

UNIVERSIDADE FEDERAL FLUMINENSE

MAIRA BEATRIZ HERNANDEZ MORAN

**METODOLOGIA PARA LOCALIZAÇÃO DE  
NÓDULOS MALIGNOS DA TIREOIDE A PARTIR  
DE IMAGENS INFRAVERMELHAS**

NITERÓI

2018

UNIVERSIDADE FEDERAL FLUMINENSE

MAIRA BEATRIZ HERNANDEZ MORAN

**METODOLOGIA PARA LOCALIZAÇÃO DE  
NÓDULOS MALIGNOS DA TIREOIDE A PARTIR  
DE IMAGENS INFRAVERMELHAS**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Computação da Universidade Federal Fluminense como requisito parcial para a obtenção do Grau de Mestre em Computação. Área de concentração: COMPUTAÇÃO VISUAL.

Orientador:  
AURA CONCI

NITERÓI

2018

Ficha catalográfica automática - SDC/BEE

M829m Moran, Maira Beatriz Hernandez  
Metodologia Para Localização de Nódulos Malignos da  
Tireoide a partir de Imagens Infravermelhas / Maira Beatriz  
Hernandez Moran ; Aura Conci, orientadora. Niterói, 2018.  
117 f. : il.

Dissertação (mestrado)-Universidade Federal Fluminense,  
Niterói, 2018.

DOI: <http://dx.doi.org/10.22409/PGC.2018.m.15418083716>

1. Termografias estáticas. 2. Processamento de imagens. 3.  
Classificação. 4. Produção intelectual. I. Título II.  
Conci,Aura, orientadora. III. Universidade Federal Fluminense.  
Escola de Engenharia.

CDD -

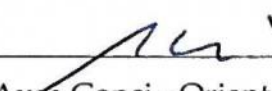
MAIRA BEATRIZ HERNANDEZ MORAN

METODOLOGIA PARA LOCALIZAÇÃO DE NÓDULOS MALIGNOS DA  
TIREOIDE A PARTIR DE IMAGENS INFRAVERMELHAS

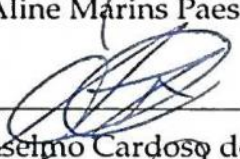
Dissertação de Mestrado apresentada ao  
Programa de Pós-Graduação em Computação  
da Universidade Federal Fluminense como  
requisito parcial para a obtenção do Grau  
de Mestre em Computação. Área de  
concentração: COMPUTAÇÃO VISUAL.

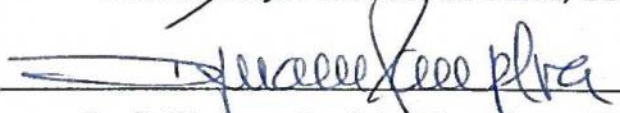
Aprovada em 23 de julho de 2018.

BANCA EXAMINADORA

  
Prof.ª Aura Conci - Orientador, UFF

  
Prof.ª Aline Marins Paes Carvalho, UFF

  
Prof. Anselmo Cardoso de Paiva, UFMA

  
Prof.ª Djenane Cordeiro Pamplona, PUC-Rio

Niterói

2018



*Dedicatória: A Deus, minha família, amigos e a todos os que me ajudaram de alguma  
forma nesta jornada.*

# Agradecimentos

A Deus que me guiou ao longo de todo este caminho.

A minha família: meus pais Luis e Maira e minha irmã Isabele, que me apoiaram e incentivaram em cada momento, especialmente nos que mais precisei. Obrigada por tudo o que me ensinaram.

Aos professores do Laboratório de Ciências Matemáticas da Universidade Estadual do Norte Fluminense, por contribuírem no meu desenvolvimento acadêmico e intelectual.

A professora Aura Conci, minha orientadora e amiga, por me auxiliar no desenvolvimento deste trabalho. Aos docentes do Instituto de Computação (IC) da Universidade Federal Fluminense, que me permitiram aprofundar meus conhecimentos para assim desenvolver este trabalho e crescer cientificamente. Aos secretários da pós-graduação do IC, Teresa e Hélio, por sua colaboração nos mais diversos momentos.

Aos colegas do VisualLab, por toda parceria nesta jornada. Aos amigos e colegas do IC, pelos momentos de alegria e pelas conversas, obrigada.

# Resumo

Anomalias na região da tireoide apresentam alta prevalência, afetando parte considerável da população. Parte dessas anomalias estão relacionada a nódulos, predominando as que apresentam características benignas. Entretanto, a incidência de nódulos malignos também é relevante, o que representa um risco à saúde do paciente, sendo sua identificação precoce essencial para um tratamento mais eficaz. Um dos meios mais comuns de detecção de nódulos é a palpação. Entretanto para melhor diagnosticar disfunções na tireoide é necessário observar seu comportamento. Diversos exames são usados atualmente no processo de diagnóstico, porém alguns deles são de difícil acesso à população ou expõe o paciente a radiação ionizante. Uma alternativa é o uso de termogramas originários de emissões infra-vermelhas, baseando-se no fato de que nódulos, em especial os que apresentam alto risco de malignidade, tendem a ser mais vascularizados, o que pode levar a um aumento de temperatura na região. Na literatura médica, diversos métodos de análise de termografias para diagnóstico já foram propostos. Neste trabalho é apresentada uma proposta de metodologia para localizar nódulos na região da tireoide, com duas fases : segmentação e classificação, utilizando técnicas de processamento de imagens e de reconhecimento de padrões respectivamente. Para isso, é realizado um estudo de possíveis parâmetros baseados em termografias visando a identificação de nódulos. A partir disso, é desenvolvido um processo de segmentação a fim encontrar regiões possivelmente nodulares. Por fim, redes neurais convolucionais são usadas para a classificação dessas regiões, buscando identificar quais delas referem-se a nódulos. Considerando os resultados da metodologia proposta conclui-se que, devido ao bom desempenho da rede CNN na fase de classificação (que obteve uma acurácia de 96%), a viabilidade da metodologia proposta depende do processo de segmentação de regiões possivelmente nodulares.

**Palavras-chave:** termografia infravermelha estática, tireoide, nódulos, classificação.

# Abstract

Anomalies in the thyroid region have high prevalence and affect a large part of the population. Most of these anomalies are nodule related and commonly benign. However, the incidence of malignant nodules is also considerable, it presents a major health risk to the patient and, its early identification is essential for a more effective treatment. In general, it is possible to identify dysfunctions in the thyroid by observing its behavior. Although, palpation is also commonly employed in the detection of nodules, several exams are currently used in the diagnostic process, and some of them are restricted to part of the population due their high costs or because they can expose the patient to ionizing radiation. An alternative is the use of infrared images, their applicability is based on the fact that nodules, especially those with a high risk of malignancy, tend to be more vascularized, which can lead to an increase in temperature in such regions. Such phenomenon can be detected through thermographies. In medical literature, several methods of diagnosis using thermographies have been proposed. In this work a methodology for the detection of nodules in the thyroid region is proposed, considering two main phases: segmentation and classification, using image processing and pattern recognition techniques. For this, firstly parameters based on thermograms for the identification of nodules are evaluated, then a segmentation process is developed in order to find possible nodular regions. Finally, convolutional neural networks are used in the classification of these regions, in order to identify which of them refer to nodules. The results of the proposed methodology suggest that its feasibility is mainly dependent on the segmentation process of possibly nodular regions, since the CNN networks have a good performance ( 96% of accuracy), in the classification phase.

**Keywords:** infrared termography, thyreoid, nodules, classification.

# Lista de Figuras

2.1	Avaliação de um paciente portador de nódulos tireoidianos . . . . .	7
3.1	Termografia usada no estudo [Mahajan and Madhe, 2014] . . . . .	12
3.2	Segmentação de lesão em termografia da tireoide [Mahajan and Madhe, 2014]	12
3.3	Página de entrada do DTNR . . . . .	16
3.4	Estrutura da rede ResNet proposta em [He et al., 2016]. . . . .	26
3.5	Estrutura da rede GoogLeNet proposta em [Szegedy et al., 2015]. . . . .	27
3.6	Estrutura de um módulo inception proposta em [Szegedy et al., 2015]. . . .	28
4.1	Termografia da paciente 2017-10-30 . . . . .	31
4.2	Termografia normalizada da paciente 2017-10-30 . . . . .	32
4.3	ROI da paciente 2017-10-30 . . . . .	33
4.4	Média de cada parâmetro para cada grupo de pacientes . . . . .	35
4.5	Média do Índice Termográfico para cada grupo de pacientes . . . . .	38
4.6	ROI do paciente 2017-09-04 . . . . .	40
4.7	Quantização em 4 tons da termografia da Figura 4.6 . . . . .	41
4.8	Quantização em 8 tons da termografia da Figura 4.6 . . . . .	41
4.9	Quantização em 16 tons da termografia da Figura 4.6 . . . . .	42
4.10	Eixos e planos anatômicos . . . . .	42
4.11	Imagem $C_{masc}$ e termografia original com marcador . . . . .	44
4.12	Metodologia proposta . . . . .	46
4.13	Fase de pré-processamento . . . . .	46
4.14	Fase de segmentação . . . . .	47
4.15	Fase de classificação . . . . .	47

5.1	Segmentação do paciente 2017-09-04 com $P_A = 0,31$ . . . . .	52
5.2	Marcação dos nódulos do paciente 2017-09-04 . . . . .	52
5.3	Segmentação do paciente 2017-09-12-3 com $\ell = 0,31$ . . . . .	53
5.4	Marcação dos nódulos do paciente 2017-09-12-3 . . . . .	53
5.5	Segmentação do paciente 2017-10-03-3 com $\ell = 0,31$ . . . . .	54
5.6	Marcação dos nódulos do paciente 2017-10-03-3 . . . . .	54
5.7	Segmentação do paciente 2017-10-10-2 com $\ell = 0,31$ . . . . .	55
5.8	Marcação dos nódulos do paciente 2017-10-10-2 . . . . .	55
5.9	Segmentação do paciente 2017-10-10-3 com $\ell = 0,31$ . . . . .	56
5.10	Marcação dos nódulos do paciente 2017-10-10-3 . . . . .	56
5.11	Segmentação do paciente 2017-10-10-1 com $\ell = 0,31$ . . . . .	57
5.12	Marcação dos nódulos do paciente 2017-10-10-1 . . . . .	57
5.13	Segmentação do paciente 2017-09-11-2 com $\ell = 0,55$ . . . . .	58
5.14	Marcação dos nódulos do paciente 2017-09-11-2 . . . . .	58
5.15	Segmentação do paciente 2017-10-10-1 com $\ell = 0,78$ . . . . .	59
5.16	Segmentação do paciente 2017-09-11-1 com $\ell = 1,02$ . . . . .	59
5.17	Marcação dos nódulos da paciente 2017-09-11-1 . . . . .	60
5.18	Segmentação do paciente 2017-10-10-1 com $\ell = 1,02$ . . . . .	60
5.19	Segmentação do paciente 2017-10-10-2 com $\ell = 1,25$ . . . . .	61
5.20	Bounding boxes em torno de possíveis nódulos considerando $P_A$ . . . . .	61
5.21	Processo de espelhamento em torno de um eixo vertical . . . . .	63
5.22	Imagem original e resultantes e seus histogramas. . . . .	63
5.23	Processo de criação de imagens sintéticas a partir de uma imagem de S1. . . . .	64
5.24	Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendi- zado 0,1, com DT1 . . . . .	67
5.25	Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendi- zado 0,01, com DT1 . . . . .	67

5.26	Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendizado 0,001, com DT1 . . . . .	68
5.27	Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendizado 0,1, com DT2 . . . . .	69
5.28	Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendizado 0,01, com DT2 . . . . .	69
5.29	Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendizado 0,001, com DT2 . . . . .	70
5.30	Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendizado 0,1, com DT3 . . . . .	71
5.31	Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendizado 0,01, com DT3 . . . . .	71
5.32	Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendizado 0,001, com DT3 . . . . .	72
5.33	Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendizado 0,1, com DT4 . . . . .	73
5.34	Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendizado 0,01, com DT4 . . . . .	73
5.35	Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendizado 0,001, com DT4 . . . . .	74
5.36	Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendizado 0,1, com DT5 . . . . .	75
5.37	Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendizado 0,01, com DT5 . . . . .	75
5.38	Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendizado 0,001, com DT5 . . . . .	76

# Lista de Tabelas

2.1	Classificação de Bethesda . . . . .	9
4.1	p-valores para os parâmetros estatísticos dos grupos de controle e com nódulos malignos . . . . .	36
4.2	p-valores para os parâmetros estatísticos dos grupos de controle e com nódulos benignos . . . . .	36
4.3	p-valores para os parâmetros estatísticos dos grupos com nódulos benignos e com nódulos malignos . . . . .	36
4.4	TDEs para os parâmetros estatísticos dos grupos de controle e com nódulos malignos . . . . .	36
4.5	TDEs para os parâmetros estatísticos dos grupos de controle e com nódulos benignos . . . . .	36
4.6	TDEs para os parâmetros estatísticos dos grupos com nódulos benignos e com nódulos malignos . . . . .	36
4.7	p-valores para os testes Wilcoxon para cada paciente do grupo maligno, considerando $P_A$ . . . . .	45
5.1	Número de nódulos por paciente no grupo maligno . . . . .	51
5.2	Valores usados como limiar $\ell$ . . . . .	51
5.3	Número de imagens para cada conjunto N . . . . .	64
5.4	Número de imagens para cada conjunto S . . . . .	65
5.5	Combinações que formam os DT usados como entrada dos CNNs . . . . .	65
5.6	Acurácia, acurácia de validação, perda e perda de validação para o experi- mento com DT1 . . . . .	66
5.7	Acurácia, acurácia de validação, perda e perda de validação para o experi- mento com DT2 . . . . .	68



5.8	Accuracy, validation accuracy, loss e validation loss para o experimento com DT3 . . . . .	70
5.9	Acurácia, acurácia de validação, perda e perda de validação para o experimento com DT4 . . . . .	72
5.10	Acurácia, acurácia de validação, perda e perda de validação para o experimento com DT5 . . . . .	74
5.11	Maiores valores de acurácia alcançados nos experimentos CNN . . . . .	76
A.1	Acúracia por época, taxa de aprendizado 0,1 . . . . .	87
A.2	Acúracia por época, taxa de aprendizado 0,01 . . . . .	87
A.3	Acúracia por época, taxa de aprendizado 0,001 . . . . .	88
A.4	Acúracia por época, taxa de aprendizado 0,1 . . . . .	89
A.5	Acúracia por época, taxa de aprendizado 0,01 . . . . .	90
A.6	Acúracia por época, taxa de aprendizado 0,001 . . . . .	91
A.7	Acúracia por época, taxa de aprendizado 0,1 . . . . .	92
A.8	Acúracia por época, taxa de aprendizado 0,01 . . . . .	93
A.9	Acúracia por época, taxa de aprendizado 0,001 . . . . .	93
A.10	Acúracia por época, taxa de aprendizado 0,1 . . . . .	94
A.11	Acúracia por época, taxa de aprendizado 0,01 . . . . .	95
A.12	Acúracia por época, taxa de aprendizado 0,001 . . . . .	96
A.13	Acúracia por época, taxa de aprendizado 0,1 . . . . .	97
A.14	Acúracia por época, taxa de aprendizado 0,01 . . . . .	98
A.15	Acúracia por época, taxa de aprendizado 0,001 . . . . .	99

# Lista de Abreviaturas e Siglas

FLIR	: Forward Looking Infrared;
$H_0$	: Hipótese nula;
$H_A$	: Hipótese alternativa;
HUAP	: Hospital Universitário Antônio Pedro;
IC	: Instituto de Computação;
I	: Imagem termográfica em cores;
M	: Mapa de temperaturas;
$max_{ROI}$	: Temperatura máxima da ROI;
$media_{ROI}$	: Temperatura média da ROI;
$mediana_{ROI}$	: Temperatura mediana da ROI;
$min_{ROI}$	: Temperatura mínima da ROI;
MLP	: Multilayer perceptron;
N	: Imagem termográfica normalizada em tons de cinza
$P_A$	: Parâmetro de assimetria;
PAAF	: Punção aspirativa por agulha fina;
RM	: Ressonância magnética;
$ROI$	: Imagem ROI normalizada em tons de cinza;
$ROI_p$	: Conjunto de pontos da ROI com valor diferente de 0;
T3	: Triiodotironina total;
T4	: Tiroxina total ou livre;
TI	: Índice termográfico;
TID	: Termografias Infravermelhas Dinâmicas;
TIE	: Termografias Infravermelhas Estáticas;
TSH	: Nível sérico de tirotropina;
TM	: Tomografia computadorizada;
UFF	: Universidade Federal Fluminense;
US	: Ultrassonografia;

# Conteúdo

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Motivação . . . . .	3
1.2	Objetivos . . . . .	3
1.3	Organização do trabalho . . . . .	3
<b>2</b>	<b>Área de interesse</b>	<b>5</b>
2.1	Tireoide . . . . .	5
2.1.1	Diagnóstico de nódulos e outras anomalias . . . . .	5
2.1.1.1	Verificação do TSH . . . . .	7
2.1.1.2	Ultrassonografia . . . . .	7
2.1.1.3	Outros exames por imagem . . . . .	8
2.1.1.4	PAAF . . . . .	8
<b>3</b>	<b>Revisão Bibliográfica</b>	<b>10</b>
3.1	Termografias . . . . .	10
3.1.1	Trabalhos relacionados . . . . .	11
3.1.2	Trabalhos do grupo VisualLab . . . . .	14
3.1.2.1	González, 2017 . . . . .	14
3.1.2.2	Fiirst, 2018 . . . . .	16
3.1.3	Testes estatísticos . . . . .	17
3.1.4	Teste Wilcoxon-Mann-Whitney . . . . .	18
3.1.5	Tamanhos de efeito . . . . .	19

3.2	Redes Neurais Convolucionais . . . . .	20
3.2.1	Modelo Residuais . . . . .	22
3.2.2	Modelo GoogLeNet . . . . .	22
3.2.3	CNNs em problemas de classificações de imagens médicas . . . . .	23
<b>4</b>	<b>Análise termográfica</b>	<b>29</b>
4.1	Representação de dados utilizada . . . . .	30
4.1.1	Imagem termográfica e mapa de temperatura . . . . .	30
4.1.2	Imagem termográfica normalizada . . . . .	31
4.1.3	ROI . . . . .	32
4.1.4	Grupo de pacientes estudado . . . . .	33
4.1.5	Análise de parâmetros estatísticos simples . . . . .	34
4.2	Índice termográfico . . . . .	37
4.3	Parâmetros de segmentação . . . . .	39
4.3.1	Agrupamento por temperatura . . . . .	39
4.3.2	Parâmetro de assimetria . . . . .	41
4.4	Metodologia proposta . . . . .	46
<b>5</b>	<b>Resultados</b>	<b>50</b>
5.1	Resultados da segmentação usando $P_A$ . . . . .	50
5.2	Resultados da classificação . . . . .	60
5.2.1	Pré-processamento das entradas . . . . .	60
5.2.2	CNNs analisadas . . . . .	62
5.2.3	Preparação dos algoritmos CNNs . . . . .	62
5.2.4	Criação de dados sintéticos . . . . .	62
5.2.5	Configuração das CNNs . . . . .	65
5.2.6	Experimentos com CNNs . . . . .	66

<b>6</b>	<b>Conclusões</b>	<b>77</b>
6.1	Trabalhos futuros . . . . .	78
	<b>Referências</b>	<b>80</b>
	<b>Apêndice A - Acurácia por CNN</b>	<b>86</b>
A.1	Resultados para o conjunto de dados DT1 . . . . .	86
A.2	Resultados para o conjunto de dados DT2 . . . . .	88
A.3	Resultados para o conjunto de dados DT3 . . . . .	91
A.4	Resultados para o conjunto de dados DT4 . . . . .	94
A.5	Resultados para o conjunto de dados DT5 . . . . .	96
	<b>Apêndice B - Protocolo de aquisição</b>	<b>100</b>

# Capítulo 1

## Introdução

Os nódulos da tireoide podem ser considerados problemas comuns na área clínica, consistindo na principal manifestação clínica de várias doenças relacionadas a tireoide. Estudos indicam que ocorrem geralmente de forma benigna, atingindo prevalentemente mulheres e população de idade mais avançada [Dean and Gharib, 2008]. Análise epidemiológicas realizadas recentemente no Brasil indicam que dentre os pacientes que apresentaram nódulos na tireoide, 93% eram mulheres e de idade média de 51 anos [Oliveira Filho, 2017].

O comportamento funcional da tireoide é essencial na maioria das doenças deste órgão e representa a base para o diagnóstico delas. Hipertireoidismo, hipotireoidismo e manifestações cancerígenas são algumas das anomalias que se caracterizam pelo mau funcionamento da tireoide. Assim como para outras doenças, o diagnóstico precoce permite um tratamento eficaz, especialmente nos casos de câncer..

A avaliação de nódulos tireoidianos é fundamental no processo de diagnóstico, especialmente se forem malignos [Bomeli et al., 2010]. Muitas vezes, os nódulos dessa região podem ser encontrados por exames físicos, através de palpação, devido a superficialidade da posição dessas estruturas e do órgão em si. É possível também que esses nódulos sejam achados através de exames de imagem.

A caracterização do nódulo permite ao especialista decidir a conveniência de uma intervenção cirúrgica no paciente. Para isso, são observados o tamanho, a posição e a malignidade do nódulo. Alguns dos exames indicados para essa avaliação são: a medição do nível sérico de tirotropina(TSH), que consiste em exame laboratorial, a ultrasonografia cervical (US) , a tomografia computadorizada (TC) e a ressonância magnética (RM), que são exames de imagem [Nguyen et al., 2015].

Os exames de imagem são uma ferramenta importante no processo de diagnóstico,

uma vez que permitem avaliar de forma mais precisa as características dos nódulos. No entanto, a TC pode expor o paciente a radiação ionizante, o que é um potencial risco para a saúde [Brenner and Hall, 2007, Smith-Bindman et al., 2009]. Além disso, o RM é caro e a população em geral não tem acesso fácil a ele [Goldbach-Mansky et al., 2003].

Neste contexto, a termografia é uma alternativa interessante no diagnóstico de doenças da tireoide, uma vez que tem um custo extremamente baixo quando comparado aos outros exames de imagem, além de não oferecer nenhum efeito colateral [Brioschi et al., 2010]. Estudos indicam que as termografias, imagens geradas a partir das medições das emissões de frequências na faixa do infravermelho dos corpos, podem ser utilizadas como exame complementar para a detecção de anormalidades em várias regiões do corpo. Isso ocorre porque, de forma genérica, lesões geralmente apresentam aumento no fluxo sanguíneo, o que consequentemente pode alterar a temperatura cutânea na região, especialmente nas lesões mais superficiais [Brioschi et al., 2010]. Esse aumento de temperatura pode ser identificado pela variação de energia eletromagnética emitida na faixa de frequência infravermelha e visualizado na imagem termográfica. Isto porque essa imagem pode ser associada às temperaturas dos pontos de uma cena capturada por sensores existentes em câmeras sensíveis a essa faixa do espectro eletromagnético, a partir do uso de programas e parâmetros adequados.

A Associação Médica Americana considera a termografia infravermelha como método diagnóstico desde 1987 [Brioschi et al., 2010]. Apesar do uso de termografias para identificação de nódulos tireoidianos ter sido proposta em trabalhos anteriores, nota-se uma escassez de propostas metodológicas práticas, predominando na literatura estudos de caso e observações qualitativas, baseadas inclusive em características visualmente identificáveis nas imagens. Isso demonstra o potencial do uso das termografias, mas não especifica um meio cientificamente válido de utilizá-las para casos gerais.

Além de anomalias relacionadas à tireoide, as termografias são utilizadas no processo de diagnóstico de outras doenças, nos quais empregam-se, na maioria das vezes, formas qualitativas de diagnósticos, e mais raramente, abordagens baseadas em análises de dados (i.e. formas quantitativas) extraídos das termografias. Dentre os parâmetros quantitativos mais comumente usados pode-se citar: mediana, temperatura máxima, média e mínima de regiões; média de distribuição da temperatura; parâmetro de assimetria contralateral; e avaliação dinâmica utilizando séries temporais [Brioschi et al., 2010].

Dessa forma, em uma primeira abordagem de uso de termografia no diagnóstico, pensa-se analisar essas características, já utilizadas em outros trabalhos e consideradas

como promotoras de um diagnóstico adequado, para avaliar quais delas permitem a identificação e segmentação de áreas que podem indicar nódulos, e como essas áreas diferem das consideradas saudáveis.

## 1.1 Motivação

Considerando que grande parte dos nódulos são inicialmente identificados em exames físicos, o que pode levar a um diagnóstico tardio, a proposta deste trabalho é a utilização das termografias como um exame complementar a ser utilizado na rotina de avaliação dos pacientes, de forma a identificar mais precocemente possíveis nódulos na região da tireoide. Partindo de um conhecimento prévio e seguro da não gravidade de casos, a demanda de exames menos acessíveis pode ser diminuída, o que leva também a uma menor exposição dos pacientes aos efeitos colaterais que esses exames oferecem, já que seriam identificados previamente os casos de menor gravidade para os quais a utilização de tais exames não seria necessária. A termografia poderia ser ainda uma opção especialmente interessante para pacientes em que os exames tradicionais não são indicados, como crianças pequenas e mulheres grávidas.

## 1.2 Objetivos

Este trabalho tem como objetivo principal o desenvolvimento de uma metodologia para a detecção de nódulos na região da tireoide, utilizando técnicas de processamento de imagens e de reconhecimento de padrões, especificamente segmentação e algoritmos de redes neurais convolucionais, em imagens infravermelhas. Os seguintes objetivos secundários podem ser citados: avaliação da viabilidade da utilização de parâmetros baseadas em termografias para a identificação de nódulos; estudo de características presentes nos termografias que possibilitem a identificação de regiões nodulares; aplicação de técnicas de processamento de imagens para segmentação de regiões possivelmente nodulares; e reconhecimento de nódulos dentre regiões extraídas das termografias.

## 1.3 Organização do trabalho

No capítulo 2, conceitos aplicados neste trabalho são revisados. Na seção 2.1 são descritos conceitos básicos relacionados à tireoide, como a anatomia e anomalias dessa região, bem como o processo de diagnóstico dessas anomalias.



O capítulo 3 mostra uma revisão da literatura das áreas que abrangem este trabalho. A seção 3.1 faz uma breve revisão do uso de termografias na área médica, e comenta alguns trabalhos relacionados. Na seção 3.1.4 é descrito o teste estatístico Wilcoxon, usado na avaliação de parâmetros de tremografias. Na seção 5.2, conceitos básicos e aplicações médicas de redes neurais convolucionais são apresentados.

No capítulo 4, é realizado um estudo dos parâmetros usados na análise de termografias, avaliando a viabilidade do uso de cada um deles no problema aqui tratado. Na seção 4.4 a metodologia proposta é apresentada. O capítulo 5 mostra os resultados da metodologia proposta, considerando diferentes variações de suas fases. Por fim, no capítulo 6 a conclusão e trabalhos futuros são apresentados

# Capítulo 2

## Área de interesse

Este capítulo apresenta uma breve revisão da área clínica de interesse deste estudo. Ele trata de uma visão mais relacionada à área médica, analisando brevemente a região da tireoide, alvo deste estudo. Aqui, são mostradas características relacionadas à anatomia desse órgão, disfunções que o atingem, especialmente nódulos, e quais as técnicas mais utilizadas no processo de diagnóstico dessas disfunções.

### 2.1 Tireoide

A tireoide é uma glândula formada por dois lobos, direito e esquerdo, e um istmo de tecido glandular que os une. Ela está localizada na parte inferior do pescoço, sobre a região da traqueia. Possui formato semelhante ao de uma borboleta. Em adultos, normalmente tem um peso que varia entre 15 e 20 gramas.

A tireoide é responsável pela produção de diversos hormônios, entre eles a tiroxina (T4) e a trii- odotironina (T3). Esses hormônios atuam em processos metabólicos importantes, além de aumentar o consumo de oxigênio em quase todos os tecidos do corpo [Wallace et al., 2003]. Os processos influenciados pelos hormônios produzidos pela tireoide são: crescimento; funcionamento do cérebro, especialmente memória; desenvolvimento neural; desenvolvimento ósseo; entre outros [Larson, 1999].

#### 2.1.1 Diagnóstico de nódulos e outras anomalias

Na área da clínica médica, o nódulo tireoidiano pode ser considerado uma manifestação comum. Entretanto, a retirada do nódulo só é recomendável caso haja manifestação de sintomas devido ao tamanho excessivo, ou no caso de suspeita de malignidade. Porém

a maior parte dos nódulos são assintomáticos, assim o resultado do diagnóstico, baseado em informações provenientes de diversos exames, é o fator determinante na decisão sobre uma possível remoção. Alguns fatores no histórico de um paciente que atuam como fortes indicativos de malignidade são [Bomeli et al., 2010]:

- contato anterior com radiação, especificamente na região da cabeça e pescoço durante o período da infância;
- histórico familiar de carcinoma medular de tireoide ou neoplasia endócrina múltipla;
- crescimento rápido;
- dificuldade de deglutir (disfagia);
- alteração na voz (disfonia);
- inflamação dos gânglios linfáticos cervicais;
- pregas vocais imobilizadas.

Há ainda algumas características físicas do nódulo em si que podem aumentar a suspeita de malignidade, entre elas: o tamanho maior que 4 cm; a consistência firme; e o grau de fixação aos tecidos adjacentes [Bomeli et al., 2010].

Na maior parte dos casos a verificação inicial da incidência de um nódulo ocorre através da palpação da região. Entretanto, estudos apontam que a incidência de nódulos palpáveis é de 4% a 7%. Além disso, para maior detalhamento da natureza do nódulo, faz-se necessário empregar outras ferramentas a fim de que mais informações sejam obtidas.

Os exames normalmente utilizados no processo de diagnóstico de nódulos malignos na tireoide são: exames laboratoriais, de imagem e baseados em punção aspirativa com agulha fina (PAAF) . A diretriz brasileira atual sugere a sequência de passos mostrado na Figura 2.1.

Note que o processo parte da avaliação do TSH, e dependendo do resultado dela, o paciente pode ser encaminhado para a realização de uma PAAF ou uma cintilografia. Como já mencionado, o resultado da PAAF é determinante para o prosseguimento do tratamento já que seu resultado pode levar a realização da cirurgia. As avaliações mencionadas são descritas nas seções a seguir.

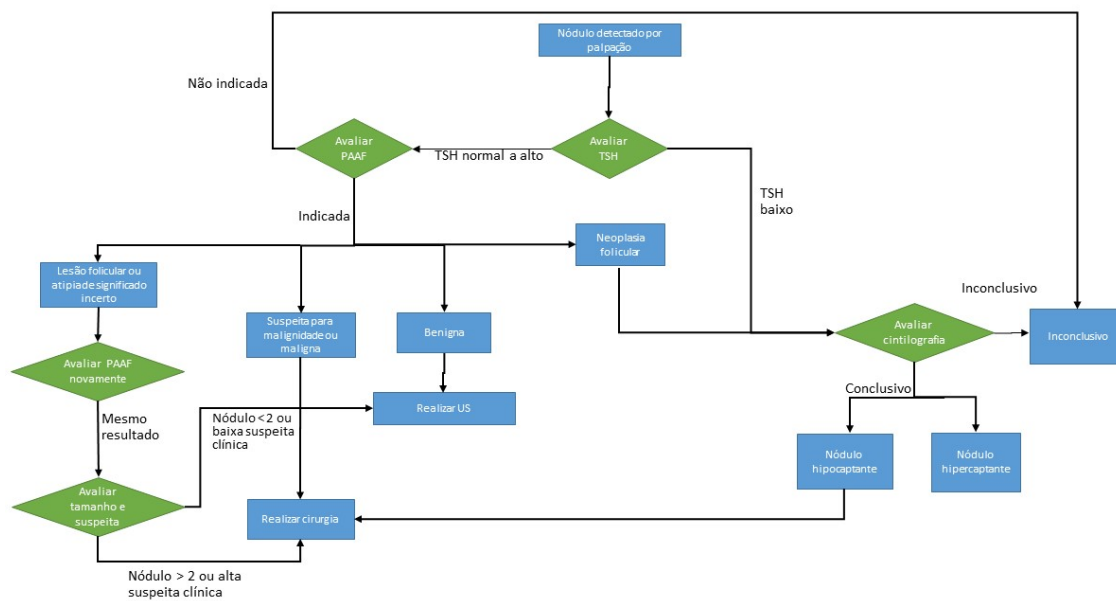


Figura 2.1: Avaliação de um paciente portador de nódulos tireoidianos

#### 2.1.1.1 Verificação do TSH

Dentre os pacientes que apresentam um nódulo somente, a maior parte apresenta alterações no funcionamento da tireoide. Tais disfunções podem ser detectadas através da verificação do TSH [Pedroza Ballesteros, 2008]. Caso esse valor esteja abaixo do esperado, o que pode indicar hipertireoidismo, é feita a verificação de tiroxina total ou livre (T4) e triiodotironina total (T3). A verificação de níveis séricos de calcitonina é indicada para pacientes de alto risco, que apresentam um histórico familiar de casos de carcinoma medular da tireoide, e não é suficiente para diferenciar nódulos malignos de benignos.

