

UNIVERSIDADE FEDERAL DE SÃO CARLOS– UFSCAR
CENTRO DE CIÊNCIAS EXATAS E DE TECNOLOGIA– CCET
DEPARTAMENTO DE COMPUTAÇÃO– DC
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO– PPGCC

Juliana Wolf Pereira

**Anotação Semântica Baseada em
Ontologia aplicada em Imagens Médicas**

São Carlos
2023

Juliana Wolf Pereira

**Anotação Semântica Baseada em
Ontologia aplicada em Imagens Médicas**

Tese apresentada ao Programa de Pós-Graduação em Ciência da Computação do Centro de Ciências Exatas e de Tecnologia da Universidade Federal de São Carlos, como parte dos requisitos para a obtenção do título de Doutora em Ciência da Computação.

Área de concentração: Metodologias e Técnicas de Computação

Orientador: Marcela Xavier Ribeiro

São Carlos

2023



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Exatas e de Tecnologia
Programa de Pós-Graduação em Ciência da Computação

Folha de Aprovação

Defesa de Tese de Doutorado da candidata Juliana Wolf Pereira, realizada em 13/07/2023.

Comissão Julgadora:

Profa. Dra. Marilde Terezinha Prado Santos (UFSCar)

Prof. Dr. Renato Bueno (UFSCar)

Prof. Dr. Diego Furtado Silva (UFSCar)

Prof. Dr. Agma Juci Machado Traina (USP)

Prof. Dr. Dilvan de Abreu Moreira (USP)

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa de Pós-Graduação em Ciência da Computação.

Este trabalho é dedicado para todas e todos que de alguma maneira contribuem com a educação pública brasileira, em especial para meninas e mulheres na ciência e tecnologia.

Agradecimentos

Agradeço a Deus.

Agradeço minha família.

Agradeço à Professora Marcela por ter me colocado este desafio ao me aceitar como sua orientada. E por ter apoiado todo este trabalho para que ele se tornasse realidade. Obrigada por todos os ensinamentos, motivação, encorajamentos e risadas.

Agradeço com muito carinho à Professora Marilde, por todo acolhimento e incentivo. Desde quando cheguei na UFSCar durante o programa de mobilidade acadêmica na graduação e tive a oportunidade de realizar uma iniciação científica, e depois o mestrado. Isso tudo foi essencial para que eu chegasse até aqui.

Agradeço aos membros da banca, Professora Agma, Professor Dilvan, Professor Diego e Professor Renato.

Agradeço ao Hakan Bulu por ter disponibilizado a ontologia MAO, ponto de partida deste trabalho.

Agradeço todos os professores e servidores do DC. Em especial, Professora Sandra Abib.

Agradeço à universidade pública, à UFSCar e ao Programa de Pós-Graduação em Ciência da Computação - PPGCC/UFSCar.

Agradeço à UFMS pelo incentivo à qualificação.

Agradeço à CAPES, ao CNPQ e à FAPESP.

“I was taught that the way of progress is neither swift nor easy.”
(Marie Curie)

Resumo

O câncer de mama é um dos tipos de câncer mais comuns entre as mulheres no Brasil e no mundo. O diagnóstico precoce de uma lesão na mama é prioridade para um tratamento eficaz e com menores custos, sendo a mamografia o exame mais utilizado. Nesta pesquisa de doutorado foi desenvolvido o método MUSA com o objetivo de classificar e anotar semanticamente imagens de mamografia, a partir da fusão de informações multimodais, fornecendo uma anotação mais completa da imagem do que o atual estado da arte. Para isso, a abordagem engloba um processo de mineração de texto, um processo de mineração de imagens, e um processo de engenharia de ontologias. Com o processo de mineração de texto é possível estabelecer as características contextuais, a partir de relações de similaridade e coocorrência entre pares de palavras presentes em um mesmo laudo. O processo de mineração de imagens permite classificar uma imagem de mamografia, usando uma estratégia *end-to-end deep learning*. No processo de engenharia de ontologias foi construída a ontologia AnotaMammo, composta pela ontologia de anotação e a ontologia de domínio Mammo. A AnotaMammo permite a anotação semântica das imagens de mamografia. Os resultados da mineração de texto foram utilizados para geração de regras semânticas, incluídas na AnotaMammo. Os resultados do processo de mineração de imagem superaram ou se compararam com estudos publicados na literatura, atingindo mais de 92% de acurácia na classificação de lesões como massa ou calcificação. Os resultados demonstram ainda que a ontologia AnotaMammo desempenhou apropriadamente o enriquecimento semântico da classificação, além de ter realizado adequadamente a fusão das informações multimodais. Por fim, o método MUSA agrega informações para tornar o resultado mais semântico, mais interpretável, reduzindo assim o gap semântico.

Palavras-chave: câncer de mama, anotação semântica, imagens médicas, aprendizado profundo, processamento de linguagem natural, ontologias, classificação.

Abstract

Breast cancer is one of the most common types of cancer among women in Brazil and worldwide. Early diagnosis of a breast lesion is a priority for effective treatment at lower costs, with mammography being the most used exam. In this doctoral research, was developed the MUSA method to classify and semantically annotate mammography images based on the fusion of multimodal information, providing a fuller image annotation than the current state of the art. For this, the approach includes a text mining process, an image mining process, and an ontology engineering process. With the text mining process, it is possible to establish the contextual features based on similarity and co-occurrence relations between pairs of words in the same medical report. The image mining process allows classifying a mammography image using an end-to-end deep learning strategy. In the ontology engineering process, we built the AnotaMammo ontology, composed of the annotation ontology and the Mammo domain ontology. AnotaMammo allows the semantic annotation of mammography images. The text mining results were used to generate semantic rules, included in AnotaMammo. The results of the image mining process surpassed or compared with studies published in the literature, reaching more than 92% accuracy for classifying lesions as mass or calcification. The results also demonstrate that the AnotaMammo ontology adequately performed the semantic enrichment of the classification, in addition to having adequately performed the fusion of multimodal information. Finally, the MUSA method adds information to make the result more semantic and interpretable, thus reducing the semantic gap.

Keywords: breast cancer, semantic annotation, medical imaging, deep learning, natural language processing, ontologies, classification.

Lista de ilustrações

Figura 1 – Projeção normal de uma mamografia	34
Figura 2 – Tipos de margem de massa	38
Figura 3 – Processo de KDD	39
Figura 4 – Extrato da ontologia Radlex	50
Figura 5 – Visão simplificada da <i>Mammography Annotation Ontology</i> (MAO) . . .	51
Figura 6 – Ferramenta de anotação de imagens AIM	53
Figura 7 – Curva ROC de dois modelos de classificação	56
Figura 8 – Região de Interesse na Mamografia	58
Figura 9 – Arquitetura do <i>Framework for Semantic Medical Image Retrieval</i> . . .	59
Figura 10 – Ontologia de Anotação	59
Figura 11 – Arquitetura do sistema para interpretação de imagem	60
Figura 12 – Relações entre as pontuações BI-RADS e características da massa, (a) lógica <i>crisp</i> e (b) lógica não <i>crisp</i>	61
Figura 13 – Extrato da ontologia RadLex	62
Figura 14 – <i>Framework</i> semântico proposto para recuperação de imagens médicas. .	63
Figura 15 – Método de anotação proposto em (KUMAR et al., 2016)	64
Figura 16 – Estratégias de classificação usada em (KUMAR et al., 2016)	64
Figura 17 – Anotação de imagem PET-CT usando o ePAD	65
Figura 18 – Método para a recomendação de cirurgia de catarata	65
Figura 19 – Ontologia geral que importa as regra TNM, axiomas, <i>moduleRadlex.owl</i> e <i>Onlira.owl</i>	67
Figura 20 – <i>Framework</i> para a anotação de imagem baseada em conhecimento e raciocinador multiestágio	68
Figura 21 – Método de classificação de imagens	68
Figura 22 – MUSA - Representação abstrata do método com as etapas <i>offline</i> e <i>online</i> para classificação e anotação semântica de imagens médicas . . .	75

Figura 23 – Representação abstrata do modelo estruturado de conhecimento. A raiz (<i>Thing</i>) da ontologia, ao lado esquerdo as classes e propriedades da ontologia de anotação e ao lado direito as classes e propriedades da ontologia de domínio.	78
Figura 24 – Arquitetura geral do método proposto. Construção de modelos de multimodalidades na etapa <i>offline</i> e utilizados na etapa <i>online</i> , para viabilizar a classificação com informação multimodal e a anotação semântica.	79
Figura 25 – Processo de mineração de texto para representação de palavras, identificação de coocorrências e a similaridade entre pares de palavras . . .	81
Figura 26 – Representação por cores de como ocorre o janelamento para identificação de coocorrência de palavras num mesmo laudo, formando os <i>Skipgrams</i>	82
Figura 27 – Processo de geração de Regras de Associação	84
Figura 28 – Processo de geração de Regras Semânticas	85
Figura 29 – Representação da arquitetura com o novo topo treinável. Camadas pré-treinadas da arquitetura CNN, com camadas treináveis e camadas congeladas	87
Figura 30 – Arquitetura após o processo <i>fine-tunning</i> com adição de novo topo . . .	88
Figura 31 – Construção do modelo e classificação binária de um ROI	89
Figura 32 – <i>Ensemble</i> produzido com os melhores modelos de cada uma das três CNNs, DenseNet121, EfficientNetV2, Xception	89
Figura 33 – Ontologia AnotaMammo	90
Figura 34 – Extrato da ontologia de anotação com as classes e propriedades para a anotação semântica	91
Figura 35 – Extrato da ontologia de domínio com classes e propriedades relacionados ao domínio — Ontologia Mammo	94
Figura 36 – Extrato da ontologia de domínio — Classe Mass	95
Figura 37 – Extrato da ontologia de domínio — Classe Calcification	96
Figura 38 – Extrato da ontologia de domínio — Classe Lymph_node	96
Figura 39 – Exemplos de ROIs de imagens de mamografia e os respectivos laudos .	102
Figura 40 – Gráfico para escolha do melhor otimizador para EfficientNetV2	113
Figura 41 – EfficientNetV2S - Gráfico com quatro valores diferentes de LR para determinar a melhor taxa de aprendizado	113
Figura 42 – Gráfico de acurácia e perda para $LR = 1e - 2$. Ao lado esquerdo da figura o gráfico de acurácia e ao lado direito o gráfico de perda	114
Figura 43 – Gráfico de acurácia e perda para $LR = 1e - 3$. Ao lado esquerdo da figura o gráfico de acurácia e ao lado direito o gráfico de perda	114
Figura 44 – <i>ROC Curve - Ensemble</i>	115
Figura 45 – EfficientNetV2 - Depois do ajuste do Otimizador e <i>Learning Rate</i> . . .	116

Figura 46 – Matriz de Confusão - classificação de benigno e maligno	118
Figura 47 – Matriz de Confusão - classificação de massa benigna (0), massa maligna (1), calcificação benigna (2), calcificação maligna (3)	119
Figura 48 – Processo de Anotação Semântica - ROI classificado no processo de mineração de imagem como massa e benigno, esse dado é entrada na AnotaMammo. O processo de anotação semântica anota o ROI com os conceitos simples obtidos na classificação inicial. Em seguida, conceitos complexos são sugeridos para possível anotação.	120
Figura 49 – Processo de Anotação Semântica - ROI classificado no processo de mineração de imagem como massa e maligno, esse dado é entrada na AnotaMammo. O processo de anotação semântica anota o ROI com os conceitos simples obtidos na classificação inicial. Em seguida, conceitos complexos são sugeridos para possível anotação.	121
Figura 50 – Anotação Semântica pela AnotaMammo do ROI_00099_LMLO_1. Anotações dos conceitos simples massa e benigno, e anotações dos conceitos complexos <i>circumscribed</i> , <i>round</i> e Birads 3	122
Figura 51 – Anotação Semântica pela AnotaMammo do ROI_00016_LMLO_1. No quadro detalhes da anotação	123
Figura 52 – Relações entre um indivíduo anotação e um indivíduo conceito. E os relacionamentos com <i>speculated_concept</i>	123

Lista de tabelas

Tabela 1 – Terminologia BI-RADS para descrição de massa e calcificação	36
Tabela 2 – Terminologia BI-RADS para descrição de outros achados	37
Tabela 3 – Categorias BI-RADS	38
Tabela 4 – Otimizadores baseados em gradiente	46
Tabela 5 – Matriz de confusão para o classificador M	55
Tabela 6 – Identificação dos critérios de comparação	71
Tabela 7 – Comparação dos trabalhos da literatura com a proposta de Doutorado.	71
Tabela 8 – Abreviações de diagnóstico histológico	103
Tabela 9 – Base Vienna. Exemplos de ocorrências de unigramas(a) e <i>skipgrams</i> (b)	105
Tabela 10 – Base CDD-CESM - <i>Mass description</i> . Exemplos de ocorrências de unigramas (a). Exemplos de ocorrências de <i>skipgrams</i> (b)	105
Tabela 11 – Base CDD-CESM - <i>Mass enhancement</i> . Exemplos de ocorrências de unigramas(a). Exemplos de ocorrências de <i>skipgrams</i> (b)	106
Tabela 12 – Base CDD-CESM - <i>Calcification</i> . Exemplos de ocorrências de unigramas (a). Exemplos de ocorrências de <i>skipgrams</i> (b)	106
Tabela 13 – Base CDD-CESM - <i>Mass Enhancement</i> . Similaridades com a palavra foco (<i>word</i>) <i>Benign</i> . Em (a) exemplos de palavras <i>context</i> e a similaridade (Sim), considerando a matriz de contagem de frequência. Em (b) os resultados de similaridade com a matriz PPMI suavizado. Em (c) os resultados com SVD	107
Tabela 14 – Base Vienna. Similaridades com a palavra foco (<i>word</i>) <i>Circumscribed</i> . Em (a) exemplos de palavras <i>context</i> e a similaridade (Sim), considerando a matriz de contagem de frequência. Em (b) os resultados de similaridade com a matriz PPMI suavizado. Em (c) os resultados com SVD	107

Tabela 15 – Base CDD-CESM — <i>Mass Enhancement</i> . Similaridades com a palavra foco (<i>word</i>) <i>Malignant</i> . Em (a) exemplos de palavras <i>context</i> e a similaridade (Sim), considerando a matriz de contagem de frequência. Em (b) os resultados de similaridade com a matriz PPMI suavizado. Em (c) os resultados com SVD.	108
Tabela 16 – Base CBIS-DDSM — <i>Mass</i> . Similaridades com a palavra foco (<i>word</i>) <i>Speculated</i> . Em (a) exemplos de palavras <i>context</i> e a similaridade (Sim), considerando a matriz de contagem de frequência. Em (b) os resultados de similaridade com a matriz PPMI suavizado. Em (c) os resultados com SVD.	108
Tabela 17 – Base CBIS-DDSM — <i>Calcification</i> . Similaridades com a palavra foco (<i>word</i>) <i>Malignant</i> . Em (a) exemplos de palavras <i>context</i> e a similaridade (Sim), considerando a matriz de contagem de frequência. Em (b) os resultados de similaridade com a matriz PPMI suavizado. Em (c) os resultados com SVD.	109
Tabela 18 – Aumento de Dados - transformação dos dados de treinamento em tempo de execução	112
Tabela 19 – Resultados de classificação	117
Tabela 20 – Acurácia e AUC	117
Tabela 21 – Acurácia - comparação com trabalhos correlatos	117
Tabela 22 – Resultados de classificação para benigno e maligno	117
Tabela 23 – Resultados de classificação	119

Lista de siglas

AIM *Annotation and Image Markup*

ACR *American College of Radiology*

AUC *Area under the ROC curve*

BCIO *Breast Cancer Imaging Ontology*

BI-RADS *Breast Imaging Reporting and Database System score*

BOW *Bag-of-Words*

BoVW *Bag-of-Visual-Word*

CAD *Computer-aided diagnosis*

CBIR *Content-based image retrieval*

CC *Cranio Caudal*

CDD-CESM *Categorized Digital Database for Low energy and Subtracted Contrast Enhanced Spectral Mammography images*

CNN *Redes Neurais Convolucionais*

DDSM *Digital Database for Screening Mammography*

DICOM *Digital imaging and communications in medicine*

DL *Deep Learning*

EHR *Electronic health records*

ES *Early Stopping*

ePAD *Electronic Physician Annotation Device*

FMA *Foundational Model of Anatomy*

GPU *Graphics Processing Unit*

HIS *Hospital information systems*

HMAX *Hierarchical Max-Pooling Model*

HSBD *Hierarchical Semantic-Based Distance*

KDD *Knowledge Discovery in Databases*

LR *Learning Rate*

MAO *Mammography Annotation Ontology*

MLO *Medio Lateral Obliquo*

MRI Magnetic resonance imaging

MUSA *Unified Method of Semantic Annotation*

NMD *National Mammography Database*

OWL *Web Ontology Language*

PACS *Picture archiving and communication system*

PET *Positron emission tomography*

PLN Processamento de Linguagem Natural

PMI *Pointwise Mutual Information*

PPMI *Positive Pointwise Mutual Information*

Radlex *Radiology Lexicon*

RDF *Resource Description Framework*

ReLU Rectified Linear Unit

RGB Red, Green and Blue

RIS *Radiology information systems*

ROI Região de Interesse

ROC *ROC Curve*

SPARQL *Protocol And RDF Query Language*

SQWRL *Semantic Query-enhanced Web Rule Language*

SVD *Singular Value Decomposition*

SVM *Support Vector Machine*

SWRL *Semantic Web Rule Language*

TC *Tomografia computadorizada*

TNM *Classification of Malignant Tumours*

W3C *World Wide Web Consortium*

WNN *weighted nearest-neighbour retrieval*

XML *Extensible Markup Language*

Sumário

1	INTRODUÇÃO	27
1.1	Contexto e Motivação	27
1.2	Questão de Pesquisa	29
1.3	Objetivos	29
1.4	Organização da Monografia	30
2	FUNDAMENTAÇÃO TEÓRICA	31
2.1	Informática na Saúde	31
2.1.1	Imagiologia Médica	31
2.1.2	Mamografia	33
2.2	Mineração de Dados Complexos	39
2.2.1	Mineração de Texto	41
2.2.2	Mineração de Imagem	42
2.3	Ontologia	47
2.3.1	Ontologias Biomédicas	49
2.4	Anotação Semântica	50
2.5	Métricas	53
2.6	Considerações Finais	56
3	TRABALHOS RELACIONADOS	57
3.1	Anotação Semântica em Imagens Médicas	57
3.2	Anotação Semântica em Outros Cenários	66
3.3	Classificação com <i>Deep Learning</i>	68
3.4	Comparação entre as Abordagens	70
3.5	Considerações Finais	72
4	MÉTODO MUSA	73

4.1	Considerações Iniciais	73
4.2	Método MUSA	74
4.2.1	Etapa <i>offline</i>	75
4.2.2	Etapa <i>online</i>	77
4.3	Arquitetura Geral da Proposta	78
4.4	Processo de Mineração de Texto	80
4.4.1	Pré-Processamento	81
4.4.2	Similaridade	82
4.4.3	Regras de Associação	83
4.4.4	Geração de Regras Semânticas	84
4.5	Processo de Mineração de Imagens	85
4.5.1	Arquiteturas CNN pré-treinadas	86
4.5.2	<i>Fine-Tuning</i>	87
4.5.3	<i>Hiperparâmetros</i>	87
4.5.4	Construção do Modelo e Classificação	88
4.5.5	<i>Ensemble</i>	89
4.6	Engenharia de Ontologia	89
4.6.1	Ontologia AnotaMammo	90
4.6.2	Ontologia de Anotação	90
4.6.3	Artefatos Semânticos	91
4.6.4	Regras Semânticas	92
4.6.5	Ontologia de Domínio	94
4.7	Considerações Finais	97
5	RESULTADOS E DISCUSSÕES	99
5.1	Implementação	99
5.2	Bases de Dados	100
5.2.1	CBIS-DDSM	100
5.2.2	CDD-CESM	100
5.2.3	MAMMOSET	100
5.2.4	VIENNA	101
5.3	Mineração de Texto	101
5.3.1	Pré-Processamento	101
5.3.2	Cálculo de Similaridade	107
5.3.3	Regras de Associação	109
5.4	Mineração de Imagem	111
5.4.1	Pré-Processamento	111
5.4.2	Hiperparametrização	111
5.4.3	<i>Ensemble</i>	115
5.4.4	Resultados	116

5.5	AnotaMammo	119
5.6	Considerações Finais	124
6	CONCLUSÃO	125
6.1	Principais Contribuições	125
6.2	Publicações	127
6.3	Trabalhos Futuros	127
6.4	Considerações Finais	128
	REFERÊNCIAS	129

Capítulo 1

Introdução

Neste capítulo são apresentados o contexto e a motivação desta pesquisa de Doutorado, o objetivo e a questão de pesquisa, o método de pesquisa e, por fim, a organização desta monografia de tese de doutorado.

1.1 Contexto e Motivação

O câncer de mama é o tipo de câncer mais comum entre as mulheres no Brasil e no mundo, atingindo cerca de 2,3 milhões de novos casos no mundo no ano de 2020. Estima-se que no Brasil surgirão mais de 73 mil novos casos de câncer de mama no Brasil no triênio 2023 – 2025, de acordo com o INCA (2022). As estimativas sugerem um aumento de mais de 10% de novos casos em relação às estimativas do triênio anterior (INCA, 2019), mostrando que as taxas de incidência são crescentes. A neoplasia de mama é o tipo de câncer com maior taxa de letalidade entre as mulheres. No entanto, possui bom prognóstico se identificado precocemente, na maioria dos casos.

A mamografia e o ultrassom são os principais exames de imagem para o diagnóstico clínico do câncer de mama, principalmente por terem menor custo e não serem exames invasivos. O diagnóstico precoce de uma lesão na mama, ajuda a reduzir os custos para a saúde pública e a melhorar as taxas de cura. Nesse sentido, é de fundamental importância a utilização de exames de imagem como estratégia de rastreamento do câncer, pois determinar o diagnóstico precocemente ajuda a tornar o tratamento mais barato. Aumenta a taxa de sucesso no tratamento e também ajuda a reduzir o número de biópsias realizadas desnecessariamente em alguns casos benignos. Políticas públicas como a de rastreamento mamográfico para mulheres de 50 a 69 anos, recomendada no Brasil, contribuem para a identificação de novos casos precocemente INCA (2022).

