

INSTITUTO FEDERAL DO ESPÍRITO SANTO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE CONTROLE E
AUTOMAÇÃO

DING YIH AN

**ESTIMATIVA DA IDADE ÓSSEA USANDO MOSAICOS DOS CENTROS DE
OSSIFICAÇÃO DE RADIOGRAFIAS CARPAIS COMO IMAGENS DE ENTRADA
PARA *DEEP LEARNING***

Serra
2017

DING YIH AN

**ESTIMATIVA DA IDADE ÓSSEA USANDO MOSAICOS DOS CENTROS DE
OSSIFICAÇÃO DE RADIOGRAFIAS CARPAIS COMO IMAGENS DE ENTRADA
PARA *DEEP LEARNING***

Dissertação apresentada ao Programa de Pós-graduação em Engenharia de Controle e Automação do Instituto Federal do Espírito Santo como requisito parcial para a obtenção do título de Mestre em Engenharia de Controle e Automação.

Orientadora: Profa. Dra. Karin Satie Komati
Orientador: Prof. Dr. Luiz Alberto Pinto

Serra
2017

Dados Internacionais de Catalogação na Publicação (CIP)

A531e An, Ding Yih
2017 Estimativa da idade óssea usando mosaicos dos centros de ossificação de radiografias carpais como imagens de entrada para *Deep Learning* / Ding Yih An. - 2017.
83 f.; il.; 30 cm

Orientadora: Profa. Dra. Karin Satie Komati.

Orientador: Prof. Dr. Luiz Alberto Pinto.

Dissertação (mestrado) - Instituto Federal do Espírito Santo, Programa de Pós-graduação em Engenharia de Controle de Automação, 2017.

1. Visão por computador. 2. Processamento de imagens. 3. Radiografia. 4. Ossos. I. Komati, Karin Satie. II. Pinto, Luiz Alberto. III. Instituto Federal do Espírito Santo. IV. Título.

CDD 006.37

Bibliotecária Rogéria Gomes Belchior - CRB6/ES 417

MINISTÉRIO DA EDUCAÇÃO
INSTITUTO FEDERAL DO ESPÍRITO SANTO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE CONTROLE E AUTOMAÇÃO

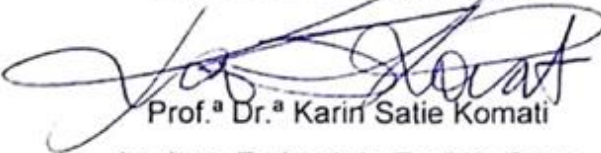
DING YIH AN

**ESTIMATIVA DA IDADE ÓSSEA USANDO MOSAICOS DOS CENTROS DE
OSSIFICAÇÃO DE RADIOGRAFIAS CARPAIS DOMO IMAGENS DE ENTRADA
PARA DEEP LEARNING**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Controle e Automação do Instituto Federal do Espírito Santo, como requisito parcial para obtenção de título de Mestre em Engenharia de Controle e Automação.

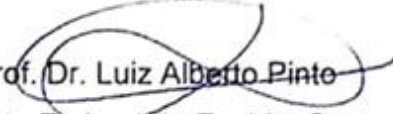
Aprovado em 21 de dezembro de 2017

COMISSÃO EXAMINADORA




Prof.ª Dr.ª Karini Satie Komati


Instituto Federal do Espírito Santo
Orientadora



Prof. Dr. Luiz Alberto Pinto
Instituto Federal do Espírito Santo
Orientador



Prof. Dr. Flávio Garcia Pereira
Instituto Federal do Espírito Santo
Membro Interno



Prof. Dr. Klaus Fabian Côco
Universidade Federal do Espírito Santo
Membro externo

A DEUS, por ser essencial em minha vida, meu guia para trilhar no caminho mais correto possível.

Ao meu tio Chang, por sempre acreditar em mim e me orientar nas decisões mais complexas da vida.

A minha família que, com muito carinho e apoio, não mediram esforços para que eu chegasse até esta etapa de minha vida, minha base.

Ao meu namorado que, além de me fazer feliz, deu-me todo apoio e incentivo para que eu conquistasse um lugar ao sol, meu porto seguro.

AGRADECIMENTOS

Agradeço a Deus por me dar a vida, a saúde e por sempre me dar forças para levantar e seguir em frente.

Ao meu tio Chang, por acreditar no meu potencial e confiar em minha pessoa.

À minha família, que me dá todo o apoio para eu ir em busca dos meus sonhos e objetivos, sem os quais eu não sou.

Ao meu namorado que permaneceu ao meu lado nos bons e maus momentos e que me dá todo o incentivo para eu ser o melhor de mim.

À minha orientadora Profa. Dra. Karin Satie Komati por ser um grande exemplo profissional, por toda paciência, compreensão e dedicação.

Ao meu orientador Prof. Dr. Luiz Alberto Pinto, pela ajuda e esforço.

Ao Prof. Dr. Flávio Garcia Pereira e ao Prof. Dr. Klaus Fabian Côco pelas colocações relevantes para a melhora do presente trabalho.

Ao Prof. Dr. Maxwell Eduardo Monteiro, pelas lindas palavras de incentivo.

A todos aqueles que acreditaram em mim, que me acrescentaram de alguma forma para chegar onde estou e pretendo seguir. Sinceros agradecimentos a todos!

RESUMO

A estimaco da idade ssea tem sido utilizada na medicina para verificar se o grau de desenvolvimento da estrutura ssea de um indivduo corresponde a sua idade cronolgica. Essa estimativa serve para prognsticos no desenvolvimento de crianas e adolescentes assim como diagnsticos de doenas endocrinolgicas. Contudo, a quase totalidade dos mtodos utilizados atualmente depende da interveno dos usurios, seja para fazer ajustes com a finalidade de melhorar a qualidade das imagens de entrada, seja para corrigir, ou mesmo indicar manualmente a localizao dos centros de ossificao. Esse trabalho prope uma metodologia totalmente automatizada para a estimaco da idade ssea que consiste na segmentao dos centros de ossificao de radiografias carpais e a aplicao de Redes Neurais Convolucionais (Convolutional Neural Networks - CNN) com a tcnica de *Deep Learning* para fins de classificao. A etapa de segmentao consistiu na identificao e recorte das regies dos dedos nas imagens radiogrficas atravs da anlise dos pontos de referncia que delimitam a regio da mo nas imagens binarizadas, e a posterior extrao dos centros de ossificao. Na etapa de classificao, CNN pr treinadas e alto desempenho foram adotadas: GoogLeNet, Alexnet e VGG-19. Para os testes foram utilizadas 1380 imagens radiogrficas carpais obtidas na base de dados da Universidade da Califrnia. A segmentao das imagens originais resultou em quatro conjuntos de imagens, sendo dois constitudos por mosaicos de imagens resultantes da segmentao dos dedos, com e sem a imagem segmentada da regio do punho, e dois formados por mosaicos de imagens obtidas na extrao dos centros de ossificao, com e sem a imagem segmentada da regio do punho. Para cada conjunto de imagens foi constituda uma verso masculina e uma verso feminina. Os desempenhos da classificao dos dedos e dos centros de ossificao com a tolerncia de 2 anos de idade foram de 81,38% e 78,17%, respectivamente. Alm dos mosaicos de imagens dos dedos e dos centros de ossificao foi utilizado tambm o conjunto constitudo pelas imagens radiogrficas carpais originais sem qualquer tipo de pr-processamento. Considerando a acurcia, os melhores resultados na etapa de classificao foram de 88,12% para indivduos masculinos, e 83,31% para indivduos femininos. Ambos foram obtidos com a rede VGG-19, para indivduos na faixa de 5-18 anos de idade, com uma tolerncia de 2 anos de idade.

Palavras Chaves: Processamento de imagens. Radiografias carpais. Segmentao dos Centros de Ossificao. Classificao. CNN. *Deep Learning*.

ABSTRACT

Bone age estimation has been used in medicine to verify whether the bone structure development degree of a person corresponds to their chronological age. Such estimative is useful for prognosis about the development of children and adolescents, as well as endocrinological diseases diagnoses. However, the majority of the methods currently used depend on user intervention, either to make adjustments for the purpose of improving the quality of the input images or to correct, or even manually indicate the location of the ossification centers in semi-automated and manual systems. This work proposes a fully automated methodology for the bone age estimation task, which consists of the ossification centers segmentation in the carpal radiographs, and Convolutional Neural Networks (CNN) application, with Deep Learning technique for classification purposes. The segmentation stage consisted of identifying and cutting the fingers regions off the radiographic images, by means of the references points' analysis which delimitate the hand region in the binarized images, and the subsequent extraction of the ossification centers. In the classification stage, high performance pre-trained CNN were adopted: GoogLeNet, Alexnet and VGG-19. Tests were carried out on 1380 radiographic images obtained from the California University database. The original images segmentation produced four images sets, two of them consisting of images mosaics resulting from fingers segmentation, with and without the wrist region segmented image, and the other two formed by images mosaics obtained from the ossification centers extraction, with and without the wrist region segmented image. For each image set it had been constituted a male and a female version. The fingers and ossification centers classification, with 2 years tolerance were 81.38% and 78.17%, respectively. Beyond the fingers and ossification centers images mosaics, an additional image set consisting of the original radiographic images without any type of preprocessing had been also used. Considering the accuracy, the best results in the classification stage were 88.12% for males and 83.31% for females. Both of them were obtained through VGG-19 network, for individuals in the range of 5-18 years of age, with a tolerance of 2 years of age.

Keywords: Image processing. Carpal Radiographs. Ossification Center Segmentation. Classification. CNN. Deep Learning.

LISTA DE FIGURAS

Figura 1 - Surtos de crescimentos em meninos e meninas.....	13
Figura 2 - Centros de ossificação carpal de acordo com o método GP	15
Figura 3 - Centros de ossificação carpal avaliados pelo método TW.....	17
Figura 4 - Representação dos 10 centros de ossificação do método de ER.....	18
Figura 5 - (a) Imagem inteira da radiografia carpal (b) Imagem mosaico dos dedos (c) Imagem mosaico dos dedos com o punho (d) Imagem mosaico dos centros de ossificação (e) Imagem mosaico dos centros de ossificação com o punho	20
Figura 6 - (a) Pontos de características disperso (b) Pontos de características denso	27
Figura 7 - Um exemplo de CNN multicamada	29
Figura 8 - Três estágios de uma camada convolucional	30
Figura 9 - Arquitetura do sistema	36
Figura 10 - Fluxograma do protocolo QAP	38
Figura 11 - Exemplos de imagens que compõem a base de dados com as dimensões de largura x altura respectivamente	39
Figura 12 - Exemplos de imagens pouco padronizadas que compõem a base de dados com as dimensões de largura x altura respectivamente.....	41
Figura 13 - Arquitetura de recorte das PROI.....	42
Figura 14 - (a) Imagem original (b) Imagem binarizada (c) Imagem rotacionada com o punho removido.....	43
Figura 15 - (a) Contorno superior da mão (b) Mão com os pontos de referência de corte determinadas.....	45
Figura 16 - (a) Ossos da mão binarizadas (b) Metacarpos obtidos da circunferência	46
Figura 17 - Imagem auxiliar do retângulo de cada um dos dedos.....	47
Figura 18 - Imagem auxiliar dos retângulos rotacionados.....	47
Figura 19 - Imagem da mão rotacionada de acordo com cada orientação	47
Figura 20 - Imagem da região dos PROI recortados.....	48
Figura 21 - Arquitetura de extração das EROI	48
Figura 22 - (a) Dedo original (b) Dedo recortado da figura 18 (c) Binarização do dedo original (d) Resultado da soma das duas binarizações sem ruídos	49
Figura 23 - Intensidade horizontal do dedo mínimo binarizado	51
Figura 24 - Metodologias das proporções áureas. (a) N-ouro-1 (b) N-ouro-2 (c) N-ouro-3 (d) N-ouro-4.....	52

Figura 25 - Os 14 centros de ossificação extraídos	53
Figura 26 - Exemplo de uma imagem da região do punho segmentada para somar à composição dos mosaicos	54
Figura 27 - Imagem radiográfica da mão sem pré-processamento	55
Figura 28 - Composição das PROI.....	56
Figura 29 - Mosaico composto pelos 14 centros de ossificação de uma imagem radiográfica.....	56
Figura 30 - Composição das PROI com o punho	57
Figura 31 - Mosaico composto pelos 14 centros de ossificação com o punho.....	57
Figura 32 - (a) Imagem redimensionada da mão esquerda (b) Imagem mosaico redimensionada da composição dos dedos (c) Imagem mosaico redimensionada da composição dos dedos com o punho (d) Imagem mosaico redimensionada dos centros de ossificação (e) Imagem mosaico redimensionada dos centros de ossificação com o punho.....	58
Figura 33 - Desempenho das 5 diferentes tipos de imagens da VGG-19	72
Figura 34 - Composição das PROI com o punho	74

LISTA DE QUADROS

Quadro 1 - Distribuição das imagens obtidas no <i>download</i>	41
Quadro 2 - Distribuição das imagens da mão por idade óssea	61
Quadro 3 - Distribuição das imagens de dedos por idade óssea	62
Quadro 4 - Distribuição das imagens de mosaico dos centros de ossificação por idade óssea.....	63
Quadro 5 - Resultados mão feminino.....	64
Quadro 6 - Resultados mão masculino	65
Quadro 7 - Resultados dedos feminino	66
Quadro 8 - Resultados dedos masculino	67
Quadro 9 - Resultados mosaico feminino	67
Quadro 10 - Resultados mosaico masculino.....	68
Quadro 11 - Resultados dedos e punho feminino	68
Quadro 12 - Resultados dedos e punho masculino.....	69
Quadro 13 - Resultados mosaico e punho feminino.....	70
Quadro 14 - Resultados mosaico e punho masculino.....	70
Quadro 15 - Melhores desempenhos do presente trabalho 0-18 anos	71
Quadro 16 - Melhores desempenhos do presente trabalho 5-18 anos	71
Quadro 17 - Comparação com o desempenho dos trabalhos correlatos	73

SUMÁRIO

1	INTRODUÇÃO	12
1.1	CONTEXTUALIZAÇÃO	12
1.2	MOTIVAÇÃO	14
1.3	PROPOSTA DO TRABALHO	17
1.4	OBJETIVOS	19
1.4.1	Objetivos Específicos	19
1.5	PUBLICAÇÕES	21
1.6	ORGANIZAÇÃO DO TRABALHO	21
2	TRABALHOS CORRELATOS	22
2.1	GRUPO DE PESQUISA DE TAIWAN	22
2.2	PESQUISA EM UNIVERSIDADES PAULISTAS	25
2.3	PESQUISA COMPARATIVA DE EXTRATORES DE CARACTERÍSTICAS	26
2.4	PESQUISAS BASEADAS EM <i>DEEP LEARNING</i>	27
2.5	CONSIDERAÇÕES SOBRE TRABALHOS CORRELATOS	33
3	MATERIAIS E MÉTODOS	35
3.1	BASE DE DADOS	36
3.2	IMAGEM MOSAICO DOS DEDOS	42
3.2.1	Busca pelos pontos de referência para realizar os cortes	44
3.2.2	Recorte dos PROI	46
3.3	IMAGEM MOSAICO DOS CENTROS DE OSSIFICAÇÃO	48
3.3.1	Localização e extração das EROI	49
3.4	IMAGEM MOSAICO COM PUNHO	53
3.5	CLASSIFICAÇÃO COM <i>DEEP LEARNING</i>	54
3.6	CRITÉRIOS DE AVALIAÇÃO	59
4	EXPERIMENTOS, RESULTADOS E DISCUSSÃO	61
4.1	RESULTADOS DA SEGMENTAÇÃO	62
4.2	RESULTADOS DA ESTIMAÇÃO DA IDADE ÓSSEA DAS IMAGENS SEM PROCESSAMENTO	64
4.3	RESULTADOS DA ESTIMAÇÃO DA IDADE ÓSSEA DAS IMAGENS MOSAICO DOS DEDOS	65
4.4	RESULTADOS DA ESTIMAÇÃO DA IDADE ÓSSEA DAS IMAGENS MOSAICO DOS CENTROS DE OSSIFICAÇÃO	67

4.5	RESULTADOS DA ESTIMAÇÃO DA IDADE ÓSSEA DAS IMAGENS MOSAICO DOS DEDOS E PUNHO	68
4.6	RESULTADOS DA ESTIMAÇÃO DA IDADE ÓSSEA DAS IMAGENS MOSAICO DOS CENTROS DE OSSIFICAÇÃO E PUNHO	69
4.7	COMPARAÇÃO ENTRE OS RESULTADOS DESTE TRABALHO	70
4.8	COMPARAÇÃO COM TRABALHOS CORRELATOS	73
4.8.1	Comparação com trabalho de Lee	74
4.8.2	Comparação com artigo de Kashif e colegas	75
4.8.3	Comparação com artigo de Spampinato e colegas	75
5	CONSIDERAÇÕES FINAIS	76
5.1	TRABALHOS FUTUROS	77
	REFERÊNCIAS	79

1 INTRODUÇÃO

1.1 CONTEXTUALIZAÇÃO

A prática médica tem passado por profundas transformações ao longo das últimas décadas, consequência direta da inclusão de novas tecnologias, notadamente da área de computação, aos procedimentos realizados pelos especialistas. Atividades, das mais simples, como a gestão de informações de pacientes com a utilização de prontuários eletrônicos, das mais complexas, como as cirurgias virtuais guiadas (GUIMARÃES et al., 2014), passando pela utilização de sistemas especialistas para auxiliar na elaboração de diagnósticos (FONSECA; VOLPASSO; VIANNA, 2016), têm sido transformadas pelos avanços nas áreas de instrumentação e computação.

Segundo Thrall (2016), a radiologia tem sido líder na era digital da medicina e a que mais acompanhou as inovações na medicina de precisão, mas ainda há espaço para melhorias na automatização da análise, na precisão das informações e em novas aplicações. Uma aplicação que ainda é tema de pesquisa para sua automatização é a avaliação da idade óssea, dada que a CLASSIFICAÇÃO manual é considerada demorada e sofre de uma variabilidade insatisfatória (RIJN; THODBERG, 2016).

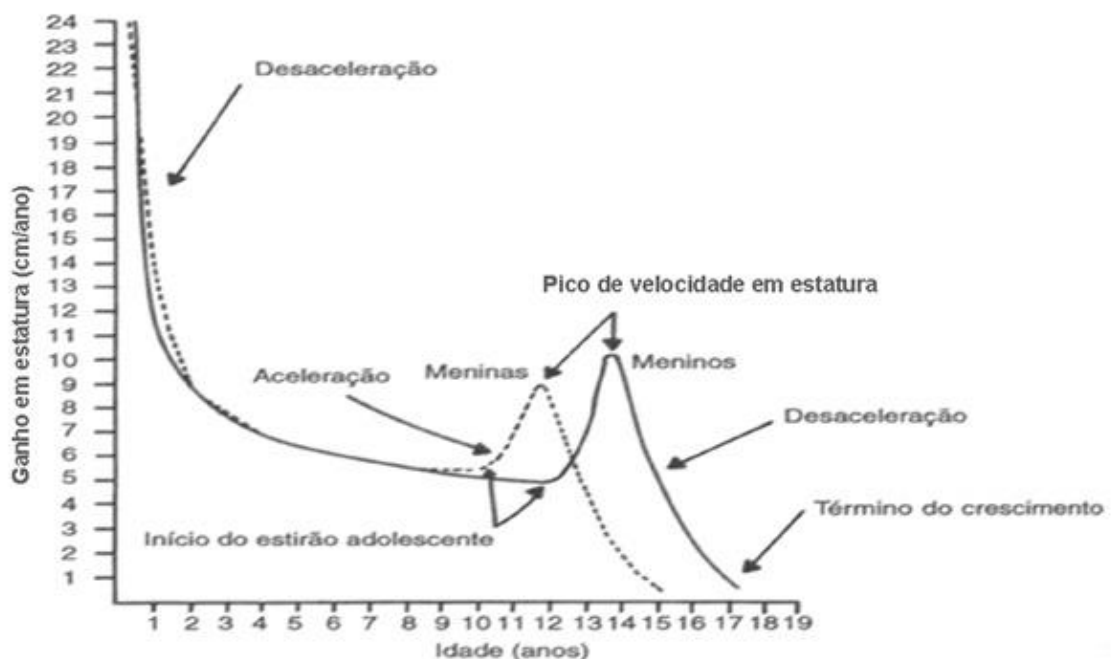
A determinação da idade óssea a partir de radiografia manual é um dos procedimentos radiográficos mais antigos, sendo que a primeira publicação sobre o tema data de 1898. A estimacão da idade óssea pode ser aplicada para prognósticos de crescimento, diagnóstico de doenças endocrinológicas, auxilia na identificação de vítima após desastres, auxilia na avaliação de requerentes de asilo que entram num país sem documentos de identificação, na prevenção da manipulação da idade em nível júnior de competições esportivas, dentre outras (STERN et al., 2014).

Quando realizado por endocrinologistas pediátricos em pacientes na faixa etária de 1 a 18 anos de idade, o exame para estimacão da idade óssea visa verificar se o

grau de desenvolvimento da estrutura óssea de um indivíduo (idade biológica) corresponde a sua idade cronológica.

Nesse contexto, esse exame é realizado para fins de avaliação do processo de desenvolvimento físico de crianças e adolescentes, permitindo estimar o seu potencial de crescimento e se o mesmo transcorre dentro dos padrões de normalidade. A dissociação entre as idades biológicas e cronológicas podem estar relacionadas a patologias do crescimento, bem como atrasos na puberdade (MENDES et al., 2010). Destacam-se três fases de surtos de crescimento, o primeiro ocorre aos três anos de idade. O segundo surto de crescimento ocorre por volta do seis-sete anos nas meninas e aos sete-nove anos nos meninos. O terceiro surto está relacionado à época da puberdade e ocorre aproximadamente aos onze-doze anos nas meninas e aos quatorze-quinze anos nos meninos, sendo denominado Surto de Crescimento Puberal (SCP). A intensidade, duração e época de início do SCP varia individualmente para cada pessoa (RODRIGUES, 2003). Assim sendo, é importante a distinção do sexo do indivíduo no processo de estimação da idade óssea. A Figura 1 apresenta a curva de crescimento em meninos e meninas.

Figura 1 - Surtos de crescimentos em meninos e meninas



Fonte: Ré (2011).

1.2 MOTIVAÇÃO

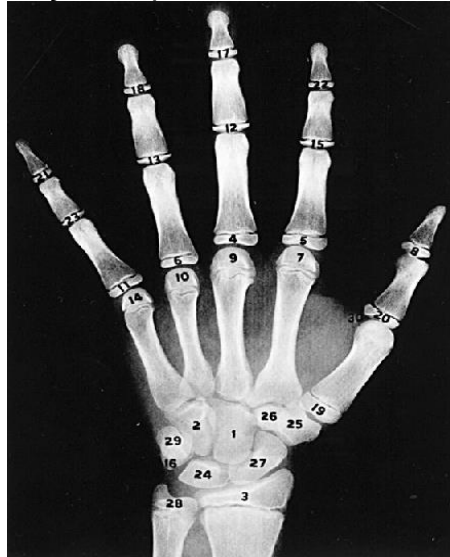
Em sua forma convencional, a estimaco da idade ssea  realizada por um especialista e consiste da inspeco visual de imagens de partes especficas do esqueleto, por exemplo, o p, o ombro, o tornozelo, o quadril, o cotovelo, as vrtebras cervicais e o conjunto mo-punho (CASANOVA et al., 2006). Tais imagens podem ser obtidas por meio de diferentes tcnicas instrumentais, entre elas, raios-X, ressonncia magntica (MOTA et al., 2012), dentre outros. A mais utilizada  o raio-x da mo e punho, a radiografia carpal, em virtude do baixo custo operacional devido ao tempo de exposio reduzido, por ser um membro de fcil acesso e por no conter rgos vitais, podendo ser submetida a doses de radiao com mais frequncia (OLIVETE; RODRIGUES, 2010).

