

Detecção de Nódulos da Tireoide em Exames de Termografia utilizando Redes Neurais Convolucionais em Cascata

Ricardo José Fernandes Anchieta Júnior¹, Italo Francyles Santos da Silva¹,
Aristófanês Corrêa Silva¹, Anselmo Cardoso de Paiva¹

¹Núcleo de Computação Aplicada (NCA) - Universidade Federal do Maranhão (UFMA)
Av. dos Portugueses, 1966 - Vila Bacanga, São Luís - MA, 65080-805

{ricardo.anchieta, francyles, ari, paiva}@nca.ufma.br

Abstract. *Thermography exams can be used as a non-invasive way for detecting diseases through the analysis of thermal images of the human body. The present work presents a method to detect nodules of the thyroid gland in thermal images using two cascade Convolutional Neural Networks (CNNs). The first CNN operates in the generation of candidate nodules, and the second CNN refines the detection, eliminating false positives. The experiments were carried out with thermal images of 20 patients. The final result reaches 97% accuracy, 95% recall and 80% f1-score using two AlexNet CNNs.*

Resumo. *Os exames de termografia podem ser utilizados como um meio não invasivo para a detecção de doenças através da análise de imagens térmicas do corpo humano. O presente trabalho apresenta um método para detectar nódulos da glândula tireoide em imagens térmicas utilizando duas Redes Neurais Convolucionais (CNN's) em cascata. A primeira CNN atua na geração de candidatos a nódulo, e a segunda faz o refinamento da detecção, eliminando falsos positivos. Os experimentos foram realizados com imagens térmicas de 20 pacientes. O resultado final atinge 97% de acurácia, 95% de sensibilidade e 80% de f1-score utilizando duas CNN's AlexNet.*

1. Introdução

Os nódulos da tireoide se tornaram um problema comum na área médica, representando a principal causa de várias doenças relacionadas a esse órgão, como o câncer, que, segundo estimativas, será o segundo mais comum entre as mulheres em 2030 [Rahib et al. 2014]. Dessa forma, a detecção precoce de tumores cancerígenos aumenta a chance de um tratamento bem-sucedido e melhora a taxa de sobrevivência ao câncer [McPhail et al. 2015].

O comportamento funcional da tireoide representa a principal informação no diagnóstico de nódulos malignos. A caracterização do nódulo permite ao especialista decidir se a biópsia de aspiração é necessária. Nesse processo, tamanho, posição e malignidade do nódulo são observados. Dois dos exames indicados para esta avaliação são a ultrassonografia cervical e ressonância magnética [Nguyen et al. 2015].

Fisiologicamente, as doenças da tireoide, incluindo nódulos, geralmente induzem uma temperatura elevada na superfície da pele do pescoço, que pode ser adotada como um índice diagnóstico potencial. Assim, a termografia poderia ser implementada como um método complementar para a detecção de nódulos, visto que a sua temperatura se

comporta de maneira diferente em relação às outras regiões do pescoço. Esse exame permite capturar a radiação infravermelha emitida por objetos e converte esta radiação em temperaturas quando as condições ambientais são bem conhecidas.

Em comparação a outros exames de imagem utilizados para a detecção de nódulos na tireoide, a termografia tem custo extremamente baixo [Brioschi et al. 2010]. Além disso, é um método rápido, sem contato e não invasivo e já tem sido empregado para o diagnóstico de tumor cerebral, câncer de pele, doenças oculares e neuropatia diabética [Bahramian and Mojra 2020].

Com o avanço dos recursos tecnológicos, e considerando a natureza digital das imagens térmicas, o desenvolvimento de métodos computacionais baseados em processamento de imagem pode auxiliar os especialistas na interpretação dos exames, garantindo mais eficiência e precisão aos diagnósticos e aprimorando a qualidade da assistência médica.

Na literatura, métodos baseados em técnicas de aprendizado profundo, como as Redes Neurais Convolucionais (CNN's), são amplamente usados para a classificação de problemas clínicos, alcançando resultados promissores no âmbito da classificação de nódulos, tumores, lesões e outras anormalidades em imagens médicas [Lévy and Jain 2016, Moran et al. 2018, Esteva et al. 2017].