Cada vez mais as imagens médicas são utilizadas para apoiar o diagnóstico clínico de neoplasia mamária. E a quantidade de dados de imagens gerada também é crescente. Na análise clínica, a leitura e a interpretação desse tipo de imagem exige especialistas treinados e com experiência conquistada ao longo dos anos. Em vista disso, é necessário investir esforços no desenvolvimento de soluções computacionais inovadoras que busquem facilitar a rotina clínica de radiologistas e médicos especialistas envolvidos no diagnóstico do câncer de mama. Além disso, métodos computacionais inovadores podem contribuir com soluções que podem ser incorporadas aos sistemas educacionais para o treinamento de novos especialistas.

Muitos desafios estão envolvidos no desenvolvimento de um arcabouço computacional que auxilie o diagnóstico por imagem, principalmente no que se refere à redução do *gap* semântico e a melhoria da interpretabilidade dos resultados. Uma maneira de lidar com esses desafios é buscar soluções que promovam o enriquecimento semântico desses resultados.

Para isso, diferentes modalidades de informação sobre a imagem investigada podem ser integradas para fornecer informações adicionais, além das obtidas por características visuais da imagem. Por exemplo, o conjunto de informações textuais provenientes de relatórios médicos e diagnósticos é muito valioso semanticamente. As informações conceituais relacionadas ao domínio dos achados em mamografia também fornecem um rico conhecimento que pode ser agregado para auxiliar a interpretação de uma imagem. Usar adequadamente essas informações, pode ajudar a enriquecer semanticamente os métodos de auxílio ao diagnóstico, melhorando a interpretação dos resultados.

No entanto, a integração de diferentes modalidades de dados pode ser uma tarefa bastante desafiadora, fazendo-se necessário a investigação e proposição de novos métodos que consideram a fusão dessas informações. Ao integrar informações multimodais provenientes de diferentes fontes, por exemplo, do conteúdo da imagem, de laudos médicos, e de uma ontologia de domínio, ocorre o enriquecimento semântico que pode aprimorar métodos para auxílio ao diagnóstico clínico.

Este trabalho de doutorado visa lidar com a semântica e a multimodalidade de informações na tarefa de classificação e anotação de imagens de mamografia, fornecendo uma anotação mais completa da imagem do que o atual estado da arte. Espera-se que isso permita um maior uso e aceitação de ferramentas computacionais de anotação semântica por parte dos especialistas de domínio, facilitando e otimizando o trabalho dos mesmos.

A fusão das informações ocorre por meio de artefatos semânticos, como os axiomas e as regras semânticas implementadas na ontologia de anotação. As seguintes informações multimodais são analisadas: (i) características visuais do conteúdo da imagem e classificações obtidas por um processo *end-to-end deep learning*; (ii) características contextuais, obtidas por meio de informações presentes nos laudos e descrições das imagens; e (iii) características conceituais, obtidas por meio da ontologia de domínio que define a estrutura

semântica do conhecimento.

Segundo Bouchakwa, Ayadi e Amous (2020), apesar dos métodos de anotação baseados exclusivamente no conteúdo da imagem ajudarem a reduzir o *gap* semântico ao fornecer descrições sobre as imagens, apresentam geralmente resultados com poucos detalhes semânticos. Dessa forma, a anotação semântica baseada na multimodalidade de informações é uma alternativa que permite enriquecer semanticamente as descrições relacionadas ao conteúdo da imagem. Além disso, a combinação de ferramentas computacionais da área de ciência de dados e inteligência artificial é cada vez mais utilizada no desenvolvimento de métodos para auxiliar a área da saúde. Os resultados produzidos contribuem ao fornecer uma segunda opinião em relação ao diagnóstico. Assim, essas inovações podem ajudar e facilitar a rotina clínica, resultando em um aumento de produtividade no processo de diagnóstico de imagens médicas, gerando menores custos para a saúde pública e um melhor prognóstico para os pacientes.

1.2 Questão de Pesquisa

Nesta tese foi elaborada a seguinte questão de pesquisa:

”Como anotar de maneira mais completa imagens de mamografia, melhorando a interpretabilidade dos resultados da classificação?”

No desenvolvimento deste trabalho, a busca de resposta para a questão anterior nos direcionou a desenvolver um método que permite reduzir o *gap* semântico, por meio de anotação semântica de imagens de mamografia, com o uso de informações multimodais, provenientes das características visuais presentes nas imagens, das informações contextuais presentes nos laudos médicos e das relações conceituais presentes no conhecimento sobre o domínio.

1.3 Objetivos

O objetivo geral desta pesquisa de doutorado é propor um método para a classificação e anotação semântica de imagens de mamografia a partir de informações multimodais. Por meio do método proposto busca-se prover o enriquecimento semântico dos resultados de classificação com vistas a produzir uma anotação mais completa da imagem, melhorando sua interpretabilidade e reduzindo o *gap* semântico.

Os objetivos específicos estabelecidos são estudar, projetar e desenvolver:

- ❑ um processo de mineração de texto para lidar com os dados textuais presentes nos laudos e registros médicos.
- ❑ um processo de mineração de imagens para classificar imagens de mamografia a partir do conteúdo visual da imagem, ou seja, de suas características visuais.

- ❑ uma ontologia de domínio para representar o conhecimento semântico associado ao domínio das imagens médicas investigadas.
- ❑ um conjunto de regras semânticas para amparar o processo de anotação semântica das imagens.
- ❑ uma ontologia de anotação que realiza a fusão das informações e promove o processo de anotação semântica das imagens.

1.4 Organização da Monografia

Esta monografia de tese de Doutorado está organizada da seguinte maneira:

- ❑ No Capítulo 2 é apresentada a fundamentação teórica, em que são discutidos os principais temas relativos com esta pesquisa de Doutorado.
- ❑ No Capítulo 3 são apresentados os trabalhos relacionados com esta pesquisa.
- ❑ No Capítulo 4 são descritos o método proposto, incluindo o processo de mineração de texto, o processo de mineração de imagem, e o processo de engenharia de ontologia.
- ❑ No Capítulo 5 são apresentados os experimentos, resultados e discussões.
- ❑ No Capítulo 6 são apresentadas as conclusões e os trabalhos futuros.

Capítulo 2

Fundamentação

Neste capítulo são apresentados os principais conceitos envolvidos nesta pesquisa de doutorado. São descritos os aspectos gerais sobre informática na saúde. Conceitos principais sobre mineração de dados, mineração de texto e mineração de imagens. Uma visão geral sobre ontologia e anotação semântica. E as métricas de avaliação.

2.1 Informática na Saúde

É crescente o número de soluções computacionais para os diferentes segmentos da biomedicina. Com isso, a quantidade de dados gerados cresce exponencialmente e surgem inúmeras fontes de dados diferentes. Várias abordagens e métodos avançados vêm sendo desenvolvidos para dar suporte na rotina clínica e também para lidar com os diferentes tipos de dados. Algumas dessas tecnologias são: a imagiologia médica e os sistemas *Computer-aided diagnosis* (CAD), *Picture archiving and communication system* (PACS), *Content-based image retrieval* (CBIR), *Hospital information systems* (HIS), *Radiology information systems* (RIS), *Electronic health records* (EHR), entre outros.

2.1.1 Imagiologia Médica

A imagiologia médica é uma área imprescindível na medicina e está constantemente em evolução. É uma ferramenta indispensável em diversas subáreas da medicina, permite a captura, transmissão, análise de dados e auxilia para um diagnóstico médico com maior precisão (SCHAEFER; HASSANIEN; JIANG, 2009). Para uma imagem médica ser capturada é necessário haver uma exposição do corpo humano a alguma forma de energia e existem diferentes modalidades de exames de imagens para o diagnóstico médico. Por

exemplo, para os casos na área de dermatologia em que é necessária uma fotografia da pele lesionada, essa imagem médica é obtida por meio da luz visível. Já no caso da radiologia, é necessário que essa energia penetre nos tecidos, isso é possível por meio de ondas eletromagnéticas conhecidas como raios-x, essas ondas são de alta frequência e estão em um intervalo o espectro eletromagnético além da luz visível. Os raios-x foram descobertos por Wilhelm Roentgen em 1895, responsável pela primeira radiografia da anatomia humana. E, cada vez mais, é utilizado em exames de diagnóstico por imagem. As modalidades mamografia e Tomografia computadorizada (TC) também utilizam os raios-x. Algumas outras modalidades de imagens médicas são: Magnetic resonance imaging (MRI) que utiliza campos magnéticos, *Positron emission tomography* (PET) realizada por emissão de pósitrons após administração no corpo de material radioativo, além das imagens de ultrassom que são uma forma de alta frequência do som capaz de gerar imagens da anatomia humana (BUSHBERG; BOONE, 2011; SUETENS, 2017).

CAD

Os sistemas CAD tem se tornado cada vez mais frequentes nos procedimentos das práticas clínicas. Segundo Azevedo-Marques et al. (2017) nos Estados Unidos são estimados que cerca de 80% dos exames mamográficos sejam estudados com CAD. As pesquisas em CAD se iniciaram na década de 1980, principalmente sob o ponto de vista de utilização para o diagnóstico auxiliado por computador e não apenas como um diagnóstico automatizado, em que o diagnóstico final é determinado pelo radiologista. Um sistema CAD auxilia na tomada de decisão em relação ao diagnóstico, pois realiza análises quantitativas das imagens médicas e gera como resultado uma segunda opinião que pode ajudar a confirmar o diagnóstico e tornar o processo mais preciso. De acordo com Azevedo-Marques (2001), existem dois tipos de sistemas CAD: um que auxilia na detecção de lesões por meio da identificação de padrões anormais como, por exemplo, agrupamento de calcificações em mamografias; e o outro que auxilia o diagnóstico por meio da extração de características e a classificação como, por exemplo, a associação da quantidade e forma das calcificações com a malignidade da lesão. Geralmente, os sistemas CAD utilizam de métodos e técnicas das áreas de visão computacional e da inteligência artificial, para o processamento de imagem e extração de características e para o reconhecimento de padrões, respectivamente. O CAD pode ser aplicado nas diferentes modalidades de imagens médicas.

PACS

Os sistemas de comunicação e arquivamento de imagens, PACS, fornecem o armazenamento, a recuperação, a manipulação e a comunicação de imagens médicas e os dados associados a essa imagem (BANKMAN, 2008). Os sistemas PACS fornecem rápido acesso às imagens e permitem a integração de diferentes modalidades de imagens médicas. Esses sistemas também são responsáveis pela comunicação entre os sistemas nos ambientes médico-hospitalar, desde a aquisição da imagem digital por meio da comunicação com

os dispositivos de captura de imagem, até a comunicação com os sistemas de informação hospitalar e sistema de informação radiológica, RIS e HIS, realizando assim a integração e disponibilização dos dados do paciente (SALOMÃO; MARQUES, 2011). Isso permite a disseminação de determinado estudo para diversas estações de trabalho por toda a instituição, possibilitando que esse estudo seja analisado por diferentes indivíduos simultaneamente (BUI; TAIRA, 2009).

DICOM

Digital imaging and communications in medicine (DICOM) é o principal padrão de comunicação na radiologia digital. Por meio do DICOM é possível fornecer interoperabilidade de dados e comunicação entre componentes do PACS. O padrão DICOM está disponível online e não especifica a implementação em detalhes, no entanto, fornece as diretrizes e estrutura para comunicação e armazenamento de imagens médicas. O padrão DICOM é constantemente atualizado e pode ser adaptado conforme necessidade do usuário (BUI; TAIRA, 2009). Um sistema que utiliza o DICOM como padrão, permite controlar a aquisição, armazenamento, recuperação das imagens médicas no ambiente PACS (BANKMAN, 2008).

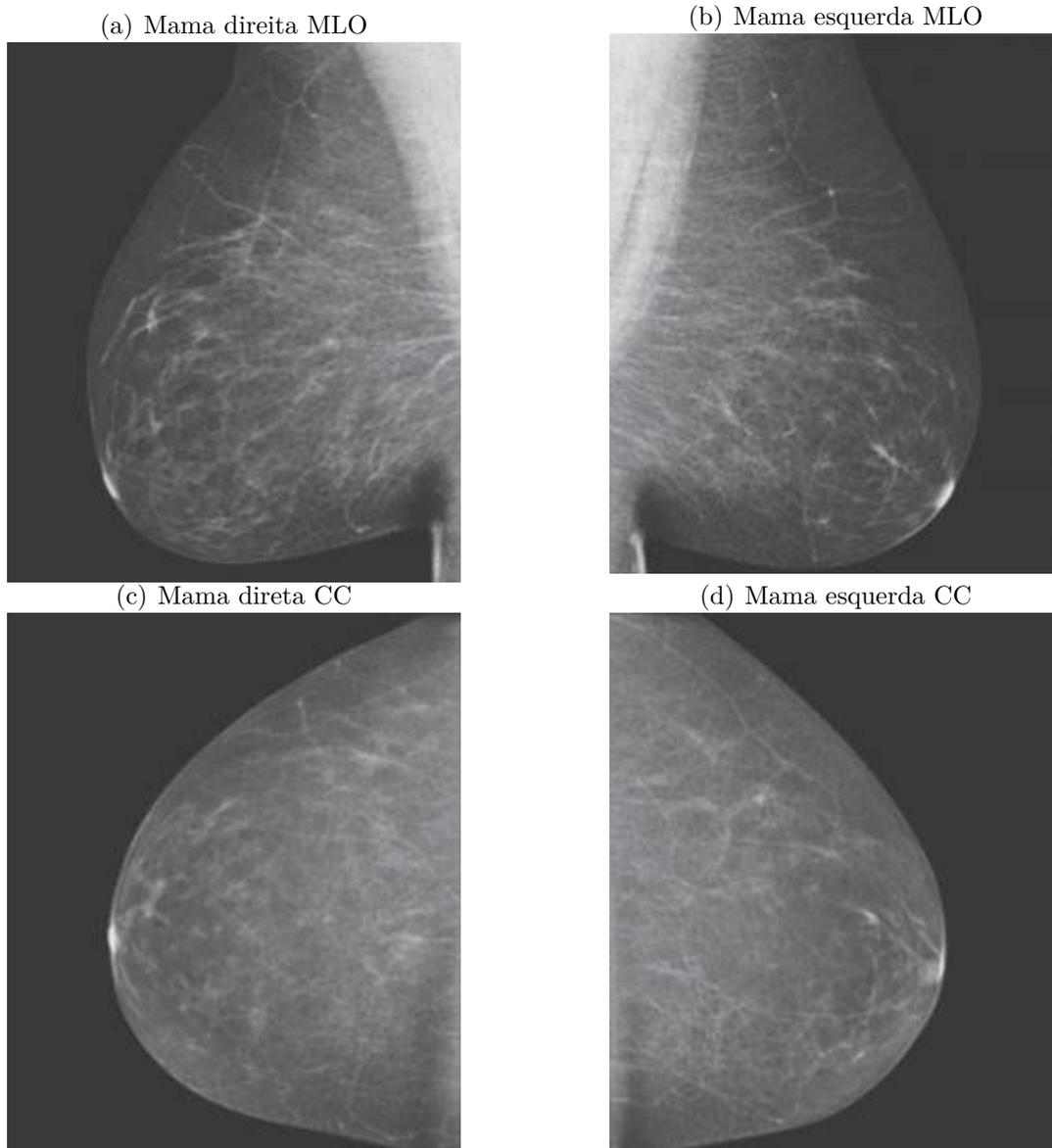
2.1.2 Mamografia

O exame de mamografia é o principal protocolo para o diagnóstico do câncer de mama, é um exame relativamente barato e ajuda na identificação de casos precoces. Colabora com a redução dos custos para a saúde pública e, principalmente, com as taxas de sucesso na resposta ao tratamento. As mamografias são um tipo de exame por imagem obtido pela exposição das mamas aos raios-x. Os dispositivos que captam essas imagens usam filme radiográfico ou produzem as imagens digitalmente, atualmente grande parte desses exames já são digitais.

Para a realização do exame é exercida uma compressão das mamas entre o receptor da imagem e as placas de compressão do dispositivo, isso é necessário para uniformizar a espessura da mama e garantir que ocorra uma expansão do tecido que passa a ser visualizado com maior nitidez e precisão. Evita-se, assim, que pequenas lesões sejam sobrepostas por tecido mamário e também que ocorra a movimentação da mama durante o exame, além disso, ajuda a reduzir a dose de radiação necessária na exposição (IKEDA; MIYAKE, 2016), (NG; WONG; CLARKE, 2018). As mamografias são realizadas no dispositivo conhecido como mamógrafo e num exame constam normalmente quatro imagens, mama esquerda e mama direita, nas visões craniocaudal (Cranio Caudal (CC)) e mediolateral-oblíqua (Medio Lateral Obliquo (MLO)), caso seja necessário o médico pode solicitar visões e exames adicionais para confirmação do diagnóstico. Um exemplo da projeção normal de uma imagem de mamografia pode ser visualizado na figura 1.

Análise das mamografias

Figura 1 – Projeção normal de uma mamografia



Fonte: Adaptado de Darby, Barron e Hyland (2011)

A análise de mamografias é uma tarefa complexa que exige vasta experiência da equipe médica. Por meio das análises das mamografias, os radiologistas detectam o câncer de mama ao identificar um tumor e achados clínicos diferentes de uma mamografia normal. Alguns dos achados mamográficos comuns ao câncer de mama são: calcificação pleomórfica, massa espiculada, massa contendo calcificação, entre outros. Para a realização da análise, os radiologistas estudam o conjunto padrão das mamografias que consistem nos pares de imagens CC e MLO.

As primeiras características a serem identificadas nas imagens são a simetria entre as mamas e a densidade da mama. Geralmente quando o tecido está normal as mamas apresentam simetria. No caso de uma assimetria em tecido normal, esse caso deve

ser acompanhado ao longo do tempo. Inclusive essa é uma estratégia essencial não só nesse caso, mas para todos os outros. O estudo de mamografias anteriores permite aos radiologistas identificar as mudanças ocorridas ano após ano.

Após analisarem a densidade e a assimetria, os radiologistas realizam uma busca sistemática em cada imagem a fim de buscar massas, calcificações e distorções. Um padrão comum utilizado é o zigue-zague ou dividir a imagens em tiras. As imagens também são divididas e ampliadas em quartis para facilitar a leitura e interpretação. Em seguida, esses achados recebem a descrição correspondente baseada no *Breast Imaging Reporting and Database System score* (BI-RADS).

Os passos descritos acima mostram alguns aspectos envolvidos no procedimento médico para a análise e diagnóstico do câncer de mama em imagens de mamografia, obviamente que se trata de uma breve descrição em alto nível desse procedimento, tendo em vista o escopo limitado apresentado neste trabalho. Muitas outras ferramentas são utilizadas pelos radiologistas para realizar a interpretação das mamografias e a tomada de decisão quanto ao diagnóstico, tais como histórico do paciente e fatores de risco, comparação com mamografias anteriores e sistemas CAD, se disponíveis (IKEDA; MIYAKE, 2016).

BI-RADS

O BI-RADS está na 5ª edição publicada em 2013 pela *American College of Radiology* (ACR) e, é o sistema padrão de descrição para os achados clínicos de patologias da mama identificados em exames de mamografia, ultrassom e ressonância magnética. O BI-RADS fornece uma terminologia padronizada para achados em imagens de mama, fornece uma padronização para relatórios médicos, estrutura de avaliação, sistema de classificação da lesão, estrutura para coleta de dados e auditoria. Essa padronização facilita a comunicação entre os médicos e instituições, permite que os médicos descrevam em seus relatórios os achados clínicos encontrados nas mamografias usando o léxico proposto, e utilizem a categorização do BI-RADS para a classificação final sobre o diagnóstico para determinado exame de mamografia (D'ORSI et al., 2013), (SPAK et al., 2017).

Na tabela 1 é apresentada a terminologia para descrição de massa e classificação. Como pode ser observado, uma massa tem forma, margem e densidade. Por exemplo, uma massa identificada no exame do paciente X possui uma margem do tipo circunscrita, isto é, uma massa que apresenta uma margem bem definida. Já os achados de calcificação podem caracterizar lesões tipicamente benignas, com morfologia suspeita e a distribuição dessa calcificação. Por exemplo, se o radiologista identifica uma calcificação do tipo pleomorfo, certamente esse caso precisa ser estudado em mais detalhes.

Na tabela 2 é apresentada a terminologia correspondente a outros achados clínicos. Tais como a composição da mama, por exemplo, a gordura e a densidade. E, também, as descrições dos achados relacionados à assimetria, características associadas, localização da lesão, informações sobre linfonodos e lesões na pele, entre outros.

Tabela 1 – Terminologia BI-RADS para descrição de massa e calcificação

<i>Masses</i>	
<i>Shape</i>	<i>Oval</i>
	<i>Round</i>
	<i>Irregular</i>
<i>Margin</i>	<i>Circumscribed</i>
	<i>Obscured</i>
	<i>Microlobulated</i>
	<i>Indistinct</i>
	<i>Spiculated</i>
<i>Density</i>	<i>High density</i>
	<i>Equal density</i>
	<i>Low density</i>
	<i>Fat containing</i>
<i>Calcifications</i>	
<i>Typically Benign</i>	<i>Skin</i>
	<i>Vascular</i>
	<i>Coarse or "popcorn-like"</i>
	<i>Large rod-like</i>
	<i>Round</i>
	<i>Rim</i>
	<i>Dystrophic</i>
	<i>Milk of calcium</i>
	<i>Suture</i>
	<i>Suspicious Morphology</i>
<i>Coarse heterogeneous</i>	
<i>Fine pleomorphic</i>	
<i>Fine linear or fine-linear branching</i>	
<i>Distribution</i>	<i>Diffuse</i>
	<i>Regional</i>
	<i>Grouped</i>
	<i>Linear</i>
	<i>Segmental</i>

Fonte: Adaptado de D'Orsi et al. (2013) e Ikeda e Miyake (2016).