No processo de estimaco, os especialistas avaliam a compatibilidade entre as imagens das partes do esqueleto dos pacientes e as imagens correspondentes que constituem os atlas de referncia de idade ssea (H. NETO et al., 2000). Um atlas de referncia  baseado na adoo de "critrios" e/ou "mtricas". Os critrios definem os parmetros a serem considerados para a atribuo da idade ssea, por exemplo, o estgio de formao de uma determinada regio de ossificao. As mtricas estabelecem as faixas de valores que os parmetros considerados pelo atlas devem assumir de acordo com a idade ssea do paciente, como, a proporo e o tamanho entre as falanges e metacarpos.

Entre os mtodos clssicos para estimaco da idade ssea que utilizam radiografias carpais, podem ser mencionados o de Greulich e Pyle (1959), o de Tanner, Whitehouse e Takaishi (1966), e o de Eklf e Ringertz (1967).

O mtodo de Greulich & Pyle (GP) fornece um atlas composto por 58 radiografias, e prope a comparao de 28 pontos de ossificao com os padres dos raios-X ilustrada na Figura 2.

Figura 2 - Centros de ossificação carpal de acordo com o método GP



Fonte: Moraes et al. (2003).

Os 28 centros de ossificação conforme a Figura 2 são: 1) capitato (osso grande); 2) hamato (ganchoso); 3) epífise distal do rádio; 4) epífise da falange distal do terceiro dedo; 5) epífise da falange proximal do segundo dedo; 6) epífise da falange proximal do quarto dedo; 7) epífise do segundo metacarpo; 8) epífise da falange distal do primeiro dedo; 9) epífise do terceiro metacarpo; 10) epífise do quarto metacarpo; 11) epífise da falange proximal do quinto dedo; 12) epífise da falange média do terceiro dedo; 13) epífise da falange média do quarto dedo; 14) epífise do quinto metacarpo; 15) epífise da falange média do segundo dedo; 16) triquetral; 17) epífise da falange distal do terceiro dedo; 18) epífise da falange distal do quinto dedo; 19) epífise do primeiro metacarpo; 20) epífise da falange proximal do primeiro dedo; 21) epífise da falange distal do quinto dedo; 22) epífise da falange distal do segundo dedo; 23) epífise da falange média do quinto dedo; 24) Semilunar; 25) Trapézio; 26) Trapezóide; 27) Escafóide; 28) Epífise distal da ulna.

O método de Tanner & Whitehouse (TW) atribui pontuações (*scores*) para 20 centros de ossificação: 1) Falange distal dedão; 2) Falange distal dedo médio; 3) Falange distal dedo mínimo; 4) Falange médio dedo médio; 5) Falange médio dedo mínimo; 6) Falange proximal dedão; 7) Falange proximal dedo médio 8) Falange proximal dedo mínimo; e a ROI Carpal mais a Ulna e o Rádio, conforme a Figura 3. Os ossos são avaliados em uma escala de acordo com pesos que lhes são

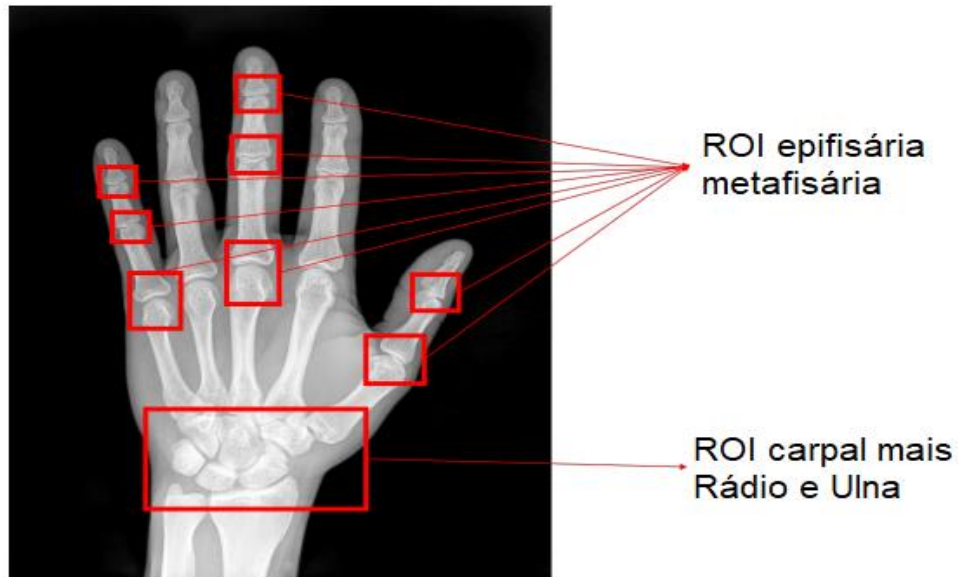
atribuídos. Para a determinação da idade óssea em TW, as somas dos pesos são comparados aos valores de referência das pontuações de uma tabela.

O método Eklof & Ringertz (ER) pondera as medidas de comprimento e/ou largura de 10 ossos carpais escolhidos: 1) Largura da epífise distal do rádio; 2) Comprimento do capitato (grande osso); 3) Largura do capitato; 4) Comprimento do hamato (ganchoso); 5) Largura do hamato; 6) Comprimento do metacarpo II; 7) Comprimento do metacarpo III; 8) Comprimento do metacarpo IV; 9) Comprimento da falange proximal do II dedo; 10) Comprimento da falange proximal do III dedo. São os que melhor representam as variações nos diversos estágios de crescimento no intervalo de 1 ano a 15 anos de idade considerados pelo método, conforme a Figura 4.

O método GP é o mais utilizado (CASTRO et al., 2009), embora os radiologistas reconheçam que a precisão desse método é inferior às dos outros métodos disponíveis. O método TW é mais preciso por ser quantitativo. O GP permanece em utilização, pois os outros métodos são mais complicados de serem aplicados manualmente. Horter et al. (2012), apresentaram a informação de que um especialista, em média, necessita de 40 segundos para fazer a análise via método de GP e 80 segundos pelo método TW.

Em todos os métodos convencionais, a estimacão da idade óssea é um procedimento visual, logo a validacão dos critérios e a obtençao das métricas se tornam questões subjetivas.

Figura 3 - Centros de ossificação carpal avaliados pelo método TW



Fonte: Adaptado de Spampinato et al. (2017).

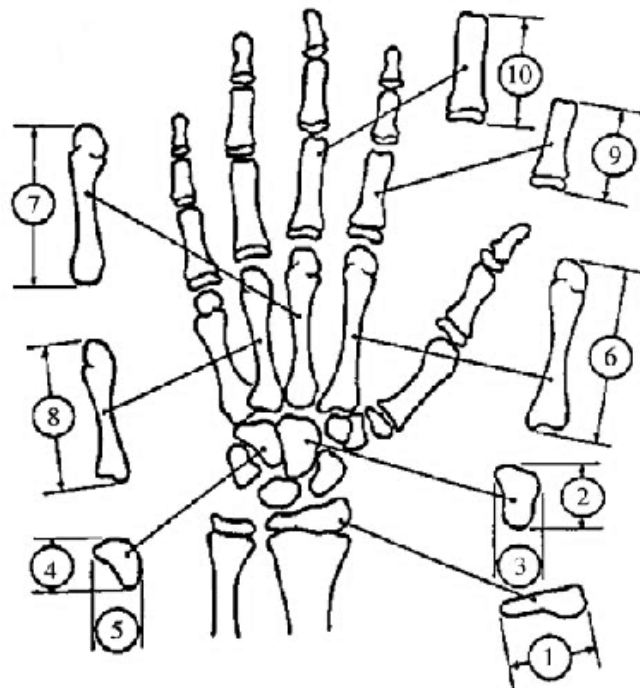
Portanto, a confiabilidade dos resultados dependerá diretamente da *expertise* do especialista (RIJN; THODBERG, 2016). Logo, é importante ter um sistema automático que apoiará o especialista, diminuindo o tempo de avaliação de cada imagem.

1.3 PROPOSTA DO TRABALHO

Os métodos convencionais utilizados por especialistas baseiam-se nos centros de ossificação das radiografias carpais no processo de avaliação da idade óssea. Várias propostas de sistemas computacionais seguiram esse mesmo procedimento, primeiro localizando e segmentando os centros de ossificação para então usar classificadores para estimar a idade óssea Kashif et al. (2016), Hsieh et al. (2017), Olivete e Rodrigues (2005). No entanto, em todos os artigos citados, ou foi realizada a localização manual dos centros de ossificação, ou há uma correção manual da localização automática dos centros de ossificação, ou há a correção manual da iluminação das imagens.

Por outro lado, verifica-se uma tendência de pesquisas recentes de estimação da idade óssea baseadas em *Deep Learning*, tais como a de Spampinato et al. (2017) e a de Lee et al. (2017). Estes estudos propõem sistemas computacionais completamente automáticos, e que usam a arquitetura clássica de *Deep Learning*, que é utilizar a imagem inteira (a imagem de raio-x inteira sem processamento) como entrada do sistema.

Figura 4 - Representação dos 10 centros de ossificação do método de ER



Fonte: Moraes et al. (2003).

A hipótese deste trabalho é que fornecer imagens pré-processadas das imagens de radiografias carpais, destacando-se os centros de ossificação das mãos (ao qual será denominada aqui de imagens mosaico), como entrada do algoritmo de Redes Neurais Convolucionais (*Convolutional Neural Networks - CNN*) baseado em *Deep Learning*, melhorará a estimativa da idade óssea em comparação com a imagem da radiografia da mão inteira sem pré-processamento. Também tem como proposta investigar a aplicação de técnicas de Processamento Digital de Imagens (PDI) (GONZALEZ; WOODS, 2010) para total automatização da etapa de localização e extração dos centros de ossificação para a criação de imagens mosaicos que serão

a entrada do algoritmo de CNN baseado em *Deep Learning* para a identificação da idade óssea.

O uso de uma base de dados própria e fechada (não disponível para *download*) impossibilita a comparação quantitativa direta dos resultados encontrados em alguns artigos. Após pesquisas, encontrou-se uma base de dados pública e anotada por especialistas da Universidade do Sul da Califórnia (GERTYCH et al., 2007), que foi a escolhida por esse trabalho. Com isso, possibilitando a comparação com os resultados obtidos em Kashif et al. (2016) e Spampinato et al. (2017).

A partir da base de dados escolhida foram criadas 4 diferentes imagens mosaico das radiografias carpais da mão esquerda (Figura 5a): apenas com as imagens dos dedos (Figura 5b); união das imagens dos dedos e do punho (Figura 5c); quatorze centros de ossificação, os mesmos utilizados pelo trabalho de Kashif et al. (2016) (Figura 5d) e por último os quatorze centros de ossificação unida com do punho (Figura 5e).

1.4 OBJETIVOS

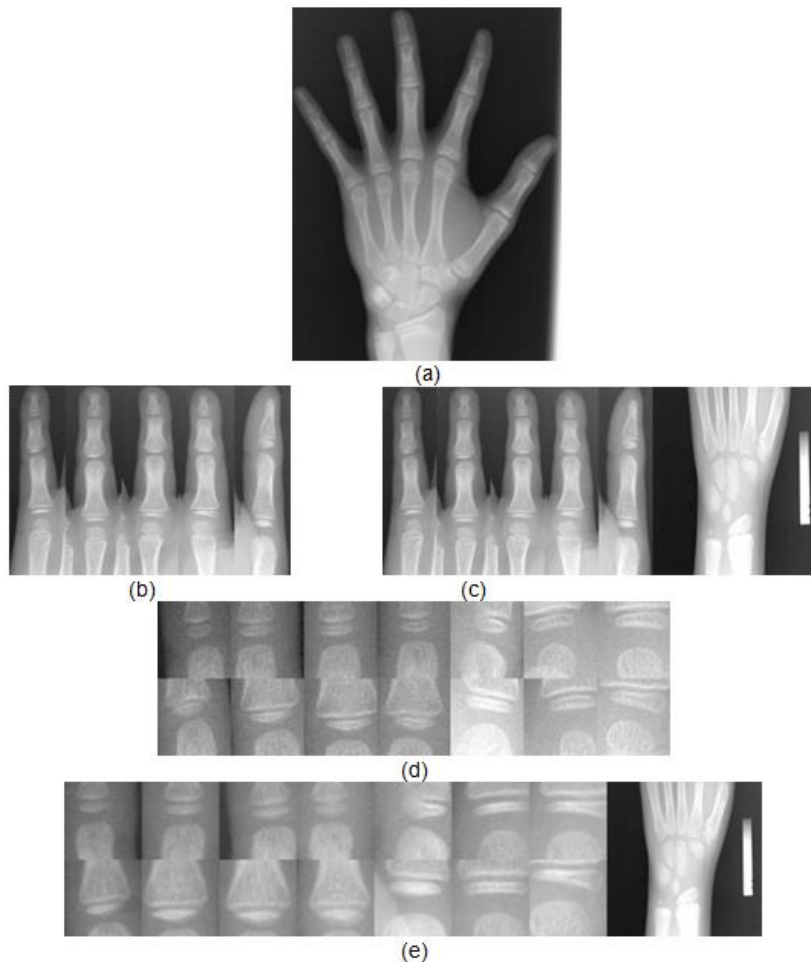
O objetivo do trabalho é a comparação dos resultados de um sistema totalmente automático para a estimacão da idade óssea através de radiografias carpais, separadas pela informacão de sexo do indivíduo, usando imagens mosaicos dos centros de ossificacão como entrada do algoritmo *Deep Learning*.

1.4.1 Objetivos Específicos

Para alcançar o objetivo principal, faz-se necessário:

- Pesquisar base de dados da radiografias carpais com anotação da estimativa da idade óssea feita por especialistas.
- Desenvolver método automático para a segmentação dos dedos.
- Desenvolver método automático para a segmentação dos centros de ossificação.
- Pesquisar e adaptar código para execução do algoritmo baseado em *Deep Learning*.
- Definir quais as métricas para comparação dos resultados.

Figura 5 - (a) Imagem inteira da radiografia carpal (b) Imagem mosaico dos dedos (c) Imagem mosaico dos dedos com o punho (d) Imagem mosaico dos centros de ossificação (e) Imagem mosaico dos centros de ossificação com o punho



Fonte: Autoria própria.

1.5 PUBLICAÇÕES

Dois trabalhos foram publicados a partir do presente trabalho:

- Um resumo com o título “Segmentação Automática das Falanges em Radiografias Carpais” na Escola Regional de Computação Aplicada à Saúde (ERCAS) 2016.
- Um artigo completo com o título “Um Estudo de Caso para a Segmentação de Falanges em Radiografias Carpais” na Escola Regional de Informática de Goiás (ERI-GO) 2016. Esse trabalho foi selecionado como melhor artigo da IV Escola Regional de Informática de Goiás.

1.6 ORGANIZAÇÃO DO TRABALHO

Para tratar do problema proposto neste capítulo, o texto desta Dissertação de Mestrado está organizado conforme descrito a seguir:

- Capítulo 2: detalha as pesquisas correlatas desenvolvidas em torno do processo de automatização da estimacão da idade óssea.
- Capítulo 3: apresenta os materiais e os métodos utilizados para o desenvolvimento do presente trabalho, que consiste nas etapas de segmentação para a extração dos centros de ossificacão e a estimacão da idade óssea com uma abordagem de *Deep Learning*.
- Capítulo 4: descreve os experimentos realizados, apresenta e discute os resultados obtidos.
- Capítulo 5: traz as consideracões finais e trabalhos futuros.

2 TRABALHOS CORRELATOS

Diversas abordagens têm sido adotadas para automatizar o processo de estimação da idade óssea. No livro publicado por Deserno (2011), um capítulo é dedicado exclusivamente sobre a segmentação baseada em região, tendo destaque as técnicas de Fuzzy e segmentação baseada em modelos. O foco do capítulo são imagens de radiografia com problemas de definição, podendo conter sombras e ruídos, que podem causar confusão em sua análise. Estes são fatores que dificultam o processo de segmentação das imagens.

Como descrição mais detalhada das pesquisas correlatas, ressaltam-se:

- Dois grupos de pesquisa: grupo de pesquisadores de Taiwan e o grupo de pesquisadores de São Paulo.
- Artigo que faz um comparativo entre 5 diferentes descritores para a extração de características.
- Um conjunto de três artigos recentes baseados em *Deep Learning*: Lee et al. (2015), Spampinato et al. (2017) e Lee et al. (2017).

2.1 GRUPO DE PESQUISA DE TAIWAN

O grupo de pesquisadores de Taiwan está há dezessete anos estudando sobre o assunto. Segue uma retrospectiva de seus artigos em ordem cronológica:

- Em Fan et al. (2001), apresenta-se uma proposta preliminar para a estimação automática da idade óssea baseada nos ossos da região do punho e contou com o pré-processamento manual das imagens radiográficas de crianças de até 7 anos, em que os ossos ainda não se encontram no estágio de sobreposição. Os resultados da classificação separadas por sexo, 114 imagens

masculinos e 47 femininas, ficou em torno dos 90% com a tolerância da variação de meio ano e de 100% para um ano.

- Hsieh et al. (2007) abordam a proposta de estimação semi-automática da idade óssea baseada em características geométricas. Foram usados um total de 909 imagens radiográficas, sendo 465 masculinas divididas em 12 grupos de dois a oito anos e, 404 femininas divididas em 14 grupos de 2,5 a 10 anos, também com o pré-processamento manual, isto é, a equalização do histograma feito manualmente nas imagens radiográficas. Chegaram a 98% de acerto no melhor caso de estimação com a variação de um ano de idade.
- Em Hsieh et al. (2008) a análise da imagem carpal foi usada para fazer o diagnóstico da síndrome de Turner. Com a técnica de regressão linear, obtiveram resultado preliminar com sensibilidade de 0,81 e a especificidade de 0,91. Depois, em Tiu et al. (2009), mostrou-se que a separação do método de triagem do artigo de 2008 em dois estágios foram satisfatórios para diagnosticar pacientes com síndrome de turner.
- Em Hsieh et al. (2010) apresenta-se um modelo de crescimento baseado em Fuzzy com seleção de análise de componentes principais para avaliação da idade óssea do carpo. Utilizou um conjunto de 550 imagens radiográficas carpais, sendo 286 masculinas e 264 femininas, com três tipos de predição, e chegou em 87% de acurácia com a tolerância de um ano e meio.
- Em Hsieh et al. (2011b) há um estudo retrospectivo da inter-relação do ângulo do carpo e a idade óssea de crianças de diferentes gerações de Taiwan e constatou que a geração de meados de 2000 apresentam o ângulo carpal maior que da geração de meados de 1960.
- O artigo de Hsieh et al. (2011a) propõe um procedimento simplificado de avaliação da idade óssea do rádio, ulna e ossos curtos usando o método agrupado de Tanner-Whitehouse. Em um conjunto de 534 imagens, a acurácia foi de 96,2% em meninos e 95% em meninas.

- Em Hsieh et al. (2011c) foi apresentada uma metodologia de segmentação rápida e totalmente automática de falange usando um algoritmo de morfologia em escala de cinza. Usaram-se 200 imagens do dedo médio de indivíduos diferentes para a segmentação, apresentando bom desempenho em questão de tempo de processamento de CPU e os resultados da segmentação foram considerados satisfatórios.
- No artigo de Hsieh, Liu e Tiu (2012) se propõe uma estimativa de crescimento com base em estatísticas de posição relativa dos ossos carpais. Os resultados indicaram que o capitato e o hamato distinguiram significativamente entre os gêneros; por outro lado, trapézio e trapezoide são diversificadas entre os sexos.
- Em Hsieh et al. (2012), apresenta-se uma proposta de segmentação automática de falanges e da região epifisária/metafisária por “Algoritmo de Melhoramento de Parâmetros Gamma”. O processamento de imagem consiste no aprimoramento da imagem inicial tendo como base a distribuição de níveis de cinza da região de contorno da falange. Dois métodos de segmentação foram aplicados separadamente em 80 imagens de indivíduos de ambos os sexos de 2 a 16 anos.
- Hsieh et al. (2017) um algoritmo hierárquico para segmentação falangeal e epifisária/metafisária, buscou automatizar o processo de localização e extração das regiões de interesse nas imagens radiográficas. Um conjunto de 100 imagens de cada sexo foi utilizado para a avaliação do resultado.

De forma geral, esse grupo de pesquisa contou com a melhoria da nitidez e com a equalização do histograma executadas manualmente pelo usuário do sistema. Assim, melhora-se manualmente a qualidade das imagens radiográficas, e com isso, alcançam-se altos índices de acurácia. A parte da segmentação automática iniciou em 2012, com uma versão melhorada em 2017, mesmo assim, manteve-se pré-processamento manual.

2.2 PESQUISA EM UNIVERSIDADES PAULISTAS

No Brasil, um grupo de pesquisadores das universidades de São Paulo vem desenvolvendo artigos na área, desde 2005:

- Olivete (2005) desenvolveu em sua dissertação uma metodologia semiautomática e simplificada para a estimação da idade óssea baseada no método de ER. Gerou um banco de dados com um conjunto de 200 imagens, sendo 100 de cada sexo, com laudos médicos respectivos. A parte semiautomática se deve ao fato de que é estimada a localização dos centros de ossificação, mas o usuário pode corrigi-las antes da avaliação da idade óssea.
- Olivete, Rodrigues e Nascimento (2005) exploraram o efeito Heel em imagens de radiografia da mão, ou seja, tentar reduzir a variabilidade na intensidade provocada pela incidência do raio-X na radiografia da mão. Houve a apresentação de uma metodologia de pré-processamento em imagens radiográficas da mão, corrigindo a iluminação não uniforme no fundo das imagens. Neste estudo foi possível comprovar a grande eficiência na correção do efeito Heel, sendo possível eliminar de forma considerável as variações de intensidade do fundo das imagens contidas no banco de dados.
- Queiroz et al. (2005) desenvolveu uma metodologia para a extração de características automáticas da mão usando a assinatura vertical. Testes foram realizadas em 50 imagens e mostraram que o método automático proposto é eficaz.
- Em Olivete e Rodrigues (2005) propôs um *software* automático para determinação da idade óssea baseado na dissertação publicada por Olivete em 2005.