#### 2.1.1.2 Ultrassonografia

A ultrassonografia é um método muito mais preciso, já que permite identificar nódulos de qualquer tamanho, chegando a detectar nódulos em 67% dos casos [Ezzat et al., 1994]. Através desse exame, é possível detectar nódulos de dimensões pequenas que não poderiam ser detectados através de palpação. O exame também propicia identificar a presença de múltiplos nódulos. A maior precisão oferecida por esse exame permite definir dimensões precisas para o diâmetro do nódulo. Outras informações também são extraídas através da ultrassonografia [Bomeli et al., 2010], como:

- hipocogenicidade (baixa reflexão de ondas de ultrassom);

- aumento da vascularização;
- presença de microcalcificações;
- margens irregulares;
- ausência de halo.

Apesar dessas características serem informações muito úteis no processo de diagnóstico, elas são subjetivas e não permitem por si só diferenciar nódulos malignos de benignos. Porém, elas facilitam o procedimento de biópsia aspirativa por agulha fina, já que ajudam a definir a melhor posição para realizar o procedimento. Além disso, tais informações podem ser usadas na seleção de um nódulo para a realização da biópsia no caso de múltiplos nódulos [Papini et al., 2002].

#### **2.1.1.3 Outros exames por imagem**

A cintilografia com isótopos é usada para determinar o nível de atividade de um nódulo [Ferreira et al., 2009]. Entretanto, ela não fornece uma medida precisa do tamanho. A cintilografia objetiva definir se o nódulo analisado é hipercaptante, o que é um indicativo de possível malignidade [Rosário et al., 2013]. Iodo radioativo ou Tc pertecnetato podem ser usados em cintilografias [Bomeli et al., 2010].

Geralmente, tanto a tomografia computadorizada quanto a ressonância magnética não são usadas no estágio inicial do processo de diagnóstico. Esses exames também não permitem diferenciar nódulos malignos de benignos. Entretanto, eles permitem identificar estruturas que podem estar comprimindo ou invadindo regiões adjacentes à tireoide [Bomeli et al., 2010, Rosário et al., 2013].

#### **2.1.1.4 PAAF**

A biópsia ou punção aspirativa por agulha fina (PAAF) é determinante no processo de diagnóstico de um nódulo tireoidiano [Torres et al., 2002]. Isso se deve ao fato de que a citologia do nódulo é o principal fator a ser considerado na avaliação de retirada total ou parcial do tireoide.

O processo de PAAF apresenta um baixo risco de complicações e pode ser realizado com ou sem orientação por ultrassom (US). Entretanto o uso do US aumenta a acurácia

diagnóstica do processo, o que diminui o número de falsos-negativos e amostras não-adequadas [Rosário et al., 2013].

Geralmente, o resultado de uma PAAF é descrito como: benigno, maligno, indeterminado ou não-diagnosticado. Entretanto, o Instituto Nacional do Câncer dos Estados Unidos propôs o uso das seis categorias de citologia para melhor descrever o risco de malignidade mostrado na Tabela 2.1, que constituem a classificação de Bethesda [Baloch et al., 2008]. Tal classificação é bastante utilizada no mundo inteiro, inclusive no Brasil.

Tabela 2.1: Classificação de Bethesda

Classe	Significado	Risco de malignidade (em %)
I	Amostra insatisfatória	1-4
II	Nódulo benigno	0-3
III	Atipia ou lesão folicular indeterminada	5-15
IV	Neoplasia folicular (ou suspeita de)	15-30
V	Lesão suspeita de malignidade	60-75
VI	Nódulo maligno	97-99

# Capítulo 3

## Revisão Bibliográfica

Este capítulo apresenta uma breve revisão bibliográfica dos tópicos relevantes a este estudo, considerando um aspecto mais tecnológico. Ele está dividido em três partes. A primeira parte mostra um aspecto mais prático, apresentando o uso de termografias para problemas médicos em geral, bem como trabalhos relacionados a este, em especial os previamente desenvolvidos pelo grupo do Laboratório de Computação Visual (VisualLab), do Instituto de Computação (IC) da Universidade Federal Fluminense (UFF).

Na segunda parte apresentam-se os testes estatísticos Wilcoxon e tamanho do efeito, posteriormente usado neste trabalho na avaliação de parâmetros baseados em termografias. A terceira parte abrange conceitos básicos de redes neurais convolucionais e apresenta brevemente algumas aplicações na área médica.

### 3.1 Termografias

A temperatura é um fator muito importante na detecção de diversas doenças e disfunções, e da integridade fisiológica e patológica de um indivíduo. Ela vem sendo analisada desde os primórdios da medicina [Brioschi et al., 2010]. Por volta de 1872, a mensuração da temperatura passou a ser incluída em rotinas de avaliações clínicas. Ela está entre os cinco sinais vitais a serem observadas nas consultas médicas, juntamente com pulso, pressão arterial, frequência respiratória e dor.

Diversos métodos podem ser utilizados na análise da temperatura de um paciente, desde as mais grosseiras, como o toque com o dorso da mão na região afetada, até o uso de instrumentos mais sofisticados, como termômetros e outros sensores térmicos. Nos anos 1940, foi desenvolvida a captura de temperaturas através de câmeras com sensores de

frequências infravermelhas. Apesar dessa técnica ter sido inicialmente utilizada para fins militares, seu uso na área médica vem ganhando espaço. Ela pode ser considerada uma modalidade de imagem não invasiva, que permite avaliar a fisiologia do paciente, medindo a distribuição de temperatura na superfície de seu corpo [Brioschi et al., 2010]. Essas imagens podem ser utilizadas para fins de diagnóstico, como uma informação adicional a ser analisada no processo [Lahiri et al., 2012, Barnes, 1963, Diakides et al., 2012]. Isso se deve ao fato de que grande parte das lesões demandam um maior fluxo sanguíneo, o que consequentemente leva a um aumento de temperatura na região[Brioschi et al., 2010]. Quando essas lesões são relativamente superficiais, esse aumento de temperatura pode influenciar a temperatura cutânea, o que pode vir a ser detectado em imagens termográficas.

Apesar da disponibilidade de outros exames de imagem já estabelecidos, como RM, US, TC, entre outros, a termografia pode constituir uma alternativa a eles. Essa alternativa, por se tratar de um exame muito barato se comparado a outros, além de não ser invasivo e não expor o paciente a nenhum tipo de onda ou substância, pode vir a ser muito interessante.

O US emite ondas sonoras para a região de exame a fim de observar seu comportamento nessa região. O CT emite raio x, que são ondas eletromagnéticas ionizantes, potencialmente danosas pelo seu efeito acumulativo, e com potencial de modificação do DNA das células, o que, entre outros problemas de saúde, pode ser cancerígeno nos organismos vivos. Os exames termográficos são baseados na energia naturalmente emitida por qualquer corpo com temperatura acima do zero absoluto.

Especificamente no caso da região da tireoide, o uso de termografias poderia fornecer informações importantes na avaliação de lesões, visto que este é um órgão superficial cuja temperatura pode influenciar na temperatura cutânea vizinha.

### 3.1.1 Trabalhos relacionados

Em Mahajan and Madhe (2014) pacientes saudáveis e com tireoides hiperfuncionantes tiveram suas termografias capturadas. Nessas termografias as temperaturas foram mapeadas para imagens em cores permitindo melhor visualização (Figura 3.1). Essas capturas foram realizadas usando uma câmera térmica Forward Looking Infrared (FLIR) modelo E-30.

As imagens foram tratadas usando filtros de Mediana [Sonka et al., 2014, pp. 74-76], e filtros Gaussianos [Sonka et al., 2014, p. 445], entre outros, e aplicando operações de



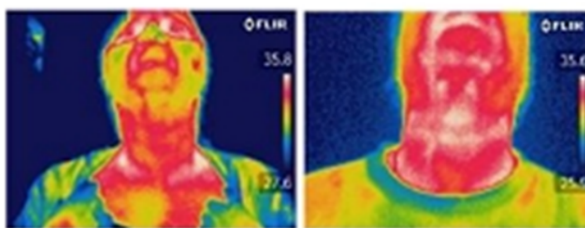


Figura 3.1: Termografia usada no estudo [Mahajan and Madhe, 2014]

equalização [Conci et al., 2008, pp. 93-96] a fim de melhorar a qualidade dessas imagens. Após isso, as imagens foram segmentadas usando o método Otsu [Otsu, 1979], a fim de separar a região da tireoide de resto da imagem (Figura 3.2). Por fim, foram extraídas as seguintes características dessas imagens: área da lesão, e média e desvio padrão dos valores na região da lesão.

Uma análise considerando essas características foi feita para 14 pacientes, a fim de identificar algum padrão de acordo com a disfunção apresentada. Cinco desses pacientes eram saudáveis e os demais apresentavam hiperfuncionamento da tireoide. Os autores propuseram valores limites para essas características, a fim de determinar se a tireoide apresenta alguma disfunção.

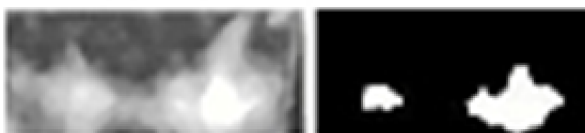


Figura 3.2: Segmentação de lesão em termografia da tireoide [Mahajan and Madhe, 2014]

Em D'Arbo et al.(1988) é realizado um estudo a fim de definir se a termografia pode ser utilizada no processo de seleção de nódulos tireoidianos com indicação cirúrgica. São estudados 110 pacientes, dos quais 92 eram mulheres e 18 eram homens. Suas idades variam entre 2 e 77 anos. Todos os pacientes apresentavam nódulos, sendo que 98 apresentavam nódulos únicos, e 12 apresentavam nódulos múltiplos, totalizando 124 lesões. Para avaliar a temperatura dessas regiões, termografias estáticas dos pacientes foram capturadas. Ao avaliar a temperatura desses nódulos, 100 deles foram considerados frios e 24 foram considerados mornos. Dentre os nódulos frios, 16 deles eram malignos. Dentre os nódulos mornos, 2 eram malignos. Esse resultado diverge um pouco do esperado, já que a princípio os nódulos frios não deviam apresentar tantos casos malignos devido a sua baixa vascularização, que é um indicativo de baixo risco de malignidade. Foram calculados: a máxima diferença de isotermas, o índice térmico e o perfil térmico. O índice térmico é definido a partir da temperatura média da área delimitada. A máxima diferença de isoter-

mas consiste no menor número de pontos de maior temperatura pertencentes a região da lesão. O perfil térmico consiste em selecionar o ponto mais representativo (pico) da lesão, e comparar com a média dos pontos mais marcantes das outras duas áreas anteriormente estabelecidas e analisadas.

Foi observado que o valor  $0,9^{\circ}\text{C}$  poderia ser usado como limite para definir a natureza da lesão. Definiu-se que caso o valor seja maior a  $0,9^{\circ}\text{C}$  a lesão é considerada maligna, caso contrário, a lesão é considerada benigna. Usando esse critério, as medidas apontaram uma acurácia média de 92%. Assim os autores concluem que as termografia podem ser usadas no processo de seleção de nódulos tireoidianos de indicação cirúrgica.

No trabalho apresentado em Brioschi et al. (2010), os autores realizam um estudo sugerindo possíveis parâmetros a serem usados na análise de anomalias da tireoide. Tais parâmetros são usados na análise apresentada nesta dissertação. Apesar de não realizar nenhum estudo quantitativo, o estudo apresenta uma proposta de como as termografias podem ser incluídas como exame complementares no processo de diagnóstico de doenças da tireoide [Brioschi et al., 2010].

No trabalho apresentado em Vaz (2014), os autores utilizam termografias para identificar disfunções da tireoide, especificamente hipotireoidismo e hipertireoidismo. Uma glândula tireoide hiperativa apresenta um maior fluxo sanguíneo, o que leva a um aumento de temperatura na região. Esse aumento pode ser observado em termografias, já que elas consistem em uma captura de temperaturas de uma área. Em Vaz (2014), são usadas termografias estáticas, as quais são tratadas com um filtro de mediana e posteriormente equalizadas [Conci et al., 2008]. A partir disso, imagens relativas a região de interesse (region of interest - ROI) são extraídas. Essas ROIs são segmentadas usando o método de Otsu [Otsu, 1979]. Características são extraídas da região segmentada. Uma classificação das termografias, considerando as características computadas, é realizada usando um classificador Multilayer Perceptron (MLP). Os experimentos usando validação cruzada 3-fold e 10-fold levaram a sensibilidade de 93,12% e 89,62%, e especificidade de 91,80% e 89,70%.

Em Aweda et al. (2012), parâmetros estatísticos básicos são calculados a partir de termografias estatísticas de 53 pacientes do Hospital Escolar de Lagos na Nigéria, sendo 16 saudáveis e 37 de pacientes com doenças da tireoide. Os parâmetros usados no estudo são a temperatura média e o desvio padrão. Os autores chegaram a conclusão de que a temperatura média pode ser utilizada na classificação de pacientes entre saudáveis e doentes.

Há ainda alguns trabalhos que usam termografias na detecção de disfunções das tireoide. Helmy et al. (2000) desenvolveram um modelo de simulação de glândulas tireoidianas, utilizando análise de elementos finitos, considerando três coeficientes de transferência de calor: condução, convecção e radiação.

Em Gavrioloia et al. (2009), termografias da tireoide são analisadas a fim de encontrar descritores a serem usados na detecção de tumores. São consideradas características das imagens, como assimetria da forma do objeto, regularidade da borda, cor, diâmetro e evolução de contorno. Dos casos estudados 89,3% dos 24 pacientes com câncer de tireoide foram identificados através das características propostas.

Já em Gavrioloia et al. (2010), análise fractal é utilizada para caracterizar tumores em termografias, quantificando as irregularidades em tamanho e forma. São usados dados da dimensão fractal e da lacunaridade, para melhor identificar padrões térmicos e de bordas.

### 3.1.2 Trabalhos do grupo VisualLab

O grupo de pesquisa do VisualLab, formado por pesquisadores do Instituto de Computação (IC) e do Hospital Universitário Antonio Pedro (HUAP), ambos da Universidade Federal Fluminense (UFF), vem estudando nos últimos anos a aplicabilidade de termografias em diagnóstico de disfunções da área da tireoide. Os trabalhos desenvolvidos pelo grupo são apresentados a seguir. Este trabalho utiliza exames capturados nos trabalhos anteriores, sendo em alguns aspectos uma continuação deles.

#### 3.1.2.1 González, 2017

Os trabalhos desenvolvidos por González (2017) [González, 2017, González et al., 2017] apresentam um estudo sobre a viabilidade do uso de termografias infravermelhas dinâmicas (TIDs) na detecção de nódulos tireoidianos.

Para isso, é proposto inicialmente um protocolo de captura de Termografias Infravermelhas Dinâmicas (TID) usando uma câmera FLIR (Forward Looking Infrared) modelo SC 620. As TIDs consistem em uma sequência de imagens em infra-vermelho adquiridas enquanto o paciente examinado entra em equilíbrio térmico com o ambiente após um stress térmico. O processo de aquisição de imagens foi realizado com acompanhamento de especialistas do Departamento de Endocrinologia do Hospital Universitário Antônio Pedro (HUAP) da Universidade Federal Fluminense. Para a captura, foram selecionados pacientes que apresentam nódulos hipofuncionantes, os quais participaram voluntaria-

mente do projeto. Esses voluntários são examinados segundo o protocolo definido, e os seus dados são incluídos no DTNR.

O protocolo de captura para termografias da região da tireoide, permite a obtenção de uma série temporal de imagens do paciente em diferentes graus de resfriamento [González et al., 2017]. Estudos indicam que no caso da tireoide anomalias são mais perceptíveis em pacientes que sofreram algum tipo de resfriamento ou choque térmico [Brioschi et al., 2010]. No protocolo, o paciente é submetido a um resfriamento prévio à captura, diminuindo a temperatura corporal do paciente até próximo de 30°C. A partir daí, uma sequência de imagens é capturadas a cada 15 segundos, durante 5 minutos, resultando num total de 20 imagens.

Uma das mais importantes contribuições do trabalho é a criação do repositório de dados e imagens DTNR, o qual está sendo desenvolvido pelo grupo de pesquisa do Laboratório de Computação Visual “VisualLab” do Instituto de Computação da Universidade Federal Fluminense [González, 2017]. Esse repositório está disponível para acesso através da página <http://visual.ic.uff.br/thyroid/>. O DTNR é parte do projeto “Avaliação da importância da termografia no auxílio à investigação diagnóstica de nódulos tireoidianos em pacientes acompanhados no HUAP-UFF” aprovado pelo Comitê de Ética em Pesquisa (CEP) do Hospital Universitário Antônio Pedro da Universidade Federal Fluminense (HUAP-UFF) que está registrado na Plataforma Brasil do Ministério de Saúde sob o número de Certificado de Apresentação para Apreciação Ética (CAAE) 57078516.8.0000.5243. Nele, encontram-se além das termografias capturadas, os mapas de temperatura correspondentes, imagens que definem a região de interesse (ROI), a informação da localização dos nódulos, dados complementares do paciente, como resultados de exames anteriores, histórico e, dependendo do nível de acesso do usuário, dados pessoais [González, 2017].

Para o estudo da viabilidade do uso de TIDs para detecção de nódulos é analisada a evolução das temperatura em termografias de pacientes, nas quais a localização e tipos dos nódulos são previamente conhecidos. Um conjunto de pontos selecionados aleatoriamente na área pertencente a nódulos e na área da região saudável da ROI tem a evolução de suas temperaturas acompanhada. As séries de intensidade são analisadas através de diferenças e integrais. Foi observado nessa análise que essas informações não permitem diferenciar regiões saudáveis de regiões com nódulo benigno. É sugerido que um estudo seja feito com termografias infravermelhas estáticas (TIE), a fim de verificar a viabilidade do uso delas na detecção de nódulos tireoidianos [González et al., 2017].



Figura 3.3: Página de entrada do DTNR

### 3.1.2.2 Fiirst, 2018

No trabalho desenvolvido por Fiirst (2018) [Fiirst, 2018, Fiirst et al., 2018], melhorias no protocolo de González (2017) foram desenvolvidas. Tais melhorias são de especial interesse do trabalho apresentado nesta dissertação, já que elas facilitam a manipulação e processamento das termografias analisadas. Além disso também é proposta uma metodologia para acompanhamento de pontos em séries temporais de termografias [Fiirst, 2018].

No protocolo proposto por González (2017), para cada captura, as imagens apresentam um intervalo de temperaturas capturadas diferente. Esse intervalo varia de acordo com a temperatura ambiente da sala no dia da captura e da temperatura corporal do paciente em análise. No trabalho realizado optou-se por padronizar os limites dos mapas de temperatura disponíveis no banco de dados, de modo a ter-se uma associação única entre o nível de cinza e a temperatura da imagem em análise [Fiirst, 2018]. Assim as temperaturas originais passaram a ser representadas através de valores obtidos ao aplicar uma transformação linear de forma a ajustá-los no intervalo de temperaturas de 20°C a 40°C. Esses limites foram definidos após a consideração dos valores máximos de temperatura registrados em todos os exames feitos até 30/10/2017. Após isso as tonalidades das imagens são definidas de acordo com a temperatura do ponto nesse intervalo [Fiirst, 2018]. Observe que, devido ao processo de padronização de limites no mapa de temperatura, a equivalência entre temperatura e os tons limites nas imagens é obtida da relação:

$$20^{\circ}\text{C} \longrightarrow \text{tom } 0$$

$$40^{\circ}\text{C} \longrightarrow \text{tom } 255$$

logo, uma variação de  $20^{\circ}\text{C}$  é dada por uma variação de 256 tons. Dessa forma, uma variação  $1^{\circ}\text{C}$  seria dada por

$$\text{variação de } 1^{\circ}\text{C} \longrightarrow \text{variação de } t'' \text{ tons}$$

$$t'' = 1^{\circ}\text{C} \times 256 \text{ tons} / 20^{\circ}\text{C} = 12,80 \simeq 13 \text{ tons}$$

Outra melhoria proposta por foi a inclusão da localização dos nódulos no protocolo de captura [Fiirst, 2018]. Essa informação é fornecida por especialistas médicos, que se baseiam na palpação e em dados de ultrassonografias previamente adquiridas para defini-las. O processo feito pelo especialista de identificar o nódulo no paciente consiste basicamente em, após a captura, apalpar o paciente na região indicada pelos dados da ultrassonografia, de forma a sentir alguma estrutura que possa ser caracterizada como nódulo. Caso essa estrutura seja encontrada, o especialista fixa um marcador no local, e uma nova imagem é capturada. Dessa forma há, para cada paciente, uma termografia na qual um marcador indica a posição do(s) nódulo(s), além da série de 20 termografias capturadas para cada paciente [Fiirst, 2018].

O trabalho desenvolvido ainda apresenta métodos para acompanhamento de pontos em termografias de uma série [Fiirst, 2018]. Cada imagem da série consiste na captura referente a um instante único de tempo, no qual diferentes variações podem ocorrer, de forma que se numa primeira termografia, um ponto referente a um detalhe do paciente está no pixel  $pix_1$ , numa segunda termografia esse mesmo ponto pode estar num pixel  $pix_2$ . Isso se deve a pequenas mudanças no posição do paciente, que naturalmente ocorrem ao longo da captura [Fiirst, 2018]. Para estudos de informações baseadas em séries temporais na série de termografias, é necessário estabelecer uma função que permite encontrar o pixel correspondente a um ponto específico em cada uma das termografias da série. Apesar da grande contribuição desse métodos para o estudo de termografias [Fiirst, 2018], no trabalho aqui apresentado, esses métodos não foram aplicados, já que aqui o estudo se baseia em termografias estáticas.

### 3.1.3 Testes estatísticos

Nesta seção são apresentadas os testes usados neste trabalho para verificação das hipóteses propostas, ou seja, para a comprovação de que as estatísticas analisadas permitem a diferenciação de pacientes, o que é mostrado no capítulo 4.

### 3.1.4 Teste Wilcoxon-Mann-Whitney

O teste de Wilcoxon-Mann-Whitney, aqui referido apenas como teste de Wilcoxon, é um teste estatístico de hipóteses não-paramétrico ou seja, é usado para avaliar conjuntos de dados cuja distribuição não é conhecida. Partindo de duas populações P1 e P2, as quais podem ter tamanhos diferentes, o teste avalia se seus valores apresentam uma diferença significativa, de forma a que possam ser consideradas diferentes. Para isso 2 amostras devem ser selecionadas, uma amostra  $A = \{A_1, A_2, \dots, A_l\}$  da população P1, e uma amostra  $B = \{B_1, B_2, \dots, B_m\}$  da população P2 são selecionadas, sendo  $l \leq m$ . Observe que  $l$  e  $m$  correspondem ao número de elementos de A e B respectivamente. Parte-se do princípio que A e B são mutualmente independentes.

Assim, considera-se que existam funções de distribuição F e G, que correspondem a P1 e a P2, de forma que  $F(z) = G(z)$ , para todo z, seja uma hipótese válida. Essa hipótese é chamada  $H_0$ , ou hipótese nula. Uma segunda hipótese, chamada hipótese alternativa  $H_A$ , sugere que A tende a ser diferente a B, ou seja,  $F(z) = G(z - \Delta)$ , para todo z. Ou seja, existem valores  $\Delta$  que determinam uma diferença entre os valores dados por F e G. Dessa forma, os seguintes cenários podem ser estabelecidos:

$$H_0 : \Delta = 0 \quad e \quad H_A : \Delta \neq 0$$

$$H_0 : \Delta = 0 \quad e \quad H_A : \Delta > 0$$

$$H_0 : \Delta = 0 \quad e \quad H_A : \Delta < 0$$

No primeiro caso, na hipótese alternativa  $H_A$ ,  $\Delta$  determina que os valores dados por F e G são diferentes. A hipótese alternativa  $H_A$ , do segundo caso indica que  $\Delta$  determina que os valores dados por F são superiores ao de G. No último caso,  $\Delta$  da  $H_A$  determina que os valores dados por F são inferiores aos dados por G.

A partir disso, os elementos de A e B são ordenados em ordem crescente. Para cada elemento tem-se a posição em que ele se encontra nessa lista ordenada. Considera-se  $S_A$  como a soma das posições dos elementos de A na lista ordenada, e  $S_B$  a soma das posições dos elementos de B na lista ordenada. Assim, W é dado pela equação 3.1.

$$W = S_b - \frac{1}{2}l(l+1) \tag{3.1}$$

Por fim o p-valor correspondente ao valor de W pode ser obtido na tabela K de Siegel [Siegel and Castellan Jr, 1988]. O p-valor determina se a hipótese  $H_A$  pode ser

considerada. Neste estudo, o nível de confiança considerado é 95%, assim caso  $p\text{-valor} < 0,05$  a  $H_A$  pode ser considerada válida.

### 3.1.5 Tamanhos de efeito

Os tamanhos de efeito (TDE) constituem uma importante informação a ser considerada em análises estatísticas, pois minimizam problemas referentes a amostras pequenas e resultados não-significativos, permitindo caracterizar melhor diferenças entre dois grupos. Segundo [Schneider and Darcy, 1984], há sete características que influenciam testes de hipóteses de significância, sendo o tamanho da amostra a característica mais determinante [Snyder and Lawson, 1993]. Dessa forma, o p-valores encontrados são altamente influenciados pelo número de elementos da amostra considerada, afetando inclusive sua validade. Os TDEs são medidas independentes do tamanho da amostra, sendo assim adequadas em casos de amostras pequenas.

Na literatura é possível encontrar diferentes definições de TDEs. TDEs podem ser entendidos como a diferença entre duas amostras segundo [Rosenthal, 1994], o grau em que a hipótese nula é falsa segundo [Cohen, 1995, pp. 9-10], ou mais especificamente como o efeito de uma variável sobre outra [Hedges, 1981]. Os TDEs pode ser divididos em duas famílias: a primeira que representa a magnitude entre a relação entre as variáveis; e a segunda que se refere a magnitude da diferença entre grupos considerando uma característica específica, realizando um comparação entre esses grupos [Espírito Santo and Daniel, 2017]. Algumas das medida pertencentes a esse segundo grupo são d de Cohen, o g de Hedges e o  $\Delta$  de Glass. Considerando a natureza do estudo aqui apresentado o g de Hedges é calculado a fim de agregar ao teste Wilcoxon também utilizado.

O g de Hedges é geralmente utilizado quando a amostra é pequena, e os grupos considerados têm tamanhos diferentes, como é o caso deste estudo [Ellis, 2010, Hedges, 1981, Rosnow et al., 2000]. Considerando novamente as amostras  $A = \{A_1, A_2, \dots, A_l\}$  da população P1, e  $B = \{B_1, B_2, \dots, B_m\}$  da população P2, valor de g é calculado através da fórmula 3.2

$$g = \frac{media_A - media_B}{DP_{combinado} * 1 - \frac{3}{4*gl-1}} \quad (3.2)$$

sendo  $gl$  o grau de liberdade dado por  $l + m - 1$ , e  $DP_{combinado}$  o desvio padrão combinado dado pela equação 3.3:



$$DP_{combinado} = \frac{(l-1)\sigma_A^2 + (m-1)\sigma_B^2}{l+m-2} \quad (3.3)$$

onde  $\sigma_A$  e  $\sigma_B$  referem-se os valores de desvio padrão das amostras A e B.

Os valores encontrados para g podem ser interpretados da seguinte forma [Cohen, 1988, Durlak, 2009]:

- $0,50 \leq g \leq 1,00$  : grande;
- $0,30 \leq g \leq 0,49$  : moderada;
- $0,10 \leq g \leq 0,29$  : pequena.

## 3.2 Redes Neurais Convolucionais

Redes Neurais Feed-Forward [Goodfellow et al., 2016] são algoritmos geralmente usados para reconhecimento de padrões baseados no funcionamento do sistema nervoso. Têm como objetivo definir uma função  $f$  de forma que através dela seja possível realizar algum processamento específico de dados. Por exemplo, considerando um problema de classificação, tem-se uma função  $y = f^*(x)$  que deve mapear uma entrada  $x$  para uma categoria  $y$ . O objetivo da rede associada a esse problema seria encontrar um conjunto de parâmetros  $\omega$  a serem usados na função  $f^*$  de forma a torná-la o mais precisa possível, ou seja, permitir que  $f^*$  retorne o  $y$  correto para o maior número de entradas  $x$  possíveis [Goodfellow et al., 2016]. Para isso, usa-se um conjunto de entradas para as quais já se sabe a saída esperada, chamado conjunto de treinamento. Tal processo pode ser entendido como um problema de otimização

Esse tipo de modelo é chamado de feedforward porque ao longo da rede diversos cálculos são realizados de forma subsequente, partindo da entrada inicial até chegar numa saída  $y$ , encontrando os parâmetros  $\omega$  e a função  $f^*$  nesse processo. Observe que essa ideia também está ligada ao conceito de rede, já que representa uma sequência de funções ligadas. Dessa forma a função  $f^*$  consiste em uma função composta [Goodfellow et al., 2016], formada por outras funções  $f^{(1)}, f^{(2)}, f^{(3)}, \dots, f^{(n)}$ , ou seja:

$$f^* = f^{(n)}(\dots f^{(3)}(f^{(2)}(f^{(1)}(x))))$$

As funções intermediárias são chamadas de camadas. As funções que recebem como

entrada a entrada inicial são chamadas camadas de entrada. As funções que realizam os últimos cálculos, retornando a saída final, são chamadas camadas de saída. As demais camadas são chamadas de intermediárias, ou ocultas.

Cada camada é formada por elementos chamados neurônios, que dada uma entrada vetorial dão uma saída escalar, de acordo com uma função de ativação [Goodfellow et al., 2016]. Diversas funções de ativação podem ser utilizadas, dentre elas a Heaviside, a Sigmóide e a Softmax. Atualmente as unidades ReLU são predominantes nos algoritmos de redes feedforward. Essas unidades são dadas pela função de ativação  $g(z) = \max\{0, z\}$ .

Nas camadas totalmente conectadas, cada elemento de entrada está conectado com todos os neurônios da camada. Isso pode tornar o processamento muito custoso.

O processo de concatenação consiste em unir as saídas das várias funções de uma camada. A técnica de Dropout consiste em uma regularização, que atua desativando os neurônios de uma camada segundo uma probabilidade definida.

O número de camadas define a profundidade da rede, sendo o termo redes neurais profundas usado para designar redes com um número alto de camadas. Uma forma de representar graficamente redes neurais feed-forward é através da estrutura de grafos, nos quais cada nó corresponde a uma camada [Goodfellow et al., 2016].

As redes neurais convolucionais consistem em um tipo especializado de redes feed-forward [Goodfellow et al., 2016]. Esse tipo de rede foi projetada para manipular informações que apresentam estrutura serial, ou seja, que consistem em uma sequência de informações ordenadas. Exemplos desse tipo de informações são imagens.

Redes Neurais Convolucionais são redes neurais que aplicam operações de convolução. Nesse processo, são considerados apenas pixels em posições vizinhas nas imagens de entrada. Em outras palavras, as camadas de convolução aplicam filtros em regiões da imagem analisada, deslizando as posições (stride) em cada iteração até percorrer a imagem toda [Fonseca, 2017].

Geralmente, as camadas de convoluções trabalham em três fases: na primeira, várias operações de convoluções são paralelamente aplicadas gerando um conjunto de funções lineares de ativação. Na segunda fase, uma função de ativação não-linear é aplicada sobre as ativações lineares. Na terceira e última fase uma função de pooling é aplicada. O pooling utiliza uma função que agrupa uma janela de elementos devolvendo um valor único, de forma que as dimensões da saída são menores que a da entrada. Essa operação sintetiza a informação de uma região em um só dado. Diferentes funções podem ser uti-

lizadas na fase de pooling, por exemplo: o max pooling, que seleciona o maior elemento da janela e o atribui como saída. No caso do average pooling, a saída é dada pela média dos valores os elementos da janela utilizada [Goodfellow et al., 2016]. O processamento realizado pelas camadas de convolução é especialmente interessante para problemas envolvendo imagens, já que os resultados podem representar informações sobre a vizinhança de um pixel [Fonseca, 2017].

### 3.2.1 Modelo Residuais

Redes Neurais Residuais, conhecidas como ResNets, propostas inicialmente em [He et al., 2016], apresentam, além das camadas já descritas, blocos residuais, os quais são compostos por duas camadas de convolução e duas funções de ativação ReLU. Sua saída é a soma de sua entrada e a saída de sua segunda camada de convolução. O modelo ResNet usado neste trabalho é o proposto em [He et al., 2016], que é composto por 34 camadas, e é mostrado na Figura 3.4.