Também são encontrados trabalhos relacionados especificamente aos exames de termografia da tireoide. Moran et al. (2018) analisaram o parâmetro de assimetria de temperatura entre os lados direito e esquerdo da tireoide para extrair regiões de nódulo nas imagens térmicas. Essas imagens foram utilizadas no treinamento de três redes neurais, com arquiteturas baseadas na GoogLeNet [Zhu et al. 2015], AlexNet [Krizhevsky et al. 2012] e VGG [Simonyan and Zisserman 2014]. A rede que obteve o melhor resultado na ocasião foi a GoogLeNet, com 86% de acurácia. Entretanto, esse método depende da imagem resultante do parâmetro de assimetria. Logo, se alguma região de nódulo for descartada no processo de geração da imagem de entrada da rede, o resultado final será afetado.

González et al. (2019) investigaram duas técnicas de análise de nódulos tireoidianos em imagens térmicas. São elas: o uso de uma única imagem do exame, verificando se existe um limiar de temperatura que diferencia os nódulos de regiões saudáveis do pescoço; e o uso de séries de temperatura, analisando o comportamento do nódulo ao longo do tempo. Os resultados mostraram que existem regiões do pescoço que possuem temperaturas similares às de nódulos, não tornando possível analisar a diferença de temperatura em ambos os casos analisados. Apesar de a análise proposta ser bastante pertinente, os autores não utilizam métodos de aprendizado para a detecção de regiões nodulares.

O trabalho de Moran et al. (2019) utiliza a rede convolucional ResNet [He et al. 2016] para classificar regiões nodulares. Os autores fazem um estudo comparativo entre técnicas responsáveis pela segmentação de regiões candidatas a nódulo como um processo anterior à classificação. Essas técnicas são: parâmetros estatísticos, índice termográfico, agrupamento de temperatura e parâmetro de assimetria. Nos experimentos, o melhor resultado alcançado pela ResNet foi 92% de acurácia usando como imagens de entrada aquelas obtidas pelo parâmetro de assimetria, sendo passível de erro quando este parâmetro descarta alguma região nodular antes da submissão de imagens à rede.

Dado esse contexto, e diferenciando-se das abordagens mencionadas, o presente trabalho propõe um método para detectar nódulos da tireoide em exames de termografia utilizando duas CNN's em cascata. Primeiro, são aplicadas técnicas nas imagens do exame de modo a construir uma série temporal de temperaturas e extrair janelas contendo todas as regiões do pescoço. Dessa forma, garante-se que o nódulos sejam mantidos, e busca-se analisar as características da temperatura ao longo do tempo. Em seguida, as janelas extraídas são submetidas à primeira CNN que é responsável por indicar candidatos a nódulo. Por fim, esses candidatos são direcionados à segunda CNN para o refinamento dos resultados com a redução de falsos positivos.

Este trabalho está dividido nas seguintes seções: na Seção 2, a base de imagens utilizada é detalhada e as etapas do método proposto são explicadas; a Seção 3 apresenta os resultados obtidos; E, na Seção 4, são expostas as conclusões e os trabalhos futuros.

2. Materiais e método

Esta seção apresenta a base de imagens utilizada no presente estudo, junto com o detalhamento do método proposto.

2.1. Base de imagens

A base de imagens utilizada neste trabalho, consiste em termografias de 20 pacientes que apresentam nódulos na região da tireoide. O protocolo utilizado para obter os exames mostra o comportamento da temperatura na superfície da região da tireoide após um estresse térmico. Esse estresse é causado pelo resfriamento da superfície do pescoço com um fluxo de ar de um ventilador. Durante o processo de resfriamento, a cada 15 segundos são capturadas as imagens térmicas. [Fiirst et al. 2018, Moran et al. 2018] . Cada exame termográfico tem 20 imagens de um mesmo paciente ao longo do tempo (Figura 1). Cada imagem da base possui um arquivo adicional indicando a marcação do especialista da localização do nódulo, e um arquivo com a matriz de temperatura. Essa matriz informa o valor da temperatura de cada pixel da imagem. É importante ressaltar que essas imagens fazem parte de um banco de dados fornecido pelo Visual Lab¹ da Universidade Federal Fluminense no âmbito do projeto *Thyroid* [Gonzalez 2017, Fiirst et al. 2019] .

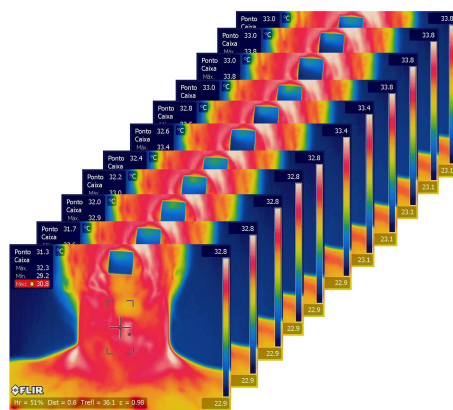


Figura 1. Resultado do exame de termografia.