Na tabela 3 podem ser observadas as categorias de avaliação da do BI-RADS de 2013. Conforme apresentado na tabela, a primeira categoria é 0, essa categoria significa que a definição está incompleta e é necessário informações adicionais para conclusão desse estudo. Na categoria 2 não ocorrem achados, na categoria 3 são os achados benignos que possuem valores menores ou iguais a 2% de probabilidade de malignância e nesses

Tabela 2 – Terminologia BI-RADS para descrição de outros achados

<i>Breast Composition</i>	<ul style="list-style-type: none"> a. <i>The breasts are almost entirely fatty</i> b. <i>There are scattered areas of fibroglandular density</i> c. <i>The breasts are heterogeneously dense, which may obscure small masses</i> d. <i>The breasts are extremely dense, which lowers the sensitivity of mammography</i>
<i>Architectural distortion</i>	
<i>Asymmetries</i>	<ul style="list-style-type: none"> <i>Asymmetry</i> <i>Global asymmetry</i> <i>Focal asymmetry</i> <i>Developing asymmetry</i>
<i>Associated features</i>	<ul style="list-style-type: none"> <i>Skin retraction</i> <i>Nipple retraction</i> <i>Skin thickening</i> <i>Trabecular thickening</i> <i>Axillary adenopathy</i> <i>Architectural distortion</i> <i>Calcifications</i>
<i>Location of lesion</i>	<ul style="list-style-type: none"> <i>Laterality</i> <i>Quadrant and clock face</i> <i>Depth</i> <i>Distance from the nipple</i>
<i>Intramammary lymph node</i>	
<i>Skin lesion</i>	
<i>Solitary dilated duct</i>	

Fonte: Adaptado de D’Orsi et al. (2013) e Ikeda e Miyake (2016).

casos, a paciente deve fazer o acompanhamento no intervalo de seis meses. A categoria 4 é subdividida em 4A, 4B e 4C, a categoria 4A apresenta baixa suspeita de malignância, a 4B suspeita moderada e a 4C alta suspeita, todas essas categorias têm recomendação de diagnóstico do tecido por meio da biópsia. A categoria 5 é altamente sugestiva para malignância e na categoria 6 a malignância já foi comprovada por biópsia, nesse caso a recomendação é cirúrgica se clinicamente apropriado.

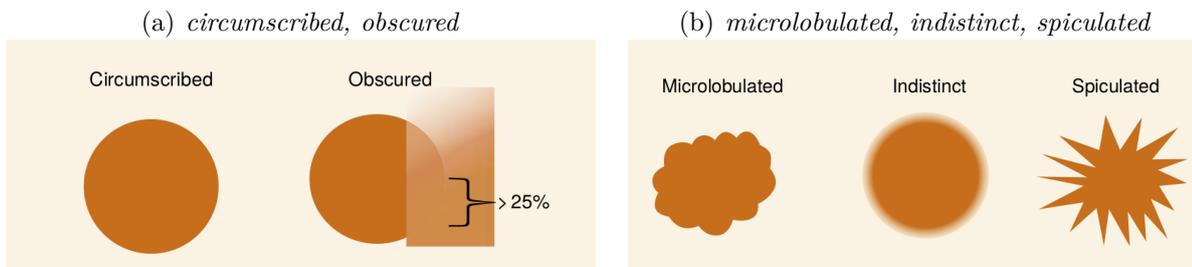
Na figura 2 é possível visualizar uma ilustração dos tipos de margem de massa.

Tabela 3 – Categorias BI-RADS

<i>BI-RADS Category</i>	<i>Definition</i>	<i>Likelihood of Malignancy</i>	<i>Management Recommendation</i>
0	<i>Incomplete</i>	N/A	<i>Recall for additional imaging and/or comparison with prior examination(s)</i>
1	<i>Negative</i>	0%	<i>Routine mammography screening</i>
2	<i>Benign</i>	0%	<i>Routine mammography screening</i>
3	<i>Probably benign</i>	>0% but ≤2%	<i>Short-interval (6-month) follow-up or continued surveillance mammography</i>
4A	<i>Low suspicion for malignancy</i>	>2% to ≤10%	<i>Tissue diagnosis</i>
4B	<i>Moderate suspicion for malignancy</i>	>10% to ≤50%	<i>Tissue diagnosis</i>
4C	<i>High suspicious for malignancy</i>	>50% to <95%	<i>Tissue diagnosis</i>
5	<i>Highly suggestive of malignancy</i>	≥95%	<i>Tissue diagnosis</i>
6	<i>Known biopsy-proven malignancy</i>	N/A	<i>Surgical excision when clinically appropriate</i>

Fonte: Adaptado de D'Orsi et al. (2013) e Ikeda e Miyake (2016).

Figura 2 – Tipos de margem de massa

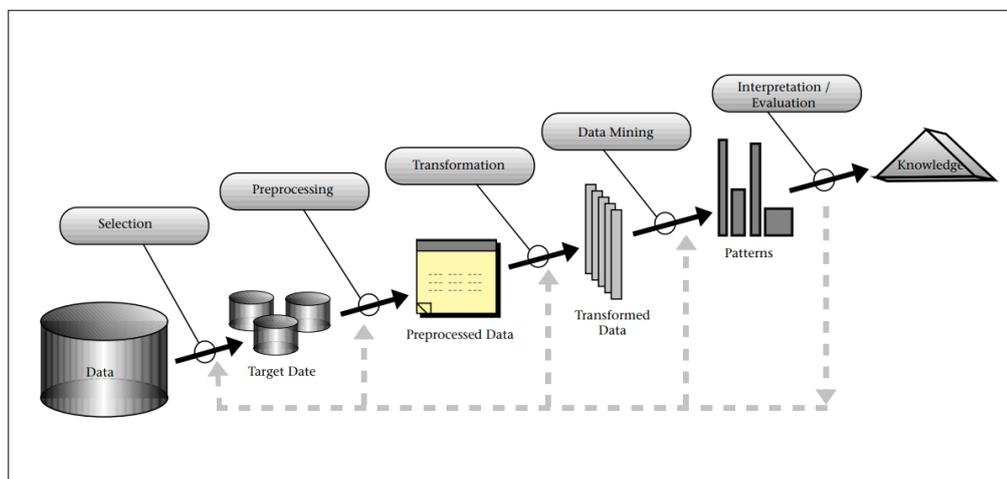


Fonte: Adaptado de Ikeda e Miyake (2016)

2.2 Mineração de Dados Complexos

A mineração de dados complexos é o processamento e descoberta de conhecimento em dados não convencionais, tais como imagem, texto, vídeo, entre outros. O processo pode utilizar de algoritmos de aprendizado de máquina para realização de tarefas como, por exemplo, a classificação. Segundo Fayyad et al. (1996), *Knowledge Discovery in Databases* (KDD) é um processo não trivial capaz de identificar padrões contidos nesses dados que sejam válidos, potencialmente úteis e compreensíveis. Conforme pode ser visualizado na figura 3 o processo conta com as seguintes etapas: seleção, pré-processamento, transformação, mineração, avaliação dos padrões e obtenção do conhecimento.

Figura 3 – Processo de KDD



Fonte: Fayyad et al. (1996).

Dados Complexos

Os dados complexos são aqueles dados que não possuem a propriedade de ordem total, portanto a utilização de operadores de comparação não é possível, operadores como '<', '>', '<=' e '>=' não podem ser utilizados nesse tipo de dado (BUENO et al., 2009). A representação dos dados complexos ocorre por meio de vetores de características produzidas por um processo da extração de características do conteúdo da imagem, texto ou vídeo. Para a comparação de um par de objetos complexos é necessária uma medida de similaridade ou função de distância. Nesse sentido, uma função de distância é aplicada sobre os vetores de características para o cálculo de similaridade entre eles. O espaço de similaridade pode ser formalizado como um espaço métrico (SAMET, 2006).

Aprendizado de Máquina

A ciência de programar computadores para que eles possam aprender a partir dos dados é conhecida como aprendizado de máquina ou *Machine Learning* (GÉRON, 2022). Conforme definição de Mitchell et al. (1997), o algoritmo aprende a partir da experiência (E), com respeito a classe de tarefas (T) e medida de desempenho (P), se o desempenho

em T medido por P melhora com E. Aprendizado de máquina é um conjunto de métodos, técnicas e ferramentas com a capacidade de aprender padrões a partir dos dados de maneira automatizada. Em seguida, utilizar os padrões descobertos para realizar previsão de dados futuros ou para outros tipos de tomada de decisão (MURPHY, 2012; ALPAYDIN, 2020). Os dois tipos principais ou categorias de aprendizado de máquina são: a abordagem preditiva ou aprendizado supervisionado e a abordagem descritiva ou aprendizado não supervisionado.

O aprendizado supervisionado é aquele na qual os dados de treinamento possuem rótulo ou classe previamente conhecida. A classificação e a regressão são tarefas de aprendizado supervisionado. Alguns dos principais algoritmos de aprendizado supervisionado são os classificadores: *k-Nearest Neighbors*, *Support Vector Machines (SVMs)*, *Decision Trees*, *Neural networks*.

No aprendizado não supervisionado os dados de treinamento não possuem rótulos previamente conhecidos. Por meio da análise não supervisionada é possível identificar padrões de interesse. Alguns dos principais algoritmos de aprendizado não supervisionado: agrupamento — *K-Means*, *DBSCAN*, *Hierarchical Cluster Analysis (HCA)*; redução de dimensionalidade — *Principal Component Analysis (PCA)*, *Kernel PCA*.

Classificação

A classificação é uma tarefa de aprendizado supervisionado cujo objetivo é atribuir uma classe aos dados cuja classe é desconhecida. A tarefa de classificação ocorre em duas etapas: o treinamento, em que ocorre a criação de um modelo; e o teste, em que esse modelo é testado. O treinamento desse classificador é supervisionado, ou seja, é necessário um conjunto de dados rotulados, o conjunto de treinamento, o qual é geralmente rotulado por especialistas de domínio. Na classificação de imagens são utilizados os vetores de características como os dados de treinamento. A partir do treinamento o modelo é criado de maneira que descreva o comportamento desses dados. Para o modelo poder ser avaliado quanto à classificação de novos dados, é realizada uma etapa de teste sobre dados rotulados que não fazem parte do conjunto de treinamento. Várias técnicas podem ser utilizadas para divisão dos dados de treinamento e teste, tais como *holdout*, *k fold*, *leave-one-out*, entre outros. Para a avaliação do modelo são utilizadas métricas de avaliação como, por exemplo, acurácia e medida F1. Se o modelo for adequado, pode ser utilizado para classificação de novos dados que não possuem um rótulo conhecido (WITTEN; EIBE; HALL, 2011), (HAN; PEI; KAMBER, 2011), (AGGARWAL, 2015). Alguns dos métodos tradicionais mais utilizados para a classificação são: *K-Nearest Neighbour*, *Logistic Regression*, *Naive Bayes*, *Neural Network*, *Random Forests*, *Support Vector Machine* e *Decision Tree*.

2.2.1 Mineração de Texto

A mineração de texto é um conjunto de métodos e ferramentas que conseguem obter conhecimento a partir de fontes de dados textuais não estruturados (REZENDE; MARCACINI; MOURA, 2011). O dado textual também é um tipo de dado complexo e necessita ser representado de forma que possa ocorrer extração de conhecimento. Inúmeras tarefas podem compor um processo de mineração de texto, uma tarefa a ser destacada é o pré-processamento do texto, pois na maioria das vezes é a tarefa mais custosa do processo. Alguns aspectos relacionados a mineração de texto e métodos de Processamento de Linguagem Natural (PLN), tais como o pré-processamento e *Pointwise Mutual Information* serão discutidos brevemente na sequência.

Pré-Processamento

A tarefa de pré-processamento é fundamental para os textos serem estruturados e possa ocorrer extração de conhecimento sobre eles. É necessário que os textos escritos em linguagem natural possam ser representados de maneira estruturada e que seja legível para algoritmos computacionais. Para isso, essa tarefa envolve um conjunto de etapas para o tratamento e padronização dos textos, seleção de palavras e representação dos dados (REZENDE; MARCACINI; MOURA, 2011).

Frequentemente os documentos que compõem a base de dados apresentam variados formatos, por isso, é necessário a padronização desses documentos, são frequentemente padronizados para o texto sem formatação. Além disso, são geralmente removidas as pontuações presentes no texto. Após isso, ocorre a tokenização das palavras, sendo a marcação e a separação de cada palavra contida no texto.

A seleção das palavras é outra tarefa essencial para lidar com a alta dimensionalidade na análise de textos. Para isso, primeiramente ocorre a eliminação dos *stop words*, que são aquelas palavras comuns que ocorrem frequentemente nos diferentes idiomas. Essas palavras não acrescentam informação semântica e devem ser excluída, esse conjunto de palavras compõe uma *stoplist*. Posteriormente, pode ser aplicada a técnica de *stemming*, a qual é uma técnica de transformação da palavra para sua raiz ou radical (KAO; POTEET, 2007), (BIRD; KLEIN; LOPER, 2009).

A seleção de palavras também pode ocorrer por meio de medidas estatísticas, como a frequência de palavra ou *term frequency* e a frequência de documento, *document frequency*. A frequência de palavra determina quantas vezes a palavra ocorre na base de dados e a frequência de documento determina o número de documentos em que determinado palavra aparece (REZENDE; MARCACINI; MOURA, 2011).

Para que os dados textuais possam ser processados por algoritmos de aprendizado de máquina, é necessário que sejam representados de maneira adequada. A representação mais utilizada é o modelo espaço vetorial. Os documentos são os vetores no espaço multidimensional e cada palavra é uma dimensão (REZENDE; MARCACINI; MOURA,

2011).

A representação também pode ser uma matriz simétrica **palavra-contexto**, ao utilizar uma abordagem tradicional de similaridade distribucional. Nesse caso, as palavras são representadas como vetores esparsos e cada dimensão representa um contexto. Ao analisar as relações entre palavra e contexto, é possível determinar a frequência de co-ocorrência entre as palavras por contagem e por *Pointwise Mutual Information* (PMI) (HOCKENMAIER, 2023).

Pointwise Mutual Information

A informação mútua é uma medida de sobreposição entre duas variáveis aleatórias; compara a probabilidade de observar dois pontos (palavras) x e y juntos (probabilidade conjunta) com a probabilidade de observar x e y independentemente, a sobreposição de informações de x e y é 0 quando as duas variáveis são independentes. O PMI é uma medida baseada na Informação Mútua que mede a probabilidade de coocorrência entre dois conceitos $p(x, y)$ considerando que são independentes. A medida é utilizada para medir a similaridade contextual entre dois conceitos, essa informação contextual pode ajudar a inferir conhecimento nas imagens. A similaridade contextual é uma medida dependente do *corpus* (CHURCH; HANKS, 1990; BOUMA, 2009; BANNOUR, 2013; LEVY; GOLDBERG; DAGAN, 2015).

Regras de Associação

As regras de associação são uma implicação do tipo antecedente e consequente. Também são conhecidos como corpo ou causa, para antecedente, e de cabeça ou consequência, para consequente. Tanto o antecedente quanto o consequente são constituídos por *itemsets*. Um *itemset* pode ser formado por um único item ou um conjunto de itens, mas nunca vazio. As métricas de suporte e a confiança são utilizadas para determinar a relevância e a frequência da regra (JUNIOR, 2018). Essas métricas são discutidas mais adiante.

O algoritmo Apriori é amplamente utilizado para a geração de regras de associação. A ideia principal do algoritmo é identificar itemsets frequentes e combiná-los internamente de maneira a produzir as regras. O algoritmo Apriori tem duas etapas principais: identificar os *itemsets* frequentes que atendam ao suporte mínimo e, a partir desses *itemsets*, gerar as regras de associação que atendam à confiança mínima. (AGRAWAL; SRIKANT et al., 1994; AGRAWAL; IMIELIŃSKI; SWAMI, 1993; SRIKANT; AGRAWAL, 1995).

2.2.2 Mineração de Imagem

As imagens são um tipo de dado complexo e necessitam de técnicas de análise, extração de características e armazenamento. Muitas ferramentas estão envolvidas no processo de mineração de imagens, tais como algoritmos de processamento de imagem para a segmentação, os extratores de características, os algoritmos de classificação, entre outros.

Nos métodos tradicionais de classificação de imagem, a representação das imagens por meio dos vetores de características permite que elas se tornem interpretáveis pelo computador e possam ser comparadas com seus pares. Alguns dos métodos tradicionais mais utilizados para a classificação de imagens são: *K-Nearest Neighbour*, *Logistic Regression*, *Naive Bayes*, *Neural Network*, *Random Forests*, *Support Vector Machine*, *Decision Tree*.

Além dos métodos tradicionais, cada vez mais as técnicas de *deep learning* são utilizadas para classificação de imagens. Num processo *end-to-end deep learning*, os algoritmos de DL realizam a extração automática das características durante o processo de treinamento. Em seguida, ocorre a classificação.

A representação computacional de uma imagem muitas vezes difere a percepção humano em relação a um conceito representado na imagem, isso cria um problema bastante conhecido na mineração de imagens que é o gap semântico.

Gap Semântico

O *gap* semântico é a diferença entre a percepção humana sobre um conceito e como ele é representado por computadores. É necessário diminuir essa lacuna envolvendo as características visuais e o conteúdo semântico (ESTRELA; HERRMANN, 2016). Algumas informações da imagem não podem ser capturadas pelos algoritmos de processamento de imagem e isso gera uma diferença entre a informação de alto nível esperada pelo usuário e a informação resultante proveniente das características de baixo nível. Os autores Akgül et al. (2011) defendem que as anotações são descritores poderosos capazes de auxiliar no preenchimento dessa lacuna. Outra técnica muito utilizada, que apresenta resultados promissores é o *feedback* de relevância fornecido pelo usuário, da mesma maneira que ocorre com as anotações, o *feedback* também relaciona conhecimento semântico para melhorar os resultados. Todavia, conforme descrevem os autores Kurtz et al. (2014b), esse problema permanece sem solução e novos métodos precisam ser desenvolvidos com o intuito de aproximar o conteúdo da imagem com informações semânticas e tornar o resultado mais satisfatório para o usuário final.

Deep Learning

Deep Learning (DL) permite que modelos computacionais multicamadas aprendam com dados em variados níveis de abstração. É uma técnica que consiste em múltiplas camadas de redes neurais que extraem características de maneira progressiva entre as camadas. Todo o sistema profundo pode ser ajustado com o *backpropagation* (LECUN; BENGIO; HINTON, 2015). Além disso, DL permite a entrada de dados brutos sem que tenham sido previamente transformados em vetores de características, ao contrário das técnicas tradicionais de aprendizado de máquina e mineração de dados.

Os algoritmos de DL possuem duas etapas, a extração automática de características durante o processo de treinamento e a classificação. A partir de uma imagem de en-

trada, várias camadas ocultas vão extrair as características dessa imagem. Essas camadas ocultas são construídas de maneira hierárquica uma sobre as outras, sendo as camadas inferiores detectam características mais simples e de mais baixo nível e as camadas superiores detectam características mais complexas. Por fim, uma camada de saída classifica e atribui um rótulo de classe para a imagem (GOODFELLOW; BENGIO; COURVILLE, 2016; WANI et al., 2020). DL pode ser utilizado para variadas aplicações, tais como: classificação de imagem médicas, reconhecimento de fala, processamento de linguagem natural, entre outras tarefas (GOODFELLOW; BENGIO; COURVILLE, 2016).

Redes Neurais Convolucionais (CNN) são redes neurais que usam convolução no lugar da multiplicação de matrizes em pelo menos uma de suas camadas (GOODFELLOW; BENGIO; COURVILLE, 2016). CNN é uma técnica de aprendizado profundo projetada para processar dados complexos, por exemplo, imagens. Uma CNN típica é composta pelas seguintes camadas:

- ❑ *Convolutional layer*
- ❑ *Activation Function layer*
- ❑ *Pooling layer*
- ❑ *Fully Connected layer*
- ❑ *Dropout layer*

A camada convolucional e a camada de função de ativação são empilhadas juntas, seguidas opcionalmente por uma camada *pooling*. Essas camadas atuam na etapa de extração de características. A camada totalmente conectada é a camada de saída que calcula a classe da imagem de entrada. A etapa de classificação pode conter uma ou mais dessas camadas, seguida por uma camada com a função softmax ou a função sigmoid, no caso de classificação binária. Opcionalmente, pode ser incluída uma camada *batch normalization* uma camada *dropout* para lidar com o *overfitting* (GOODFELLOW; BENGIO; COURVILLE, 2016).

A camada convolucional usa a operação de convolução. Seus parâmetros são um conjunto de filtros conhecidos como *kernels*, e sua tarefa é detectar características locais e gerar o mapa de características. A camada de função de ativação é responsável por gerar o mapa de ativação a partir do mapa de características.

Rectified Linear Unit (ReLU) é a função de ativação mais utilizada do DL. A camada *pooling* é opcionalmente usada para reduzir a dimensionalidade da entrada e, assim, reduzir o número de parâmetros. A camada totalmente conectada calcula a probabilidade das características detectadas pertencerem a uma classe específica.

Batch normalization normaliza as entradas para a próxima camada da rede. Normalmente ocorrem modificações nos valores de saída da camada anterior alterando sua

distribuição. Portanto, uma camada *batch normalization* pode melhorar o tempo de treinamento, ao evitar que esses valores se propaguem e causem mudanças significativas na distribuição das entradas para a próxima camada (WANI et al., 2020; GOODFELLOW; BENGIO; COURVILLE, 2016).

Dropout é uma técnica de regularização que visa eliminar neurônios de uma rede de treinamento. Desconectar uma unidade da rede também desconecta todas as suas conexões. Eventualmente, uma camada *dropout* pode descartar aleatoriamente alguns neurônios e suas conexões para atenuar o *overfitting* e melhorar a generalização do modelo (GOODFELLOW; BENGIO; COURVILLE, 2016). Essa estratégia reduz a rede e força a rede a aprender características mais robustas (WANI et al., 2020).

Hiperparametrização

A hiperparametrização é o ajuste de parâmetros fornecidos como argumento para a arquitetura da CNN. Naturalmente, os resultados diferem para diferentes valores de parâmetros, e um ajuste adequado pode melhorar o desempenho do treinamento e classificação (TAJBAKHSI et al., 2016).

Existem vários hiperparâmetros descritos na literatura. Entre eles, existe o otimizador baseado em gradiente que é um parâmetro de treinamento da rede. Na Tabela 4 alguns otimizadores baseados em gradiente são descritos, Adadelta, Adagrad, Adam, Adamax, Ftrl, Nadam, RMSprop e SGD.

Transfer Learning e Fine-Tuning

Transfer Learning ou aprendizagem por transferência permite usar o conhecimento descrito por meio de características, pesos ou parâmetros de modelos previamente treinados para treinar novos modelos. Os modelos são geralmente pré-treinados a partir de grandes conjuntos de imagens naturais, como o ImageNet (DENG et al., 2009). Com a aprendizagem por transferência, os pesos aprendidos na rede pré-treinada são transferidos para a rede não treinada (WANI et al., 2020).

O modelo pré-treinado pode ser utilizado como um extrator de características. Para isso, obtém-se o vetor de características da saída de uma das camadas da rede que seja anterior à camada de saída. Também é possível substituir a camada de classificação totalmente conectada por uma nova camada de classificação, mantendo o restante da rede como um extrator de características e apenas a nova camada é treinada para o novo conjunto de dados. Nesse caso, apenas o classificador é ajustado.

