

UNIVERSIDADE FEDERAL DO PARANÁ

PEDRO HENRIQUE RAMOS PRADO

DETECÇÃO DO VENTRÍCULO ESQUERDO DO CORAÇÃO PARA
GERAÇÃO DE IMAGENS HÍBRIDAS MULTIMODAIS COM
AQUISIÇÃO ATEMPORAL

CURITIBA

2018

PEDRO HENRIQUE RAMOS PRADO

DETECÇÃO DO VENTRÍCULO ESQUERDO DO CORAÇÃO PARA
GERAÇÃO DE IMAGENS HÍBRIDAS MULTIMODAIS COM
AQUISIÇÃO ATEMPORAL

Dissertação apresentada como requisito parcial à obtenção do grau de Mestre em Informática no Programa de Pós-Graduação em Informática, Setor de Ciências Exatas, da Universidade Federal do Paraná.
Orientador: Prof. Dr. Luciano Silva

CURITIBA

2018

FICHA CATALOGRÁFICA ELABORADA PELO SISTEMA DE BIBLIOTECAS/UFPR
BIBLIOTECA DE CIÊNCIA E TECNOLOGIA

P896d Prado , Pedro Henrique Ramos
Detecção do ventrículo esquerdo do coração para geração de imagens híbridas multimodais com
aquisição atemporal / Pedro Henrique Ramos Prado . – Curitiba, 2018.
63 p. : il. color. ; 30 cm.

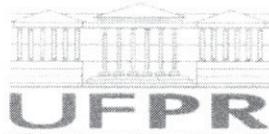
Dissertação - Universidade Federal do Paraná, Setor de Ciências Exatas, Programa de Pós-
Graduação em Informática, 2018.

Orientador: Luciano Silva.

1. Segmentos cardíacos. 2. Aprendizagem de máquina. 3. Imagens médicas híbridas. 4. Fusão
multimodal atemporal. I. Universidade Federal do Paraná. II. Silva, Luciano. III. Título.

CDD: 004.19

Bibliotecária: Romilda Santos - CRB-9/1214



MINISTÉRIO DA EDUCAÇÃO
SETOR CIÊNCIAS EXATAS
UNIVERSIDADE FEDERAL DO PARANÁ
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO INFORMÁTICA

TERMO DE APROVAÇÃO

Os membros da Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em INFORMÁTICA da Universidade Federal do Paraná foram convocados para realizar a arguição da Dissertação de Mestrado de **PEDRO HENRIQUE RAMOS PRADO** intitulada: **Deteção do Ventrículo Esquerdo do Coração para Geração de Imagens Híbridas Multimodais com Aquisição Atemporal**, após terem inquirido o aluno e realizado a avaliação do trabalho, são de parecer pela sua APROVAÇÃO no rito de defesa.

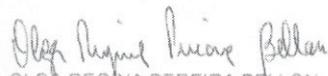
A outorga do título de mestre está sujeita à homologação pelo colegiado, ao atendimento de todas as indicações e correções solicitadas pela banca e ao pleno atendimento das demandas regimentais do Programa de Pós-Graduação.

Curitiba, 09 de Março de 2018.



LUCIANO SILVA

Presidente da Banca Examinadora (UFPR)



OLGA REGINA PEREIRA BELLON

Avaliador Interno (UFPR)



MIGUEL MORITA FERNANDES DA SILVA

Avaliador Externo (PUC/PR)



RESUMO

Este trabalho apresenta um estudo sobre o processamento de imagens cardíacas multimodais com aquisição em equipamentos e datas distintos para geração de imagens híbridas. Foi desenvolvida uma abordagem não supervisionada, baseada em aprendizagem profunda computacional, para a detecção e reorientação do ventrículo esquerdo (VE) do coração em imagens anatômicas e funcionais das técnicas SPECT e CT. Esta abordagem permite a geração precisa dos eixos curto, longo vertical e longo horizontal do VE, em ambas as modalidades, fundamentais no auxílio do diagnóstico médico. A abordagem proposta se mostrou promissora quando comparada com métodos estado da arte, atingindo 99% de acurácia na detecção e 100% na reorientação do VE usando uma base de imagens com diferentes casos clínicos, anotadas por médicos especialistas das áreas de medicina nuclear e radiologia. Dentro do nosso conhecimento, este é o primeiro trabalho que utiliza métodos deep learning para detecção do ápice e base do ventrículo esquerdo, com consequente reorientação e geração de imagens médicas híbridas, utilizando cortes transaxiais das técnicas de imagem médica, SPECT e CT. Resultados experimentais na geração de imagens híbridas se mostram efetivos devido a precisão obtida pela abordagem proposta.

Palavras-chave: detecção segmentos cardíacos; aprendizagem de máquina; imagens médicas híbridas; fusão multimodal atemporal

ABSTRACT

We present a study on the processing of multimodal cardiac images with equipment acquisition and different dates for hybrid image generation. An unsupervised approach, based on deep computational learning, was developed for the detection and reorientation of the left ventricle (LV) of the heart in anatomical and functional images of the techniques: SPECT and CT. This approach allows the precise generation of the short, long vertical and long horizontal axes of the LV, in both modalities, fundamental in the aid of the medical diagnosis. The proposed approach was promising when compared to state-of-the-art methods, reaching 99% accuracy in detection and 100% in LV reorientation using a database of different clinical cases, noted by specialist physicians in the areas of nuclear medicine and radiology. As far as we know, this is the first work that uses deep learning methods to detect the apex and base of the left ventricle, with consequent reorientation and generation of hybrid medical images, using transaxial sections of the medical imaging techniques, SPECT and CT. Experimental results in the generation of hybrid images are shown to be effective due to the precision obtained by the proposed approach.

Keywords: cardiac segments detection; machine learning; hybrid medical images; timeless multimodal fusion

LISTA DE FIGURAS

2.1	Exemplo da Faster R-CNN Ren et al. (2017) para detecção de objetos em conjunto com a rede RPN. É iniciado com a extração das características, RPN gerando as regiões propostas e rede CNN para classificação e regressão das regiões e dos objetos. Adaptado de Ren et al. (2015)	20
2.2	O limiar inicial do <i>threshold</i> utilizado no processo de agrupamento é baseado no valor de intensidade do voxel no quadrante superior direito do volume das imagens transaxiais, evidenciado em vermelho. Adaptado de Germano et al. (1995).	21
3.1	Gráfico com acurácia da rede para detecção do VE nas imagens SPECT reconstruídas com diferentes filtros e interpolações	24
4.1	Interface da ferramenta desenvolvida para anotação das imagens SPECT. As <i>bounding boxes</i> vermelha, amarela e verde, representam o ápice, extremidade da base esquerda e direita respectivamente.	25
4.2	Parâmetros para geração/construção de volumes tridimensionais. A) Campo de visão em mm obtidos no arquivo DICOM; B) Largura da matriz obtidos no arquivo DICOM ou dimensões da imagem; C) Representação do <i>voxel</i> , espessura do <i>slice</i> , ou (<i>thickness</i>).	28
4.3	Ilustração das regiões de predição da rede. Em azul <i>bounding-box</i> do VE e verde <i>bounding-box</i> do ápice obtidas pelos modelos treinados. Os pontos vermelhos representam o centro da BB do ápice e extremidade inferior esquerda do VE, respectivamente. A linha amarela representa o eixo longo do VE obtido pelos pontos A e B.	30

4.4	Exemplo de <i>slices</i> feitos perpendicularmente ao eixo longo do VE em diferentes espessuras/direções do volume. Na figura, podemos ver os eixos: A) curto, B) longo vertical e C) longo horizontal das imagens tomográficas axiais da técnica SPECT. As imagens ilustradas utilizam <i>colormap</i> , ou mapa de cores, para coloração dos <i>pixels</i> de acordo com a intensidade dos mesmos. Os níveis de cinza da imagem, são definidos pela quantidade de fármaco na aquisição.	31
4.5	Gráfico com os erros totais (classificação e regressão) nas imagens SPECT	32
4.6	O IoU é simplesmente calculado dividindo a área de sobreposição entre <i>bouding boxes</i> , pela área de união.	33
4.7	Gráfico com resultados da sobreposição entre as imagens da base <i>ground truth</i> e regiões preditas. O eixo IoU contempla a média da soma dos segmentos ápice e base no decorrer das épocas.	34
4.8	Gráfico com distribuição dos resultados entre GT e predição do ápice utilizando a métrica IoU. Apesar de alguns resultados ficarem próximos de 0.5 (resultado bom), ficam evidenciados resultados onde o modelo obteve uma ótima predição de até 0.97 de sobreposição.	34
4.9	Exemplo de retorno de falsos positivos em imagens SPECT. A seta vermelha aponta a detecção errada, enquanto que a verde a detecção correta, já que seu classificador teve um maior valor.	36
4.10	Ilustração dos eixos. A) <i>Slices</i> do eixo longo vertical ilustrando a extremidade do ápice em 0°. B) <i>Slices</i> do eixo longo horizontal ilustrando a extremidade do ápice em 90°. C) Círculo com graus e radianos. Para melhor visualização, foram evidenciados os ângulo 90° em azul e 0° em laranja.	37
4.11	Gráfico de correlação entre RMS (distância <i>pixels</i>) e IoU das imagens SPECT.	38

4.12	Exemplo da análise de resultados do eixo longo horizontal. A) Respe- tivamente, corte perpendicular (<i>Slice</i>) do eixo longo horizontal após reo- rientação, segmentação do VE, e ângulo do ápice calculado da base GT; B) Respetivamente, corte perpendicular (<i>Slice</i>) do eixo longo horizontal após reorientação, segmentação do VE, e ângulo do ápice calculado da base de predição; C) Representação dos ângulos encontrados em graus, levando como base o maior eixo da elipse detectada na imagem segmentada do VE. Em vermelho ângulo da base GT, e em verde da reorientação realizada com parâmetros do modelo treinado.	39
4.13	Gráfico de Bland e Altman. Concordância entre base anotada por espe- cialistas e reorientação proposta, considerando os resultados dos ângulos obtidos após reorientação.	40
5.1	Em azul <i>bouding-box</i> em torno do VE. Em vermelho, reta paralela traçadas com base nos pontos amarelos gerados a partir das coordenadas $A(x_1, y_1)$ e coordenadas do ápice $B(x_2, y_2)$ representando o eixo longo.	42
5.2	A) Imagem CT tomográfica do VE no corte axial. B) Predição da <i>bouding</i> <i>box</i> em azul, em torno do VE, <i>bouding box</i> amarela em torno do ápice e em vermelho o centro da <i>bouding box</i> do ápice. Em rosa eixo longo do VE.	43
5.3	Exemplo de <i>slices</i> dos eixos A) curto, B) longo vertical e C) longo horizontal obtidos após reorientação do volume CT.	43
5.4	A seta vermela indica a predição com menor classificador, enquanto a seta verde indica a detecção correta.	44
5.5	Gráfico com os erros totais (classificação e regressão) nas imagens CT . . .	45
5.6	Gráfico com melhores resultados da interseção da sobreposição entre base <i>ground truth</i> e regiões preditas das imagens CT. Escala IoU médio para 50 épocas 100 iterações. O eixo IoU contempla a média da soma do ápice e base no decorrer das épocas.	46

5.7	Gráfico com distribuição dos resultados entre GT e predição do ápice utilizando a métrica IoU. Apesar de alguns resultados ficarem abaixo de 0.5 (resultado bom/considerado correto), ficam evidenciados resultados onde o modelo obteve uma ótima predição de até 0.91 de sobreposição entre GT e a predição.	47
5.8	Gráfico de correlação	47
6.1	Pipeline para fusão 2D das imagens das modalidades SPECT e CT	49
6.2	Registro entre as imagens utilizando método rígido	50
6.3	Registro das imagens SPECT e CT utilizando utilizando método não rígido	52
6.4	Gráfico de correlação entre RSM e IoU de ambas técnicas SPECT e CT. A correlação do SPECT, foi dividida em duas classes, casos normais e casos anormais.	53

LISTA DE TABELAS

4.1	Resultados de detecção nas imagens SPECT. Taxa de acerto detecções corretas.	35
4.2	Resultados de detecção do Ápice nas imagens CT. Taxa de acerto para detecções corretas.	35
4.3	Resultados médio da reorientação HLA e VLA SPECT (10°)	37
5.1	Resultados de detecção do VE nas imagens CT. Taxa de acerto para detecções corretas.	44
6.1	Média dos resultados de erro de distância dos <i>pixels</i> entre modalidades . . .	51

SUMÁRIO

1	INTRODUÇÃO	11
2	TRABALHOS RELACIONADOS	16
2.1	Faster R-CNN	19
2.2	Segmentação e reorientação do VE em imagens SPECT	21
3	RECONSTRUÇÃO DE IMAGENS DA TÉCNICA SPECT	23
3.1	Experimentos	23
4	DETECÇÃO DO VE E ÁPICE NAS IMAGENS SPECT	25
4.1	<i>Ground truth</i>	25
4.2	Treinamento e otimização da <i>CNN</i>	26
4.3	Reorientação do VE	27
4.4	Experimentos de detecção do VE	32
5	DETECÇÃO DO VE E ÁPICE NAS IMAGENS CT	41
5.1	<i>Ground truth</i>	41
5.2	Treinamento e otimização das <i>CNNs</i>	41
5.3	Reorientação do VE nas imagens CT	42
5.4	Experimentos	44
6	FUSÃO 2D DAS IMAGENS FUNCIONAIS E ANATÔMICAS	49
6.1	Experimentos realizados	50
7	CONCLUSÃO	54
7.1	Trabalhos futuros	55
	REFERÊNCIAS	56

CAPÍTULO 1

INTRODUÇÃO

A DAC (Doença Arterial Coronariana) é uma das principais causas de morte no Brasil e no mundo. Dados publicados pela OMS, demonstraram que somente no ano de 2015, aproximadamente 15 milhões de pessoas morreram por decorrências de problemas cardíacos e derrames, colocando essas doenças em primeiro lugar do ranking de mortalidade mundial (World Health Organization) (2017).

Um diagnóstico precoce, por meio de avaliação clínica dos fatores de risco em associação aos métodos não invasivos, como técnicas de diagnóstico por imagem, são de fundamental importância. Esses métodos viabilizam a detecção de doenças cardíacas que levam à morte. Quando essas doenças apresentam comprometimento da oferta de sangue ao miocárdio, o tratamento em tempo hábil pode reverter ou melhorar o baixo fluxo sanguíneo no VE (ventrículo esquerdo).

É por meio da observação do VE, que várias doenças cardíacas são diagnosticadas, sendo, no entanto, necessária uma análise de imagens, para que se percebam as suas modificações em comparação a um saudável Otto (2013).

Para realização deste trabalho, foram selecionadas imagens de duas técnicas tomográficas amplamente utilizadas no diagnóstico por imagem. Para a técnica de Tomografia Computadorizada por Emissão de Fóton Único ou SPECT, do inglês, Single Photon Emission Computed Tomography Mettler and Guiberteau (2011), foram utilizadas imagens do exame de cintilografia do miocárdio. Para a técnica de tomografia computadorizada ou CT, do inglês, Computed Tomography Goldman (2007), imagens de exames de angiogramia coronariana, foram utilizadas.

Os parâmetros de aquisição de cada exame/estudo utilizados, são armazenados em arquivos DICOM (Digital Imaging and Communications in Medicine) Mildemberger et al. (2002), arquivos médicos com informações de configuração, dados de paciente (para este

trabalho, anonimizados) e *raw data* das imagens. Ainda, os equipamentos de aquisição de imagens médicas e protocolos de aquisição, são constantemente calibrados e revisados, seguindo boas práticas e normas vistas em Henzlova et al. (2016).

As imagens utilizadas no trabalho foram adquiridas em equipamentos e datas distintos, desta maneira, o fator tempo entre a realização dos exames não é primordial na geração das imagens híbridas.

Os exames de cintilografia do miocárdio, comumente, são feitos em duas etapas, stress e repouso. O objetivo de fusão das técnicas SPECT e CT é demonstrar se há ou não repercussão funcional, ou isquemia, recorrente de uma lesão anatômica. Essa repercussão funcional é demonstrada na imagem da etapa de stress, por isso, apenas imagens dessa etapa foram utilizadas no trabalho, pois dessa etapa, especialistas de medicina nuclear e radiologia podem evidenciar os riscos miocárdicos.

Equipamentos de imagens médicas híbridas tornam mais caro e de difícil acompanhamento tecnológico a difusão na utilização de imagens fusionadas funcionais e anatômicas na prática clínica Kaufmann and Carli (2009). De maneira isolada, as técnicas de imagens funcionais e anatômicas oferecem importantes valores prognósticos para médicos especialistas e clínicos. Porém, a fusão dessas técnicas oferece valores adicionais, tornando-se amplamente aceitas e utilizadas na prática clínica.

Os valores prognósticos fornecidos por imagens médicas cardíacas híbridas SPECT/CT auxiliam na evidenciação de eventos cardíacos adversos maiores (ECAM). Realizar um tratamento de intervenção coronária nos achados fisiológicos encontrados nos exames de cintilografia, que foram identificados com auxílio da imagem de fusão, tem uma chance de reduzir a severidade do ECAM no paciente e melhorar o prognóstico Gaemperli et al. (2007).