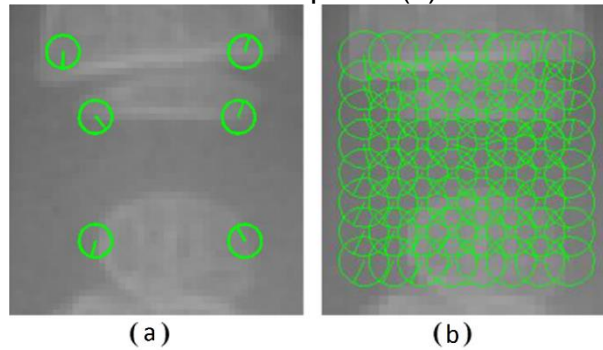
2.3 PESQUISA COMPARATIVA DE EXTRATORES DE CARACTERÍSTICAS

Kashif et al. (2016) fizeram uma comparação de 5 diferentes descritores para a extração de características no processo de estimação da idade óssea. Para isso, desenvolveram o trabalho em quatro etapas:

- Extração dos centros de ossificação: extraíram as imagens dos centros de ossificação a partir do processo manual de cliques simples de *mouse* nos centros relevantes das epífises, as imagens então foram rotacionadas e normalizadas em um tamanho de 32x32 *pixels* para a etapa posterior.
- Especificação de pontos de características: duas formas de especificação de pontos de características foram propostas. A primeira é a extração de pontos de características dispersos em que definem 6 pontos de características demarcando os limites do osso (Figura 6 (a)). A segunda é uma grade de características densa distribuída pela imagem de centro de ossificação (Figura 6 (b)).
- Descrição de características: SIFT, SURF, BRIEF, BRISK e FREAK foram os cinco descritores adotados para a extração de características. Aplicou a grade densa e dispersa para os descritores SIFT e SURF, já para o restante, apenas a grade densa foi usada.
- Classificação: usaram o classificador SVM com a terceira ordem polinomial para a classificação. Com isso, 30 classes foram criadas para 4 diferentes faixas ósseas de acordo com o grau de desenvolvimento das epífises.

Kashif et al. (2016) obtiveram a acurácia de 38.33% a 45.69%, de 81.74% a 89.37% para a tolerância de um ano e 95.64 a 98.36% com a tolerância de dois anos. No entanto, toda a parte de segmentação dos centros de ossificação é feita manualmente, pois é o usuário que indica as localizações dos centros.

Figura 6 - (a) Pontos de características disperso (b) Pontos de características denso



Fonte: Kashif et al. (2016).

2.4 PESQUISAS BASEADAS EM *DEEP LEARNING*

O *Deep Learning* ou Aprendizado Profundo usado com Redes Neurais está bem difundido atualmente (VARGAS; PAES; VASCONCELOS, 2016), sendo que a Rede Neural Convolutiva (CNN, do inglês *Convolutional Neural Network*) é uma de suas principais abordagens.

Deep Learning se tornou uma técnica famosa quando uma arquitetura CNN, ou ConvNet, ganhou o "Imagenet 2012" ou "ILSVRC 2012" (*ImageNet Large-Scale Visual Recognition Challenge*) pela primeira vez e por uma margem ampla, reduzindo a taxa de erro do estado da arte de 26.1% para 15.3% (HINTON et al., 2012). Posteriormente, alcançou resultados melhores que os métodos clássicos em muitas tarefas de visão computacional, incluindo classificação de objetos, classificação de imagens, detecção e segmentação semântica (LEE et al., 2017).

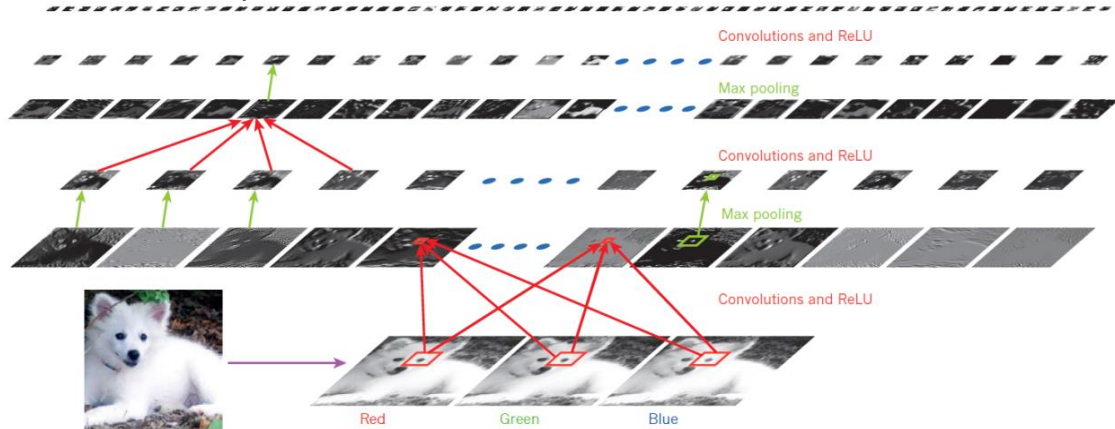
Esse método permite modelos computacionais com abstrações de alto nível de dados usando um grafo profundo com várias camadas de processamento e compostas de várias transformações lineares e não lineares. *Deep Learning* melhorou o estado da arte no reconhecimento automático de fala, visão computacional, reconhecimento de linguagem natural e em vários outros domínios (LECUN; BENGIO; HINTON, 2015).

Historicamente, o conceito de *Deep Learning* foi originado da pesquisa de RNA (Redes Neurais Artificiais), desta forma, é comum a denominação ocasional de "redes neurais de nova geração". As redes neurais *feed-forward* com camadas ocultas são, de fato, um bom exemplo para os modelos de arquitetura com diferentes níveis de aprendizagem. Essa técnica reconhece os padrões em grandes conjuntos de dados usando o algoritmo de *backpropagation*, popularizado em 1980, que aprende os pesos das redes de *Deep Learning*, alterando a representação em cada camada de representação de acordo com a camada anterior (DENG, 2014).

Uma RNA usual liga cada pixel à um neurônio. No caso do CNN, aplica convoluções para abstrair muitas dessas conexões pouco significantes passando a considerar a proximidade dos pixels. De tal forma que cada neurônio aceite como entrada uma certa subseção da camada anterior. Portanto, cada neurônio fica responsável para processar uma certa porção da imagem (RAVINDRA, 2017). A Figura 7 mostra o processo de trabalho de uma rede neural Convolutiva.

As CNN consistem em alternar camadas de convolução e de agrupamento, aprendendo de forma hierárquica a cada camada. A partir de imagens de entrada, as camadas de convolução extraem características representativas que são seguidas por camadas de classificação totalmente conectadas que são então formadas com os vetores de recursos extraídos das camadas anteriores. Assim, o aprendizado das arquiteturas que implementam *Deep Learning* busca identificar padrões nos dados, do nível mais baixo para os mais altos, exigindo uma grande quantidade de dados de treinamento rotulados para convergência estável e alcançar alta precisão de classificação (JOST, 2015).

Figura 7 - Um exemplo de CNN multicamada

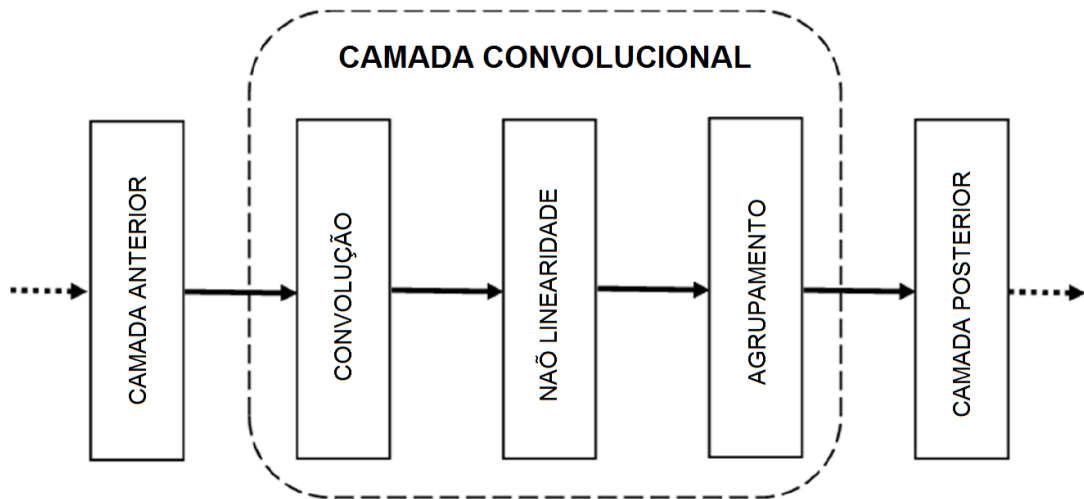


Fonte: Lecun; Bengio; Hinton (2015).

A Figura 8 apresenta os três estágios de uma típica camada convolucional. Na primeira etapa, a camada executa convoluções em paralelo para produzir um conjunto de ativações lineares. No segundo estágio, cada ativação linear é executada através de uma função de ativação não-linear, como a função de ativação linear retificada, com a finalidade de tornar os seus dados mais separáveis. Esse estágio às vezes é chamado de estágio detector. Na terceira etapa, usa-se uma função *pool* para modificar ainda mais a camada de saída (F. JÚNIOR, 2016).

Uma arquitetura de CNN é composta por essa e outras camadas como a de agrupamento (*Pooling*) que visa fazer agrupamentos para reduzir progressivamente o tamanho espacial de representação, reduzir a quantidade de parâmetros e computação na rede, assim como controlar a superposição. A camada de unidade linear retificada ou RELU (do inglês *Rectified Linear Unit*) que aplica uma função de ativação elementar com limite em zero. E a camada totalmente conectadas (*Fully Connected*) que possui conexões completas para todas as ativações da camada anterior, suas ativações podem, portanto, ser computadas com uma multiplicação de matriz seguida de um deslocamento de polarização (KARPATHY, 2016).

Figura 8 - Três estágios de uma camada convolucional



Fonte: Adaptado de F. Júnior (2016).

Na prática, poucas pessoas treinam uma rede convolutiva a partir do zero, porque é relativamente raro ter um conjunto de dados de tamanho suficiente. Em vez disso, é comum prever uma ConvNet em um conjunto de dados muito grande, por exemplo, ImageNet (RUSSAKOVSKY et al., 2015), que contém 1,2 milhão de imagens com 1000 categorias, e, em seguida, usar o ConvNet como uma inicialização ou um extrator de características fixas para o conjunto de dados de interesse. Os três principais cenários de transferência de aprendizagem são os seguintes (KARPATHY, 2016):

- ConvNet como extrator de características fixas: a partir de um ConvNet pré-treinado, remove-se a última camada *Fully Connected* e, em seguida, trata o resto do ConvNet como um extrator de características fixas para o novo conjunto de dados de interesse. Depois de extrair as características para todas as imagens, treina um classificador linear (por exemplo, SVM Linear) para esse novo conjunto de dados.
- ConvNet com ajuste fino (*Fine-tuning*): A segunda estratégia é não só substituir e reconverter o classificador em cima da ConvNet sobre um novo conjunto de dados, mas também ajustar os pesos da rede pré-treinada. É possível ajustar e afinar todas as camadas da ConvNet, como é possível manter algumas das

camadas anteriores fixadas e ajustar apenas uma parte do nível mais alto da rede.

- Modelos pré-treinados: levando em consideração que as ConvNets modernos levam 2-3 semanas para treinar em múltiplos GPUs no ImageNet, é comum ver as pessoas liberarem suas versões finais de ConvNet para o benefício de outros que podem usar essas redes para fazer os seus ajustes.

Algumas das redes pré-treinadas de alto desempenho:

- Alexnet (KRIZHEVSKY; SUTSKEVER; HINTON, 2012): o primeiro trabalho que popularizou as redes convolucionais em visão computacional, é composta por 25 camadas e requer imagens de tamanho 227x227x3 como entrada.
- GoogLeNet (SZEGEDY et al., 2015): com a contribuição no desenvolvimento de um Módulo de Iniciação que reduziu significativamente o número de parâmetros na rede. Além disso, usa o *Average Pooling* em vez das camadas *Fully Connected* na parte superior do ConvNet, eliminando uma grande quantidade de parâmetros que não parecem importar muito. É composta por 144 camadas as imagens de entrada devem ser de tamanho 224x224x3.
- VGG-19 (SIMONYAN; ZISSERMAN, 2014): a principal contribuição dessa rede foi demonstrar que a profundidade da rede é um componente crítico para um bom desempenho, é composta por 47 camadas, as imagens de entrada são do mesmo tamanho exigido pela GoogLeNet.

Dentre os os trabalhos que adotam o uso de *Deep Learning*, o sistema FingerNet (LEE et al., 2015) propõe uma abordagem para detecção de todas as 15 articulações dos dedos de uma imagem de radiografia, composta por 3 etapas:

- Pré-processamento: a imagem radiográfica passou pelo processamento de ajuste de contraste, redimensionamento, eliminação do ruído (*denoising*), detecção de bordas, binarização pelo método de Otsu e segmentação.

- Extração dos dedos: nessa etapa requer inicialmente a interação manual do usuário para indicar o final do dedo polegar e mínimo assim como os seus respectivos vales para a posterior extração dos dedos usando *watershed*.
- Extração das juntas: uma combinação de aprendizagem por CNN com a detecção de pico de junta baseada em processamento de sinal foi adotada. A arquitetura de CNN usada em FingerNet foi otimizada e customizada baseada em LeNet-5 (LECUN et al., 1998) usando uma janela que desliza pela imagem em um tamanho similar ao de entrada de LeNet-5. Picos das juntas foram detectadas por uma abordagem de processamento de sinal de imagem que analisa cada imagem de dedo da ponta para base, para encontrar picos locais em que a intensidade muda bruscamente. Então, o resultado é combinado com o resultado de CNN.

No trabalho de (LEE et al., 2015), apresenta-se uma taxa de acerto de 98,02% para um conjunto de 130 imagens de teste contendo 1.950 juntas dos dedos. A demarcação da localização dos limites do dedo polegar e mínimo é feita pelo usuário. No FingerNet não se faz a estimativa da idade óssea, chega-se apenas até à localização e extração dos centros de ossificação.

Spampinato et al. (2017) aplicou três métodos diferentes de *Deep Learning* para o processo automático de estimação da idade óssea:

- Usa CNN como um extrator de características para a posterior treinamento de um classificador mais simples como SVM.
- Usa um modelo pré-treinado para realizar a transferência de aprendizado (do inglês *Transfer Learning*) para o conjunto de base de interesse.
- Uma CNN pré-treinada em imagens de radiografia.

Nesse trabalho, fizeram a validação cruzada *5-fold* e computaram o Erro Absoluto Médio (em inglês *Mean Absolute Error* - MAE) entre as duas leituras manuais de especialistas com a sua idade estimada correspondente. Os resultados mostraram uma discrepância média entre avaliação manual e automática de cerca de 0,8 anos.

Os autores de Lee et al. (2017) também propuseram um sistema automático para a estimacão da idade óssea com uma base privada coletadas de 2005 a 2015. Usaram CNN em duas etapas do sistema:

CNN de detecção: nessa fase de pré-processamento, o CNN de detecção identificou e segmentou a mão e criou uma máscara correspondente seguida de ajustes da visão para padronizar e maximizar os recursos invariantes das imagens.

CNN de classificacão: AlexNet, GoogLeNet e VGG-16 foram os três CNN considerados nessa etapa de classificacão com o uso de *Data Augmentation* para aumentar sinteticamente o tamanho do conjunto de dados de treinamento com transformacões geométricas, transformacões fotométricas, injeções de ruído e alteracões de cor. Com isso, o VGG-16 apresentou o melhor desempenho e o AlexNet o pior.

A classificacão foi feita com radiografias de indivíduos de 5-18 anos e chegaram a acurácia de: 39.06% a 61.40% para a classificacão exata; 75.54% a 94.18% com um ano de tolerância; 90.08% a 99.00% com dois anos de tolerância.

2.5 CONSIDERAÇÕES SOBRE TRABALHOS CORRELATOS

Resumindo, no trabalho do FingerNet (LEE et al., 2015) faz-se a localizacão semiautomática dos centros de ossificacão e não se faz a estimativa da idade óssea, já no trabalho de Kashif et al. (2016), faz-se a segmentacão de forma manual que é a entrada para a etapa de classificacão automática. Nos trabalhos dos grupos de Taiwan e São Paulo, o sistema é completo, mas há sempre uma correção manual, ou da segmentacão dos centros de ossificacão ou da iluminaçao das imagens. Apenas os trabalhos de Lee et al. (2017) e de Spampinato et al. (2017) implementaram sistemas completamente automáticos.

Além disso, as bases de dados usadas pelo grupo de Taiwan, pelos pesquisadores de São Paulo, no trabalho FingerNet e pelo trabalho do Lee et al. (2017) foram diferentes e não estão disponíveis para *download* e uso livre, dificultando

comparações quantitativas. Apenas o trabalho de Kashif et al. (2016) e de Spampinato et al. (2017) usam a base de dados pública e anotada da Universidade do Sul da Califórnia (GERTYCH et al., 2007), que foi a escolhida por esse trabalho. Portanto, apenas o trabalho de Spampinato et al. (2017) serve como referência para a comparação dos resultados obtidos.

3 MATERIAIS E MÉTODOS

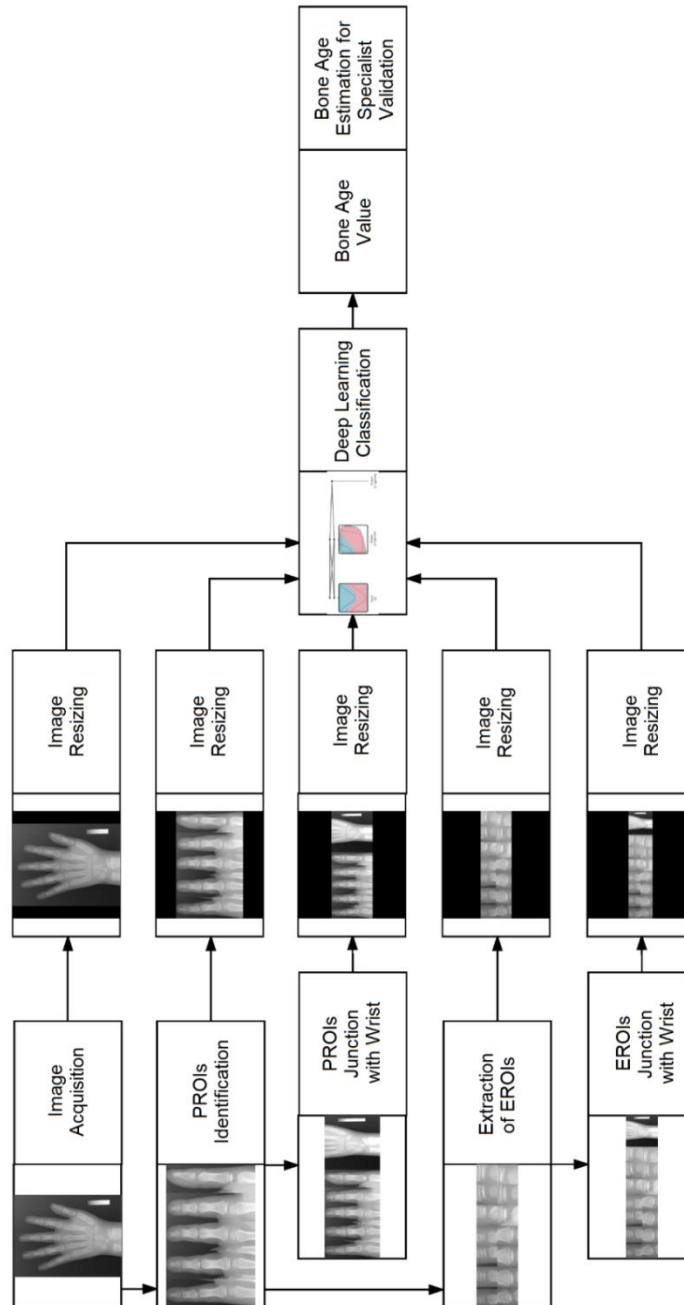
A arquitetura utilizada para a execução do trabalho, em uma visão de alto nível é constituída pelas fases de segmentação dos centros de ossificação e classificação. A Figura 9 mostra a arquitetura em blocos das etapas, representando as fases de segmentação e classificação.

A fase de segmentação é constituída pelas etapas de segmentação dos dedos, ou o recorte angular das Regiões Falangeais de Interesse (em inglês *Phalangeal bone Region Of Interests* - PROI) gerando as imagens mosaicos dos dedos e extração dos Centros de Ossificação das Epífises (em inglês *Epiphyseal Region Of Interests* - EROI) gerando as imagens mosaico dos centros de ossificação. Além disso, faz-se a segmentação dos punhos, que é unida à imagem mosaico dos dedos, gerando as imagens mosaico dos dedos e punho; e unida à imagem dos centros de ossificação gerando as imagens mosaico dos centros de ossificação e punhos. Dessa forma, criando os 4 tipos de imagem mosaico. Além das imagens mosaico, também se testa o sistema com a imagem original da radiografia carpal da mão esquerda.

A fase de classificação é realizada com a técnica de *Deep Learning* usando *Transfer Learning* de modelos pré-treinados. Todas as imagens são redimensionadas de acordo com a ConvNet, que são a GoogLeNet, a Alexnet e a VGG-19. Ao final, tem-se a estimativa da idade óssea que será validada por especialista humano.

Assim, há 10 resultados de testes, pois são 5 tipos de imagens de entrada para cada sexo. Além disso, utilizou-se a validação cruzada, que é uma ferramenta padrão para análise de classificadores e é um importante recurso para a validade a robustez do classificador. O conceito central das técnicas de validação cruzada é o particionamento do conjunto de dados em subconjuntos mutualmente exclusivos, e posteriormente, utiliza-se alguns destes subconjuntos para a estimação dos parâmetros do modelo (dados de treinamento) e o restante dos subconjuntos (dados de validação ou de teste) são empregados na validação do modelo. A validação cruzada usada nos testes foi o *k-fold*, com $k=10$, desta forma, totalizando 100 execuções dos experimentos.

Figura 9 - Arquitetura do sistema



Fonte: Autoria Própria.

3.1 BASE DE DADOS

A base de dados utilizada para o desenvolvimento das pesquisas foi obtida da Universidade do Sul da Califórnia (GERTYCH et al., 2007). O conjunto de dados

foram coletados em dois ciclos no decorrer de 10 anos, com início no ano de 1997 (CAO et al., 2000) indo até o ano de 2008, todos os indivíduos saudáveis da faixa etária de 0 a 18 anos. Com isso, foi desenvolvido uma base digital de radiografias carpais (CAO et al., 2003) composto por 1.390 imagens de indivíduos de quatro etnias (caucasiana, asiática, afro-americana e hispânica). Neste trabalho não foram realizados testes diferenciados ou separados por etnia.

Para cada uma das imagens da base de dados foram atribuídas duas leituras de referência de radiologistas experientes. No acesso à base, foram baixadas 1.380 imagens com sucesso. As 10 imagens restantes encontravam-se indisponíveis para *download*. Essas imagens radiográficas foram obtidas por diferentes formas de aquisição para imagens de 2k x 2k:

- Radiografias computadorizadas (CR): os sistemas Fuji FCR-9000 foram usados para imagens CR de resolução total (2 x 2.5K x 10 bits); ou
- *Scanner* de filme a *laser*: um *scanner* de filme 2 x 2.5K fornecido pela Abe Sekkei Inc. (Tóquio, Japão) foi usado para digitalização de filme com distância de amostragem de 100 μm e resolução de 12 bits.