Esse modelo apresentou ótimos resultados para problemas similares ao apresentado neste trabalho [Lotter et al., 2017, Kwajiri and Tezuka, 2017, Dogdas et al., 2018].

### 3.2.2 Modelo GoogLeNet

O segundo modelo de rede usado neste trabalho é o GoogLeNet [Szegedy et al., 2015]. Esse modelo é formado por módulos chamados de inception, resultando em um total de 27 camadas. É amplamente conhecido por seus resultados para problemas de classificação estabelecidos na literatura [Zhong et al., 2015, Ballester and de Araújo, 2016]. Ele consegue resultados superiores em comparação com outras arquiteturas CNN, mesmo considerando que ele usa menos parâmetros e tem menos camadas. A arquitetura GoogLeNet (Figura 3.5) tem um total de 9 módulos inception e algumas camadas adicionais externas a nesses módulos.

Podemos ver os módulos inception como redes, de modo que toda a arquitetura pode ser vista como uma grande rede formada por redes menores. Os módulos inception são formados por camadas de convolução de 1x1, 3x3 e 5x5 e uma camada de max pooling de 3x3, como mostrado na Figura 3.6.

### 3.2.3 CNNs em problemas de classificações de imagens médicas

Algoritmos de redes neurais têm mostrado bons resultados em uma ampla gama de problemas clínicos. Mais recentemente, abordagens baseadas em deep learning usando CNNs alcançaram desempenho impressionante em tarefas médicas relacionadas a imagens [Lévy and Jain, 2016].

A viabilidade do uso de CNNs para classificação de massas mamárias pré-segmentadas em mamografias é considerada no trabalho [Lévy and Jain, 2016]. O objetivo é classificar essas lesões em malignas ou benignas. Três arquiteturas de rede foram avaliadas: um modelo CNN simplificado (usado como modelo base), um modelo AlexNet e um modelo GoogLeNet. O banco de dados digital para mamografia de rastreamento (DDSM [Heath et al., 2000]) foi utilizado nos experimentos realizados. Os resultados mostram que a rede AlexNet supera significativamente o modelo base na eficácia da aprendizagem de transferência. Considerando o desempenho, o GoogLeNet supera os outros modelos por uma margem ampla. Durante uma comparação final, os resultados de todos os modelos avaliados superam os de radiologistas treinados [Lévy and Jain, 2016].

Em outro trabalho são classificadas lesões de pele usando uma única CNN, treinada a partir de imagens diretamente, usando apenas pixels e rótulos de doenças como entradas [Esteva et al., 2017]. A CNN é treinada usando um conjunto de dados de 129450 imagens clínicas consistindo de 2032 doenças diferentes. Os resultados são comparados com a avaliação de 21 dermatologistas. Eles compararam carcinomas de queratinócitos versus queratoses seborreicas benignas e melanomas malignos versus nevos benignos. A CNN classifica o câncer de pele com um nível de competência comparável ao dos dermatologistas.

Outro trabalho que visa avaliar e classificar as lesões cutâneas é apresentado em [Jafari et al., 2016], no qual um CNN é utilizado para identificar regiões com lesões de melanoma. A entrada consiste em imagens capturadas por câmeras digitais, portanto, alguns fatores devem ser considerados no processamento, como por exemplo, variações de iluminação e reflexos de luz. Informações contextuais locais e globais são consideradas pela CNN. A saída consiste em um rótulo para cada pixel, indicando se é parte da lesão ou não. Os resultados mostram que o método proposto supera os algoritmos existentes no estado-da-arte em termos de precisão de segmentação.

Uma CNN é usada para encontrar características genéricas em imagens de endoscopia, a fim de detectar lesões [Zhu et al., 2015]. As imagens utilizadas são extraídas de

exames gastrointestinais obtidos por gastroscopia. Os autores usam os recursos extraídos pelo CNN desenvolvido para alimentar uma máquina de vetores de suporte (SVM) para melhorar a capacidade de generalização do método. O teste realizado mostra que o método proposto supera outros métodos baseados em características de cor e textura [Zhu et al., 2015].

Uma CNN é desenvolvida para classificar lesões de pele a partir de imagens de diferentes resoluções, as quais são usadas na etapa de treinamento por Kawahara e Hamarneh (2016). Para isso, os autores propõem o uso de uma CNN pré-treinada em um conjunto de imagens de resolução única e, em seguida, a utilizam como entradas de múltiplas resoluções. A solução proposta supera outras abordagens quando comparada com um conjunto de dados público de lesões de pele.

Em Halicek et al. (2017) um algoritmo CNN é desenvolvido para classificar o câncer de cabeça e pescoço usando imagens de amostras de tecido hiperespectral. Os resultados da CNN foram validados por anotação manual de um patologista. Os resultados preliminares de 50 pacientes indicam o potencial do uso de aprendizagem profunda para a rotulagem automática de tecidos de espécimes cirúrgicos de pacientes de cabeça e pescoço.

Uma avaliação do impacto do tamanho do conjunto de dados usado no treinamento de uma CNN é feita em [Cho et al., 2015]. Uma CNN é implementada para classificar imagens de tomografia computadorizada (TC) em seis classes. As imagens usadas foram fornecidas pelo Massachusetts General Hospital. São utilizados seis tamanhos diferentes de conjunto de dados: 5, 10, 20, 50, 100 e 200. Os resultados dos experimentos apontam que apesar de que o aumento no tamanho do conjunto de dados de treinamento leva a um melhor desempenho, existe uma tendência de que a partir de um certo ponto, essa melhora não é tão significativa.

No trabalho apresentado por Ridell e Spett (2017), uma rede Inception (GoogLeNet) previamente treinada é usada para classificar lesões da pele em malignas ou benignas. Assim como em Cho et al. (2015), a rede é inicialmente treinada novamente com vários conjuntos de diferentes tamanhos. Após isso, experimentos foram realizados a fim de verificar se o tamanho de conjunto de dados usado melhora o desempenho da rede. Os resultados encontrados confirmam essa hipótese. Era esperado encontrar um ponto no qual o aumento no número de dados usados não levaria a um aumento considerável no desempenho [Cho et al., 2015]. Entretanto tal ponto não foi encontrado, o que sugere que quanto maior o tamanho do conjunto de dados usado melhor.

Em Brestel et al. (2018) é desenvolvida uma CNN baseada no modelo inception para

a detecção de anomalias em radiografias da região do tórax. As anomalias consideradas são: opacidades pulmonares focais, opacidade pulmonar difusa, cardiomegalia e proeminência hilar anormal. Na comparação de resultados realizadas, a rede apresentou resultados ligeiramente melhores que os radiologistas, sugerindo um bom desempenho da ferramenta proposta.

No trabalho desenvolvido em Wang et al. (2016), os autores desenvolvem CNNs para a identificação de lesões cancerígenas em estágio de metástase na mama. Os resultados desse trabalho para a Camelyon Grand Challenge 2016 ([Bejnordi et al., 2017]) mostraram desempenho próximo do do nível humano. Foram avaliados quatro CNNs baseados nos modelos GoogLeNet, AlexNet, VGG16 e a rede para reconhecimento de faces apresentada [Wang et al., 2015].

Existem alguns trabalhos na área de processamento de imagens que são baseados em termogramas da tireoide [Gonzáles, 2017, Mahajan and Madhe, 2014], no entanto, nenhum deles aplica CNNs para classificação. Neste trabalho, espera-se que as áreas encontradas pela metodologia proposta apresentem alguma correspondência com as posições apontadas pelos especialistas como pertencentes a nódulos.

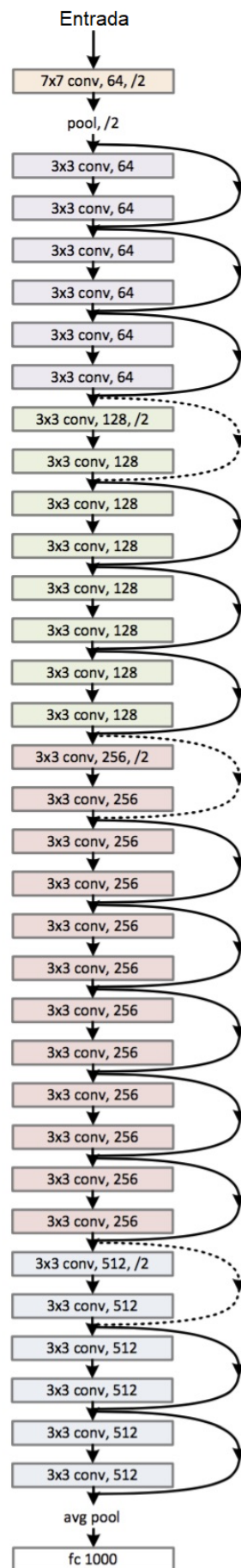


Figura 3.4: Estrutura da rede ResNet proposta em [He et al., 2016].



Figura 3.5: Estrutura da rede GoogLeNet proposta em [Szegedy et al., 2015].



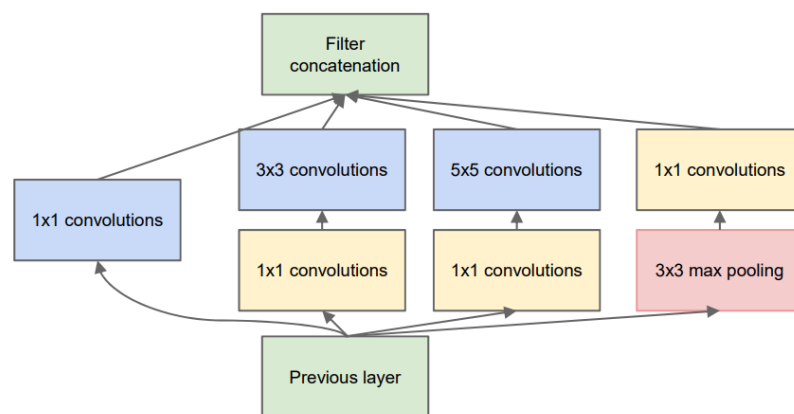


Figura 3.6: Estrutura de um módulo inception proposta em [Szegedy et al., 2015].

# Capítulo 4

## Análise termográfica

Na literatura médica é possível encontrar diversos estudos que sugerem a aplicabilidade de termografias no processo de diagnóstico de doenças relacionadas à tireoide, especialmente na detecção de nódulos [Alves et al., 1988, Aweda et al., , Gavrioloia et al., 2009, González et al., 2017, Mahajan and Madhe, 2014]. Para realizar a identificação de nódulos tireoidianos primeiro deve-se mostrar que em termografias há algum parâmetro quantitativo relativo a temperaturas que indique alguma diferença entre as regiões referentes a nódulos e as regiões consideradas saudáveis.

Nódulos tireoidianos podem apresentar diferentes características, sendo que algumas delas podem influenciar diretamente na temperatura da região. De forma geral, nódulos benignos tendem a ser menos vascularizados, e conseqüentemente, tendem a apresentar uma temperatura mais baixa [Brioschi et al., 2010]. Por outro lado, os nódulos malignos necessitam de grande vascularização para se manter [Borchardt, 2013a]. Dessa forma, tendem a apresentar temperaturas mais elevadas e espera-se que sejam detectados mais facilmente nas termografias [Brioschi et al., 2010]. Outro ponto a ser observado é se o nódulo é sólido ou preenchido com fluido (cisto). Este fator está relacionado também a vascularização da região. Deve-se considerar o tamanho do nódulo, já que nódulos demasiadamente pequenos dificilmente impactarão na temperatura a ponto de se destacarem na termografia.

Uma análise preliminar é realizada a fim de avaliar os parâmetros quantitativos propostos em trabalhos da área médica [Brioschi et al., 2010]. Nessa análise, assim como em todo o estudo apresentado neste trabalho, utilizou-se parte do conjunto de imagens disponível no repositório de dados e imagens DTNR, o qual está sendo desenvolvido pelo grupo de pesquisa do Laboratório de Computação Visual “VisualLab” do Instituto de Compu-

tação da Universidade Federal Fluminense. Esse repositório pode ser acessado pela URL <http://visual.ic.uff.br/thyroid/>. Nela, encontram-se além das termografias capturadas, os mapas de temperatura correspondentes, imagens que definem a região de interesse (ROI), informações da localização dos nódulos e dados dos complementares do paciente, como resultados de exames anteriores, histórico e dados pessoais (dependendo do nível de acesso ao banco) [Gonzáles, 2017].

## 4.1 Representação de dados utilizada

Análises de termografias podem ser feitas considerando uma sequência de imagens, no caso de estudos baseados em séries temporais [Gonzáles, 2017], ou apenas uma imagem estática do paciente, como é feito neste estudo. Embora os dados usados neste trabalho sejam provenientes do banco DTNR, as termografias estáticas usadas aqui se referem a temperaturas capturadas pelos sensores da câmera em um único instante de tempo. Dentre as imagens da série obtidas através do protocolo dinâmico [Gonzáles, 2017], selecionou-se a terceira, que corresponde a 15 segundos após o resfriamento e início das capturas.

### 4.1.1 Imagem termográfica e mapa de temperatura

As imagens termográficas mostram a distribuição discreta de temperaturas da cena abrangida pelo quadro de captura da câmera no momento da aquisição. Essa distribuição de temperaturas é representada pela imagem  $I$  (Figura 4.1), com número de pixels definido pela resolução da câmera e consequentemente, pelo número de sensores infravermelhos do modelo de câmera utilizado. No caso da FLIR SC 620, as imagens resultantes têm tamanho 640 x 480. Os dados de  $I$  podem ser representados pela função  $I(x,y) = i$ , que retorna um tom RGB, definido por um vetor de três valores  $i=(r,g,b)$ , dada uma posição  $(x,y)$ , sendo  $x \in [0,639]$  e  $y \in [0,479]$ . O sistema de coordenadas considerado tem como posição  $(0,0)$  o extremo superior esquerdo. A função  $I(x,y) = i=(r,g,b)$  pode ser entendida como um mapeamento da temperatura capturada pela câmera para um tom definida numa escala geralmente mostrada no extremo direito da imagem, como mostrado na Figura 4.1.

A partir do quadro capturado pela câmera, uma matriz  $M$  de duas dimensões, também de resolução 640 x 480, pode ser obtida com o auxílio da ferramenta ThermoCad [Borchardt, 2013b]. Tal matriz pode ser representada pela função  $M(x,y) = m$ , que re-

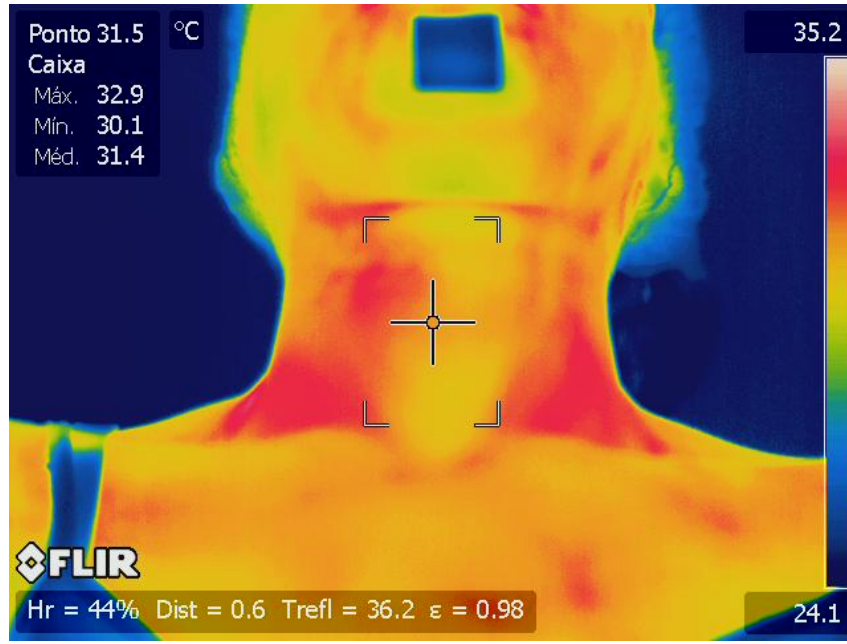


Figura 4.1: Termografia da paciente 2017-10-30

torna um valor de temperatura em graus Celsius dada uma posição  $(x,y)$ , sendo  $x \in [0, 639]$  e  $y \in [0, 479]$ . Ou seja ambas as representações I e M apresentam 480 linhas e 640 colunas.

#### 4.1.2 Imagem termográfica normalizada

Como mencionado na seção 3.1, o processo desenvolvido anteriormente pelo grupo [Fiirst, 2018, Fiirst et al., 2018] sugere uma normalização dos valores de temperatura para gerar uma imagem  $N(x,y) = n = (r=g=b = f(m))$  em tons de cinza, o que facilita a compreensão e análise de todas as imagens usadas pois em cada nível de cinza sempre se terá a mesma temperatura para todas as imagens. Para isso, após uma verificação de todas as amplitudes térmicas das imagens adquiridas de outubro de 2017 pelo projeto, os valores de temperatura mínima (min) e máxima (max) usadas foram fixados em  $20^{\circ}\text{C}$  e  $40^{\circ}\text{C}$  [Fiirst, 2018]. Assim a imagem  $N(x,y)$  pode ser representada pela função  $f(m)$  que dado um ponto de coordenadas  $(x,y)$  retorna um valor  $n$  entre 0 e 255, que corresponde a intensidade do ponto nessa imagem em tons de cinza. Dessa forma,  $N$  é calculado a partir das temperaturas  $m$  de  $M$ , de forma que seus valores de intensidade são definidos por:

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

sendo  $C_{max}$  a constante de valor do tom máximo possível na imagem, ou seja 255. A

Figura 4.2, mostra a imagem resultante ao aplicar essa normalização na termografia da Figura 4.1.

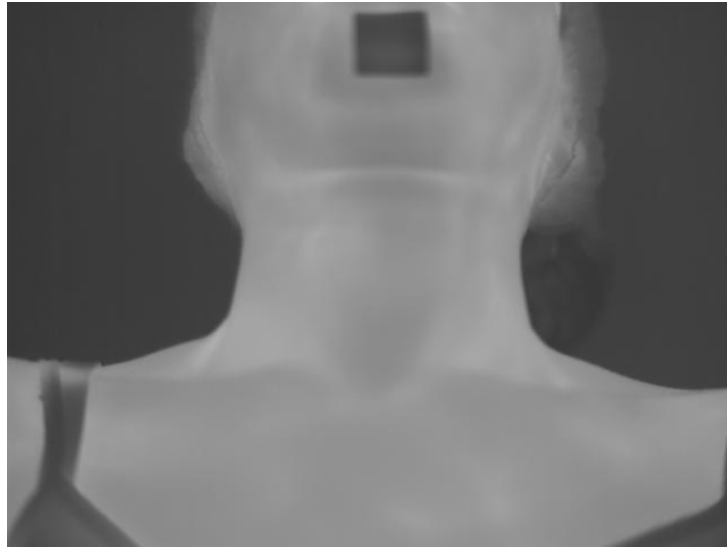


Figura 4.2: Termografia normalizada da paciente 2017-10-30

### 4.1.3 ROI

As termografias normalizadas ( $N$ ) são processadas de forma a definir uma região de interesse (region of interest - ROI) que considere apenas a área referente à tireoide do paciente (Figura 4.3). Na ROI são mantidos os pontos de  $N$  que estão na área que corresponde a região da tireoide, mas os demais pontos tem seus valores de intensidade definidos como 0, passando a fazer parte da região de fundo (na cor preta). Pelos valores mínimos de temperatura usados no mapeamento de normalização não haverá pontos com esse tom na região interior da ROI.

O processo de segmentação da ROI identifica a parte mais estreita da região que define o corpo da paciente na imagem. Ou seja, para cada linha da imagem, verifica-se onde a região não-fundo (com intensidade diferente de 0, isto é, não-preto) começa e termina. A linha em que o fim e o início da região não-fundo estejam mais próximas é considerada como pertencente à ROI. A partir dessa linha, regiões acima, abaixo, e nas laterais são incluídas na ROI, usando limites definidos pelo algoritmo usado, descrito em [Fiirst, 2018]. Esses limites foram obtidos através de observações e estudos das aquisições feitas [González et al., 2017, Fiirst, 2018]. A ROI resultante da termografia normalizada da Figura 4.2, é mostrada na Figura 4.3.

Para fins práticos, neste trabalho o termo  $ROI_p$  refere-se ao conjunto de pontos per-

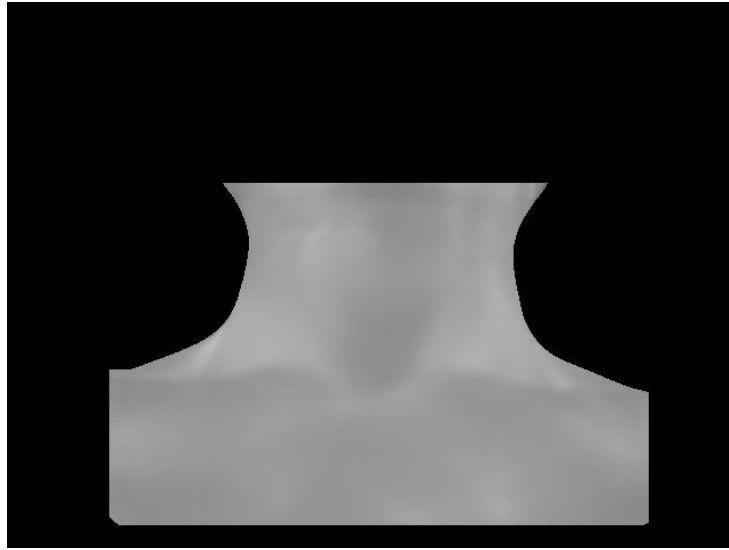


Figura 4.3: ROI da paciente 2017-10-30

tencientes à área da ROI, desconsiderando o fundo (área em preto). Ou seja,  $ROI_p$  consiste no conjunto de pontos para os quais  $ROI(x, y) \neq 0$ . Define-se  $n_p$  como o número de pontos  $p = (x, y)$  para pertencentes a  $ROI_p$ , isto é,  $n_p$  denota a cardinalidade do conjunto  $ROI_p$ .

#### 4.1.4 Grupo de pacientes estudado

Para verificar se há características que permitem diferenciar pacientes com nódulos malignos, benignos e pacientes saudáveis, foi definido um grupo de pacientes para os quais alguns parâmetros propostos na literatura médica [Brioschi et al., 2010] foram extraídos a partir de seus mapas de temperatura e imagens da ROI. Para esse estudo da relevância estatística das características propostas na literatura, foi usado um subconjunto dos dados do repositório, formado por dezoito (18) pacientes que apresentavam nódulos malignos; um conjunto de oitenta e cinco (85) pacientes com nódulos benignos; e outro conjunto de dois (2) pacientes saudáveis, chamado grupo de controle.

O grupo de controle é formado por uma mulher e um homem que não apresentavam nenhuma anomalia na tireoide. Esse grupo não consta no DTNR, e suas termografias foram capturadas a parte, usando o mesmo protocolo proposto em [Fiirst, 2018]. Os pacientes do grupo de controle são referidos neste estudo como GC1 e GC2.

O grupo de pacientes com nódulos benignos é formado por 6 homens e 118 mulheres, com idades entre 18 e 77 anos. Já o grupo de pacientes que apresentavam nódulos malignos é formado por mulheres com idades entre 17 e 84 anos.

Os pacientes que apresentam nódulos são identificados por uma sequência de números, que se refere a data da captura. Nos casos em que mais de uma captura tenha sido realizada na mesma data, um dígito a mais é acrescentado no fim da sequência, identificando a ordem de captura de cada paciente no dia.

#### 4.1.5 Análise de parâmetros estatísticos simples

A fim de verificar se as temperaturas na região da tireoide são influenciadas pela presença de nódulos, sejam eles malignos ou benignos, os seguintes parâmetros estatísticos foram calculados para cada ROI: temperaturas mínima, média, máxima e mediana.

Os valores de temperatura máxima ( $max_{ROI}$ ) e mínima ( $min_{ROI}$ ) da ROI correspondem ao maior e menor valor no mapa de temperatura nas posições correspondentes aos pontos pertencentes a  $ROI_p$ .

$$max_{ROI} = max(M(x, y)) \mid p = (x, y) \in ROI_p$$

$$min_{ROI} = min(M(x, y)) \mid p = (x, y) \in ROI_p$$

A característica  $media_{ROI}$  refere-se à média aritmética de todos os valores no mapa de temperatura dos pontos  $p = (x, y)$  que pertencem a  $ROI_p$ , sendo definida pela equação 4.1.

$$media_{ROI} = \frac{1}{n_p} \sum_{j=1}^{n_p} M(x, y) \mid p = (x, y) \in ROI_p \quad (4.1)$$

A mediana, por sua vez, é calculada ao ordenar os valores de temperatura do mapa dessas mesmas posições e selecionar o valor central dessa lista de valores.

Estudos anteriores [Aweda et al., ] considerando outras partes do corpo, mostram que os valores desses parâmetros são diferentes para regiões saudáveis e regiões que apresentem alguma anomalia, o que sugere que os valores destes parâmetros para tireoides que apresentam nódulos malignos sejam diferentes dos valores referentes a tireoides saudáveis [Brioschi et al., 2010] e de tireoides com nódulos benignos.

Para verificar as hipóteses de que esses grupos sejam identificados por essas características, estes parâmetros foram calculados para as termografias do conjunto de pacientes

aqui estudados e testados pelo teste de significância comentado no capítulo anterior. A média dos valores desses parâmetros para cada conjunto pode ser visto na Figura 4.4.

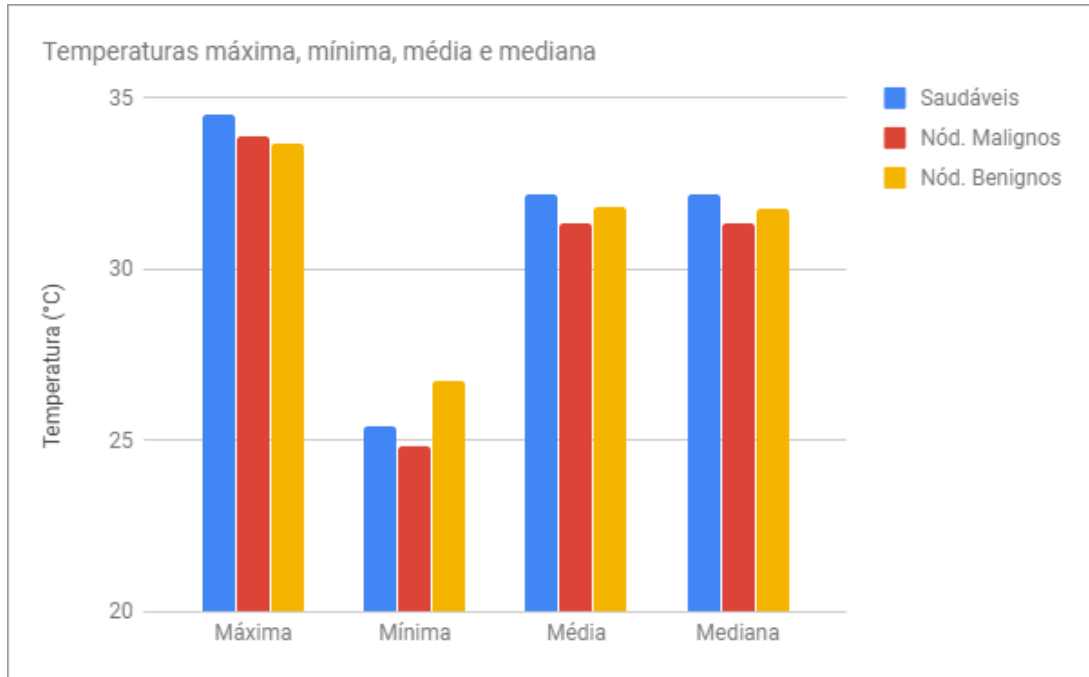


Figura 4.4: Média de cada parâmetro para cada grupo de pacientes

Para a análise desses grupos, foi utilizado o teste Wilcoxon bicaudal para amostras independentes com nível de confiança 95% (ver 3.1.4), usando os resultados de cada um desses parâmetros a fim de verificar se os valores obtidos são significativamente diferentes para combinações dos 3 grupos de pacientes : saudáveis, com nódulos malignos e com nódulos benignos. Como mencionado na seção 3.1.4, o teste de Wilcoxon é um teste utilizado para analisar significância entre dados não-paramétricos que não apresentam uma distribuição conhecida [Whitley and Ball, 2002].

O teste de Wilcoxon usa como dados de entrada dois conjuntos de dados a serem analisados, e fornece como saída valores  $W$  com os quais se obtém o  $p$ -valor que indica o quão significativamente se pode considerar as hipóteses nula e alternativa definidas. A Hipótese Nula  $H_0$  usada para cada parâmetro é que não há diferença nos dados desse parâmetro entre os grupos analisados. A Hipótese Alternativa  $H_A$  usada foi que há diferença entre os dados de cada grupo. Sendo 95% o nível de confiança utilizado, considera-se o valor  $\alpha = 0,05$ , ou seja, apenas os  $p$ -valores menores que  $\alpha$  comprovam  $H_A$ . Os  $p$ -valores encontrados nas análises considerando cada par de grupos estão na Tabelas 4.1, 4.2 e 4.3.

Os  $p$ -valores encontrados indicam que as temperaturas mínima, máxima, média e mediana não podem ser utilizadas para diferenciar pacientes saudáveis dos que apresentam



Tabela 4.1: p-valores para os parâmetros estatísticos dos grupos de controle e com nódulos malignos

Medida	Máxima	Mínima	Média	Mediana
p-valor	0,28	0,52	0,17	0,17

Tabela 4.2: p-valores para os parâmetros estatísticos dos grupos de controle e com nódulos benignos

Medida	Máxima	Mínima	Média	Mediana
p-valor	0,19	0,98	0,39	0,39

Tabela 4.3: p-valores para os parâmetros estatísticos dos grupos com nódulos benignos e com nódulos malignos

Medida	Máxima	Mínima	Média	Mediana
p-valor	0,52	0,99	0,80	0,77

nódulos malignos ou benignos considerando apenas os casos deste estudo.

Para complementar esta análise, calculou-se o tamanho do efeito para essas estatísticas considerando os mesmos grupos. Tal teste foi aplicado devido ao número pequeno de elementos no grupo de controle. Os resultados podem ser vistos nas Tabelas 4.4, 4.5 e 4.6.

Tabela 4.4: TDEs para os parâmetros estatísticos dos grupos de controle e com nódulos malignos

Medida	Máxima	Mínima	Média	Mediana
TDE	0,88	0,62	1,06	1,12

Tabela 4.5: TDEs para os parâmetros estatísticos dos grupos de controle e com nódulos benignos

Medida	Máxima	Mínima	Média	Mediana
TDE	0,78	1,35	0,34	0,39

Tabela 4.6: TDEs para os parâmetros estatísticos dos grupos com nódulos benignos e com nódulos malignos

Medida	Máxima	Mínima	Média	Mediana
TDE	0,28	2,33	0,54	0,52

Considerando agora os valores encontrados para TDE, apesar dos valores para temperatura mínima na análise dos grupos saudável X grupo de controle e malignos X benignos serem considerados altos, nenhuma das estatísticas demonstram uma alta diferenciação para todos os grupos.

Assim, o uso de apenas esses parâmetros para a identificação de regiões nodulares neste estudo está descartado, embora trabalhos similares os tenham utilizados [Aweda et al., ].

## 4.2 Índice termográfico

Há ainda mais um parâmetro baseado na ROI a ser estudado: o índice termográfico (TI). O TI foi inicialmente definido por [Collins et al., 1974]. Ele se baseia em regiões chamadas áreas isotérmicas, referidas como ISO. Para encontrar essas áreas em cada  $ROI$ , as temperaturas em  $M(x,y)$  dos pontos pertencentes a  $ROI_p$ , são separadas em intervalos de  $0,50^\circ\text{C}$  iniciando em  $min_{ROI}$  e finalizando em  $max_{ROI}$ . Para cada um desses intervalos há uma área isotérmica correspondente, que é definida pelos pontos cuja temperatura se encontra nesse intervalo. Assim, tem-se um conjunto de áreas isotérmicas, denotado como ISO, definido por:

$$\begin{aligned} \text{ISO} = \{ \\ & ISO_1 = M(x,y) \mid p = (x,y) \in ROI_p \quad min_{ROI} < M(x,y) < min_{ROI} + 0,5 \\ & ISO_2 = M(x,y) \mid p = (x,y) \in ROI_p \quad min_{ROI} + 0,5 < M(x,y) < min_{ROI} + 2 * 0,5 \\ & ISO_3 = M(x,y) \mid p = (x,y) \in ROI_p \quad min_{ROI} + 2 * 0,5 < M(x,y) < min_{ROI} + 3 * 0,5 \\ & \dots \\ & ISO_{n_{intervalos}} = M(x,y) \mid p = (x,y) \in ROI_p \quad min_{ROI} + (n_{intervalos} - 1) * 0,5 < M(x,y) \\ & < max_{ROI} \\ & \} \end{aligned}$$

sendo  $n_{intervalos}$  o número de intervalos, dado por:

$$n_{intervalos} = \frac{max_{ROI} - min_{ROI}}{0,50}$$

Observe que o número de elementos em cada  $ISO_q \in ISO$ , denotado por  $|ISO_q|$ , com  $q \in [1, 2, \dots, n_{intervalos}]$ , pode ser determinado através do histograma do conjunto de temperaturas dado por  $M(x,y)$  considerando os pontos  $p = (x,y)$  que pertencem a  $ROI_p$ .

