neste cenário –, portanto, será adotado o termo traduzido e separado por hífen quando estivermos nos referindo ao tipo da CNN.

2.2 Extração das Regiões de Interesse

A partir da obtenção do conjunto de dados a próxima etapa a ser realizada é a extração das regiões de interesse. Chamados de *Patches*, os pontos de interesse são extraídos através de suas coordenadas em formato de subáreas – de tamanho 64x64 (píxeis) – de forma centralizada. Entretanto, a extração direta dessas áreas traz consigo distorções inerentes à obtenção das imagens de TC, ocasionando problemas para os sistemas automáticos de detecção.

A palavra Tomografia é uma junção dos termos gregos *Tomos* – que significa “Fatia” ou “Seção” – e *Grphe* – que significa “Desenho”. Quando falamos em Tomografia estamos nos referindo a um procedimento de imagem de Raio-X Computadorizado que nos permite visualizar seções de um objeto, normalmente, um órgão. Com essa técnica é possível obter ensaios não destrutivos (e/ou invasivos) que possibilitam a visualização do interior de um corpo sem a necessidade de violá-lo (SEBASTIÃO, 2022). Devido à complexidade do método e os princípios substanciais com que ele interage em um corpo acaba que por muitas vezes ocorre distorções nas imagens obtidas – os chamados ruídos, a Figura 3 ilustra alguns *patches* com nódulo, neles são apresentados distorções do tipo Borramento e “Sal e Pimenta” (em inglês, *Salt-And-Pepper*).

Figura 3 - Exemplos de distorções/ruídos apresentados nos *patches*.



Fonte: O autor (2023).

2.2.1 Filtragem Espacial

Visando contornar esse tipo de empecilho – que corrobora para distorções das informações em uma imagem – e muitos outros – surgiram os métodos de Processamento Digital de Imagens. Esses métodos provêm técnicas que possibilitam a melhoria das informações visuais de uma imagem para interpretação humana e/ou para “percepções” automáticas realizadas por máquinas. Por isso, na busca por tornar as imagens mais nítidas e

com isso alcançar melhores resultados com a rede foi optado pela utilização de métodos de transformação de intensidade com filtragem espacial, conhecido também como Domínio Espacial, métodos que atuam diretamente nos píxeis da imagem e que são, em geral, computacionalmente mais eficientes e requerem menos recursos de processamento para serem realizados (GONZALEZ; WOODS, 2010).

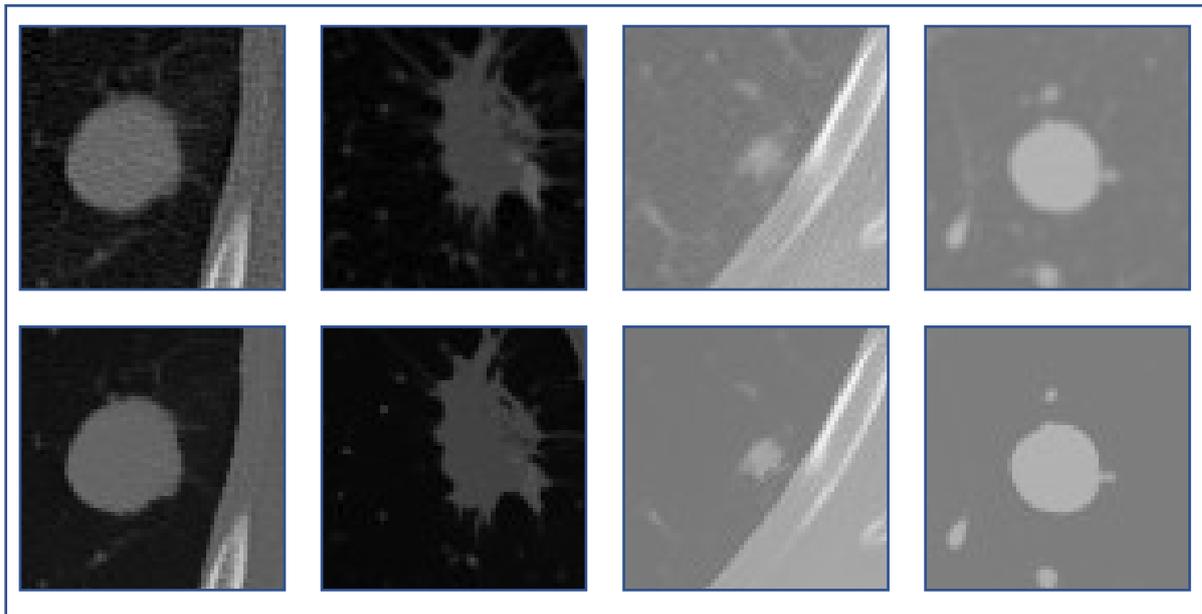
Devido às altas taxas de ruídos presentes nas imagens de TC e em busca do melhor método para esta atividade, foi optado pela utilização de filtros de dois segmentos: Lineares de Suavização e Não Lineares.