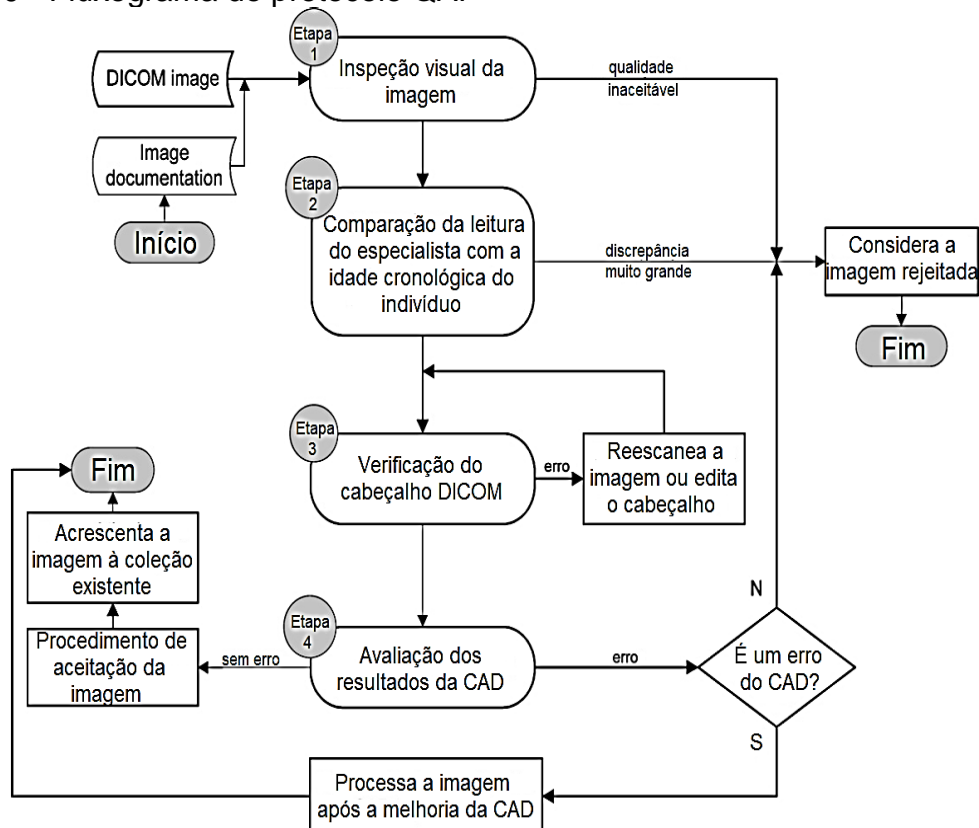
Com a finalidade do desenvolvimento de um módulo para o diagnóstico assistido por computador (CAD, do inglês *Computer Aided Design*), a base de dados da Universidade do Sul da Califórnia foi elaborada seguindo um protocolo de garantia de qualidade (*Quality Assurance Protocol - QAP*). Esse protocolo é composto por 4 etapas que abrangem a inspeção visual da imagem e a avaliação de sua adequação para o processo de CAD automático. A Figura 10 demonstra o fluxograma adotado por Gertych et al. (2007) para o desenvolvimento da base de dados:

- Etapa 1: pode ser realizada pelo operador utilizando a interface gráfica do usuário do CAD. Se a imagem da mão não estiver alinhada adequadamente em relação ao plano da imagem, ela será imediatamente rejeitada;
- Etapa 2: executado de forma não interativa, nessa etapa é realizada uma comparação entre as leituras do radiologista e a idade cronológica. A imagem é

rejeitada se a diferença entre a idade cronológica e a leitura da idade óssea do radiologista for superior a 3 anos;

- Etapa 3: também executado de forma não interativa pelo CAD, o conteúdo do cabeçalho da imagem DICOM que contém os dados demográficos do sujeito é comparado com os dados na documentação utilizada durante o procedimento de digitalização. Quaisquer inconsistências entre a documentação e o registro do paciente retornavam para correções;
- Etapa 4: são realizados os procedimentos de pré-processamento para verificar se a imagem pode revelar artefatos causados por fundo não uniforme, bordas de filme subex-postas, riscos e marcadores radiológicos que podem causar dificuldades no processamento automático de imagens. Feita essa avaliação, a imagem é submetida ao procedimento de aceitação e anexada à coleta de dados existente.

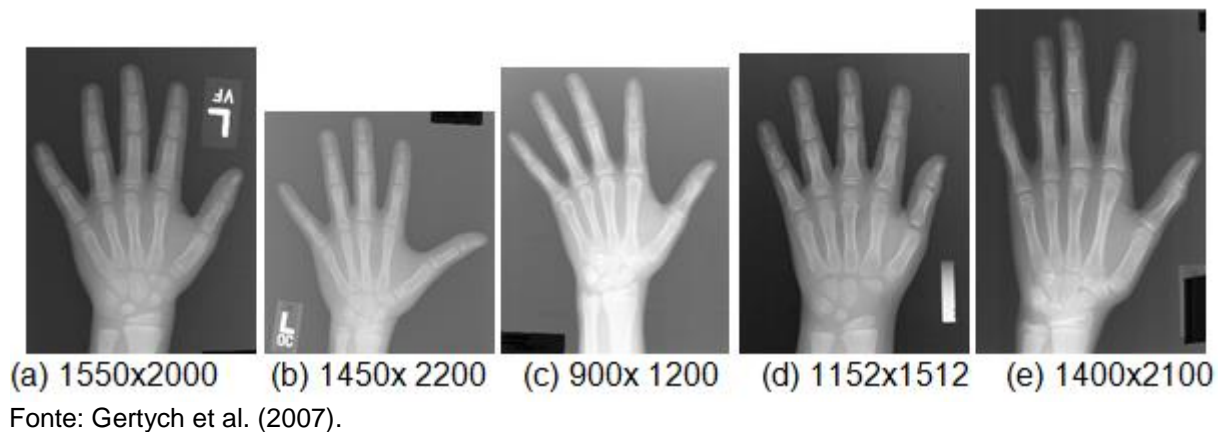
Figura 10 - Fluxograma do protocolo QAP



Fonte: Adaptado de Gertych et al. (2007).

A Figura 11 apresenta alguns exemplos de imagens padronizadas com menor incidência de ruídos nas imagens que afetam o pré-processamento das mesmas. Todas elas possuem a região da mão bem expressa dentro da imagem radiográfica, uma distribuição de intensidades em que consegue se definir bem a região de etiqueta, do tecido e dos ossos sem nenhuma sobreposição das etiquetas sobre o osso carpal.

Figura 11 - Exemplos de imagens que compõem a base de dados com as dimensões de largura x altura respectivamente

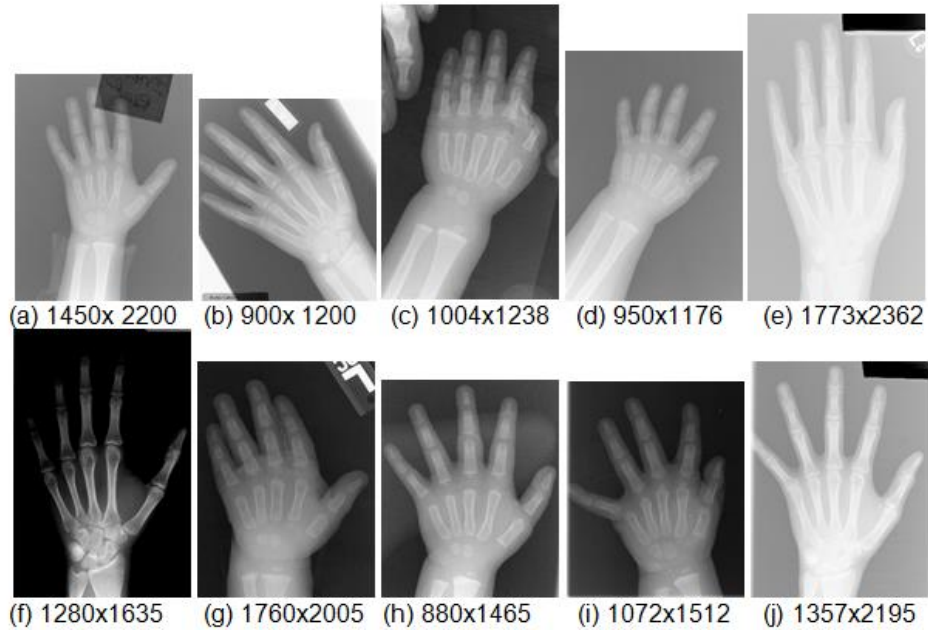


Mesmo com as medidas de padronização, a diferença de *scanners* de raio-x ao longo dos 10 anos gerou a não padronização de ruídos e também diferentes localizações das etiquetas, incidências de efeitos Heels que resultou em uma larga variação de intensidades (MENZE et al., 2011) nas imagens da base de dados. Além disso, diferenças significantes de dimensões de altura e largura das imagens foram verificadas conforme mostra a Figura 12:

- A Figura 12 (a) apresenta um bom contraste dos ossos em relação ao fundo, porém, como se pode observar, perdeu praticamente todas as falanges distais dos dedos mínimo, anelar, médio e indicador.
- A Figura 12 (b) possui os dedos mínimo e anelar juntos assim como o dedo médio colado ao indicador, que também prejudica na determinação dos pontos de referência para a segmentação.

- Pela Figura 12 (c) pode-se perceber claramente um borrão de intensidades afetando a região da mão, o que prejudica também na binarização da imagem.
- Pela Figura 12 (d) observa-se que o dedo mínimo está tangendo a lateral da imagem radiográfica, o que também afeta no pré-processamento da imagem.
- Com a Figura 12 (e) o dedo mínimo ultrapassa da imagem radiográfica da mão além de uma tira que vai do topo da imagem até o final com forte incidência de raios-x.
- Na Figura 12 (f) pode-se notar a etiqueta sobrepondo a região dos dedos da mão.
- Pela Figura 12 (g) além da imagem estar em um posicionamento ruim, tem o dedo anelar tangendo a lateral da imagem.
- Na Figura 12 (h) tem-se a dificuldade de detectar o polegar por estar em uma posição dobrada e a falange distal do mesmo é sobreposto pela falange proximal do dedo indicador.
- A Figura 12 (i) demonstra um excesso dos ossos radio e ulna que ocupam proporcionalmente e praticamente a mesma área que a mão ocupa na imagem.
- Por fim, a Figura 12 (j) mostra a forte intensidade dos raios-x sobre a imagem, dificultando a distinção dos ossos do tecido assim como do fundo. Além disso, a identificação da etiqueta que foi recortada afeta fortemente na hora de se determinar um limiar para a binarização, em que a imagem binarizada passa a ser toda branca só com a região da etiqueta preto.

Figura 12 - Exemplos de imagens pouco padronizadas que compõem a base de dados com as dimensões de largura x altura respectivamente



Fonte: Gertych et al. (2007).

O Quadro 1 apresenta a distribuição das imagens conforme a faixa etária, etnia e o sexo.

Quadro 1 - Distribuição das imagens obtidas no *download*

Distribuição das Imagens por Idade Cronológica Gênero e Raça								
Idade	Masculino				Feminino			
	A	B	C	H	A	B	C	H
0	2	5	3	4	1	4	3	1
1	5	5	5	5	5	5	4	5
2	5	5	5	5	5	5	5	5
3	5	5	5	5	5	5	5	5
4	5	5	5	5	5	5	5	5
5	9	9	10	9	7	9	7	10
6	6	7	9	9	6	8	7	10
7	7	8	9	10	7	9	8	10
8	5	10	10	10	9	10	9	9
9	7	10	7	10	7	9	8	10
10	14	15	11	12	15	12	12	14
11	15	15	13	14	12	10	13	15
12	15	15	12	15	14	14	15	15
13	15	14	12	15	15	15	13	15
14	12	14	10	14	13	12	10	14
15	10	10	9	10	10	10	10	10
16	10	10	10	10	10	10	10	10
17	10	9	10	10	10	10	10	10
18	10	10	10	10	10	10	10	10
	167	181	165	182	166	172	164	183
		695			685			
Total:				1380				

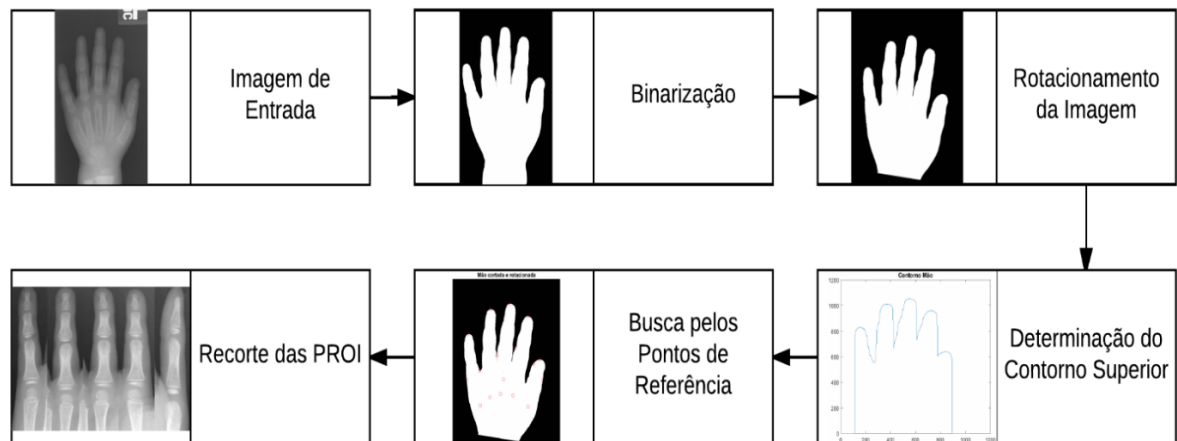
Nota: A: Asiática; B: Afro-Americana; C: Caucasiana; H: Hispânica.

Fonte: Autoria Própria.

3.2 IMAGEM MOSAICO DOS DEDOS

Para a criação das imagens mosaico dos dedos, ou o recorte angular dos PROI, foi preciso primeiramente pré-processar a imagem radiográfica para a identificação da divisão entre os dedos, ou dos pontos de referência. Após ter os pontos de divisão devidamente identificados, prossegue-se com a segmentação. A Figura 13 mostra o fluxo das etapas de recorte das PROI.

Figura 13 - Arquitetura de recorte das PROI



Fonte: Autoria Própria.

Para realizar a etapa de recorte angular dos PROI, sabe-se que existem 5 tipos de objetos em uma imagem de radiografia carpal: ossos, tecidos, fundo, Efeito Heel (BEHIELS et al., 2002) e etiqueta de identificação do indivíduo. Assim sendo, primeiramente, as imagens passaram por uma etapa de melhoria de contraste para a correção da distribuição irregular dos raios-X na imagem radiográfica (Efeito Heel). Como consequência desse efeito, algumas regiões da imagem são pouco sensibilizadas, enquanto outras são sensibilizadas em excesso, comprometendo o contraste para a obtenção de imagens de qualidade (OLIVETE; RODRIGUES; NASCIMENTO, 2005). Além disso, em muitos casos, as imagens apresentam apenas diferenças sutis entre as estruturas ósseas que se desejam segmentar e o tecido mole, como pode ser visto na Figura 14 (a). Dessa forma, o resultado final da

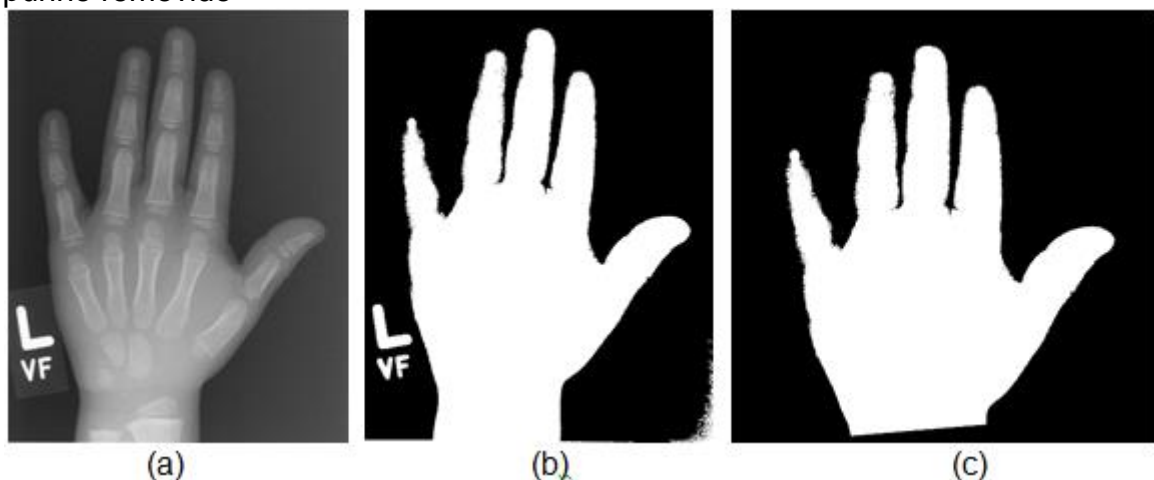
segmentação das falanges pode apresentar erros consideráveis, em relação à região de interesse.

Em seguida, determinou-se o limiar global usando o método de Otsu para a binarização e a remoção das etiquetas e/ou ruídos. Considerando que a maior região binarizada da imagem seja a da mão, e que nesse caso é a região de interesse, todas as outras regiões com menor quantidade de *pixels* foram considerados como ruídos e eliminadas da imagem, mantendo dessa forma apenas região da mão. As Figuras 14 (a) e 14 (b) mostram respectivamente a imagem original e a imagem binarizada de uma mão.

A partir da imagem binarizada, fez-se a busca pelo centroide da mão, a partir desse ponto foi realizada uma busca horizontal descendente na imagem a fim de encontrar a linha de menor quantidade de *pixels* para realizar o corte horizontal com o objetivo de eliminar o excesso de punho para as etapas posteriores.

Uma vez realizado o corte, rotaciona-se a imagem segundo a sua orientação com o objetivo de padronizar e garantir que os dedos estejam voltados para cima para facilitar a extração do melhor contorno superior. A orientação da imagem foi obtida traçando uma elipse na imagem e rotaciona-se a mesma a partir do ângulo dado para o grau zero, conforme a Figura 14 (c).

Figura 14 - (a) Imagem original (b) Imagem binarizada (c) Imagem rotacionada com o punho removido



Fonte: (a) Cao et al. (2003). (b) e (c) Autoria própria.

3.2.1 Busca pelos pontos de referência para realizar os cortes

Para determinar o contorno, o procedimento consistiu em realizar buscas na imagem de cima para baixo. A posição do contorno corresponde ao primeiro *pixel* diferente de 0 de cada uma das colunas, que foi armazenado em um vetor de posições. Esse procedimento é mais simples para traçar o contorno necessário para a localização dos pontos de referências se comparado ao método de um extrator de bordas ou soma vertical de intensidades de pontos brancos.

O resultado do extrator de bordas é descontínuo em função da imagem resultante no processo de binarização o que prejudica na busca por picos e vales a partir dele. Quanto à soma vertical de intensidades de pontos brancos tem a sua precisão afetada pelo posicionamento da mão na imagem de raio-x. Além disso, testes usando técnicas de morfologia e esqueletização foram realizados e não ajudaram para a obtenção dos pontos de referência. No entanto, adotou-se a determinação do contorno superior da mão guardando o posicionamento do primeiro *pixel* 1. A Figura 14 (c) mostra a imagem da mão cortada e rotacionada e a Figura 15 (a) o contorno obtido a partir da Figura 14 (c).

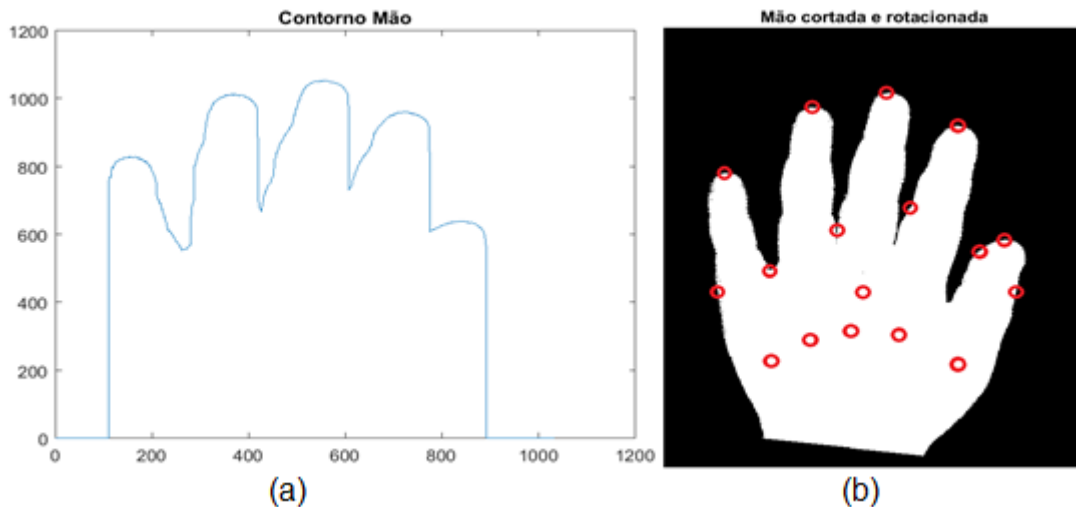
Uma vez obtido o contorno definem-se os máximos e os vales entre os dedos partindo do fato de que o ponto máximo do contorno será o dedo médio. A partir desse máximo os primeiros mínimos à direita e à esquerda foram localizados. Será um vale quando o valor mínimo estiver localizado em uma posição correspondente a 15 *pixels* abaixo das posições posteriores do contorno. Sendo essa busca realizada da esquerda para direita para encontrar o vale direito e da direita para esquerda para encontrar o vale esquerdo do dedo médio.

A partir do vale direito, faz-se a busca do máximo pelo lado direito para obter o ponto máximo do dedo indicador e a partir do vale esquerdo faz-se o mesmo para obter o máximo do dedo anelar. Com a mesma lógica, aplica-se para obter o vale direito do dedo indicador a partir do seu máximo e fazendo a busca à direita, e para o vale esquerdo do dedo anelar, busca-se da direita para esquerda. Desses dois novos

vales obtidos, buscam-se pelos máximos do dedo mínimo e polegar. Os resultados dessas buscas são mostrados na imagem da Figura 15 (b).

Pode-se observar na Figura 15 (b) que existem mais dois pontos demarcando os limites do dedo polegar e mínimo, esses pontos podem ser identificados se for traçado uma linha horizontal passando pelo centroide. Eles foram obtidos buscando pelo primeiro *pixel* diferente de zero no sentido das laterais da imagem para o meio tendo como referência o centroide da mão.

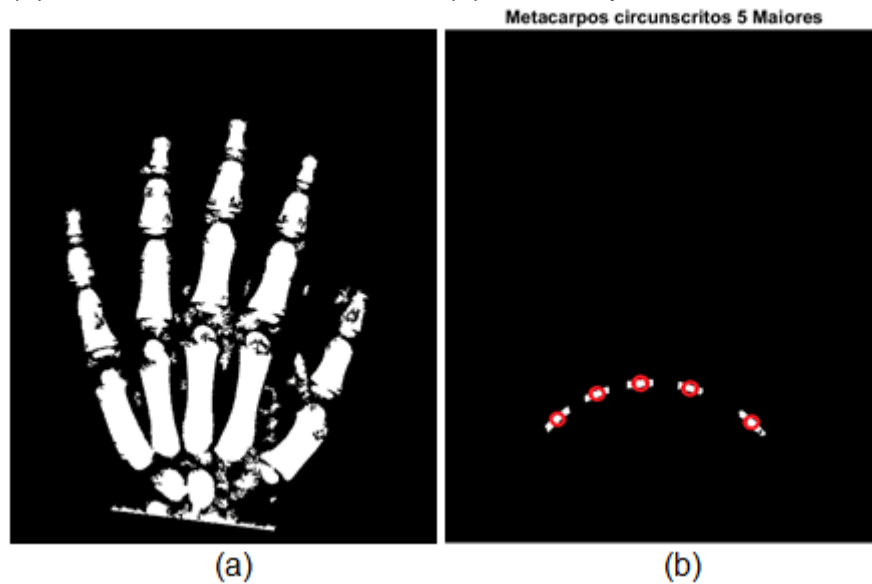
Figura 15 - (a) Contorno superior da mão (b) Mão com os pontos de referência de corte determinadas



Fonte: Autoria própria.