Assim, tendo o número de elementos de cada conjunto em ISO, que corresponde por definição ao índice termográfico é dado pela equação 4.2.

$$TI = \sum_{q=1}^{n_{intervalos}} D(ISO_q) |ROI_q| \quad (4.2)$$

onde  $D(ISO_q)$  é definido pelo produto do número de elementos de  $ISO_q$ , dado por  $|ISO_q|$ , e a diferença entre a temperatura do ponto  $p$  e 26, dado por  $\Delta t$ , para cada  $p$  no conjunto  $ISO_q$ , para cada  $ISO_q \in ISO$ .

$$D(ISO_q) = \prod_{i=1}^{|ISO_q|} |ISO_q| * \Delta t(p) \quad p = (x_p, y_p) \in ISO_q \quad (4.3)$$

$$\Delta t = M(x_p, y_p) - C, C = 26 \text{ [Collins et al., 1974]}$$

Para estudos similares, considerando outras partes do corpo [Brioschi et al., 2010], o uso do TI permitiu diferenciar regiões saudáveis de regiões que apresentam alguma disfunção. Deseja-se, portanto, avaliar se as ROIs referentes a tireoides que apresentam nódulos, sejam eles benignos ou malignos, têm valores diferentes de TI das ROIs referentes a tireoides saudáveis, para o grupo de pacientes analisado neste estudo. Assim, os valores TI foram calculados para cada termografia, considerando novamente os mesmos conjuntos de imagens. As médias de cada grupo são mostradas na Figura 4.5.

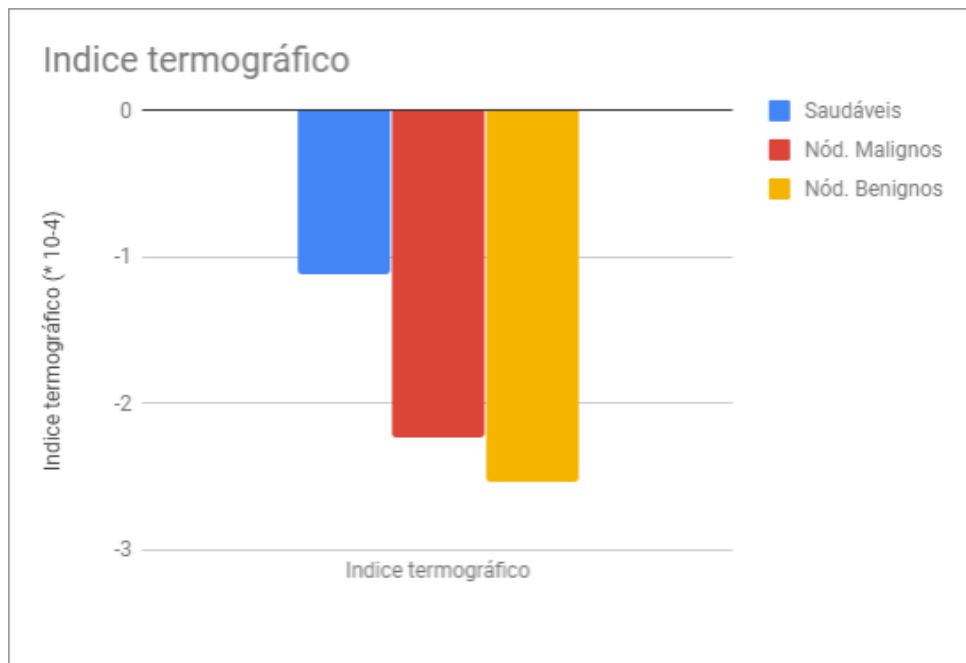


Figura 4.5: Média do Índice Termográfico para cada grupo de pacientes

Novamente realiza-se o teste Wilcoxon para esse parâmetro, a fim de avaliar se os valores de TI para cada um dos grupos são significativamente diferentes entre eles. O p-valor resultante desse teste é 0,26 para os grupos de controle e com nódulos malignos. Já para os grupos de controle e com nódulos benignos o p-valor encontrado foi de 0,97. Por fim o p-valor para o teste com os grupos com nódulos malignos e com nódulos benignos foi de 0,99. Esses valores indicam novamente que essa medida não pode ser usada para diferenciar pacientes de diferentes grupos considerando um nível de confiança de 95

Considerando agora os TDEs encontrados nos testes realizado para cada combinação de grupos, os valores encontrados estiveram abaixo de 0,001, o que pode ser considerado muito pequeno, sugerindo que o TI não é apropriado para diferenciar os grupos.

## 4.3 Parâmetros de segmentação

Pode-se ainda utilizar parâmetros para avaliar cada pixel das termografias dos pacientes que apresentam nódulos, a fim de identificar os pixels que se referem a regiões saudáveis e os que se referem a regiões nodulares. Considerando as informações presentes numa termografia, pode-se realizar diferentes análises considerando como o valor de um ponto específico se relaciona com os valores dos outros. Por exemplo pode-se avaliar o quão diferente é o valor desse ponto em relação aos demais da *ROI*, o que pode ser analisado através de agrupamentos por temperatura, como proposto neste trabalho.

Também é possível comparar pontos com outros de uma região considerada equivalente, que no caso de órgão simétricos, corresponde a mesma região no lado oposto em relação a um eixo de simetria. Essa análise pode ser feita através do parâmetro de assimetria ( $P_A$ ). Ambas abordagens foram analisadas e seus resultados são mostrados nas seções a seguir.

### 4.3.1 Agrupamento por temperatura

Para avaliar se apenas a informação da temperatura é suficiente para identificar os pontos referentes à região nodular, uma requantização em níveis de cinza da *ROI*, denominada posterização por alguns autores [Conci et al., 2008] é aplicada para cada  $ROI_p$  a fim de avaliar visualmente agrupamentos de temperatura. Ou seja, considerando que os tons nessas imagens refletem as temperaturas capturadas no mapa de temperatura, é possível definir regiões considerando pontos que apresentam tonalidades próximas. Caso haja um padrão visual de tons que correspondem às regiões nodulares, então poderia haver

um padrão de temperatura que indicaria a presença de nódulos. Tal conclusão é sugerida pela literatura médica [Brioschi et al., 2010, Mahajan and Madhe, 2014], e está de acordo com o fato de que nódulos, especialmente os malignos, geralmente apresentam grande vascularização, e portanto espera-se que tenham temperaturas mais elevadas em  $M(x,y)$  e tons que se destaquem na imagem da *ROI*. Neste trabalho propõe-se criar imagens que agrupem visualmente esse valores. Assim, foram geradas quantizações nas ROIs dos dezoito (18) pacientes com nódulos malignos e dos oitenta e cinco ( 85 ) pacientes com nódulos benignos, usando níveis de cinza 4, 8 e 16. Dessa forma, as temperaturas seriam visualmente agrupadas em 4, 8 e 16 grupos respectivamente. O resultado desse processo, para a termografia do paciente 2017-09-04 é mostrado nas Figuras 4.7 a 4.9. Essa técnica é baseada também na análise qualitativa que os médicos fazem nos exames visuais das imagens térmicas.

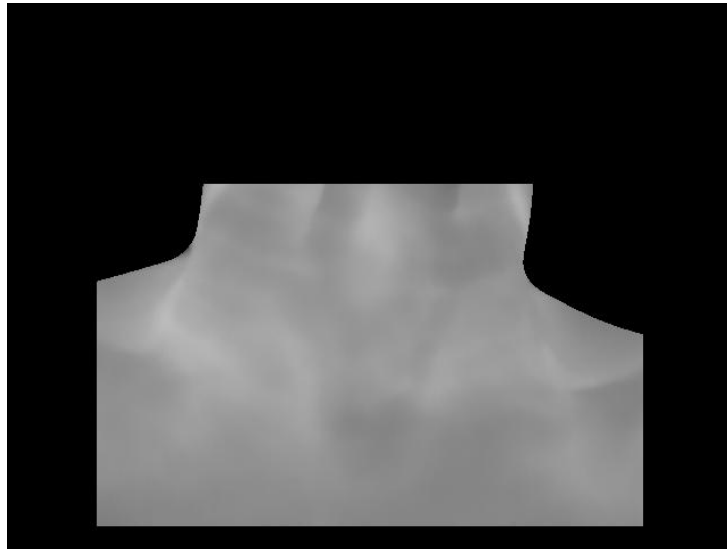


Figura 4.6: ROI do paciente 2017-09-04

Avaliando visualmente os resultados de todos os pacientes com nódulos e comparando-os com as localizações dos nódulos de cada imagem identificados pelos especialistas, observa-se que para nenhum dos agrupamentos de tons utilizados é possível definir um grupo que indique regiões que se aproximassem das identificadas pelos especialistas como nodulares para maior parte das termografias avaliadas, tanto as de nódulos malignos, quanto as de nódulos benignos. Isso sugere que apenas a informação de temperatura não é suficiente para identificar tais regiões.



Figura 4.7: Quantização em 4 tons da termografia da Figura 4.6

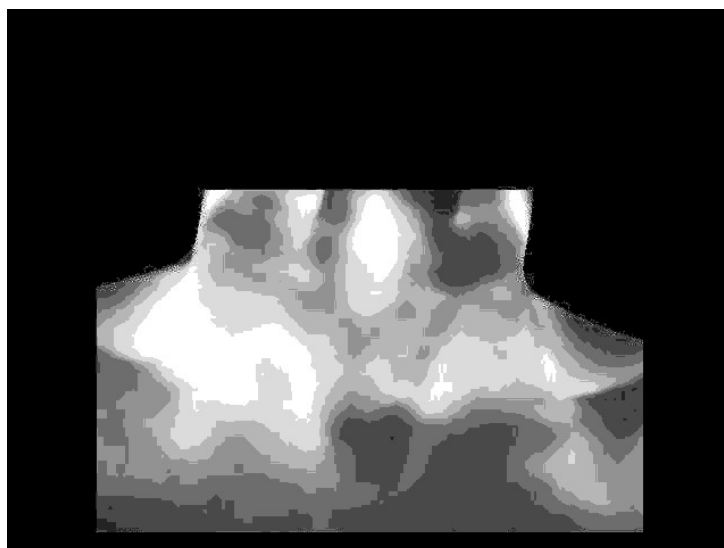


Figura 4.8: Quantização em 8 tons da termografia da Figura 4.6

### 4.3.2 Parâmetro de assimetria

Uma outra abordagem a ser considerada é o estudo das diferenças de temperatura de pontos simétricos na direção horizontal, chamado de parâmetro de assimetria ( $P_A$ ). Como já comentado, a tireoide é um órgão bastante simétrico horizontalmente, considerando a anatomia de um corpo no eixo crânio-caudal (Figura 4.10). Além disso, é um órgão superficial o suficiente para permitir a detecção de calor causado por hipermetabolismo de nódulos ou da glândula em si. Segundo [Brioschi et al., 2010], o parâmetro de assimetria é definido pela diferença de temperatura de uma região em relação ao seu lado contralateral, considerando um eixo vertical de simetria do corpo. Valores do  $P_A$  acima de  $0,3^{\circ}\text{C}$  podem

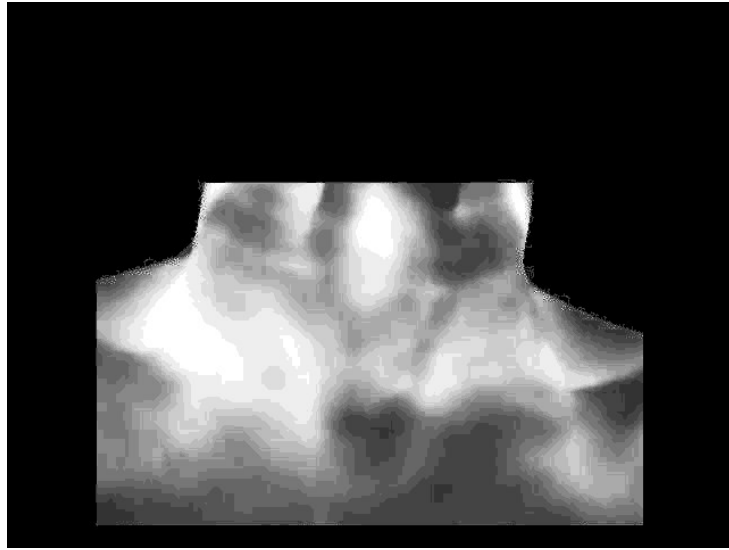


Figura 4.9: Quantização em 16 tons da termografia da Figura 4.6

indicar alguma disfunção, sendo que, de forma geral, se o valor ultrapassar  $1^{\circ}\text{C}$  algum problema significativo pode estar ocorrendo na região [Brioschi et al., 2010].

A câmera FLIR modelo SC620, usada na captura das imagens utilizadas neste estudo, tem sensibilidade de  $0,04^{\circ}\text{C}$ , o que permite detectar diferenças sutis de temperatura.

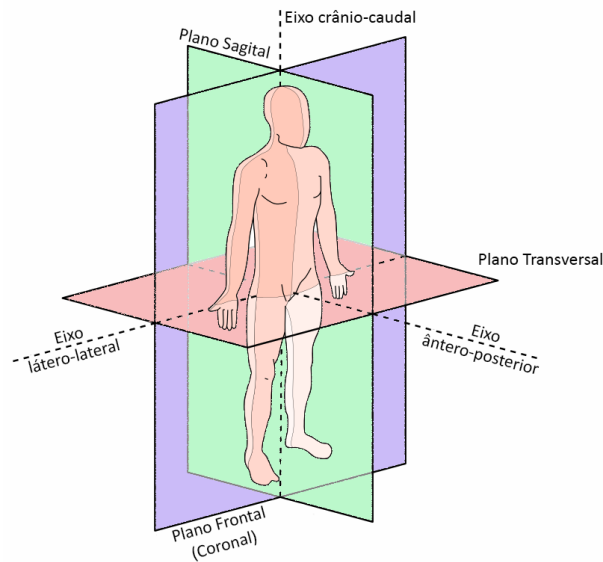


Figura 4.10: Eixos e planos anatômicos

Considerando apenas a *ROI* de cada paciente, o eixo vertical usado para o cálculo de  $P_A$  e foi estabelecido com base no centro geométrico  $O$  da *ROI*, que tem tamanho  $640 \times 480$ .  $O$  é definido pelas coordenadas  $(o_x, o_y)$  [Fiirst, 2018] :

$$O = (o_x, o_y)$$

sendo  $o_x$  e  $o_y$  definidos pela equação 4.4.

$$o_x = \frac{\sum_{u=0}^{479} \sum_{v=0}^{639} u * ROI(u, v)}{I(u, v)} \quad o_y = \frac{\sum_{u=0}^{479} \sum_{v=0}^{639} v * ROI(u, v)}{I(u, v)} \quad (4.4)$$

Assim o eixo vertical e a ser considerado no cálculo de  $P_A$  é definido pela coordenada horizontal de  $O$ :

$$e(y) = o_x$$

Após a definição do eixo vertical, o parâmetro de assimetria ( $P_A$ ) de cada ponto em  $ROI_p$  pode ser calculado. O  $P_A$  considera a diferença de temperatura de pontos na mesma altura (ou seja mesma coordenada  $y$ ) em lados opostos ao eixo  $e$ . Para realizar esse cálculo, consideramos apenas os pontos que pertencem a  $ROI_p$ , assim como seus simétricos.

O  $P_A$  de um ponto é dado por:

$$P_A(x, y) = \begin{cases} M(x, y) - M(x + (2 * o_x - x), y), & \text{se } x < o_x \text{ e } p_1 = (x, y) \in ROI_p \text{ e} \\ & p_2 = (x + (2 * o_x - x), y) \in ROI_p \\ M(x, y) - M(x - (2 * x - o_x), y), & \text{se } x \geq o_x \text{ e } p_1 = (x, y) \in ROI_p \text{ e} \\ & p_2 = (x - (2 * x - o_x), y) \in ROI_p \end{cases}$$

onde  $(x, y)$  são coordenadas de pontos pertencentes à  $ROI_p$ . Observe que o sinal do valor encontrado de  $P_A$  indica o lado mais quente. Se o valor é positivo, o ponto  $p_1$  definido pelas coordenadas  $(x, y)$  é o mais quente. Se o valor for negativo, o ponto  $p_2$  oposto a  $p_1$  é o mais quente.

Esse parâmetro se baseia no fato de que há partes da região da tireoide que são altamente vascularizadas, portanto apresentam temperaturas maiores. Estas porém geralmente estão presentes de forma praticamente simétrica em ambos os lados da tireoide, assim se for considerada a simetria para avaliar a temperatura de cada ponto, pontos de temperatura mais alta devido à anatomia normal da região não seriam considerados [Brioschi et al., 2010].

Para verificar se  $P_A$  é adequado para identificar regiões nodulares, testes estatísticos Wilcoxon, descritos na seção 3.1.4, foram aplicados. No primeiro teste, para cada um dos dezoito pacientes com nódulos malignos, todos os pixels da ROI tiveram  $P_A$  calculados.



Para cada paciente, os valores positivos de  $P_A$  são agrupados em dois conjuntos:  $P_{AN}$  que corresponde aos valores de  $P_A$  dos pontos em  $ROI_p$  que estão na região nodular; e  $P_{AS}$  que corresponde aos valores de  $P_A$  dos demais pontos em  $ROI_p$ . Apenas os valores positivos foram considerados, evitando a contagem repetida de um mesmo  $P_A$ , já que os  $P_{AS}$  de ponto simétricos têm mesmo valor absoluto, diferindo apenas no sinal.

A informação da localização dos nódulos foi fornecida por especialistas médicos, que se basearam na palpação e em dados de ultrassonografias previamente adquiridas para defini-las. O processo feito pelo especialista de identificar o nódulo no paciente está definido em [Fiirst, 2018], e consiste basicamente em, após a captura, apalpar o paciente na região indicada pelos dados da ultrassonografia, de forma a sentir alguma estrutura que possa ser caracterizada como nódulo. Caso essa estrutura seja encontrada, o especialista fixa um marcador no local, e uma nova imagem é capturada. Dessa forma, há para cada paciente uma termografia na qual marcadores indicam a posição do(s) nódulo(s).

A fim de obter as temperaturas da região nodular, a imagem da termografia com marcador é processada, de forma a gerar uma nova imagem  $C_{masc}$  binária de fundo preto com círculos brancos de 35 pixels de diâmetro nas regiões onde os marcadores são posicionados. Esse valor de diâmetro foi escolhido porque leva a círculos de tamanho próximo aos marcadores usados. Um exemplo do resultado desse processo pode ser visto na Figura 4.11.

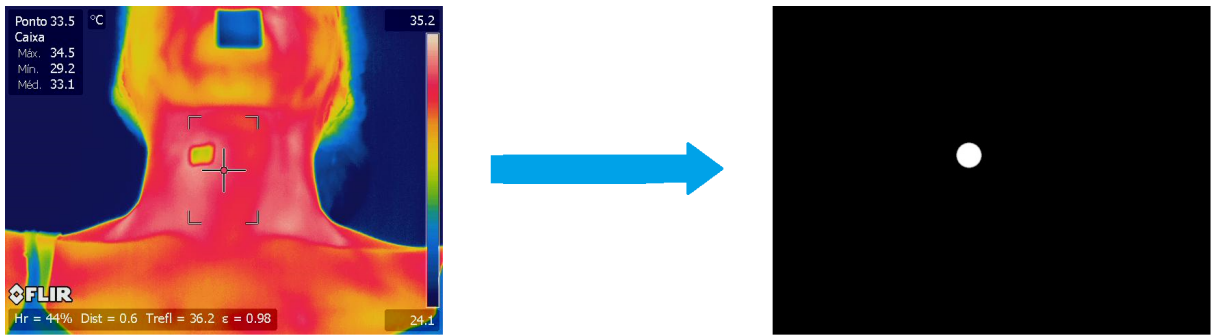


Figura 4.11: Imagem  $C_{masc}$  e termografia original com marcador

Dessa forma  $P_{AN}$  e  $P_{AS}$  são definidos como:

$$P_{AN} = P_A(x,y) \mid p(x,y) \in ROI_p \text{ e } C_{masc}(x, y) = 1$$

$$P_{AS} = P_A(x,y) \mid p(x,y) \in ROI_p \text{ e } C_{masc}(x, y) = 0$$

Assim, é verificado, com o teste de Wilcoxon, se os valores de  $P_A$  das regiões nodulares ( $P_{AN}$ ) são significativamente superiores em relação aos das demais regiões ( $P_{AS}$ ). Para cada paciente do grupo maligno os p-valores do teste são apresentados na Tabela 4.7.

Tabela 4.7: p-valores para os testes Wilcoxon para cada paciente do grupo maligno, considerando  $P_A$

Paciente	p-valor
2017-09-04	$2,64 * 10^{-8}$
2017-09-05-1	$6,15 * 10^{-6}$
2017-09-05-2	$1,22 * 10^{-9}$
2017-09-11-1	$2,66 * 10^{-2}$
2017-09-11-2	$9,00 * 10^{-12}$
2017-09-12-1	$7,45 * 10^{-2}$
2017-09-12-2	$1,04 * 10^{-1}$
2017-09-12-3	$1,12 * 10^{-1}$
2017-09-12-4	$3,04 * 10^{-4}$
2017-09-18	$9,65 * 10^{-2}$
2017-09-25	$1,11 * 10^{-3}$
2017-10-03-1	$7,36 * 10^{-2}$
2017-10-03-2	$4,89 * 10^{-3}$
2017-10-03-3	$9,60 * 10^{-1}$
2017-10-10-1	$8,49 * 10^{-5}$
2017-10-10-2	$5,36 * 10^{-2}$
2017-10-10-3	$1,11 * 10^{-2}$
2017-10-30	$1,37 * 10^{-4}$

Observe que para a maioria dos casos, os p-valores concordam, ou se aproximam da hipótese de que  $P_A$  pode ser utilizado na detecção de regiões referentes a nódulos malignos. Dessa forma, o  $P_A$  foi o parâmetro escolhido para detecção de regiões nodulares malignas neste trabalho.

Entretanto, nesse primeiro teste há casos em que o p-valor se distancia bastante do valor que comprovaria a hipótese, que seria  $> 0,05$ , como comentado na seção 3.1.4. Segmentações usando  $P_A$ , podem ser feitas para avaliar melhor o comportamento desse parâmetro nos casos aqui estudados, a fim de verificar se a partir dele pode-se encontrar regiões nodulares.

De forma similar, no segundo teste para cada um dos oitenta e cinco pacientes com nódulos benignos, todos os pixels da ROI tiveram seus  $P_A$  calculados e agrupados nos conjuntos  $P_{AN}$  e  $P_{AS}$ , como feito no teste anterior. Observa-se que apenas 7 dentre as 85 termografias apresentaram p-valor que concorda com a hipótese de que as regiões nodulares apresentam maior temperatura que o resto da ROI, lembrando que para confirmar essa hipótese o p-valor deveria ser menor que 0,05. Esse resultado está de acordo com o esperado, já que nódulos benignos não apresentam características que levariam a um aumento de temperatura [Rosário et al., 2013].

## 4.4 Metodologia proposta

Neste trabalho, propõe-se uma metodologia para encontrar áreas em termografias que apresentem alguma correspondência com nódulos indicados pelos especialistas. Para isso uma sequência de passos foi definida, baseando-se no estudo prévio realizado no capítulo 3, e em trabalhos anteriores. A Figura 4.12 mostra a metodologia proposta, que pode ser agrupada em 3 etapas: pré-processamento (Figura 4.13), segmentação dos possíveis nódulos (Figura 4.14) e classificação das regiões segmentadas (Figura 4.15).



Figura 4.12: Metodologia proposta

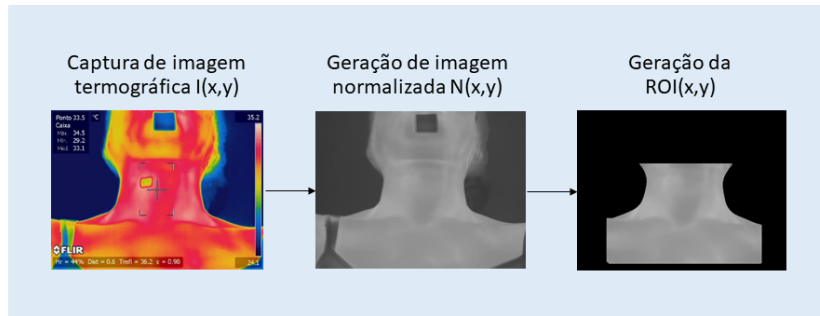


Figura 4.13: Fase de pré-processamento

Tem-se inicialmente a imagem termográfica  $I$  de cada paciente capturado através do protocolo utilizado nos exames [González, 2017]. A partir dessa imagem, o mapa de temperatura ( $M$ ), a  $ROI$  e a  $ROI_p$  são obtidos. Tendo previamente analisado e comparado as características a serem extraídas das termografias, definiu-se o uso do parâmetro de assimetria ( $P_A$ ) para realizar uma segmentação a fim de encontrar regiões referentes a nódulos. Para isso, o processo já definido na seção 4.3.2, foi realizado. Assim, para cada paciente, o eixo  $e$  foi definido e os valores de  $P_A$  para  $ROI_p$  foram calculados. Após isso, a segmentação é realizada: os pontos de  $ROI_p$  com  $P_A$  acima de um limiar  $\ell$  são considerados como pertencentes a uma região possivelmente nodular.

No processo de segmentação, há regiões encontradas que se não se referem a nódulos. Isso se deve a particularidades da própria anatomia da região, visto que dobras na pele,

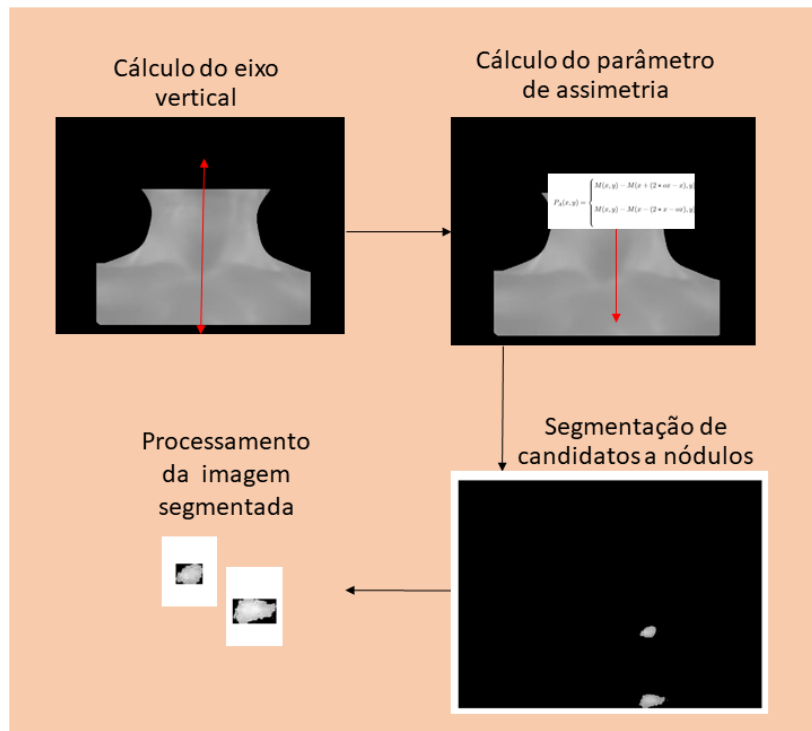


Figura 4.14: Fase de segmentação

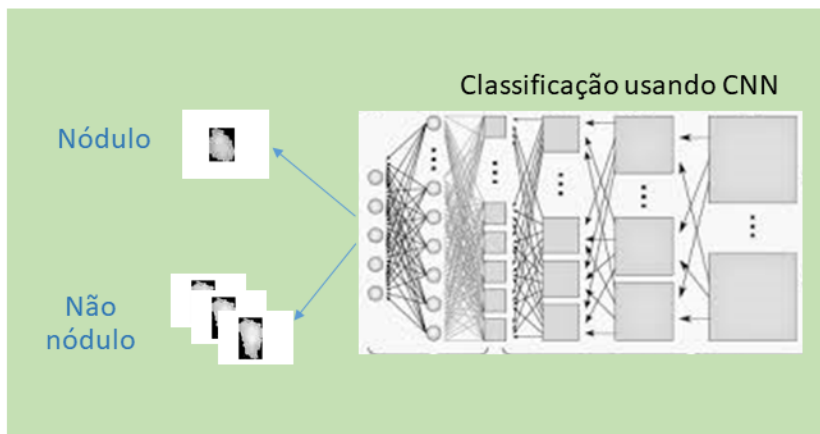


Figura 4.15: Fase de classificação

vasos sanguíneos e outras formações podem apresentar temperaturas alteradas. Além disso, a área indicada na ROI não é totalmente simétrica, especialmente nas bordas, devido a posição da paciente e limitações no processo. Os candidatos a nódulo que são resultantes desses fatores tendem a ter forma alongada e orientação próxima à vertical. Para eliminar esses falsos nódulos, pode-se aplicar a operação morfológica de abertura, que consiste na combinação das operações de erosão e dilatação [Sonka et al., 2014].

Na operação de erosão, o elemento estruturante é sucessivamente posicionado de forma a ficar centralizado em cada um dos pixels da imagem. Para cada um desses passos,

seleciona-se na imagem os pixels que estão na área em que o elemento estruturante está posicionado. Esses pixels têm seus valores de intensidade alterados para a intensidade mínima dentre eles. A operação de dilatação é executada de forma similar, com a diferença de que os pixels selecionados têm suas intensidades alteradas para a intensidade mínima dentre eles [Sonka et al., 2014].

Observe que caso se opte por usar um elemento estruturante maior, as regiões alongadas descritas anteriormente serão apagadas durante a aplicação da operação de abertura. Dessa forma, após esse processamento, os possíveis nódulos restantes de acordo com a segmentação seriam apenas os de formas mais arredondadas, as quais apresentam aspecto mais semelhante a nódulos reais. Os resultados alcançados através do processo de segmentação são detalhados na seção 5.1.

Para utilizar essas imagens como entrada para os algoritmos CNNs, um segundo processo de segmentação é feito através da definição de bounding boxes que abrangem cada uma das regiões restantes na imagem resultante da segmentação anterior baseada em  $P_A$ .

Por fim, essas regiões são então classificadas como nodulares ou não utilizando um algoritmo CNN. Uma comparação dos resultados alcançados com cada variação de CNN considerada é mostrada na seção 5.2.

Para alcançar um bom desempenho com os algoritmos CNNs, é necessário utilizar um volume razoável de dados de treinamento [Goodfellow et al., 2016]. Entretanto, neste estudo, assim como na área médica em geral, o número de imagens disponíveis para análise é bem restrito. Isso se deve às dificuldades em se obter dados dessa natureza, sendo inclusive inviável em alguns casos [Shaikhina and Khovanova, 2017]. Uma alternativa a ser considerada nesse contexto é a geração de dados sintéticos, aumentando assim o número de dados para treinamento. Diversas técnicas podem ser aplicadas no conjunto de imagens iniciais a fim de gerar novas imagens que também possam ser utilizadas como entrada [Shaikhina and Khovanova, 2017].

Porém, é importante considerar se o processamento aplicado nessas imagens não resultaria em distorções que comprometem características importantes para a definição de padrões a serem utilizados na tarefa de classificação realizada pelos algoritmos CNNs. Dessa forma, neste estudo em particular, não seria interessante aplicar operação de escala por exemplo, feita em diversos trabalhos, já que o tamanho da anomalia pode ser um indicativo da presença de nódulos. Também o processo de inclusão de ruído nas imagens, proposto em algumas aplicações de CNN, acrescenta informações que não correspondem

à presença ou não de nódulos, o que pode dificultar na identificação de padrões.

O fato de que a estrutura da tireoide apresenta grande simetria contralateral, propicia o uso da operação de espelhamento para geração de dados sintéticos. Se considerarmos que também os nódulos na tireoide podem apresentar aparência similar quando estão posicionados tanto no lóbulo direito quanto no esquerdo, apresentando inclusive inclinações e formatos similares segundo o lado em que estão, o espelhamento em torno de um eixo vertical é um ideia bem interessante a ser aplicada.