¹Disponível em: <http://visual.ic.uff.br/thyroid>

2.2. Método Proposto

Neste artigo é proposto um método de detecção de nódulos da tireoide utilizando duas CNN's em cascata. As duas CNN's utilizadas estão associadas, como mostra a Figura 2.

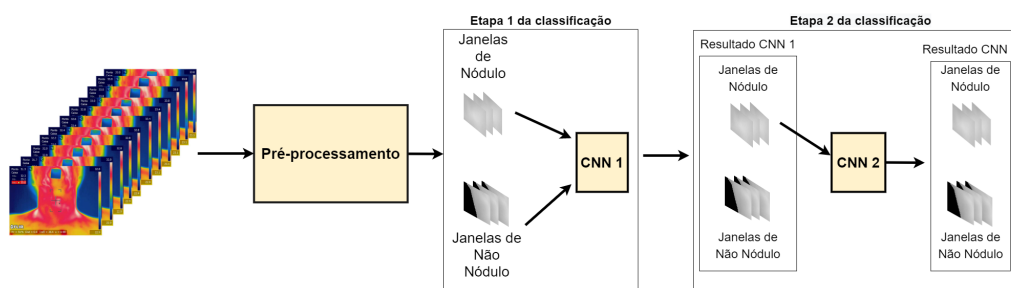


Figura 2. Demonstração do método de detecção em cascata proposto neste estudo.

Antes de serem submetidas para as CNN's, as imagens passam por cinco etapas de pré-processamento: transformação das imagens em níveis de cinza, registro, extração da região de interesse, melhoramento de contraste e janelamento. Depois dessas etapas, as imagens chegam até a CNN 1 como janelas. Nessa CNN serão geradas regiões candidatas a nódulo. Essas regiões serão passadas para a CNN 2 e então classificadas novamente, com o intuito de retirar possíveis falsos positivos e detectar definitivamente as regiões nodulares. A seguir, cada etapa do método proposto será explanada.

2.3. Pré-processamento

As matrizes de temperatura da base de imagens, contêm as temperaturas em graus Celsius de cada pixel das imagens térmicas. Para poder utilizar as informações das matrizes, é feito como primeira etapa do pré-processamento a transformação das matrizes em imagens em níveis de cinza, em um processo de normalização como mostra a Figura 3.

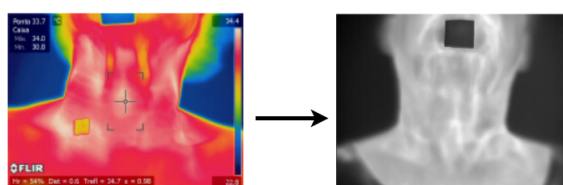


Figura 3. Transformação da matriz de temperatura em imagens em escala de cinza

Para o processo de normalização, utiliza-se a Equação 1, onde M é a matriz de temperatura, e (x,y) são as coordenadas.

$$N(x, y) = \frac{M(x, y) - \min(M)}{\max(M) - \min(M)} * 255 \quad (1)$$

Após a transformação em escala de cinza ser feita nas imagens, elas passam pelo processo de registro. Ao longo do exame de termografia, a movimentação do paciente pode influenciar a retirada correta das imagens. Quando isso ocorre, a região do nódulo pode mudar de posição ao longo das imagens do exame. Para corrigir esse erro, usa-se o

registro de sub-pixel [Guizar-Sicairos et al. 2008]. Essa técnica toma a primeira imagem como base e corrige as demais utilizando uma correlação cruzada entre a imagem base e a imagem a ser corrigida.

A tireoide está localizada no pescoço, então apenas essa região deve ser considerada na análise das imagens. Sendo assim, para extrair a região de interesse, é feito um recorte manual a ser utilizado em todas as imagens, que foram previamente registradas, de modo que nenhuma região nodular de nenhum paciente fosse perdida. No final desse processo, todas as imagens que antes tinham 640x480 terão 320x320 de dimensão. Em seguida é retirado o fundo dos recortes, deixando apenas a região do pescoço do paciente (Figura 4).