Além disso, foi preciso fazer uma nova segmentação na imagem da mão para obter a localização dos cinco metacarpos com a finalidade de determinar a orientação de cada um dos dedos. Para isso, foi adotada como referência a propriedade geométrica utilizada por (HSIEH et al., 2017), que determina o centro do punho como o ponto de referência para traçar uma circunferência com o raio de um terço da distância desse ponto até o máximo do dedo médio. As Figuras 16 (a) e 16 (b) ilustram a imagem dos ossos da mão binarizada e os ossos dos metacarpos encontrados na circunferência traçada. Os centroides dos metacarpos obtidos nessa etapa também podem ser verificados na Figura 15b.

Figura 16 - (a) Ossos da mão binarizadas (b) Metacarpos obtidos da circunferência



Fonte: Autoria própria.

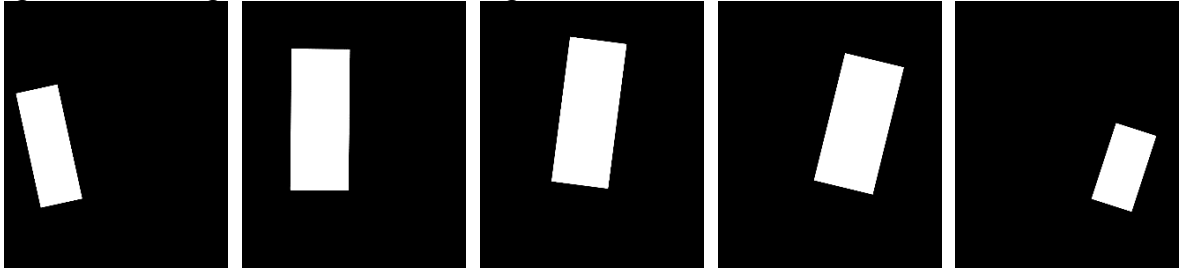
3.2.2 Recorte dos PROI

Para recortar cada um dos dedos foi usada como referência para largura a distância entre dois vales vizinhos que delimitam cada dedo, ou entre um vale e o limite da mão no caso do dedo mínimo e polegar. E como altura a distância entre o centroide de cada metacarpo (identificados pelos círculos vermelhos localizados sobre o raio da circunferência) e o ponto máximo de cada dedo respectivamente.

Assim sendo, foi definido um retângulo de acordo com a orientação angular de cada PROI. Posteriormente, foi gerada uma imagem auxiliar desse retângulo obtido do tamanho da imagem da mão. Essa imagem auxiliar serve para rotacionar a imagem original e obter o recorte do dedo. As Figuras 17, 18 e 19 demonstram a imagem auxiliar do retângulo, a imagem auxiliar rotacionada e a imagem da mão rotacionada sobre a mesma orientação.

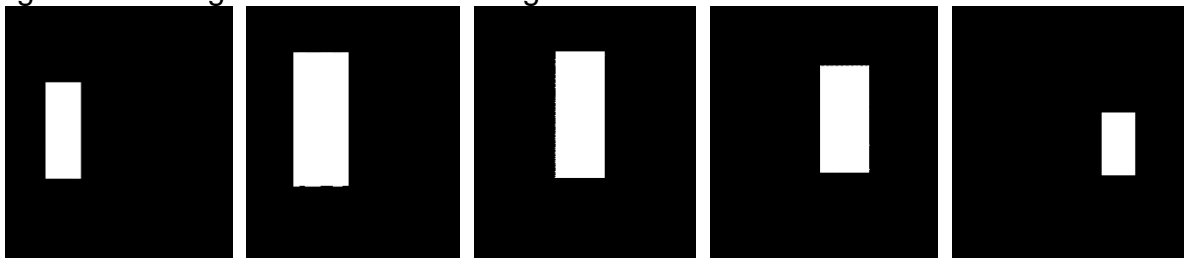
Com cada uma das imagens auxiliares rotacionadas, faz-se o recorte de cada um dos dedos segundo os limites do retângulo da imagem auxiliar. Os resultados do recorte de cada um dos dedos podem ser observados na Figura 20.

Figura 17 - Imagem auxiliar do retângulo de cada um dos dedos



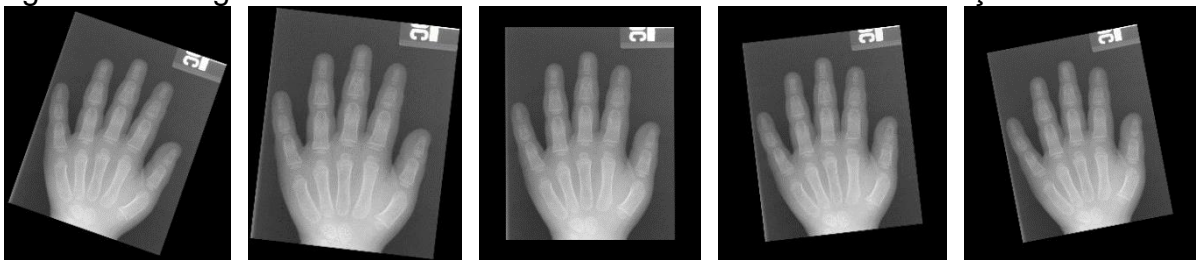
Fonte: Autoria própria.

Figura 18 - Imagem auxiliar dos retângulos rotacionados



Fonte: Autoria própria.

Figura 19 - Imagem da mão rotacionada de acordo com cada orientação

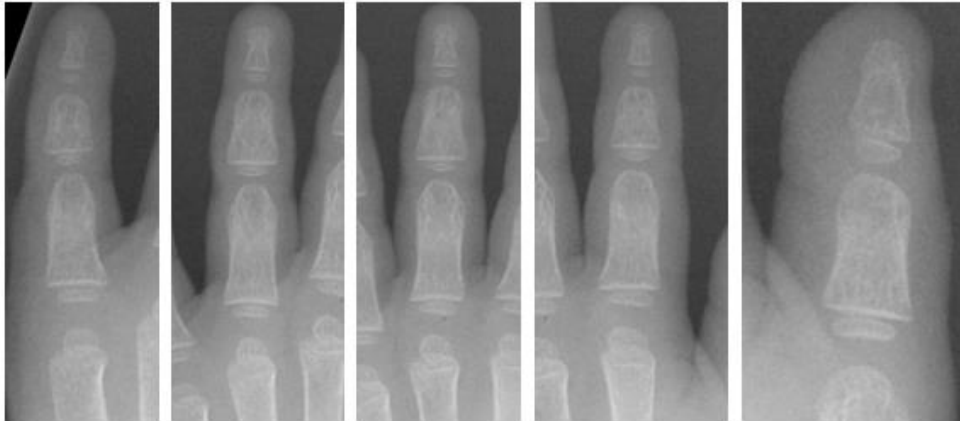


Fonte: Autoria própria.

Nessa etapa de recorte angular das PROI foi adaptada a propriedade geométrica utilizada no trabalho de Hsieh et al. (2017) para determinar os limites inferiores de cada um dos dedos e também para a remoção do excesso de punho. Além disso, foi utilizada a ideia de buscas pelos pontos de referência superiores a partir do trabalho de Castro et al. (2009). A ideia de rotacionamento da imagem para traçar o contorno superior da mão, assim como de delimitar um retângulo auxiliar para a rotação e o recorte de cada um dos dedos foram do presente trabalho.

Caso o processo de segmentação dos dedos não encontre cinco regiões, a imagem do raio-x é automaticamente descartada e não comporá a base de dados de treinamento e testes.

Figura 20 - Imagem da região dos PROI recortados

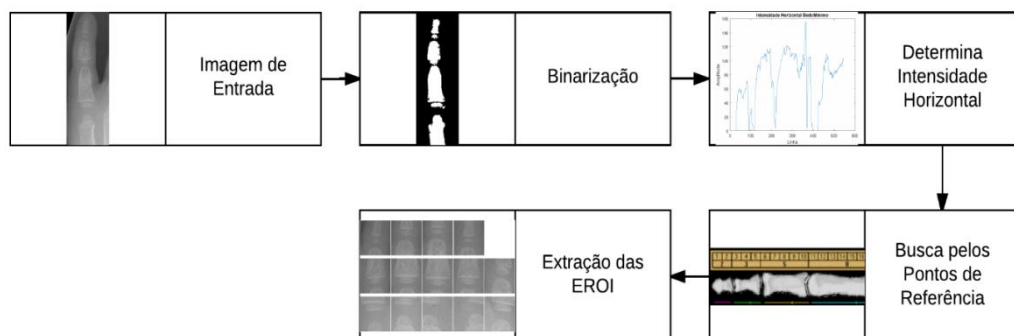


Fonte: Autoria própria.

3.3 IMAGEM MOSAICO DOS CENTROS DE OSSIFICAÇÃO

Na etapa de extração das EROI, centros de ossificação, fez-se novamente a binarização da imagem do dedo de interesse para a posterior detecção e extração das mesmas. A Figura 21 mostra o fluxo de etapas da extração das EROI.

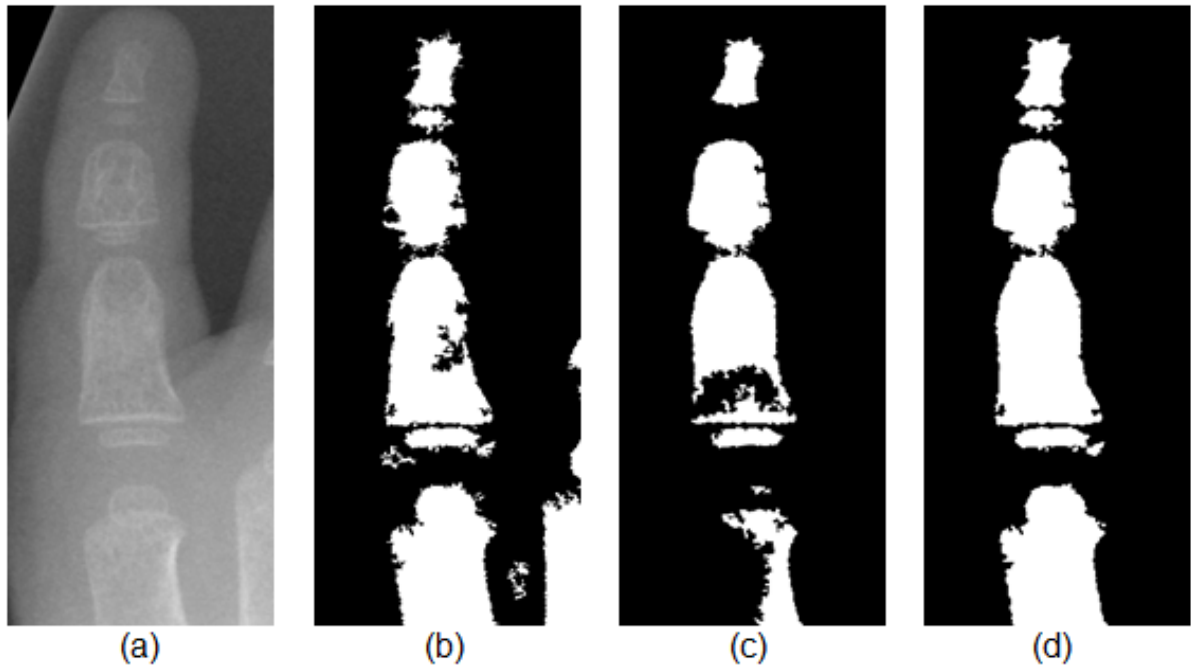
Figura 21 - Arquitetura de extração das EROI



Fonte: Autoria Própria.

Para essa etapa de segmentação dos centros de ossificação, usaram-se como entrada as imagens recortadas dos dedos juntamente com o mesmo recorte do dedo respectivo da Figura 18, que serviu como complemento para a binarização da imagem do dedo. As Figuras 22 (a) e 22 (b) mostram o recorte realizado da imagem original e realizado da imagem dos ossos da mão binarizadas respectivamente, seguindo o mesmo procedimento de rotação e extração das imagens da região dos dedos. A Figura 22 (c) representa a nova binarização realizada a partir do dedo cortado e a Figura 22 (d) o resultado da soma das duas binarizações já sem ruídos. Para a determinação dos ruídos na imagem do dedo binarizado, foi traçado uma linha mediana na imagem e toda região de *pixels* brancos que não tangenciam essa linha foi considerada como ruído e eliminado.

Figura 22 - (a) Dedo original (b) Dedo recortado da figura 18 (c) Binarização do dedo original (d) Resultado da soma das duas binarizações sem ruídos



Fonte: Autoria Própria.

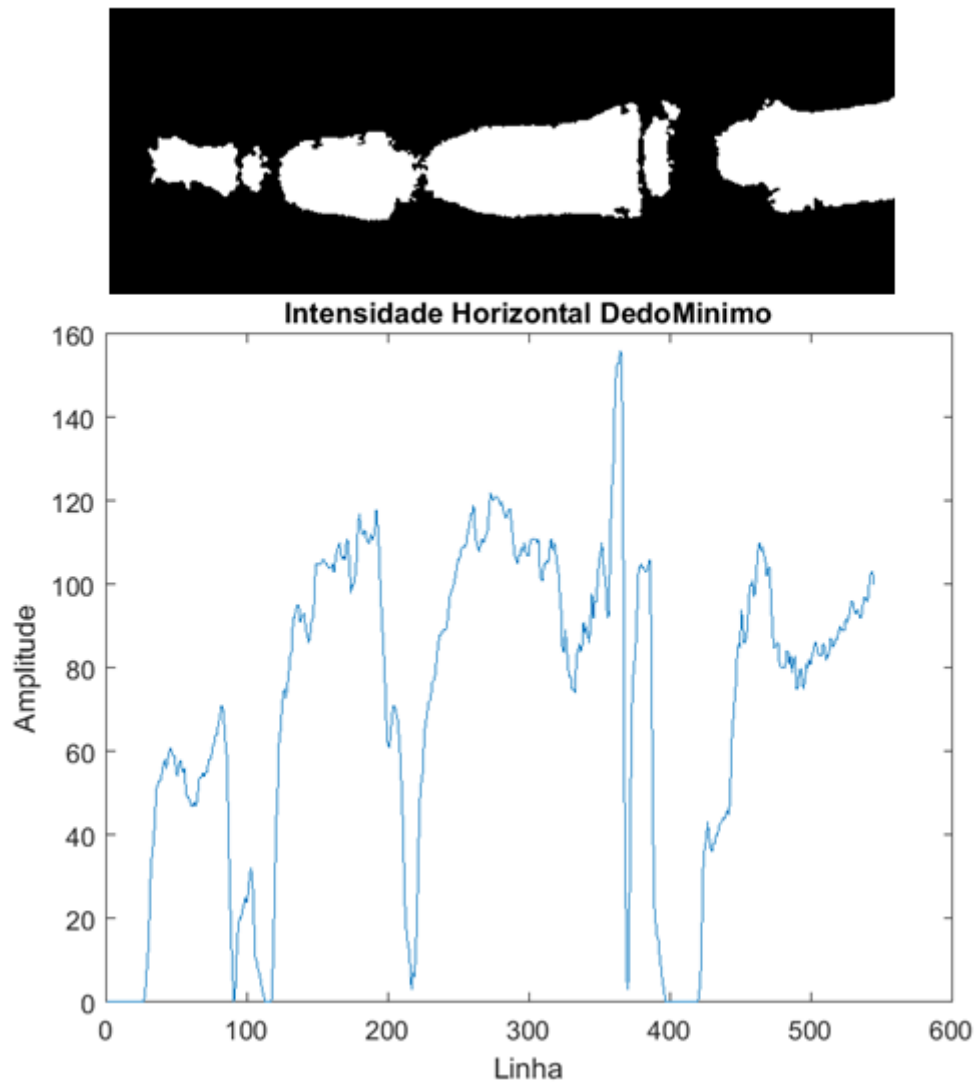
3.3.1 Localização e extração das EROI

Após a binarização da imagem do dedo, foi preciso determinar a localização dos centros de ossificação para possibilitar a realização da extração dos mesmos. Para tanto, fez-se a soma de *pixels* brancos em cada linha da imagem binarizada. Ou seja, é como se girasse a imagem a -90° em que ela ficasse conforme a parte superior da Figura 23 e fizesse a soma da quantidade de *pixels* brancos em cada uma das colunas. Este conjunto de somas, gera um vetor de inteiros que pode ser plotada num gráfico como da parte inferior da Figura 23.

Em Olivete et al. (2012) verificou-se a presença da proporção áurea para o processo de estimação da idade óssea. A Figura 24 ilustra as quatro metodologias adotadas pelo artigo para a análise das razões existentes entre os ossos carpais:

- N-ouro-1: nas falanges e metacarpos do terceiro dedo.
- N-ouro-2: entre o comprimento da terceira falange proximal mais a terceira falange distal com o comprimento do terceiro metacarpo.
- N-ouro-3: soma da terceira falange proximal mais a terceira falange média e a soma da terceira falange média mais a terceira falange distal.
- N-ouro-4: existência da sequência de Fibonacci nos ossos da mão.

Figura 23 - Intensidade horizontal do dedo mínimo binarizado

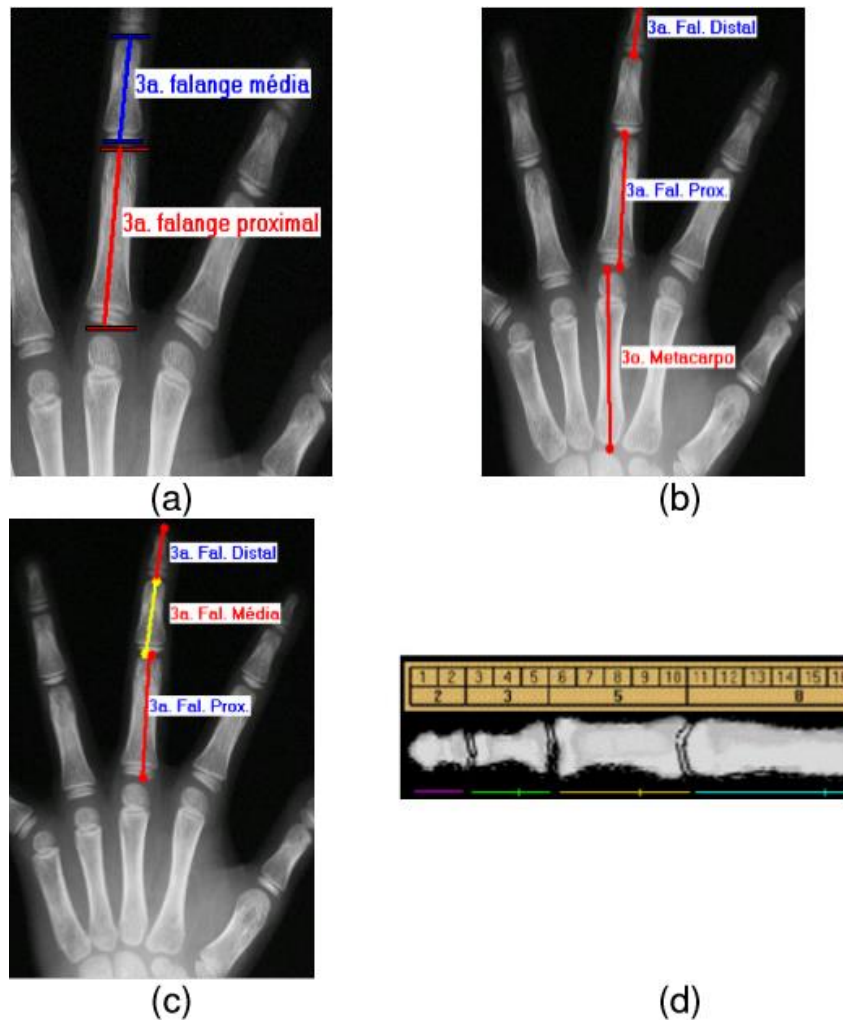


Fonte: Autoria Própria.

Assim sendo, o N-ouro-4 é obtido através do cálculo da existência da sequência de Fibonacci em que o tamanho da falange proximal é estimada através da soma entre os dois ossos anteriores, a falange média e distal, respectivamente, conforme a Figura 24 (d). Apesar da sequência de Fibonacci não ter apresentado o melhor desempenho no trabalho de Olivete et al. (2012), teve um desvio de até 12mm em comparação às outras abordagens, ela chegou a 90,3% de concordância no melhor caso. Assim sendo, adotou-se esse número ouro como referência para auxiliar na busca das regiões prováveis dos centros de ossificação de cada um dos dedos. E do valor encontrado segundo essa proporção da sequência de Fibonacci, é determinada uma janela de tamanho 100, ou seja, 50 posições à direita e à

esquerda, e a partir desses 100 valores, extrai-se o mínimo local para ser adotado como o ponto central do centro de ossificação.

Figura 24 - Metodologias das proporções áureas. (a) N-ouro-1 (b) N-ouro-2 (c) N-ouro-3 (d) N-ouro-4

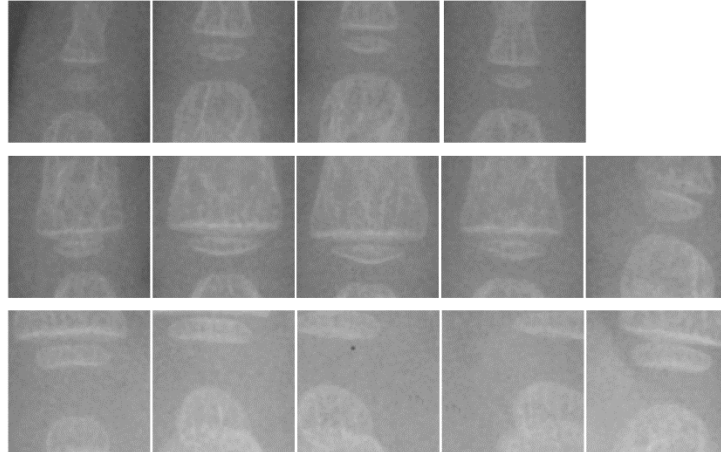


Fonte: Olivete et al. (2012).

A Figura 25 demonstra as 14 imagens resultantes dos centros de ossificação da região das falanges da mão esquerda. Foi adotado o tamanho de 99x99 para a obtenção de cada região do centro de ossificação. Esse tamanho de imagem foi escolhida, pois, precisava determinar e padronizar uma dimensão para delimitar todos os centros de ossificação da base, e, em vez de estipular esse valor

empiricamente, usou-se o tamanho da máscara adotada no trabalho de Hsieh et al. (2012) que delimitou bem todos os centros de ossificação do presente trabalho.

Figura 25 - Os 14 centros de ossificação extraídos



Fonte: Autoria Própria.