Uma outra transformação que a princípio não impactaria na alteração de características determinantes na classificação é a mudança de tons nas imagens das regiões. Após as segmentações usando PA, as regiões que estão sendo analisadas apresentam temperaturas similares em cada caso, já que elas foram selecionadas justamente por terem tons acima de um limiar. Elas também tem seu valor máximo definido pelo processamento aplicado em [Fiirst, 2018]. Assim sendo, a esse estágio do processo, o que se tem são diferentes regiões com temperaturas similares mas com outras características que as diferenciam, ou seja, a temperatura em si não permite diferenciar nódulos de outras regiões que também apresentam aumento de temperatura. Portanto, modificar igualmente as tonalidades das imagens analisadas tanto para as que se referem a nódulos, quanto às referentes a outras anomalias, não impactaria a princípio no processo de classificação.

O processo aplicado de mudança de tons consiste em usar uma mudança de escala ou alteração da amplitude das temperaturas, e uma em uma translação de seus valores, mapeando os valores de intensidade de cada pixel da imagem para novos valores definidos em outro intervalo diferente do original. Dessa forma, tendo originalmente uma imagem de candidato a nódulo com seus tons definidos no intervalo 100 como mínimo e 155 como máximo numa escala de cinza, a translação consistiria em atribuir novos valores de intensidade definidos num intervalo com mesma amplitude e escala, mas iniciando ou finalizando em outros valores limites, como por exemplo de 0 a 55, limitando pelo valor mínimo e ou de 200 a 255 limitando pelo valor máximo.

Observe que a operação resulta em possíveis deslocamentos e alargamentos ou estreitamentos da amplitude do histograma da imagem original, dependendo dos novos valores limites definidos.

# Capítulo 5

## Resultados

Dada a metodologia proposta no capítulo 4, os resultados dessa metodologia para o grupo de pacientes analisado, descrito na seção 4.1.4, são avaliados a seguir, considerando as variações testadas das fases de segmentação e classificação.

### 5.1 Resultados da segmentação usando $P_A$

Apresenta-se a seguir os resultados para o processo de segmentação usando diferentes valores de limiar  $\ell$ . São utilizadas na análise, ao todo, 18 termografias de diferentes pacientes, alguns deles apresentavam mais de um nódulo, totalizando 25 nódulos. Na Tabela 5.1 estão listados os pacientes com nódulos malignos e o número de nódulos que eles apresentam. Todos esses nódulos tinham grandes indícios de malignidade.

Dos pacientes com nódulos benignos, dos 124 nódulos apresentados pelos 85 pacientes, apenas 18 foram encontrados. Isso sugere a não aplicabilidade dessa técnica para esses casos.

Para cada imagem analisada, o processo descrito na seção 4.3.2 foi realizado. A  $ROI_p$  de cada imagem teve seu eixo  $e$  definido e a partir dele o  $P_A$  de cada ponto de  $ROI_p$  foi calculado. Os pontos com  $P_A$  acima de um limiar definido são considerados como pertencentes a uma região possivelmente nodular. Os valores utilizados como limiar são os definidos na Tabela 5.2.

O uso de valores positivos como limiar garante que os pontos incluídos na região tenham temperatura superior aos posicionados no lado oposto. O resultado das segmentações foi avaliado por um especialista médico, que comparou visualmente esses resultados com os dados da ultrassonografia e as termografias com marcadores indicando a região

Tabela 5.1: Número de nódulos por paciente no grupo maligno

Paciente ID	Nódulos reais	Nódulos achados pela segm.
2017-09-04	1	0
2017-09-05-1	1	1
2017-09-05-2	2	1
2017-09-11-1	1	1
2017-09-11-2	2	1
2017-09-12-1	1	1
2017-09-12-2	2	1
2017-09-12-3	1	0
2017-09-12-4	2	1
2017-09-18	1	1
2017-09-25	2	2
2017-10-03-1	1	1
2017-10-03-2	2	1
2017-10-03-3	1	0
2017-10-10-1	1	1
2017-10-10-2	1	1
2017-10-10-3	2	1
2017-10-30	1	1

Tabela 5.2: Valores usados como limiar  $\ell$ 

Limiar $\ell$ (em °C)
0,31
0,55
0,78
1,02
1,25
1,48

dos nódulos. Para três pacientes não foi possível encontrar regiões que apresentassem alguma correspondência com nenhum dos nódulos apontados pelos especialistas com nenhum valor de  $\ell$  definido na Tabela 5.2. As segmentações destes pacientes, e as suas termografias com os marcadores colocados pelos especialistas são mostradas nas Figuras 5.1 a 5.6. Além disso, dentre os pacientes que apresentam mais nódulos, houve casos que um dos nódulos não foi detectado.

Nas termografias das Figuras 5.1 a 5.6, pode ser observado que as pacientes apresentam diversas regiões com temperatura mais elevada. Isso pode estar relacionado à fisiologia do paciente, às condições do paciente prévias à captura, às particularidades da anatomia do próprio paciente, ou ainda a outras anomalias na região da tireoide e adjacências,



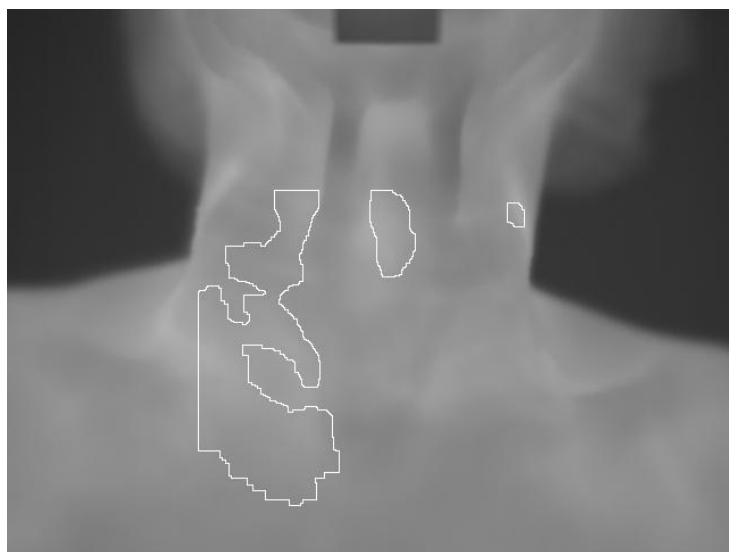


Figura 5.1: Segmentação do paciente 2017-09-04 com  $P_A = 0,31$

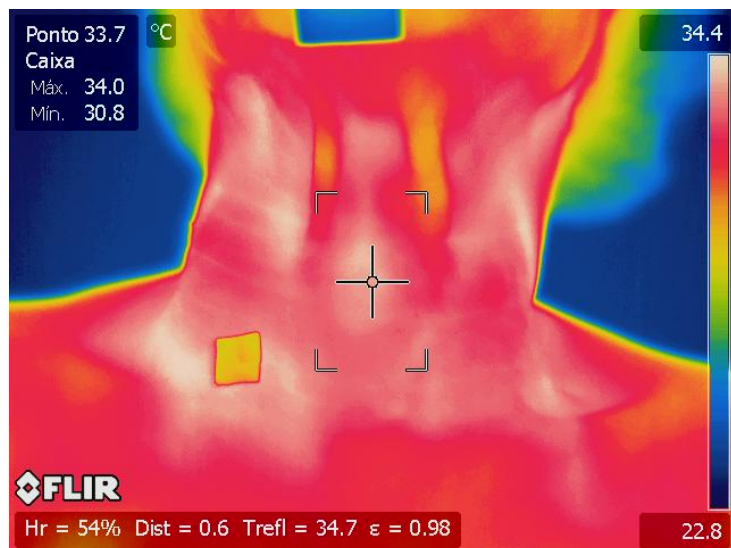


Figura 5.2: Marcação dos nódulos do paciente 2017-09-04

que podem influenciar na temperatura da região como um todo. Outro fator que pode influenciar na identificação de regiões candidatas a nódulos usando  $P_A$ , é a posição do nódulo. Quando um nódulo está muito centralizado na área do istmo, ele pode acabar apresentando grande simetria, dificultando sua detecção, como no caso mostrado nas Figuras 5.3 e 5.4.

Para as quinze imagens restantes, as segmentações com  $\ell = 0,31$  apontaram regiões que abrangiam, pelo menos parcialmente, a posição apontada pelo especialista para pelo menos um dos nódulos. O fato de que as segmentações com  $\ell = 0,31$  permitem identificar regiões relacionadas a grande parte dos nódulos, exceto para os casos de nódulos que não apresentam temperaturas diferenciadas na região oposta, está de acordo com o descrito na

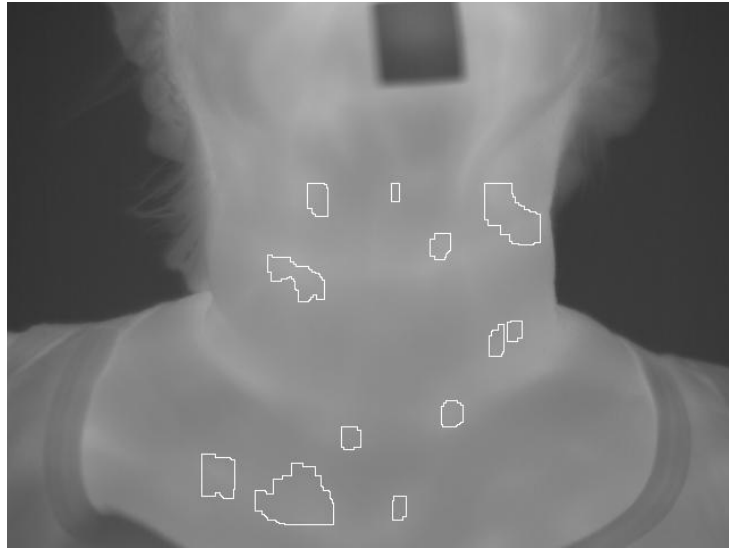


Figura 5.3: Segmentação do paciente 2017-09-12-3 com  $\ell = 0,31$

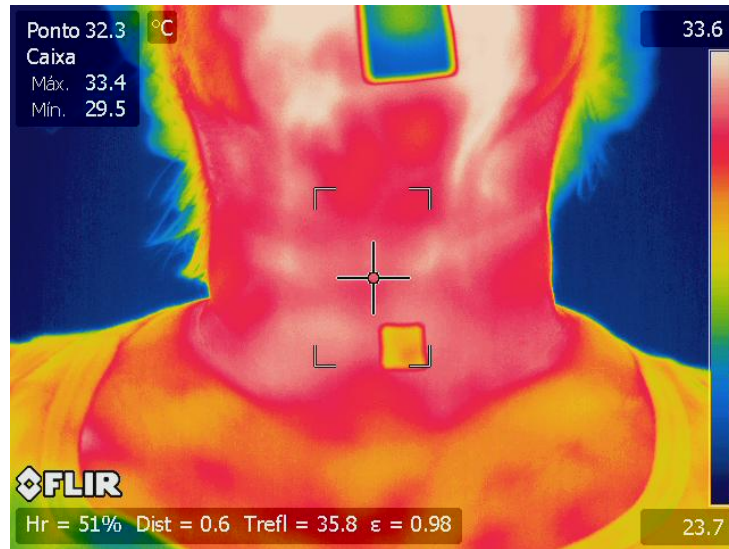


Figura 5.4: Marcação dos nódulos do paciente 2017-09-12-3

literatura médica que define que diferenças de temperatura a partir de  $0,3^{\circ}\text{C}$  já constituem indicativos de disfunções [Brioschi et al., 2010].

Entretanto, apenas para quatro (4) das termografias analisadas, as regiões que correspondiam à localização dos nódulos também apresentavam dimensões próximas às descritas nas informações da ultrassonografia, como no caso das Figuras 5.7 e 5.8. Para duas (2) das quatro termografias, a segmentação só permitiu encontrar um dos dois nódulos da paciente, como mostrado nas Figuras 5.9 e 5.10.

Para as demais onze termografias, esse valor de  $\ell$  leva a regiões que englobam tanto o nódulo quanto suas adjacências (Figuras 5.11 e 5.12), o que sugere que a presença do

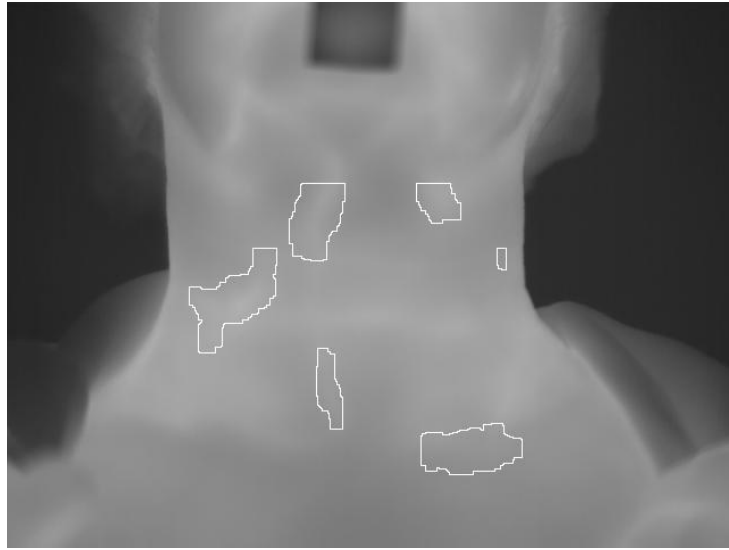


Figura 5.5: Segmentação do paciente 2017-10-03-3 com  $\ell = 0,31$

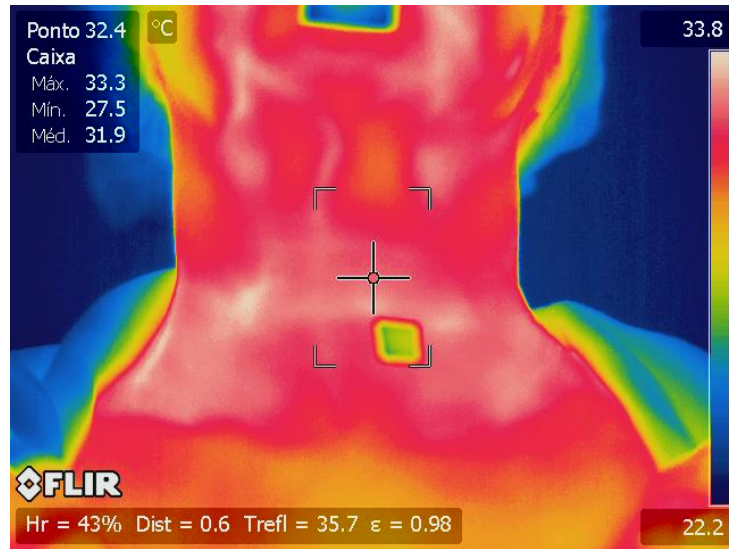


Figura 5.6: Marcação dos nódulos do paciente 2017-10-03-3

nódulo pode chegar a influenciar a temperatura de regiões próximas a ele.

Partindo do princípio de que a temperatura do nódulo influencia regiões próximas a ele, e que a região do nódulo em si apresenta temperaturas mais elevadas, espera-se que ao usar um valor mais alto de  $\ell$  na segmentação seja possível encontrar regiões com dimensões mais próximas às do nódulo.

Para as segmentações com  $\ell = 0,55$ , observa-se que para sete (7) casos a região apontada pela segmentação corresponde à localização do nódulo dada pelos especialistas e também apresenta dimensões que se aproximam das descritas pelo resultado da ultrassonografia. Entretanto, para três (3) delas, a segmentação só permite encontrar um dos dois

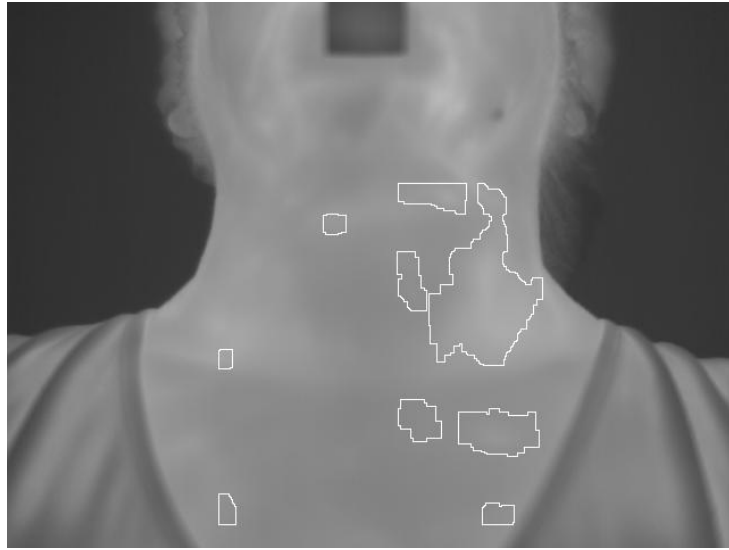


Figura 5.7: Segmentação do paciente 2017-10-10-2 com  $\ell = 0,31$

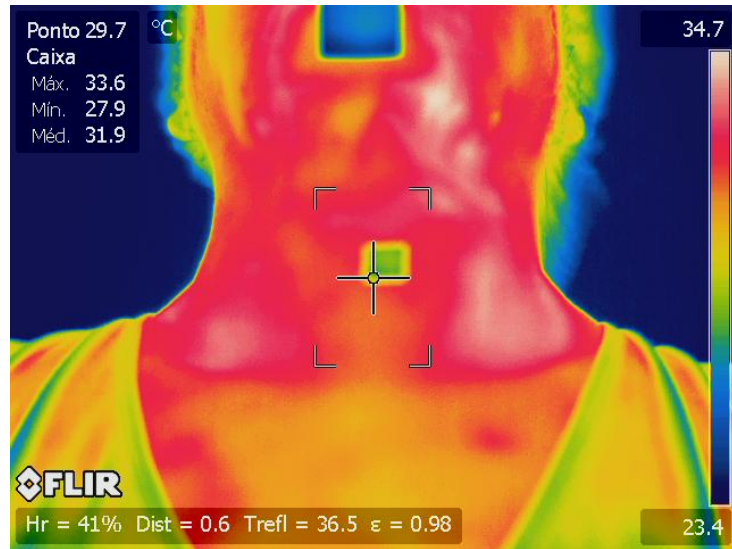


Figura 5.8: Marcação dos nódulos do paciente 2017-10-10-2

nódulos da paciente, como mostrado nas Figuras 5.13 e 5.14. As imagens em que  $\ell = 0,31$  já aponta regiões próximas aos nódulos, a segmentação com  $\ell = 0,55$  leva a regiões muito pequenas ou não permite encontrar nenhuma região que corresponda à posição apontada pelos especialistas.

Ao aumentar o valor de  $\ell$  usado na segmentação para 0,78, foi possível encontrar regiões com dimensões próximas ao esperado para três (3) casos em que os valores mais baixos de  $\ell$  não foram apropriados, como observado nas Figuras 5.11 e 5.12 e 5.15.

O uso do valor  $\ell = 1,02$  nas segmentações gerou regiões com dimensões próximas ao indicado para apenas um (1) dos casos, mostrado nas Figuras 5.16 e 5.17. Para esse

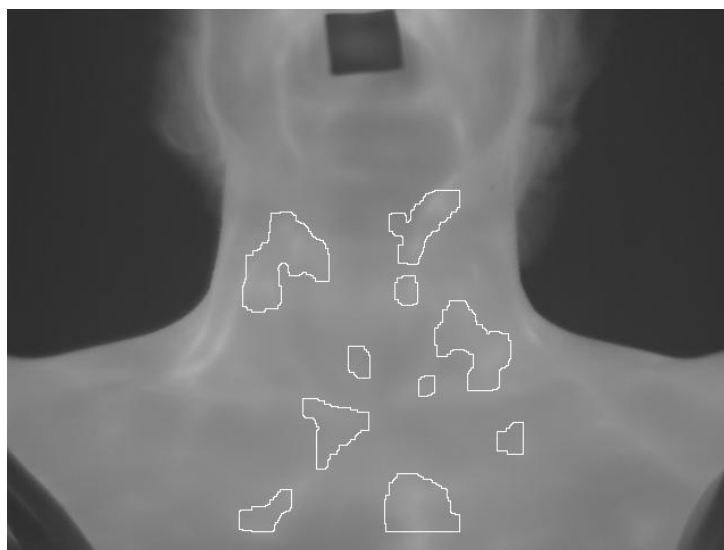


Figura 5.9: Segmentação do paciente 2017-10-10-3 com  $\ell = 0,31$

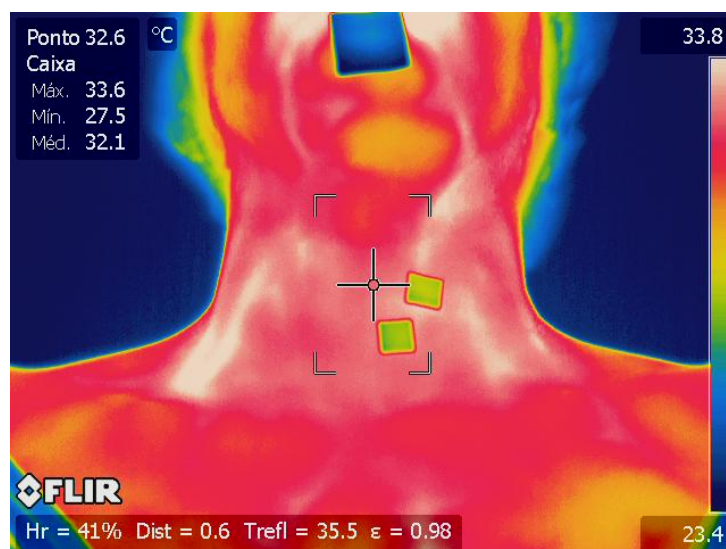


Figura 5.10: Marcação dos nódulos do paciente 2017-10-10-3

caso, as segmentações anteriores geraram regiões que abrangiam além do nódulo, uma área muito maior em torno dele. Isso pode ser um indicativo de que esse nódulo pode estar com tanta vascularização que chega a influenciar grande parte da área a sua volta. Entretanto, tal fenômeno pode também estar relacionado com a anatomia da paciente, que pode ter alguma estrutura particularmente quente nas proximidades do nódulo, como artérias por exemplo.

As demais regiões encontradas com  $\ell = 1,02$  apresentam dimensões muito menores que o esperado, como pode ser visto nas Figuras 5.12, 5.15 e 5.18. Houve ainda regiões nodulares que não foram detectadas com esse valor de  $P_A$ , o que sugere que ele seja um valor muito alto para os casos deste estudo. Isso diverge do encontrado na literatura médica,



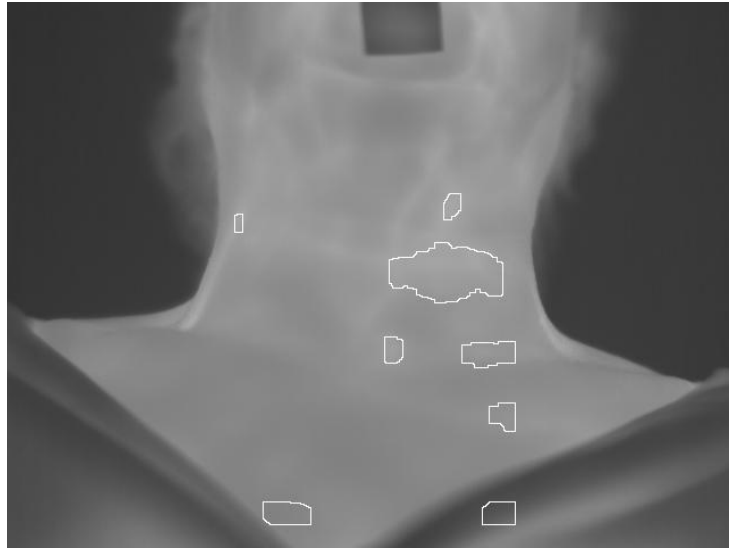


Figura 5.11: Segmentação do paciente 2017-10-10-1 com  $\ell = 0,31$

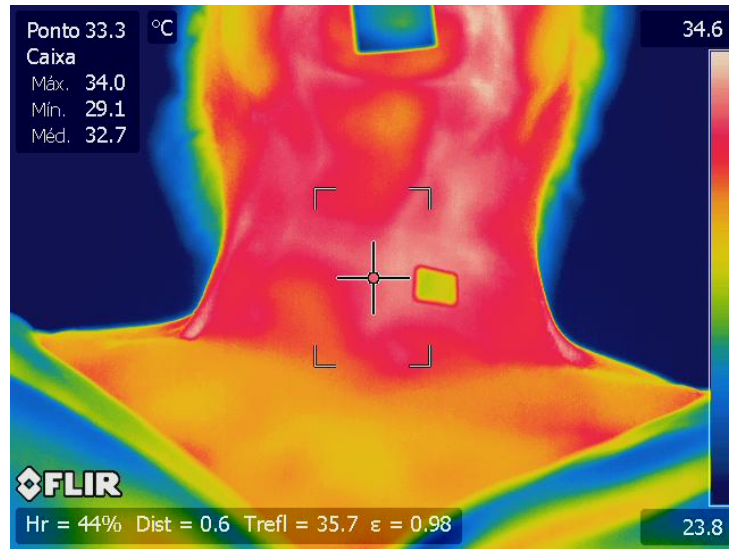


Figura 5.12: Marcação dos nódulos do paciente 2017-10-10-1

visto que segundo trabalhos nessa área, diferenças de  $1^{\circ}\text{C}$  são fortes caracterizadores de anomalias na região [Brioschi et al., 2010] .

Seguindo a tendência da segmentação com  $\ell = 1,02$ , as segmentações com  $\ell = 1,25$  e  $\ell = 1,48$  também geraram regiões com dimensões muito menores que o esperado, ou ainda não permitiram detectar as regiões nodulares, como no exemplo da Figura 5.19. Portanto, esses também são valores muito altos para serem usados como limiar  $\ell$  nos casos deste estudo.

De forma geral, para as termografias analisadas neste estudo, o  $\ell$  permite, para grande parte delas, identificar regiões com localização correspondente ao indicado pelos especi-

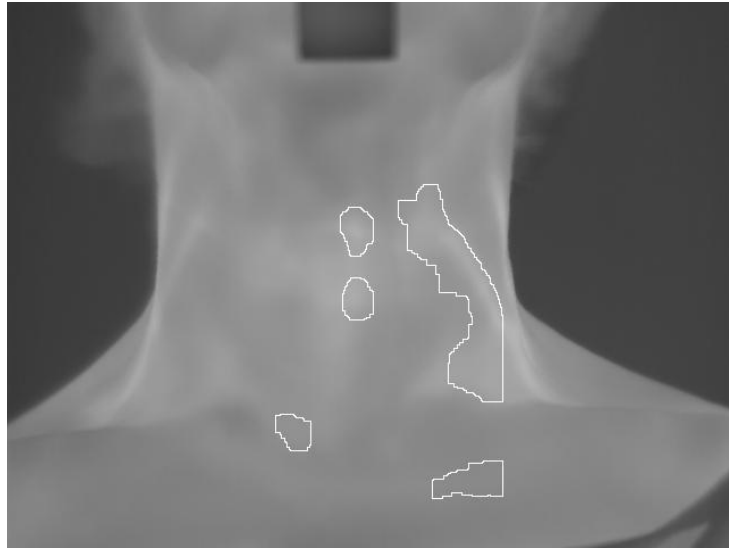


Figura 5.13: Segmentação do paciente 2017-09-11-2 com  $\ell = 0,55$

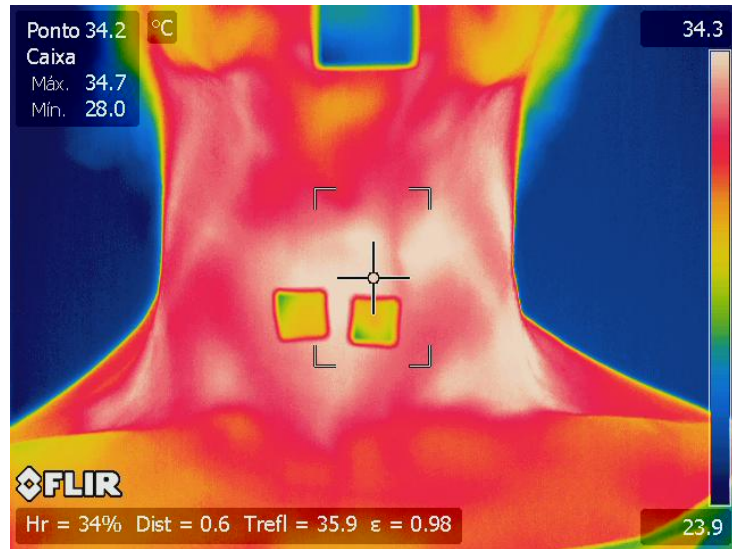


Figura 5.14: Marcação dos nódulos do paciente 2017-09-11-2

alistas, e dimensões próximas ao indicado nas ultrassonografias. Porém, não houve um único valor que gerasse essas regiões para todos os casos, o que indica que a diferença de temperatura está relacionada a cada caso, considerando o paciente e as características do nódulo em si. Há entretanto casos que o  $\ell$  não permitiu encontrar nenhuma correspondência com o indicado pelo especialista, o que pode indicar que há características dos nódulos que contribuem para que não haja uma maior vascularização da sua região e consequentemente um aumento de temperatura. Isso já era esperado, já que há casos de nódulos que apresentam formações diferenciadas. Assim, apesar do  $P_A$  ser a melhor alternativa de medida baseada em temperatura, nas termografias estáticas estudada neste trabalho, ela deveria ser usada em conjunto com alguma outra informação complementar

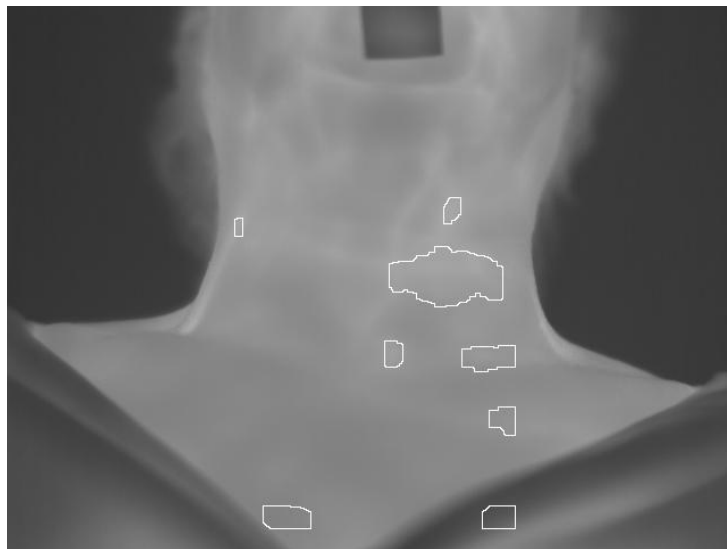


Figura 5.15: Segmentação do paciente 2017-10-10-1 com  $\ell = 0,78$

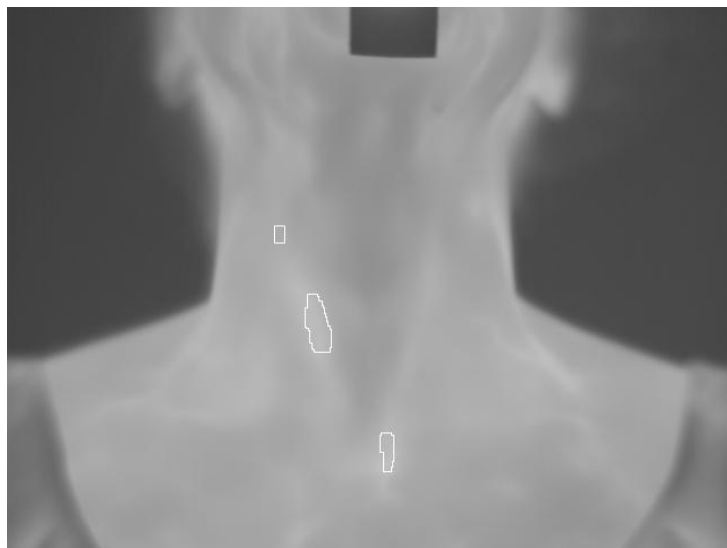


Figura 5.16: Segmentação do paciente 2017-09-11-1 com  $\ell = 1,02$

a fim de abranger casos em que a temperatura por si só não é fator determinante.

Além das regiões nodulares, outras regiões são encontradas na segmentação utilizando o parâmetro de assimetria  $P_A$ . Para eliminar falsos nódulos, aplica-se a operação morfológica de abertura com elemento estruturante de forma retangular de maior largura [Sonka et al., 2014], como descrito na seção 4.4.