O processo de *fine-tuning* permite ajustar um modelo pré-treinado para um novo modelo. Diferentes abordagens são possíveis para manipulação da rede pré-treinada. Entre elas, é possível ajustar todos os pesos de toda a rede para o novo conjunto de dados. Também é possível congelar algumas camadas da rede e treinar somente as outras camadas para o novo conjunto de dados. E, também, é possível criar novas camadas com diferentes filtros (PONTI; COSTA, 2018).

Tabela 4 – Otimizadores baseados em gradiente

Otimizador	Descrição
Adadelta (ZEILER, 2012)	É uma extensão do método Adagrad. É um método de descida de gradiente estocástico baseado na taxa de aprendizado por dimensão. Utiliza apenas informações de primeira ordem e a sobrecarga computacional é trivial, provando ser um método robusto.
Adagrad (DUCHI; HAZAN; SINGER, 2011)	É um método que incorpora dinamicamente o conhecimento obtido na iterações anteriores. As taxas de aprendizado são adaptadas durante o treinamento.
Adam (KINGMA; BA, 2014)	É um método de otimização estocástica que usa gradientes de primeira ordem e é computacionalmente eficiente, com baixo consumo de memória.
Adamax (KINGMA; BA, 2014)	É uma extensão do método de Adam, baseado na norma do infinito. Esse otimizador geralmente é superior ao Adam.
Ftrl (MCMAHAN et al., 2013)	É um método de otimização baseado em descida de gradiente online e suporta regularização L2 online, ideal para espaços de características esparsos. O modelo foi desenvolvido pelo Google e é computacionalmente eficiente.
Nadam (DOZAT, 2016)	É uma extensão do método Adam com momento de Nesterov. Este otimizador demonstra geralmente resultados superiores a Adam.
RMSprop (HINTON; SRIVASTAVA; SWERSKY, 2012)	É um método que usa momento simples. A ideia é manter uma média móvel do quadrado do gradiente e dividir o gradiente pela raiz da média.
SGD (BOTTOU, 2010)	É um método estocástico de otimização de gradiente descendente com momento. Associa o risco de convergir para o mínimo local em vez do mínimo global, que pode ser corrigido com uma taxa de aprendizado adequada.

Ao congelar as camadas, ou seja, ao atribuir como argumento uma taxa de aprendizado igual a 0, essas camadas não são retreinadas, ocorrendo a transferência de pesos aprendidos a partir do modelo pré-treinado. Dessa forma, os pesos extraídos de uma rede pré-treinada são transferidos para a rede não treinada, evitando assim que a rede seja treinada do zero (WANI et al., 2020). Isso ajuda a reduzir o tempo de treinamento e pode contribuir para melhorar os resultados de classificação.

Geralmente ao substituir a camada classificadora e ajustar as últimas camadas da rede é suficiente para bons resultados. Contudo, pode ser necessário o ajuste das camadas iniciais, no caso de ser grande a distância entre o novo conjunto de dados e a base utilizada no treinamento. O uso do modelo pré-treinado com um apropriado *fine-tuning* pode igualar ou superar o desempenho da rede treinada do zero (TAJBAKSHI et al., 2016).

2.3 Ontologia

Na Filosofia, Ontologia é definida como a ciência que estuda a natureza do ser e o significado da existência. A natureza e as estruturas envolvidas são descritas em termos de categorias, relações e atributos. Para a Ciência da Computação, uma ontologia é um tipo especial de artefato computacional ou objeto de informação (GUARINO; OBERLE; STAAB, 2009). Uma ontologia define um conjunto de primitivas representacionais com as quais é possível modelar um domínio de conhecimento ou discurso (GRUBER, 2009). De modo geral, tais primitivas são hierarquias de conceitos (taxonomias), atributos, relacionamentos, indivíduos e restrições (axiomas), que descrevem o conhecimento de um determinado domínio de modo consensual e compartilhado (GUARINO, 1998).

Pela definição de Gruber (1993), Borst (1997) e Studer, Benjamins e Fensel (1998): "Uma ontologia é uma especificação formal e explícita de uma conceitualização compartilhada". Formal pois é descrita usando mecanismo formal e lógica matemática, assim permite a interpretação por máquinas. Explícita, pois conceitos, propriedades, relacionamentos e axiomas devem ser explicitamente definidos. A conceitualização é o modelo abstrato do mundo real. E é compartilhada, pois esse conhecimento é consensual entre a comunidade de usuários da ontologia.

De acordo com Noy e McGuinness (2001), uma ontologia é uma descrição formal de conceitos ou classes em um domínio de discurso. As classes podem estar organizadas na forma de uma taxonomia, permitindo subclasses e superclasses. As propriedades descrevem as características e os atributos de cada conceito, e, podem ser divididas em propriedades de objetos (*Object Property*) e propriedades de tipos dados (*Data Type Property*). A primeira determina relações binárias entre indivíduos ou classes. A segunda relaciona um indivíduo a um dado literal. As restrições sobre os conceitos e propriedades expressam significado de forma que máquinas consigam interpretar por meio de raciocínio automático. Por fim, os indivíduos ou instâncias de classes, junto à ontologia, constituem a base de conhecimento.

Por representarem informação semântica e permitirem inferência de conhecimento, ontologias tem sido utilizadas em diversas aplicações para facilitar a comunicação tanto entre humanos quanto entre os sistemas computacionais (YAGUINUMA; SANTOS; BI-AJIZ, 2007).

Representação de ontologias

Resource Description Framework (RDF) é a especificação proposta pelo *World Wide Web Consortium* (W3C) para descrever metadados ou recursos. É uma extensão de *Extensible Markup Language* (XML), portanto, herda sua sintaxe. Recursos são quaisquer coisas, tanto concretas quanto abstratas. Uma afirmação RDF consiste em três elementos (uma tripla), com a seguinte estrutura: sujeito – predicado – objeto.

O Esquema RDF é uma extensão do RDF, conhecido como *RDF Schema* ou *RDF-*

S que possibilita descrição semântica. Amplia a expressividade do RDF para prover mecanismos de descrição de taxonomias entre recursos e suas propriedades. Embora possa ser usado para descrever ontologias, tem algumas limitações, especialmente para apoiar o raciocínio computacional (LASSILA; SWICK et al., 1998).

A linguagem de representação de ontologia, *Web Ontology Language* (OWL), também é uma linguagem desenvolvida e aprovada pelo W3C, OWL 1 (2004) e OWL 2 (2012). É a linguagem mais utilizada para representar ontologias formalmente e baseada nas especificações do RDF/RDF-S.

OWL é uma linguagem declarativa que descreve um determinado universo do discurso de forma lógica. OWL é baseada em *Description Logics*. *Description Logics* são lógicas e possuem uma semântica formal. A semântica formal permite que os seres humanos e sistemas de computador possam intercambiar ontologias *description logics* sem ambiguidade quanto ao seu significado. Possibilita usar dedução lógica para inferir informações adicionais dos fatos expostos explicitamente em uma ontologia.

Uma ontologia especificada em OWL consiste em um conjunto de afirmações (axiomas), separadas em três grupos:

- A **Tbox** descreve os relacionamentos entre as classes e os relacionamentos entre os indivíduos.
 - Por exemplo, a classe "Person" é equivalente à classe "Human" significa que as duas classes possuem o mesmo conjunto de indivíduos.
- A **Abox** captura conhecimento sobre os indivíduos, ou seja, as classes as quais eles pertencem e como eles se relacionam entre si.
 - Por exemplo, o indivíduo "Mary" é da classe "Person".
- A **Rbox** contém as afirmações sobre as propriedades, ou seja, metapropriedades.
 - Por exemplo, transitividade, simetria, etc. Podem permitir que inferências sejam feitas.

Para a realização de consultas em ontologias a linguagem padrão é *Protocol And RDF Query Language* (SPARQL). É uma recomendação da W3C e é a linguagem padrão para consulta em dados baseados em grafos com triplas RDF. Porém, o SPARQL permite apenas consulta e não inferência sobre a ontologia.

Para realizar inferência é necessário um raciocinador, que permite inferir consequências lógicas a partir de um conjunto de fatos ou axiomas. Checa a consistência da ontologia, se alguma classe não pode ser instanciada ou se é contraditória. Infere tipos ao computar equivalência entre classes a partir das propriedades. E, suporta regras que podem ser escritas nas linguagens *Semantic Web Rule Language* (SWRL) e *Semantic Query-enhanced*

Web Rule Language (SQWRL). Exemplos de raciocinadores são *HermiT*, *Pellet*, entre outros.

As regras desenvolvidas na linguagem SWRL são um complemento para ontologia e permitem expressar e inferir conhecimentos diferentes daqueles que são possíveis de serem expressos na ontologia desenvolvida em OWL. SWRL adiciona o suporte a cláusulas no formato de Horn e permite inferir novos conhecimentos sobre os indivíduos OWL. Já SQWRL fornece operações *SQL-like* para recuperar informações a partir de uma ontologia OWL (HORROCKS et al., 2004), (O’CONNOR; DAS, 2009).

O *framework* Protégé¹, é uma ferramenta *open source* para desenvolver, gerenciar e realizar inferência em ontologias. Trabalha com várias linguagens, principalmente OWL. Foi desenvolvida pelo *Stanford Center for Biomedical Informatics Research* da escola de medicina da Universidade de Stanford. Seu código é em Java e pode ser facilmente estendido por *plugins*, conta com uma grande comunidade de desenvolvedores.

2.3.1 Ontologias Biomédicas

As ontologias são muito utilizadas no domínio biomédico, nesse domínio existe a necessidade de determinar com clareza terminologias, vocabulários médicos e procedimentos rotineiros na prática clínica. Assim, o uso de ontologias permite expressar essa riqueza de terminologias e permite a interoperabilidade semântica entre terminologia médica e sistemas clínicos.

Segundo Ivanović e Budimac (2014) além das ontologias, a classificação da terminologia na medicina pode ocorrer de outras maneiras, entre elas o vocabulário controlado, que é uma lista fechada de termos, em que cada termo específico é atribuído a um determinado conceito. A taxonomia representa uma estrutura abstrata e permite a classificação dos termos em uma hierarquia com relacionamentos pai-filho. O tesauro é uma extensão da taxonomia e permite relacionamentos associativos, fornece um vocabulário mais rico que as taxonomias.

Existem inúmeras ontologias desenvolvidas para o domínio biomédico, muitas delas disponíveis no BioPortal². Alguns exemplos são descritos a seguir.

Radlex

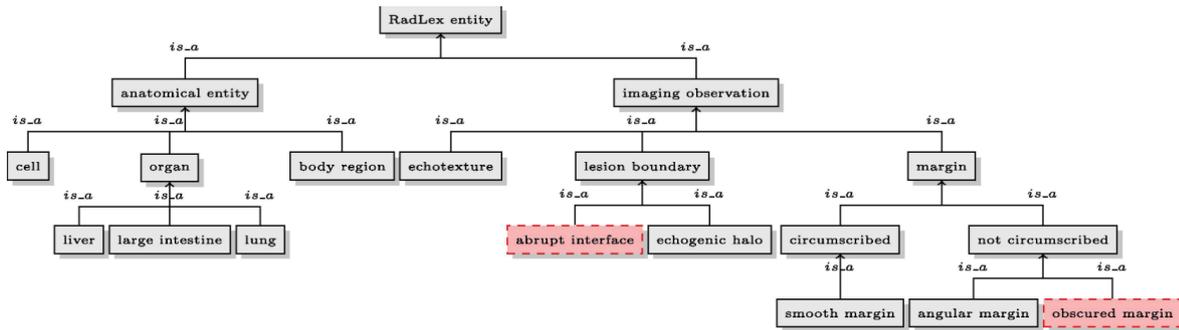
O projeto Radlex³ da *Radiological Society of North America’s* (RSNA’s) foi desenvolvido para expressar o índice para diagnósticos radiológicos da ACR, conhecido como *ACR Index*. RadLex é uma ontologia de termos radiológicos e terminologia da prática radiológica, contém mais de 8.000 termos anatômicos e patológicos e possibilita a indexação de imagens com conceitos específicos. Na figura 4 é possível visualizar um extrato da ontologia Radlex.

¹ protege.stanford.edu/

² bioportal.bioontology.org/

³ radlex.org/

Figura 4 – Extrato da ontologia Radlex



Fonte: Kurtz et al. (2014a).

Foundational Model of Anatomy Ontology (FMA)

A *Foundational Model of Anatomy* (FMA) é uma ontologia abrangente do domínio biomédico que determina conceitos e relações anatômicas. A ontologia FMA possui quatro componentes inter-relacionados: *Anatomy taxonomy* que classifica entidades anatômicas; *Anatomical Structural Abstraction* que especifica alguns relacionamentos existentes entre as entidades; *Anatomical Transformation Abstraction* que especifica a transformação morfológica; *Metaknowledge* que especifica os princípios, regras e definições para a representação dos três componentes anteriores (ROSSE; MEJINO, 2008).

Mammography Annotation Ontology (MAO)

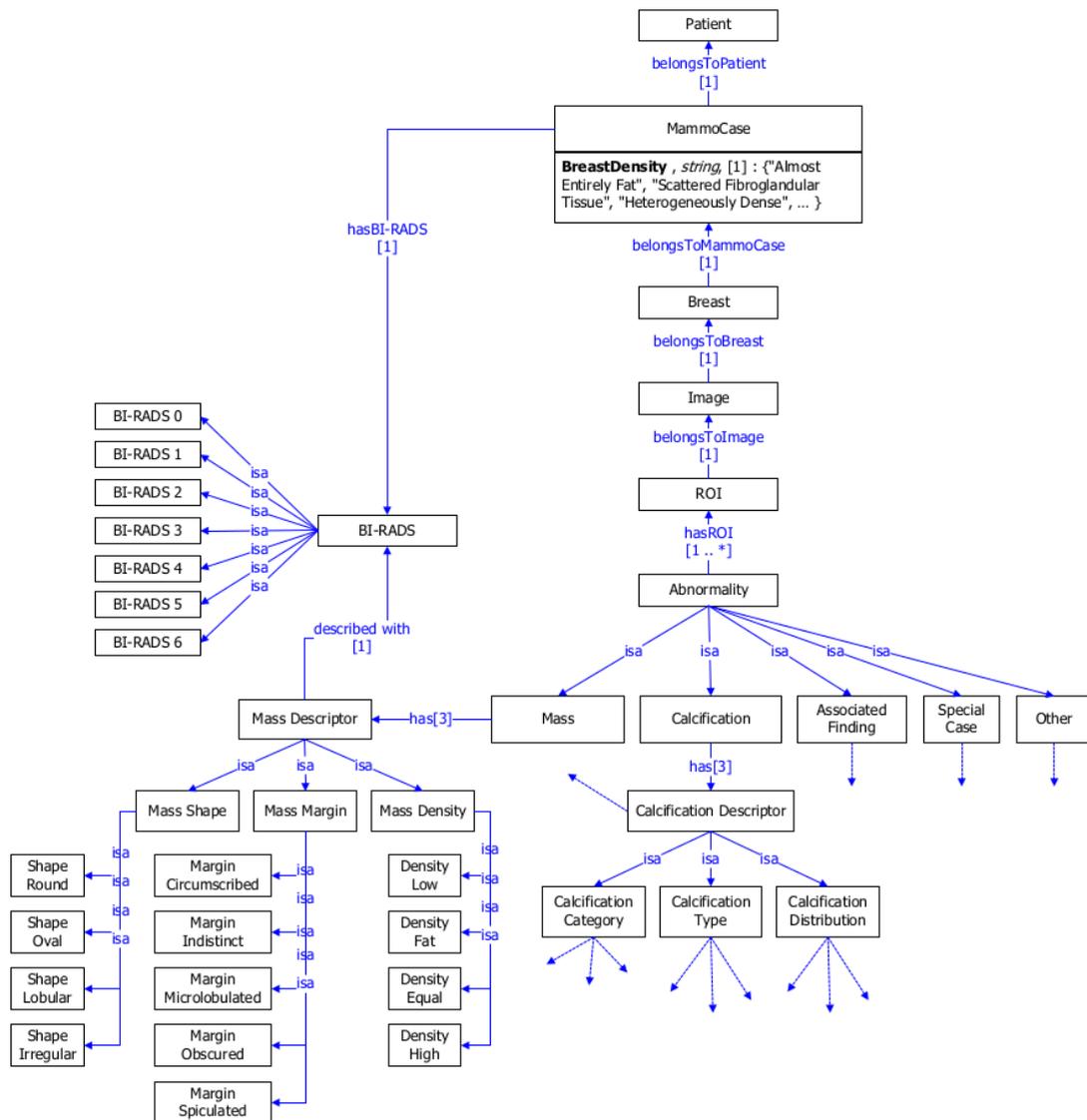
Mammography Annotation Ontology (MAO) é uma ontologia de domínio proposta por Bulu, Alpkocak e Balci (2012). Essa ontologia fornece um vocabulário compartilhado para interpretação de mamografia, permite a anotação de anormalidade observadas em mamogramas. Uma visão simplificada da ontologia pode ser observada na figura 5.

NCI Thesaurus

NCI Thesaurus é uma terminologia baseada em lógica de descrição. Fornece definições, sinônimos e informações relevantes sobre tipos de câncer e doenças relacionadas, terapias combinadas e outros tópicos relacionados com câncer e domínio biomédico (IVANOVIĆ; BUDIMAC, 2014).

2.4 Anotação Semântica

O processo de vincular modelos semânticos e linguagem natural é conhecido como anotação semântica. Segundo Bhagat e Choudhary (2018) a anotação de imagem é um processo de vinculação de descrições à imagem. Essas descrições representam o conteúdo dessa imagem, isso ajuda na recuperação inteligente de imagens relevantes. A anotação de imagem significa atribuir descrições a uma imagem considerando o seu conteúdo (STANESCU et al., 2011). Através da anotação semântica é possível referenciar as entidades às descrições formais presentes na ontologia. Esse processo cria inter-relações entre ontologias e documentos ou imagens. A anotação semântica é a atribuição de *links* para

Figura 5 – Visão simplificada da *Mammography Annotation Ontology* (MAO)

Fonte: Bulu (2013).

a descrição semântica de cada entidade, (KIRYAKOV et al., 2004). Segundo (UREN et al., 2006), as ferramentas que permitem a produção de anotações semânticas podem ser manuais, semiautomática ou automáticas.

A anotação manual é feita por ferramentas que fornecem um editor de texto ou imagem e suporte para ontologias, podendo possuir recursos que apoiam a anotação semântica, como o *Electronic Physician Annotation Device* (ePAD)⁴ que permite anotação semântica e marcação de imagens radiológicas (RUBIN et al., 2008), (RUBIN et al., 2019). A anotação semiautomática pode ser realizada por ferramentas que proporcionam sugestões para as anotações, porém, que ainda necessitam da intervenção de especialistas. As anotações automáticas podem ser conduzidas por ferramentas que realizam as anotações automaticamente, principalmente decorrentes da aplicação de algoritmos de aprendizado

⁴ epad.stanford.edu

de máquina.

A alta complexidade da tarefa de anotação de imagens torna o processo desafiador, esse processo é estudado no campo de classificação e recuperação de imagens. A anotação de imagem ajuda na recuperação, organização, classificação, entre outros. As ontologias podem ser usadas para o processo de anotação semântica por modelarem as relações entre os conceitos e fornecem o conhecimento estruturado que permite atribuir descrições semânticas às imagens. Os vocabulários controlados também permitem a anotação de imagens, ao fornecerem termos predefinidos com as respectivas definições que podem facilitar a anotação de grandes conjuntos de imagens, além disso, fornecem termos padrão que descrevem as características das imagens (KURTZ et al., 2014a).

Geralmente o processo de anotação é apoiado por ambientes que incluem ferramentas para manipulação das imagens, principalmente nos casos de anotação manual das imagens em que um especialista de domínio realiza as anotações. Nesses sistemas, as ontologias proporcionam a estrutura conceitual de um domínio específico, apoiando o processo de anotação semântica. Diversas abordagens foram desenvolvidas para apoiar o processo de anotação semântica manual de imagens médicas como, por exemplo, a abordagem *Annotation and Image Markup* (AIM) (RUBIN et al., 2009), o ePAD (RUBIN et al., 2019) e a abordagem no formato de *plugin* para o ePAD para marcação 3D em imagens radiológicas (MOREIRA et al., 2015).

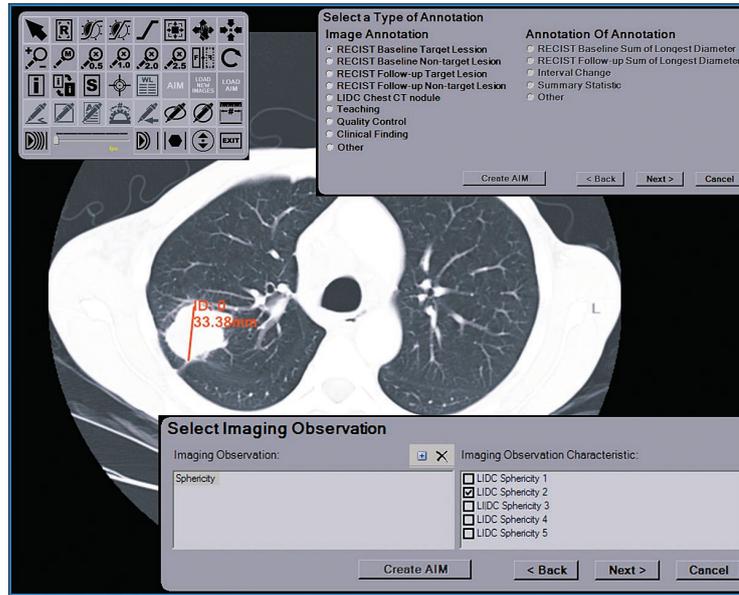
O AIM é uma abordagem que permite marcar graficamente e anotar semanticamente imagens médicas. Permite estabelecer padrões para o registro de informações semânticas de imagens médicas. Para fornecer uma terminologia controlada, descrever o conteúdo de imagens médicas e fornecer um modelo padrão para as informações das imagens e para o registro das anotações semânticas foi desenvolvida a ontologia AIM em OWL. A ontologia representa o conhecimento sobre estruturas anatômicas, observações dos radiologistas sobre as imagens, regiões espaciais da imagem e os requisitos de anotação, as estruturas anatômicas e observações foram obtidas da ontologia RadLex.

Para o processo de anotação e marcação no AIM utiliza-se uma ferramenta na linguagem Java que permite coletar anotações dos usuários enquanto as imagens são analisadas, conforme apresentado na Figura 6.

A ferramenta de anotação utiliza a ontologia para orientar os usuários quanto aos requisitos de anotação durante o processo de marcação e anotação. O AIM e a ferramenta de anotação juntos permitem vincular conceitualmente a imagem e seu conteúdo semântico.