Encontra-se também, em outros trabalhos, a comprovação que as imagens híbridas têm vantagens como permitir a identificação de uma isquemia em pacientes com doença cardíaca isquêmica. A avaliação de que a isquemia está associada à estenose coronária pela correspondência de imagem de fusão é extremamente importante na consideração da estratégia de tratamento de intervenção coronariana Matsuo et al. (2009); Pazhenkottil

et al. (2011).

Um outro ponto importante para discussão é a qualidade das imagens geradas nos equipamentos híbridos. Com intuito de minimizar a radiação exposta ao paciente, protocolos de aquisição nesses equipamentos, utilizam menores doses de radiação, porém, aumentam a possibilidade de desalinhamento entre dados de emissão e transmissão. Esse desvio representa o risco de geração de artefatos e ruídos artificiais na reconstrução das imagens e exigem maiores cuidados no controle de qualidade das mesmas Flotats et al. (2011). Esse fato também potencializa a necessidade no desenvolvimento de novos métodos para geração de imagens híbridas, cuja imagens de cada técnica SPECT e CT foram adquiridas em datas diferentes.

Já foram expostas algumas vantagens na utilização de técnicas de imagens híbridas no diagnóstico por imagem de exames cardíacos. Outros fatores importantes para realização desses diagnósticos são os filtros, métodos de interpolação e reorientação do ventrículo esquerdo nas imagens. Atualmente na literatura médica, mesmo que esses fatores possam ter impacto na acurácia diagnóstica Haddad and Porenta (1998); Figueiredo and Frago Costa (2014), observa-se que algumas etapas no processamento das imagens seguem critérios subjetivos. Apesar de não dispor de métricas para mensuração se a reorientação está adequada, seja ela realizada de maneira automática por algoritmos ou manual por técnicos e/ou médicos especialistas, podemos encontrar normas e padrões em Cerqueira et al. (2002). Padrões esses utilizados para o desenvolvimento da abordagem proposta.

Diversos fatores e etapas não triviais são importantes na geração de imagens híbridas, sejam elas adquiridas de maneira simultânea em equipamentos híbridos ou de maneira atemporal, em datas distintas. Como visto em Kaufmann and Di Carli (2009); Tang et al. (2006), pode-se ressaltar problemas como artefatos e ruídos nas imagens, movimentação do paciente durante aquisição das projeções e intervenção excessiva por técnicos radiologistas e de medicina nuclear, para o processamento das imagens. Verifica-se ainda outros problemas relacionados, como a detecção/segmentação das estruturas ou definição da ROI (*region of interest*), e registro entre as técnicas de imagens Oliveira and Tavares (2014). Problemas esses que o método proposto pretende sanar em ambas técnicas SPECT e CT.

A detecção do VE (ápice e base) define a localização do mesmo em imagens do plano transaxial das técnicas SPECT e CT, como ilustrados nas figuras 4.3 e 5.1. Essa detecção consiste basicamente em um processo de classificação/extração de padrões nas imagens. Quando entramos no mérito de métodos para extração de padrões e detecção de objetos, podemos observar na literatura diversos trabalhos que abordam o assunto, como visto em Zavan (2016), onde o autor utiliza o método para detecção do nariz. Ainda em Zhang et al. (2016) para detecção de células cancerígenas e Akselrod-Ballin et al. (2016) na detecção e classificação de tumores em mamografias.

Quando falamos em detecção de padrões, extração de características, regressão (previsão de resultados) e classificação de objetos, as técnicas de *deep learning*, traduzindo, aprendizado profundo, têm demonstrado importantes avanços na área de visão computacional. As CNNs (Convolutional Neural Networks), em português, redes neurais convolucionais, são usadas com sucesso em vários problemas envolvendo o reconhecimento de imagens e objetos, como visto em Curiale et al. (2017), Girshick et al. (2014), Sermanet et al. (2013) e Chéron et al. (2015).

A motivação para usar arquiteturas CNN, surgiu como uma possibilidade de tratar abstrações de alto nível de maneira eficiente, como o caso da identificação do VE em imagens médicas. Como esse é um dos passos importantes para a fusão, optou-se em utilizar uma das técnicas de redes convolucionais com bons resultados de classificação e desempenho, o método estado-da-arte Faster-RCNN Ren et al. (2017). Um método que utiliza o conceito RPN (*Region Proposal Networks*) em conjunto com CNN de detecção.

Embora o atual estado da arte para reorientação do VE em imagens da técnica SPECT Germano et al. (1995) tenha uma boa acurácia na detecção e segmentação, o mesmo tem limitações como a localização do VE na imagem e intensidade das mesmas, como demonstrado no capítulo 2. Limitações essas que podem ser minimizadas ou eliminadas com o uso de CNNs. Ainda, o método proposto por Guido *et al.* 1995, contempla somente a detecção e reorientação para uma única técnica de imagem médica e não fornece parâmetros que possibilitam a fusão multimodal para geração das imagens híbridas.

O presente trabalho propõe realizar a reconstrução das projeções SPECT analisando

quais melhores filtros e parâmetros para treinamento de redes neurais CNN. Também a detecção do ápice e base do ventrículo esquerdo nas imagens tomográficas transaxiais das técnicas SPECT e CT, conseqüente definição do eixo longo do ventrículo esquerdo. Ainda, a geração dos *slices* do eixos curto, longo horizontal e longo vertical de maneira perpendicular ao eixo longo, e calcular o ângulo necessário para geração de *slices* reorientados, tanto para as imagens da técnica SPECT quanto para CT. Por final, oferecer parâmetros iniciais para a fusão/registo das imagens entre as técnicas SPECT e CT, como resultado a geração de imagens médicas híbridas.

Como contribuições, podemos citar: resultados de quais melhores métodos de reconstrução de imagens SPECT para utilização em treinamento de redes CNN; um novo método para reorientação do VE em exames de cintilografia do miocárdio e angiotomografia coronária; detecção automática do ápice e VE utilizando *deep learning*; registo 2D de imagens médicas multimodais; disponibilização de uma base GT (*ground truth*) anonimizada e um *dataset*, com 1000 imagens dos exames de cintilografia do miocárdio da técnica de imagem SPECT e angiotomografia coronárea da técnica CT, com anotações feitas por médicos e técnicos especialistas das áreas de medicina nuclear e radiologia.

Este trabalho está estruturado da seguinte maneira, no capítulo 2 são introduzidos os trabalhos relacionados a reorientação e registo de imagens médicas. O capítulo 3 relata a reconstrução de imagens da técnica SPECT e apresenta os *datasets* e experimentos realizados. No capítulo 4 é descrita a detecção do VE e ápice nas imagens da técnica SPECT e apresenta os experimentos realizados. No capítulo 5 é exposto a detecção do VE e ápice nas imagens da técnica CT e apresentado os experimentos realizados. No capítulo 6, são detalhados experimentos realizados para fusão/registo 2D entre as imagens funcionais e anatômicas. E por fim, no capítulo 7 são expostas as conclusões em função dos resultados alcançados e sugestões para trabalhos futuros.

CAPÍTULO 2

TRABALHOS RELACIONADOS

Neste capítulo são descritos os principais métodos e abordagens relacionados aos temas desenvolvidos neste trabalho.

Na literatura podemos encontrar alguns desafios e métodos para prevenir o problema na falha da detecção do ventrículo esquerdo e geração de imagens híbridas multimodais, como visto a seguir. Para cada técnica de imagem médica estudada, são encontrados problemas e desafios diferentes ligados a aquisição das projeções, resolução espacial, processamento e visão computacional, como demonstrado no decorrer deste capítulo.

Em Motwani et al. (2016), os autores citam quão desafiadora pode ser a segmentação do miocárdio em imagens SPECT, fatores como defeitos na aquisição das imagens e distribuição do traçador do miocárdio em diferentes tecidos do corpo do paciente, como por exemplo intestino, são apresentados. Outros fatores como sincronização dos batimentos cardíacos e ruídos nas imagens também são mencionados como impeditivos de uma correta segmentação do VE.

Para Tsadok *et al.* Tsadok et al. (2013), os trabalhos encontrados na literatura normalmente realizam a detecção e segmentação do VE em imagens do eixo curto, e não contemplam abordagens para segmentação do eixo longo, onde a mesma também é desafiadora. Para os autores, os limites endocárdiais e epicárdicos do coração tem uma significativa variância entre os pacientes. Ainda, fatores como orientação e posição durante aquisição, e condições cardíacas, vistas em pacientes pós-infarto, também são problemas durante o processo de segmentação.

Em Sonesson et al. (2009), os autores concluem e apresentam algumas dificuldades encontradas durante os processos para detecção e segmentação de segmentos cardíacos, como por exemplo a baixa resolução espacial em imagens SPECT. Outro aspecto evidenciado pelos autores, foi a complexidade em detectar e estimar os limites das paredes

miocárdicas, fato esse, que os mesmos afirmam ser um obstáculo a ser vencido com os atuais métodos de visão computacional.

Em Moro et al. (1994) são apresentadas duas tentativas para segmentação do ventrículo esquerdo em imagens da técnica SPECT: a primeira aplicando técnicas como *threshold* e a segunda pela utilização de um método descrito em Ezekiel et al. (1991), que consiste na identificação automática e geração de mapas polares cardíacos. Em outras abordagens, podemos verificar as tentativas de Yang et al. (2006), onde é utilizado o método de contornos ativos chamado Modelo de Contorno Carregado (CCM) como alternativa de segmentação dos segmentos do coração.

Outros trabalhos estudados abordam a utilidade das imagens médicas híbridas. Na grande maioria dos trabalhos sobre fusão multimodal encontrados na literatura, o estudo não é para fusão de imagens cardíacas, porém apresentam justificativas, técnicas e abordagens com resultados satisfatórios.

Declerck *et al.* Declerck et al. (1997), utilizam métodos não rígidos de registro entre as imagens das etapas stress e repouso dos exames de cintilografia do miocárdio da técnica SPECT.

Em Flotats et al. (2011), antes mesmo das questões computacionais, são encontrados problemas relacionados aos valores prognósticos que imagens híbridas podem fornecer. Os autores apresentam fatos, como a importante utilização de imagens híbridas em oncologia, no entanto, afirmam que para imagens cardíacas, os equipamentos híbridos não são rotineiramente utilizados, pela atual dificuldade em prever, *a priori*, quais pacientes se beneficiariam da técnica.

De maneira contraditória ao trabalho visto em Flotats et al. (2011), os autores entendem que a análise de imagens híbridas agregam na realização do diagnóstico médico. Nas imagens híbridas, as informações quantitativas encontradas nas imagens funcionais da técnica SPECT, com as anatômicas da CT, oferecem a oportunidade de uma avaliação abrangente e não invasiva dos problemas anatômicos e suas consequências fisiológicas do coração, mesmo que, sejam para descartar a presença de doenças cardíacas.

Nas ciências computacionais de visão computacional, aprendizagem de máquina e

processamento de imagens, pode-se encontrar abordagens que ressaltam os desafios e problemas na geração de imagens híbridas. No trabalho visto em Hill et al. (2001), os autores utilizam algoritmos como *Rigid body registration* e ICP (*Iterative closest point*) como alternativas para o registro/fusão de imagens multimodais. Apesar de serem abordagens desenvolvidas, inicialmente, sem o foco em imagens médicas, foram aplicadas posteriormente em conjunto com outros métodos com sucesso considerável Hill et al. (2001). Atualmente esses algoritmos são largamente utilizados para correspondência de superfície entre imagens e transformações geométricas com base em *landmarks* Stewart et al. (2003); Hutton et al. (2006).

Em Guetter et al. (2005) é apresentada uma nova abordagem de registro de imagens multimodais não rígidas. Também são salientadas as necessidades de utilização de registro multimodal utilizando algoritmos não rígidos e como o problema pode ser complexo.

Num estudo recente, visto em Li and Fan (2017), são utilizadas redes de aprendizagem profunda para a realização de registro não rígido. Neste estudo, os autores apresentam resultados promissores utilizando uma arquitetura totalmente convolucional, a FNC (Fully Convolutional Network), demonstrando melhor desempenho, tanto na precisão do registro, quanto no desempenho computacional. Para comparação dos resultados, entre o método proposto pelos autores e algoritmos estado da arte, foi utilizado *DICE* (Dice (1945)) como métrica para comparação da similaridade entre as imagens após registro. Como resultados: antes do registro $0,654 \pm 0,062$, após o registro ANTs (Avants et al. (2011)) $0,762 \pm 0,057$ e método proposto $0,798 \pm 0,033$. Concluindo que o método foi capaz de identificar melhor a correspondência entre as imagens.

Uma das vantagens da utilização de métodos não rígidos é que além da translação e rotação vista em algoritmos rígidos, as escalas entre as imagens também podem ser ajustadas, fornecendo melhores resultados quantitativos. Por esse motivo, as técnicas de registro não rígidas têm se tornando cada vez mais comuns e necessárias nas aplicações clínicas Crum et al. (2004).

Na continuação do capítulo, serão abordados os principais estudos que serviram como base para realização deste trabalho.

2.1 Faster R-CNN

A Faster R-CNN Ren et al. (2017), demonstrada na Figura 2.1 é composta por dois módulos principais. O primeiro módulo é uma rede convolucional profunda que propõe regiões de interesse (RPN) e o segundo módulo é o detector rápido que usa as regiões propostas pelo primeiro módulo. O primeiro módulo foi modelado com uma CNN Long et al. (2015), com o objetivo de diminuir o tempo de processamento computacional, compartilhando mapas de características e demais dados da rede de regiões, com a rede de detecção de objetos Girshick (2015).

Neste trabalho, para reconhecimento e detecção dos objetos, o modelo ResNet50 He et al. (2015a), uma rede neural residual profunda com 50 camadas, foi utilizado.

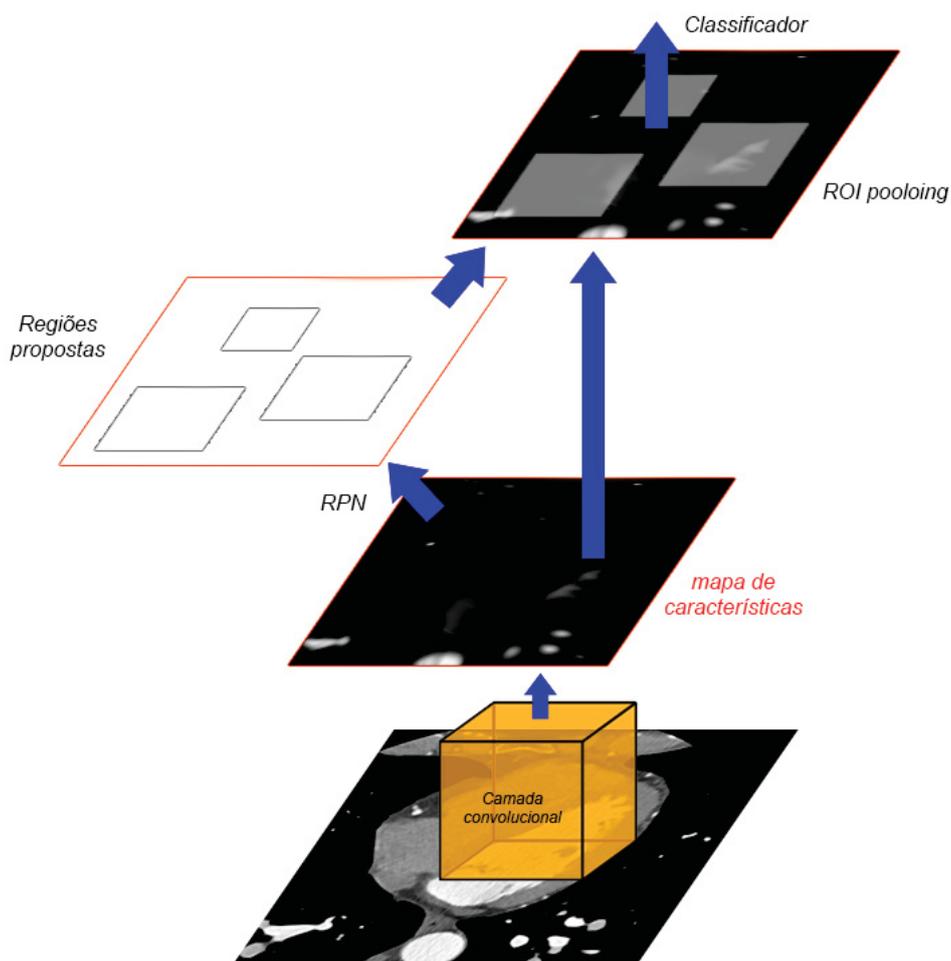


Figura 2.1: Exemplo da Faster R-CNN Ren et al. (2017) para detecção de objetos em conjunto com a rede RPN. É iniciado com a extração das características, RPN gerando as regiões propostas e rede CNN para classificação e regressão das regiões e dos objetos. Adaptado de Ren et al. (2015)

Uma RPN é uma rede convolucional que prevê, de maneira simultânea, limites e pontuação de objetos em cada posição. RPNs são treinadas de ponta a ponta para gerar regiões propostas de alta qualidade, que são utilizados posteriormente na detecção.