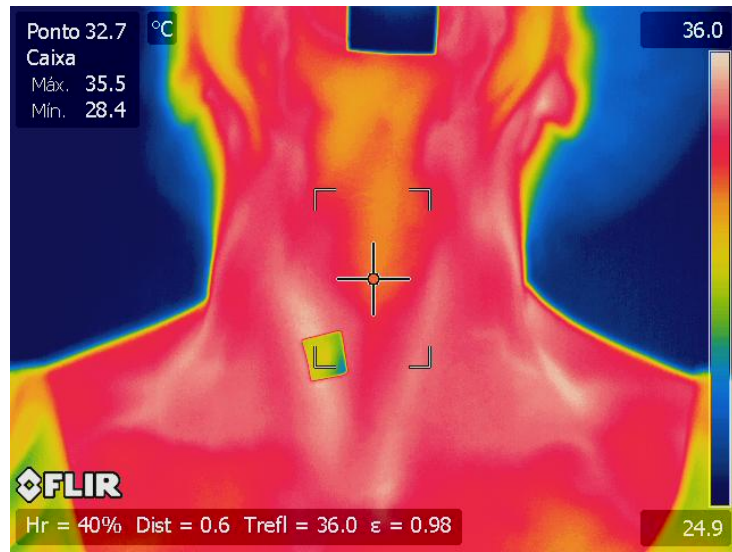


Figura 5.17: Marcação dos nódulos da paciente 2017-09-11-1

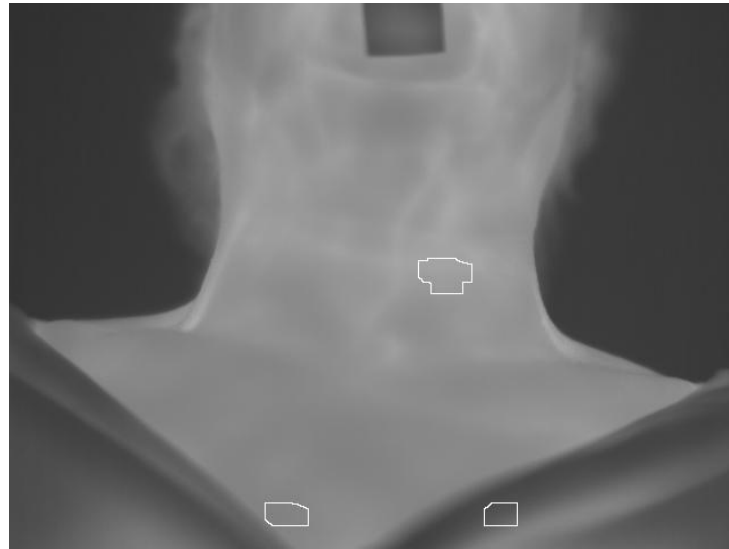


Figura 5.18: Segmentação do paciente 2017-10-10-1 com  $\ell = 1,02$

## 5.2 Resultados da classificação

A fim de definir o melhor CNN para a fase de segmentação, algumas variações de CNNs foram criadas, considerando diferentes parâmetros e dados de entrada. Esse processo e os resultados alcançados são mostrados a seguir.

### 5.2.1 Pré-processamento das entradas

Visando analisar cada uma dessas estruturas individualmente, novas imagens são criadas a partir delas, como é mostrado na Figura 5.20. Esse segundo processo de segmentação

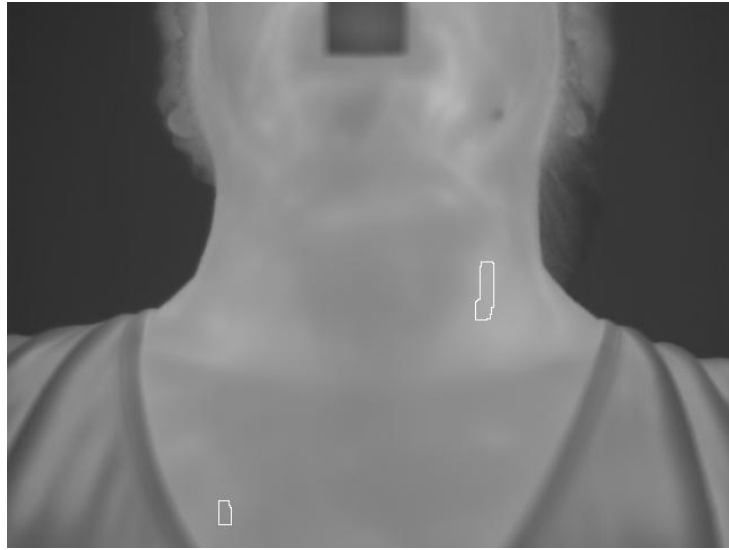


Figura 5.19: Segmentação do paciente 2017-10-10-2 com  $\ell = 1,25$

usa bounding boxes, como descrito na seção 4.4.



Figura 5.20: Bounding boxes em torno de possíveis nódulos considerando  $P_A$

Observou-se que apesar dessas regiões resultantes terem formas semelhantes, há ainda algumas características, principalmente relativas a contorno, que as diferenciam de modo que seja possível definir padrões que pudessem indicar quais se referem a regiões nodulares. Estudos na área médica indicam que alguns dos fatores a serem considerados no processos de identificação de nódulos tireoidianos são a aparência e o tamanho da anomalia [Bomeli et al., 2010] .

Reconhecer padrões em imagens com base nessas características é uma tarefa na qual algoritmos CNNs tem mostrado ótimos resultados. Assim, experimentos foram realizados a fim de avaliar a viabilidade do uso desse tipo de algoritmo para a classificação de imagens nodulares e não-nodulares encontradas na segmentação baseada em  $P_A$ . Entretanto, é

necessário ressaltar que este se trata de um estudo preliminar, já que o volume de dados disponível é muito restrito para levar a conclusões definitivas.

### 5.2.2 CNNs analisadas

Neste estudo foram selecionados duas arquiteturas de CNNs que apresentam bons resultados em problemas similares: ResNet e GoogLeNet, a fim de avaliar o desempenho de cada um deles no problema de identificação de nódulos tireoidianos a partir de candidatos a nódulos obtidos a partir da segmentação das imagens térmicas.

Estudos previamente realizados para este problema [Moran et al., 2018], apontam que as redes VGG e AlexNet, que também apresentaram bons resultados em trabalhos similares, não foram bem sucedidas no problema aqui tratado, chegando a acurácias de apenas, 0,75 e 0,77, respectivamente.

### 5.2.3 Preparação dos algoritmos CNNs

Como descrito na seção 4.4, é necessário processar as imagens obtidas na segmentação a fim de criar um maior volume de dados a serem utilizados como entradas dos algoritmos CNNs na fase de treinamento, visando melhorar o desempenho das redes. As operações de espelhamento e mudança de tons, usadas para criação de dados sintéticos, são descritas na seção 4.5, e um exemplo dos resultados obtidos através delas é mostrado a seguir.

### 5.2.4 Criação de dados sintéticos

Nas Figuras 5.21 e 5.22 é possível ver o resultado dessas operações para um exemplo de imagem de candidato a nódulo. Na Figura 5.21 a operação de espelhamento foi aplicada. Já na Figura 5.22 é mostrada a operação de mudança de tons considerando novos limites. Esse processo de definição dos novos possíveis limites pode ser melhor observado a partir da análise dos histogramas das imagens original e resultante, como mostrado na Figura 5.22, na qual à esquerda está a imagem original (inicialmente definida num intervalo de 0 a 255) e seu histograma, e à direita está a imagem resultante possível do processo de mudança de intensidade, usando como novos valores limite 0 e 127 e seu histograma.

Dessa forma, um grande número de imagens podem ser geradas através de espelhamento e mudança de tons usando diferentes valores limite sequencialmente, como mostrado no exemplo da Figura 5.23. Porém, deve-se considerar que quanto maior o volume de da-

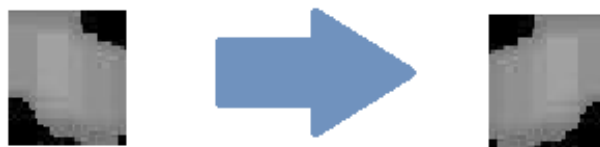


Figura 5.21: Processo de espelhamento em torno de um eixo vertical

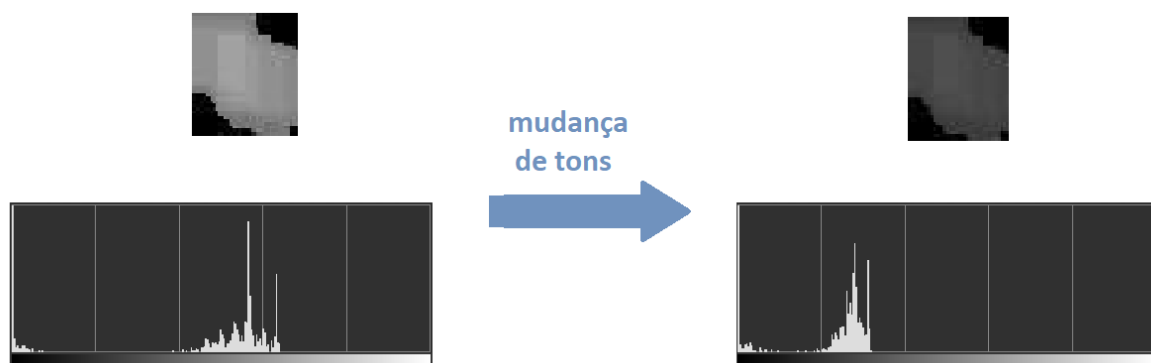


Figura 5.22: Imagem original e resultantes e seus histogramas.

dos, maior o tempo de processamento resultante. É necessário portanto definir quantas imagens são suficientes para garantir resultados satisfatórios com o algoritmo, considerando que um pequeno acréscimo no número de imagens não resultaria em grande melhora. Segundo [Goodfellow et al., 2016], uma forma de estimar um bom número de dados de entrada é realizar testes com diferentes tamanhos de entradas e observar a acurácia obtida pelo algoritmos, e a partir desses resultados observar o número de entradas necessárias para chegar a um valor de acurácia desejável. Ainda segundo [Goodfellow et al., 2016], é recomendável variar os tamanhos de entrada em uma escala logarítmica, dobrando o número de imagens a cada experimento.

Considerando que inicialmente tem-se apenas 16 imagens de nódulos malignos para os experimentos, as quais foram obtidas na fase de segmentação, definiu-se esse conjunto de imagens originais como N1. Após a aplicação da operação de espelhamento em cada uma das imagens de N1, tem-se um conjunto N2 formado pelas imagens de N1 e as novas imagens geradas através do espelhamento, resultando num total de 32 imagens.

Ao aplicar a operação de mudança de intensidade tonal em todas as imagens de N2, obtém-se um conjunto N3, com todas as imagens de N2 mais as novas imagens geradas com os valores limite 25 e 210, resultando em um total de 64 imagens. De forma similar, o conjunto N4 é obtido ao aplicar a operação de mudança de intensidade em N3 usando

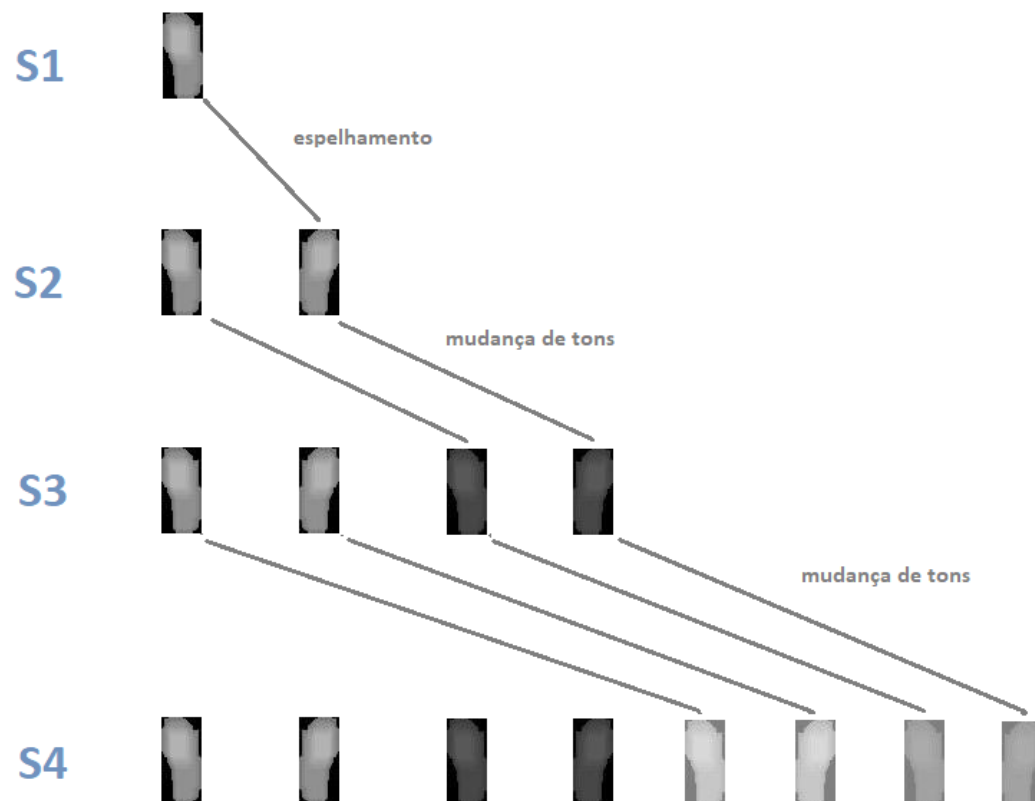


Figura 5.23: Processo de criação de imagens sintéticas a partir de uma imagem de S1.

os valores limites 70 e 197 resultando em um total de 128 imagens. Assim, são obtidos os conjuntos N5 e N6 com os valores limites 0 e 127, e 127 e 255, respectivamente. Observe que cada novo conjuntos de imagens criado tem o dobro de elementos do conjunto anterior. A Tabela 5.3 mostra o número de imagens para cada um dos subconjuntos N.

Tabela 5.3: Número de imagens para cada conjunto N

Conjunto	Operações realizadas	textbf Número de imagens
N1	originais	16
N2	espelhamento	32
N3	mudança de intensidade tonal (limites 25 -210)	64
N4	mudança de intensidade tonal (limites 70 -197)	128
N5	mudança de intensidade tonal (limites 0 -127)	256
N6	mudança de intensidade tonal (limites 127 -255)	512

Como o número inicial de imagens de regiões não nodulares é maior que os de nodulares, foi necessário gerar apenas 4 novos conjuntos (S2, S3 e S4) a partir do conjunto de imagens iniciais (S1). O conjunto S2 foi criado através da operação de espelhamento. Os

conjuntos S3 e S4 foram criados com a operação de mudança de tons os valores limites 0 e 127, e 127 e 255, respectivamente. O número de imagens para cada um desses conjuntos está na Tabela 5.4. O processo de criação desses conjuntos é mostrado na Figura 5.23.

Tabela 5.4: Número de imagens para cada conjunto S

Conjunto	Operações realizadas	Número de imagens
S1	originais	68
S2	espelhamento	136
S3	mudança de intensidade tonal (limites 0 -127)	272
S4	mudança de intensidade tonal (limites 127 - 255)	544

Os conjuntos de entrada a serem usadas nos CNNs devem ser compostos por imagens de todas as classes definidas. Assim, combinações dos conjuntos de regiões nodulares (N) e não-nodulares (S) foram criadas para serem usadas como entrada das redes. Essas combinações são mostradas na Tabela 5.5. Observe que os conjuntos de dados DT2 a DT5 foram definidos de forma a equilibrar o número de imagens em cada classe.

Tabela 5.5: Combinações que formam os DT usados como entrada dos CNNs

DT	N	S
DT1	N1	S1
DT2	N3	S1
DT3	N4	S2
DT4	N5	S3
DT5	N6	S4

### 5.2.5 Configuração das CNNs

Um fator a ser analisado é a escolha de valores para cada parâmetros do algoritmo. A taxa de aprendizado é um dos fatores que mais impactam no resultado final [Goodfellow et al., 2016]. Como o valor ideal para esse parâmetro varia bastante de acordo com o problema, é importante realizar testes com diferentes valores a fim de observar qual seria o valor mais apropriado que resultaria num algoritmo mais preciso. Geralmente os valores para taxa de aprendizado utilizados em CNNs são da forma  $10^{-k}$ , sendo k um número natural. Assim, nos experimentos a serem realizados, definiu-se que os valores de taxa de aprendizados a serem testados seriam 0,1, 0,01 e 0,001.

Em relação às técnicas de validação, cogitou-se a utilização da técnica de cross-validation a fim de garantir resultados mais confiáveis, entretanto tal possibilidade foi

inicialmente descartada devido ao alto custo computacional requerido.

Outro fator que possivelmente levaria a melhores resultados seria o uso de modelos pré-treinados, implementando a técnica de fine-tuning. Um teste prévio usando o conjunto DT1 foi realizado, usando a rede de arquitetura GoogLeNet com um modelo pré-treinado com o conjunto CIFAR-10 [Krizhevsky and Hinton, 2009]. Entretanto os resultados alcançados foram bem similares aos apresentados pelos modelos sem pré-treinamento, assim o uso dessa técnica foi descartado para os demais experimentos.

### 5.2.6 Experimentos com CNNs

Os experimentos foram realizados da seguinte forma: considerando as arquiteturas de CNN GoogLeNet e ResNet, para cada combinação de conjunto de dados descrita na Tabela 5.5, o treinamento das redes foi feito considerando as taxas de aprendizado já definidas. Todas as redes foram treinadas por 50 épocas e a cada época testou-se o resultado da rede com o conjunto de teste. Os subconjuntos de treinamento e teste correspondem a 80% e 20% do conjunto inicial, respectivamente. Estes são formados selecionando aleatoriamente no conjunto inicial a cada época. Os resultados dos experimentos são apresentados no Anexo A.

O primeiro experimento usou o conjunto de dados (dataset) DT1, que consistia no conjunto de imagens iniciais de cada classe. Os valores de acurácia (accuracy) de cada CNN alcançados ao longo das épocas do experimento, para as taxa de aprendizado 0,1, 0,01 e 0,001 são mostrados nas Figuras 5.24, 5.25 e 5.26, respectivamente. Na Tabela 5.6, são mostrados valores de acurácia (accuracy), acurácia de validação (validation accuracy), perda (loss) e perda de validação (validation loss) alcançados após o fim do treinamento.

Tabela 5.6: Acurácia, acurácia de validação, perda e perda de validação para o experimento com DT1

Arquitetura	Taxa de aprendizado	Accuracy	Val. accuracy	Loss	Val. loss
ResNet	0,1	0,82	0,57	0,47	0,80
ResNet	0,01	0,64	0,43	0,65	0,72
ResNet	0,001	0,71	0,43	0,61	0,99
GoogLeNet	0,1	0,49	0,71	0,73	0,64
GoogLeNet	0,01	0,61	0,57	0,67	0,69
GoogLeNet	0,001	0,60	0,57	0,68	0,68

De forma geral, a rede GoogLeNet apresenta melhores valores nas primeiras épocas,

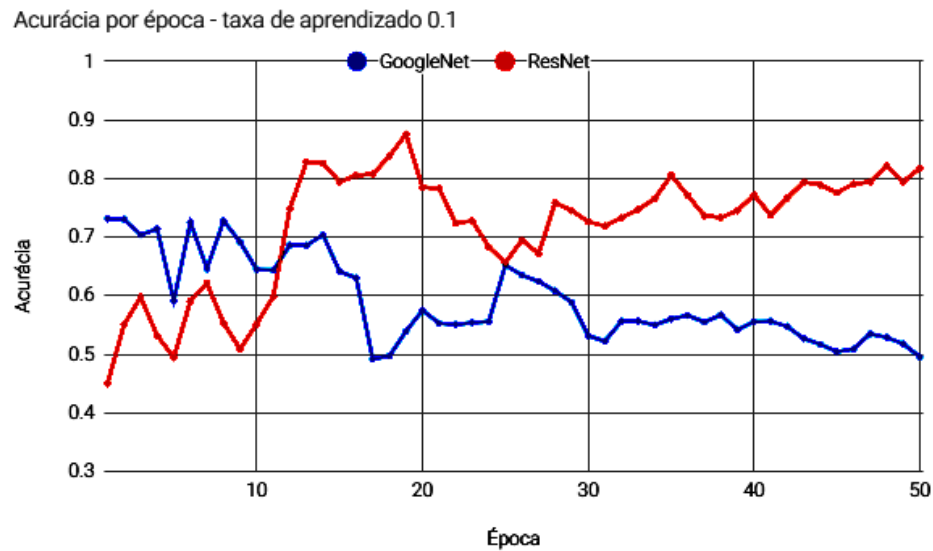


Figura 5.24: Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendizado 0,1, com DT1

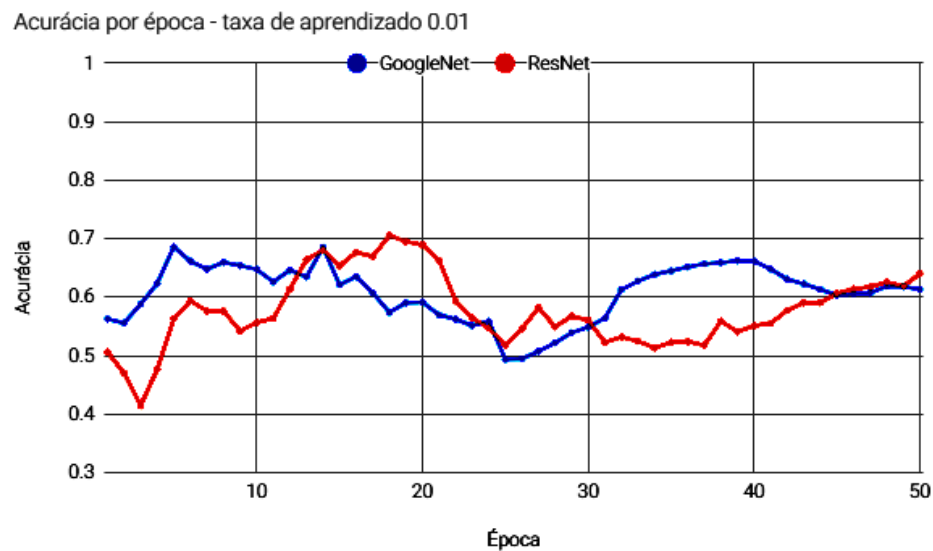


Figura 5.25: Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendizado 0,01, com DT1

sendo depois superada pela rede ResNet. Exceto para a taxa de aprendizado 0,01, a rede ResNet permanece acima da GoogLeNet a partir do ponto em que a supera. Para a ResNet os melhores resultados aparecem ao usar a taxa de aprendizado 0,1. Para a rede GoogLeNet os resultados são bastante oscilantes, de forma a que dentre as três taxas de aprendizado não há uma configuração cujos resultados sejam notoriamente superiores. Dentre todas as configurações a que apresenta melhores resultados com o conjunto DT1 é a rede ResNet com taxa de aprendizado 0,1.



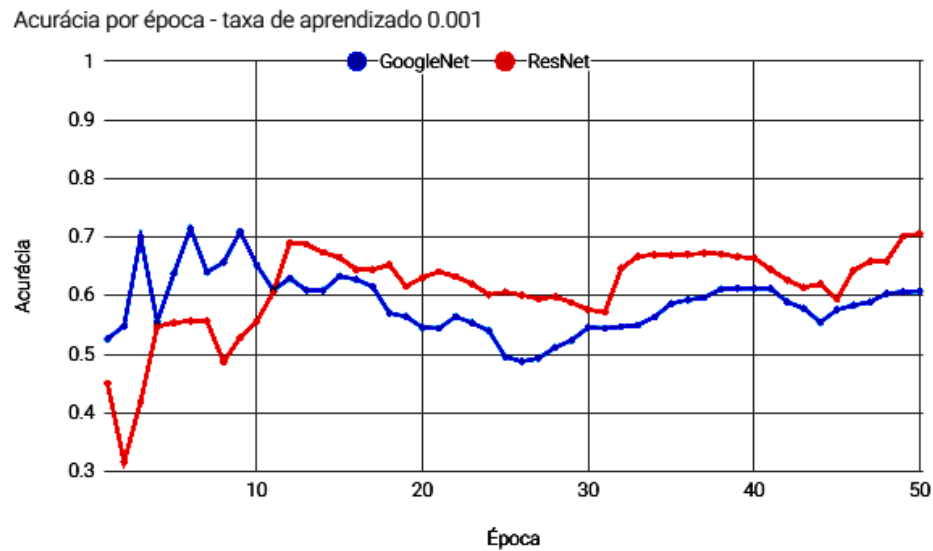


Figura 5.26: Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendizado 0,001, com DT1

O segundo experimento usa o conjunto DT2, que consiste no conjunto DT1 mais as imagens obtidas através a operação de espelhamento. Os valores de acurácia de cada CNN para as taxa de aprendizado definidas são mostrados nas Figuras 5.27, 5.28 e 5.29, respectivamente. Os valores de acurácia, acurácia de validação, perda e perda de validação obtidos no fim do treinamento são mostrados na Tabela 5.7.

Tabela 5.7: Acurácia, acurácia de validação, perda e perda de validação para o experimento com DT2

Rede	Taxa de aprendizado	Accuracy	Val. accuracy	Loss	Val. loss
ResNet	0,1	0,82	0,92	0,38	0,26
ResNet	0,01	0,67	0,50	0,63	0,65
ResNet	0,001	0,76	0,79	0,52	0,62
GoogLeNet	0,1	0,45	0,48	0,78	0,74
GoogLeNet	0,01	0,66	0,67	0,64	0,60
GoogLeNet	0,001	0,55	0,54	0,69	0,69

Para a entrada DT2, os melhores resultados são apresentados pela CNN ResNet com taxa de aprendizado 0,1. Para as taxas 0,1 e 0,001 a rede ResNet supera a GoogLeNet em quase todas as épocas. Já para a taxa 0,01, ambas redes apresentam resultados bem próximos, apesar de algumas oscilações. A CNN GoogLeNet apresenta seus piores resultados com a taxa 0,1. Novamente, os resultados sugerem que a rede ResNet seja melhor para o problema aqui apresentado.

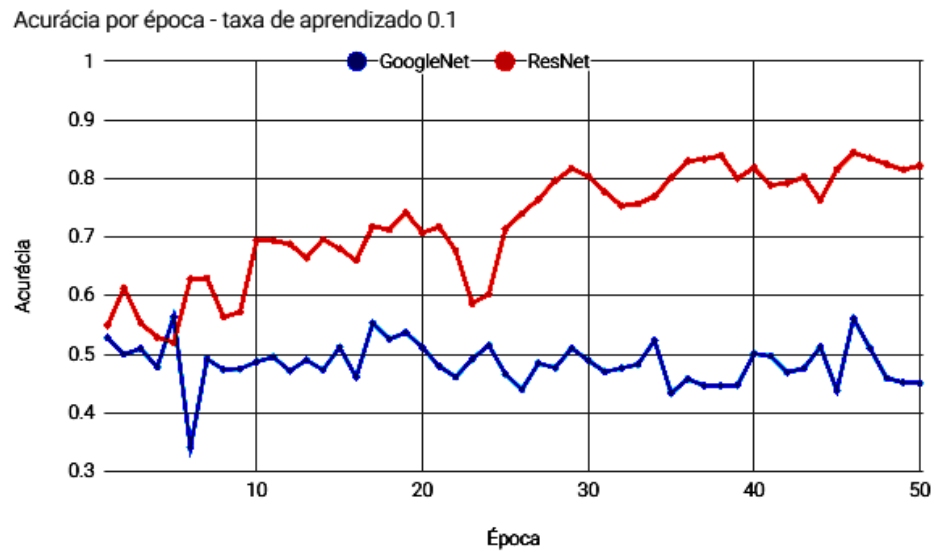


Figura 5.27: Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendizado 0,1, com DT2

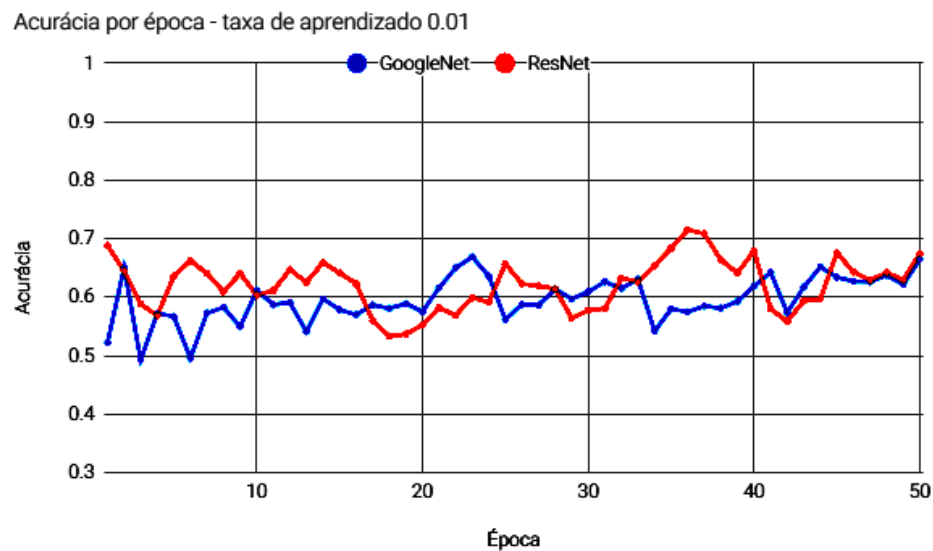


Figura 5.28: Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendizado 0,01, com DT2

O terceiro experimento usa o conjunto DT3, que consiste no conjunto DT2 mais as imagens obtidas através a operação de mudança de intensidade tonal. Os valores de acurácia obtidos são mostrados nas Figuras 5.30, 5.31 e 5.32 , respectivamente. Os valores de acurácia, acurácia de validação, perda e perda de validação obtidos no fim do treinamento são mostrados na Tabela 5.8.

No experimento usando DT3, a rede ResNet supera com considerável diferença a rede

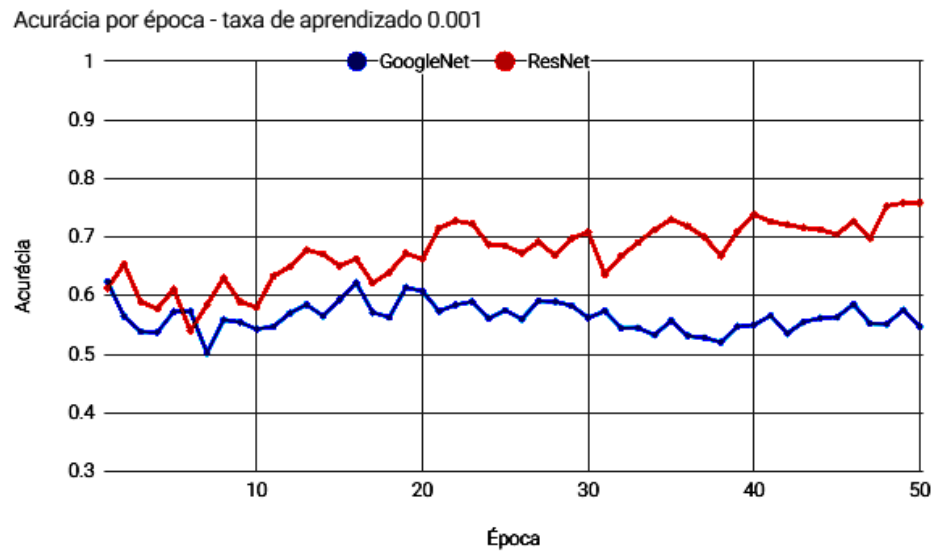


Figura 5.29: Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendizado 0,001, com DT2

Tabela 5.8: Accuracy, validation accuracy, loss e validation loss para o experimento com DT3

Rede	Taxa de aprendizado	Accuracy	Val. accuracy	Loss	Val. loss
ResNet	0,1	0,89	0,83	0,28	0,31
ResNet	0,01	0,85	0,81	0,36	0,53
ResNet	0,001	0,79	0,79	0,45	0,81
GoogLeNet	0,1	0,62	0,44	0,67	0,75
GoogLeNet	0,01	0,73	0,56	0,59	0,75
GoogLeNet	0,001	0,59	0,42	0,68	0,75

GoogLeNet. Os resultados da ResNet para esse conjunto são iguais ou superiores aos anteriores. Novamente a taxa 0,1 leva a melhores resultados para a ResNet. Já para a GoogLeNet, os resultados se mantêm entre 0,50 e 0,60, com algumas oscilações. A taxa 0,01 leva a resultados levemente superiores para a GoogLeNet.

O quarto experimento usa o conjunto DT4, que é formado pelo DT3, e imagens geradas por uma nova operação de mudança de espectro, com novos limites. Os resultados desse experimento estão nas Figuras 5.33, 5.34 e 5.35. Na Tabela 5.9 são listados os valores de acurácia, acurácia de validação, perda e perda de validação atingidos no fim do treinamento.