Abordagens para anotação automática geralmente incluem processamento de imagem e extração de características do conteúdo da imagem. Além disso, essas abordagens realizam o processo de anotação por meio de algoritmos de aprendizado de máquina (BHAGAT; CHOUDHARY, 2018). No entanto, frequentemente não consideram as relações semânticas entre os conceitos para realizar a anotação. A anotação semântica de imagens

Figura 6 – Ferramenta de anotação de imagens AIM



Fonte: Rubin et al. (2009).

com o uso de ontologias é uma área de pesquisa que ainda precisa ser melhor explorada e é um dos focos deste trabalho de doutorado.

2.5 Métricas

As métricas são importantes para comparar abordagens e/ou conceitos.

Medida de Similaridade do Cosseno

A similaridade entre dois vetores de palavras pode ser calculada usando a similaridade do cosseno, na equação 1. O produto escalar dividido pelo produto de seus comprimentos é o cosseno de dois vetores (HOCKENMAIER, 2023).

$$\text{Simcos}(\vec{x}, \vec{y}) = \frac{\sum_{i=1}^n x_i y_i}{\sqrt{\sum_{i=1}^n (x_i)^2} \sqrt{\sum_{i=1}^n (y_i)^2}} = \frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\| \|\vec{y}\|} \quad (1)$$

A similaridade do cosseno é utilizada para o cálculo de similaridade entre os vetores de palavra e contexto pertencentes às matrizes de coocorrência.

Medidas de Interesse

Para calcular a qualidade de uma regra de associação, são aplicadas medidas de interesse objetivas que medem estatisticamente a força de uma regra (GONÇALVES, 2005; CAMILO et al., 2010). O modelo suporte e confiança é o modelo frequentemente utilizado na mineração de regras de associação (AGRAWAL; SRIKANT et al., 1994). O suporte mínimo (*minsup*) e a confiança mínima (*minconf*) são especificados pelo usuário.

O suporte calcula a fração de transações que contém X e Y. Para uma regra $X \rightarrow Y$, sendo T o total de transações. O suporte é dado pela equação 2:

$$\text{Suporte} = \frac{\text{sup}(X \cup Y)}{T} \quad (2)$$

A confiança calcula a frequência com que o item Y ocorre em transações com o item X. A confiança é dada pela equação 3:

$$\text{Confiança} = \frac{\text{sup}(X \cup Y)}{\text{sup}(X)} \quad (3)$$

Além do suporte e da confiança, outras medidas são utilizadas. A medida *Lift* avalia a dependência entre o antecedente e o consequente (BRIN et al., 1997). Para a regra $X \rightarrow Y$, quanto mais frequente é Y quando X ocorre. A medida *lift* é demonstrada na equação 4:

$$\text{Lift}(X \rightarrow Y) = \frac{\text{conf}(X \rightarrow Y)}{\text{sup}(Y)} \quad (4)$$

Se $\text{Lift}(X \rightarrow Y) = 1$, X e Y são independentes. Se $\text{Lift}(X \rightarrow Y) > 1$ a dependência entre X e Y é positiva. Se $\text{Lift}(X \rightarrow Y) < 1$ a dependência entre X e Y é negativa. A regra é mais interessante para maiores valores de *lift*.

A convicção é a medida que avalia a regra como uma implicação. Mede a importância da associação entre X e Y. A medida considera o suporte do antecedente e do consequente. A medida *conv* é apresentada na equação 5:

$$\text{Conv}(X \rightarrow Y) = \frac{\text{sup}(X) * \text{sup}(\neg Y)}{\text{sup}(X \cup Y)} \quad (5)$$

Se $\text{Conv}(X \rightarrow Y) = 1$, X e Y são independentes. O valor da convicção tende ao infinito caso X e Y sempre ocorram. Um valor adequado para regras interessantes é entre 1,01 e 5, de acordo com Brin et al. (1997).

Medidas de qualidade do Classificador

Para a avaliação do classificador das imagens são utilizadas algumas medidas discutidas a seguir. (HAN; PEI; KAMBER, 2011; POWERS, 2020; WITTEN; EIBE; HALL, 2011; AGGARWAL, 2015).

A acurácia de um classificador M, $\text{accuracy}(M)$ é a porcentagem de imagens do conjunto de teste classificadas corretamente pelo modelo M. Uma matriz de confusão representa como o classificador se comportou na separação dessas imagens por classes. Um exemplo é demonstrado na tabela 5.

Em uma classificação binária, sendo uma classe denominada positivo (p.ex. classe 1 ou *mass*) e a outra negativo (p.ex. classe 2 ou *calcification*), os achados verdadeiro-positivo são aqueles que tem a classe positivo e foram classificados corretamente. Os achados verdadeiro-negativo são aqueles que tem a classe negativo e foram classificados corretamente. Os falso-positivo são aqueles que tem a classe negativo e foram classificados como positivo. E, os falso-negativo são aqueles que tem a classe positivo e foram classificados como negativo.

Tabela 5 – Matriz de confusão para o classificador M

		<i>Predito</i>	
		Positivo	Negativo
<i>Verdadeiro</i>	Positivo	verdadeiro-positivo	falso-negativo
	Negativo	falso-positivo	verdadeiro-negativo

A sensibilidade ou revocação, ou taxa de verdadeiro positivo, é a proporção de positivos preditos corretamente como positivo. A medida é demonstrada na equação 6.

$$\text{sensibilidade} = \frac{\text{verdadeiro} - \text{positivo}}{\text{positivo}} \quad (6)$$

A especificidade ou revocação inversa, ou taxa de verdadeiro negativo, é a proporção de negativos preditos corretamente como negativo. A medida é demonstrada na equação 7.

$$\text{especificidade} = \frac{\text{verdadeiro} - \text{negativo}}{\text{negativo}} \quad (7)$$

A precisão ou confiança é a proporção dos classificados como positivo que são verdadeiros positivo. A precisão é definida na equação 8.

$$\text{precisão} = \frac{\text{verdadeiro} - \text{positivo}}{\text{verdadeiro} - \text{positivo} + \text{falso} - \text{positivo}} \quad (8)$$

A acurácia em função da sensibilidade e especificidade é definida na equação 9.

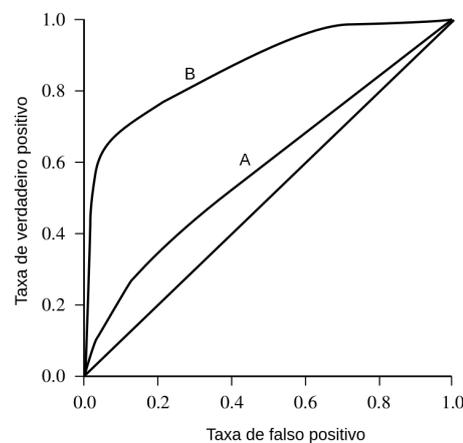
$$\text{acurácia} = \text{sensibilidade} \frac{\text{positivo}}{(\text{positivo} + \text{negativo})} + \text{especificidade} \frac{\text{negativo}}{(\text{positivo} + \text{negativo})} \quad (9)$$

A medida *F1-measure* é a média harmônica entre a precisão e a revocação, definida na equação 10.

$$acurácia = \frac{2 \cdot \text{precisão} \cdot \text{revocação}}{\text{precisão} + \text{revocação}} \quad (10)$$

A *ROC Curve* (ROC) é a representação gráfica da fração de verdadeiro positivos contra a fração de falso positivos, apresentada na figura 7.

Figura 7 – Curva ROC de dois modelos de classificação



Fonte: Adaptada de (HAN; PEI; KAMBER, 2011)

A *Area under the ROC curve* (AUC) é área que se encontra abaixo da curva ROC no gráfico. AUC fornece um quantitativo de avaliação da eficácia de um determinado classificador. No gráfico da figura 7, é possível observar que a área abaixo da curva produzida pelo classificador B é maior que a área do classificador A. Assim, o classificador B apresenta melhor desempenho que o classificador A.

2.6 Considerações Finais

Neste trabalho as imagens são exploradas usando *deep learning* para uma classificação a ser enriquecida pela ontologia. A mineração de texto e de regras de associação enriquecem a ontologia. Métricas são usadas para avaliar a classificação, a mineração de texto e de regras. A ontologia enriquecida pelos resultados da mineração de texto e de regras é utilizada para expandir a anotação das imagens. Assim, este capítulo abordou os principais temas relacionados com esta pesquisa de Doutorado, tais como os aspectos gerais, imagiologia médica, mineração de dados complexos, mineração de imagens, *deep learning*, mineração de texto, uma visão geral sobre ontologia, a anotação semântica e as métricas de avaliação. No próximo capítulo serão discutidos os trabalhos relacionados.

Capítulo 3

Trabalhos Relacionados

Neste Capítulo são apresentados trabalhos relacionados com esta pesquisa de doutorado. São discutidos trabalhos sobre anotação semântica de imagens médicas e em outros cenários, além dos trabalhos relacionados com classificação de imagens e deep learning.

3.1 Anotação Semântica em Imagens Médicas

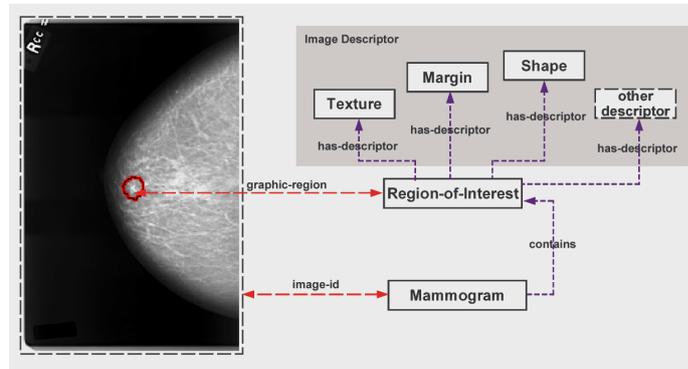
Vários trabalhos abordam o uso de anotação semântica em imagens médicas, entre eles:

Ontology-based medical image annotation with description logics

No artigo de Hu et al. (2003) é proposto um sistema baseado em ontologia para anotação das imagens de mamografia. Os autores desenvolveram a ontologia *Breast Cancer Imaging Ontology* (BCIO) para fornecer um vocabulário comum e definições formais para representar as imagens de mamografia, anormalidades e a avaliação médica. A BCIO é baseada em BI-RADS, desenvolvido por ACR e contém um léxico abrangente sobre dados mamográficos, tais como, descritores de imagem, tipo de lesão e termos médicos de diagnóstico e tratamento. A ontologia BCIO é um modelo de três níveis contendo um conjunto de descritores das características gráficas, assuntos da imagem e termos médicos relacionados aos assuntos da imagem. Por exemplo, como mostrado na Figura 8, as características de uma região delimitada Região de Interesse (ROI) são representadas usando descritores como *Texture*, *Shape*, *Margin*, entre outros. No trabalho, é utilizada a linguagem de lógica descritiva por permitir fazer inferência e descrever formalmente o conhecimento de domínio presente em cada uma das imagens. O sistema utiliza uma interface gráfica para entrada das anotações dos médicos em relação a um determinado ROI. Somente é possível a inserção de dados de instância na ontologia. O sistema não

permite a modificação da estrutura da ontologia para evitar inconsistências nas informações. O sistema de inferência baseado na lógica descritiva pode responder a consultas baseadas em semântica.

Figura 8 – Região de Interesse na Mamografia



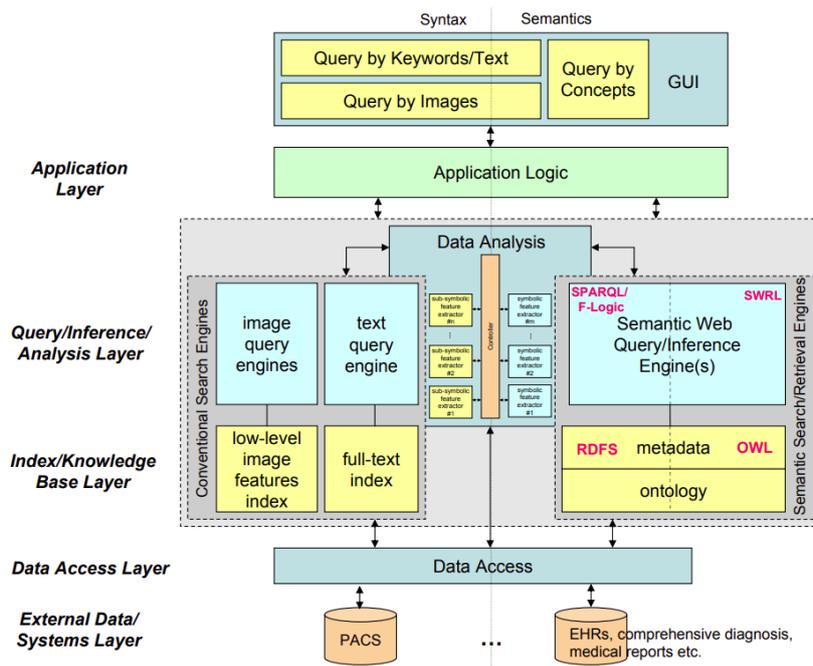
Fonte: Hu et al. (2003)

A Generic Framework for Semantic Medical Image Retrieval

No trabalho apresentado pelos autores Möller e Sintek (2007), é proposto um sistema automático para interpretação de imagens médicas que visa realizar uma fusão formal dos resultados obtidos a partir de técnicas de extração de características da representação de baixo nível da imagem e do conhecimento prévio representado por meio de ontologias. No trabalho é apresentada uma arquitetura que visa a integração das características de alto nível e de baixo nível em vários níveis de abstração para permitir a recuperação multimodal. A consulta é realizada baseada na semelhança visual por meio de recuperação de imagem baseada em conteúdo (CBIR) e por palavra-chave ou texto. Os autores indicam a consulta por conceito, na qual ocorre o mapeamento para os conceitos na ontologia, mas não ocorre uma relação entre os conceitos e, essa parte não é implementada no trabalho. A estrutura da ontologia modela o conhecimento do domínio e visa reutilizar e mapear ontologias existentes no domínio biomédico, a ontologia é implementada na linguagem RDF. A arquitetura proposta pelos autores pode ser visualizada na figura 9.

Radsem: Semantic annotation and retrieval for medical images

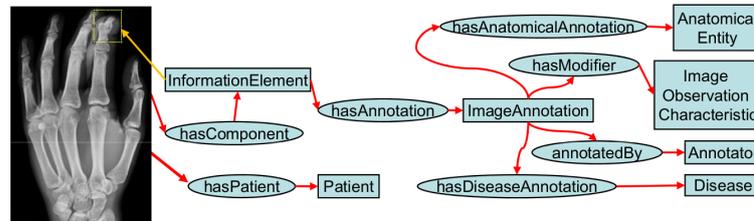
No trabalho de Möller, Regel e Sintek (2009) foi desenvolvida uma ferramenta para anotação e recuperação semântica de imagens médicas. A ferramenta utiliza a ontologia MEDICO que abrange alguns conceitos básicos de outras ontologias, tais como FMA, terminologias como ICD-10 e RadLex, além dos aspectos de procedimentos clínicos. A ontologia foi utilizada nas etapas de anotação e recuperação para desenvolver um extrator de metadados orientado a ontologias para o formato DICOM. Como saída contém informações como nome da pessoa, região do corpo, entre outros. Essa saída é usada para simplificar a anotação manual, fornecendo um subconjunto de conceitos de anotação. Além disso, os metadados extraídos são vinculados às anotações. As buscas são realizadas por expansão da consulta, com base na estrutura das ontologias médicas. A

Figura 9 – Arquitetura do *Framework for Semantic Medical Image Retrieval*

Fonte: Möller e Sintek (2007).

ontologia pode gerenciar os dados clínicos ao permitir vincular e combinar pacientes, imagens médicas e anotações. A ontologia de anotação pode ser observada na Figura 10.

Figura 10 – Ontologia de Anotação

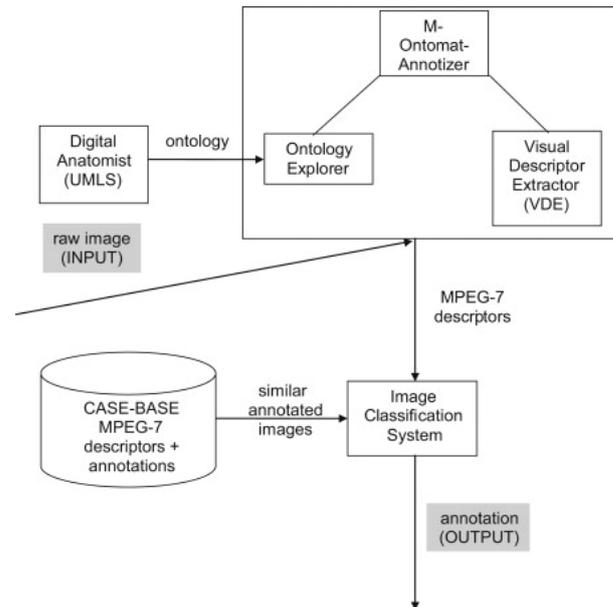


Fonte: Möller, Regel e Sintek (2009).

Automatic semantic indexing of medical images using a web ontology language for case-based image retrieval

No trabalho de Allampalli-Nagaraj e Bichindaritz (2009), os autores apresentam uma abordagem para a recuperação de imagens médicas. Essas imagens são indexadas semanticamente por meio de uma combinação de características de baixo nível da imagem e semântica de alto nível. Diversos tipos de características de baixo nível são extraídas automaticamente, tais como cor, forma e textura. Em seguida, essas características são associadas aos termos relacionados, gerando os metadados semânticos. Os testes foram realizados em um conjunto com 90 imagens e o sistema atingiu 93.4% de precisão na recuperação no nível semântico. A arquitetura da abordagem proposta é apresentada na figura 11.

Figura 11 – Arquitetura do sistema para interpretação de imagem



Fonte: Allampalli-Nagaraj e Bichindaritz (2009).

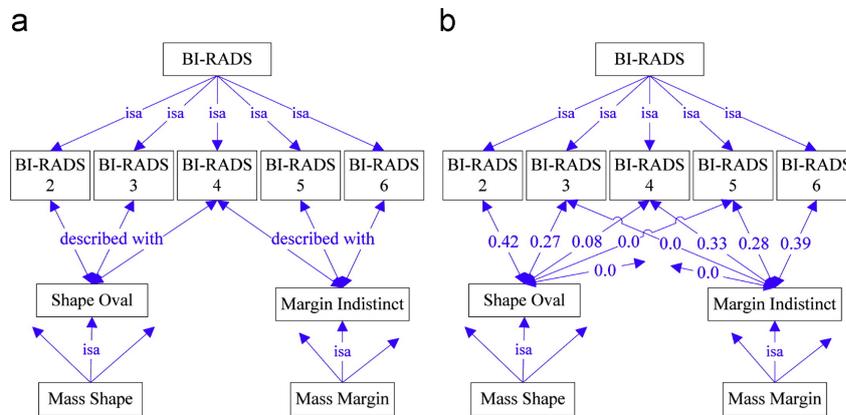
Uncertainty modeling for ontology-based mammography annotation with intelligent BI-RADS scoring

No trabalho de Bulu, Alpkocak e Balci (2013) os autores apresentam um sistema para anotação baseada em ontologia e raciocínio para determinar a pontuação BI-RADS e o sistema utiliza a ontologia MAO. Para determinar a pontuação BI-RADS, os autores propõem uma abordagem bayesiana para modelar a incerteza na determinação da probabilidade de ocorrer uma pontuação BI-RADS em uma dada imagem. Essas informações são incluídas na ontologia e representadas por regras com o uso das linguagens SWRL e SQWRL. Os experimentos foram realizados em dois conjuntos de dados: *Dokuz Eylul University Mammography Set* e *Digital Database for Screening Mammography (DDSM)*, os resultados mostram que existe de fato incerteza na interpretação da pontuação de BI-RADS e o uso da probabilidade bayesiana e lógica de descrição ajudam a lidar com as incertezas entre os conceitos. A figura 12 é uma ilustração das relações *crisp* e não *crisp* entre a pontuação de BI-RADS e as características da massa.

A hierarchical knowledge-based approach for retrieving similar medical images described with semantic annotations

No trabalho de Kurtz et al. (2014a), os autores apresentam uma estrutura que permite incluir características semânticas em ROIs de imagens médicas para possibilitar a recuperação de imagens semelhantes por meio de suas relações semânticas. Essas características são os termos semânticos provenientes de uma ontologia e podem descrever variadas informações sobre o conteúdo da imagem como a forma da lesão. Essa ontologia é um extrato da ontologia RadLex. As imagens são descritas por meio de *Bag-of-Words* (BOW). Os vetores de características permitem o cálculo de similaridade entre eles. Para o cálculo da

Figura 12 – Relações entre as pontuações BI-RADS e características da massa, (a) lógica *crisp* e (b) lógica não *crisp*



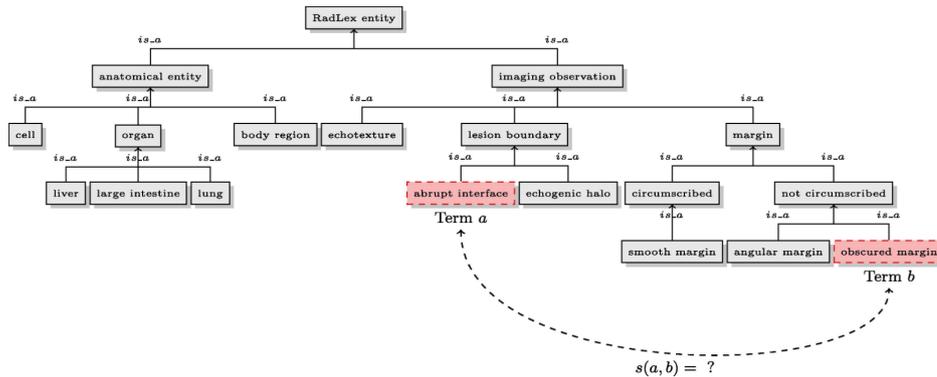
Fonte: Bulu, Alpkocak e Balci (2013).

similaridade, os autores desenvolvem uma nova medida de distância nomeada *Hierarchical Semantic-Based Distance* (HSBD), essa medida considera a distância semântica entre os termos na taxonomia da ontologia. Uma ilustração dessa proposta pode ser observada na figura 13. A proposta ocorre em duas etapas, primeiro a etapa de anotação das imagens com os termos semânticos, em seguida, o cálculo de similaridade dentre as imagens a partir da similaridade entre os termos usados na anotação. A implementação dessas etapas, segundo os autores, lida com o gap semântico ao capturar as relações semânticas entre os termos usados na caracterização da imagem. Para validar a abordagem são realizadas as tarefas de recuperação e de classificação das imagens. A acurácia na recuperação foi maior que 96% num conjunto de 30 imagens de tomografia computadorizada do fígado. Na tarefa de classificação, as taxas de acurácia foram maiores que 93%, em um conjunto de dados com 77 imagens. Ao unir as estratégias de modelar a relação entre os termos, quantificar a proximidade semântica entre eles e calcular a distância entre essas imagens descritas com BOW, a estrutura proposta pode levar a relatórios e sugestões de diagnóstico mais precisos, conforme apontam os autores. O trabalho relatado considera apenas o relacionamento taxonômico entre os termos e não como os mesmos coocorrem em laudos médicos, além disso, não consideram as características de baixo nível das imagens e as anotações são realizadas manualmente por meio da ferramenta ePAD.