Os Filtros Lineares recebem essa nomenclatura por seu processo – denominado de Convolução – consistir em, a partir de um ponto central efetuar uma combinação linear entre as intensidades dos píxeis da vizinhança e um filtro (máscara), em geral, esse processo reduz ruídos de alta frequência suavizando falsos contornos e reduzindo detalhes irrelevantes (Passa-Baixa), porém, acarreta perda de nitidez nas bordas. Desta categoria foram utilizados dois filtros: (1) Filtro de Média, efetua a troca do píxel central pela média aritmética dos píxeis da vizinhança definida pela máscara; (2) Filtro Gaussiano, ele recebe este nome por ser método está calcado na teoria da probabilidade em forma de curva de distribuição Gaussiana, de uma forma bem grosseira, esse filtro efetua uma média ponderada na vizinhança de forma que seus pesos diminuem ao se afastar do centro (LI *et al.*, 2021).

Já a segunda categoria, os filtros Não Lineares, são baseados em algum tipo de classificação, substituindo o valor do píxel central de uma região pelo valor determinado pelo resultado da classificação. Sua principal vantagem é uma remoção de ruído mais seletiva que “força” pontos com valores distintos a ficarem mais próximos da sua vizinhança, ao passo que preserva as arestas enquanto suaviza com menos borramento que os Lineares. Desse, também foram utilizados dois filtros: (1) Filtro de Mediana, consiste em substituir o ponto central pelo valor da mediana da vizinhança; (2) Bilateral, este filtro trabalha semelhantemente ao filtro Gaussiano, com o acréscimo de operações para agrupamento de cor, a ideia é que para um píxel influenciar outro não basta ele apenas existir e ocupar um espaço, mas também possuir um valor semelhante (TOMASI; MANDUCHI, 1998; GONZALEZ; WOODS, 2010).

No presente estudo a filtragem é realizada após a extração dos *patches* e são utilizadas as implementações dos filtros providas pela Biblioteca Multiplataforma de Visão Computacional OpenCV (JOSHI, 2015). A Figura 4 ilustra um exemplo de transformação utilizando como base os *patches* apresentados na Figura 3, na parte superior temos o antes e na inferior o depois da aplicação do Filtro Bilateral.

Figura 4 - Comparativo de *patches* com e sem a aplicação do Filtro Bilateral.