Para a entrada DT4, os resultados seguem a tendência de forma que a ResNet supera a GoogLeNet para todas as taxas de aprendizado, apesar de que para a taxa 0,01 os

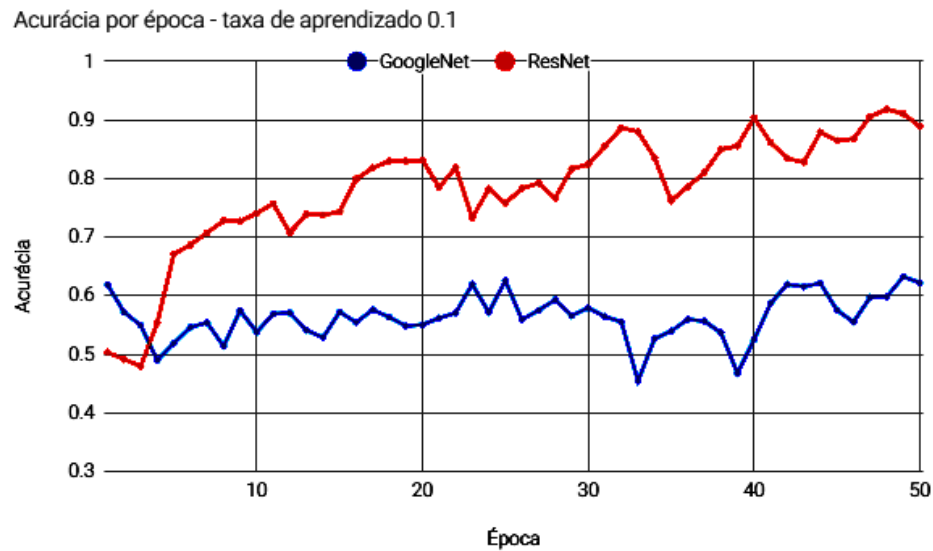


Figura 5.30: Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendizado 0,1, com DT3

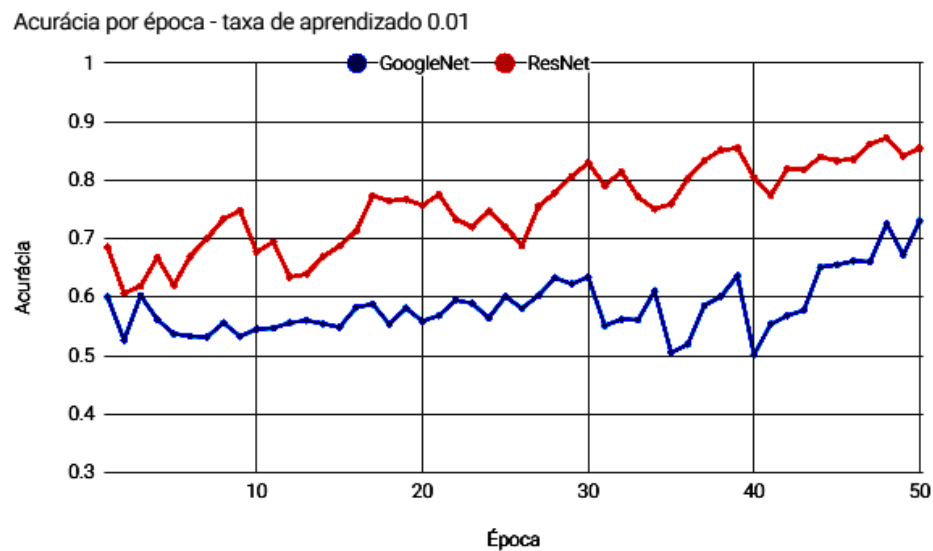


Figura 5.31: Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendizado 0,01, com DT3

resultados se aproximam bastante. Entretanto para as taxas 0,1 e 0,001 a diferenças entre as redes é significativa. Em relação a ResNet, os resultados são superiores aos dos experimentos anteriores. Já para a GoogLeNet, os melhores resultados neste experimento são com taxa 0,01, sendo eles inclusive os melhores resultados da rede GoogLeNet dentre os experimentos realizados. Já os resultados da ResNet nesse experimento também estão entre os melhores do estudo.

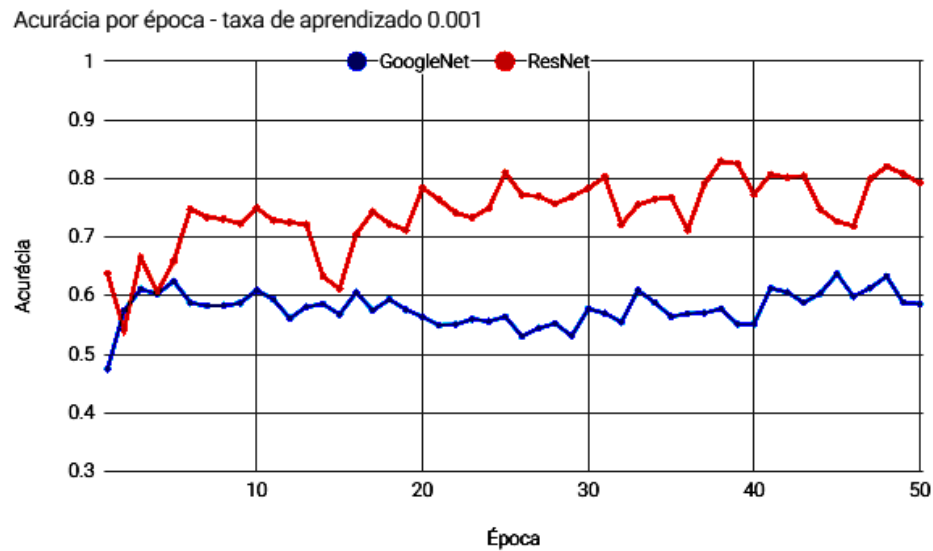


Figura 5.32: Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendizado 0,001, com DT3

Tabela 5.9: Acurácia, acurácia de validação, perda e perda de validação para o experimento com DT4

Rede	Taxa de aprendizado	Accuracy	Val. accuracy	Loss	Val. loss
ResNet	0,1	0,96	0,95	0,14	0,08
ResNet	0,01	0,94	0,90	0,16	0,29
ResNet	0,001	0,92	0,87	0,18	0,25
GoogLeNet	0,1	0,44	0,60	0,70	0,69
GoogLeNet	0,01	0,86	0,90	0,38	0,30
GoogLeNet	0,001	0,76	0,80	0,52	0,46

O quinto e último experimento usou o conjunto DT5, por DT4, e imagens obtidas através de outra mudança de tons. Os resultados desse experimento estão nas Figuras 5.36, 5.37 e 5.38. As informações de acurácia, acurácia de validação, perda e perda de validação obtidos no fim do treinamento são mostrados na Tabela 5.10.

Considerando por fim o experimento com DT5, observa-se que para o CCN ResNet os resultados diminuem um pouco em relação ao experimento anterior. Assim como na maioria dos casos dos experimentos anteriores a ResNet superou a GoogLeNet, neste experimento com ampla diferença. Para a CNN ResNet os resultados se mantêm relativamente similares para as três taxas de aprendizado, sendo os valores para taxa 0,001 superiores. Ao analisar o desempenho do GoogLeNet é possível observar que este se mantêm entre 0,50 e 0,60 para as taxas 0,1 e 0,001. Essa tendência também se apresenta em grande

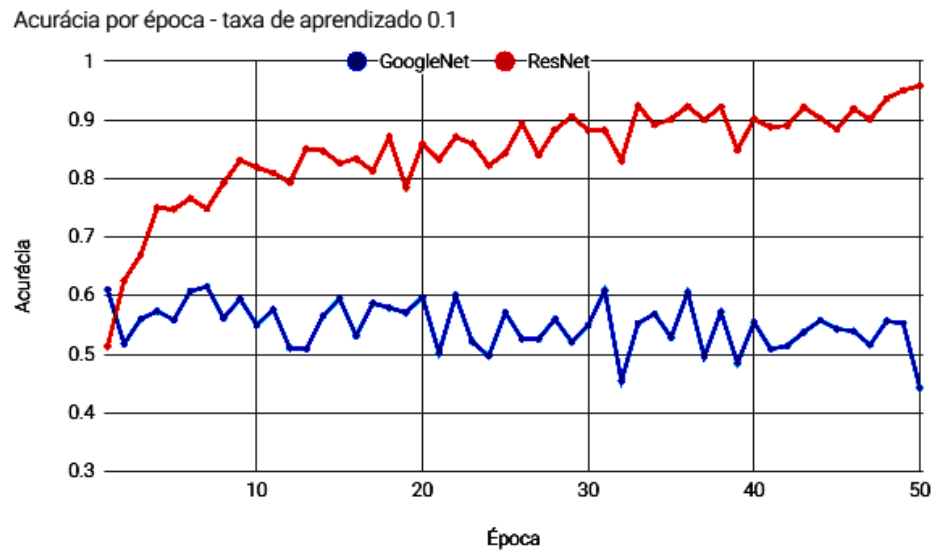


Figura 5.33: Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendizado 0,1, com DT4

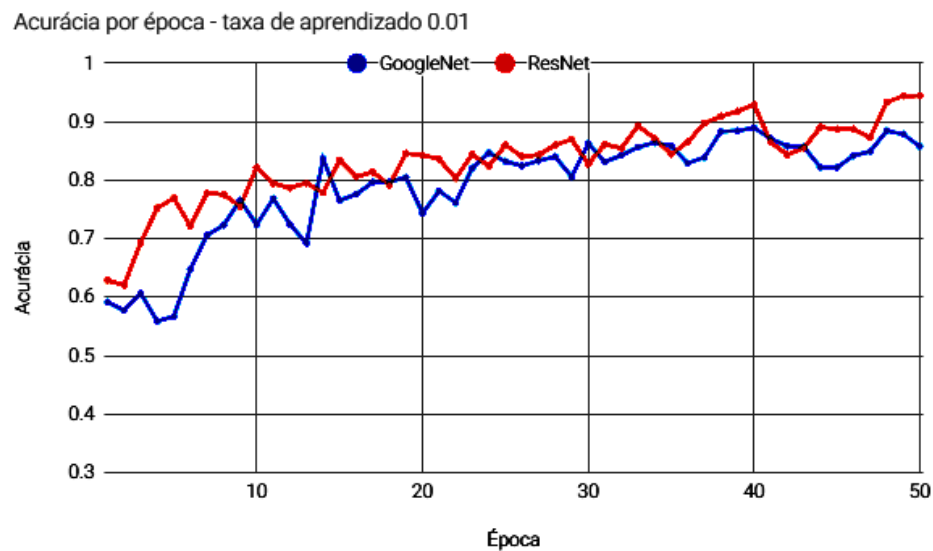


Figura 5.34: Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendizado 0,01, com DT4

parte das épocas para a taxa 0,01, porém as épocas finais os valores apresentam um leve crescimento.

De forma geral, na maioria das execuções, a rede ResNet mantém uma tendência crescente e apresenta melhores resultados que a GoogLeNet. Para a rede GoogLeNet não é possível identificar uma tendência. Como já comentado, os melhores resultados são os obtidos com a rede ResNet nos experimentos DT4, o que leva a concluir que o acréscimo

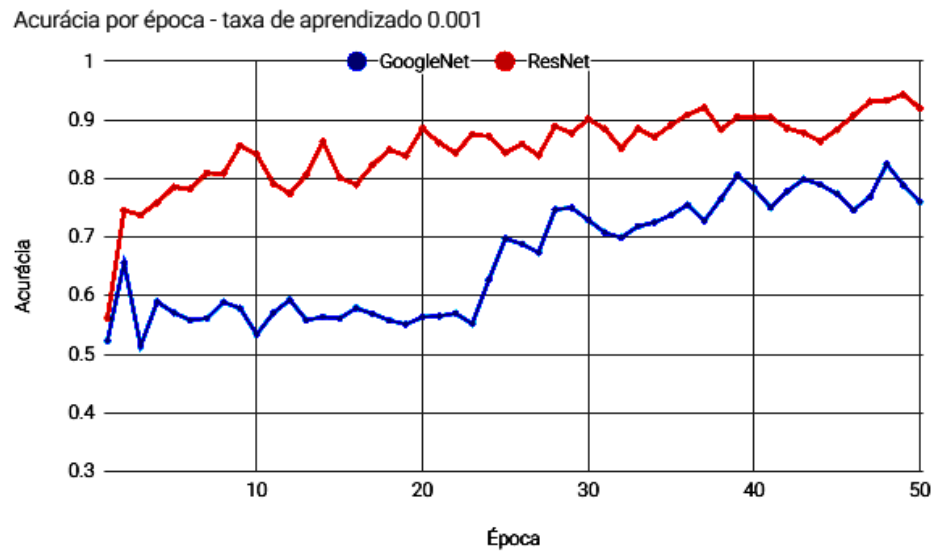


Figura 5.35: Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendizado 0,001, com DT4

Tabela 5.10: Acurácia, acurácia de validação, perda e perda de validação para o experimento com DT5

Rede	Taxa de aprendizado	Accuracy	Val. accuracy	Loss	Val. loss
ResNet	0,1	0,89	0,78	0,31	0,55
ResNet	0,01	0,93	0,85	0,22	0.40608
ResNet	0,001	0,94	0,91	0,17	0,26
GoogLeNet	0,1	0,54	0,56	0,70	0,69
GoogLeNet	0,01	0.6834	0.6522	0.60	0,57
GoogLeNet	0,001	0,54	0,51	0,68	0,69

de imagens do conjunto DT4 para o DT5 não introduz melhora no desempenho da rede.

Dentre os melhores resultados da GoogLeNet estão os do experimento com DT4 com taxa 0,01. Assim, de forma geral, o conjunto que permite melhor desempenho é o DT4. Os resultados com DT1, DT2 e DT3 são bastante oscilantes e baixos, o que é esperado devido ao volume de dados da entrada. Os cinco maiores valores de acurácia alcançados nos experimentos estão na Tabela 5.11. Observe que todos eles se referem a CNNs do modelo ResNet.

Ao observar também as demais métricas de avaliação do desempenho das redes (acurácia de validação, perda e perda de validação) conclui-se que as redes ResNet com DT4 e taxas de aprendizado 0,1 e 0,01 apresentam os melhores resultados, já que o valores de acurácia de validação se aproximam dos valores de acurácia, e o valores de perda e

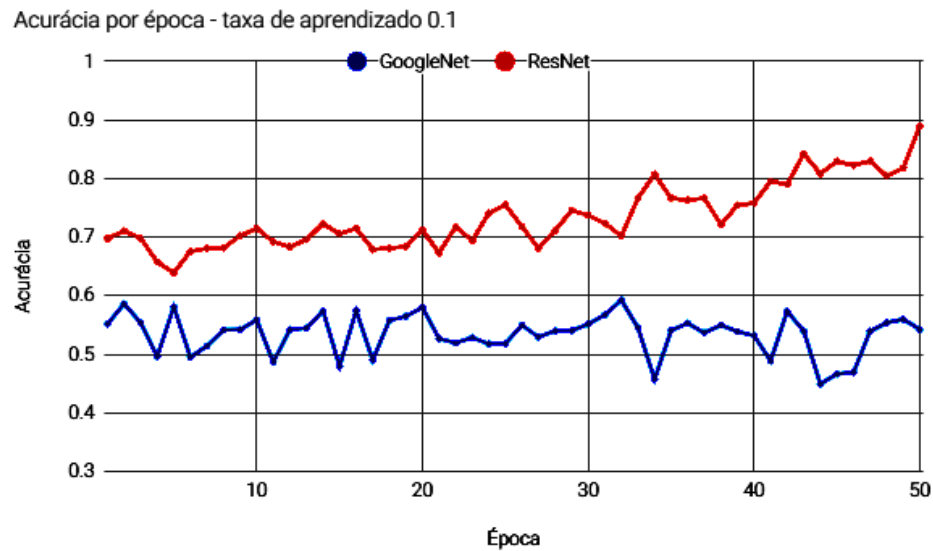


Figura 5.36: Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendizado 0,1, com DT5

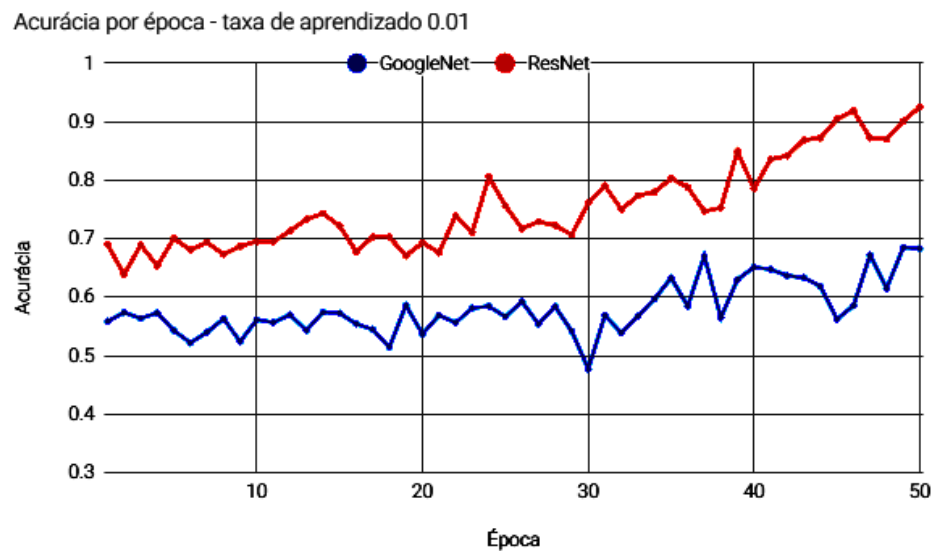


Figura 5.37: Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendizado 0,01, com DT5

perda de validação são relativamente baixos e se aproximam, o que aponta um bom fitting. Observa-se em geral, que diversas variações testadas nos experimentos apresentam valores de perda altos. Também foi observado que diversos testes levaram a casos de overfitting e underfitting, sendo esse modelos considerados portanto inadequados para o problema aqui apresentado.

Considerando a configuração com melhor desempenho, a adequação da metodologia



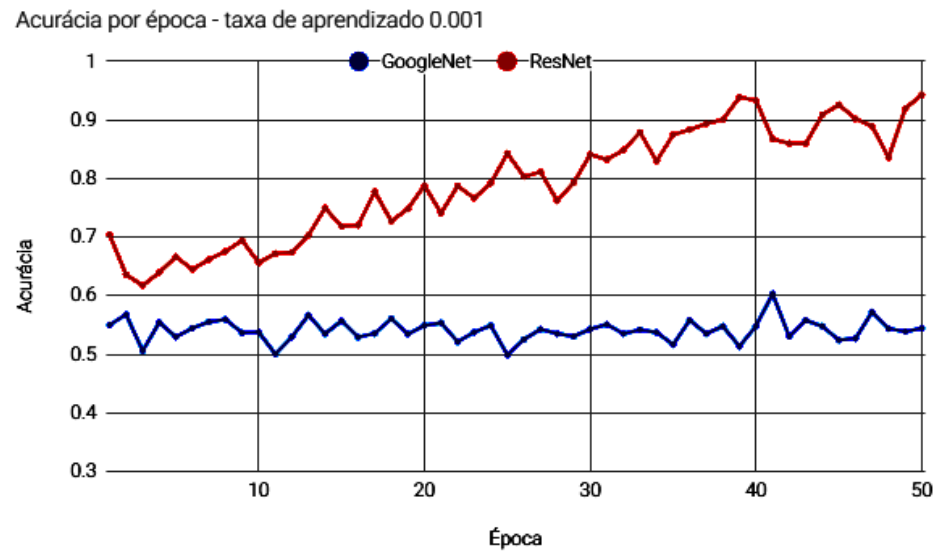


Figura 5.38: Acurácia por época dos CNNs GoogLeNet e ResNet com taxa de aprendizado 0,001, com DT5

Tabela 5.11: Maiores valores de acurácia alcançados nos experimentos CNN

Rede	Conjunto de entrada	Taxa de aprendizado	Acurácia
ResNet	DT4	0,1	0,96
ResNet	DT4	0,01	0,94
ResNet	DT5	0,001	0,94
ResNet	DT5	0,01	0,93
ResNet	DT4	0,001	0,92

aqui proposta fica principalmente dependente do processo de extração de regiões possivelmente nodulares, já que ainda há regiões nodulares que não conseguem ser detectadas por esse método.

# Capítulo 6

## Conclusões

Nódulos tireoidianos apresentam grande prevalência na área clínica, e sua malignidade representa um grande risco à saúde. O câncer de tireoide é a doença maligna mais comum do sistema endócrino e é o oitavo tipo de câncer diagnosticado em mulheres em todo o mundo. A identificação precoce desse tipo de lesão permite um tratamento mais rápido e eficaz, aumentando as chances de cura. No trabalho apresentado nesta dissertação, uma metodologia é proposta, com o objetivo de localizar nódulos malignos tireoidianos.

A realização deste estudo utiliza imagens de termografias em infravermelho estáticas. Tais imagens foram adquiridas usando o protocolo desenvolvido em trabalhos prévios do grupo VisualLab [Gonzáles, 2017, Fiirst, 2018].

A metodologia proposta aqui, envolve as fases de segmentação de candidatos a nódulos e sua classificação. Para o processo de segmentação dos possíveis nódulos, um estudo prévio dos parâmetros proposto na literatura médica é realizado. Como resultado, observa-se que o parâmetro de assimetria  $P_A$  foi o que melhor permitiu diferenciar regiões nodulares de regiões saudáveis. Entretanto, não se chegou a um valor único definitivo de  $P_A$  que permitisse a segmentação aproximada de nódulos para todos os pacientes. Para três das dezoito pacientes não foi possível encontrar regiões correspondentes a nódulos com o  $P_A$ , o que se deve a particularidades do nódulo, que pode estar localizado no istmo, e limitação no processo.

O processo de classificação consistiu em utilizar CNNs, para diferenciar as regiões nodulares de regiões referentes a outras anomalias, dentre as regiões encontradas na fase de segmentação de candidatos a nódulos. A escolha dessa técnica deve-se ao bom desempenho dela em problemas similares. Dois modelos de CNNs são selecionados para comparação de desempenho, considerando os valores de acurácia obtidos. Os algoritmos ResNet e

GoogLeNet foram selecionados com base em trabalhos anteriores.

Três taxas de aprendizado e cinco conjuntos de dados são testados, a fim de maximizar o desempenho dos algoritmos. Os dados da segmentação são processados com técnicas de processamento de imagens a fim de utilizá-los como entrada das redes CNN. A rede de arquitetura ResNet com taxa de aprendizado 0,1 é o que apresenta melhor acurácia. O conjunto de dados DT4 leva a melhores desempenhos. Como trabalho futuro, consideramos implementar uma nova arquitetura CNN, considerando aspectos específicos do problema. Entretanto, essa fase deve ser melhor avaliada, quando houver mais dados disponíveis, de forma a tornar o processo mais robusto.

Considerando o bom desempenho da rede CNN na fase de classificação, a adequação da metodologia aqui proposta fica principalmente dependente do processo de segmentação de regiões possivelmente nodulares, já que ainda há regiões nodulares que não conseguem ser detectadas por esse método.

Com isso, foi alcançado o desenvolvimento das técnicas de segmentação e classificação usando algoritmos de redes neurais convolucionais, como foi definido na proposta inicial deste trabalho.

## 6.1 Trabalhos futuros

A fim de dar continuidade ao trabalho aqui apresentado, sugere-se a realização de:

- Aprimorar o processo de cálculo do parâmetro de assimetria, considerando outros eixos no cálculo, como por exemplo, eixos paralelos ao contorno do pescoço, o que amenizaria o problema de mal-posicionamento dos pacientes durante a captura;
- Estudo de outros parâmetros baseados em termografias, em especial, termografias dinâmicas, usando séries de temperatura;
- Avaliação de técnicas visando automatizar o processo de identificação de valores usados na segmentação com parâmetros de assimetria;
- Dar continuidade na aquisição de termografias e desenvolvimento do repositório de imagens disponíveis
- Usar como entrada dos CNNs um recorte (bounding box) considerando a região na imagem original, não a região na segmentação com o fundo preto;

- Avaliar o desempenho das redes das arquiteturas estudadas definindo um conjunto de teste não usado no treinamento;
- Avaliar o desempenho de outros modelos de CNN propostos na literatura na fase de classificação;
- Testar o uso de outros modelos para o processo de pré-treinamento além do CIFAR-10, como , os quais podem ser mais adequados ao problema aqui estudado;
- Analisar a variação de outros parâmetros nos algoritmos CNNs a fim de melhorar o desempenho, como a aplicação de uma taxa de aprendizado decrescente;
- Desenvolver um processo de segmentação utilizando algoritmos de inteligência artificial, e avaliar seu desempenho;
- Estudar a inclusão de outras informações do pacientes ligadas a possíveis diagnósticos, de forma a melhorar o resultado da metodologia proposta

# Referências

- [Alves et al., 1988] Alves, M. L. D., Andrade, J., Cherri, J., Moriya, T., Piccinato, C. E., Okano, N., Llorach Velludo, M. A., and Iazigi, N. (1988). Papel da termografia da seleção de nódulos tireoideanos de indicação cirúrgica. *Arq. bras. endocrinol. metab*, 32(4):97–9.
- [Aweda et al., ] Aweda, M., Adeyomoye, A., and Abe, G. Thermographic analysis of thyroid diseases at the lagos university teaching hospital, nigeria.
- [Ballester and de Araújo, 2016] Ballester, P. and de Araújo, R. M. (2016). On the performance of googlenet and alexnet applied to sketches. In *AAAI*, pages 1124–1128.
- [Baloch et al., 2008] Baloch, Z. W., LiVolsi, V. A., Asa, S. L., Rosai, J., Merino, M. J., Randolph, G., Vielh, P., DeMay, R. M., Sidawy, M. K., and Frable, W. J. (2008). Diagnostic terminology and morphologic criteria for cytologic diagnosis of thyroid lesions: a synopsis of the national cancer institute thyroid fine-needle aspiration state of the science conference. *Diagnostic cytopathology*, 36(6):425–437.
- [Barnes, 1963] Barnes, R. B. (1963). Thermography of the human body: Infrared-radiant energy provides new concepts and instrumentation for medical diagnosis. *Science*, 140(3569):870–877.
- [Bejnordi et al., 2017] Bejnordi, B. E., Veta, M., Van Diest, P. J., Van Ginneken, B., Karssemeijer, N., Litjens, G., Van Der Laak, J. A., Hermesen, M., Manson, Q. F., Balkenhol, M., et al. (2017). Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *Jama*, 318(22):2199–2210.
- [Bomeli et al., 2010] Bomeli, S. R., LeBeau, S. O., and Ferris, R. L. (2010). Evaluation of a thyroid nodule. *Otolaryngologic clinics of North America*, 43(2):229–238.
- [Borchartt, 2013a] Borchartt, T. B. (2013a). *Análise de Imagens Termográficas para a Classificação de Alteração na Mama*. PhD thesis, Instituto de Computação, Universidade Federal Fluminense, Niterói, RJ, Brasil.
- [Borchartt, 2013b] Borchartt, T. B. (2013b). *Análise de Imagens Termográficas para a Classificação de Alteração na Mama*. PhD thesis, Instituto de Computação, Universidade Federal Fluminense, Niterói, RJ, Brasil.
- [Brenner and Hall, 2007] Brenner, D. J. and Hall, E. J. (2007). Computed tomography—an increasing source of radiation exposure. *New England Journal of Medicine*, 357(22):2277–2284.

- [Brioschi et al., 2010] Brioschi, M., Teixeira, M., Silva, F., and Colman, D. (2010). Medical thermography textbook: principles and applications. *Editores e Livraria Andreoli*.
- [Cho et al., 2015] Cho, J., Lee, K., Shin, E., Choy, G., and Do, S. (2015). Medical image deep learning with hospital pacs dataset. *Arxiv Preprint Arxiv*, 1511.
- [Cohen, 1988] Cohen, J. (1988). Statistical power analysis for the behavioral sciences. 1988, hillsdale, nj: L. *Lawrence Earlbaum Associates*, 2.
- [Cohen, 1995] Cohen, J. (1995). The earth is round ( $p < .05$ ): Rejoinder.
- [Collins et al., 1974] Collins, A., Ring, E., Cosh, J., and Bacon, P. (1974). Quantitation of thermography in arthritis using multi-isothermal analysis. i. the thermographic index. *Annals of the rheumatic diseases*, 33(2):113.
- [Conci et al., 2008] Conci, A., Azevedo, E., and Gráfica, F. L.-C. (2008). volume 2 (processamento e análise de imagens digitais). *Campus/Elsevier*.
- [Dean and Gharib, 2008] Dean, D. S. and Gharib, H. (2008). Epidemiology of thyroid nodules. *Best practice & research Clinical endocrinology & metabolism*, 22(6):901–911.
- [Diakides et al., 2012] Diakides, M., Bronzino, J. D., and Peterson, D. R. (2012). *Medical infrared imaging: principles and practices*. CRC press.
- [Dogdas et al., 2018] Dogdas, B., Forbes, J., Sampath, S., Bagchi, A., and Chin, C.-L. (2018). Transfer learning for the fully automatic segmentation of left ventricle myocardium in porcine cardiac cine mr images. In *Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges: 8th International Workshop, STACOM 2017, Held in Conjunction with MICCAI 2017, Quebec City, Canada, September 10-14, 2017, Revised Selected Papers*, volume 10663, page 21. Springer.
- [Durlak, 2009] Durlak, J. A. (2009). How to select, calculate, and interpret effect sizes. *Journal of pediatric psychology*, 34(9):917–928.
- [Ellis, 2010] Ellis, P. D. (2010). *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results*. Cambridge University Press.
- [Espírito Santo and Daniel, 2017] Espírito Santo, H. and Daniel, F. (2017). Calcular e apresentar tamanhos do efeito em trabalhos científicos (1): As limitações do  $p < 0,05$  na análise de diferenças de médias de dois grupos (calculating and reporting effect sizes on scientific papers (1):  $P < 0.05$  limitations in the analysis of mean differences of two groups).
- [Esteva et al., 2017] 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.
- [Ezzat et al., 1994] Ezzat, S., Sarti, D. A., Cain, D. R., and Braunstein, G. D. (1994). Thyroid incidentalomas: prevalence by palpation and ultrasonography. *Archives of internal medicine*, 154(16):1838–1840.

- [Ferreira et al., 2009] Ferreira, F., Peixoto, J., Maciel, M., Cunha, C., and Souza, D. (2009). Development of a simulator with thyroid nodules for nuclear medicine. *Scientia Plena*, 5(10).
- [Fiirst et al., 2018] Fiirst, W., Conci, A., Moran, M. B. H., Damiao, C., Saade, D. C. M., Araujo, A. S., Filho, R., Gonzalez, J. R., Lima, G., and Resmini, R. (2018). Monomodal image registration by tensor analysis: going beyond the brightness constancy assumption. In *ACS/IEEE International Conference on Computer Systems and Applications (AICCSA)*. ACS/IEEE.
- [Fiirst, 2018] Fiirst, W. G. (2018). Desenvolvimento e comparação de métodos para geração de séries temporais para imagens térmicas da tireoide. Master’s thesis, Instituto de Computação, Universidade Federal Fluminense, Niterói, RJ, Brasil.
- [Fonseca, 2017] Fonseca, L. M. V. A. (2017). Classificação de Objetos em Contexto Real por Redes Neurais Convolutivas. Master’s thesis, PUC-Rio, Rio de Janeiro, RJ, Brasil.
- [Gavriiloaia et al., 2009] Gavriiloaia, G., Ghemigian, A.-M., and Gavriiloaia, M.-R. (2009). Infrared signature analysis of the thyroid tumors. In *European Conference on Biomedical Optics*, page 7371\_1F. Optical Society of America.
- [Goldbach-Mansky et al., 2003] Goldbach-Mansky, R., Woodburn, J., Yao, L., and Lipsky, P. E. (2003). Magnetic resonance imaging in the evaluation of bone damage in rheumatoid arthritis: a more precise image or just a more expensive one? *Arthritis & Rheumatism*, 48(3):585–589.
- [González et al., 2017] González, J., Rodrigues, É., Damião, C., Fontes, C., Silva, A., Paiva, A., Li, H., Du, C., and Conci, A. (2017). An approach for thyroid nodule analysis using thermographic images. In *Application of Infrared to Biomedical Sciences*, pages 451–475. Springer.
- [Gonzáles, 2017] Gonzáles, J. R. (2017). TUm estudo sobre a possibilidade do uso de imagens infravermelhas na análise de nódulos de tireoide. Master’s thesis, Instituto de Computação, Universidade Federal Fluminense, Niterói, RJ, Brasil.
- [Goodfellow et al., 2016] Goodfellow, I., Bengio, Y., Courville, A., and Bengio, Y. (2016). *Deep learning*, volume 1. MIT press Cambridge.
- [He et al., 2016] 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.
- [Heath et al., 2000] Heath, M., Bowyer, K., Kopans, D., Moore, R., and Kegelmeyer, P. (2000). The digital database for screening mammography. *Digital mammography*, pages 431–434.
- [Hedges, 1981] Hedges, L. V. (1981). Distribution theory for glass’s estimator of effect size and related estimators. *journal of Educational Statistics*, 6(2):107–128.
- [Jafari et al., 2016] Jafari, M. H., Karimi, N., Nasr-Esfahani, E., Samavi, S., Soroushmehr, S. M. R., Ward, K., and Najarian, K. (2016). Skin lesion segmentation in clinical images using deep learning. In *Pattern Recognition (ICPR), 2016 23rd International Conference on*, pages 337–342. IEEE.