Na etapa de extração das EROI combinou-se a abordagem de buscar mínimos locais de acordo com a intensidade de *pixels* brancos com a proporção áurea da sequência de Fibonacci abordado no trabalho de Olivete et al. (2012) para aumentar a precisão na identificação das EROI. E a máscara de tamanho 99x99 adotada no trabalho de Hsieh et al. (2012) para a extração de cada um dos centros de ossificação. Os demais detalhes foram decisões adotadas pelo presente trabalho.

Caso o processo de segmentação dos centros de ossificação não encontrasse catorze regiões, a imagem do raio-x é automaticamente descartada e não comporá a base de dados de treinamento e testes.

3.4 IMAGEM MOSAICO COM PUNHO

A região do punho da mão possui um conjunto significativo de ossos carpais que expressam bem o desenvolvimento ósseo de um indivíduo e adotados nos métodos

clássicos para a estimativa da idade óssea. Por essa razão, essa parte do raio-x foi incorporada no mosaico dos dedos e dos centros de ossificação.

A segmentação do punho da mão foi realizada a partir do centroide obtido na imagem binarizada da mão (Figura 14 (b)) até o final da imagem carpal. E depois foi unida às imagens de mosaico de dedos e dos centros de ossificação para compor um novo conjunto de imagens para a classificação. Espera-se, que, com um conjunto maior dos ossos carpais, possa aumentar o desempenho de aprendizado das CNN. A Figura 26 mostra o exemplo de uma imagem de punho segmentada unida às imagens de mosaico.

Figura 26 - Exemplo de uma imagem da região do punho segmentada para somar à composição dos mosaicos



Fonte: Autoria Própria.

3.5 CLASSIFICAÇÃO COM *DEEP LEARNING*

Para a etapa da classificação adotou-se o uso de *Deep Learning* no método de transferência de aprendizado (*Transfer Learning*). Essa escolha foi realizada por não possuir uma base com quantidades significantes de imagens para se treinar uma CNN própria. Assim sendo, consideraram-se três CNN bastante conhecidas e

utilizadas por outros trabalhos para a realização dessa etapa: a GoogLeNet, a Alexnet e a VGG-19.

Para fins de comparação, as imagens de radiografia da mão como as imagens obtidas nas duas etapas de segmentação foram adotadas como entrada para da classificação.

E como já abordado anteriormente, a distinção do sexo do indivíduo é relevante em função dos diferentes surtos de crescimento que ocorrem entre meninos e meninas. Dessa forma, 5 tipos de imagens separados pelo sexo foram utilizados como entradas para a classificação, ou seja, para cada tipo de imagem foram gerados 2 bases, uma do sexo masculino e outra feminino e cada uma das bases foram colocado como entrada para o treinamento e classificação com *Deep Learning*:

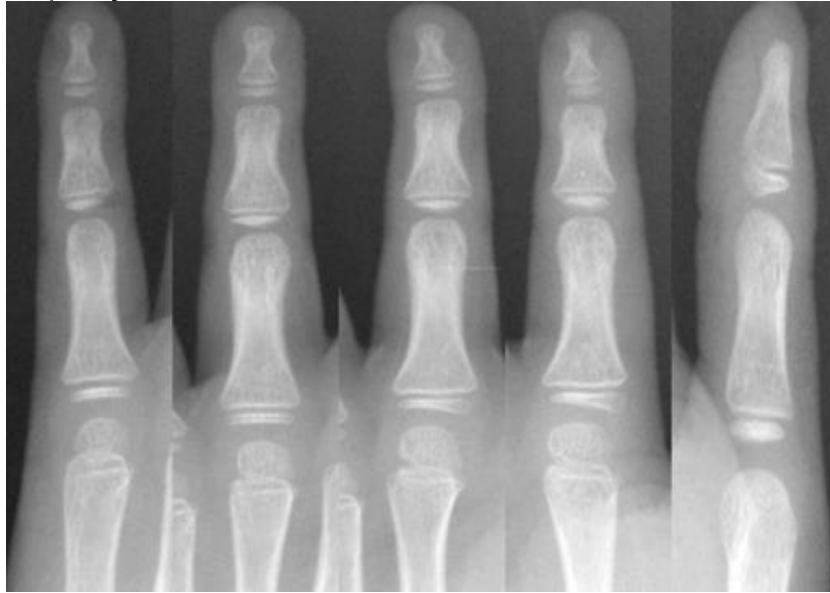
- A imagem radiográfica da mão sem nenhum pré-processamento conforme a Figura 27.
- A composição das imagens dos 5 dedos de uma mão como a Figura 28.

Figura 27 - Imagem radiográfica da mão sem pré-processamento



Fonte: Gertych et al. (2007).

Figura 28 - Composição das PROI



Fonte: Autoria Própria.

- O mosaico com as imagens dos 14 centros de ossificação em 2 linhas e 7 colunas segundo a Figura 29.
- A composição de imagens dos cinco dedos com o punho como a Figura 30.
- O mosaico com imagens dos 14 centros de ossificação em 2 linhas e 7 colunas segundo a Figura 31.

Figura 29 - Mosaico composto pelos 14 centros de ossificação de uma imagem radiográfica



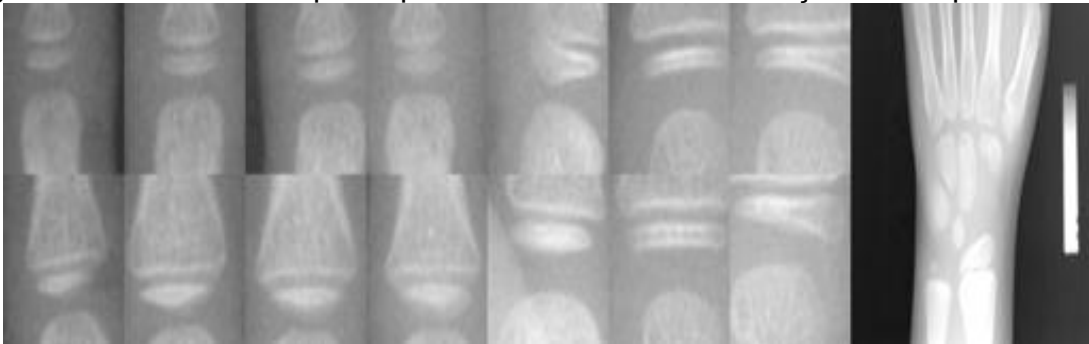
Fonte: Autoria Própria.

Figura 30 - Composição das PROI com o punho



Fonte: Autoria Própria.

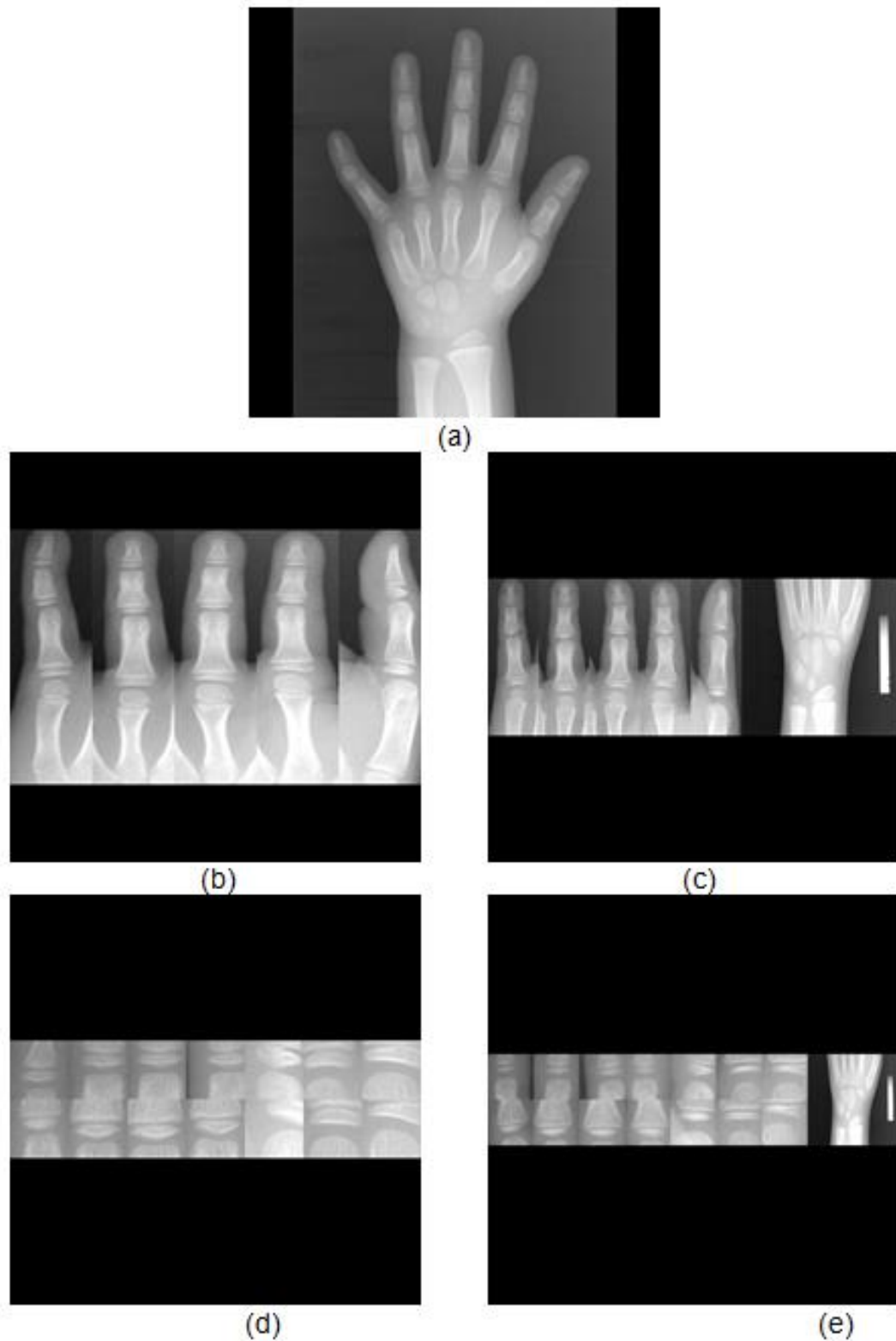
Figura 31 - Mosaico composto pelos 14 centros de ossificação com o punho



Fonte: Autoria Própria.

Para fins de normalização das imagens para servirem como entrada das CNN, foi adotado o padrão de tamanho 512x512 conforme Lee et al. (2017) e apresentada na Figura 32. As imagens foram redimensionadas com *zero-padding* para preservar as suas proporções de aspecto, como mostrado nas Figuras 27 a 31. Feito isso, as imagens são redimensionadas em lote no carregamento para o processo da classificação para o tamanho específico de cada CNN. E foi realizada a validação cruzada no método *10-fold* para estimar a precisão na classificação em cada um dos conjuntos de imagens.

Figura 32 - (a) Imagem redimensionada da mão esquerda (b) Imagem mosaico redimensionada da composição dos dedos (c) Imagem Mosaico redimensionada da composição dos dedos com o punho (d) imagem mosaico redimensionada dos centros de ossificação (e) Imagem mosaico redimensionada dos centros de ossificação com o punho



Fonte: Autoria própria.

3.6 CRITÉRIOS DE AVALIAÇÃO

Além da acurácia percentual dos acertos dos resultados obtidos na classificação, 4 outras métricas foram adotadas para estimar o erro dos valores preditos (faixa etária atribuída) com os valores de referência (idade óssea de referência correspondente) (HYNDMAN; ATHANASOPOULOS, 2014). Essas métricas foram calculadas para cada par de valores dentro de cada *fold*:

- *Mean Absolute Error (MAE)*:

$$MAE = \sum_{i=1}^n \frac{|r_i - p_i|}{n} \quad (1)$$

- *Mean Squared Error (MSE)*:

$$MSE = \frac{1}{n} \sum_{i=1}^n (r_i - p_i)^2 \quad (2)$$

- *Root Mean Square Error (RMSE)*:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (r_i - p_i)^2} \quad (3)$$

- *Mean Absolute Percentage Error (MAPE)*:

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{r_i - p_i}{r_i} \right| \quad (4)$$

onde r é o valor da idade óssea, p o valor predito e n a quantidade de imagens do *fold* equivalente. O valor final da métrica foi definido pela média dos valores das métricas obtidas nos 10-*fold* para cada conjunto de imagens. Quanto menor o valor dessas métricas, melhor o desempenho obtido.

O desempenho de cada uma das CNN foi avaliado de 3 formas: exata; com a tolerância da variação de 1 ano de idade; e com a tolerância da variação de 2 anos de idade. Ou seja, para um mesmo conjunto de dados, cada uma das métricas foi calculado na sua forma precisa, com tolerância de 1 ano e com tolerância de 2 anos. Além disso, ela foi repetida duas vezes, uma vez para o grupo de 0-18 anos e outra vez para o grupo de 5-18 anos para fins de comparação com os resultados dos trabalhos correlatos.

Dessa forma, as tabelas são apresentadas com os resultados dos 2 conjuntos de imagens: as 5 linhas superiores de 0-18 anos e as 5 linhas inferiores expressam valores correspondentes ao conjunto de indivíduos de 5-18 anos.

As CNN são colocadas lado a lado como colunas, sendo que para cada rede são projetadas 3 colunas: primeira sem tolerância, segunda com a tolerância de 1 ano e terceira com a tolerância de 2 anos.

A classificação para os indivíduos de 5-18 anos foi feita para possibilitar a comparação com o trabalho de Lee et al. (2017), que adotou essa faixa etária, pois em sua base tinha poucos casos de crianças de 0-4 anos. Este fato é comum, pois evita-se submeter crianças dessa faixa etária à radiação. Além disso, para essa faixa etária as imagens radiográficas costumam ter uma variação maior quanto ao posicionamento da mão, afetando a qualidade da imagem.

4 EXPERIMENTOS, RESULTADOS E DISCUSSÃO

Este capítulo descreve os experimentos realizados com a segmentação e classificação, de acordo com a descrição do Capítulo 3. Além da imagem da mão sem pré-processamento, cujo resultado será mostrado na seção 4.2, a etapa de classificação utilizou quatro outros conjuntos de imagens resultantes da etapa de segmentação. As seções 4.3 e 4.4 descrevem os resultados obtidos na classificação das imagens mosaico dos dedos, e das imagens de mosaico constituídas pelos centros de ossificação. As seções 4.5 e 4.6 descrevem, respectivamente, os resultados da classificação das imagens mosaico dos dedos, e do mosaico constituído pelos centros de ossificação acrescidas da imagem da região do punho.

Como já destacado anteriormente, os estágios de crescimento ocorrem de forma diferente em indivíduos do sexo masculino e feminino, deste modo, os conjuntos de imagens foram separadas para cada sexo, masculino e feminino. O Quadro 2 apresenta a distribuição das imagens do Quadro 1 de acordo com as suas idades ósseas num total de 1380 imagens, sendo 695 de indivíduos do sexo masculino e 685 de indivíduos do sexo feminino.

Quadro 2 - Distribuição das imagens da mão por idade óssea

Distribuição por Idade Óssea		
Idade	Masculino	Feminino
0	14	7
1	20	15
2	19	21
3	18	33
4	25	24
5	32	33
6	38	41
7	35	31
8	28	35
9	35	15
10	41	43
11	54	35
12	42	39
13	65	52
14	42	32
15	30	75
16	31	54
17	51	78
18	75	22
Subtotal:	695	685
Total:	1380	

Fonte: Autoria Própria.

Para a definição da idade óssea, foi levado em consideração o valor médio das idades ósseas feitas pelas leituras dos dois especialistas referente à cada uma das imagens como foi adotado no trabalho de Kashif et al. (2016). Considerou-se cada uma das idades como uma classe. A base de dados não é balanceada, tendo-se classe com apenas 7 imagens e outra com 78 imagens.

4.1 RESULTADOS DA SEGMENTAÇÃO

O sistema automaticamente retira da base de dados as imagens cuja segmentação não encontraram 5 dedos, isto é, não encontraram pontos de referência o suficiente para o recorte de 5 regiões. Assim, após o processo de segmentação das imagens dos dedos, a quantidade total de imagens da base diminuiu para 991 imagens segmentadas, sendo 486 imagens de indivíduos do sexo masculino e 505 imagens de indivíduos do sexo feminino. Havendo uma redução de 28.18% do total de imagens, sendo 30.07% do conjunto masculino e 26.27% no conjunto feminino. A distribuição das imagens de acordo com a idade óssea após a etapa de recorte dos dedos está apresentada no Quadro 3.

Quadro 3 - Distribuição das imagens de dedos por idade óssea

Distribuição por Idade Óssea		
Idade	Masculino	Feminino
0	8	4
1	12	12
2	11	17
3	13	22
4	20	19
5	20	22
6	24	30
7	29	26
8	21	29
9	26	14
10	30	34
11	42	25
12	26	23
13	43	37
14	29	23
15	19	47
16	22	43
17	35	61
18	56	17
Subtotal:	486	505
Total:	991	

Fonte: Autoria Própria.

Esta diminuição do total de imagens ocorreu devido à dificuldade em padronizar um método para o pré-processamento das imagens com as variações de intensidades apresentadas pelas imagens, mesmo sendo calculado um limiar individualmente para cada imagem, afetando todas as etapas posteriores.

Na fase da segmentação dos centros de ossificação da mão utilizou-se como entrada o recorte dos dedos, com a quantidade de imagens inicial descrito no Quadro 3. De novo, o sistema automaticamente descarta as imagens nas quais não se encontrou 14 regiões de centros de ossificação. Com isso, o total de imagens diminuiu de 991 para 945 imagens, sendo 461 do sexo masculino e 484 do feminino. Nesta etapa de validação, 4.64% das imagens foram descartadas, sendo 5.14% do sexo masculino e 4.15% do sexo feminino. O Quadro 4 apresenta a distribuição do conjunto das imagens validadas resultantes.

Quadro 4 - Distribuição das imagens de mosaico dos centros de ossificação por idade óssea

Distribuição por Idade Óssea		
Idade	Masculino	Feminino
0	8	4
1	12	10
2	11	16
3	13	22
4	20	19
5	20	22
6	23	29
7	29	23
8	18	28
9	25	13
10	28	31
11	38	23
12	23	21
13	40	36
14	27	21
15	17	47
16	21	43
17	34	61
18	54	15
Subtotal:	461	484
Total:	945	

Fonte: Autoria Própria.

As imagens da composição do mosaico de dedos com o punho possuem o mesmo conjunto de imagens da distribuição da etapa de segmentação dos dedos. Assim

como as imagens da composição do mosaico dos centros de ossificação com o punho possuem a mesma quantidade de imagens conforme o conjunto das imagens de mosaico dos centros de ossificação.

4.2 RESULTADOS DA ESTIMAÇÃO DA IDADE ÓSSEA DAS IMAGENS SEM PROCESSAMENTO

A classificação baseada no método de *Deep Learning* foi realizada no conjunto de imagens originais sem nenhum pré-processamento para servir de comparação com os resultados das imagens obtidas no processo da segmentação. Como já descrito anteriormente, este método de classificação exige uma grande quantidade de imagens para o bom desempenho da classificação. Os resultados obtidos em um conjunto de imagens sem aumento de dados (*Data Augmentation*) foram considerados boas dentro de suas limitações.

O Quadro 5 apresenta os resultados da classificação do conjunto de imagens da mão do sexo feminino. Pode-se perceber que para o conjunto das imagens de 0-18 anos obteve a melhor acurácia exata (sem tolerância) com 25,85% na rede GoogLeNet. A VGG-19 apresentou melhores resultados para o cálculo dos desvios, sendo 1,46 para MAE, 3,92 para MSE, 1,98 para RMSE e 18,45 para MAPE.

Quadro 5 - Resultados mão feminino

		GoogLeNet			VGG-19			Alexnet		
		Sem Tol.	Tol. 1 Ano	Tol. 2 Anos	Sem Tol.	Tol. 1 Ano	Tol. 2 Anos	Sem Tol.	Tol. 1 Ano	Tol. 2 Anos
0 - 18 anos	Acurácia (%)	25,85	56,36	78,12	25,45	58,46	81,85	21,17	53,07	71,76
	MAE	1,61	0,87	0,43	1,46	0,71	0,29	1,95	1,16	0,69
	MSE	5,10	2,62	1,33	3,92	1,76	0,75	7,50	4,39	2,54
	RMSE	2,26	1,62	1,15	1,98	1,32	0,87	2,74	2,10	1,59
	MAPE	22,24	11,63	5,70	18,45	8,06	3,07	23,72	13,14	7,43
5 - 18 anos	Acurácia (%)	25,67	52,49	74,05	28,80	62,53	83,31	22,57	56,10	75,21
	MAE	1,74	1,00	0,52	1,37	0,66	0,29	1,78	1,01	0,57
	MSE	5,78	3,04	1,51	3,65	1,61	0,67	6,22	3,44	1,86
	RMSE	2,40	1,74	1,23	1,91	1,27	0,82	2,49	1,85	1,37
	MAPE	16,80	9,84	5,29	13,11	6,43	2,99	16,70	9,38	5,36

Fonte: Autoria Própria.

Quando se considerou a tolerância de variação de 1 e 2 anos, a VGG-19 apresentou melhor desempenho, chegando a 81,85% no segundo caso, com a MAE em 0,29.

Já para o conjunto das imagens de 5-18 anos observa-se uma melhora nos valores obtidos, aumento na acurácia e redução nos desvios. Sendo 28,80% o valor da acurácia exata e 1,37 o do MAE.

O Quadro 6 mostra que o conjunto de dados do sexo masculino apresentou resultados me-lhores comparado aos obtidos para o sexo feminino e com o maior desempenho apresentado pela VGG-19 e a pior pela Alexnet. A acurácia exata para 0-18 anos apresentou 26,03% na VGG-19 e 21,71% na Alexnet. Com os melhores valores exatos das métricas em 1,32 para MAE, 3,49 para MSE, 1,87 para RMSE e 18,19 para MAPE. Quando observado a tolerância, o valor do MAE chega a 0.17 na de 2 anos.

Com a redução para a faixa etária de 5-18 anos, a acurácia sobe para 29,52% na VGG-19 sem tolerância de idade e 88,12 com 2 anos de tolerância. Assim como uma redução em todas outras métricas se comparado com a de 0-18 anos.

Quadro 6 - Resultados mão masculino

	GoogLeNet			VGG-19			Alexnet			
	Sem Tol.	Tol. 1 Ano	Tol. 2 Anos	Sem Tol.	Tol. 1 Ano	Tol. 2 Anos	Sem Tol.	Tol. 1 Ano	Tol. 2 Anos	
0 - 18 anos	Acurácia (%)	24,90	59,12	78,86	26,03	66,43	86,68	21,71	52,62	73,18
	MAE	1,58	0,83	0,42	1,32	0,58	0,24	1,85	1,06	0,59
	MSE	5,09	2,68	1,44	3,49	1,59	0,77	6,67	3,77	2,11
	RMSE	2,26	1,64	1,20	1,87	1,26	0,88	2,58	1,94	1,45
	MAPE	22,03	10,91	5,27	18,19	7,26	2,54	26,83	15,81	9,34
5 - 18 anos	Acurácia (%)	26,20	60,27	79,29	29,52	70,44	88,12	24,40	52,26	72,61
	MAE	1,52	0,78	0,39	1,17	0,46	0,17	1,76	1,00	0,52
	MSE	4,55	2,24	1,07	2,60	0,97	0,34	5,65	2,89	1,37
	RMSE	2,13	1,50	1,03	1,61	0,98	0,58	2,38	1,70	1,17
	MAPE	14,99	7,80	3,91	11,34	4,65	1,77	18,14	10,73	5,91

Fonte: Autoria Própria.