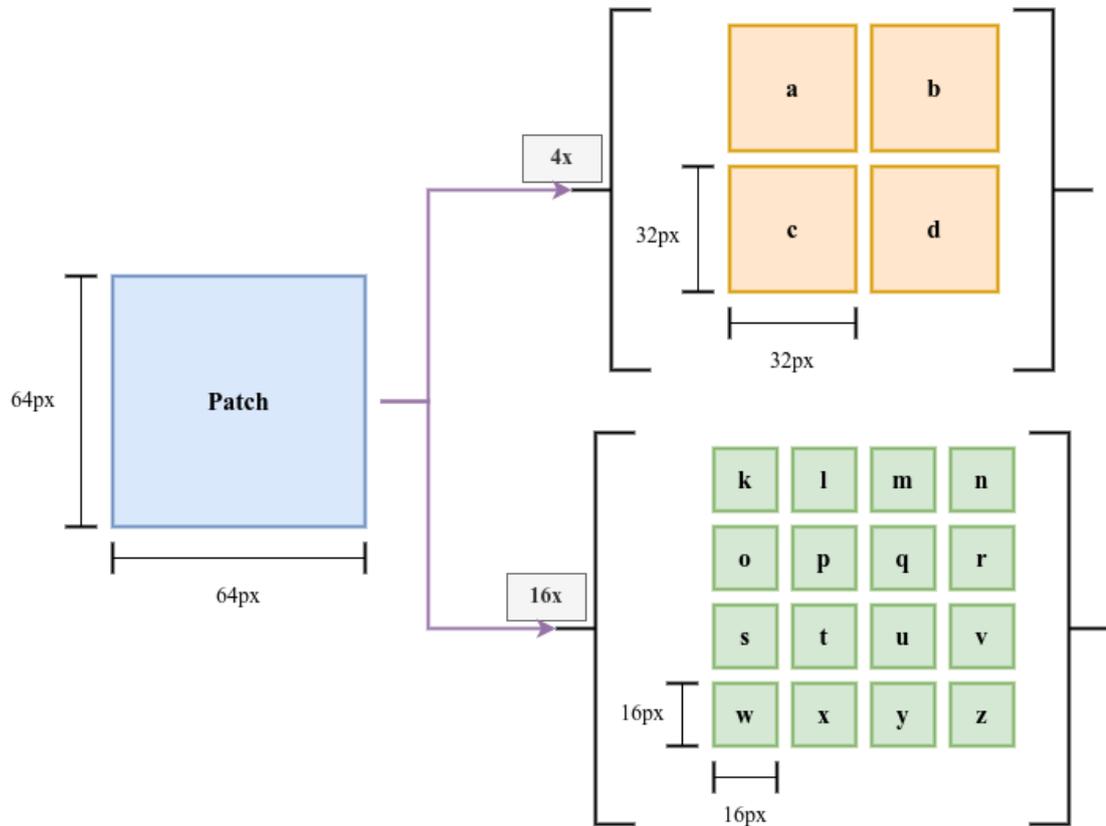
Fonte: O autor (2023).

2.3 Rede Neural Convolutacional

A terceira parte do projeto se concentra propriamente na construção da rede que será utilizada para identificação dos nódulos. Como mencionado anteriormente, neste estudo será adotada uma Rede Neural Convolutacional Multi-Escala. Esta técnica envolve a criação de várias camadas convolucionais com diferentes resoluções de entradas das imagens e de tamanhos de filtros. Cada camada processa as informações em uma determinada escala, permitindo que a CNN seja capaz de reconhecer e capturar detalhes e padrões discriminatórios em objetos de diferentes tamanhos e resoluções, capturando assim diferentes níveis de detalhes e produzindo mapas receptivos que conseguem captar recursos de baixo e alto nível que são combinados em uma camada de fusão (ELIZAR *et al.*, 2022).

Sendo assim, a rede empregada neste estudo é constituída de 21 entradas, sendo: uma de tamanho 64x64 (px), quatro de 32x32 (px) e dezesseis de 16x16 (px). Como ilustrado na Figura 5, a alimentação de dados na rede é efetuada utilizando os *patches* produzidos pelo processo descrito na Seção 2.2 (Extração das Regiões de Interesse) – em azul –, que antes de entrar passam por um processo de segmentação em *grid* (ou grade) que dará origem a quatro *patches* de 32x32 (px) – em amarelo – e dezesseis *patches* de 16x16 (px) – em verde.

Figura 5 - *Patches* – região de interesse e segmentação em *grid*.