On combining image-based and ontological semantic dissimilarities for medical image retrieval applications

Dando continuidade ao trabalho exposto anteriormente, nessa nova abordagem Kurtz et al. (2014b) propõem uma estratégia para prever os termos semânticos por meio do aprendizado automático de assinaturas de termos visuais a partir das características de textura extraídas dos ROIs das imagens. Para isso, utiliza o algoritmo *Support Vector Machine* (SVM) e *wavelets* de Riesz. Segundo os autores, as propriedades multidirecionais e multiescala das *wavelets* de Riesz ajudam na modelagem de características específicas

Figura 13 – Extrato da ontologia RadLex



Fonte: Kurtz et al. (2014a).

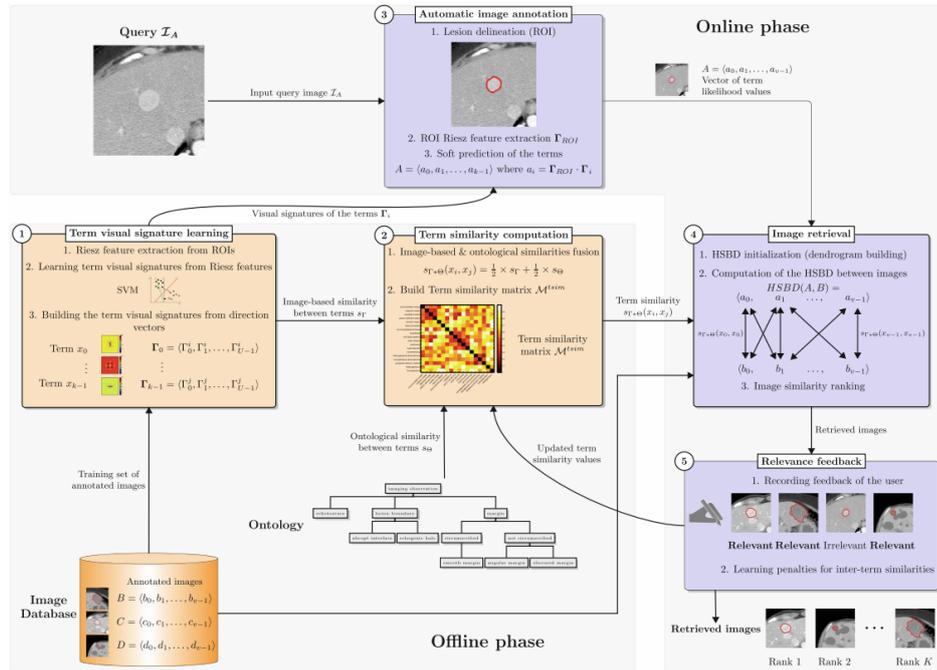
de cada termo semântico. Nesse sentido, ajudam a modelar com maior precisão as assinaturas visuais dos termos semânticos e, assim, permitir uma predição "soft" dos termos ontológicos que descrevem o conteúdo da imagem. Os valores de AUC são de 77% na recuperação das imagens, relativas ao diagnóstico, considerando todas as 72 imagens da base de dados de tomografias computadorizadas do fígado.

Semantic Retrieval of Radiological Images with Relevance Feedback

Num trabalho posterior, (KURTZ et al., 2015) apresentam uma estratégia para considerar também o *feedback* de relevância fornecido pelo usuário, utilizado para capturar semelhanças e relações entre os termos que descrevem o conteúdo da imagem. O mecanismo permite ao usuário marcar as respostas interativamente informando quais delas são relevantes para a consulta, o processo é repetido até satisfazer ao usuário. Os experimentos foram realizados num conjunto de dados composto por 72 imagens de tomografia computadorizada do fígado, e mostram que sucessivas rodadas do *feedback* de relevância permitem aumentar continuamente número de lesões recuperadas para o termo investigado. Na figura 14 pode ser observada a arquitetura proposta pelos autores, no trabalho anterior (KURTZ et al., 2014b) a mesma arquitetura é proposta sem considerar o processo de *Relevance Feedback*.

Adapting content-based image retrieval techniques for the semantic annotation of medical images

Neste trabalho é apresentado um método para anotação semântica automática de imagens médicas que aprimora técnicas de recuperação de imagens baseada em conteúdo CBIR (KUMAR et al., 2016). O método proposto estende as técnicas de CBIR para recuperar um conjunto de imagens rotuladas que possuem características de baixo nível similares. Esse conjunto de imagens rotuladas é então utilizado para definir as melhores anotações semânticas de alto nível para a imagem de consulta. Para demonstrar o método de anotação proposto, os autores implementam duas diferentes estratégias para a recuperação *weighted nearest-neighbour retrieval* (WNN) e *multi-class classification*, conforme pode ser observado na figura 15. Para a estratégia de classificação em dois estágios é uti-

Figura 14 – *Framework* semântico proposto para recuperação de imagens médicas.

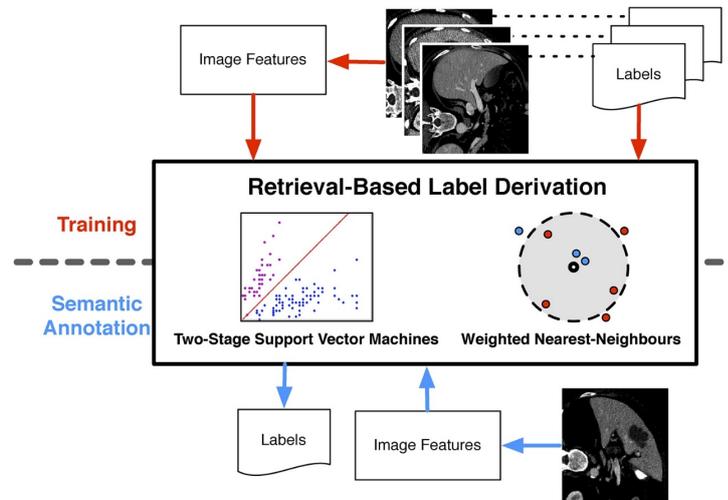
Fonte: Kurtz et al. (2015).

lizada a técnica de SVM. SVMs são modelos de aprendizado supervisionado que podem ser usados para classificação binária. A técnica divide os dados de treinamento rotulados em duas categorias e classifica novas amostras em uma dessas categorias. Essa ideia foi adaptada e utilizada na estratégia de classificação, o problema é dividido em dois estágios, o primeiro identifica a coleção de rótulos que representam os grupos de imagens similares e no segundo estágio é avaliado cada elemento do grupo para selecionar a melhor resposta. A segunda estratégia de recuperação é a técnica WNN, a ideia central é que a consulta seja anotada com o mesmo rótulo atribuído a uma coleção de imagens com características similares. A visão geral das estratégias pode ser observada na figura 16. Para a realização do experimento foi utilizada a validação cruzada e as estratégias foram comparadas com métodos estabelecidos na área. Os resultados mostram que as estratégias superaram os outros métodos e a abordagem WNN superou a precisão entre todos os métodos de comparação. Este trabalho é uma extensão de um relatório submetido ao *Liver Annotation Challenge* do ImageCLEF, que superou os outros métodos de comparação no ano de 2014.

Building and Querying RDF/OWL Database of Semantically Annotated Nuclear Medicine Images

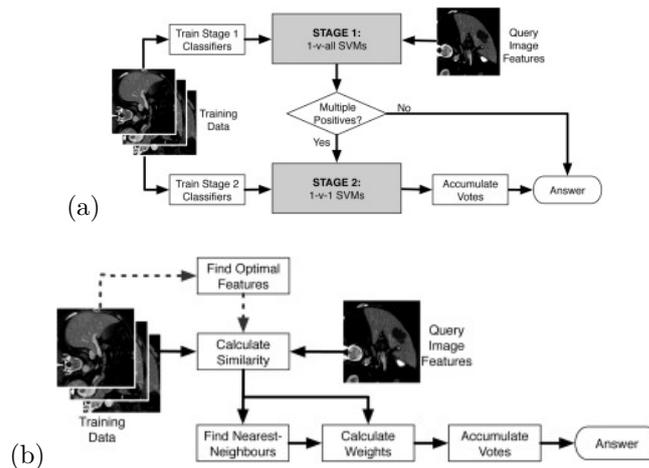
Neste trabalho, os autores constroem um banco de dados semanticamente estruturado de imagens de medicina nuclear (TC) usando o formato AIM, além disso, avaliam a capacidade das anotações do AIM de melhorar a pesquisa de imagens (HWANG et al., 2017). Ao todo, 100 estudos de TC foram anotados no formato AIM, usando vocabulário controlado, e a avaliação da recuperação de imagens foi realizada em 20 consultas clínicas específicas. Dois médicos produziram o *gold standard* com anotações manuais para serem

Figura 15 – Método de anotação proposto em (KUMAR et al., 2016)



Fonte: Kumar et al. (2016).

Figura 16 – Estratégias de classificação usada em (KUMAR et al., 2016)

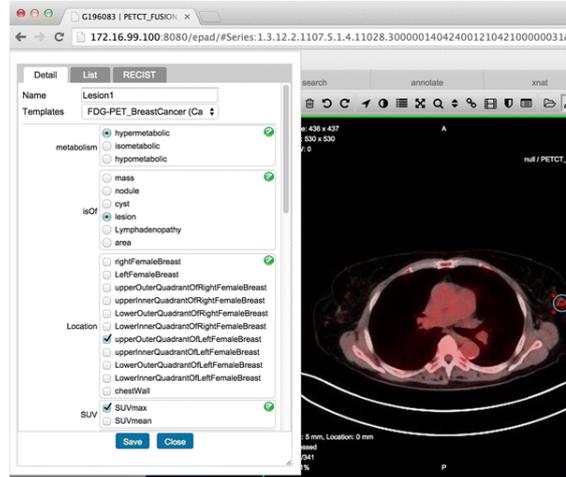


Fonte: Kumar et al. (2016).

utilizadas na avaliação. Para prover as anotações, foram desenvolvidos modelos de AIM específicos para as imagens nucleares e a ferramenta utilizada para anotação foi ePAD. Na figura 17 é possível observar a realização da anotação de uma imagem TC usando o ePAD. No trabalho foram desenvolvidos métodos para a conversão das anotações em AIM em dados RDF/OWL e a linguagem de consulta SPARQL foi usada para realizar a recuperação. A ferramenta ePAD e as anotações AIM permitem vincular informações sobre a lesão e a imagem. Isso possibilita a criação do banco de dados semântico de imagens médicas, que pode ser consultado para encontrar imagens similares e avaliar alterações temporais das lesões. Na avaliação foram comparados os resultados da consulta com os resultados obtidos manualmente pelos médicos. O desempenho da consulta indicou um *recall* de 98% para consultas simples e um *recall* de 89% para consultas complexas. Nas consultas complexas, o sistema apresentou alta precisão na comparação com o *gold standard*,

exceto nas consultas que exigem raciocínio. Como trabalho futuro, os autores sugerem desenvolver o sistema de inferência para melhorar a precisão nesse tipo de consulta.

Figura 17 – Anotação de imagem PET-CT usando o ePAD

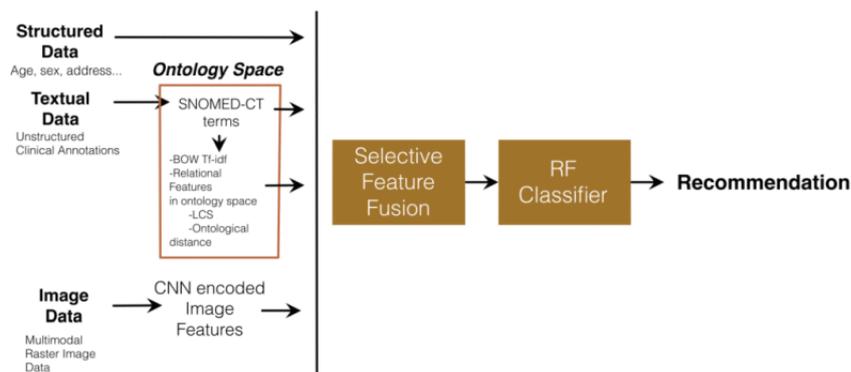


Fonte: Hwang et al. (2017).

Image and Ontological Information Fusion for Cataract Surgery Recommendation

No trabalho desenvolvido por Galveia, Cruz e Travassos (2019), é proposto um sistema automático que combina anotações clínicas em formato de texto, informações demográficas e imagens oftalmológicas visando a recomendação de cirurgia de catarata. As anotações textuais são codificadas utilizando a terminologia contida na ontologia médica SNOMED-CT. As características das imagens são extraídas por redes neurais convolucionais. Esses dois tipos de dados são associados às informações demográficas e inseridos em um classificador *random forest* para determinar se a cirurgia é ou não recomendada. Os resultados do método proposto obtiveram valores de aproximadamente 90% para precisão e F1. Na figura 18 pode ser observado o modelo para recomendação de cirurgia de catarata.

Figura 18 – Método para a recomendação de cirurgia de catarata



Fonte: Galveia, Cruz e Travassos (2019)

Automatic Staging of Cancer Tumors Using AIM Image Annotations and Ontologies

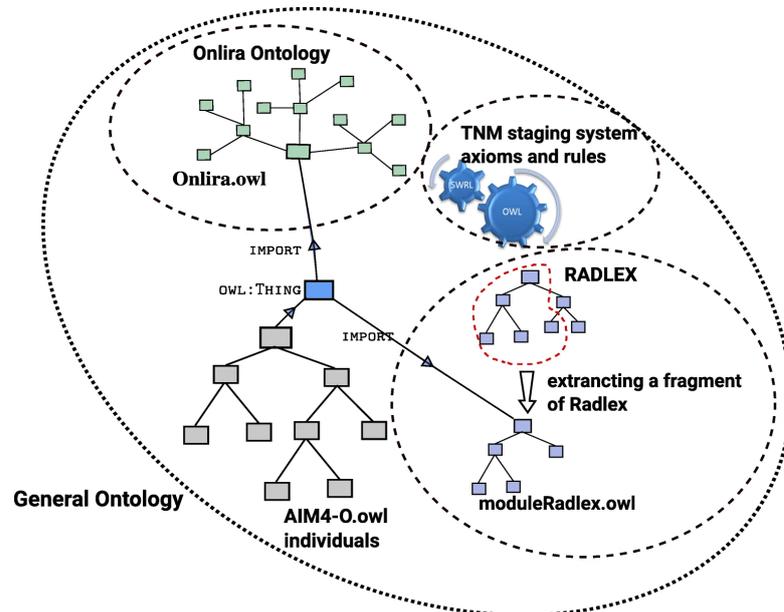
No trabalho de Luque, Rubin e Moreira (2015) e Luque et al. (2019), os autores propõem um classificador que fornece uma segunda opinião para o diagnóstico do estágio do câncer de fígado. Esse classificador segue os padrões para classificação de tumores malignos Classification of Malignant Tumours (TNM). Os radiologistas realizam anotações semânticas nas imagens utilizando a ferramenta ePAD. Essas anotações são armazenadas no formato AIM e por meio de axiomas e regras. Essas anotações são transformadas em instâncias da ontologia AIM4-O, também proposta no mesmo trabalho. A partir disso, por meio de uma abordagem baseada em raciocínio é possível determinar automaticamente o estadiamento TNM do câncer de fígado. Os resultados do classificador demonstram precisão de 85,7% e revocação de 81% na classificação do estágio de câncer quando comparados com o *gold standard*. No trabalho é implementada uma ontologia geral que especifica a semântica das observações das imagens e permite a classificação. A ontologia foi desenvolvida na linguagem OWL e as regras na linguagem SWRL. Nessa ontologia geral são importadas as regras TNM e os axiomas que definem os mecanismos formais de raciocínio, o módulo Radlex e Onlira, uma ontologia do fígado para radiologia. A ontologia geral pode ser observada na figura 19. O trabalho não considera as características de baixo nível das imagens, no entanto, de certa maneira, o trabalho considera o contexto da informação ao incluir, por exemplo, o conjunto de órgãos adjacentes a um órgão principal. Essa informação é necessária para a definição do estágio do câncer, tendo sido modelada por duas superclasses *adjacentOrganGroup* e *noadjacentOrganGroup*. Os autores concluem que o classificador pode auxiliar para a automatização no fluxo de trabalho clínico e, além disso, pode ser aplicado a outros tipos de câncer apenas modificando as regras e axiomas.

3.2 Anotação Semântica em Outros Cenários

Building and using fuzzy multimedia ontologies for semantic image annotation

Neste artigo, os autores Bannour e Hudelot (2014) propõem uma metodologia de construção de ontologia multimídia *fuzzy* para a anotação de imagens. Essa ontologia modela a semântica da imagem ao incorporar o conhecimento visual, conceitual, contextual e espacial da imagem. As informações visuais e conceituais são utilizadas para construir a hierarquia semântica da ontologia, enquanto as informações contextuais e espaciais dos conceitos que descrevem uma imagem são incorporados na ontologia para modelar os relacionamentos semânticos entre esses conceitos. A lógica de descrição *fuzzy* é utilizada para formalizar a incerteza e a imprecisão da informação. As características visuais da imagem são representadas por *Bag-of-Visual-Word* (BoVW). As informações contextuais

Figura 19 – Ontologia geral que importa as regra TNM, axiomas, *moduleRadlex.owl* e Onlira.owl



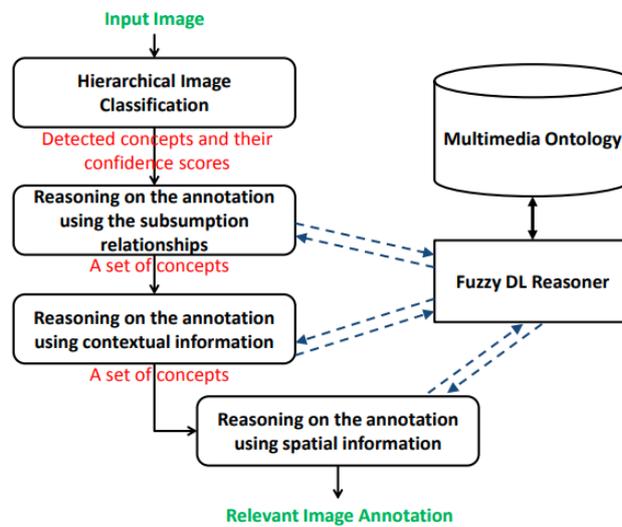
Fonte: Luque et al. (2019).

são obtidas pela distribuição dos conceitos no conjunto de dados, portanto são dependentes do *corpus*. Uma dessas informações é a frequência de coocorrência de pares de conceitos como, por exemplo, se uma imagem contém "televisão" e "sofá" provavelmente ela representa uma "sala de estar". A anotação das imagens ocorre por meio da classificação hierárquica e raciocínio ontológico *fuzzy* em vários estágios para lidar com a consistência da anotação produzida. As bases de dados utilizadas na avaliação são Pascal VOC'2009 e Pascal VOC'2010, essas bases são compostas por imagens que contém coisas comuns ao dia a dia, tais como, pessoa, carro, bicicleta, bola, televisão, etc. Esses conjuntos de dados são geralmente utilizados para validação de propostas de reconhecimento de objetos comuns em imagens. Segundo os autores, os experimentos realizados tiveram melhores resultados de precisão quando comparados com outros métodos. Na figura 20 pode ser observado o *framework* para anotação de imagem baseada em conhecimento e raciocínio em multiestágio.

Ontology and HMAX Features-based Image Classification using Merged Classifiers

No trabalho apresentado em Filali, Zghal e Martinet (2019), os autores propõem uma abordagem para classificação de imagem baseada em ontologia. Essa ontologia é construída a partir das relações taxonômicas encontradas na WordNet, para isso são consideradas as categorias envolvidas nas relações de hiperonímia e hiponímia no domínio "animal doméstico". Para a extração de características da imagem é utilizado o modelo *Hierarchical Max-Pooling Model* (HMAX). Os classificadores são treinados para cada uma das classes da ontologia. Na classificação, uma nova imagem recebe a probabilidade

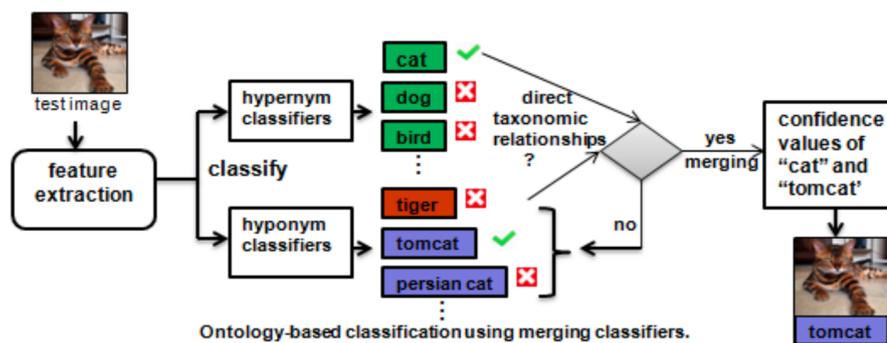
Figura 20 – *Framework* para a anotação de imagem baseada em conhecimento e raciocinador multiestágio



Fonte: Bannour e Hudelot (2014).

de pertencer a determinada classe, tanto para as classes de hiperonímia quanto as classes de hiponímia, e quando uma relação direta entre essas classes é detectada a saída dos classificadores são mescladas para atribuir a melhor classe para a imagem de teste. Um exemplo pode ser visualizado na figura 21. Segundo os autores, os resultados demonstram um desempenho superior para o método proposto quando comparado com os métodos de *baseline*.

Figura 21 – Método de classificação de imagens



Fonte: Filali, Zghal e Martinet (2019).