Figura 4. Processo de extração da região de interesse

Nas regiões de interesse já extraídas, são aplicadas duas técnicas. São elas: o filtro Gaussiano, para a suavização; e o *Contrast Limited Adaptive Histogram Equalization* (CLAHE). Esse último foi utilizado com o intuito de realçar as regiões nodulares através da equalização do histograma que produz um melhoramento no contraste. As imagens resultantes desta etapa serão utilizadas no processo da criação das séries temporais.

Como dito anteriormente, o nódulo geralmente apresenta uma temperatura mais elevada do que as demais regiões do pescoço. Baseado nisso, foi proposto um método para levar em consideração o comportamento da temperatura dos nódulos perante ao tempo através das séries temporais. A termografia de um paciente resulta em 20 imagens térmicas capturadas em tempos diferentes ao longo do exame. Neste trabalho, para construir a série temporal de um paciente é utilizada a primeira (T1), a décima (T10) e a vigésima (T20) imagem do exame de um paciente, essas foram escolhidas por representarem respectivamente o início, o meio e o fim do exame de termografia. As imagens selecionadas são concatenadas para construir uma única imagem com 3 dimensões, uma demonstração é mostrada na Figura 5.

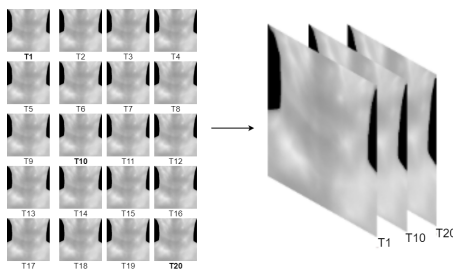


Figura 5. Demonstração da série temporal, à esquerda as 20 imagens de um paciente organizadas pelo tempo T em que foram tiradas, à direita a concatenação da primeira (T1), décima (T10) e vigésima (T20) imagem para a criação da série temporal

Para conseguir identificar as regiões nodulares nas imagens térmicas, as imagens foram divididas em janelas menores, esse processo de janelamento recebe a imagem resultante da série temporal de um paciente e a divide em janelas sendo suas dimensões de 64x64, o que resulta em 81 janelas por imagem. Dentre as janelas obtidas existirão as de regiões de nódulo e de não nódulo. Para uma janela ser considerada de uma região de nódulo, o mesmo precisa abranger mais que 50% da área da janela ou a janela precisa abranger mais que 50% da área do nódulo, caso ele seja menor que janela. No processo de janelamento as janelas extraídas se sobrepõem, como por exemplo a segunda janela começa 32 pixels após a primeira como demonstra a Figura 6. Esta é a última etapa do pré-processamento. As janelas geradas aqui serão utilizadas para as detecções de nódulos feitas pelas CNN's.

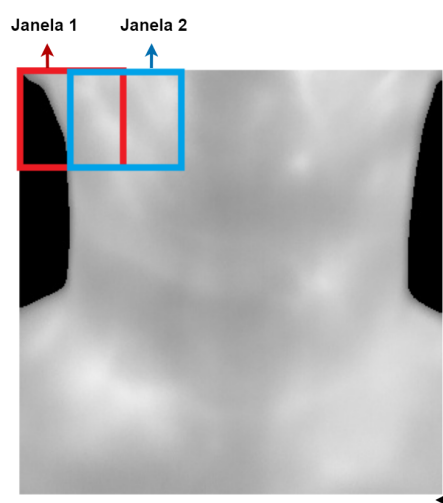


Figura 6. Demonstração da sobreposição no janelamento

2.4. Detecção usando CNNs em cascata

Neste trabalho, o método utilizado para a detecção se dá pela classificação das janelas em nódulo e não nódulo. Este método foi dividido em duas etapas, de modo que cada qual utiliza uma CNN.

Na etapa 1, para classificar primariamente as janelas, utiliza-se uma CNN com a arquitetura baseada na AlexNet. A AlexNet [Krizhevsky et al. 2012] é uma rede neural profundamente convolucional, e consiste em oito camadas: cinco camadas convolucionais, algumas das quais são seguidas por camadas de max-pooling e três camadas totalmente conectadas com a camada final. A camada final, para este trabalho, foi alterada para uma camada sigmóide, já que a classificação é binária, isto é, entre regiões de nódulo e não nódulo.

Na etapa 2 da detecção em cascata proposta, utiliza-se a CNN que melhor performou na primeira classificação, porém o treino desta nova CNN será com o todas as janelas de regiões de nódulo utilizadas no primeiro treino, adjunto a todos os falsos positivos gerados pelo treino da CNN 1. A CNN 2 será testada com as janelas que a CNN 1 classificou como sendo de nódulo, com a finalidade de diminuir a quantidade de falsos positivos gerados. Dessa forma, no resultado final, serão mantidas apenas as regiões nodulares.