A Faster R-CNN Ren et al. (2017) permite que um sistema de detecção de objetos compartilhado com sistema de regiões propostas, funcione em taxas próximas ao tempo real. As características aprendidas, também melhoram a qualidade das regiões propostas e consequente, precisão da detecção de objetos. Para isso, é utilizando o conceito *bouding box regression*, com o objetivo de: reduzir os erros na detecção do objeto; treinar um modelo de regressão linear para predição das *bouding boxes* de detecção; ajustar a *bouding*

box de detecção para o tamanho ideal da imagem;

2.2 Segmentação e reorientação do VE em imagens SPECT

Um dos principais métodos para segmentação do miocárdio, é o encontrado em Germano et al. (1995). No trabalho, os autores apresentam um algoritmo com taxas de acerto de 98% e 96% para segmentação e reorientação respectivamente.

O algoritmo proposto por Guido *et al.* é composto das seguintes etapas: inicia-se o processo pela segmentação do VE, onde o valor de contagem do voxel ($C_{\text{máx}}$) no quadrante superior direito da imagem transaxial $64 \times 64 \times L$ é calculado como demonstrado em 2.2. Se o estudo tiver sido corretamente adquirido e reconstruído, o método assume que neste quadrante está o miocárdio.

Depois, todo volume transaxial é limiarizado para 50% do $C_{\text{máx}}$, e são feitos agrupamentos binários. Cada agrupamento (*cluster*), de voxels conectados é analisado, onde o primeiro *seed voxel* (primeiro voxel diferente de zero), tem sua localização marcada. Após análise de todos os *clusters*, são verificados os *cluster* muito pequenos (menor que 50 milímetros), e estes são eliminados. Se apenas um *cluster* permanecer e seu volume for maior ou igual a 250 milímetros, o *cluster* é assumido como a posição do VE.

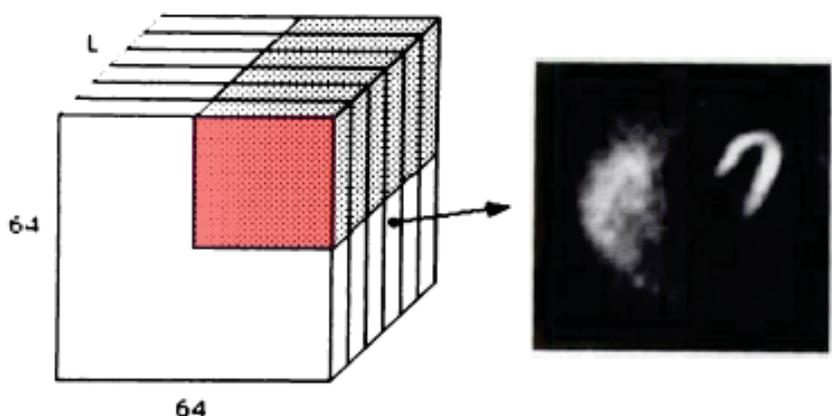


Figura 2.2: O limiar inicial do *threshold* utilizado no processo de agrupamento é baseado no valor de intensidade do voxel no quadrante superior direito do volume das imagens transaxiais, evidenciado em vermelho. Adaptado de Germano et al. (1995).

Algumas desvantagens do método proposto por Guido *et al.* são que o mesmo não leva em consideração alguns fatores vistos na prática diária de aquisição de imagens do

coração. Podemos levantar questões técnicas, metabólicas e anatômicas. Por exemplo, o posicionamento do paciente pode interferir na questão da posição do VE no quadrante definido pelo método, ainda, outro órgão como fígado ou alça estomacal podem estar próximos, assim, também ficando com *cluster* iguais ou maiores que o VE, retornando dois ou mais *clusters*.

CAPÍTULO 3

RECONSTRUÇÃO DE IMAGENS DA TÉCNICA SPECT

Os dados adquiridos pela técnica SPECT consistem em um conjunto de projeções Germano (2001) que representam a distribuição da radioatividade no corpo do paciente Mettler and Guiberteau (2011). Para esse trabalho, foram utilizadas imagens de cortes tomográficos transaxiais obtidas por gama-câmaras Khalil (2010), isótopo e fármaco $^{99m}\text{Tc-sestamibi}$ e aquisição com protocolo com matriz de aquisição 64x64, Waterstram-Rich and Gilmore (2016).

Após aquisição, são utilizados os dados da *tag pixelData* do arquivo *raw data* DICOM. Os dados *pixelData*, são projeções utilizadas para reconstrução por meio de técnicas matemáticas, descritos pelo matemático J. Radon, como visto em Kak and Slaney (2001), Deans (1981). Radon demonstrou que projeções de imagens bidimensionais, que tenham intensidade conhecida, podem ser reconstruídas gerando imagens tomográficas.

Um dos objetivos deste trabalho, foi avaliar diferentes técnicas de reconstrução de imagens SPECT. Essa avaliação foi feita para verificar quais melhores filtros e técnicas de interpolação para reconstrução de imagens SPECT, oferecem melhores características para o treino dos modelos.

3.1 Experimentos

O gráfico de acurácia do classificador das imagens reconstruídas, figura(3.1), apresenta os resultados obtidos com os modelos usando os mesmos parâmetros para treino. Foram realizados experimentos com imagens de SPECT reconstruídas com diferentes filtros e técnicas de interpolação. Em média, os treinos finalizam a primeira época de 100 iterações com aproximadamente 92% de acurácia para predição do VE e ápice na imagem.

Destaca-se o modelo treinado com a base reconstruída utilizando a técnica FBP (Filtered Back Projection) com o filtro *butterworth*, com os parâmetros de *cutoff* = 0.5 e *order*

= 10, comumente utilizados em imagens médicas cardíacas Lyra and Ploussi (2011). Os resultados com *butterworth*, apresentaram resultados de detecção e classificação do ápice e VE de até 98%. Por esse motivo, os experimentos de reorientação, para treino do modelo, melhor detalhados no capítulo 4, foram realizados com as imagens anotadas reconstruídas com esse filtro e parâmetros.

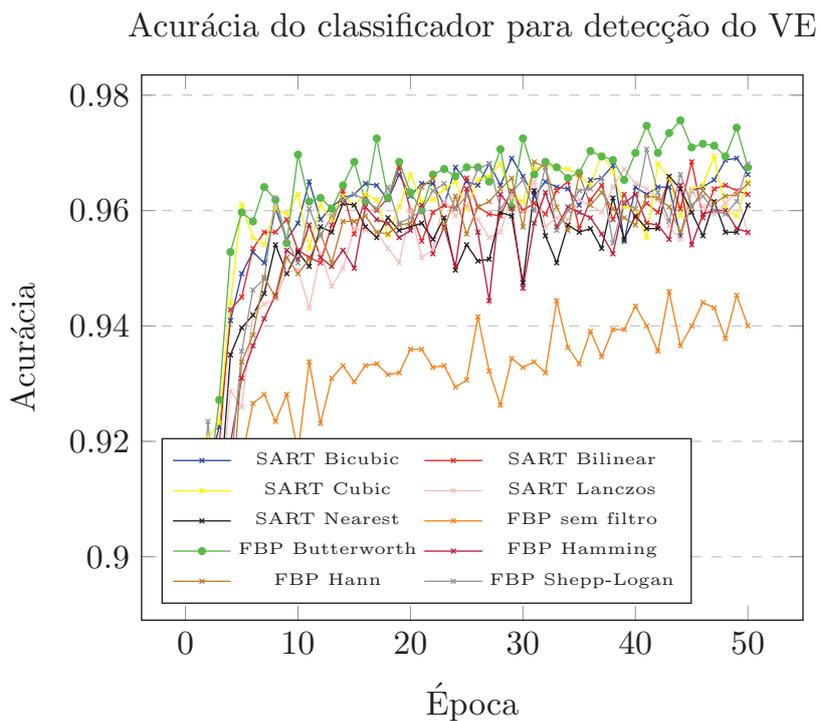


Figura 3.1: Gráfico com acurácia da rede para detecção do VE nas imagens SPECT reconstruídas com diferentes filtros e interpolações

CAPÍTULO 4

DETECÇÃO DO VE E ÁPICE NAS IMAGES SPECT

4.1 *Ground truth*

Nas técnicas de imagens SPECT, após reconstrução das projeções adquiridas pelo equipamento médico, são geradas imagens tomográficas transaxiais, comumente chamadas de imagens axiais. Uma base *ground truth* e um conjunto *dataset* foram gerados utilizando essas imagens e anotações das coordenadas do ápice e base do VE. As anotações foram geradas por cinco médicos e técnicos especialistas em medicina nuclear, totalizando 500 imagens e dados prognósticos obtidos de 500 exames anonimizados de cintilografia do miocárdio.

Para a geração da base, foi necessário o desenvolvimento de uma ferramenta para anotação das mesmas, demonstrada na figura 4.1.

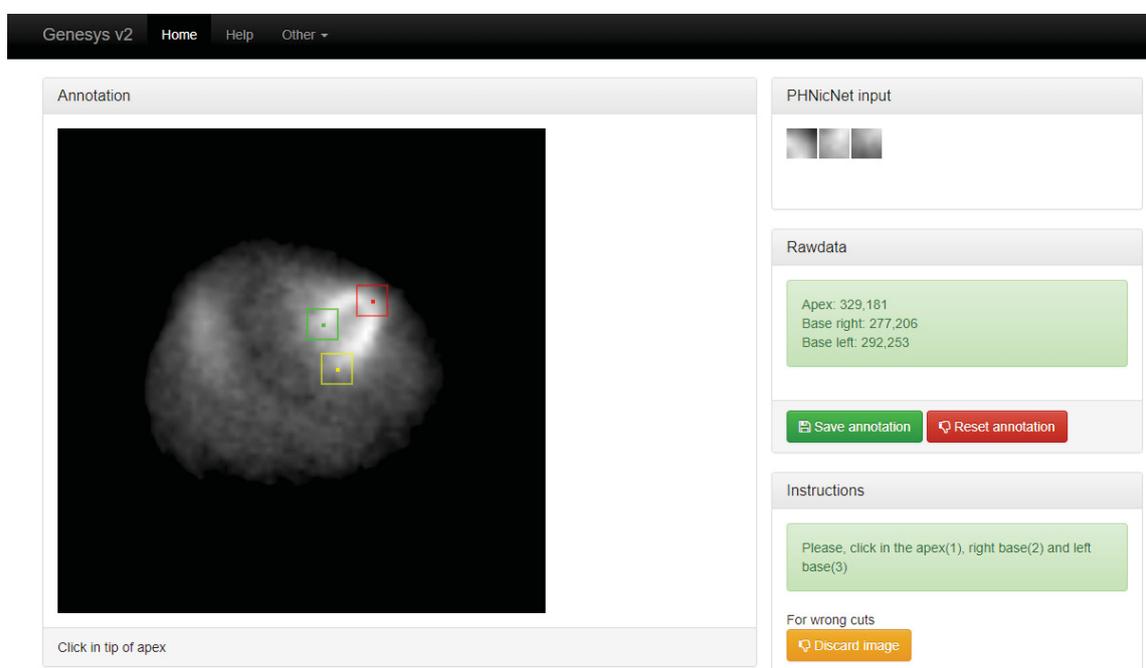


Figura 4.1: Interface da ferramenta desenvolvida para anotação das imagens SPECT. As *bounding boxes* vermelha, amarela e verde, representam o ápice, extremidade da base esquerda e direita respectivamente.

A ferramenta apresenta ao especialista imagens tomográficas transaxiais de exames de cintilografia do miocárdio de maneira aleatória, para a realização da anotação das *landmarks* no ápice e base do VE. As coordenadas das anotações foram armazenadas, vinculadas ao *ID* de identificação de cada exame. Inicialmente a base contemplava um valor maior de exames, porém as imagens com coordenadas muito divergentes, com distância das *landmarks* consideradas grandes (3 ou mais *pixels*), entre especialistas, foram descartadas.

4.2 Treinamento e otimização da *CNN*

Para a detecção do VE nas imagens SPECT, foram treinados dois modelos, um para detecção do VE e outro para detecção do ápice na imagem. Foi utilizado a Faster R-CNN Ren et al. (2017), utilizando uma rede ResNet50 He et al. (2015a) para detecção de objetos compartilhando recursos convolucionais com a rede RPN He et al. (2015b) para extração das regiões candidatas.

Para otimização, foi utilizado o otimizador ADAM Kingma and Ba (2014), com parâmetros de taxa de aprendizagem de 0.001, $\beta_1 = 0.9$ e $\beta_2 = 0.999$, 50 épocas de 100 iterações. No pseudo-código 1 temos a representação lógica de detecção.

Algorithm 1 Detecção do ápice e VE. A função *detectBB* carrega os modelos para detecção do ápice e VE de acordo com o técnica definida pelos parâmetro *tec*(1 é SPECT e 2 para CT) e *seg*(1 para ápice e 2 para VE). Tem como saída as coordenadas da *bouding-box* com maior probabilidade de ser o segmento escolhido.

```

função DETECTBB(slice, tec, seg)
  carregaModeloResNet50()
  se tec = 1 então
    se seg = 1 então
      carregaPesosModeloSPECTApice()
      resultado ← detectSPECTApice(slice)
    senão
      carregaPesosModeloSPECTVE()
      resultado ← detectSPECTVE(slice)
    fim se
  senão
    se seg = 1 então
      carregaPesosModeloCTApice()
      resultado ← detectCTApice(slice)
    senão
      carregaPesosModeloSPECTVE()
      resultado ← detectCTVE(slice)
    fim se
  fim se
  devolve resultado
fim função

```

4.3 Reorientação do VE

A primeira etapa da abordagem proposta é a utilização das imagens tomográficas axiais para geração de um volume tridimensional por meio do "empilhamento", sobreposição das mesmas.

São utilizados dados como FOV e dimensões da matriz nessa etapa, ilustrados na figura 4.2, e podem ser obtidos no arquivo DICOM (*raw data*) de cada exame/estudo utilizado nos experimentos.

Para geração dos volumes tridimensionais da técnica de imagem SPECT, foram levados em consideração boas práticas, padrões e protocolos definidos em Henzlova et al. (2016).

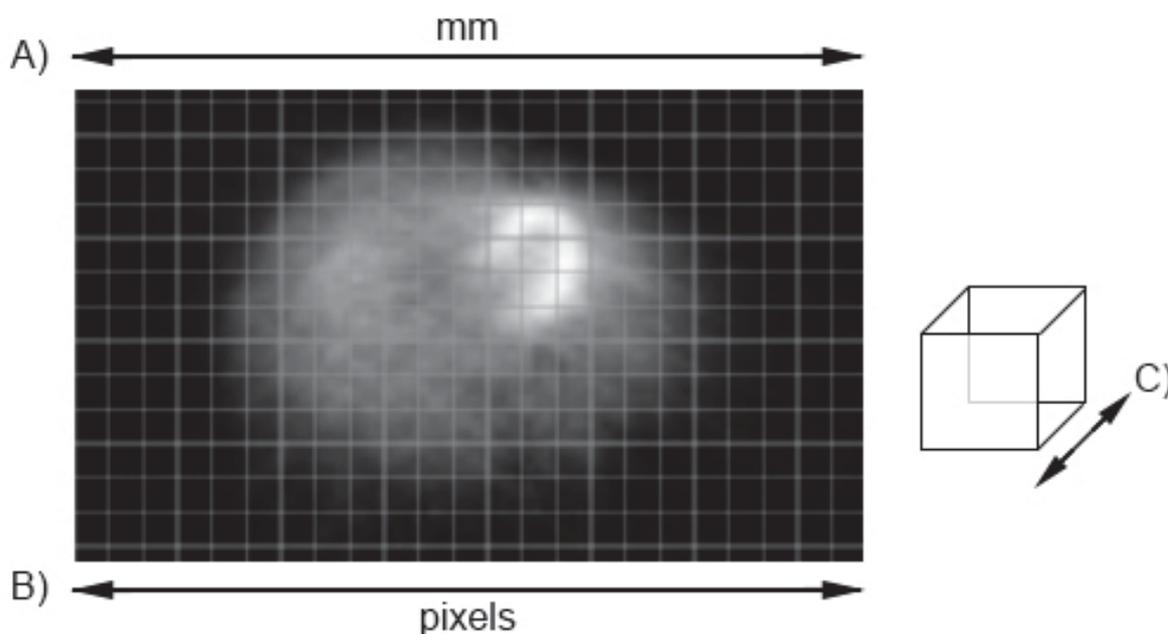


Figura 4.2: Parâmetros para geração/construção de volumes tridimensionais. A) Campo de visão em mm obtidos no arquivo DICOM; B) Largura da matriz obtidos no arquivo DICOM ou dimensões da imagem; C) Representação do *voxel*, espessura do *slice*, ou (*thickness*).

O pseudo-algoritmo 2, demonstra a lógica de geração do volume tridimensional. O algoritmo inicia o processo lendo os dados dos eixos x, y de cada *slice* tomográfico, e gera uma nova dimensão z fazendo a sobreposição das imagens.