- [Krizhevsky and Hinton, 2009] Krizhevsky, A. and Hinton, G. (2009). Learning multiple layers of features from tiny images.
- [Kwajiri and Tezuka, 2017] Kwajiri, T. and Tezuka, T. (2017). Classification of lung nodules using deep learning. *Transactions of Japanese Society for Medical and Biological Engineering*, 55(Proc):516–517.
- [Lahiri et al., 2012] Lahiri, B., Bagavathiappan, S., Jayakumar, T., and Philip, J. (2012). Medical applications of infrared thermography: a review. *Infrared Physics & Technology*, 55(4):221–235.
- [Larson, 1999] Larson, C. (1999). Williams’ textbook of endocrinology. *The Endocrinologist*, 9(1):69–70.
- [Lévy and Jain, 2016] Lévy, D. and Jain, A. (2016). Breast mass classification from mammograms using deep convolutional neural networks. *arXiv preprint arXiv:1612.00542*.
- [Lotter et al., 2017] Lotter, W., Sorensen, G., and Cox, D. (2017). A multi-scale cnn and curriculum learning strategy for mammogram classification. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pages 169–177. Springer.
- [Mahajan and Madhe, 2014] Mahajan, P. and Madhe, S. (2014). Morphological feature extraction of thermal images for thyroid detection. *Int. J. Electron. Comput. Eng*, 5(4):11–14.
- [Moran et al., 2018] Moran, M. B. H., Conci, A., Gonzalez, J. R., Araujo, A. S., Fiirst, W., and Damiao, C. L. G. F. R. (2018). Identification of thyroid nodules in infrared images by convolutional neural networks. In *31st International Joint Conference on Neural Networks (IJCNN)*. IEEE.
- [Nguyen et al., 2015] 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.
- [Oliveira Filho, 2017] Oliveira Filho, J. S. d. (2017). Aspectos epidemiológicos e anatomopatológicos de nódulos da tireoide. Master’s thesis, Universidade Federal de Pernambuco.
- [Otsu, 1979] Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, 9(1):62–66.
- [Papini et al., 2002] Papini, E., Guglielmi, R., Bianchini, A., Crescenzi, A., Taccogna, S., Nardi, F., Panunzi, C., Rinaldi, R., Toscano, V., and Pacella, C. M. (2002). Risk of malignancy in nonpalpable thyroid nodules: predictive value of ultrasound and color-doppler features. *The Journal of Clinical Endocrinology & Metabolism*, 87(5):1941–1946.
- [Pedroza Ballesteros, 2008] Pedroza Ballesteros, A. (2008). Manejo del nódulo tiroideo: revisión de la literatura. *Revista Colombiana de Cirugía*, 23(2).



- [Rosário et al., 2013] Rosário, P. W., Ward, L. S., Carvalho, G. A. d., Graf, H., Maciel, R. M. d. B., Maciel, L. M. Z., Maia, A. L. S., and Vaisman, M. (2013). Nódulo tireoidiano e câncer diferenciado de tireoide: atualização do consenso brasileiro. *Arquivos brasileiros de endocrinologia & metabologia*. São Paulo. Vol. 57, n. 4 (jul. 2013), p. 240-264.
- [Rosenthal, 1994] Rosenthal, R. (1994). Parametric measures of effect size. *The handbook of research synthesis*, 621:231–244.
- [Rosnow et al., 2000] Rosnow, R. L., Rosenthal, R., and Rubin, D. B. (2000). Contrasts and correlations in effect-size estimation. *Psychological science*, 11(6):446–453.
- [Schneider and Darcy, 1984] Schneider, A. L. and Darcy, R. E. (1984). Policy implications of using significance tests in evaluation research. *Evaluation Review*, 8(4):573–582.
- [Shaikhina and Khovanova, 2017] Shaikhina, T. and Khovanova, N. A. (2017). Handling limited datasets with neural networks in medical applications: A small-data approach. *Artificial intelligence in medicine*, 75:51–63.
- [Siegel and Castellan Jr, 1988] Siegel, S. and Castellan Jr, N. (1988). Non parametric studies. *New York: McGraw*.
- [Smith-Bindman et al., 2009] Smith-Bindman, R., Lipson, J., Marcus, R., Kim, K.-P., Mahesh, M., Gould, R., De González, A. B., and Miglioretti, D. L. (2009). Radiation dose associated with common computed tomography examinations and the associated lifetime attributable risk of cancer. *Archives of internal medicine*, 169(22):2078–2086.
- [Snyder and Lawson, 1993] Snyder, P. and Lawson, S. (1993). Evaluating results using corrected and uncorrected effect size estimates. *The Journal of Experimental Education*, 61(4):334–349.
- [Sonka et al., 2014] Sonka, M., Hlavac, V., and Boyle, R. (2014). *Image processing, analysis, and machine vision*. Cengage Learning.
- [Szegedy et al., 2015] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A., et al. (2015). Going deeper with convolutions. *Cvpr*.
- [Torres et al., 2002] Torres, O. J. M., de Alcântara Caldas, L. R., Palácio, R. L., de Azevedo, R. P., Pacheco, J. S., Neto, J. L., and Macau, R. P. (2002). Punção aspirativa com agulha fina (paaf) em nódulo da tireóide: análise de 61 casos. *Revista Brasileira de Cancerologia*, 48(4):511–15.
- [Wallace et al., 2003] Wallace, R. B., Stone, M. B., et al. (2003). *Medicare coverage of routine screening for thyroid dysfunction*. National Academies Press.
- [Wang et al., 2015] Wang, D., Otto, C., and Jain, A. K. (2015). Face search at scale: 80 million gallery. *arXiv preprint arXiv:1507.07242*.
- [Whitley and Ball, 2002] Whitley, E. and Ball, J. (2002). Statistics review 6: Nonparametric methods. *Critical care*, 6(6):509.

- [Zhong et al., 2015] Zhong, Z., Jin, L., and Xie, Z. (2015). High performance offline handwritten chinese character recognition using googlenet and directional feature maps. In *Document Analysis and Recognition (ICDAR), 2015 13th International Conference on*, pages 846–850. IEEE.
- [Zhu et al., 2015] Zhu, R., Zhang, R., and Xue, D. (2015). Lesion detection of endoscopy images based on convolutional neural network features. In *Image and Signal Processing (CISP), 2015 8th International Congress on*, pages 372–376. IEEE.

## APÊNDICE A - Acurácia por CNN

Nesta seção são apresentados os resultados de acurácia por época nos experimentos utilizados.

### A.1 Resultados para o conjunto de dados DT1

Época	GoogLeNet	ResNet	Época	GoogLeNet	ResNet
1	0,7312	0,45	26	0,635	0,6947
2	0,7302	0,5508	27	0,6243	0,672
3	0,7039	0,5973	28	0,6082	0,7584
4	0,7141	0,5316	29	0,5885	0,7453
5	0,5915	0,4944	30	0,5316	0,726
6	0,7253	0,5908	31	0,5222	0,7189
7	0,6464	0,6214	32	0,5563	0,7328
8	0,7278	0,5534	33	0,5568	0,7469
9	0,6915	0,5088	34	0,55	0,7653
10	0,6444	0,5517	35	0,56	0,8052
11	0,6434	0,5987	36	0,5661	0,7711
12	0,6861	0,7489	37	0,5551	0,7365
13	0,6856	0,8281	38	0,5671	0,7331
14	0,7038	0,8261	39	0,5417	0,7453
15	0,6412	0,7943	40	0,5555	0,7719
16	0,6302	0,805	41	0,5562	0,7375
17	0,4918	0,8077	42	0,5472	0,7663
18	0,4969	0,8378	43	0,5269	0,7938
19	0,5391	0,8755	44	0,5168	0,7892
20	0,575	0,7849	45	0,5041	0,7759
21	0,5528	0,783	46	0,5089	0,7907

22	0,5507	0,7237	47	0,5347	0,7942
23	0,554	0,7276	48	0,5287	0,8222
24	0,556	0,683	49	0,5175	0,7945
25	0,652	0,6562	50	0,4955	0,8173

Tabela A.1: Acúracia por época, taxa de aprendizado 0,1

Época	GoogLeNet	ResNet	Época	GoogLeNet	ResNet
1	0.5625	0.5062	26	0.4947	0.5459
2	0.5559	0.47	27	0.5073	0.5817
3	0.5882	0.4146	28	0.522	0.5495
4	0.6235	0.4765	29	0.5391	0.5672
5	0.6859	0.5631	30	0.5488	0.5606
6	0.6614	0.5933	31	0.5639	0.5224
7	0.6478	0.5763	32	0.6128	0.5318
8	0.6595	0.5761	33	0.6275	0.5245
9	0.6539	0.542	34	0.6387	0.5131
10	0.6475	0.5567	35	0.6448	0.5224
11	0.6262	0.5636	36	0.6516	0.5239
12	0.646	0.6141	37	0.657	0.5175
13	0.6347	0.664	38	0.659	0.5587
14	0.6853	0.6802	39	0.6625	0.5412
15	0.6213	0.6538	40	0.6615	0.5511
16	0.6348	0.6766	41	0.6482	0.555
17	0.6075	0.6696	42	0.6306	0.5773
18	0.5739	0.7055	43	0.6226	0.5904
19	0.5902	0.695	44	0.6131	0.5907
20	0.5914	0.6892	45	0.6032	0.6063
21	0.5695	0.663	46	0.6059	0.6139
22	0.5621	0.5933	47	0.607	0.6172
23	0.5514	0.5641	48	0.6184	0.6255
24	0.5586	0.5474	49	0.6182	0.6185
25	0.4936	0.5177	50	0.6132	0.6397

Tabela A.2: Acúracia por época, taxa de aprendizado 0,01

Época	GoogLeNet	ResNet	Época	GoogLeNet	ResNet
1	0.5261	0.45	26	0.4878	0.6011
2	0.5482	0.3159	27	0.4931	0.5947
3	0.6993	0.419	28	0.5116	0.598
4	0.5555	0.5475	29	0.5239	0.588
5	0.6372	0.5537	30	0.5456	0.5761
6	0.7147	0.5568	31	0.5443	0.5721
7	0.6403	0.5563	32	0.547	0.6468
8	0.6566	0.487	33	0.5497	0.6671
9	0.7092	0.5285	34	0.5636	0.6695
10	0.6524	0.5556	35	0.5862	0.6689
11	0.6098	0.606	36	0.5928	0.6705
12	0.6296	0.6901	37	0.5968	0.672
13	0.6091	0.6877	38	0.611	0.6709
14	0.6085	0.6738	39	0.6123	0.6663
15	0.6331	0.6651	40	0.6122	0.6646
16	0.628	0.6445	41	0.6122	0.6445
17	0.6149	0.6441	42	0.5892	0.6265
18	0.5701	0.6525	43	0.5781	0.6138
19	0.5644	0.6158	44	0.5545	0.6196
20	0.5454	0.631	45	0.5758	0.5941
21	0.5442	0.6408	46	0.5837	0.6433
22	0.5641	0.6328	47	0.5891	0.6585
23	0.5536	0.6197	48	0.6038	0.659
24	0.54	0.6017	49	0.6064	0.7014
25	0.4952	0.6056	50	0.6074	0.7056

Tabela A.3: Acúrcia por época, taxa de aprendizado 0,001

## A.2 Resultados para o conjunto de dados DT2

Época	GoogLeNet	ResNet	Época	GoogLeNet	ResNet
1	0.5284	0.5489	26	0.4396	0.74
2	0.5	0.6128	27	0.4847	0.7647

3	0.5095	0.5533	28	0.4767	0.7958
4	0.478	0.529	29	0.5102	0.8171
5	0.5642	0.5202	30	0.4894	0.8038
6	0.3405	0.6281	31	0.47	0.7773
7	0.4918	0.6298	32	0.4762	0.7535
8	0.4738	0.5635	33	0.4828	0.7568
9	0.4749	0.5722	34	0.5235	0.7692
10	0.4872	0.694	35	0.4334	0.8015
11	0.4952	0.6941	36	0.4577	0.83
12	0.4717	0.688	37	0.4463	0.8326
13	0.4899	0.6647	38	0.4462	0.8391
14	0.4731	0.6963	39	0.4471	0.8002
15	0.5117	0.6801	40	0.5008	0.8184
16	0.4609	0.6603	41	0.497	0.7886
17	0.5532	0.7182	42	0.4688	0.7918
18	0.5255	0.7127	43	0.4758	0.8025
19	0.5371	0.7414	44	0.5121	0.7628
20	0.5112	0.7074	45	0.438	0.8154
21	0.4796	0.7176	46	0.5609	0.8439
22	0.461	0.6771	47	0.5097	0.8345
23	0.4915	0.5866	48	0.4592	0.8245
24	0.515	0.6025	49	0.4519	0.8151
25	0.4663	0.7138	50	0.4507	0.8216

Tabela A.4: Acúrcia por época, taxa de aprendizado 0,1

Época	GoogLeNet	ResNet	Época	GoogLeNet	ResNet
1	0.5221	0.6886	26	0.5875	0.6225
2	0.6504	0.6444	27	0.5862	0.6195
3	0.4927	0.588	28	0.614	0.6141
4	0.5722	0.5683	29	0.5968	0.5641
5	0.5667	0.6358	30	0.609	0.5774
6	0.4955	0.6623	31	0.6263	0.5808
7	0.5727	0.6406	32	0.615	0.632
8	0.5825	0.6092	33	0.6307	0.6264
9	0.5494	0.6409	34	0.5419	0.6539

10	0.6112	0.6032	35	0.5798	0.6837
11	0.5874	0.6112	36	0.5748	0.7154
12	0.5905	0.6469	37	0.5849	0.7083
13	0.541	0.6252	38	0.5811	0.6633
14	0.597	0.6592	39	0.5926	0.6412
15	0.5784	0.6412	40	0.6186	0.6792
16	0.5702	0.6228	41	0.6426	0.5798
17	0.5864	0.5597	42	0.5735	0.5585
18	0.5806	0.5334	43	0.6174	0.5933
19	0.5886	0.5367	44	0.6517	0.5973
20	0.5753	0.5524	45	0.6337	0.6757
21	0.616	0.5819	46	0.627	0.643
22	0.6501	0.5687	47	0.6262	0.6285
23	0.6688	0.599	48	0.6372	0.642
24	0.636	0.5913	49	0.6218	0.6294
25	0.5615	0.6569	50	0.6646	0.6747

Tabela A.5: Acúracia por época, taxa de aprendizado  
0,01

Época	GoogLeNet	ResNet	Época	GoogLeNet	ResNet
1	0.6234	0.6129	26	0.56	0.6721
2	0.5645	0.6536	27	0.5911	0.6916
3	0.5386	0.5899	28	0.5897	0.6687
4	0.5373	0.5775	29	0.5828	0.6976
5	0.5727	0.6108	30	0.5621	0.7085
6	0.5729	0.5398	31	0.5737	0.6361
7	0.5023	0.5842	32	0.5445	0.668
8	0.5588	0.6302	33	0.5447	0.6903
9	0.5545	0.5895	34	0.5329	0.7124
10	0.5426	0.5803	35	0.5571	0.7302
11	0.5471	0.6331	36	0.5313	0.7185
12	0.5699	0.6493	37	0.5281	0.7001
13	0.5846	0.6777	38	0.5203	0.6679
14	0.5656	0.671	39	0.5474	0.7093
15	0.593	0.651	40	0.5499	0.7383

16	0.622	0.6627	41	0.5659	0.7264
17	0.5711	0.6216	42	0.5355	0.7211
18	0.5631	0.6398	43	0.5551	0.7159
19	0.6139	0.6723	44	0.5614	0.7128
20	0.6075	0.6622	45	0.5632	0.7046
21	0.5735	0.7154	46	0.5851	0.7266
22	0.5842	0.7277	47	0.5525	0.6979
23	0.5895	0.7228	48	0.5517	0.7529
24	0.5612	0.6872	49	0.5751	0.7585
25	0.5747	0.6849	50	0.5476	0.7583

Tabela A.6: Acúracia por época, taxa de aprendizado  
0,001

### A.3 Resultados para o conjunto de dados DT3

Época	GoogLeNet	ResNet	Época	GoogLeNet	ResNet
1	0.6186	0.5031	26	0.5593	0.783
2	0.5717	0.4914	27	0.5747	0.7919
3	0.5495	0.4793	28	0.5928	0.7667
4	0.4897	0.5536	29	0.5658	0.8167
5	0.5185	0.6708	30	0.579	0.824
6	0.5462	0.6869	31	0.5642	0.855
7	0.5539	0.707	32	0.5554	0.8863
8	0.5137	0.7281	33	0.4539	0.8801
9	0.5745	0.7269	34	0.5263	0.8357
10	0.5378	0.7404	35	0.5395	0.762
11	0.5689	0.757	36	0.5595	0.7858
12	0.5707	0.7071	37	0.5564	0.8099
13	0.5411	0.7392	38	0.5371	0.8495
14	0.5285	0.7378	39	0.4677	0.8553
15	0.5719	0.7431	40	0.5247	0.9037
16	0.5542	0.7994	41	0.5866	0.8614
17	0.5758	0.818	42	0.6192	0.8343
18	0.5632	0.8298	43	0.6152	0.8276
19	0.5482	0.8296	44	0.6213	0.8788



20	0.5504	0.8314	45	0.575	0.8648
21	0.5615	0.785	46	0.5551	0.8679
22	0.5707	0.8189	47	0.597	0.9055
23	0.6196	0.7329	48	0.5978	0.9184
24	0.5716	0.7824	49	0.6324	0.9107
25	0.6257	0.7577	50	0.6221	0.8886

Tabela A.7: Acúracia por época, taxa de aprendizado 0,1

Época	GoogLeNet	ResNet	Época	GoogLeNet	ResNet
1	0.6006	0.6856	26	0.5806	0.6878
2	0.5265	0.6065	27	0.6028	0.7551
3	0.6026	0.6187	28	0.6331	0.7786
4	0.5618	0.6677	29	0.6226	0.8062
5	0.537	0.6203	30	0.6341	0.8297
6	0.5331	0.6702	31	0.5511	0.7909
7	0.5319	0.7002	32	0.5619	0.8139
8	0.5559	0.734	33	0.5609	0.7713
9	0.5327	0.7475	34	0.6101	0.7507
10	0.5455	0.6774	35	0.505	0.7593
11	0.5471	0.6942	36	0.519	0.8027
12	0.5563	0.6341	37	0.5855	0.8336
13	0.5601	0.6395	38	0.6011	0.8516
14	0.5543	0.669	39	0.6365	0.8548
15	0.5484	0.6875	40	0.5018	0.8044
16	0.5825	0.713	41	0.5535	0.7737
17	0.5878	0.7735	42	0.5682	0.8191
18	0.5537	0.7646	43	0.5772	0.8176
19	0.581	0.7673	44	0.6514	0.8397
20	0.5587	0.7569	45	0.6552	0.8333
21	0.5684	0.7757	46	0.6618	0.8354
22	0.5947	0.7334	47	0.6603	0.8617
23	0.5894	0.7201	48	0.7255	0.8723
24	0.5644	0.7469	49	0.672	0.8414
25	0.6012	0.7202	50	0.7304	0.8547

Tabela A.8: Acúracia por época, taxa de aprendizado  
0,01

Época	GoogLeNet	ResNet	Época	GoogLeNet	ResNet
1	0.4745	0.6374	26	0.5305	0.7722
2	0.5745	0.5418	27	0.5444	0.7696
3	0.611	0.6652	28	0.5524	0.7569
4	0.6032	0.6055	29	0.5313	0.769
5	0.6244	0.659	30	0.5774	0.7832
6	0.5878	0.7479	31	0.57	0.8025
7	0.5829	0.734	32	0.5546	0.7213
8	0.5831	0.7309	33	0.6092	0.7551
9	0.5872	0.7227	34	0.5882	0.7644
10	0.6092	0.7499	35	0.5644	0.7676
11	0.5937	0.7287	36	0.5692	0.7121
12	0.5612	0.7251	37	0.5701	0.7905
13	0.5808	0.7215	38	0.5775	0.8294
14	0.5857	0.6327	39	0.5513	0.8248
15	0.5678	0.6118	40	0.5516	0.7731
16	0.6055	0.705	41	0.6127	0.8069
17	0.5748	0.7433	42	0.6056	0.8015
18	0.5935	0.723	43	0.5883	0.8045
19	0.5762	0.7119	44	0.604	0.7471
20	0.5638	0.7836	45	0.6373	0.7266
21	0.5495	0.7641	46	0.5985	0.7186
22	0.5506	0.7415	47	0.613	0.7991
23	0.5598	0.7331	48	0.633	0.8204
24	0.556	0.7496	49	0.5884	0.8082
25	0.5638	0.8097	50	0.5858	0.7923

Tabela A.9: Acúracia por época, taxa de aprendizado  
0,001

## A.4 Resultados para o conjunto de dados DT4

Época	GoogLeNet	ResNet	Época	GoogLeNet	ResNet
1	0.6101	0.5128	26	0.5264	0.8946
2	0.5176	0.6259	27	0.5258	0.8396
3	0.5601	0.6697	28	0.5597	0.8832
4	0.5735	0.7497	29	0.5211	0.9058
5	0.5586	0.7475	30	0.5499	0.8827
6	0.6077	0.7658	31	0.609	0.8823
7	0.6154	0.7483	32	0.4542	0.8303
8	0.5616	0.7923	33	0.5528	0.9242
9	0.5943	0.8308	34	0.5686	0.8918
10	0.5496	0.8188	35	0.5287	0.9014
11	0.5763	0.8094	36	0.6053	0.9228
12	0.5102	0.7935	37	0.4953	0.9004
13	0.5093	0.8509	38	0.5721	0.9223
14	0.5656	0.8473	39	0.4848	0.8491
15	0.5946	0.8262	40	0.5548	0.9012
16	0.5314	0.8337	41	0.5088	0.8875
17	0.5871	0.8126	42	0.5142	0.8903
18	0.5792	0.8709	43	0.5376	0.9214
19	0.5709	0.7847	44	0.5578	0.9028
20	0.5975	0.8598	45	0.5431	0.8839
21	0.5025	0.8324	46	0.5393	0.919
22	0.5997	0.8706	47	0.516	0.9008
23	0.5213	0.8594	48	0.5566	0.9365
24	0.4969	0.8221	49	0.5527	0.9504
25	0.5714	0.8429	50	0.4431	0.9584

Tabela A.10: Acúracia por época, taxa de aprendizado

0,1

Época	GoogLeNet	ResNet	Época	GoogLeNet	ResNet
1	0.5914	0.6286	26	0.8244	0.8404
2	0.5773	0.6205	27	0.8331	0.8433
3	0.6069	0.693	28	0.8403	0.8602

4	0.5588	0.7522	29	0.8059	0.8697
5	0.5672	0.7692	30	0.8627	0.8276
6	0.6481	0.7215	31	0.8313	0.8611
7	0.7059	0.7778	32	0.8425	0.8547
8	0.7229	0.7755	33	0.8564	0.8929
9	0.7657	0.7553	34	0.8641	0.8716
10	0.7238	0.8219	35	0.8586	0.8443
11	0.7684	0.7946	36	0.829	0.865
12	0.7246	0.7865	37	0.839	0.8973
13	0.692	0.7951	38	0.8829	0.9091
14	0.8368	0.7786	39	0.8844	0.9178
15	0.7654	0.8343	40	0.8894	0.9291
16	0.7761	0.8058	41	0.8711	0.8651
17	0.7965	0.8141	42	0.8582	0.8435
18	0.7975	0.7917	43	0.8568	0.855
19	0.8045	0.8457	44	0.8221	0.8906
20	0.7436	0.8431	45	0.8214	0.887
21	0.7815	0.836	46	0.8419	0.8873
22	0.7616	0.8032	47	0.8492	0.8732
23	0.8208	0.8441	48	0.8844	0.933
24	0.8466	0.8244	49	0.8788	0.9442
25	0.832	0.8603	50	0.8577	0.9444

Tabela A.11: Acúrcia por época, taxa de aprendizado  
0,01

Época	GoogLeNet	ResNet	Época	GoogLeNet	ResNet
1	0.5232	0.5621	26	0.6881	0.8588
2	0.6562	0.7456	27	0.6735	0.8392
3	0.514	0.7377	28	0.747	0.8891
4	0.5894	0.7591	29	0.7499	0.8772
5	0.5707	0.7854	30	0.7292	0.9019
6	0.558	0.7821	31	0.7074	0.8844
7	0.5612	0.8095	32	0.6991	0.8514
8	0.5889	0.8087	33	0.7184	0.8848
9	0.5782	0.8559	34	0.7248	0.8708

10	0.5338	0.8409	35	0.738	0.8915
11	0.5703	0.7917	36	0.7547	0.9087
12	0.5925	0.7743	37	0.7275	0.9213
13	0.558	0.8067	38	0.7653	0.8836
14	0.5636	0.8632	39	0.8056	0.9048
15	0.5611	0.8015	40	0.7827	0.9043
16	0.5787	0.7897	41	0.7507	0.9044
17	0.5689	0.8233	42	0.7783	0.8855
18	0.5577	0.8488	43	0.7986	0.8778
19	0.551	0.8389	44	0.7897	0.8637
20	0.5636	0.8851	45	0.7738	0.8834
21	0.565	0.8612	46	0.7458	0.9079
22	0.5696	0.8431	47	0.7686	0.9311
23	0.5521	0.8754	48	0.8245	0.9331
24	0.6265	0.8718	49	0.7887	0.9434
25	0.6973	0.8445	50	0.7607	0.92

Tabela A.12: Acúracia por época, taxa de aprendizado 0,001

## A.5 Resultados para o conjunto de dados DT5

Época	GoogLeNet	ResNet	Época	GoogLeNet	ResNet
1	0.5514	0.6974	26	0.5497	0.7176
2	0.5859	0.7102	27	0.5292	0.6808
3	0.5534	0.6981	28	0.5398	0.7113
4	0.4962	0.6575	29	0.5404	0.7455
5	0.5811	0.6389	30	0.5516	0.7371
6	0.4948	0.6754	31	0.5673	0.7235
7	0.514	0.6805	32	0.5922	0.7018
8	0.5415	0.6815	33	0.545	0.7668
9	0.5419	0.7022	34	0.4573	0.8069
10	0.5585	0.7147	35	0.5405	0.7666
11	0.4865	0.6921	36	0.5529	0.7632
12	0.5419	0.6834	37	0.5366	0.7664
13	0.5445	0.6959	38	0.5497	0.7211

14	0.5738	0.7229	39	0.5387	0.7542
15	0.4791	0.7055	40	0.5316	0.7583
16	0.5745	0.7152	41	0.4886	0.7954
17	0.4901	0.6791	42	0.5731	0.7899
18	0.5579	0.6808	43	0.5396	0.8422
19	0.5645	0.684	44	0.4493	0.8081
20	0.5798	0.7117	45	0.4663	0.8292
21	0.5265	0.6729	46	0.4689	0.8227
22	0.5192	0.7175	47	0.5394	0.8298
23	0.5282	0.6937	48	0.5538	0.8043
24	0.5177	0.7408	49	0.5596	0.818
25	0.518	0.7549	50	0.5424	0.8898

Tabela A.13: Acúracia por época, taxa de aprendizado

0,1

Época	GoogLeNet	ResNet	Época	GoogLeNet	ResNet
1	0.5586	0.6899	26	0.5923	0.7172
2	0.5739	0.6385	27	0.5539	0.7291
3	0.5637	0.6892	28	0.5839	0.7225
4	0.5729	0.6533	29	0.5419	0.7068
5	0.5429	0.7002	30	0.4768	0.7607
6	0.5217	0.6807	31	0.5688	0.7907
7	0.5398	0.6937	32	0.5388	0.75
8	0.5629	0.6733	33	0.5673	0.7737
9	0.5239	0.6867	34	0.5971	0.7797
10	0.5611	0.6951	35	0.6329	0.8034
11	0.5565	0.6945	36	0.5847	0.7878
12	0.5697	0.713	37	0.6705	0.7468
13	0.5431	0.7332	38	0.5651	0.753
14	0.5743	0.743	39	0.6299	0.8492
15	0.5724	0.722	40	0.6513	0.7855
16	0.5543	0.6778	41	0.6474	0.8355
17	0.5449	0.7029	42	0.6366	0.8418
18	0.5143	0.7028	43	0.6335	0.8686
19	0.5859	0.6709	44	0.6186	0.8721

20	0.5369	0.6926	45	0.5621	0.9042
21	0.5689	0.6762	46	0.5855	0.9185
22	0.5561	0.7394	47	0.672	0.8725
23	0.5812	0.7109	48	0.6149	0.8702
24	0.5844	0.8054	49	0.685	0.901
25	0.5665	0.7565	50	0.6834	0.9253

Tabela A.14: Acúracia por época, taxa de aprendizado  
0,01

Época	GoogLeNet	ResNet	Época	GoogLeNet	ResNet
1	0.5498	0.7043	26	0.5252	0.8034
2	0.5677	0.636	27	0.5424	0.8111
3	0.5058	0.6175	28	0.535	0.7621
4	0.5542	0.6395	29	0.5304	0.7928
5	0.5294	0.6664	30	0.5427	0.8411
6	0.5444	0.6446	31	0.5507	0.8325
7	0.5553	0.6619	32	0.5352	0.8484
8	0.5593	0.6754	33	0.5417	0.8787
9	0.5365	0.6941	34	0.5372	0.8295
10	0.5377	0.6561	35	0.5164	0.8747
11	0.5003	0.6717	36	0.5579	0.8835
12	0.5294	0.674	37	0.5353	0.8936
13	0.5666	0.7032	38	0.5476	0.9005
14	0.5351	0.7495	39	0.5132	0.9389
15	0.557	0.7183	40	0.5483	0.9333
16	0.529	0.7199	41	0.6028	0.8673
17	0.5354	0.7778	42	0.5304	0.8596
18	0.5607	0.7269	43	0.558	0.8597
19	0.5344	0.7483	44	0.5476	0.9087
20	0.5489	0.7881	45	0.5241	0.9257
21	0.5534	0.7407	46	0.5268	0.9018
22	0.5212	0.7872	47	0.572	0.8894
23	0.5375	0.7664	48	0.5437	0.8357
24	0.5487	0.7919	49	0.5388	0.9194

---

25	0.4986	0.8428	50	0.5444	0.943
----	--------	--------	----	--------	-------

Tabela A.15: Acúracia por época, taxa de aprendizado  
0,001



## APÊNDICE B - Protocolo de aquisição

Os termogramas usados neste estudo foram adquiridos através do protocolo definido em trabalhos anteriores do grupo VisualLab [Gonzáles, 2017, Fiirst, 2018], . A definição desse protocolo tem como objetivo padronizar o processo de captura, minimizando possíveis erros e variações. A aquisição desses dados foi previamente aprovado pelo Comitê de Ética do Hospital Universitário Antônio Pedro (HUAP) e pelo Ministério da Saúde do Brasil, sob o número do Certificado de Apresentação para Apreciação Ética (CAAE) 57078516.8.0000.5243. A seguir uma visão geral do protocolo de aquisição é mostrada.

O primeiro passo do protocolo consiste no preenchimento do Termo de Consentimento Livre e Esclarecido (TCLE) por parte do paciente. Através dele o paciente concorda com sua participação no estudo. O local da captura tem sua temperatura controlada entre 20 e 25°C [Gonzáles, 2017].

O paciente permanece no local de captura por 10 minutos, e qualquer acessório que possa vir a interferir na captura é removido, como por exemplo brincos, colares, etc. Além disso, o cabelo do paciente deve ser preso se necessário a fim de não atrapalhar a captura. Após isso, com o paciente sentado, seu pescoço é posicionado para atrás, apoiando a nuca num suporte desenvolvido para este processo, de forma a destacar a região da tireoide [Fiirst, 2018]. A câmera é posicionada a cerca de 0,5 a 0,6 metros de distancia do paciente, apontando para seu pescoço. A temperatura do pescoço do paciente é reduzida a 31,5 °C utilizando um ventilador direcionado a essa região. Tendo atingido essa temperatura, o ventilador é desligado, e a captura é imediatamente iniciada. Uma imagem é capturada a cada 15 segundos durante 5 minutos, resultando numa sequencia de 20 imagens [Gonzáles, 2017, Fiirst, 2018].