3. Resultados e Discussão

Neste trabalho, os processos de pré-processamento e detecção foram implementados em linguagem Python 3.8, utilizando as bibliotecas OpenCV² versão 4.1.2, Scikit-learn³ versão 0.22.1, Keras⁴ versão 2.4.0 e Tensorflow⁵ versão 2.4.1. O hardware utilizado possui um processador AMD Ryzen 5 1600 de seis núcleos e 3.20 GHz, 16 gigabytes de memória RAM e uma placa de vídeo NVIDIA GeForce GTX 1050 Ti.

Nesse experimento, três CNN's foram testadas na etapa 1 da classificação, são elas: a AlexNet, a ResNet50 [He et al. 2016] e a VGG16 [Simonyan and Zisserman 2014]; e foram treinadas utilizando um tamanho do batch de 8, uma taxa de aprendizagem de 0,001 e 100 épocas. Na etapa 2 a CNN foi treinada utilizando um tamanho de batch de 4, taxa de aprendizagem de 0,001 e 50 épocas. Foi decidido usar os mesmos parâmetros para todas as três CNN's que foram testadas na etapa 1 a fim de verificar melhor o comportamento de cada CNN na resolução do problema, excluindo variações causadas pelo uso de parâmetros diferentes. Após o processo de janelamento foram obtidas 1620 janelas. 60% das imagens do conjunto de imagens foram usadas na fase de treinamento, totalizando 972 janelas, sendo 85 de regiões de nódulo e 887 de não nódulo. Os 40% restantes foram usados na fase de teste, que consiste em 648 janelas, sendo 37 de regiões de nódulo e 611 de não nódulo.

Todas as CNN's que foram testadas nas etapa 1 e 2 da classificação, foram treinadas usando duas funções de custo distintas: a *binary cross-entropy* e a *binary focal loss*, essa última utilizada pela sua eficiência quando se tem um conjunto de treino desbalanceado [Lin et al. 2018].

As métricas usadas para avaliar o resultado das CNN's são: acurácia (Acc), precisão (Prec.), sensibilidade (Sens.), f1-score (F1); essas são representadas respectivamente pelas equações 2,3,4,5, onde VP significa verdadeiro positivo, VN é verdadeiro negativo, FP é falso positivo e FN é falso negativo. Também foram usadas mais duas métricas importantes para o contexto da detecção. A primeira é o número de falsos positivos por imagem (FP/Img), que consiste na média entre o número de falsos positivos dividido pela quantidade de imagens. A segunda é o número de detecções (ND), que calcula o número de pacientes que tiveram nódulos detectados.

$$DP = \frac{VP + VN}{VP + VN + FP + FN} \quad (2)$$

$$Prec = \frac{VP}{VP + FP} \quad (3)$$

$$Sens = \frac{VP}{VP + FN} \quad (4)$$

$$F1 = 2 * \frac{Prec * Sens}{Prec + Sens} \quad (5)$$

²Disponível em: <https://opencv.org/>

³Disponível em: <https://scikit-learn.org/stable/>

⁴Disponível em: <https://keras.io/>

⁵Disponível em: <https://www.tensorflow.org/>

Como a proposta do trabalho trata de detecção, isto é, indicar em qual região o nódulo se encontra, as métricas de avaliação consideram que se as janelas classificadas como região de nódulo pelas CNN's preencherem no mínimo 50% da região nodular marcada pelo especialista, isso será considerado como um acerto, pois assim a maior parte do nódulo já terá sido detectada e a região que ele se encontra já estará explícita.

Nos resultados para a CNN 1, mostrados na Tabela 1, a AlexNet com a função de custo *binary cross-entropy* foi a que obteve o melhor resultado dentre as CNN's testadas. Mesmo a VGG16 apresentando um menor valor de falsos positivos por imagem, ela detecta nódulos em menos pacientes do que a AlexNet.

CNN	Função de custo	Acc	Prec.	Sens.	F1	FP/Img	ND
AlexNet	B. Cross-Entropy	0,96	0,60	0,95	0,74	2,875	8/8
ResNet50	B. Cross-Entropy	0,93	0,35	0,37	0,36	4,000	3/8
VGG16	B. Focal Loss	0,94	0,42	0,62	0,50	1,750	5/8

Tabela 1. Resultados das CNN's testadas na etapa 1 com a função de custo que obteve o melhor resultado em cada uma.