Algorithm 2 Geração do volume tridimensional. A função *volume3D*, representa a geração do volume tridimensional utilizando os *slices* tomográficos 2D.

```

função VOLUME3D(slices2D)
   $i = 0$ 
  para eixo  $x \leftarrow$ 
    para até  $\leftarrow$  até e façaixo  $y$  faça
      enquanto slices2D faça
         $volume3D[x, y, i] \leftarrow slices2D[x, y]$ 
         $i++$ 
      fim enquanto
    fim para
  fim para
  devolve volume3D
fim função

```

É importante mencionar que não existe o problema de superposição desalinhada no

eixo axial, já que esse problema é tratado por softwares embarcados nos equipamentos médicos, durante a etapa de aquisição. Essas informações e parâmetros como posição do paciente em relação ao detector, tamanho da matriz de detecção, tempo de aquisição, FOV (Field of view), em português, campo de visão, ficam armazenadas nos arquivos DICOM utilizados para reconstrução das imagens tomográficas utilizadas neste trabalho.

Os modelos treinados com melhores pesos têm como saídas coordenadas no formato (x, y, l, a) , coordenadas x, y , largura e altura respectivamente, tanto para o ápice, quanto VE. Essas coordenadas são utilizadas para geração de pontos, por exemplo, A para ápice e B para extremidade do VE, e que serão utilizados para determinação do eixo longo.

Na abordagem proposta, o eixo longo do VE é determinado pela linha obtida pelos pontos $A(x_1, y_1)$ e $B(x_2, y_2)$, ilustrados na figura 4.3. Os pontos A e B foram obtidos pelo centro da *bouding-box* predita pelo modelo do ápice e extremidade inferior esquerda da *bouding-box* predita pelo modelo do VE, respectivamente, detectados na imagem axial central do volume SPECT tridimensional. Além do eixo longo, outros parâmetros são necessários para geração de *slices* perpendiculares, um deles é o ângulo de rotação, e é obtido pela equação 4.1.

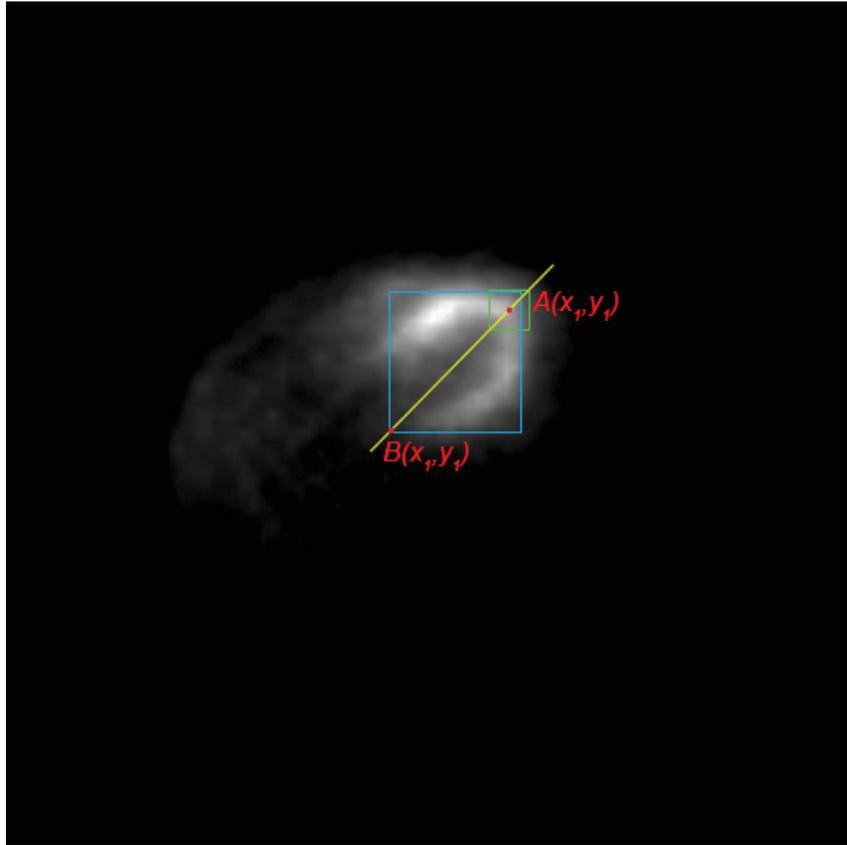


Figura 4.3: Ilustração das regiões de predição da rede. Em azul *bouding-box* do VE e verde *bouding-box* do ápice obtidas pelos modelos treinados. Os pontos vermelhos representam o centro da BB do ápice e extremidade inferior esquerda do VE, respectivamente. A linha amarela representa o eixo longo do VE obtido pelos pontos A e B.

Já que temos que $\beta=90^\circ-\alpha$, resta suficiente que precisamos do ângulo α para obtermos β . Temos que a medida do ângulo α pode ser obtida por meio dos pontos $A(x_1, y_1)$ e $B(x_2, y_2)$ arbitrados através do cálculo do coeficiente angular a da reta \overline{AB} , como podemos visualizar em 4.1:

$$\beta = 90^\circ - \alpha = 90^\circ - \arctan \frac{(y_2 - y_1)}{(x_2 - x_1)} \quad (4.1)$$

onde:

$$a = \frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1} \rightarrow a = \tan \alpha$$

e:

$$\alpha = \arctan a \Rightarrow \alpha = \arctan \frac{(y_2 - y_1)}{(x_2 - x_1)}$$

O pseudo-algoritmo 3, representa o raciocínio da equação 4.1:

Algorithm 3 O cálculo do ângulo para rotação, pode ser obtido pelas coordenadas resultantes da detecção do ápice e VE das redes treinadas.

```

função CALCULAANGULO(coordApice, coordVE)
   $x1, y1 \leftarrow extractPointApice(coordApice)$ 
   $x2, y2 \leftarrow extractPointVE(coordVE)$ 
   $\beta = 90^\circ - \arctan((y2 - y1)/(x2 - x1))$ 
  devolve  $\beta$ 
fim função

```

Definido o eixo longo, já temos bases para reorientação, já que entende-se por reorientação as imagens feitas de maneira perpendicular ao eixo longo do VE. O próximo passo é a geração efetiva dos *slices* com o VE reorientado utilizando padrões estabelecidos em Cerqueira et al. (2002).

Na prática clínica, essas imagens ou *slices*, são comumente conhecidos por, HLA (horizontal longe axis), VLA (vertical longe axis) e SHORT axis, em português, eixo longo horizontal, eixo longo vertical e eixo curto, respectivamente, e são demonstrados na figura 4.4.

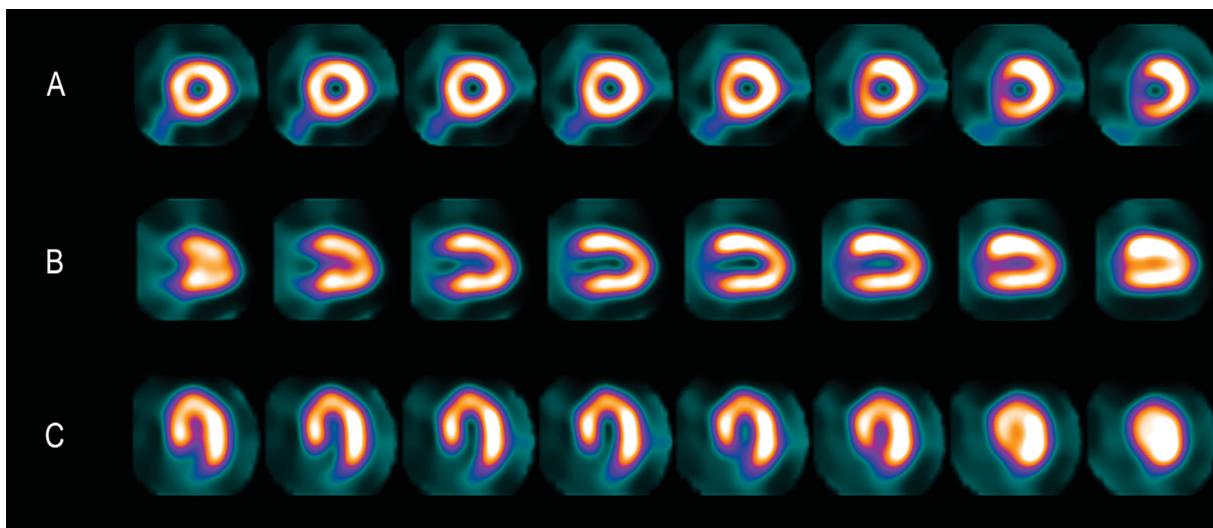


Figura 4.4: Exemplo de *slices* feitos perpendicularmente ao eixo longo do VE em diferentes espessuras/direções do volume. Na figura, podemos ver os eixos: A) curto, B) longo vertical e C) longo horizontal das imagens tomográficas axiais da técnica SPECT. As imagens ilustradas utilizam *colormap*, ou mapa de cores, para coloração dos *pixels* de acordo com a intensidade dos mesmos. Os níveis de cinza da imagem, são definidos pela quantidade de fármaco na aquisição.

4.4 Experimentos de detecção do VE

A perda geral da Faster R-CNN é uma combinação da perda do classificador e regressão, a figura 4.5, demonstra os resultados da perda total no decorrer das épocas.

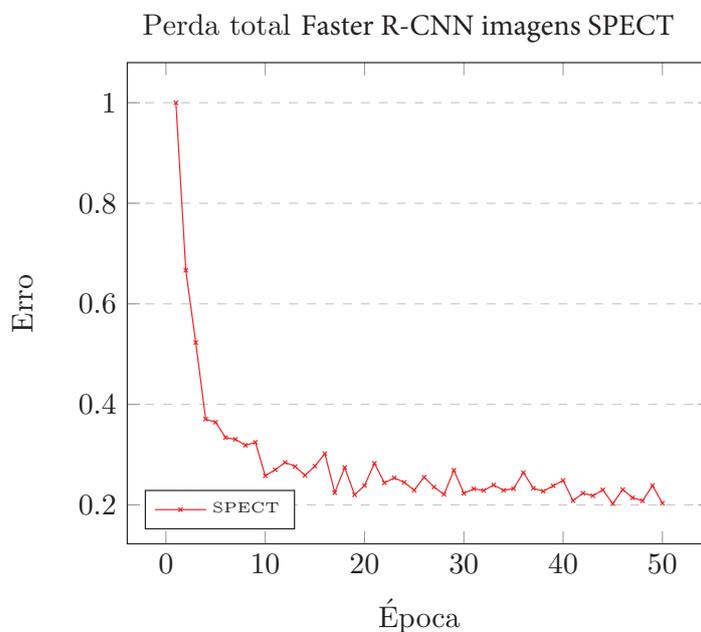


Figura 4.5: Gráfico com os erros totais (classificação e regressão) nas imagens SPECT

Podemos observar no gráfico 4.5 que a rede inicia com altos índices de erro, porém ao decorrer das iterações e épocas de treinamento, tendem a melhorar os pesos e os erros diminuem, partindo de 1.0 na primeira época, até 0.2 na última.

O gráfico demonstrado na figura 4.7 apresenta uma média dos resultados de treinamento da IoU (Intersection over Union), interseção da sobreposição das *bouding-boxes* da base *ground truth* e *bouding-boxes* preditas pela rede CNN e detecção. O coeficiente de interseção pode ser visto em trabalhos como Lin et al. (2014) e Girshick (2015), e é calculado pela equação 4.2, exemplificado na figura 4.6.



Figura 4.6: O IoU é simplesmente calculado dividindo a área de sobreposição entre *bounding boxes*, pela área de união.

$$\text{IoU}(A, B) = \frac{A \cap B}{A \cup B} \quad (4.2)$$

Um IoU com valores $\geq 0,5$ são consideradas 'boas' predições, ou predições corretas. Quanto maior o valor da métrica, com uma pontuação máxima de 1.0, maior é a sobreposição entre as *bounding-boxes* e regiões preditas, e melhor foi a predição da rede. Assim sendo, a métrica IoU é ótima para avaliar a precisão de detecção dos objetos.

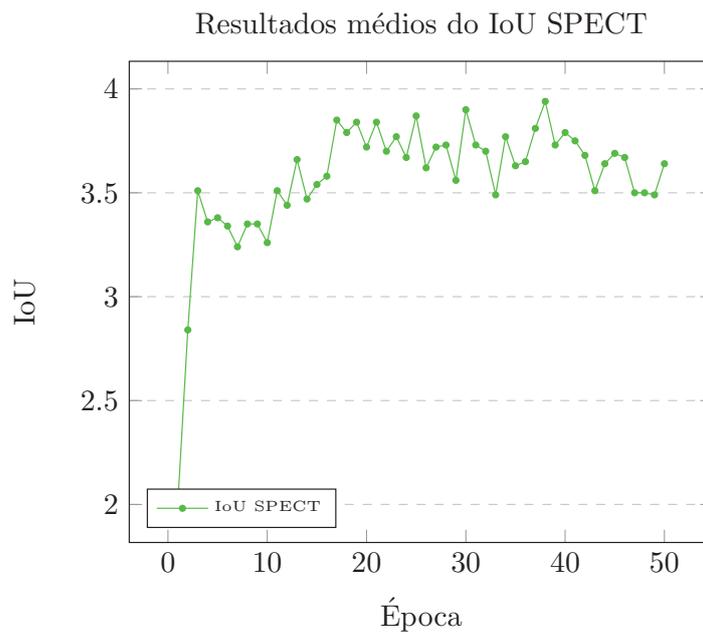


Figura 4.7: Gráfico com resultados da sobreposição entre as imagens da base *ground truth* e regiões preditas. O eixo IoU contempla a média da soma dos segmentos ápice e base no decorrer das épocas.

Durante a fase de treino, a Faster R-CNN gera e salva modelos de acordo com melhores classificadores e regressões encontrados, com médias do IoU, demonstradas no gráfico 4.7.

Já no gráfico da figura 4.8, podemos ver a distribuição dos resultados de sobreposição entre GT e predição utilizando o modelo com melhores pesos.

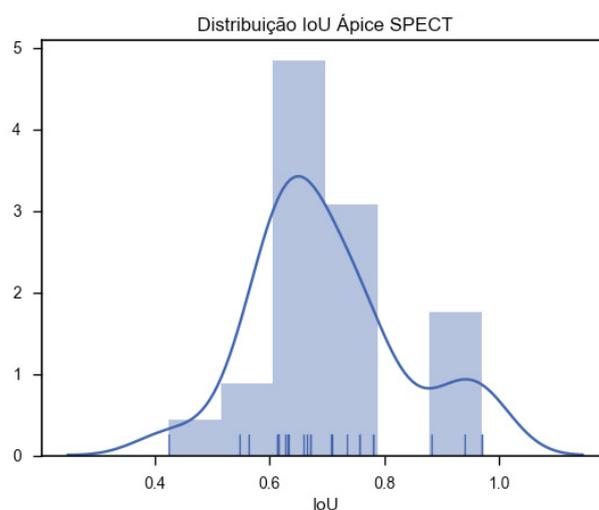


Figura 4.8: Gráfico com distribuição dos resultados entre GT e predição do ápice utilizando a métrica IoU. Apesar de alguns resultados ficarem próximos de 0.5 (resultado bom), ficam evidenciados resultados onde o modelo obteve uma ótima predição de até 0.97 de sobreposição.

A tabela 4.1 apresenta os resultados de detecção do VE obtidos utilizando os modelos treinados, comparados com o baseline. A tabela apresenta a porcentagem de acerto, se foi, ou não, encontrado o ápice e VE na imagem tomográfica.

Tabela 4.1: Resultados de detecção nas imagens SPECT. Taxa de acerto detecções corretas.

Casos clínicos			
	Normal	Anormal	Média
Baseline	-	-	98.5%
Método proposto	100%	94%	97%

A tabela 4.2, apresenta os resultados dos experimentos para detecção nas imagens SPECT. Pode-se destacar a acúrcia para detecção do ápice para os casos normais, com uma acertabilidade de 100%. E um bom desempenho para as imagens de casos clínicos anormais com 94% de detecção.

Tabela 4.2: Resultados de detecção do Ápice nas imagens CT. Taxa de acerto para detecções corretas.

Detecção do ápice nas imagens SPECT / Caso clínico		
	Anormais	Normais
Falsos negativos	6%	0%
Falsos positivos	12%	6%
Corretos	94%	100%

Tanto para casos normais, quanto para os anormais, em um baixo número de imagens como visto na tabela 4.2, o modelo retornou falso positivos, além do ápice, nas imagens, como ilustrado na figura 4.9.

O interessante de analisar na imagem 4.9 é a semelhança das estruturas entre o falso positivo e a *bouding-box* do ápice correto.