3.3 Classificação com *Deep Learning*

Vários trabalhos foram desenvolvidos para tratar da classificação de lesões a partir de exames de mamografia. Esta seção apresenta alguns trabalhos relacionados que usam técnicas de aprendizado profundo para classificar imagens de mamografia. Além disso,

trabalhos que adaptaram hiperparâmetros para classificar imagens médicas.

Multi-View Feature Fusion Based Four Views Model for Mammogram Classification Using Convolutional Neural Network

O artigo (KHAN et al., 2019) propõe uma abordagem baseada em *multi-view feature fusion* que usa a técnica de fusão para combinar os vetores de características das quatro diferentes visualizações de uma mamografia em um único vetor de características. A ferramenta CADx conta com três estágios, no primeiro ocorre a classificação em normal e anormal. No segundo, a classificação é para massa e calcificação e no terceiro estágio, a classificação é de benigno e maligno. A extração de características baseada em CNN ocorre para cada visão da imagem separadamente. Essas características são então fundidas na camada final de classificação. Os autores utilizam aumento de dados e realizam o ajuste dos hiperparâmetros para melhorar a classificação. Os experimentos foram realizados com o conjunto de dados CBIS-DDSM e mini-Mias. A abordagem melhorou de 2% a 4% a classificação das mamografias com a estratégia de múltiplas visões em comparação com a classificação de uma única visão. Alcançou 93,73% de precisão e área sob a curva ROC (AUC) de 0,932 para classificação de massa e calcificação no conjunto de teste.

Discrimination of Breast Cancer with Microcalcifications on Mammography by Deep Learning

O artigo (WANG et al., 2016) propõe uma abordagem usando *deep learning* para classificar lesões em imagens de mamografia. Um modelo classificador de discriminação foi desenvolvido para classificar microcalcificações e massas, em conjunto e separadas. No estudo, ao invés da imagem original, 15 características de microcalcificações e 26 características de massa são entradas para o modelo stacked autoencoder. Na camada de topo da rede foi adicionada uma camada softmax para classificar as lesões. Os autores comparam a abordagem *deep learning* com os classificadores tradicionais *support vector machines*, *k-nearest neighbor*, *linear discriminant analysis*. O modelo proposto obteve melhor desempenho, com uma precisão de 87,3% ao analisar apenas microcalcificações e 89,7% ao analisar simultaneamente microcalcificações com massa.

Using Deep Convolutional Neural Networks to Predict Semantic Features of Lesions in Mammograms

Em outro estudo (AGARWAL; CARSON, 2015), foi treinada uma CNN baseada em *conv-nets*, a base DDSM foi utilizada nos experimentos. A abordagem realiza duas tarefas de classificação, para detectar uma lesão como massa ou calcificação, e para detectar o achado da lesão, como benigno ou maligno. Para cada tarefa de classificação binária, a rede *conv-net* foi treinada separadamente. Os hiperparâmetros foram ajustados empiricamente, a regularização e a taxa de aprendizado foram melhoradas com uma abordagem

de varredura de parâmetros. Para a tarefa de classificação de massa e calcificação, a *conv-net* consiste em seis camadas e, para classificação da patologia, cinco camadas. O modelo atingiu a acurácia de 87% para classificação de massa e calcificação. Para benigno e maligno, o modelo atingiu um valor de 90% na precisão no conjunto de treinamento, no conjunto de validação a acurácia obtida foi de 69,8%.

Diagnosis of breast cancer based on modern mammography using hybrid transfer learning

No trabalho de Khamparia et al. (2021), é proposto um processo com foco em *transfer learning* para a predição do câncer de mama. O modelo híbrido proposto modifica o VGG16 ao ajustar as camadas densas finais para uma camada de 32 nós, seguida por uma camada de saída sigmoid para classificação binária. Para melhorar o desempenho, os autores adicionam o aumento de dados e *transfer learning*. A base de dados DDSM é utilizada nos experimentos, e a classificação visa distinguir as imagens patológicas (massa e calcificação), das não patológicas (normais). Os autores comparam o modelo proposto com outras redes, incluindo AlexNet, ResNet50, MobileNet, VGG16, VGG19. O modelo proposto superou os outros modelos, atingindo uma acurácia de 86,9% e uma AUC de 0.933 para o conjunto de teste.

Efficient Hyperparameter Optimization of Convolutional Neural Networks on Classification of Early Pulmonary Nodules

Em Lima, Ferreira e Oliveira (2019), os autores propõem o ajuste de hiperparâmetros da rede para identificar a melhor arquitetura CNN para a detecção precoce de nódulos pulmonares. No trabalho, os autores utilizam o algoritmo TPE para determinar a melhor configuração da rede. Diferentes configurações foram avaliadas, com alteração do número de camadas convolucionais e de camadas totalmente conectadas. Nos experimentos, os autores limitaram em três o número de camadas convolucionais. Ao todo foram avaliadas seis arquiteturas e quatro abordagens diferentes. A melhor arquitetura foi a composta por duas camadas convolucionais e o otimizador SGD. O melhor modelo atingiu 88% de acurácia, a menor taxa de erro foi de 12%.

3.4 Comparação entre as Abordagens

Na tabela 6 estão representados os identificadores de cada critério de comparação. São os seguintes: se o trabalho estudado utiliza as características visuais das imagens; se utiliza informações de contexto obtidas por meio do processamento de informações textuais e das relações de similaridade e coocorrência entre os conceitos envolvidos na anotação dessas imagens; se utiliza ontologias como representação do conhecimento, informação conceitual ou arquitetura para anotação; se utiliza raciocínio ontológico no processo de anotação,

recuperação ou classificação; se utiliza regras de associação; e, por fim, se utiliza técnicas de *deep learning*. A comparação dos trabalhos relacionados é apresentada na tabela 7.

Tabela 6 – Identificação dos critérios de comparação

Identificador	Critério
A	Características Visuais
B	Características de Contexto
C	Ontologia
D	Raciocínio Ontológico
E	Regra de Associação
F	<i>Deep Learning</i>

Tabela 7 – Comparação dos trabalhos da literatura com a proposta de Doutorado.

Trabalho	A	B	C	D	E	F
Hu et al. (2003)	◇		◇	◇		
Möller e Sintek (2007)	◇		◇			
Möller, Regel e Sintek (2009)			◇			
Allampalli-Nagaraj e Bichindaritz (2009)	◇		◇			
Bulu, Alpkocak e Balci (2013)			◇	◇		
Bannour e Hudelot (2014)	◇	◇	◇	◇		
Kurtz et al. (2014a)		◇	◇			
Kurtz et al. (2014b)	◇	◇	◇			
Kurtz et al. (2015)	◇	◇	◇			
Luque, Rubin e Moreira (2015)			◇	◇		
Kumar et al. (2016)	◇		◇			
Hwang et al. (2017)			◇			
Galveia, Cruz e Travassos (2019)	◇	◇	◇			
Filali, Zghal e Martinet (2019)	◇		◇			
Luque et al. (2019)			◇	◇		
Agarwal e Carson (2015)	◇					◇
Wang et al. (2016)	◇					◇
Lima, Ferreira e Oliveira (2019)	◇					◇
Khan et al. (2019)	◇					◇
Khamparia et al. (2021)	◇					◇
Proposta	◇	◇	◇	◇	◇	◇

3.5 Considerações Finais

Neste capítulo foram apresentados trabalhos relacionados com esta pesquisa de doutorado. Primeiramente são discutidos trabalhos sobre anotação semântica de imagens médicas. Logo depois, trabalhos sobre anotação semântica em outros cenários. Por fim, são apresentados os trabalhos relacionados com classificação de imagens e *deep learning*.

Os trabalhos sobre anotação semântica de imagens foram apresentados para demonstrar a evolução da pesquisa relacionado a esse tema. A maioria das abordagens lida com as características visuais para o processo de anotação, e todas incluem o uso de ontologias como parte do processo de anotação. Apenas nos trabalhos de Hu et al. (2003), Bulu, Alpkocak e Balci (2013), Bannour e Hudelot (2014), Luque, Rubin e Moreira (2015), Luque et al. (2019) foi considerado o raciocínio lógico. Nos trabalhos de (BANNOUR; HUDELLOT, 2014; KURTZ et al., 2014b; KURTZ et al., 2014a; KURTZ et al., 2015; GALVEIA; CRUZ; TRAVASSOS, 2019), foi considerado o uso das características de contexto a partir de dados textuais.

Nos trabalhos relacionados com classificação por meio de técnicas de *deep learning*, todos utilizaram as características visuais. No entanto, nenhuma das abordagens utilizou os outros critérios analisados. Além disso, no levantamento realizado, não foram identificados trabalhos que tenham integrado a utilização de *deep learning* para realizar a tarefa de classificação, como foi feito nesta tese. Também não foram usados a mineração de texto e regras de associação visando produzir recursos para a construção de regras semânticas. Além do uso de ontologias para representar o conhecimento e proporcionar o processo de anotação semântica.

Sendo assim, o método proposto neste trabalho é inovador ao agregar diferentes perspectivas de análise de dados, a partir de diferentes fontes de dados. Ao unir as distintas fontes de informação, o método é enriquecido semanticamente e pode aplicar esse conhecimento no processo de anotação. Como vantagem, o método desenvolvido aqui, possibilita uma anotação mais rica semanticamente. E, como consequência, melhoria da interpretabilidade dos resultados e redução do *gap* semântico.

No próximo capítulo é apresentado o método *Unified Method of Semantic Annotation* (MUSA), desenvolvido nesta pesquisa de doutorado.

Capítulo 4

Método Proposto

Neste capítulo são apresentados os materiais e métodos utilizados nesta pesquisa de doutorado. São apresentados o método MUSA, sua arquitetura e seus processos.

4.1 Considerações Iniciais

Com a evolução de tecnologias para obtenção de imagens médicas, e com o aumento do número de exames de imagem realizados, a quantidade desse tipo de dado cresce exponencialmente. Para realizar a leitura e interpretação dessas imagens exige-se do especialista larga experiência na área, conquistada com anos de treinamento e prática. Ferramentas computacionais de anotação semântica e classificação podem auxiliar nesse processo, tanto no dia-a-dia do ambiente médico como no treinamento de residentes e especialistas.

No entanto, grande parte das abordagens consideram apenas as características visuais (*features*) ao nível de *pixels* das imagens. Além disso, muitas vezes os resultados gerados podem ter baixa interpretabilidade e dificultar o entendimento desse resultado por parte dos especialistas. Ao considerar, não apenas as características visuais, mas também os conceitos relacionados ao domínio de conhecimento, ocorre uma melhora na interpretabilidade dos resultados devido à semântica. A multimodalidade de informações muito tem a contribuir para melhorar a leitura, a interpretação e fornecer uma segunda opinião para o diagnóstico de imagens médicas.

As informações multimodais provenientes de características ao nível de *pixel* da imagem, obtidas por meio de técnicas de *deep learning*, compõem um modelo de características visuais. Características de coocorrência e similaridade entre os conceitos provenientes

de laudos médicos compõem um modelo contextual. As características intrínsecas ao domínio e sua representação ontológica compõem o modelo conceitual.

As informações dos modelos contextual e conceitual podem ajudar a melhorar a interpretação dos resultados de classificação e prover anotação semântica de conceitos simples e conceitos complexos em imagens médicas. A construção de artefatos semânticos que traduzem os relacionamentos envolvidos entre as informações de diferentes modalidades que enriquecem o processo de análise de imagens. Ademais, esses artefatos colaboram com a interpretabilidade dos resultados de classificação, realizam a anotação semântica e propõem conceitos como sugestão para um possível laudo, promovendo uma segunda opinião computacional para o diagnóstico médico.

Conforme apresentado no capítulo 3, diversos trabalhos da literatura propõem diferentes abordagens para lidar com a semântica e a multimodalidade de informações na tarefa de classificação e anotação de imagens. Neste projeto de Doutorado serão utilizadas técnicas de mineração de imagem e de texto, e a engenharia de ontologias, para construção dos modelos que apoiam o processo de classificação e a anotação semântica de imagens médicas.

Para isso, serão utilizadas as seguintes informações multimodais: (i) características simples, intermediárias e complexas das imagens, e classificação binária, obtidas por um processo *end-to-end deep learning*; (ii) características contextuais, obtidas por meio de informações presentes nos laudos ou descrições de uma imagem; e (iii) características conceituais, obtidas por meio da ontologia de domínio que define a estrutura semântica do conhecimento. A fusão dessas informações é feita por meio dos artefatos semânticos, tais como os axiomas e as regras na ontologia, que visam contribuir para a redução do "gap semântico" existente entre as características visuais das imagens e sua interpretação.

Este capítulo está organizado da seguinte maneira: primeiramente, na seção 4.2 é apresentado o método proposto, chamado MUSA. Em seguida, na seção 4.3 é apresentada a arquitetura geral que instancia o método proposto. Na seção 4.4 é apresentado o processo de mineração de texto. Na seção 4.5 é apresentado o processo de mineração de imagem. Na sequência, na seção 4.6 é apresentada a engenharia de ontologia utilizada para a construção da estrutura de conhecimento (AnotaMammo), uma ontologia que engloba a ontologia de anotação e a ontologia de domínio.

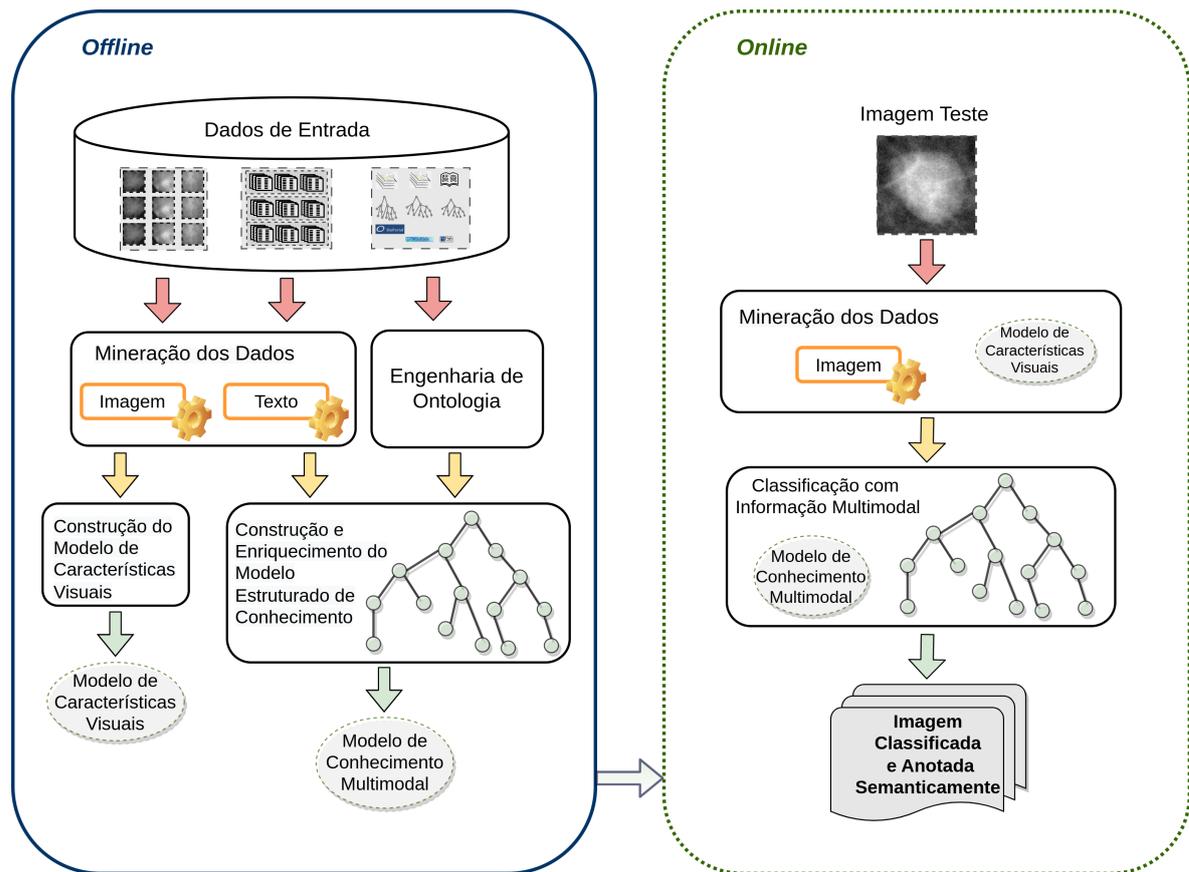
4.2 Método MUSA

O método MUSA (*Unified Method of Semantic Annotation*), é uma proposta para classificar e anotar semanticamente imagens médicas. A partir da mineração de dados de imagem e de texto, e da engenharia de ontologia, ocorre a construção dos modelos que possibilitam a classificação e anotação semântica. O método propõe a fusão de informações multimodais com a vantagem de melhorar a interpretabilidade do resultado da minera-

ção ao incluir novas informações semânticas ao resultado da classificação características visuais.

A abordagem proposta é apresentada na figura 22. Como pode ser observado, a abordagem engloba uma etapa *offline* (treinamento) e uma etapa *online* (teste).

Figura 22 – MUSA - Representação abstrata do método com as etapas *offline* e *online* para classificação e anotação semântica de imagens médicas



Fonte: Elaborada pela autora.

A etapa *offline* tem por objetivo a construção do Modelo de Características Visuais e do Modelo de Conhecimento Multimodal. Esses modelos são utilizados na etapa *online* que tem por objetivo para realizar a classificação e a anotação semântica da imagem de teste.

4.2.1 Etapa *offline*

Na etapa *offline* ocorre a construção do modelo de conhecimento multimodal e do modelo de características visuais. O modelo de características visuais é construído a partir das técnicas de mineração de imagens, em especial, técnicas de *deep learning*, e visa representar as características visuais das imagens médicas. O modelo de conhecimento multimodal visa representar o conhecimento embutido na classificação e anotação semân-

tica das imagens médicas com vistas ao diagnóstico médico. O modelo de conhecimento multimodal é composto pelo modelo contextual obtido a partir de técnicas de mineração de texto, e pelo modelo conceitual obtido a partir da engenharia de ontologias.

Esse conjunto de informações multimodais permite a construção de artefatos semânticos que proveem o enriquecimento do modelo estruturado de conhecimento, para que ele seja utilizado para apoiar o processo de classificação e anotação de conceitos simples e conceitos complexos em imagens médicas.

Na etapa *offline*, é primeiramente inserido um conjunto de dados brutos, um conjunto de treinamento com imagens médicas, um conjunto de laudos médicos e outros documentos que contenham metadados, e uma coleção com ontologias e publicações acadêmicas. Logo após, um processo de mineração é realizado sobre esses dados de imagem e texto. Ocorre também o processo de engenharia de ontologias para construção da estrutura de conhecimento. A mineração de imagem ou de texto é aplicada em cada respectivo tipo de dado, e a engenharia de ontologias é aplicada sobre a coleção com ontologias e publicações acadêmicas. O objetivo dessa etapa é a construção do modelo de características visuais e do modelo de conhecimento multimodal.

A mineração de imagens é realizada sobre os dados de imagem. Primeiramente, pode ocorrer ou não um processo de segmentação para localizar as regiões de interesse de investigação ROI. Após isso, um processo de extração de características é realizado sobre essas regiões de interesse. Essas características são representadas na forma de vetores de características para a aplicação de algoritmos de aprendizado de máquina sobre eles. Aqui é produzido um modelo de nível de características visuais (*features*), obtido a partir dos *pixels* da imagem. A partir da mineração de imagens é possível determinar se uma dada imagem de teste pertence ou não a uma classe, as classes identificadas por esse processo correspondem a um conceito simples.

A mineração de texto é realizada sobre os dados textuais provenientes dos laudos médicos e metadados da imagem médica. O objetivo principal é a construção de regras semânticas que representam esse conhecimento. Inicialmente ocorre o pré-processamento para transformação dos dados em um formato interpretável por máquina. A tokenização e a remoção de *stopwords* são aplicadas neste momento. Em seguida, os dados de texto são transformados para serem processados por algoritmos de aprendizado de máquina. Os algoritmos são utilizados para identificar a similaridade e coocorrência entre as palavras, e também para gerar regras de associação a partir das ocorrências de itens frequentes. A partir da mineração de texto é possível obter informações sobre a correlação entre conceitos simples e como as suas relações de contexto ocorrem para formação de conceitos complexos. Esse é o modelo contextual.

A engenharia de ontologias é realizada sobre os dados provenientes de outras ontologias e publicações científicas. O objetivo é a construção do modelo estruturado de conhecimento, uma ontologia que possibilita a anotação semântica de uma imagem, e, também,

que representa o domínio da aplicação. No caso deste trabalho de doutorado, as imagens de mamografia e o domínio envolvido na interpretação e diagnóstico médico.

O modelo estruturado de conhecimento que realiza a fusão multimodal de informação é uma ontologia. Essa ontologia apoia o processo de anotação semântica e classificação das imagens. A ontologia é composta pela parte relacionada ao processo de anotação, com as informações sobre os metadados das imagens, além dos axiomas e das regras que correspondem aos artefatos semânticos desenvolvidos. E, a outra parte é formada pela ontologia de domínio. Uma representação abstrata dessa estrutura pode ser observada na figura 23. Como em toda ontologia a classe raiz é *Thing*, ao lado esquerdo da imagem são representadas as classes e propriedades da ontologia de anotação, ao lado direito da imagem as classes e propriedades da ontologia de domínio.

A parte da ontologia que apoia o processo de anotação tem como base o modelo estruturado de conhecimento desenvolvido por Bannour e Hudelot (2014). A parte que corresponde ao domínio é baseada em um conjunto de recursos já estabelecidos pela comunidade médica. Principalmente, a MAO (BULU; ALPKOCAK; BALCI, 2012), o léxico BI-RADS (D'ORSI et al., 2013), e variáveis pertencentes ao formato *National Mammography Database* (NMD) (OSUCH et al., 1995; BURNSIDE et al., 2009; CHHATWAL et al., 2009; PATTI, 2011). Além da ontologia *Radiology Lexicon* (Radlex) (LANGLOTZ, 2006) e a ontologia FMA (ROSSE; MEJINO, 2008).

A ontologia de domínio representa o conhecimento semântico associado ao domínio das imagens médicas investigadas, o conhecimento compartilhado entre a comunidade envolvida. No caso desta pesquisa de doutorado é analisado o domínio da interpretação de achados em imagens de mamografia. As informações sobre o conhecimento relacionado ao domínio e determinadas pelas relações de hierarquia, propriedades e axiomas presentes nessa ontologia, formam o modelo conceitual.