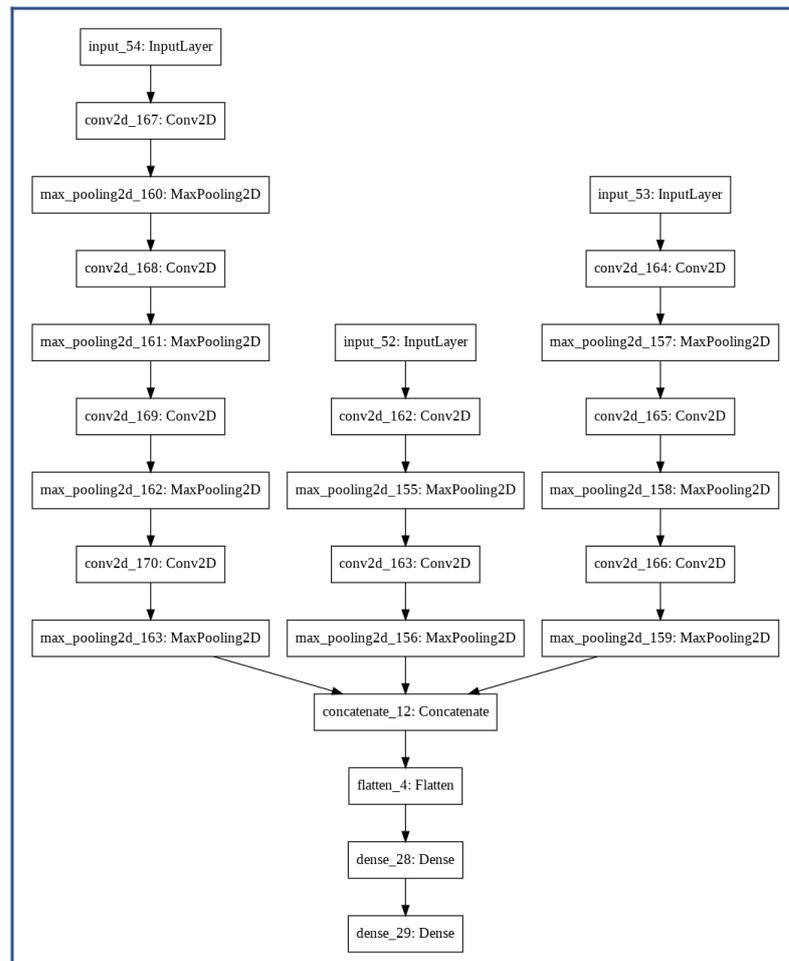


Fonte: O autor (2023).

Como demonstrado na Figura 6, exemplificando a entrada de um lote de *patches* na rede, cada entrada é responsável por uma região da imagem com base em sua localização em relação ao “*Patch Base*” que a originou. Esse procedimento tem o intuito de instruir cada entrada da rede a se especializar em uma única região para aprender quais características a compõe, como possíveis semicírculos em fragmentos, tonalidades, intensidades e texturas, além de reduzir a quantidade de elementos que serão analisados por região e consequentemente possíveis informações ruidosas.

Para construção da rede foi utilizado a Biblioteca Keras, uma API (*Application Programming Interface*) de aprendizado profundo escrita em Python, executada sobre a plataforma de aprendizado de máquina TensorFlow (KERAS, [2020]). Na Figura 7 é mostrado uma visão geral e simplificada da arquitetura da CNN construída, onde os termos “input_52: InputLayer”, “input_53: InputLayer” e “input_54: InputLayer” representam respectivamente os conjuntos de entradas de tamanho 16x16, 32x32 e 64x64 (píxeis). As camadas de convolução (*Conv2D*) são configuradas com *padding=same* e *activation=relu*. Já as camadas de agrupamento (*MaxPooling2D*) são configuradas com *pool_size=(2, 2)*, *strides=(2, 2)* e *padding=valid*. As camadas Densa 1 e 2 (saída) possuem *activation=relu* e

Figura 7 - Arquitetura (resumida) da CNN Multi-Escala.