Também foram feitos testes para avaliar o uso da função de custo *binary focal loss* junto à AlexNet no contexto da CNN 1 da cascata. Contudo, os resultados não foram satisfatórios, visto que a CNN não conseguiu identificar regiões nodulares nos pacientes de teste como mostra o resultado das métricas na Tabela 2.

Função de custo	Acc	Prec.	Sens.	F1	FP/Img	ND
B. Cross-Entropy	0,96	0,60	0,95	0,74	2,875	8/8
B. Focal Loss	0,94	0,00	0,00	0,00	0,000	0/8

Tabela 2. Resultados da AlexNet da CNN 1 com funções de custo diferentes.

Como o melhor resultado da CNN 1 foi obtido pela AlexNet, ela também foi utilizada como a CNN 2 do método em cascata. Os resultados desse teste podem ser vistos na Tabela 3.

CNN	Função de custo	Acc	Prec.	Sens.	F1	FP/Img	ND
AlexNet	B. Cross-Entropy	0,97	0,72	0,70	0,74	1,250	6/8
AlexNet	B. Focal Loss	0,97	0,69	0,95	0,80	2,000	8/8

Tabela 3. Resultados após a CNN 2 com funções de custo diferentes.

A CNN 2 que foi utilizada com a finalidade de reduzir os falsos positivos gerados pelo resultado da CNN 1, se mostrou capaz de realizar a tarefa ao utilizar a *binary focal loss* como função de custo para o seu treinamento. Isso se reflete na redução do valor obtido para a métrica de falso positivo por imagem, bem como no aumento da precisão e do F1-score. Ressalta-se que a sensibilidade se manteve, o que indica que nenhum nódulo foi perdido no processo de detecção em cascata.

Das 648 janelas no conjunto de treinamento da CNN 1, apenas 50 são janelas de regiões nodulares. Isso revela um desbalanceamento notável nesse conjunto, o que afeta

diretamente o treinamento das CNN's da primeira etapa da detecção. Esse desbalanceamento persiste até a CNN 2 onde teremos 59 janelas no total para o treinamento e apenas 9 janelas de regiões não nodulares. Porém, como observado, o uso da *binary focal loss* foi importante para melhorar os resultados, corroborando sua aplicação no contexto de dados desbalanceados. O resultado final pode ser conferido na Tabela 4. A AlexNet se mostrou a CNN mais efetiva, dentre as testadas neste trabalho, para a classificação de janelas no âmbito da detecção de nódulos.

CNN	Função de custo	Acc	Prec.	Sens.	F1	FP/Img	ND
1 - AlexNet	B. Cross-Entropy	0,96	0,60	0,95	0,74	2,875	8/8
2 - AlexNet	B. Focal Loss	0,97	0,69	0,95	0,80	2,000	8/8

Tabela 4. Resultado final do método proposto.

Entre os testes, também foi realizada uma análise comparativa entre duas abordagens de aplicação do método proposto, uma usando na entrada uma série temporal com 3 imagens do paciente, conforme descrito na Seção 2.3, e a outra utilizando todas as 20 imagens capturadas durante o exame. Os resultados podem ser visualizados na Tabela 5. Com esse teste, verifica-se que mesmo possuindo mais informações por causa da adição de imagens, o método proposto produz mais falsos positivos em comparação a utilização de somente 3 imagens na série temporal, além de não detectar nódulos em todos os pacientes testados. Dessa forma, infere-se que o uso de somente 3 imagens para compor a série, representando os tempos T1, T10 e T20, é a abordagem mais promissora.

Série Temporal	Acc.	Prec.	Sens.	F1	FP/Img	ND
Com 3 imagens	0,97	0,69	0,95	0,80	2,000	8/8
Com 20 imagens	0,95	0,55	0,82	0,69	3,625	7/8

Tabela 5. Comparação do resultado final utilizando séries temporais de tamanhos diferentes.

Foram feitos estudos de caso para dois pacientes do conjunto de teste. Analisando o primeiro paciente, apresentado na Figura 8, é possível verificar ainda a presença de um falso positivo, mesmo após o uso da CNN 2. No caso apresentado, a região da clavícula, onde se encontra o falso positivo remanescente, apresenta uma temperatura mais elevada em comparação ao pescoço, similar a um nódulo. Essas regiões similares trazem uma desvantagem ao método proposto uma vez que elas influenciam a aparição de mais falsos positivos. Os falsos positivos podem levar a uma leitura errada do quadro de um paciente, tornando sua condição pior do que deveria ser.

