

1.03.03 – Ciência da Computação / Metodologia e Técnicas da Computação

AUXÍLIO COMPUTADORIZADO À IDENTIFICAÇÃO DE NÓDULOS PULMONARES BASEADO EM TÉCNICAS DE APRENDIZADO PROFUNDO

José Y. R. A. Gonçalves*, Eduardo V. Silva, Rodolfo C. Cavalcante

1. Estudante do curso de Ciência da Computação na Universidade Federal de Alagoas (UFAL)
2. Estudante do curso de Ciência da Computação na UFAL
3. Professor do curso de Ciência da Computação na UFAL – Orientador

Resumo:

O câncer de pulmão é o tipo de câncer que mais causa mortes no mundo e sua principal manifestação ocorre devido ao aparecimento de lesões no tecido pulmonar. Dada a sua alta taxa de mortalidade é extremamente importante que seu diagnóstico seja feito o mais rápido possível. Perante esta realidade, esse tipo de câncer tem sido cada vez mais estudado e na tentativa de automatizar o diagnóstico, nos deparamos com diversos desafios que afetam a identificação de nódulos presentes, em sua maioria, relacionados aos ruídos presentes nas imagens tomográficas computadorizadas (TC).

Diante disso, este estudo tem como objetivo principal o desenvolvimento de uma rede neural profunda multiescala para detecção automática de possíveis nódulos presentes nas imagens e como objetivo secundário, a aplicação de métodos de pré-processamento sobre as imagens TC para melhoramento das informações presentes.

Palavras-chave: Aprendizado Profundo, Detecção Nodular, Multiescala.

Apoio financeiro: UFAL.

Trabalho selecionado para a JNIC: UFAL.

Introdução

O câncer de pulmão é o tipo de câncer que mais causa mortes no mundo e sua principal manifestação ocorre devido ao aparecimento de lesões no tecido pulmonar (ALA, 2021). Segundo a Organização Mundial da Saúde, apenas em 2020 foram mais de 2.20 milhões de casos em todo o mundo, levando a óbito cerca de 1.80 milhões de pessoas – ocupando a primeira posição do ranking de mortes por câncer (WHO, 2021). Devido a sua alta taxa de mortalidade é de suma importância que esse tipo de câncer seja identificado e diagnosticado o mais rápido possível, aumentando assim as chances de cura por parte da pessoa afetada.

A maior dificuldade em sua identificação está em distinguir nódulos verdadeiros de outras estruturas pulmonares, como brônquios, traqueia e outros tecidos. Além disso, nesta tarefa, aspectos humanos podem interferir na identificação dos nódulos por parte dos especialistas, tais como: cansaço, pressão, inexperiência e fatores psicológicos. E isso é extremamente ruim, pois em caso de falhas na identificação de nódulos, a fase de classificação está comprometida, pois a indicação da inexistência de nódulos em um exame (falsos negativos) dispensará a fase de classificação posterior, o que por sua vez resultará em um diagnóstico falho do paciente.

É neste sentido que algoritmos e técnicas de visão computacional e aprendizagem de máquina têm sido utilizados no escaneamento de imagens TC de modo a identificar regiões onde estão localizados prováveis nódulos. Portanto, este estudo tem como objetivo principal o desenvolvimento de técnicas computacionais baseadas em redes neurais profundas multiescala que assistam um médico no processo de detecção de nódulos pulmonares em imagens de TC, indicando-o pontos de possíveis nódulos. Além disso, tem como objetivo secundário a aplicação de métodos de pré-processamento sobre as imagens TC, com o intuito de melhorar as informações presentes e facilitar o processo de detecção e classificação de seus objetos.

Metodologia

A utilização de aprendizagem de máquina (1) para classificação de imagens médicas comumente enfrenta alguns desafios, sendo os principais: a quantidade e a qualidade das imagens, além das regiões de interesse (ROIs)(2) que são relativamente pequenas. Com o propósito de sanar parte desses problemas, presentemente foi adotada a utilização de uma rede neural convolucional (CNN – *Convolutional Neural Network*) com múltiplas entradas, chamadas de multiescala (3). Esse tipo de rede permite uma maior sensibilidade sobre os dados examinados, principalmente em ocasiões que há uma grande variedade no tamanho dos objetos, isso porque ela consegue extrair informações independentes da escala, além de conseguir melhores representações das estruturas hierárquicas da região.

Diante disso, a arquitetura empregada é constituída de três grupos de tamanho de patch que totalizam 21 entradas e recebem fragmentos da imagem TC, sendo: 1 de tamanho 64x64(px), 4 de 32x32(px) e 16 de 16x16(px). (Representação da arquitetura na Figura 1)