Fonte: O autor (2023).

3 EXPERIMENTOS E RESULTADOS

3.1 Experimentos

Nos experimentos conduzidos neste estudo foram utilizados um total de 700 áreas de interesse, como mencionado anteriormente, obtidas a partir da base de dados LUNA16, sendo 350 áreas com nódulo e 350 sem nódulo, ou seja, um conjunto balanceado. Os resultados foram aferidos com limiar de classificação/predição (*threshold*) da saída da rede de 50%, e o resultado final é dado por uma média aritmética de três testes sobre o lote utilizando como métrica a acurácia, que mede a capacidade da rede de classificar corretamente se a região contém ou não um nódulo. A acurácia é calculada da seguinte forma:

$$\text{Acurácia} = \frac{\text{VP} + \text{VN}}{\text{VP} + \text{FN} + \text{FP} + \text{VN}}$$

A fórmula de cálculo da acurácia é obtida através da Matriz de Confusão apresentada na Figura 8. Podemos definir a Matriz de Confusão como uma tabela de frequência de classificação para uma classe do modelo, sendo ela um “casamento” entre o valor esperado/real e valor predito (SOUZA, 2019). Para o modelo apresentado existem quatro tipos de resultados possíveis: Verdadeiro Positivo (VP), Verdadeiro Negativo (VN), Falso Positivo (FP) e Falso Negativo (FN).

Figura 8 - Matriz de Confusão.

		VALOR PREDITO	
		Positivo	Negativo
VALOR REAL	Positivo	VP	FN
	Negativo	FP	VN

Fonte: O autor (2023).

3.2 Resultados

A Tabela 1 descreve os resultados obtidos através dos experimentos realizados. A partir desses resultados é possível constatar que a CNN alcançou uma acurácia de 86% no experimento sem o uso de filtro. Tomaremos esse valor como resultado de controle para fins comparativos.

Além do resultado de controle, a tabela mostra também os resultados com o uso dos filtros, apresentando-os de forma decrescente, ou seja, do melhor para o pior. A partir deles constatamos que alguns tipos de filtros promoveram de fato uma melhora na acurácia da rede, destacando-se nessa tarefa os filtros “Não Lineares”, com o Filtro Bilateral sendo o melhor ao alcançar a marca de 90%, seguido pelo Filtro de Mediana com 89%, representando respectivamente um ganho de 4% e 3% de acurácia em relação ao resultado de controle. Seguindo eles – na terceira posição – vem o Filtro Gaussiano com 87%. Por último, ficou o Filtro de Média com somente 85% de acurácia, o que representa uma perda de 1% em relação ao resultado de controle.

Tabela 1 - Resultados dos experimentos com a CNN.

Filtro	Acurácia	Tipo
Sem (Controle)	86%	
Bilateral	90%	Não Linear
Mediana	89%	Não Linear
Gaussiano	87%	Linear
Média	85%	Linear

Fonte: O autor (2020).

4 CONSIDERAÇÕES FINAIS

Como elucidado durante todo esse estudo, o processo de automatização da identificação de nódulos pulmonares traz consigo diversos desafios que não possuem métodos triviais de serem resolvidos. Neste estudo propomos uma investigação de um sistema baseado em Redes Neurais Convolucionais do tipo Multi-Escala, para resolver o problema da identificação dos objetos com tamanhos e resoluções variadas. Com este método, conseguimos alcançar a marca de 86% de acurácia, e com o acréscimo dos Filtros Espaciais elevamos os resultados para até 90%, o que podemos considerar como resultados satisfatórios diante das dificuldades enfrentadas nesse contexto.

A partir dos métodos explorados neste estudo obtivemos resultados promissores que demonstram, a princípio, que o uso de Redes Neurais Convolucionais Multi-Escala podem ser de fato uma solução para predição de imagens médicas que apresentam uma grande variabilidade no tamanho dos objetos. Além disso, o emprego de técnicas de Processamento Digital de Imagens – com o uso dos filtros – se mostraram um grande aliado no realce das informações presentes nessas imagens ao conseguirem aumentar a capacidade da rede de classificar os objetos.