4.3 RESULTADOS DA ESTIMAÇÃO DA IDADE ÓSSEA DAS IMAGENS MOSAICO DOS DEDOS

Importante lembrar que houve uma redução de 28.18% do total de imagens para essa etapa da classificação.

Pode-se perceber que para os resultados obtidos no sexo feminino apresentado no Quadro 7, a VGG-19 apresentou os melhores resultados, mas teve uma redução de 1.46% na acurácia exata, resultado da subtração de 25,45% para o conjunto de mão feminina (valor referente à linha 01, coluna 06 do Quadro 5) menos 23,99% da mão inteira (valor referente à linha 01, coluna 06 do Quadro 7).

O resultado do mosaico de dedos femininos no conjunto reduzido, classes de 5-18 anos, apresentou melhor desempenho na acurácia do que para o conjunto de 0-18 anos. Isso pode ser resultado da forma como os *fold* foram divididos para a classificação, caso a divisão não tenha quantidade de imagens de todas as classes para o treinamento e teste. As melhores métricas obtidas foram no conjunto de 5-18 anos com MAE 1,49, MSE 4,14, RMSE 2,03 e MAPE 14,26.

Quadro 7 - Resultados dedos feminino

	GoogLeNet			VGG-19			Alexnet		
	Sem Tol.	Tol. 1 Ano	Tol. 2 Anos	Sem Tol.	Tol. 1 Ano	Tol. 2 Anos	Sem Tol.	Tol. 1 Ano	Tol. 2 Anos
Acurácia (%)	20,80	60,06	78,69	23,99	55,30	76,95	19,07	47,52	69,52
MAE	1,64	0,85	0,45	1,74	0,98	0,54	2,27	1,46	0,94
MSE	5,40	2,90	1,59	6,75	4,02	2,50	11,35	7,61	5,21
RMSE	2,32	1,70	1,26	2,60	2,00	1,58	3,37	2,76	2,28
MAPE	21,19	10,50	5,72	23,76	13,21	7,29	37,27	25,73	18,48
Acurácia (%)	23,26	57,14	76,10	23,60	57,97	81,38	21,36	48,98	67,51
MAE	1,64	0,89	0,47	1,49	0,74	0,33	2,03	1,26	0,76
MSE	5,37	2,84	1,48	4,14	1,91	0,85	8,06	4,77	2,76
RMSE	2,32	1,69	1,22	2,03	1,38	0,92	2,84	2,18	1,66
MAPE	15,56	8,59	4,70	14,26	7,17	3,39	19,69	12,36	7,68

Fonte: Autoria Própria.

O Quadro 8 mostra o valor das métricas obtidas para o conjunto do sexo masculino. A acurácia na classificação exata foi melhor para as duas faixas etárias em comparação com as da mão inteira do mesmo sexo com 27.34% de 0-18 anos e 30.63% de 5-18 anos, mas perde nas métricas com tolerância. Em sua maioria, as métricas foram melhores que o Quadro 7 feminino.

Quadro 8 - Resultados dedos masculino

		GoogLeNet			VGG-19			Alexnet		
		Sem Tol.	Tol. 1 Ano	Tol. 2 Anos	Sem Tol.	Tol. 1 Ano	Tol. 2 Anos	Sem Tol.	Tol. 1 Ano	Tol. 2 Anos
0 - 18 anos	Acurácia (%)	23,62	53,49	71,79	27,34	58,13	78,36	19,57	46,40	64,91
	MAE	1,85	1,09	0,62	1,63	0,91	0,49	2,18	1,37	0,84
	MSE	6,97	4,03	2,32	5,86	3,32	1,92	8,91	5,36	3,14
	RMSE	2,64	2,01	1,52	2,42	1,82	1,38	2,98	2,31	1,77
	MAPE	24,30	14,09	7,71	24,72	14,65	8,48	30,50	19,31	11,94
5 - 18 anos	Acurácia (%)	25,96	56,91	75,96	30,63	61,82	79,46	22,54	53,92	71,90
	MAE	1,65	0,91	0,48	1,46	0,76	0,38	1,86	1,09	0,63
	MSE	5,31	2,74	1,35	4,55	2,32	1,17	6,79	3,84	2,12
	RMSE	2,30	1,66	1,16	2,13	1,52	1,08	2,61	1,96	1,46
	MAPE	16,06	8,88	4,60	13,60	6,92	3,28	18,71	11,19	6,59

Fonte: Autoria Própria.

4.4 RESULTADOS DA ESTIMAÇÃO DA IDADE ÓSSEA DAS IMAGENS MOSAICO DOS CENTROS DE OSSIFICAÇÃO

Relembrando que a base de dados diminuiu em mais de 4.64% das imagens, por descartar mosaicos que não tinham 14 centros de ossificação.

O Quadro 9 mostra mais uma vez que o melhor desempenho foi da rede VGG-19 com a acurácia no conjunto de 0-18 anos de 22,36% sem tolerância e 75,09% com tolerância de 2 anos. Melhor ainda para o grupo de 5-18 anos com 25,41% e 78,17%.

Quadro 9 - Resultados mosaico feminino

		GoogLeNet			VGG-19			Alexnet		
		Sem Tol.	Tol. 1 Ano	Tol. 2 Anos	Sem Tol.	Tol. 1 Ano	Tol. 2 Anos	Sem Tol.	Tol. 1 Ano	Tol. 2 Anos
0 - 18 anos	Acurácia (%)	18,53	48,30	68,42	22,36	55,49	75,09	17,23	46,42	63,94
	MAE	2,42	1,61	1,09	2,11	1,33	0,89	2,57	1,74	1,20
	MSE	14,58	10,56	7,86	12,36	8,92	6,70	14,47	10,17	7,23
	RMSE	3,82	3,25	2,80	3,52	2,99	2,59	3,80	3,19	2,69
	MAPE	53,32	41,84	34,24	42,33	31,92	25,55	43,46	31,63	23,99
5 - 18 anos	Acurácia (%)	20,82	49,93	70,19	25,41	58,37	78,17	17,73	46,03	62,44
	MAE	1,99	1,20	0,70	1,68	0,93	0,52	2,51	1,69	1,15
	MSE	7,62	4,43	2,53	5,95	3,34	1,89	12,67	8,47	5,64
	RMSE	2,76	2,10	1,59	2,44	1,83	1,37	3,56	2,91	2,37
	MAPE	19,47	11,97	7,29	16,96	9,96	6,00	25,85	18,04	12,88

Fonte: Autoria Própria.

Para as imagens do mosaico dos centros de ossificação do sexo masculino os resultados foram um pouco inferiores, como apresentada no Quadro 10. Entretanto, mostrou menores desvios com os valores de 2,04 no MAE, 10,24 no MSE, 3,2 no

RMSE e 35,55 no MAPE na rede de VGG-19 de 0-18 anos. Nesse conjunto, a classificação de 5-18 anos também apresentou melhores desempenhos em questão de acurácia que foi de 22,77% para 27,43%, assim como foi de 71,85% para 77,86% com a tolerância de 2 anos.

Quadro 10 - Resultados mosaico masculino

	GoogLeNet			VGG-19			Alexnet			
	Sem Tol.	Tol. 1 Ano	Tol. 2 Anos	Sem Tol.	Tol. 1 Ano	Tol. 2 Anos	Sem Tol.	Tol. 1 Ano	Tol. 2 Anos	
0 - 18 anos	Acurácia (%)	22,28	50,88	66,73	22,77	55,24	71,85	21,28	46,55	63,04
	MAE	2,57	1,79	1,30	2,04	1,27	0,82	2,55	1,77	1,23
	MSE	17,11	12,75	9,65	10,24	6,93	4,83	15,03	10,71	7,71
	RMSE	4,14	3,57	3,11	3,20	2,63	2,20	3,88	3,27	2,78
	MAPE	48,72	37,46	29,89	35,55	24,52	17,91	42,21	30,81	23,00
5 - 18 anos	Acurácia (%)	24,04	53,25	71,05	27,43	60,52	77,86	21,25	42,88	59,89
	MAE	2,08	1,32	0,85	1,66	0,93	0,54	2,49	1,71	1,14
	MSE	9,84	6,44	4,27	6,32	3,73	2,26	12,05	7,85	5,01
	RMSE	3,14	2,54	2,07	2,51	1,93	1,50	3,47	2,80	2,24
	MAPE	21,60	14,26	9,62	16,36	9,26	5,53	24,63	17,09	11,55

Fonte: Autoria Própria.

4.5 RESULTADOS DA ESTIMAÇÃO DA IDADE ÓSSEA DAS IMAGENS MOSAICO DOS DEDOS E PUNHO

O Quadro 11 mostra os resultados da classificação da composição da imagem de mosaico dos dedos com o punho. Apresentando a melhor acurácia de 72,96% com a tolerância de 2 anos no grupo de 5-18 anos e a melhor MAE de 0,76 também na tolerância de 2 anos no grupo de 0-18 anos, ambos na CNN VGG-19.

Quadro 11 - Resultados dedos e punho feminino

	GoogLeNet			VGG-19			Alexnet			
	Sem Tol.	Tol. 1 Ano	Tol. 2 Anos	Sem Tol.	Tol. 1 Ano	Tol. 2 Anos	Sem Tol.	Tol. 1 Ano	Tol. 2 Anos	
0 - 18 anos	Acurácia (%)	18,99	46,77	66,76	17,95	46,08	67,89	19,83	49,91	66,77
	MAE	2,22	1,41	0,87	2,12	1,30	0,76	2,33	1,53	1,03
	MSE	9,63	6,00	3,72	8,66	5,24	3,18	11,96	8,09	5,53
	RMSE	3,10	2,45	1,93	2,94	2,29	1,78	3,46	2,84	2,35
	MAPE	31,98	20,67	13,22	32,46	20,58	12,78	38,44	26,88	19,03
5 - 18 anos	Acurácia (%)	17,43	43,23	63,62	21,12	54,62	72,96	17,57	50,70	68,59
	MAE	2,27	1,46	0,90	1,79	1,01	0,57	2,17	1,36	0,87
	MSE	9,61	5,89	3,53	6,26	3,45	1,87	9,81	6,29	4,06
	RMSE	3,10	2,43	1,88	2,50	1,86	1,37	3,13	2,51	2,01
	MAPE	22,38	14,67	9,47	17,53	10,18	5,92	21,00	13,36	8,91

Fonte: Autoria Própria.

Com a adição da imagem do punho aos mosaicos, esperou-se que pudesse aumentar o desempenho da classificação às imagens de mosaico. Porém, os resultados acabaram não atendendo às expectativas. Acredita-se que isso se deu em função do redimensionamento que ocorreu nessa imagem para compor à altura das imagens do mosaico e por não ter tratado a remoção do excesso de ossos de rádio e ulna que compõem os braços dos indivíduos.

Os resultados do sexo masculino são apresentados no Quadro 12. A melhor acurácia foi no grupo de 0-18 anos com a tolerância de 2 anos e a melhor MAE nesse mesmo grupo foi de 1,91 sem tolerância. Já as métricas de MSE, RMSE e MAPE foram melhores no grupo de 5-18 anos com os valores de 7,14, 2,67 e 18,98 respectivamente e sem tolerância de idade.

Quadro 12 - Resultados dedos e punho masculino

	GoogLeNet			VGG-19			Alexnet			
	Sem Tol.	Tol. 1 Ano	Tol. 2 Anos	Sem Tol.	Tol. 1 Ano	Tol. 2 Anos	Sem Tol.	Tol. 1 Ano	Tol. 2 Anos	
0 - 18 anos	Acurácia (%)	19,91	43,64	63,22	24,49	54,10	72,00	17,30	41,81	57,89
	MAE	2,33	1,53	0,97	1,91	1,16	0,70	2,55	1,72	1,14
	MSE	10,53	6,66	4,16	7,89	4,82	2,97	12,33	8,06	5,21
	RMSE	3,24	2,58	2,04	2,81	2,20	1,72	3,51	2,84	2,28
	MAPE	36,03	25,05	17,04	29,51	18,81	11,94	40,16	28,68	19,93
5 - 18 anos	Acurácia (%)	20,78	47,72	64,83	21,03	49,12	67,95	19,73	42,96	61,23
	MAE	2,21	1,42	0,89	1,98	1,19	0,68	2,35	1,54	0,97
	MSE	9,21	5,58	3,27	7,14	3,98	2,11	9,99	6,10	3,59
	RMSE	3,04	2,36	1,81	2,67	1,99	1,45	3,16	2,47	1,89
	MAPE	21,09	13,40	8,30	18,98	11,38	6,48	22,89	15,24	9,73

Fonte: Autoria Própria.

4.6 RESULTADOS DA ESTIMAÇÃO DA IDADE ÓSSEA DAS IMAGENS MOSAICO DOS CENTROS DE OSSIFICAÇÃO E PUNHO

O último conjunto de imagens utilizados para a classificação foi o da composição do mosaico dos centros de ossificação com o punho. No sexo feminino, como mostra o Quadro 13 teve uma melhora significativa nas métricas obtidas pelo conjunto de 5-18 anos na melhor CNN, a GoogLeNet, que foi de 2,04, 8,30, 2,88 e 20,99 para MAE, MSE, RMSE e MAPE respectivamente. A melhor acurácia está nesse mesmo conjunto de idades com o valor de 68,51% com a tolerância de 2 anos.

Quadro 13 - Resultados mosaico e punho feminino

		GoogLeNet			VGG-19			Alexnet		
		Sem Tol.	Tol. 1 Ano %	Tol. 2 Anos %	Sem Tol.	Tol. 1 Ano %	Tol. 2 Anos %	Sem Tol.	Tol. 1 Ano %	Tol. 2 Anos %
0 - 18 anos	Acurácia (%)	16,79	41,87	60,59	21,70	49,21	67,95	14,51	38,71	54,13
	MAE	3,17	2,04	1,61	2,60	1,59	1,22	3,38	2,17	1,72
	MSE	20,09	13,49	11,00	15,83	10,60	9,09	20,18	13,03	10,43
	RMSE	4,48	3,67	3,32	3,98	3,26	3,02	4,49	3,61	3,23
	MAPE	59,73	42,88	38,15	52,23	37,01	33,19	60,04	41,85	34,79
5 - 18 anos	Acurácia (%)	18,35	43,85	63,98	21,71	49,58	68,51	15,73	39,58	55,57
	MAE	2,41	1,59	1,03	2,04	1,26	0,76	2,83	1,99	1,38
	MSE	11,17	7,17	4,55	8,30	4,99	2,97	14,95	10,14	6,77
	RMSE	3,34	2,68	2,13	2,88	2,23	1,72	3,87	3,18	2,60
	MAPE	25,35	17,49	11,93	20,99	13,63	8,77	28,83	20,67	14,82

Fonte: Autoria Própria.

O conjunto de resultados de mosaico dos centros de ossificação com punho do sexo masculino é apresentado no Quadro 14, com uma boa acurácia exata de 25,20% no grupo de 5-18 anos e 70,66% para a tolerância de 2 anos.

Quadro 14 - Resultados mosaico e punho masculino

		GoogLeNet			VGG-19			Alexnet		
		Sem Tol.	Tol. 1 Ano	Tol. 2 Anos	Sem Tol.	Tol. 1 Ano	Tol. 2 Anos	Sem Tol.	Tol. 1 Ano	Tol. 2 Anos
0 - 18 anos	Acurácia (%)	20,00	45,69	60,80	22,74	53,93	70,35	16,52	43,98	58,03
	MAE	2,87	2,07	1,53	2,22	1,44	0,98	2,62	1,78	1,22
	MSE	19,15	14,20	10,61	12,83	9,17	6,74	13,72	9,32	6,31
	RMSE	4,38	3,77	3,26	3,58	3,03	2,60	3,70	3,05	2,51
	MAPE	52,16	40,60	32,39	40,78	29,80	23,05	43,44	31,48	22,72
5 - 18 anos	Acurácia (%)	21,85	50,14	63,70	25,20	52,28	70,66	17,16	43,84	58,46
	MAE	2,48	1,69	1,20	2,06	1,31	0,83	2,56	1,74	1,17
	MSE	12,86	8,69	5,80	9,45	6,08	3,94	12,47	8,18	5,27
	RMSE	3,59	2,95	2,41	3,07	2,47	1,99	3,53	2,86	2,30
	MAPE	25,74	18,16	13,12	21,87	14,53	9,72	25,55	17,59	12,09

Fonte: Autoria Própria.

4.7 COMPARAÇÃO ENTRE OS RESULTADOS DESTE TRABALHO

A maior parte dos resultados se mostraram satisfatórios com as acurácias e métricas comparáveis ao estado da arte.

A sumarização dos melhores valores encontrados com os experimentos do presente trabalho está apresentada nos Quadros 15 e 16.

Quadro 15 - Melhores desempenhos do presente trabalho 0-18 anos

Sexo	Variação	Acurácia %	Base	MAE	Base	MSE	Base	RMSE	Base	MAPE	Base
Feminino	Sem Tol.	25.45	MI	1.46	MI	3.92	MI	1.98	MI	18.45	MI
	Tol. 1 Ano	58.46	MI	0.71	MI	1.76	MI	1.32	MI	8.06	MI
	Tol. 2 Anos	81.85	MI	0.29	MI	0.75	MI	0.87	MI	3.07	MI
Masculino	Sem Tol.	27.34	MD	1.32	MI	3.49	MI	1.87	MI	18.19	MI
	Tol. 1 Ano	66.43	MI	0.58	MI	1.59	MI	1.26	MI	7.26	MI
	Tol. 2 Anos	86.68	MI	0.24	MI	0.77	MI	0.88	MI	2.54	MI

Nota: MI: Mão Inteira; MD: Mosaico de Dedos.

Fonte: Autoria Própria.

Quadro 16 - Melhores desempenhos do presente trabalho 5-18 anos

Sexo	Variação	Acurácia %	Base	MAE	Base	MSE	Base	RMSE	Base	MAPE	Base
Feminino	Sem Tol.	28.8	MI	1.37	MI	3.65	MI	1.97	MI	13.11	MI
	Tol. 1 Ano	62.53	MI	0.66	MI	1.61	MI	1.27	MI	6.43	MI
	Tol. 2 Anos	83.31	MI	0.29	MI	0.67	MI	0.82	MI	2.99	MI
Masculino	Sem Tol.	30.63	MD	1.17	MI	2.60	MI	1.61	MI	11.34	MI
	Tol. 1 Ano	70.44	MI	0.46	MI	0.97	MI	0.98	MI	4.65	MI
	Tol. 2 Anos	88.12	MI	0.17	MI	0.34	MI	0.58	MI	1.77	MI

Nota: MI: Mão Inteira; MD: Mosaico de Dedos.

Fonte: Autoria Própria.

O conjunto de imagens da mão masculina teve um destaque nas melhores métricas obtidas nos experimentos realizados. Com exceção da melhor acurácia exata que foi gerada pelo conjunto de imagens do mosaico de dedos masculinos de 5-18 anos com 30,63%. E de todos os experimentos, o conjunto de mosaico dos dedos com punho obtiveram as acurácias mais baixas considerada a tolerância de 2 anos com o valor de 67,89% numa diferença de 20,23% para a melhor acurácia encontrada.

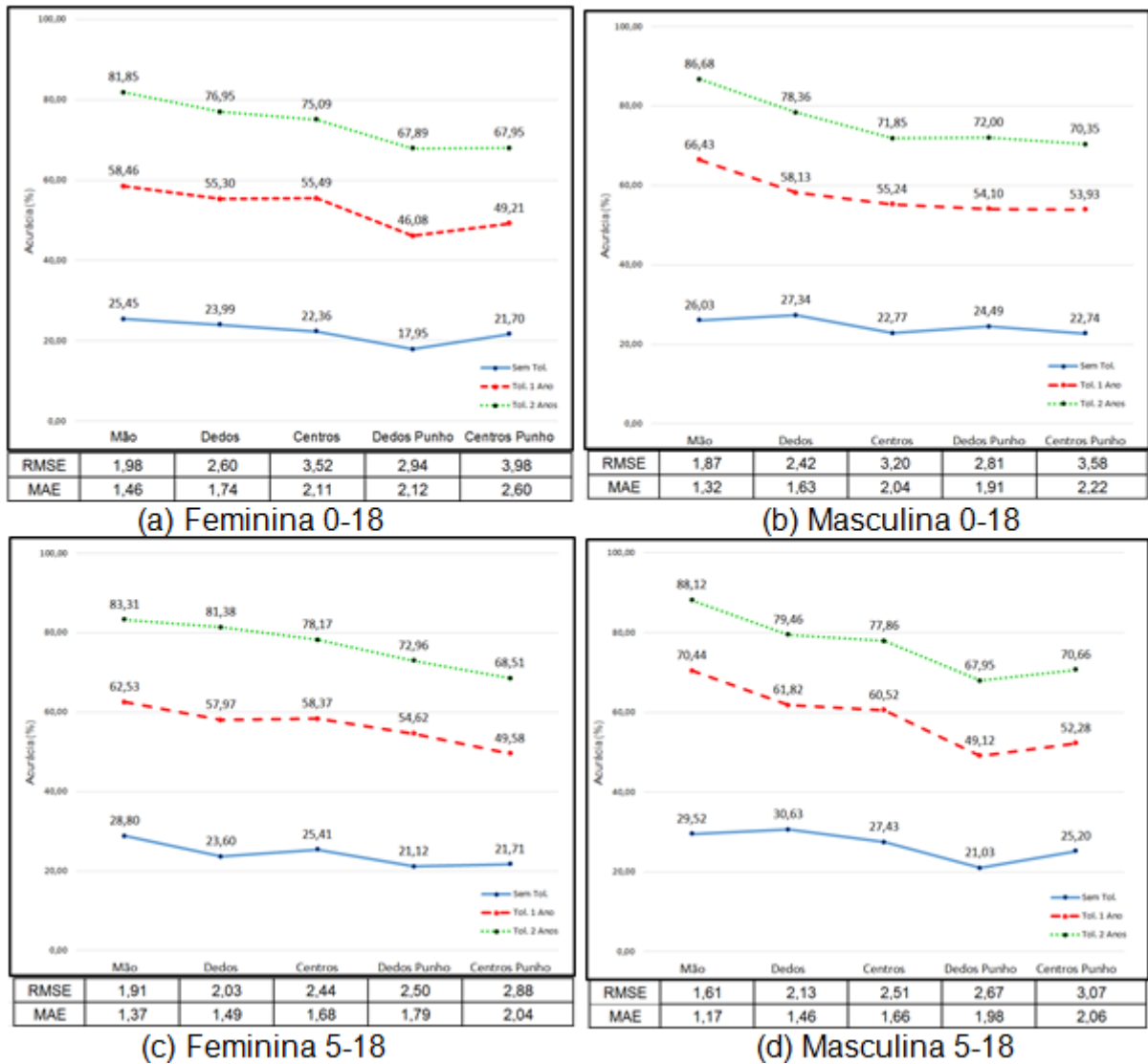
Como pode ser percebido com todos os resultados do Quadro 5 à Quadro 14, VGG-19 apresentou o melhor desempenho em quase todos dos tipos de imagem. Dessa forma, as acurácias obtidas por essa CNN é demonstrada conforme a Figura 33.

A Figura 33 (a) e a Figura 33 (b) apresentam as acurácias dos conjuntos de imagens femininas e masculinas de 0-18 anos exatas, com 1 e 2 anos de tolerância, respectivamente. Assim como os seus valores de RMSE e MAE exatos de cada conjunto.