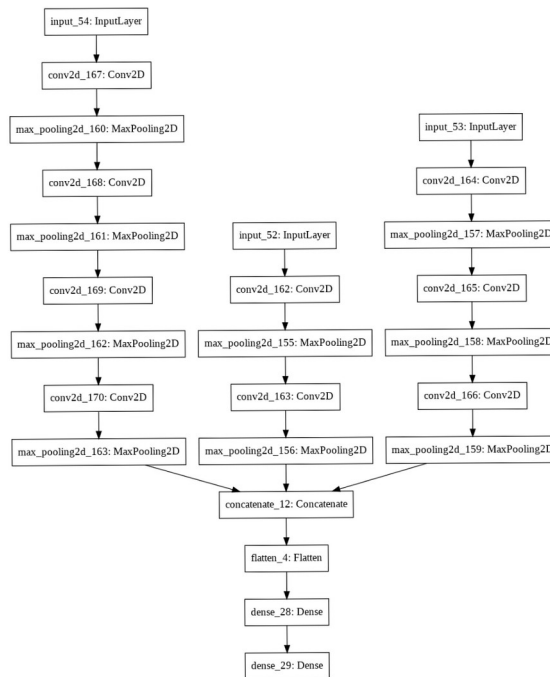


Figura 1: Arquitetura resumida da rede

Para a fazer a entrada da rede é efetuado a extração de regiões de interesses das imagens TC – sendo essas regiões áreas de tamanho 64x64(px), chamadas de *patch* base –, que possuem ou não a presença de nódulo. Posteriormente, os patches base (denotados por $P_1, P_2 \dots P_\alpha$) são segmentadas em grid (Figura 2) e os fragmentos são agrupados por sua posição em relação ao patch base que o originou, como mostrado na Figura 3. Para então serem inseridas na rede. Esse procedimento tem como objetivo ensinar a rede a identificar porções da imagem, encontrando possíveis objeto e formatos.

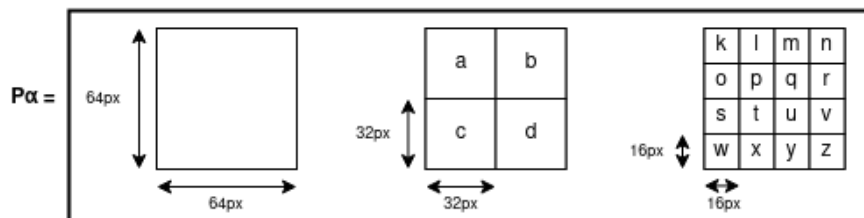


Figura 2: Segmentação de uma regiões de interesse em formato de grid.

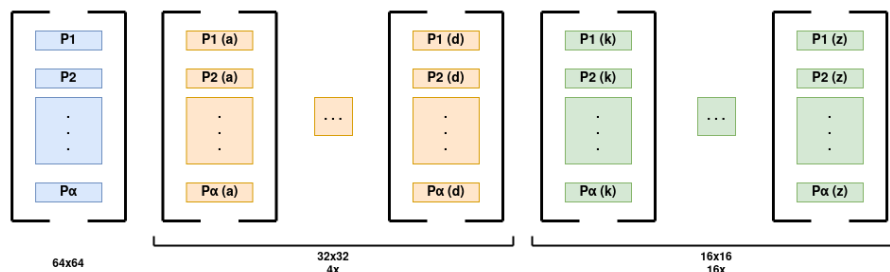


Figura 3: Agrupamento das regiões de interesse e seus fragmentos.

Ainda assim, apenas a arquitetura da rede não soluciona todos os desafios supracitados, restando entre eles o problema da qualidade das imagens, que ocorrem inerentemente a obtenção das imagens TC, os chamados ruídos. Com a finalidade de contornar esse empecilho foi optado pelo uso de métodos de processamento de imagens. Esses métodos consistem em técnicas para melhoria das informações visuais de imagens para a interpretação humana ou para percepção automática feita por máquinas. Sendo utilizado aqui, em especial, métodos de transformações de intensidade – filtros.

Devido alta quantidade de ruído presente nas imagens TC, foram utilizados filtros de duas categorias: lineares e não lineares. Os filtros lineares recebem essa nomenclatura por seu processo, denominado convolução, poder ser descrito como uma função linear das intensidades dos pixels. Em geral, esse processo

reduz ruído de alta frequência, suavizando a imagem, porém pode acarretar perda de nitidez. Já os filtros denominados não lineares, é qualquer tipo de filtro que não se encache na definição anterior. Sua principal vantagem é uma remoção de ruído seletiva – enfatizando características específicas locais da imagem –, por exemplo, bordas. Sendo assim, foram utilizados dois filtros da categoria dos lineares, sendo eles: Blur e Gaussian Blur; e dois filtros dos não lineares: Median Blur e Bilateral. (4-6)

Resultados e Discussão

Nos experimentos a métrica utilizada como parâmetro de avaliação dos resultados foi a acurácia da rede, e o resultado final é constituído da média obtida a partir de três testes. Em cada teste foram utilizados um total de 700 regiões de interesses, obtidas a partir do *dataset* LUNA16 (7), um subconjunto DICOM baseada no banco de dados *Lung Image Database Consortium (LIDC-IDRI)*, sendo, 350 imagens com nódulos e 350 sem nódulos. A configuração da rede utilizada nos experimento foi: camada de convolução sendo 2, no máximo 64 filtros de extração de características, otimizador *Adam* e a função de erro entropia cruzada binária (*Binary Cross Entropy*). E no treinamento configuramos o *Batch_Size* como 32 e 20 épocas.