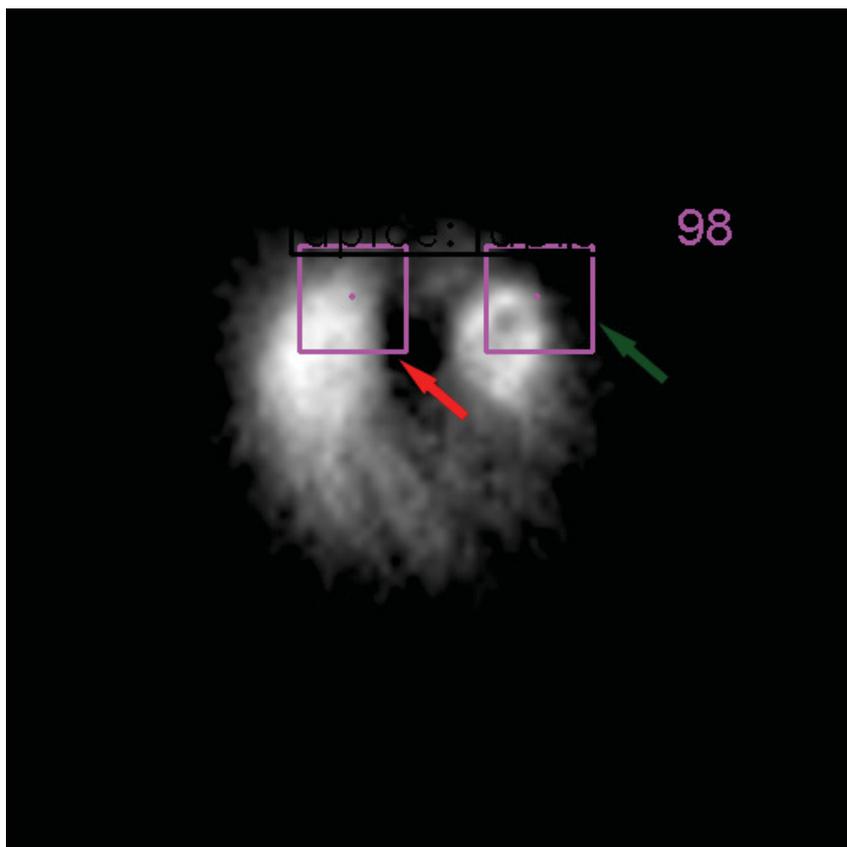


Figura 4.9: Exemplo de retorno de falsos positivos em imagens SPECT. A seta vermelha aponta a detecção errada, enquanto que a verde a detecção correta, já que seu classificador teve um maior valor.

Já na tabela 4.3, podemos verificar os resultados de reorientação utilizando o ângulo calculado pela abordagem proposta, comparados ao *baseline*. Nenhum estudo científico, até o momento, analisou precisamente qual a margem de erro aceitável no ângulo de rotação dos eixos longo vertical horizontal e longo vertical vertical. Isso se dá pelo fato que as imagens são consideradas corretamente reorientadas, caso os *slices* tenham sido feitos de maneira perpendicular ao eixo longo do VE Cerqueira et al. (2002). Com isso, na abordagem proposta, levou-se em consideração um grau de liberdade de até 10° de erro, baseado em opiniões empíricas de especialistas. Para a abordagem proposta, definiu-se que a extremidade do ápice deveria estar em 90° para as imagens do eixo longo horizontal, e em 0° para o eixo longo vertical, como ilustrado na figura 4.10, onde foi respeitada a margem de 90° definida.

Por não poder simular a base do estado da arte, foi considerado que a taxa de acerto da reorientação de Guido *et al.* foi de 96,3% descrita no trabalho Germano et al. (1995).

Tabela 4.3: Resultados médio da reorientação HLA e VLA SPECT (10°)

Casos clínicos			
	Normal	Anormal	Média
Baseline	-	-	96.3%
Abordagem proposta	100%	95%	98.3%
GT	100%	100%	100%

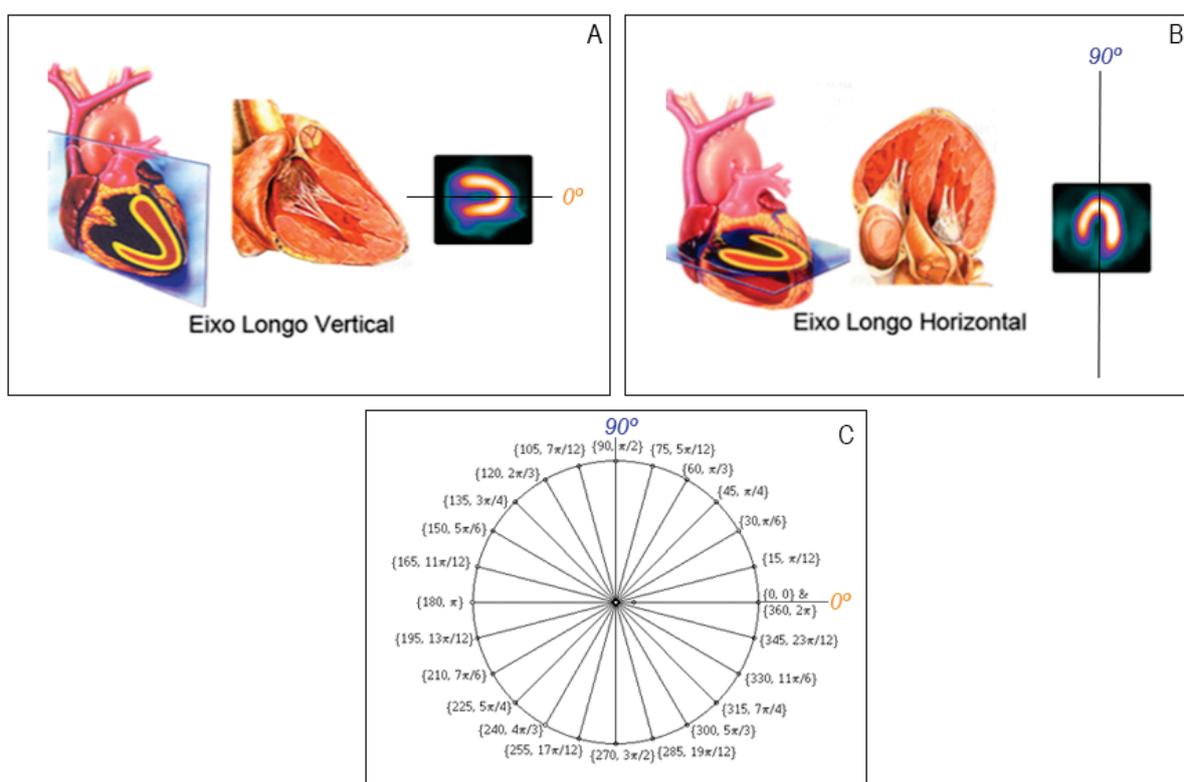


Figura 4.10: Ilustração dos eixos. A) *Slices* do eixo longo vertical ilustrando a extremidade do ápice em 0°. B) *Slices* do eixo longo horizontal ilustrando a extremidade do ápice em 90°. C) Círculo com graus e radianos. Para melhor visualização, foram evidenciados os ângulo 90° em azul e 0° em laranja.

A figura 4.11 apresenta a correlação entre os resultados das *landmarks* preditas, quando comparados com a similaridade entre as regiões da base GT e preditas pelo modelo treinado. Para essa análise, foram utilizadas as *landmarks* obtidas utilizando o centro da *bouding-box* de predição, comparados com as *landmarks* anotadas manualmente pelos médicos especialistas, calculando o erro RMS da distância em *pixels* para cada imagem.

Podemos analisar que quanto maior a sobreposição entre as *bouding-boxes*, menor a

distância entre as *landmarks* da base *ground-truth*, calculadas e previstas pela rede. Em média, RMS com valores de 11, ou seja, com até 11 *pixels* de distância entre GT e predição, obtiveram uma $\text{IoU} \geq 0.65$.

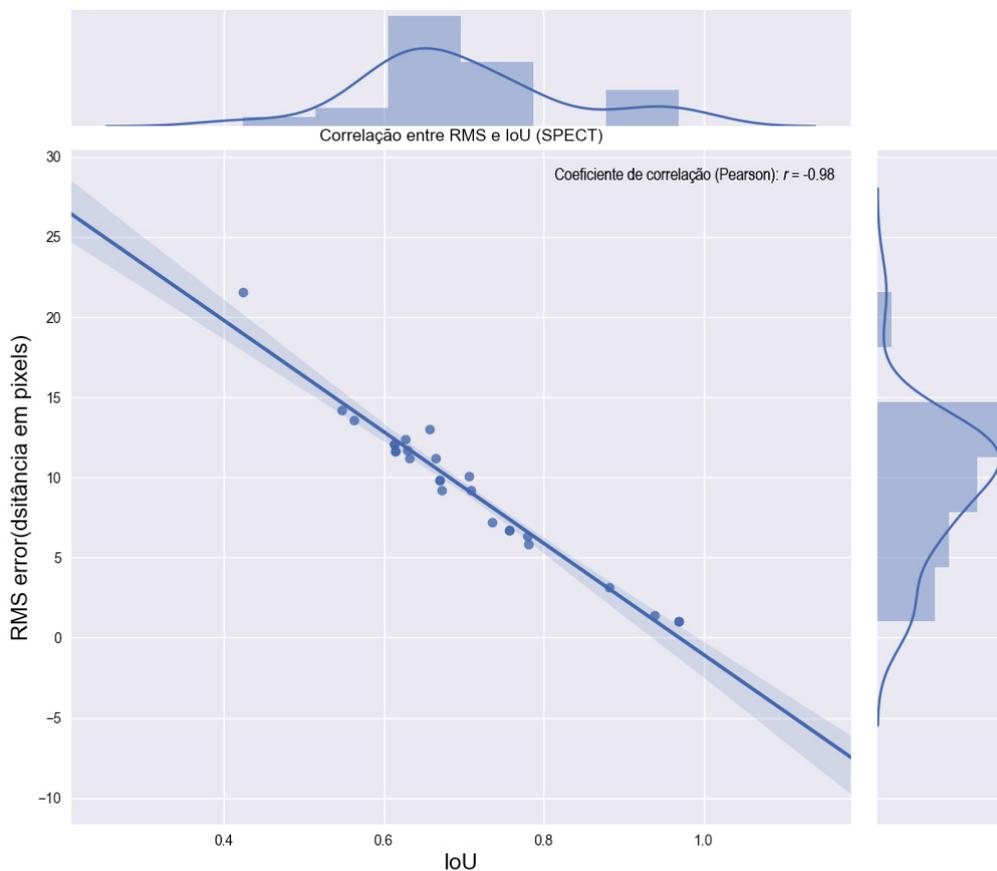


Figura 4.11: Gráfico de correlação entre RMS (distância *pixels*) e IoU das imagens SPECT.

Após reorientadas as imagens, foram obtidos os ângulos da extremidade do ápice em relação ao centro da imagem, conforme demonstrado na figura 4.12. De maneira similar ao que o estado da arte utiliza nas imagens transaxiais, é realizada a detecção de uma elipse na imagem e definido o ângulo do maior eixo da elipse, como o ângulo do ápice no *slice*. Para detecção da elipse foi utilizada a técnica descrita em Fitzgibbon et al. (1996).

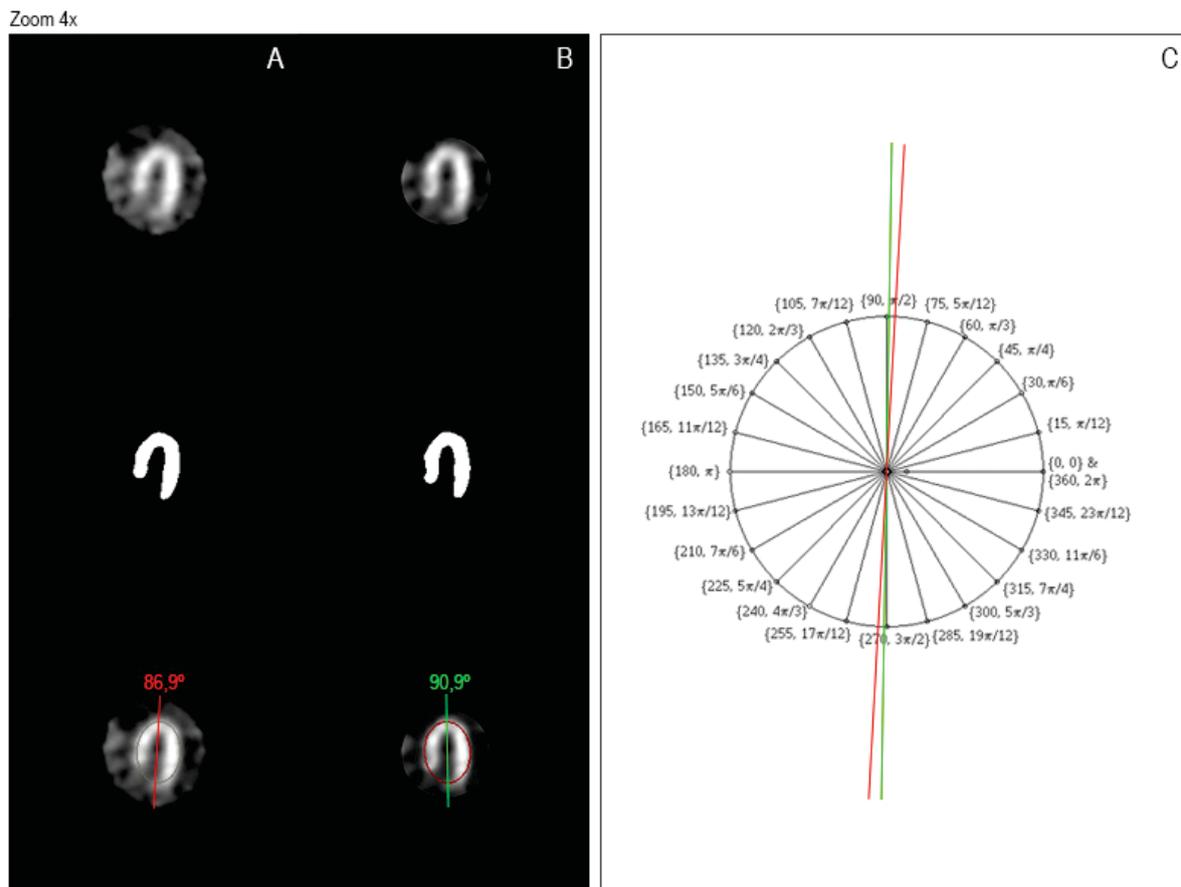


Figura 4.12: Exemplo da análise de resultados do eixo longo horizontal. A) Respectivamente, corte perpendicular (*Slice*) do eixo longo horizontal após reorientação, segmentação do VE, e ângulo do ápice calculado da base GT; B) Respectivamente, corte perpendicular (*Slice*) do eixo longo horizontal após reorientação, segmentação do VE, e ângulo do ápice calculado da base de predição; C) Representação dos ângulos encontrados em graus, levando como base o maior eixo da elipse detectada na imagem segmentada do VE. Em vermelho ângulo da base GT, e em verde da reorientação realizada com parâmetros do modelo treinado.

Como resultados principais da reorientação, podemos destacar a precisão da reorientação da abordagem proposta, tanto para eixo longo vertical, quando eixo longo horizontal. Tanto a abordagem proposta, quanto especialistas, obtiveram resultados de 0° para o eixo vertical e 90° para o eixo horizontal. Como piores resultados da abordagem proposta, podemos mencionar um erro de 15° na reorientação do eixo longo horizontal, em uma imagem de um exame anormal, com defeito de perfusão.

Para a verificação de concordância entre a base anotada por especialistas e a abordagem proposta, foi gerada um análise proposta por Altman *et al.* Altman and Bland (1983), do resultados da reorientação do eixo longo horizontal (4.13). Na análise, foi

possível observar uma diferença média de apenas $-0,450^\circ$ (IC 95% $-2,065$ até $1,165$), muito abaixo do clinicamente aceitável de 10° . Onde os limites de concordância ficaram entre $-10,5^\circ$ e $9,6^\circ$.

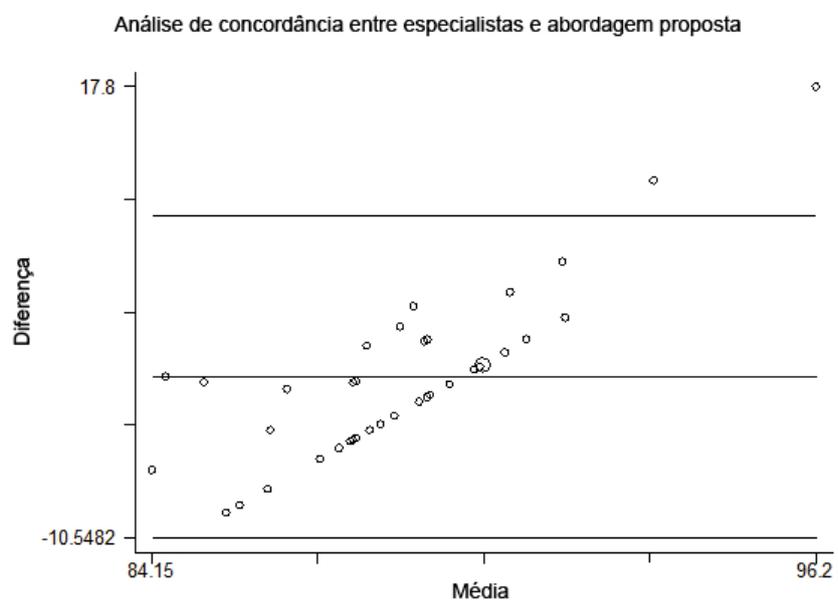


Figura 4.13: Gráfico de Bland e Altman. Concordância entre base anotada por especialistas e reorientação proposta, considerando os resultados dos ângulos obtidos após reorientação.