No segundo paciente, mostrado na Figura 9, verifica-se que a CNN 2 não retira nenhuma detecção antes feita pela primeira CNN. Nesse caso, a CNN 1 já fornece o resultado ideal para a detecção, não gerando nenhum falso positivo, o resultado se mantém após a passagem pela CNN 2. Os resultados mostram a vantagem do uso da segunda CNN, uma vez que ela não atrapalha os resultados ideais obtidos, alterando somente os resultados que necessitam da redução de falsos positivos.

A partir dessa análise, é possível apontar que o método proposto apresenta resultados promissores. No entanto, acredita-se também que o acréscimo de novos pacientes à amostra pode contribuir para que os resultados sejam melhorados.

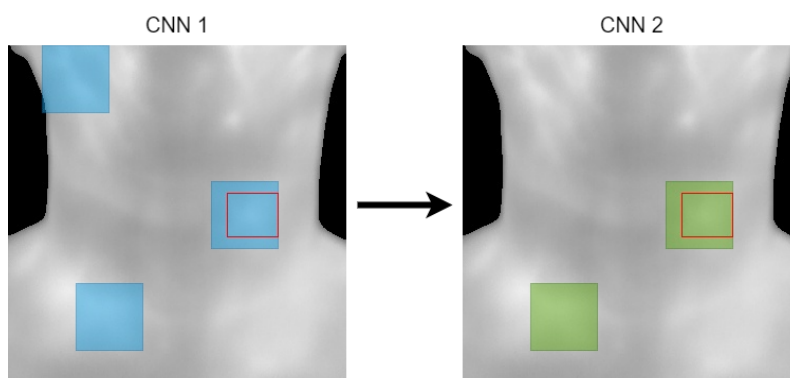


Figura 7. Resultados para o estudo de caso do primeiro paciente analisado. O nódulo está marcado em vermelho; à esquerda o resultado da CNN 1, com as janelas de nódulo previstas marcadas em azul. À direita, o resultado após a passagem da CNN 2, com as regiões previstas marcadas em verde.

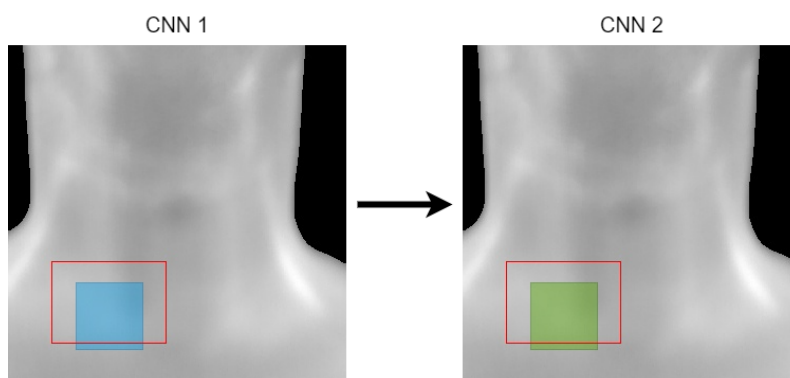


Figura 8. Resultados para o estudo de caso do segundo paciente analisado. O nódulo está marcado em vermelho; à esquerda o resultado da CNN 1, com as janelas de nódulo previstas pela CNN marcadas em azul. À direita o resultado após a passagem da CNN 2, com as regiões previstas marcadas em verde.

4. Conclusão

As termografias, por serem não invasivas e de baixo custo, se mostram uma boa alternativa de exame a ser usado para detectar nódulos na glândula tireóide. Utilizando o método proposto que usa CNN's em cascata foi possível detectar regiões nodulares, podendo colaborar com o especialista na obtenção de um diagnóstico mais preciso, e também com o processo de triagem médica para que a assistência seja prestada de forma célere.

O melhor resultado alcançado pelo método proposto se deu com o uso de duas CNN's AlexNet, obtendo a acurácia de 0,97. O método apresentou uma taxa de falsos positivos por imagem relativamente baixa (2,000 FP/img) em comparação com o uso de uma única CNN, que foi de 2,875 FP/img. O intuito da utilização de CNN's em cascata foi validado, uma vez que a CNN 2 reduziu os falsos positivos gerados pela CNN 1. Ressalta-se também que os nódulos foram detectados em todos os pacientes do conjunto de testes. Dessa forma, é possível apontar que arquiteturas mais simples foram as mais indicadas para a tarefa proposta neste trabalho, e isso é corroborado pelos resultados obtidos que são satisfatórios.