Sobre os resultados, partimos de um experimento de controle, ou seja, sem o uso de filtros, no qual a rede conseguiu alcançar um índice de 86% de acurácia. A partir dos métodos explorados foi possível constatar que as técnicas de processamento de imagem proveram uma melhora significativa na acurácia, demonstrando ser uma a bordagem promissoras. Destacando-se nessa tarefa os filtros de estatística de ordem (não lineares), que são caracterizados pela classificação dos pixels contidos em uma região e substituindo o pixel central pelo valor determinado pelo resultado da classificação, assim realçando a bordas dos objetos. Com isso, os filtros Bilateral e Median Blur foram os que apresentaram os maiores ganhos, alcançando a índice de 90% e 89% de acurácia, respectivamente; Seguidos pelo filtro Gaussian Blur com 87% de acurácia. Em relação ao filtro linear e de desfoque Blur, em todos os testes ocorreu perda de desempenho no modelo. Resultados mostrado de forma organizada na tabela 1.

Filtro	Acurácia	Tipo
Sem (Controle)	86%	
Bilateral	90%	Não Linear
Median Blur	89%	Não Linear
Gaussian Blur	87%	Linear
Blur	85%	Linear

Tabela 1: Resultado obtidos nos testes.

Conclusões

As soluções para classificação de imagens médicas vêm sendo cada vez mais estudadas, principalmente devido à necessidade social e, também, o avanço tecnológicos e metodológico do ramo de aprendizagem de máquina ocorridos nos últimos anos. Sendo a classificação em imagens médicas um dos problemas mais desafiadores devido a limitada quantidade de dados e a dificuldade intrínseca do problema.

A partir de todos os experimentos realizados obtivemos resultados promissores que demonstram, a princípio, que a aplicação de uma rede multiescala pode ser uma possível solução para este tipo de classificação. Além disso, ficou demonstrado que as técnicas de processamento digital de imagem podem ser grandes aliados no realce de informações presente nas imagens tomográfica computadorizadas, tais como: tecidos e estruturas morfológicas. Cabe ressaltar que, devido à sensibilidade das redes neurais esses resultados podem apresentar variações de acordo com o modelo utilizado. Portanto, depois de todos os estudos ocorridos 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 o estudo de novas técnicas e abordagem. Para pesquisas futuras, fica como desafio a busca por técnica para o aumento artificial da base de dados (*Data Augmentation*)(8), algo que foi constante problema no decorrer do projeto, sendo essa uma das prováveis técnicas que serão exploradas futuramente.

Referências bibliográficas

1. APRENDIZAGEM DE MÁQUINA. In: WIKIPÉDIA, a enciclopédia livre. Flórida: Wikimedia Foundation, 2021. Disponível em: <https://pt.wikipedia.org/wiki/Aprendizado_de_m%C3%A1quina>. Acesso em: 27 de abr. de 2021.
 2. REGION OF INTEREST. In: WIKIPÉDIA, a enciclopédia livre. Flórida: Wikimedia Foundation, 2021. Disponível em: <https://en.wikipedia.org/wiki/Region_of_interest>. Acesso em: 19 de abr. de 2021.
 3. SHEN, W.; ZHOU, M.; Yang, F.; Yang, C; Tian, J. Multi-scale Convolutional Neural Networks for Lung Nodule Classification. In: Ourselin S., Alexander D., Westin CF., Cardoso M. (eds) Information Processing in Medical Imaging. IPMI 2015. Lecture Notes in Computer Science, vol 9123. Springer, Cham.
 4. GONZALEZ, R. C.; WOODS, Richard E. Processamento Digital de Imagem. 3ª ed. São Paulo: Pearson Prentice Hall, 2010.
 5. NIXON, M. S.; AGUADO, A. S. Feature Extraction and Image Processing. 1ª st. Oxford: Newnes, 2002.
 6. BILATERAL FILTERING. In: WIKIPÉDIA, a enciclopédia livre. Flórida: Wikimedia Foundation, 2021. Disponível em: <https://en.wikipedia.org/wiki/Bilateral_filter>. Acesso em: 19 de abr. de 2021.
 7. LUng Nodule Analysis, 2016. Disponível em: <<https://luna16.grand-challenge.org/>>. Acesso em: 19 de abr. de 2021.
 8. Claro, M.; Vogado, L.; Santos, J.; Veras, R.; Utilização de Técnicas de Data Augmentation em Imagens: Teoria e Prática. In: Sociedade Brasileira da Computação - SBC. Disponível em: <<https://sol.sbc.org.br/livros/index.php/sbc/catalog/download/48/217/455-1?inline=1>>. Acesso em: 27 de abr. de 2021.
- Lung Cancer Basics. American Lung Association, 2021. Disponível em: <<https://www.lung.org/lung-health-diseases/lung-disease-lookup/lung-cancer/learn-about-lung-cancer/lung-cancer-basics>>. Acesso em: 22 de abr. de 2021.
- Cancer. World Health Organization, 2021. Disponível em: <<https://www.who.int/news-room/fact-sheets/detail/cancer>>. Acesso em: 19 de abr. de 2021.