CAPÍTULO 5

DETECÇÃO DO VE E ÁPICE NAS IMAGES CT

5.1 *Ground truth*

De maneira semelhante a base SPECT, uma base *ground truth* e um conjunto *dataset*, com imagens e anotações das coordenadas do ápice e base do VE, foram geradas por um médico radiologista, totalizando 500 imagens e dados prognósticos obtidos de 500 exames anonimizados de angiotomografia coronária. Para a geração da base, foi utilizada a mesma ferramenta para anotação do SPECT, demonstrada na figura 4.1, com o único diferencial que em vez de imagens de 64x64, a ferramenta estava configurada para imagens de 512x512.

A ferramenta apresenta ao especialista imagens tomográficas transaxiais dos exames de angiotomografia coronária de maneira aleatória, para a realização da anotação das *landmarks* no ápice e base do VE. As coordenadas das anotações foram armazenadas, vinculadas ao *ID* de identificação de cada exame.

5.2 **Treinamento e otimização das CNNs**

De maneira semelhante a abordagem utilizada no SPECT, para as imagens CT, para a detecção, foram treinados dois modelos, um para detecção do VE e outro para detecção do ápice na imagem. Foi utilizado o método Faster R-CNN, utilizando uma rede ResNet50 He et al. (2015a) para detecção de objetos compartilhando recursos convolucionais com a rede RPN He et al. (2015b) para extração das regiões candidatas.

Para otimização, foi utilizado o otimizador ADAM Kingma and Ba (2014), com parâmetros de taxa de aprendizagem de 0.001, $\beta_1 = 0.9$ e $\beta_2 = 0.999$, 50 épocas de 100 iterações. No algoritmo 1 temos a representação do pseudo-código de detecção.

5.3 Reorientação do VE nas imagens CT

O mesmo algoritmo descrito na reorientação do VE nas imagens SPECT é utilizado para a reorientação nas imagens do CT. Com o diferencial que os parâmetros de entrada para as coordenadas são obtidos pela rede treinada com as imagens CT. Também, podemos analisar um exemplo mais ilustrativo encontrado na figura 5.1.

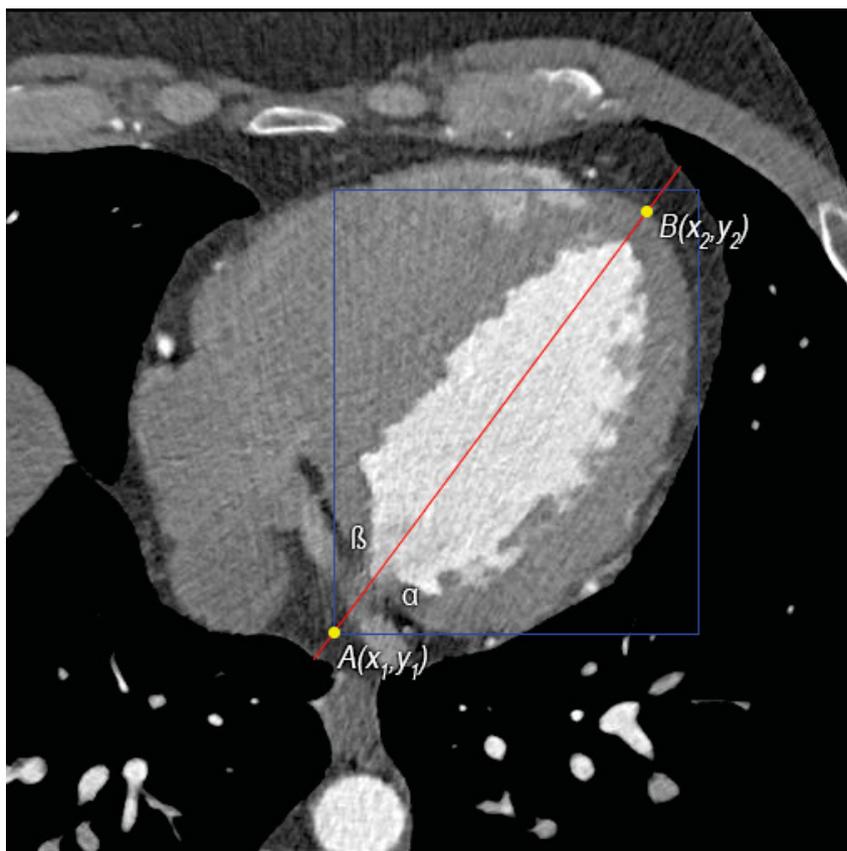


Figura 5.1: Em azul *bounding-box* em torno do VE. Em vermelho, reta paralela traçada com base nos pontos amarelos gerados a partir das coordenadas $A(x_1, y_1)$ e coordenadas do ápice $B(x_2, y_2)$ representando o eixo longo.

No exemplo 5.2, está ilustrado a entrada e saída da rede utilizando os pesos do modelo treinado com as imagens e *dataset* da base CT.

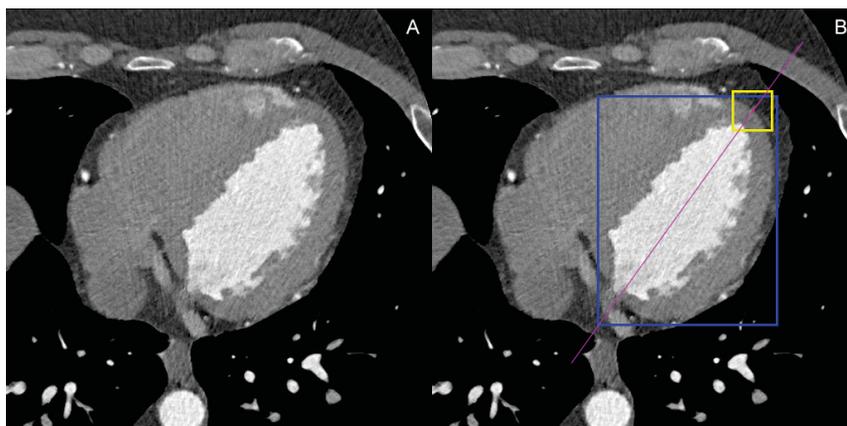


Figura 5.2: A) Imagem CT tomográfica do VE no corte axial. B) Predição da *bounding box* em azul, em torno do VE, *bounding box* amarela em torno do ápice e em vermelho o centro da *bounding box* do ápice. Em rosa eixo longo do VE.

A figura 5.3, apresenta os eixos curto, longo vertical e longo horizontal obtidos após reorientado do volume das imagens transaxiais da técnica CT.

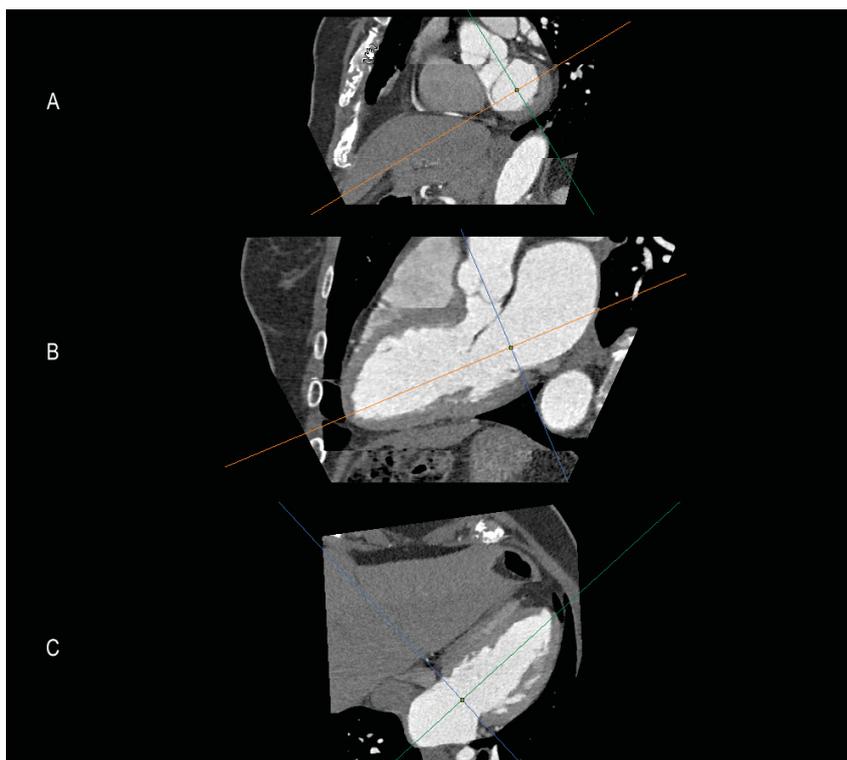


Figura 5.3: Exemplo de *slices* dos eixos A) curto, B) longo vertical e C) longo horizontal obtidos após reorientação do volume CT.

5.4 Experimentos

Até a finalização do levantamento do estado da arte, não foram encontrados métodos para detecção do VE em imagens axiais e reorientação do VE em imagens CT, o que não permitiu a realização de um *benchmark*.

De qualquer forma, são apresentados na tabela 5.1, os resultados obtidos pelo método proposto.

Tabela 5.1: Resultados de detecção do VE nas imagens CT. Taxa de acerto para detecções corretas.

Detecção do VE e ápice nas imagens CT	
Falsos negativos	0%
Falsos positivos	30%
Corretos	100%

A tabela 5.1, apresenta os resultados para detecção do ápice e VE nas imagens CT. Apesar de apresentar uma taxa de falso positivos de 30%, o desempenho da rede foi satisfatório, porque em todos os casos de falsos positivos, onde dois ou mais classificadores foram encontrados, também foi detectado o VE de maneira correta, conforme demonstrado na figura 5.4, nesses casos, foram utilizadas as *bouding boxes* com maiores valores do classificador.

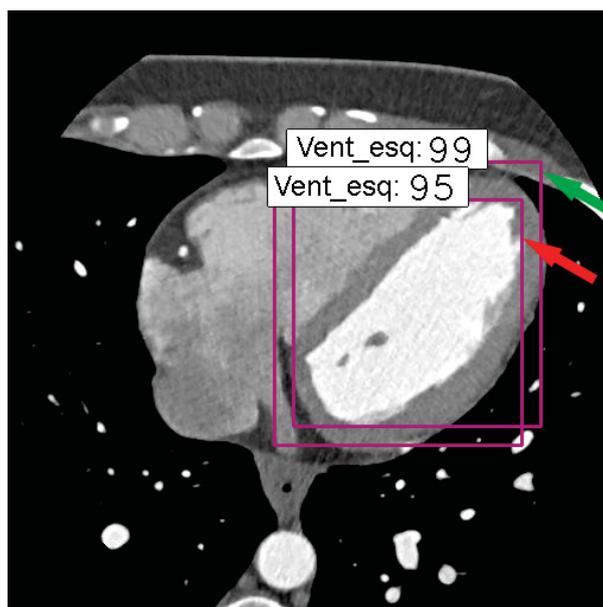


Figura 5.4: A seta vermelha indica a predição com menor classificador, enquanto a seta verde indica a detecção correta.

Os resultados de falsos positivos nas imagens de CT, surpreenderam pelo fato de imagens anatômicas aparentemente oferecerem maiores características para as rede de predição das regiões propostas e detecção/classificação. Uma das explicações possíveis é o fato da anatomia humana apresentar semelhanças consideráveis, porém, novos estudos devem ser feitos utilizando maiores bases de treino e diferentes parâmetros nas redes, para confirmar esse fato.

No gráfico 5.5, de maneira semelhante aos resultados de erros no SPECT, a rede inicia com erros de 1.0, porém ao decorrer das iterações e épocas de treinamento, tendem a melhorar os pesos e os erros diminuíam, finalizando em 0.3.

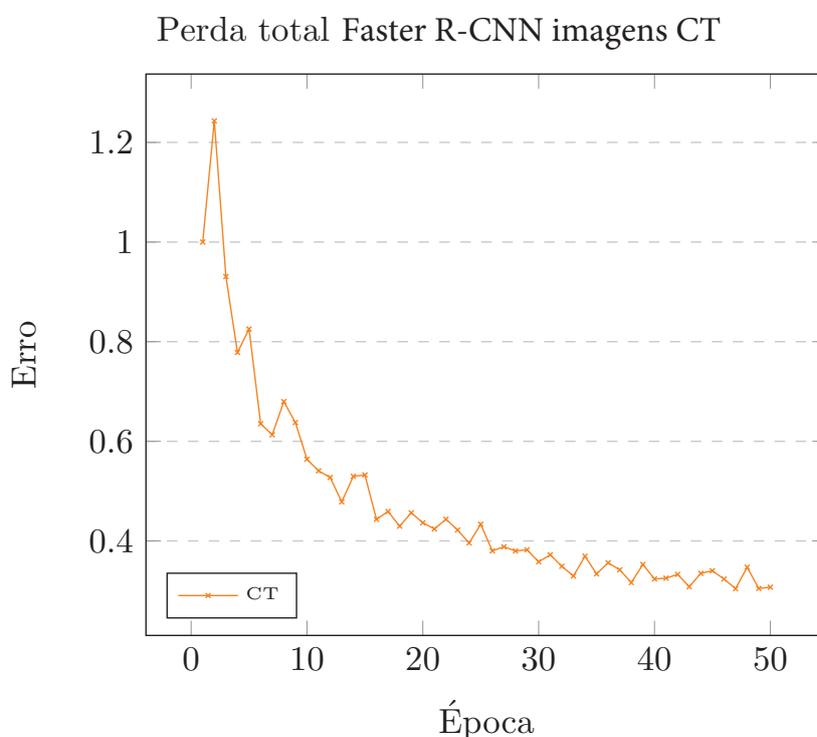


Figura 5.5: Gráfico com os erros totais (classificação e regressão) nas imagens CT

O gráfico demonstrado na figura 5.6, apresenta os resultados médios de IoU, interseção da sobreposição das regiões, da base *ground truth* e preditas pela Faster R-CNN, nas imagens CT para cada época de treinamento.

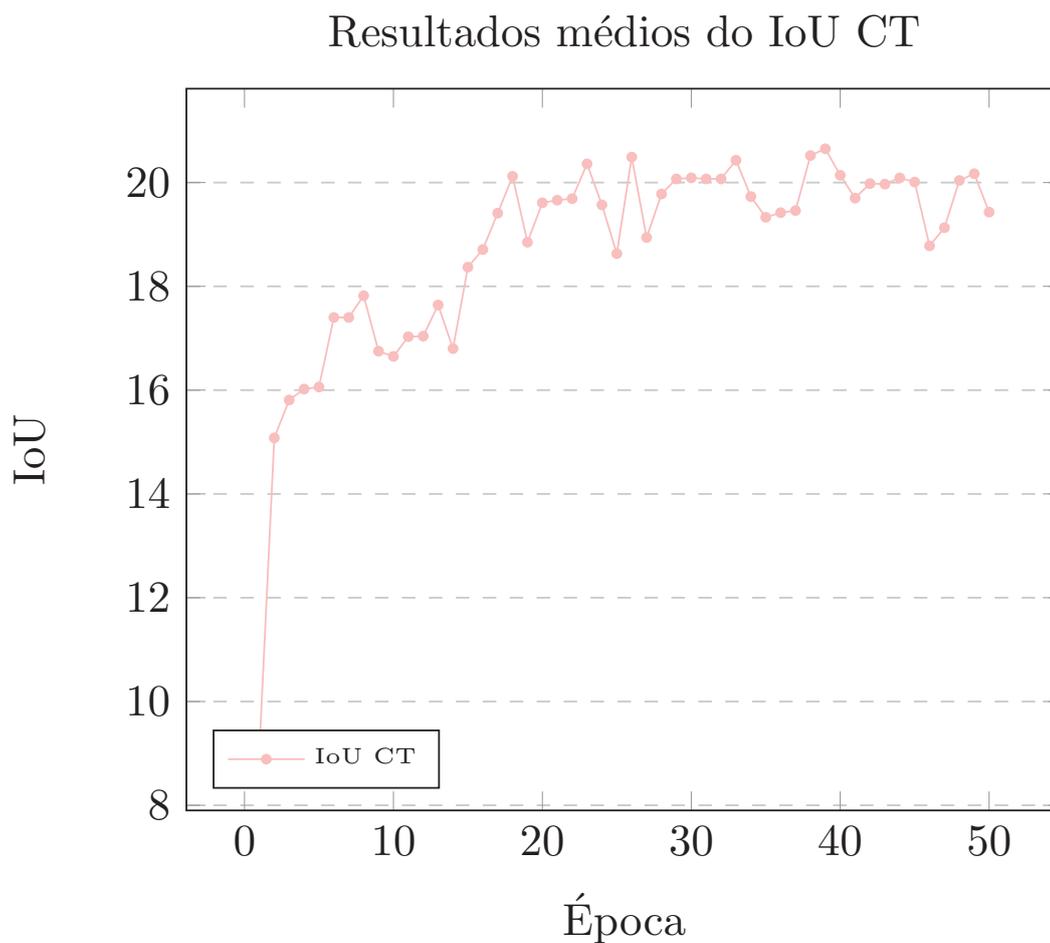


Figura 5.6: Gráfico com melhores resultados da interseção da sobreposição entre base *ground truth* e regiões preditas das imagens CT. Escala IoU médio para 50 épocas 100 iterações. O eixo IoU contempla a média da soma do ápice e base no decorrer das épocas.