Pretende-se, como trabalho futuro, melhorar a fase de janelamento da série temporal, fazendo com que mais janelas sejam geradas sobrepondo as já extraídas, de modo que se tenha mais janelas para o treinamento das CNN's. Além disso, serão testadas outras arquiteturas de CNN, como a EfficientNet [Tan and Le 2019]; utilizando-se também a abordagem de *data augmentation* como forma de reduzir o desbalanceamento, buscando também uma maior redução de falsos positivos. Os autores agradecem á Fundação de Amparo à Pesquisa e ao Desenvolvimento Científico e Tecnológico do Maranhão (FAPEMA) pelo apoio financeiro.

Referências

- Bahramian, F. and Mojra, A. (2020). Thyroid cancer estimation using infrared thermography data. *Infrared Physics & Technology*, 104:103126.
- Brioschi, M., Teixeira, M., Silva, F., and Colman, D. (2010). Medical thermography textbook: principles and applications. *Editora e Livraria Andreoli*.
- Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., and Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *nature*, 542(7639):115–118.
- Fiirst, W., Montenegro, A., Machenry, T., Li, H., Du, C., and Conci, A. (2018). Introducing tensor analysis for biomedical image registration. *6o. Encontro Nacional de Engenharia Biomecânica (ENEBI 2018), Aguas de Lindoia, SP (proc. ABCM, SBB)*.
- Fiirst, W. G., Resmini, R., and Conci, A. (2019). Desenvolvimento e comparação de métodos para geração de séries temporais para imagens térmicas da tireoide. In *Anais Estendidos do XIX Simpósio Brasileiro de Computação Aplicada à Saúde*, pages 43–48. SBC.
- Gonzalez, J. (2017). Um estudo sobre a possibilidade do uso de imagens infravermelhas na análise de nódulos de tireoide. *Dissertation, IC, Universidade Federal Fluminense*.
- González, J., Conci, A., Moran, M., Araujo, A., Paes, A., Damião, C., and Fiirst, W. (2019). Analysis of static and dynamic infrared images for thyroid nodules investigation. In *2019 IEEE/ACS 16th International Conference on Computer Systems and Applications (AICCSA)*, pages 1–7. IEEE.
- Guizar-Sicairos, M., Thurman, S., and Fienup, J. (2008). Efficient subpixel image registration algorithms. *Optics letters*, 33 2:156–8.
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778.
- Hernandez Moran, M., Conci, A., and Araújo, A. (2019). Evaluation of quantitative features and convolutional neural networks for nodule identification in thyroid thermographies. pages 747–751.
- Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25:1097–1105.
- Lévy, D. and Jain, A. (2016). Breast mass classification from mammograms using deep convolutional neural networks. *arXiv preprint arXiv:1612.00542*.

- Lin, T.-Y., Goyal, P., Girshick, R., He, K., and Dollár, P. (2018). Focal loss for dense object detection.
- McPhail, S., Johnson, S., Greenberg, D., Peake, M., and Rous, B. (2015). Stage at diagnosis and early mortality from cancer in england. *British journal of cancer*, 112(1):S108–S115.
- Moran, M. B., Conci, A., González, J. R., Araújo, A. S., Fiirst, W., Damião, C. P., Lima, G. A., and da Cruz Filho, R. A. (2018). Identification of thyroid nodules in infrared images by convolutional neural networks. In *2018 International joint conference on neural networks (IJCNN)*, pages 1–7. IEEE.
- Nguyen, Q. T., Lee, E. J., Huang, M. G., Park, Y. I., Khullar, A., and Plodkowski, R. A. (2015). Diagnosis and treatment of patients with thyroid cancer. *American health & drug benefits*, 8(1):30.
- Rahib, L., Smith, B. D., Aizenberg, R., Rosenzweig, A. B., Fleshman, J. M., and Matriasian, L. M. (2014). Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the united states. *Cancer research*, 74(11):2913–2921.
- Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Tan, M. and Le, Q. V. (2019). Efficientnet: Rethinking model scaling for convolutional neural networks. *arXiv preprint arXiv:1905.11946*.
- Zhu, R., Zhang, R., and Xue, D. (2015). Lesion detection of endoscopy images based on convolutional neural network features. In *2015 8th International Congress on Image and Signal Processing (CISP)*, pages 372–376. IEEE.