Portanto, depois de todos os experimentos realizados até aqui, é possível ter uma ideia do que pode ou não funcionar para o aprimoramento do modelo de aprendizado. Sendo os resultados totalmente providos por empiria, é de suma importância que haja mais estudos neste campo buscando aprimorar os modelos existentes e em busca de novas técnicas e abordagens. Para pesquisas futuras, fica como desafio o estudo de técnicas de aumento artificial da base de dados, popularmente conhecida como *Data Augmentation*, em busca de resolver o problema da falta de dados para treinamento da rede, tema que foi um constante empecilho no decorrer do projeto (SHAH, 2022).

REFERÊNCIAS

- DING, Jia *et al.* Accurate pulmonary nodule detection in computed tomography images using deep convolutional neural networks. **ArXiv**, [S. l.], 2017. Disponível em: <https://doi.org/10.48550/ARXIV.1706.04303>. Acesso em: 10 jun. 2023.
- ELIZAR, Elizar *et al.* A review on multiscale-deep-learning applications. **Sensors**, [S. l.], v. 22, n. 19, p. 7384, 2022. Disponível em: <https://doi.org/10.3390/s22197384>. Acesso em: 10 jun. 2023.
- GONZALEZ, Rafael C.; WOODS, Richard E. **Processamento digital de imagens**. 3. ed. São Paulo: Pearson Education, 2010.
- JOSHI, Prateek. **OpenCV with Python by example: build real-world computer vision applications and develop cool demos using OpenCV for Python**. Birmingham: Packt Publishing, 2015.
- KERAS. **About Keras**. [S. l.]: Keras Team, [2020]. Disponível em: <https://keras.io/about>. Acesso em: 1 maio 2023.
- LI, Pei *et al.* Overview of image smoothing algorithms. **Journal of Physics: Conference Series**, Wuhan, v. 1883, p. 012024, 2021. Disponível em: <https://doi.org/10.1088/1742-6596/1883/1/012024>. Acesso em: 10 jun. 2023.
- LUNG Nodule Analysis 2016. **Grand Challenge**. [S. l.], 2016. Disponível em: <https://luna16.grand-challenge.org>. Acesso em: 1 abr. 2023.
- PFIZER. **Câncer de pulmão**. São Paulo: Pfizer, [ca. 2020]. Disponível em: <https://www.pfizer.com.br/sua-saude/oncologia/cancer-de-pulmao>. Acesso em: 1 abr. 2023.
- SEBASTIÃO, Amanda. **Principais recursos avançados do PACS para tomografia computadorizada**. [S. l.], 5 maio 2022. Disponível em: <https://www.pixeon.com/blog/tomografia-computadorizada>. Acesso em: 1 abr. 2023.
- SHAH, Deval. **The essential guide to data augmentation in deep learning**. [S. l.], 6 maio 2022. Disponível em: <https://www.v7labs.com/blog/data-augmentation-guide>. Acesso em: 1 abr. 2023.
- SOUZA, Emanuel G. de. **Entendendo o que é Matriz de Confusão com Python**. [S. l.], 27 mar. 2019. Disponível em: <https://medium.com/data-hackers/entendendo-o-que-é-matriz-de-confusão-com-python-114e683ec509>. Acesso em: 1 maio 2023.
- TOMASI, C.; MANDUCHI, R. Bilateral filtering for gray and color images. **Sixth International Conference on Computer Vision (IEEE)**, Bombay, 7 jan. 1998. Disponível em: <https://doi.org/10.1109/ICCV.1998.710815>. Acesso em: 10 jun. 2023.
- WHO. **Cancer**. Genebra: WHO, 3 fev. 2022. Disponível em: <https://www.who.int/news-room/fact-sheets/detail/cancer>. Acesso em: 1 abr. 2023.