No gráfico da figura 5.7, pode-se analisar a distribuição dos resultados de IoU entre GT e predição das imagens do CT do melhor modelo treinado. Podemos verificar que em aproximadamente 50% das imagens a predição obteve resultados maiores ou iguais a 0.6 de IoU, resultados considerados bons ou ainda, corretos. O modelo se mostrou eficaz com IoU de até 0.91 para o ápice.

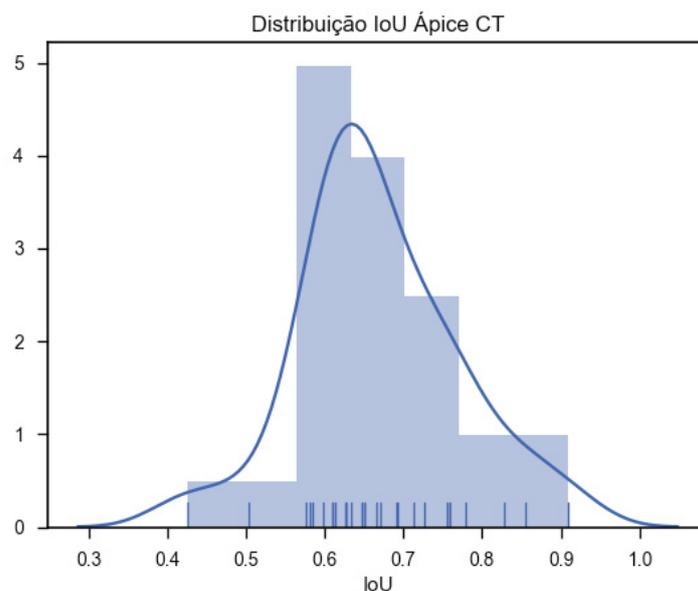


Figura 5.7: Gráfico com distribuição dos resultados entre GT e predição do ápice utilizando a métrica IoU. Apesar de alguns resultados ficarem abaixo de 0.5 (resultado bom/considerado correto), ficam evidenciados resultados onde o modelo obteve uma ótima predição de até 0.91 de sobreposição entre GT e a predição.

Já no gráfico da figura 5.8, podemos visualizar a correlação entre o IoU e RMS da distância dos *pixels*, entre a base GT e predição do modelo treinado com melhores pesos.

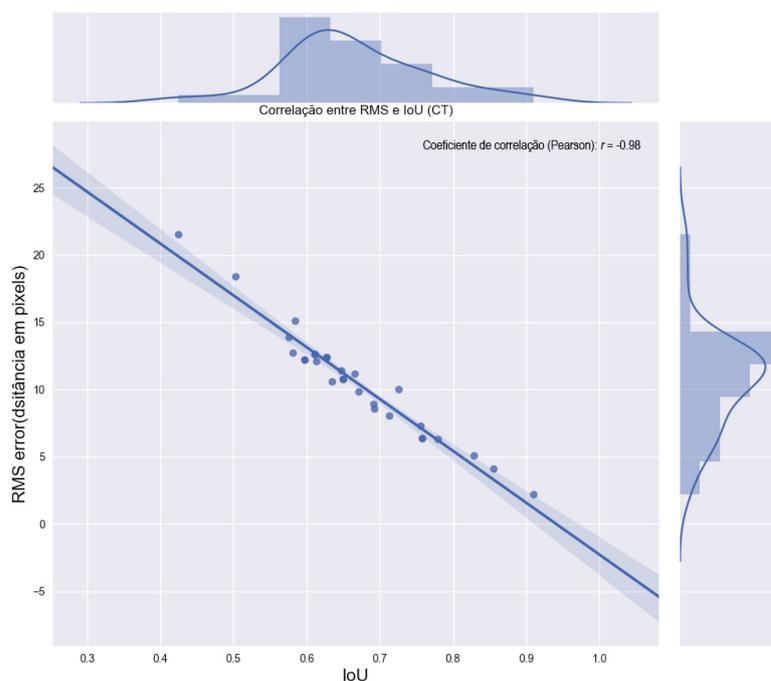


Figura 5.8: Gráfico de correlação

Pode-se analisar no gráfico que quanto menor a distância entre as *landmarks* da base GT e base predita, melhores os resultados de similaridade entre as regiões da base GT e regiões preditas.

CAPÍTULO 6

FUSÃO 2D DAS IMAGENS FUNCIONAIS E ANATÔMICAS

A figura 6.1 apresenta o diagrama da abordagem proposta para geração das imagens híbridas. No *pipeline*, leva-se em consideração que as imagens CT e SPECT são do mesmo paciente e que foram devidamente reorientadas utilizando o ângulo calculado para cada uma das técnicas.

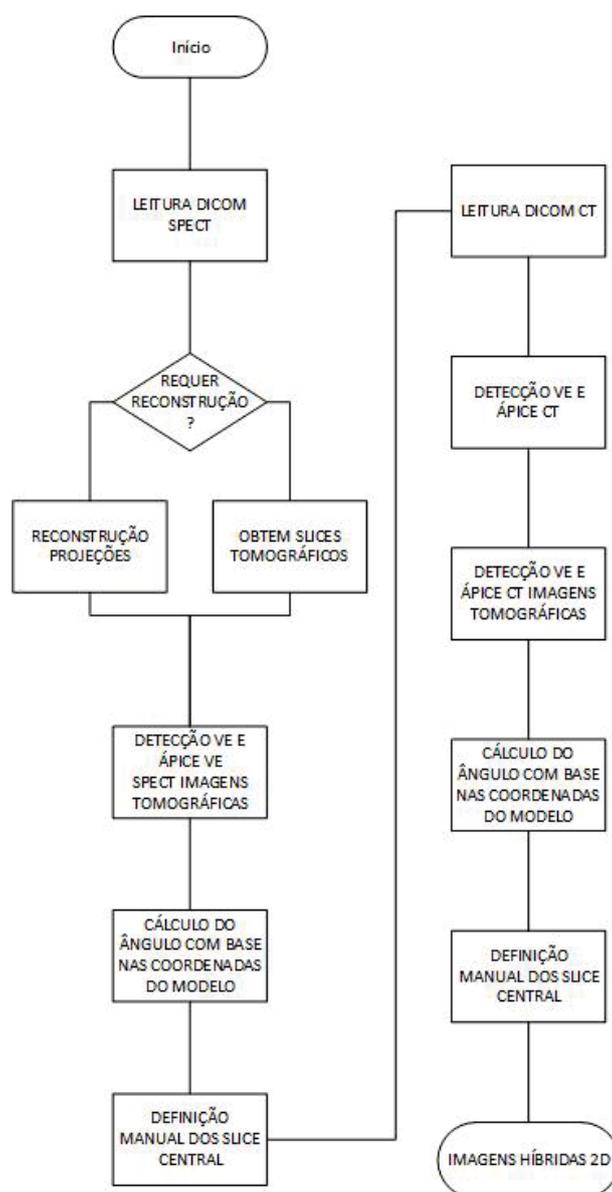


Figura 6.1: Pipeline para fusão 2D das imagens das modalidades SPECT e CT

6.1 Experimentos realizados

Após definido os *slices* centrais do volume de cada modalidade, é realizado o registro entre as imagens SPECT e CT alterando a translação e rotação, utilizando algoritmos *base point registration* Ma et al. (2016); Fitzpatrick et al. (1998); Mani et al. (2013); Gold et al. (1995) e coordenadas obtidos com os modelos treinados, demonstrados na figura 5.1. Na figura 6.2 e 6.3 não foi utilizado ICP Besl et al. (1992), pois nos experimentos realizados, a ordem das *landmarks* tem que ser levadas em consideração.

A figura 6.2, apresenta um registro de imagens das modalidades SPECT e CT do mesmo paciente, porém de maneira rígida, ou seja, apenas ajustando translação e rotação, preservando a diferença de escala/dimensões entre as imagens.

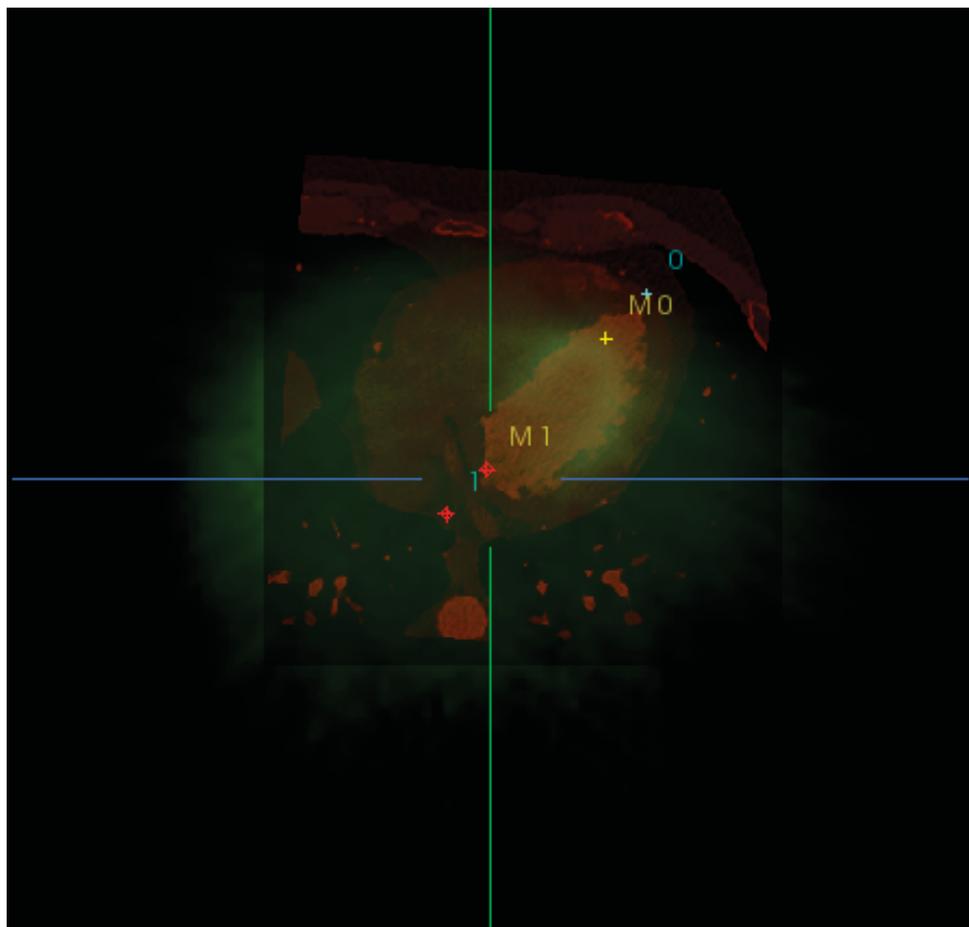


Figura 6.2: Registro entre as imagens utilizando método rígido

Os resultados da tabela 6.1 apresentam as médias de erro, em *pixels*, entre as técnicas SPECT e CT. Alguns trabalhos da literatura, utilizam esse tipo de medida para avaliar o

registro, como visto em Minaeizaeim et al. (2017); Pelizzari et al. (1989); Tomaevič et al. (2002); van Herk et al. (1998). Nos experimentos, foram utilizados algoritmos baseados em pontos rígidos e não rígidos por similaridade Myronenko et al. (2007).

Os melhores resultados foram obtidos com o método por similaridade, pois além de translação e rotação, as escalas das imagens também são ajustadas para um melhor registro.

Tabela 6.1: Média dos resultados de erro de distância dos *pixels* entre modalidades

Erro médio de distância entre pontos n=50	
Antes do registro	175.49
Método rígido	30.17
Método não rígido	2.19

Como visto na tabela 6.1, os valores iniciais médios de distância partem de valores altos (175), e diminuem de acordo com a técnica de registro utilizada. Baixos valores de RMS, demonstram uma boa relação no registro, pois é pequena a distância restante entre cada par de *landmarks* após o registro.

A figura 6.3, apresenta um exemplo do uso do método não rígido Myronenko and Song (2009) para registro das imagens.

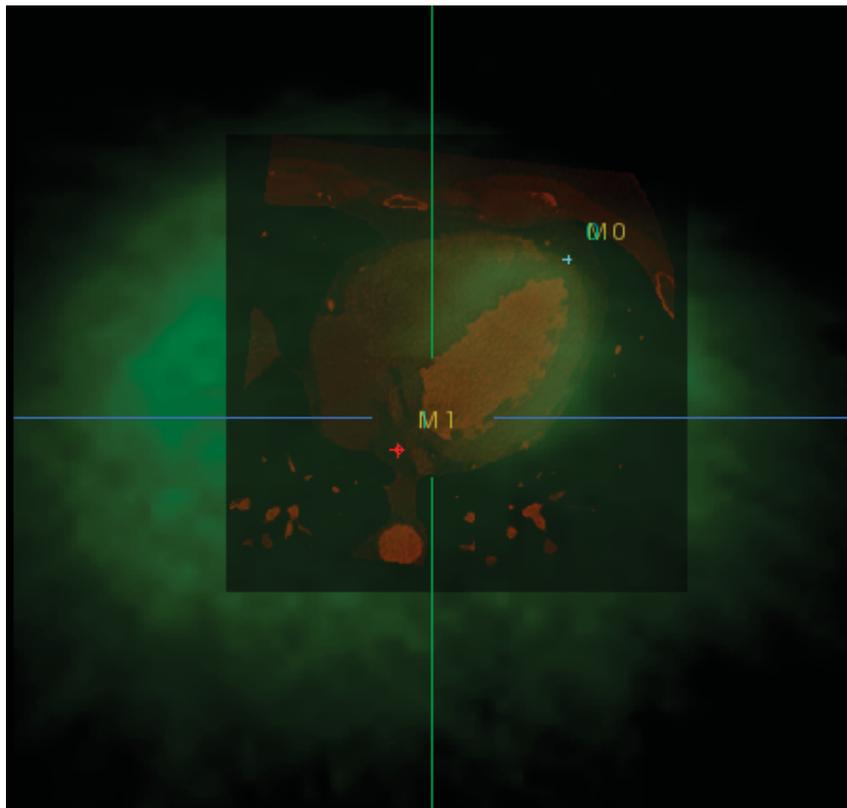


Figura 6.3: Registro das imagens SPECT e CT utilizando utilizando método não rígido

Para comparação, o gráfico da figura 6.4, podemos visualizar a correlação entre o IoU e RMS de distância dos *pixels*, entre a base GT e predição dos melhores modelos de ambas técnicas de imagens estudadas.

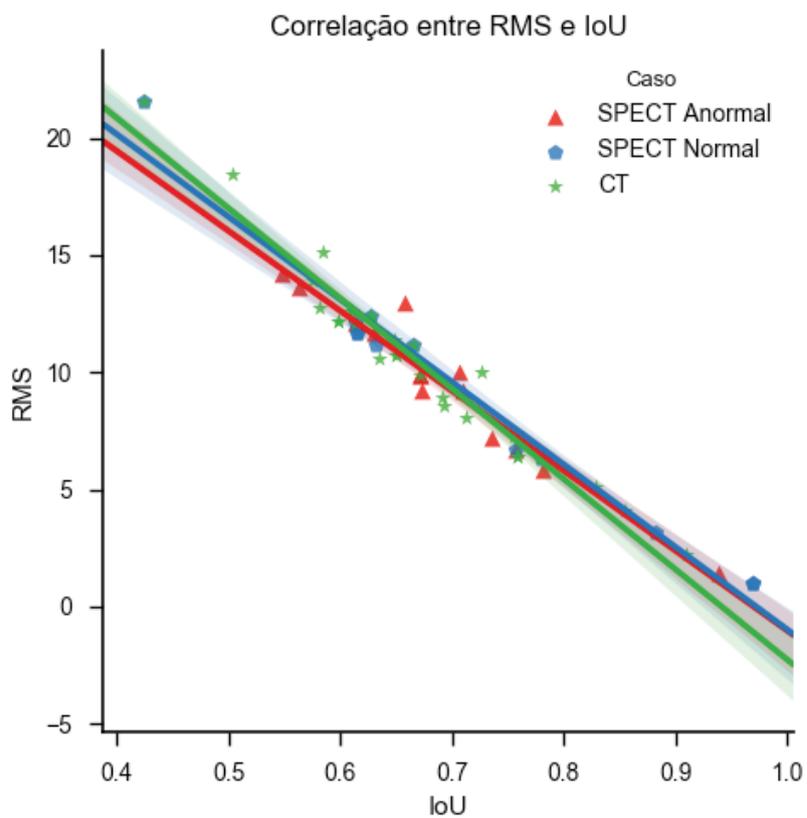


Figura 6.4: Gráfico de correlação entre RSM e IoU de ambas técnicas SPECT e CT. A correlação do SPECT, foi dividida em duas classes, casos normais e casos anormais.