Já a Figura 33 (c) e a Figura 33 (d) apresentam as acurácias dos conjuntos de imagens femininas e masculinas de 5-18 anos exatas, com 1 e 2 anos de tolerância,

respectivamente. Assim como os valores de RMSE e MAE sem tolerância de cada conjunto.

Figura 33 - Desempenho das 5 diferentes tipos de imagens da VGG-19



Fonte: Autoria Própria.

Através da Figura 33, concluímos que a hipótese de que a composição do mosaico dos dedos e dos centros de ossificação não melhorou a estimativa da idade óssea, seja pela redução do conjunto de imagens que compõem para o treinamento e classificação das mesmas.

4.8 COMPARAÇÃO COM TRABALHOS CORRELATOS

A comparação com os trabalhos correlatos é apresentada na Tabela 17. Dentre eles, dois são sistemas totalmente automatizados: Lee et al. (2017) e Spampinato et al. (2017). Os demais apresentaram, em alguma fase do sistema, a intervenção manual do usuário, seja para a melhoria da imagem ou para a indicação da localização das regiões de interesse.

Quadro 17 - Comparação com o desempenho dos trabalhos correlatos

Método	Base de Dados	# Imagens	Faixa Etária	Sexo	Acurácia %	Tol. 0.5 Ano	Tol. 1 Ano %	Tol. 2 Anos %	MAE	RMSE
Lee et al. (2017)	Privada	4047 M	5-18	M	40.60	-	75.54	92.35	-	1.45
		4278 F		F	39.06	-	75.59	90.08	-	1.51
Spampinato et al. (2017)	Pública	1391	0-18	All	-	-	-	-	1.15	-
Kashif et al. (2016)	Pública	1101	0-18	All	45.69	-	89.38	98.36	0.6	-
Hsieh et al. (2007)	Privada	465	2-8	M	-	64	84	-	-	-
		444		2.5-10	F	-	59	80	-	-
Olivete e Rodrigues (2005)	Privada	100	6-18	M	30.00	100	100	100	-	-
		100		F	23.33	100	100	100	-	-
		695		M	29.52	-	66.43	86.68	1.32	1.87
Presente Trabalho	Pública	685	0-18	F	25.45	-	58.46	81.85	1.46	1.98
		599		M	30.63	-	70.44	88.31	1.17	1.61
		585		F	28.80	-	62.53	83.31	1.37	1.91

Fonte: Autoria Própria.

Pode-se perceber que, na comparação da acurácia exata obtida, o presente trabalho apresentou melhor desempenho que a do artigo de Olivete e Rodrigues (2005) (com a intervenção manual do usuário para ajustar a localização dos limites dos ossos de interesse). Sendo a melhor acurácia exata deles foi de 30% enquanto a do presente trabalho foi 30.63%.

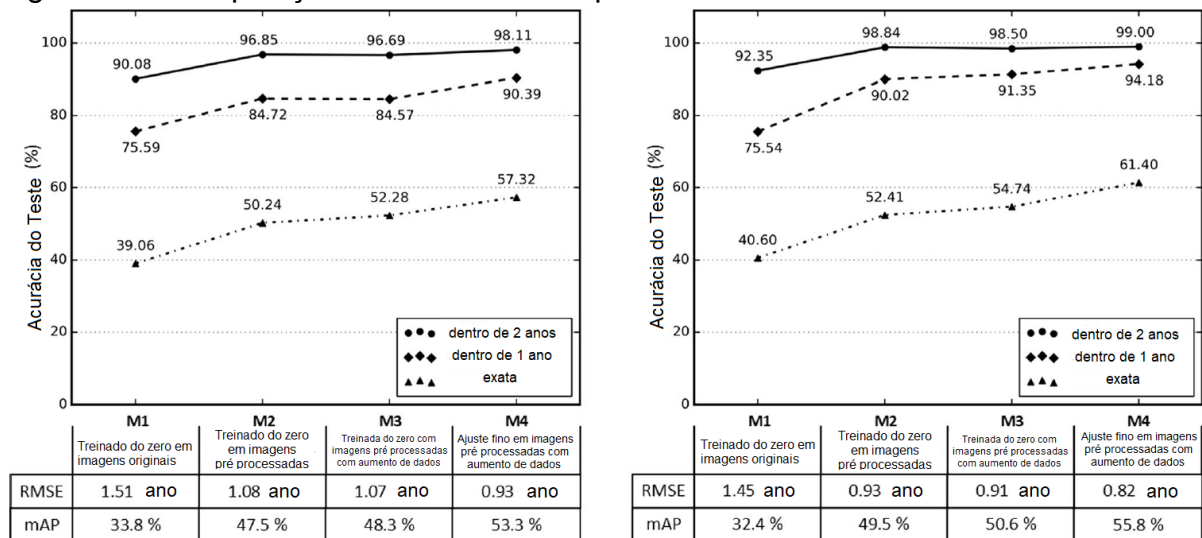
Na acurácia com 1 e 2 anos de tolerância, considerando a mesma faixa etária de 5-18 anos, o presente trabalho também conseguiu aproximar bem do artigo de Lee et al. (2017) que teve uma base 5 vezes maior que base da Universidade do Sul da Califórnia, além do uso de *data augmentation*.

Os artigos de Hsieh et al. (2007) e de Kashif et al. (2016) obtiveram bom desempenho nas acurácias, porém vale ressaltar que ambos tiveram intervenção manual no processo da estimaco da idade óssea.

4.8.1 Comparaco com trabalho de Lee

Lee et al. (2017) apresentou um bom desempenho por se tratar de um sistema totalmente automatizado e obteve os resultados conforme a Figura 34. Porém, deve-se considerar principalmente que a base de dados é privada, podendo ter qualidade muito melhor que a base de dados pública da Universidade do Sul da Califórnia, além da restrico da faixa etária de 5-18 anos. A base tinha uma quantidade muito grande de imagens e foi usada a técnica *Data Augmentation*, em que cada uma das 8.325 imagens pode ser transformada em tempo de execuo para 1.107.150.000 imagens.

Figura 34 - Composico das PROI com o punho



(a) Feminino
Fonte: Adaptado de Lee et al. (2017).

4.8.2 Comparação com artigo de Kashif e colegas

Kashif et al. (2016) utilizou a mesma base de dados, porém, eles utilizaram apenas 1.101 imagens da base. O presente trabalho se inspirou no artigo de Kashif et al. (2016) para a extração dos 14 centros de ossificação, automatizando o que foi feito por eles manualmente na etapa de segmentação com a indicação da localização desses 14 centros através de cliques de *mouse*.

Apesar de não ter chegado a mesma acurácia de 45,69% obtido pelo artigo de Kashif e colegas, considera-se que o presente trabalho alcançou o objetivo proposto e se aproximou do resultado apresentado por Spampinato et al. (2017).

4.8.3 Comparação com artigo de Spampinato e colegas

Spampinato et al. (2017) automatizou o processo da estimativa da idade óssea com a mesma base de dados do presente trabalho. Chegou ao valor de MAE em 1,15 com a utilização de *Data Augmentation*, mas não apresenta informação sobre acurácia.

O presente trabalho chegou ao MAE de 1,17, valor próximo ao do artigo de Spampinato e colegas para o conjunto de imagens de mão do sexo masculino. Vale destacar que, para os demais conjuntos de dados, mesmo havendo uma redução significativa na quantidade de imagens que os compõem, mantiveram-se com o desempenho considerável.

5 CONSIDERAÇÕES FINAIS

A automatização do processo de estimação da idade óssea busca auxiliar os trabalhos de especialistas que ainda usam os métodos clássicos e manuais para tal tarefa. A automatização desse processo irá otimizar o tempo dos especialistas em que, antes, precisavam pegar a radiografia de um indivíduo de 0-18 anos comparar visualmente os centros de ossificação com um atlas de referências como o de GP para então atribuir uma idade óssea ao indivíduo e avaliar se o seu desenvolvimento ocorre dentro das normalidades.

Para a compreensão do processo de automatização da estimação da idade óssea, vários trabalhos correlatos no âmbito Brasil e no mundo, que expressam o estado da arte nesse campo foram estudados. Foram extraídas várias ideias dos trabalhos correlatos, adaptados e implementados para se chegar aos resultados obtidos pelo presente trabalho e foram considerados bastante satisfatórios dentro das suas limitações.

Para a estimativa da idade óssea usando *Deep Learning*, duas macro etapas foram implementadas: a criação das imagens mosaico e a classificação da idade óssea. Na primeira etapa, montou-se 4 diferentes bases de dados: uma com os 5 dedos; uma com os 5 dedos e o punho; uma com 14 centros de ossificação; e por último, uma com 14 centros de ossificação e o punho.

O objetivo do trabalho foi a comparação dos resultados de um sistema totalmente automático para a estimação da idade óssea através de radiografias carpais, separadas pela informação de sexo do indivíduo, usando imagens mosaicos dos centros de ossificação como entrada do algoritmo de CNN baseado em *Deep Learning* foi alcançado.

No entanto, os resultados não corroboraram com a hipótese de que fornecer imagens mosaico, destacando-se os centros de ossificação das mãos, como entrada do algoritmo de *Deep Learning* melhoraria a estimativa da idade óssea. Supõe-se que para a técnica *Deep Learning*, a informação espacial entre os ossos (correlação

espacial entre os ossos), que é perdida quando se monta as imagens mosaico, seja importante para se ter um melhor resultado. Outro resultado é que incluir a imagem do punho não melhorou os resultados dos mosaicos dos dedos e centros de ossificação.

Ao final, pode-se avaliar que a melhor ConvNet para esta aplicação foi a VGG-19 sobre a AlexNet e a GoogLeNet.

Diante dos experimentos realizados neste trabalho foi observado que apesar de ser um procedimento antigo, ainda é uma aplicação pouco automatizada na área médica e que tem bastante a melhorar e agregar soluções mais sofisticadas para auxiliar no trabalho dos especialistas.

O sistema focou a automatização do processo de estimativa da idade óssea usando a composição do mosaico dos centros de ossificação da mão. Não tentou propor melhorias para essa estimativa. Ressalta-se que o presente trabalho conseguiu automatizar todo o processo de estimativa da idade óssea mantendo a qualidade da classificação.

5.1 TRABALHOS FUTUROS

Melhorias podem ser feitas para aumentar a desempenho do presente trabalho:

- Determinar uma heurística em que se a imagem, em tempo de processamento, for descartada por não conseguir fazer a sua estimativa da idade óssea. Essa imagem radiográfica será encaminhada para o especialista humano avaliar.
- Adotar o uso da Transformada de Hough para tentar melhorar a identificação dos pontos de referências necessários para a fase da segmentação das imagens de mosaicos.

- Acrescentar uma etapa de regularização das imagens com a técnica de *Deep Learning* verificar a qualidade do conteúdo da segmentação e adequação das imagens para a posterior classificação das mesmas. Isto é, se os recortes realmente continham a forma esperada, ficando dentro de um padrão esperado.
- Fazer a etapa de segmentação dos centros de ossificação adaptando o método de pré-processamento também baseada na técnica de *Deep Learning* como proposto no trabalho de Lee et al. (2017).
- Adotar o *data augmentation* para aumentar a quantidade de imagens necessárias para o *Deep Learning*.
- Fazer a separação também das etnias para comparar os resultados.
- Utilizar uma base de dados de radiografias do Brasil.

REFERÊNCIAS

- BEHIELS, G. et al. Retrospective correction of the heel effect in hand radiographs. **Medical Image Analysis**, v. 6, n. 3, p. 183-190, 2002. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S1361841502000786>>.
- CAO, F. et al. Digital hand atlas and web-based bone age assessment: system design and implementation. **Computerized medical imaging and graphics**, Elsevier, v. 24, n. 5, p. 297-307, 2000.
- _____. et al. Image database for digital hand atlas. In: HUANG, H. K.; RATIB, O. M. (Eds.). **Medical Imaging 2003: PACS and Integrated Medical Information Systems: Design and Evaluation**. 2003, v. 5033, p. 461-470.
- CASANOVA, M. S. et al. Análise comparativa da maturação óssea determinada pelo método de Grave-Brown entre imagens convencionais e digitalizadas. **Revista Dental Press de Ortodontia e Ortopedia Facial**, Scielo, v. 11, p. 104-109, out. 2006. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1415-54192006000500011&nrm=iso>.
- CASTRO, Fernando Cordeiro de. et al. **Localização, segmentação e classificação automáticas de regiões de interesse para a determinação de maturidade óssea utilizando o método de Tanner-Whitehouse**. 2009. 120f. Dissertação (Mestrado em Engenharia Elétrica) - Universidade Federal de Uberlândia, Uberlândia, 2009.
- DENG, L. A tutorial survey of architectures, algorithms, and applications for deep learning. **APSIPA Transactions on Signal and Information Processing**, Cambridge University Press, v. 3, 2014.
- DESERNO, T. M. **Biomedical image processing**. Springer Science & Business Media, 2011.
- EKLÖF, O.; RINGERTZ, H. A method for assessment of skeletal maturity. **Annales de radiologie**, v. 10, n. 3, p. 330-336, 1967.
- FAN, B. C. et al. Automatic bone age estimation based on carpal-bone image – a preliminary report. **Chin Med J (Taipei)**, v. 64, p. 203-208, 2001.
- FONSECA, V. de L.; VOLPASSO, L. H.; VIANNA, G. K. Utilização de sistema especialista para diagnósticos de doenças transmitidas pelo aedes aegypti. ESCOLA REGIONAL DE SISTEMAS DE INFORMAÇÃO DO RIO DE JANEIRO. 3., 2016, Seropédica. p. 89-94, nov. 2016.
- GERTYCH, A. et al. Bone age assessment of children using a digital hand atlas. **Computerized Medical Imaging and Graphics**, Elsevier, v. 31, n. 4, p. 322-331, 2007.

GONZALEZ, R. C.; WOODS, R. **Processamento digital de imagens**. São Paulo: Pearson Prentice Hall, 2010.

GREULICH, W. W.; PYLE, S. I. Radiographic atlas of skeletal development of the hand and wrist. **The American Journal of the Medical Sciences**, v. 238, n. 3, p. 393, 1959.

GUIMARÃES, C. M. et al. Um novo sistema simplificado para cirurgia virtual guiada com implantes dentários-relato de caso clínico com acompanhamento clínico e tomográfico de 11 meses. **ImplantNews**, v. 11, n. 6, p. 803-811, 2014.

HINTON, G. E. et al. **Improving neural networks by preventing co-adaptation of feature detectors**. arXiv preprint arXiv:1207.0580, 2012.

HORTER, M. et al. Determination of skeletal age: comparison of the methods of greulich and pyle and tanner and whitehouse. **Der Orthopade**, v. 41, n. 12, p. 966-976, 2012.

HSIEH, C. et al. Automatic segmentation of phalanx and epiphyseal/metaphyseal region by gamma parameter enhancement algorithm. **Measurement Science Review**, v. 12, n. 1, p. 21-27, 2012.

HSIEH, C.; LIU, T.; TIU, C. Growth estimation based on statistics of relative position of carpals. **Advanced Science Letters**, American Scientific Publishers, v. 7, n. 1, p. 579-582, 2012.

HSIEH, C. W. et al. Computerized geometric features of carpal bone for bone age estimation. **Chinese Medical Journal-Beijing-English Edition**, Chinese Medical Association, v. 120, n. 9, p. 767, 2007.

_____. et al. A hierarchical algorithm for phalangeal and epiphyseal/metaphyseal segmentation. **Multimedia Tools and Applications**, Springer, v. 76, n. 2, p. 3047-3063, 2017.

_____. et al. Turner's syndrome screening based on hand image assessment. In: IEEE. INFORMATION TECHNOLOGY AND APPLICATIONS IN BIOMEDICINE, 2008. ITAB 2008. INTERNATIONAL CONFERENCE, 2008. p. 378-379.

_____. et al. A fuzzy-based growth model with principle component analysis selection for carpal bone-age assessment. **Medical & biological engineering & computing**, Springer, v. 48, n. 6, p. 579-588, 2010.

_____. et al. Fast and fully automatic phalanx segmentation using a grayscale-histogram morphology algorithm. **Optical Engineering**, International Society for Optics and Photonics, v. 50, n. 8, p. 087007-087007, 2011.

_____. et al. Interrelationship of carpal angle and bone age in children of different generations: a retrospective study. **Journal of Pediatric Endocrinology and Metabolism**, v. 24, n. 11-12, p. 1015-1018, 2011.

_____. et al. Simplified radius, ulna, and short bone-age assessment procedure using grouped-tanner-whitehouse method. **Pediatrics International**, Wiley Online Library, v. 53, n. 4, p. 567-575, 2011.

HYNDMAN, R. J.; ATHANASOPOULOS, G. **Forecasting**: principles and practice. OTexts, 2014.

JOST, Ingo. **Aplicação de Deep Learning em dados refinados para mineração de opiniões**. 2015. 88 f. Dissertação (Mestrado em Computação aplicada) - Universidade do Vale do Rio dos Sinos, São Leopoldo, 2015.

FREITAS JÚNIOR, E. de. **A robust deep convolutional neural network model for text categorization**. 2016. 157 f. Dissertação (Mestrado em Ciência da Computação) - Universidade Federal de Minas Gerais, Belo Horizonte, 2016.

KARPATHY, A. Cs231n: convolutional neural networks for visual recognition. **Neural networks**, v. 1, 2016.

KASHIF, M. et al. Feature description with sift, surf, brief, brisk, or freak? a general question answered for bone age assessment. **Computers in biology and medicine**, Elsevier, v. 68, p. 67-75, 2016.

KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: **ADVANCES in neural information processing systems**. 2012. p. 1097-1105.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **Nature**, Nature Research, v. 521, n. 7553, p. 436-444, 2015.

LECUN, Y. et al. Gradient-based learning applied to document recognition. **Proceedings of the IEEE**, IEEE, v. 86, n. 11, p. 2278-2324, 1998.

LEE, H. et al. Fully automated deep learning system for bone age assessment. **Journal of Digital Imaging**, Springer, p. 1-15, 2017.

LEE, S. et al. Fingernet: Deep learning-based robust finger joint detection from radiographs. In: IEEE. **Biomedical Circuits and Systems Conference (BioCAS), 2015 IEEE**. 2015. p. 1-4.

MENDES, Y. B. E. et al. Analysis of bone maturation in patients 13 to 20 years old by means of wrist radiographies. **Dental Press Journal of Orthodontics**, SciELO Brasil, v. 15, n. 1, p. 74-79, 2010.

MENZE, B. et al. **Medical computer vision**: recognition techniques and applications in medical imaging. Springer, v. 6533, 2011.

MORAES, M. E. L. de. et al. Fidelidade dos métodos greulich & pyle e eklöf & ringertz para avaliação da idade óssea em crianças brasileiras. **Rev Odontol UNESP**, v. 32, n. 1, p. 9-17, 2003.

MOTA, L. M. H. d. et al. Diagnóstico por imagem da artrite reumatoide inicial. **Revista Brasileira de Reumatologia**, Elsevier, v. 52, n. 5, p. 761-766, 2012.

H. NETO, F. et al. Estudo comparativo dos métodos de estimativa da idade óssea de greulich & pyle e tanner & whitehouse. **Pesquisa Odontológica Brasileira**, SciELO Brasil, 2000.

OLIVETE, C. J. **Estimativa da idade óssea através da análise carpal baseada na simplificação do método de Eklof & Ringertz**. 2005. Dissertação (Mestrado em Engenharia Elétrica) - Universidade de São Paulo, São Carlos, 2005.

_____. et al. Utilização na proporção divina no processo de estimação da maturidade óssea. In: WORKSHOP DE VISÃO COMPUTACIONAL. 8., 2012, Goiânia/GO.

_____. RODRIGUES, E. Simplificações do método de eklof & ringertz para automatizar o processo de estimação da maturidade óssea. WORKSHOP DE VISÃO COMPUTACIONAL. 5., 2005.

_____. _____. Maturidade óssea: estimação por simplificações do método de eklof e ringertz. **Radiologia Brasileira**, Colégio Brasileiro de Radiologia e Diagnóstico por Imagem, v. 43, n. 1, p. 13-18, fev. 2010. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-9842010000100006&nrm=iso>.

_____. _____. NASCIMENTO, M. do. O efeito da correção do “efeito heel” em imagens radiográficas da mão. **Revista Brasileira de Física Médica**, v. 1, n. 1, p. 38-50, 2005.

QUEIROZ, A. et al. Metodologia para extração de características automáticas da mão usando assinatura vertical. WORKSHOP DE VISÃO COMPUTACIONAL. 1., 2005, Piracicaba, p. 60-63.

RAVINDRA, S. **What is image recognition and why is it used?** 2017. Disponível em: <<https://www.kdnuggets.com/2017/08/convolutional-neural-networks-image-recognition.html>>.

RÉ, A. H. N. Growth, maturation and development during childhood and adolescence: Implications for sports practice. **Motricidade**, v. 7, n. 3, p. 55-67, 2011.

RIJN, R. R. V.; THODBERG, H. H. Bone age assessment: automated techniques coming of age? **ACTA Radiology**, SAGE Publications Sage UK: London, England, dec. 2016.

RODRIGUES, Maria de Fátima R. **Avaliação da idade e maturação óssea através de radiografias**. 2003. 61 f. Monografia (Especialização em Radiologia Odontológica) - Faculdade de Odontologia de Piracicaba, Universidade Estadual de Campinas, Piracicaba, 2003.

RUSSAKOVSKY, O. et al. Imagenet large scale visual recognition challenge. **International Journal of Computer Vision**, Springer, v. 115, n. 3, p. 211-252, 2015.

SIMONYAN, K.; ZISSERMAN, A. **Very deep convolutional networks for large-scale image recognition**. arXiv preprint arXiv:1409.1556, 2014.

SPAMPINATO, C. et al. Deep learning for automated skeletal bone age assessment in x-ray images. **Medical image analysis**, Elsevier, v. 36, p. 41-51, 2017.

STERN, D. et al. Fully automatic bone age estimation from left hand mr images. In: SPRINGER. INTERNATIONAL CONFERENCE ON MEDICAL IMAGE COMPUTING AND COMPUTER-ASSISTED INTERVENTION. 2014. p. 220-227.

SZEGEDY, C. et al. Going deeper with convolutions. In: *PROCEEDINGS OF THE IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION*. 2015. p. 1-9.

TANNER, J. M.; WHITEHOUSE, R.; TAKAISHI, M. Standards from birth to maturity for height, weight, height velocity, and weight velocity: British children, 1965. **Archives of Disease in Childhood**, BMJ Group, v. 41, n. 219, p. 454, 1966.

THRALL, J. H. Trends and developments shaping the future of diagnostic medical imaging: 2015 annual oration in diagnostic radiology. **Radiology**, v. 279, n. 3, p. 660-666, 2016. Disponível em: <<https://doi.org/10.1148/radiol.2016160293>>.

TIU, C. M. et al. Turner syndrome phalangeal screening based on a two-stage linear regression concept. **Pediatrics International**, Wiley Online Library, v. 51, n. 4, p. 453-459.

VARGAS, A. C. G.; PAES, A.; VASCONCELOS, C. N. Um estudo sobre redes neurais convolucionais e sua aplicação em detecção de pedestres. In: CONFERENCE ON GRAPHICS, PATTERNS AND IMAGES. 29., 2016. **Proceedings...** 2016. p. 1-4.