CAPÍTULO 7

CONCLUSÃO

A detecção e segmentação de segmentos do coração humano não é uma tarefa trivial. Foram encontradas na literatura diferentes técnicas e abordagens, sobre como realizar de maneira satisfatória esses passos.

Um dos objetivos do trabalho, foi determinar qual o melhor filtro e técnica de reconstrução e interpolação da técnica SPECT, oferece as melhores informações para uma boa classificação das imagens SPECT. Esse objetivo foi alcançado, onde as imagens SPECT reconstruídas utilizando o filtro *butterworth*, apresentaram melhor acurácia na classificação de até 98% dos segmentos cardíacos estudados. Por esse fato, para detecção e reorientação das imagens SPECT, esse modelo treinado foi utilizado.

Outros resultados importantes obtidos, foram as de sobreposição das regiões preditas em comparação com as da base *growth-truth*. Resultados de IoU 0.97 e IoU 0.91 de sobreposição nas imagens SPECT e CT, respectivamente, indicam uma alta precisão de detecção dos segmentos nas imagens.

Outros importantes resultados foram a acurácia do classificador, em alguns casos observados de até 99% (validação) e 98% (testes), tanto para imagens da técnica SPECT, quanto CT.

Outro ponto que podemos concluir é a concordância entre a base anotada por especialistas e a abordagem proposta, do eixo longo horizontal, diferença média de apenas $-0,450^\circ$ (IC 95% -2.065 até 1.165), muito abaixo do clinicamente aceitável de 10° .

Resultados preliminares demonstraram que a abordagem de detecção do VE nas imagens SPECT e CT tem um futuro promissor. Os experimentos realizados de registro obtiveram, apesar de utilizar apenas dois pontos de referência em cada modalidade, um registro satisfatório utilizando métodos não rígidos, com pequenas distâncias entre as *landmarks*, com RMS de 2.19 *pixels* entre as modalidades. É importante salientar que as

imagens foram adquiridas em equipamentos médicos distintos e em datas diferentes.

Como limitação do registro multimodal, o algoritmo define o *slice* central de cada modalidade, do volume gerado com as imagens tomográficas, utilizando o centro de massa de cada um, com efeitos experimentais, inicialmente sem fins diagnósticos.

A base de imagens médicas utilizadas nos treinos e experimentos não foi tão ampla (com milhares de imagens anotadas), como comumente visto em trabalhos de outras áreas que utilizaram Faster R-CNN. De qualquer forma, a abordagem mostrou-se promissora na detecção e reorientação do VE das imagens médicas cardíacas, com objetivo de geração de imagens híbridas.

7.1 Trabalhos futuros

Como trabalhos futuros, pretende-se aumentar a base de treinamento das modalidades para melhorar a acurácia dos classificações e predição das regiões de predição. Ainda, gerar uma base com imagens da técnica SPECT adquiridas de maneira sincronizada por eletrocardiogramas.

Também pretende-se desenvolver um método para escolha automática dos *slices* com maiores correspondências entre cada técnica de imagem SPECT e CT, já que a resolução espacial das técnicas são diferentes.

Experimentos para geração de volumes tridimensionais híbridos, com correspondência visual entre problemas funcionais decorrentes dos problemas anatômicos, serão realizados. Por fim, pretende-se gerar um método que tenha validação médica, da avaliação desses volumes.

Outros trabalhos, como geração de um métrica para mensuração da reorientação do VE seja ela automática, ou feita por especialistas e validação clínica da fusão 2D, estão em andamento.

REFERÊNCIAS

- Ayelet Akselrod-Ballin, Leonid Karlinsky, Sharon Alpert, Sharbell Hasoul, Rami Ben-Ari, and Ella Barkan. A region based convolutional network for tumor detection and classification in breast mammography. In *Deep Learning and Data Labeling for Medical Applications*, pages 197–205. Springer, 2016. 14
- Douglas G Altman and J Martin Bland. Measurement in medicine: the analysis of method comparison studies. *The statistician*, pages 307–317, 1983. 39
- Brian B Avants, Nicholas J Tustison, Gang Song, Philip A Cook, Arno Klein, and James C Gee. A reproducible evaluation of ants similarity metric performance in brain image registration. *Neuroimage*, 54(3):2033–2044, 2011. 18
- Paul J Besl, Neil D McKay, et al. A method for registration of 3-d shapes. *IEEE Transactions on pattern analysis and machine intelligence*, 14(2):239–256, 1992. 50
- Manuel D Cerqueira, Neil J Weissman, Vasken Dilsizian, Alice K Jacobs, Sanjiv Kaul, Warren K Laskey, Dudley J Pennell, John A Rumberger, Thomas Ryan, Mario S Verani, et al. Standardized myocardial segmentation and nomenclature for tomographic imaging of the heart. *Circulation*, 105(4):539–542, 2002. 13, 31, 36
- Guilhem Chéron, Ivan Laptev, and Cordelia Schmid. P-CNN: pose-based CNN features for action recognition. *CoRR*, abs/1506.03607, 2015. 14
- William R Crum, Thomas Hartkens, and DLG Hill. Non-rigid image registration: theory and practice. *The British journal of radiology*, 77(suppl.2):S140–S153, 2004. 18
- Ariel Hernán Curiale, Flavio D. Colavecchia, Pablo Kaluza, Roberto A. Isoardi, and Germán Mato. Automatic myocardial segmentation by using A deep learning network in cardiac MRI. *CoRR*, abs/1708.07452, 2017. 14

- Stanley R Deans. Hough transform from the radon transform. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (2):185–188, 1981. 23
- Jérôme Declerck, Jacques Feldmar, Michael L Goris, and Fabienne Betting. Automatic registration and alignment on a template of cardiac stress and rest reoriented spect images. *IEEE transactions on medical imaging*, 16(6):727–737, 1997. 17
- Lee R Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302, 1945. 18
- A Ezekiel, K Van Train, D Berman, G Silagan, J Maddahi, and EV Garcia. Automatic determination of quantitation parameters from tc-sestamibi myocardial tomograms. In *Computers in Cardiology 1991, Proceedings.*, pages 237–240. IEEE, 1991. 17
- Sérgio Figueiredo and Pedro Fragoso Costa. *Myocardial Perfusion Imaging - Image Processing and Software*. 10 2014. ISBN 978-3-902785-09-1. 13
- Andrew W Fitzgibbon, Robert B Fisher, et al. A buyer’s guide to conic fitting. *DAI Research paper*, 1996. 38
- J Michael Fitzpatrick, Jay B West, and Calvin R Maurer. Predicting error in rigid-body point-based registration. *IEEE transactions on medical imaging*, 17(5):694–702, 1998. 50
- Albert Flotats, Juhani Knuuti, Matthias Gutberlet, Claudio Marcassa, Frank M Bengel, Philippe A Kaufmann, Michael R Rees, Birger Hesse, et al. Hybrid cardiac imaging: Spect/ct and pet/ct. a joint position statement by the european association of nuclear medicine (eanm), the european society of cardiac radiology (escr) and the european council of nuclear cardiology (ecnc). *European journal of nuclear medicine and molecular imaging*, 38(1):201–212, 2011. 13, 17
- Oliver Gaemperli, Tiziano Schepis, Ines Valenta, Lars Husmann, Hans Scheffel, Victor Duerst, Franz R Eberli, Thomas F Luscher, Hatem Alkadhi, and Philipp A Kaufmann.

- Cardiac image fusion from stand-alone spect and ct: clinical experience. *Journal of Nuclear Medicine*, 48(5):696–703, 2007. 12
- Guido Germano. Technical aspects of myocardial spect imaging. *Journal of Nuclear Medicine*, 42(10):1499–1507, 2001. 23
- Guido Germano, Paul B Kavanagh, Hsiao-Te Su, Marco Mazzanti, Hosen Kiat, Rory Hachamovitch, Kenneth F Van Train, Joseph S Areeda, and Daniel S Berman. Automatic reorientation of three-dimensional, transaxial myocardial perfusion spect images. *Journal of Nuclear Medicine*, 36(6):1107–1114, 1995. , 14, 21, 36
- Ross Girshick. Fast r-cnn. *arXiv preprint arXiv:1504.08083*, 2015. 19, 32
- Ross B. Girshick, Forrest N. Iandola, Trevor Darrell, and Jitendra Malik. Deformable part models are convolutional neural networks. *CoRR*, abs/1409.5403, 2014. 14
- Steven Gold, Chien-Ping Lu, Anand Rangarajan, Suguna Pappu, and Eric Mjolsness. New algorithms for 2d and 3d point matching: Pose estimation and correspondence. In *Advances in Neural Information Processing Systems*, pages 957–964, 1995. 50
- Lee W Goldman. Principles of ct and ct technology. *Journal of nuclear medicine technology*, 35(3):115–128, 2007. 11
- Christoph Guetter, Chenyang Xu, Frank Sauer, and Joachim Hornegger. Learning based non-rigid multi-modal image registration using kullback-leibler divergence. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 255–262. Springer, 2005. 18
- Mojgan Haddad and Gerold Porenta. Impact of reorientation algorithms on quantitative myocardial spect perfusion imaging. *The Journal of Nuclear Medicine*, 39(11):1864, 1998. 13
- Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *CoRR*, abs/1512.03385, 2015a. 19, 26, 41

- Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *corr abs/1512.03385* (2015), 2015b. 26, 41
- Milena J Henzlova, W Lane Duvall, Andrew J Einstein, Mark I Travin, and Hein J Verberne. Asnc imaging guidelines for spect nuclear cardiology procedures: Stress, protocols, and tracers. *Journal of Nuclear Cardiology*, 23(3):606–639, 2016. 12, 27
- Derek LG Hill, Philipp G Batchelor, Mark Holden, and David J Hawkes. Medical image registration. *Physics in medicine & biology*, 46(3):R1, 2001. 18
- BF Hutton, M Braun, and P Slomka. Image registration techniques in nuclear medicine imaging. In *Quantitative analysis in nuclear Medicine imaging*, pages 272–307. Springer, 2006. 18
- Avinash C Kak and Malcolm Slaney. *Principles of computerized tomographic imaging*. SIAM, 2001. 23
- Philipp A. Kaufmann and Marcelo F. Di Carli. Hybrid spect/ct and pet/ct imaging: The next step in noninvasive cardiac imaging. *Seminars in Nuclear Medicine*, 39(5):341 – 347, 2009. ISSN 0001-2998. Hybrid Imaging Anniversary Issue (Part II). 12
- Philipp A Kaufmann and Marcelo F Di Carli. Hybrid spect/ct and pet/ct imaging: the next step in noninvasive cardiac imaging. In *Seminars in nuclear medicine*, volume 39, pages 341–347. Elsevier, 2009. 13
- Magdy M Khalil. Elements of gamma camera and spect systems. In *Basic Sciences of Nuclear Medicine*, pages 155–178. Springer, 2010. 23
- Diederik Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. 26, 41
- Hongming Li and Yong Fan. Non-rigid image registration using fully convolutional networks with deep self-supervision. *arXiv preprint arXiv:1709.00799*, 2017. 18

- Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In *European conference on computer vision*, pages 740–755. Springer, 2014. 32
- Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440, 2015. 19
- Maria Lyra and Agapi Ploussi. Filtering in spect image reconstruction. *Journal of Biomedical Imaging*, 2011:10, 2011. 24
- Jiayi Ma, Ji Zhao, and Alan L Yuille. Non-rigid point set registration by preserving global and local structures. *IEEE Transactions on image Processing*, 25(1):53–64, 2016. 50
- VRS Mani et al. Survey of medical image registration. *Journal of Biomedical Engineering and Technology*, 1(2):8–25, 2013. 50
- Shinro Matsuo, Kenichi Nakajima, Nasima Akhter, Hiroshi Wakabayashi, Junichi Taki, Koichi Okuda, and Seigo Kinuya. Clinical usefulness of novel cardiac mdct/spect fusion image. *Annals of nuclear medicine*, 23(6):579–586, 2009. 12
- Fred A Mettler and Milton J Guiberteau. *Essentials of Nuclear Medicine Imaging E-Book*. Elsevier Health Sciences, 2011. 11, 23
- Peter Mildemberger, Marco Eichelberg, and Eric Martin. Introduction to the dicom standard. *European radiology*, 12(4):920–927, 2002. 11
- Hamed Minaeizaeim, Haribalan Kumar, Merryn H Tawhai, and Alys R Clark. Lung image registration between postures for disease tracking, 2017. 51
- Claudia Maria Cabral Moro et al. Metodo automatico para construção de mapas polares. 1994. 17
- Manish Motwani, Daniel S Berman, Guido Germano, and Piotr Slomka. Automated quantitative nuclear cardiology methods. *Cardiology clinics*, 34(1):47–57, 2016. 16

- Andriy Myronenko and Xubo B. Song. Point-set registration: Coherent point drift. *CoRR*, abs/0905.2635, 2009. URL <http://arxiv.org/abs/0905.2635>. 51
- Andriy Myronenko, Xubo Song, and Miguel A Carreira-Perpinán. Non-rigid point set registration: Coherent point drift. In *Advances in Neural Information Processing Systems*, pages 1009–1016, 2007. 51
- Francisco PM Oliveira and Joao Manuel RS Tavares. Medical image registration: a review. *Computer methods in biomechanics and biomedical engineering*, 17(2):73–93, 2014. 13
- OMS (World Health Organization). Who the top 10 causes of death, 2017. URL <http://www.who.int/mediacentre/factsheets/fs310/en/>. 11
- Catherine M. Otto. *Textbook of Clinical Echocardiography, 5th Edition*. Elsevier, 2013. 11
- Aju P Pazhenkottil, Rene N Nkoulou, Jelena-Rima Ghadri, Bernhard A Herzog, Silke M Küest, Lars Husmann, Mathias Wolfrum, Robert Goetti, Ronny R Buechel, Oliver Gaemperli, et al. Impact of cardiac hybrid single-photon emission computed tomography/computed tomography imaging on choice of treatment strategy in coronary artery disease. *European heart journal*, 32(22):2824–2829, 2011. 12
- Charles A Pelizzari, George TY Chen, Danny R Spelbring, Ralph R Weichselbaum, and Chin-Tu Chen. Accurate three-dimensional registration of ct, pet, and/or mr images of the brain. *Journal of computer assisted tomography*, 13(1):20–26, 1989. 51
- Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in neural information processing systems*, pages 91–99, 2015. , 20
- Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. *IEEE transactions on pattern analysis and machine intelligence*, 39(6):1137–1149, 2017. , 14, 19, 20, 26

- Pierre Sermanet, David Eigen, Xiang Zhang, Michaël Mathieu, Rob Fergus, and Yann LeCun. Overfeat: Integrated recognition, localization and detection using convolutional networks. *CoRR*, abs/1312.6229, 2013. 14
- Helen Sonesson, Joey FA Ubachs, Martin Ugander, Håkan Arheden, and Einar Heiberg. An improved method for automatic segmentation of the left ventricle in myocardial perfusion spect. *Journal of Nuclear Medicine*, 50(2):205–213, 2009. 16
- Charles V Stewart, Chia-Ling Tsai, and Badrinath Roysam. The dual-bootstrap iterative closest point algorithm with application to retinal image registration. *IEEE transactions on medical imaging*, 22(11):1379–1394, 2003. 18
- Lisa Tang, Ghassan Hamarneh, and Anna Celler. Co-registration of bone ct and spect images using mutual information. In *Signal Processing and Information Technology, 2006 IEEE International Symposium on*, pages 116–121. IEEE, 2006. 13
- Dejan Tomaevič, Boštjan Likar, and Franjo Pernuš. “gold standard” 2d/3d registration of x-ray to ct and mr images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 461–468. Springer, 2002. 51
- Yossi Tsadok, Yael Petrank, Sebastian Sarvari, Thor Edvardsen, and Dan Adam. Automatic segmentation of cardiac mri cines validated for long axis views. *Computerized Medical Imaging and Graphics*, 37(7-8):500–511, 2013. 16
- Marcel van Herk, Jan C de Munck, Joos V Lebesque, Saar Muller, Coen Rasch, and Adriaan Touw. Automatic registration of pelvic computed tomography data and magnetic resonance scans including a full circle method for quantitative accuracy evaluation. *Medical physics*, 25(10):2054–2067, 1998. 51
- Kristen M Waterstram-Rich and David Gilmore. *Nuclear Medicine and PET/CT-E-Book: Technology and Techniques*. Elsevier Health Sciences, 2016. 23
- Ronghua Yang, Majid Mirmehdi, and David Hall. A charged contour model for cardiac spect segmentation. *Department of Computer Science, University of Bristol*, 2006. 17

F. H. B. Zavan. Nose pose estimation in the wild and its applications on nose tracking and 3d face alignment. In *Dissertação de Mestrado, Universidade Federal do Paraná - UFPR*, 2016. 14

Junkang Zhang, Haigen Hu, Shengyong Chen, Yujiao Huang, and Qiu Guan. Cancer cells detection in phase-contrast microscopy images based on faster r-cnn. In *Computational Intelligence and Design (ISCID), 2016 9th International Symposium on*, volume 1, pages 363–367. IEEE, 2016. 14