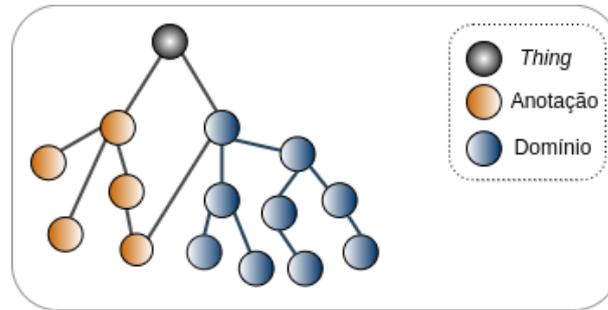
A informação obtida pelo modelo contextual e modelo conceitual, formam o modelo de conhecimento multimodal. Esse modelo é o *core* do modelo estruturado de conhecimento que apoia o processo de anotação e classificação. Por meio dessa estrutura, junto ao modelo de conhecimento multimodal, é realizada a anotação semântica das imagens médicas. Isso permite a atribuição de conceitos simples e complexos às imagens, contribuindo com uma segunda opinião para o diagnóstico de uma imagem.

4.2.2 Etapa *online*

A etapa *online* tem por objetivo realizar a classificação e a anotação semântica da imagem teste.

Na etapa *online*, uma dada imagem ou um conjunto de imagens de teste correspondem aos dados de entrada. Em seguida, ocorre a extração de características dessas imagens e a classificação usando o modelo de características visuais. É uma classificação binária em que é identificado se o tipo de lesão é massa ou calcificação, e outra classificação

Figura 23 – Representação abstrata do modelo estruturado de conhecimento. A raiz (*Thing*) da ontologia, ao lado esquerdo as classes e propriedades da ontologia de anotação e ao lado direito as classes e propriedades da ontologia de domínio.



Fonte: Elaborada pela autora.

binária para identificar se é benigno ou maligno, cada um desses tipos corresponderá a um conceito simples na próxima etapa.

Os resultados obtidos pela classificação de uma imagem de consulta são inseridos na estrutura de conhecimento e por meio da classificação de informação multimodal é possível inferir novos conhecimentos sobre a imagem. Dessa forma, é possível atribuir conceitos simples e conceitos complexos para a nova imagem.

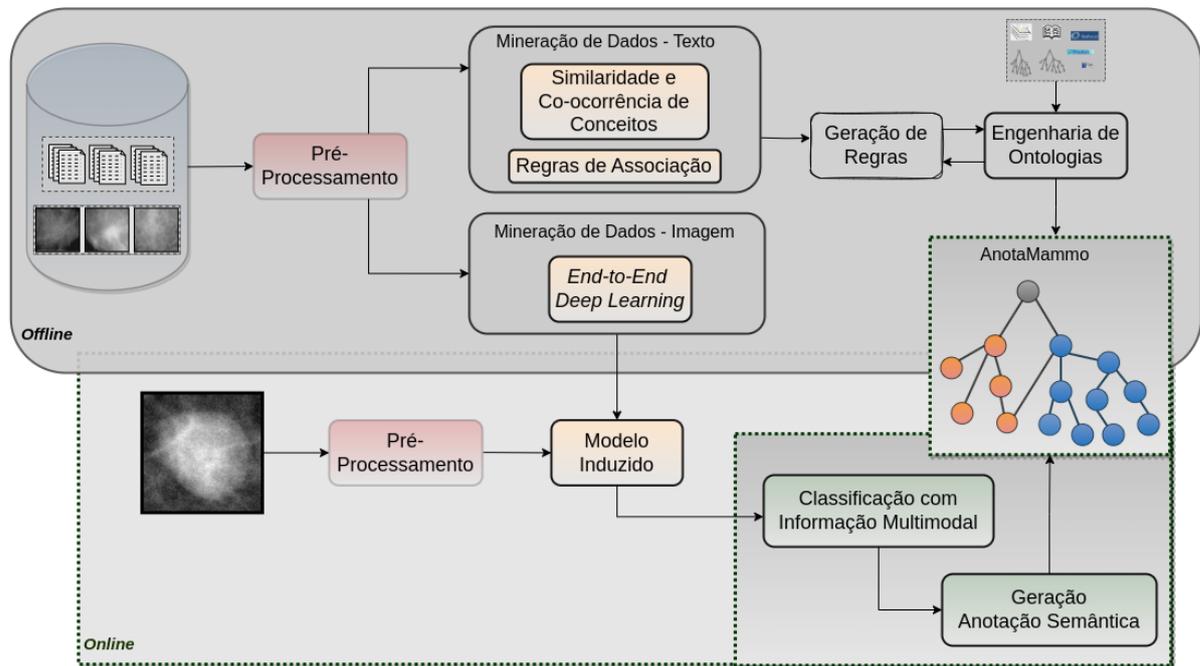
Portanto, ao executar esse processo, uma nova imagem é classificada e anotada semanticamente, além da oportunidade de sugestão de conceitos para compor um laudo para o diagnóstico da imagem analisada, propiciando uma segunda opinião computacional aos especialistas médicos.

4.3 Arquitetura Geral da Proposta

A arquitetura geral discutida nesta seção é uma instanciação da abordagem proposta. Além disso, são abordadas as técnicas e ferramentas utilizadas em cada etapa. Conforme pode ser observado na figura 24, a abordagem requer duas etapas: *offline* e *online*. O objetivo é a construção dos modelos de cada modalidade na etapa *offline*, para que possam ser utilizados na etapa *online*. Com isso, viabilizar a classificação com informação multimodal e a anotação semântica, que possibilita melhorar a interpretabilidade dos resultados da classificação de imagem.

Na etapa *offline* as imagens de treinamento e os laudos são inseridos no processo de mineração de dados, cada tipo de dado em seu respectivo processo. Na mineração de texto algumas técnicas e métodos são aplicados sobre o conjunto de laudos com os diagnósticos das imagens, as informações sobre a similaridade e coocorrência entre os conceitos obtidas

Figura 24 – Arquitetura geral do método proposto. Construção de modelos de multimodalidades na etapa *offline* e utilizados na etapa *online*, para viabilizar a classificação com informação multimodal e a anotação semântica.



Fonte: Elaborada pela autora.

nesse processo contribuem na geração de regras sobre o contexto envolvido. Na mineração de dados de imagens, as imagens de treinamento podem ou não passar por um processo de segmentação, em seguida, passam por um processo de extração de características e por algoritmos de classificação responsáveis pela criação dos modelos de classificação binária. A classificação gerada aqui será associada posteriormente ao conceito simples correspondente.

No processo de geração de regras são utilizadas as informações obtidas por meio de mineração sobre os dados de texto, e as informações conceituais compreendidas na ontologia de domínio. Além das informações contidas na literatura médica e no conhecimento procedural dos especialistas de domínio. Essas regras são implementadas nas linguagens SWRL e SQWRL e são embutidas no modelo estruturado de conhecimento, construindo assim o *core* requerido para a classificação e anotação das imagens.

Na engenharia de ontologias ocorre o processamento dos recursos utilizados para a construção e enriquecimento da estrutura de conhecimento, a ontologia AnotaMammo. A ontologia é composta por duas ramificações que serão nomeadas de ontologia de anotação e ontologia de domínio para melhor compreensão. A ontologia de anotação representa as classes e relacionamentos envolvidos na identificação da imagem e suas informações obtidas pelo processo de mineração. Além disso, também é formada pelas classes e relações necessárias ao processo de classificação com a informação multimodal e ao processo de anotação semântica. A ontologia de domínio representa a estrutura de conhecimento re-

lacionado ao objeto investigado, no caso deste trabalho de doutorado são as mamografias, sua interpretação e diagnóstico.

Na etapa *online* a imagem de teste passa pelo pré-processamento, na sequência, mediante aos modelos induzidos, a imagem é então classificada como massa ou calcificação, e como benigno ou maligno. Essas classes correspondem com os conceitos simples no processo de anotação semântica. Após isso, essas informações são a entrada para o processo de classificação multimodal, aqui ocorre a análise dessas informações pela estrutura de conhecimento que permite a partir da classificação, a anotação semântica das imagens. Como resultado, obtêm-se as imagens classificadas e anotadas semanticamente. Além da sugestão de conceitos simples e conceitos complexos possíveis para a imagem, auxiliando o laudo médico com uma segunda opinião.

4.4 Processo de Mineração de Texto

Para que os laudos se tornem legíveis para o computador é necessário transformá-los para um formato de dados que permita a leitura e interpretação por algoritmos de mineração de dados, essa representação é dada por vetores de palavras. Nesta pesquisa de doutorado é utilizada uma representação clássica de palavras do PLN, uma matriz simétrica de coocorrência palavra-palavra *word-context matrix*, em que a linha representa um vetor de palavras e a coluna representa um vetor de contexto.

Na abordagem tradicional, *vector semantics (distributional similarities)*, o modelo distributivo utiliza o método de contagem para calcular as probabilidades de coocorrência e similaridade entre as palavras (TURNERY; PANTEL, 2010). Conforme apresentado em Levy, Goldberg e Dagan (2015), os métodos tradicionais podem alcançar resultados tão significativos quanto os *neural embeddings*, inspirados em redes neurais. Os autores mostram que parte de ganhos de desempenho nesses métodos ocorre por escolhas de projeto e otimização de hiperparâmetros, e ao realizar tais modificações nos modelos tradicionais é possível atingir desempenho semelhante.

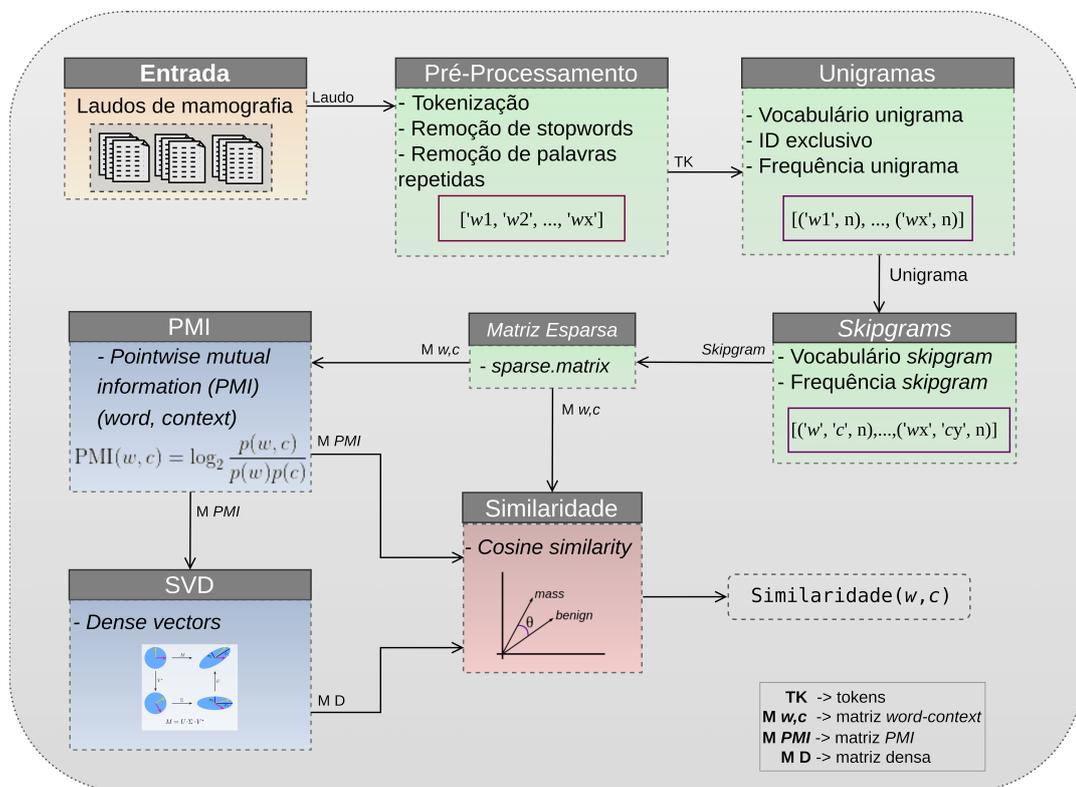
Além disso, o método baseado em contagem é apropriado para pequenas bases de dados, por exigirem menor custo computacional para a execução. Sendo assim, optou-se pelo processo de mineração de texto baseado em contagem para ser usado neste trabalho de doutorado.

O processo de mineração de texto utilizado nesta pesquisa é um processo clássico da PLN, que se baseia em uma matriz *word-context* para cálculo de similaridade e segue algumas recomendações discutidas em Levy, Goldberg e Dagan (2015). O processo é representado na figura 25. O objetivo do processo é identificar coocorrências e a similaridade entre pares de palavras presentes um mesmo laudo. Ou seja, identificar a probabilidade de uma palavra (*word*) e um contexto (*context*) ocorrerem juntos. O processo é composto por diferentes etapas: pré-processamento e tokenização das palavras em cada laudo, iden-

tificação de cada token e contagem de unigramas, além da contagem de coocorrência entre os unigramas formando os *skipgrams*. Essas informações são armazenadas em uma matriz *word-context*.

Essas matrizes têm alta dimensionalidade e são esparsas. A medida utilizada para o cálculo da similaridade entre uma palavra (w) e um contexto (c) é a similaridade do cosseno. Essa medida é aplicada sobre a matriz esparsa com contagem simples de coocorrência entre w e c , e também na matriz esparsa resultante da medida de associação entre palavras PMI (CHURCH; HANKS, 1990). Além disso, o cálculo de similaridade também ocorre sobre os vetores das matrizes densas produzidas por *Singular Value Decomposition* (SVD) (DEERWESTER et al., 1990).

Figura 25 – Processo de mineração de texto para representação de palavras, identificação de coocorrências e a similaridade entre pares de palavras



Fonte: Elaborada pela autora.

A seguir são descritas as etapas do processo de mineração de texto.

4.4.1 Pré-Processamento

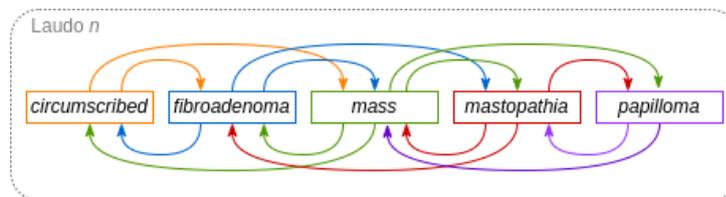
O pré-processamento é a primeira etapa do processo de mineração. A entrada do processo são os laudos de mamografia obtidos a partir de diferentes bases de dados, as informações contidas nos laudos são utilizadas para a obtenção de recursos para construção das regras semânticas. Os documentos de entrada passam pela etapa de pré-processamento em que ocorrem a tokenização e remoção de *stopwords*.

A saída dessa etapa são os tokens separados para cada laudo. Essa saída alimenta a próxima etapa que é a identificação dos unigramas. Cada unigrama, ou cada token, recebe uma identificação única e esse mapeamento é armazenado em um dicionário. Além disso, é calculada a frequência de cada unigrama no corpus. Os unigramas identificados e contabilizados são a entrada para a próxima etapa que define o vocabulário de *skipgrams*.

Na etapa de *skipgram* é definido o vocabulário e a frequência de cada par de coocorrências, esse cálculo é feito por contagem simples de quantas vezes w e c ocorrem juntos em um mesmo laudo. O cálculo ocorre por meio de uma janela dinâmica, sendo duas janelas, uma para frente e outra para trás. O tamanho da janela é definido como o tamanho máximo de tokens em um laudo, para garantir que todas as palavras contidas no mesmo laudo formem pares *word*, *context*, tanto sendo *word* quanto sendo *context*.

A janela desliza sobre os unigramas e aqueles próximos a ele conforme o tamanho e direção da janela. Na figura 26 é possível observar uma representação desse comportamento pelas diferentes cores. No exemplo, considera-se um laudo n contendo cinco unigramas e uma janela de tamanho $janela_frente = janela_traz = 2$. Os *skipgrams* são computados para cada laudo do corpus. Com isso, é possível estabelecer a frequência de pares de palavras em todo um corpus. A matriz *word-context* com a frequência de cada *skipgram* é a saída da etapa.

Figura 26 – Representação por cores de como ocorre o janelamento para identificação de coocorrência de palavras num mesmo laudo, formando os *Skipgrams*



Fonte: Elaborada pela autora.

As matrizes *word-context* podem apresentar alta dimensionalidade, pois sua dimensão é do tamanho do vocabulário. Por isso, é adequado utilizar uma representação de matriz esparsa que otimiza memória no processamento de algoritmos de mineração (IM; YELICK, 2001). Nessa etapa ocorre a transformação das matrizes para matrizes esparsas.

4.4.2 Similaridade

A similaridade entre um par w e c é calculada com a medida de similaridade do cosseno, que permite aproximar termos com correlações parecidas. Inicialmente a similaridade é calculada sobre os vetores que compõem a matriz produzida pela contagem bruta de coocorrências de cada *skipgram*. A matriz *word-context* com os valores de similaridade é a saída dessa etapa.

A similaridade também é calculada sobre os vetores da matriz produzida pelo algoritmo PMI. Nessa etapa é produzida uma matriz *word-context* com valores de PMI (w, c). A probabilidade é definida pela razão logarítmica entre a probabilidade conjunta dos termos, obtida pelo *skipgram*, pelo produto das probabilidades independentes de cada termo, obtidas pela contagem de unigramas.

$$\text{PMI}(w, c) = \log_2 \frac{p(w, c)}{p(w)p(c)} \quad (11)$$

PMI pode resultar em valores positivos e valores negativos, que são quando os pares estão ocorrendo como menos frequência que ao acaso. Na variante *Positive Pointwise Mutual Information* (PPMI) esses valores negativos são substituídos por zero. Aqui também foi substituído por zero os valores no qual a contagem de coocorrência é zero, resolvendo o problema de $-\infty$ de $\log_2(0)$.

$$\text{PPMI}(w, c) = \max(\log_2 \frac{p(w, c)}{p(w)p(c)}, 0) \quad (12)$$

Além disso, para lidar com o viés da PMI em relação aos termos infrequentes, é possível usar uma suavização da distribuição unigrama de contexto. A suavização é realizada por uma função modificada que altera a probabilidade do contexto ao elevar à potência de α . O valor de $\alpha = 0,75$ ou $\alpha = 3/4$ é a configuração estabelecida no trabalho de Mikolov et al. (2013) e preservada em Levy, Goldberg e Dagan (2015). A suavização produzida por α aumenta a probabilidade do contexto e, conseqüentemente, diminui o $\text{PMI}(w, c)$.

$$\text{PPMI}(w, c) = \max(\log_2 \frac{p(w, c)}{p(w)p_\alpha(c)}, 0) \quad (13)$$

$$p_\alpha(c) = \frac{\text{count}(c)^\alpha}{\sum_c \text{count}(c)^\alpha} \quad (14)$$

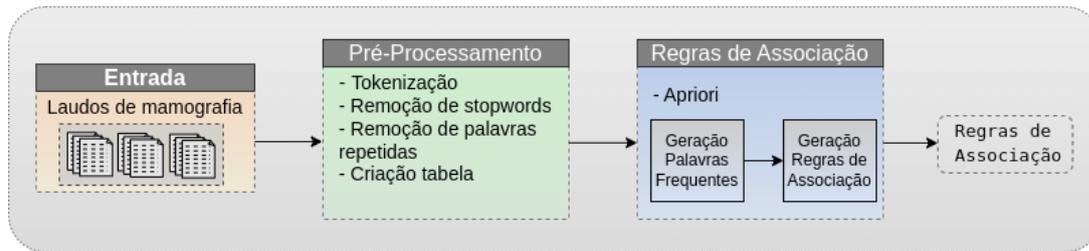
Considerando o exposto acima, a etapa de PMI possibilita a criação de quatro matrizes diferentes. A matriz com os resultados de PMI, PPMI, PMI suavizada, PPMI suavizada.

As matrizes resultantes possuem alta dimensionalidade e são esparsas, ou seja, possuem muitos valores iguais a zero. Para a redução da dimensionalidade as matrizes foram transformadas em matrizes densas. Para isso, foi utilizada a SVD. Com SVD fatorou-se a matriz *word-context* no produto de três matrizes $U \cdot \Sigma \cdot V^T$, sendo U e V matrizes ortonormais e Σ matriz diagonal de autovalores.

4.4.3 Regras de Associação

Os laudos passam por um processo de geração de regras de associação para identificação de novas regras não encontradas anteriormente na mineração de texto. E também para descobrir novas regras geradas pela correlação entre os *itemsets* mais frequentes. A arquitetura do processo pode ser observada na figura 27.

Figura 27 – Processo de geração de Regras de Associação



Fonte: Elaborada pela autora.

O processo recebe como entrada os laudos de mamografia, esses laudos podem contar com dados únicos em cada coluna ou podem compreender conteúdos descritivos. Nem todos os laudos descritivos são facilmente adaptados para a tarefa de geração de regras de associação baseada no algoritmo *Apriori* (AGRAWAL; SRIKANT et al., 1994).

Para laudos disponibilizados como tabela com colunas de valor único, ocorre a seleção de colunas na etapa de pré-processamento. Para os laudos com conteúdo descritivo que podem ser adaptados para colunas, ocorrem a tokenização, remoção de *stopwords* e remoção de palavras repetidas. Após isso, cada token é alocado em uma única coluna, sendo que o tamanho da tabela é o número de tokens do maior laudo pré-processado.

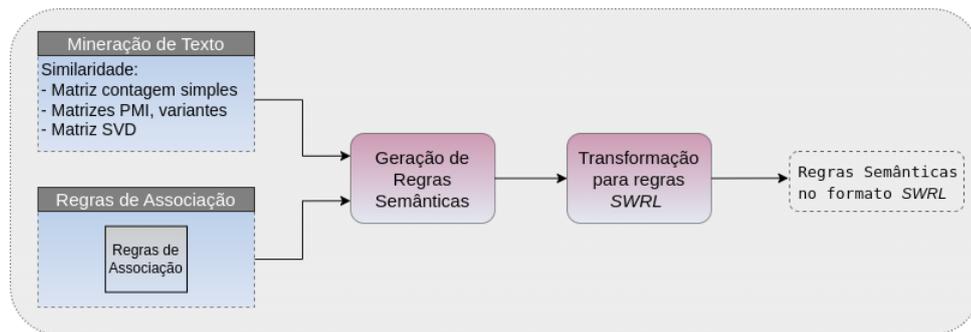
A tabela resultante é a entrada para a etapa de geração de regras de associação por meio da execução do algoritmo *Apriori* (AGRAWAL; SRIKANT et al., 1994). Primeiramente ocorre a identificação dos *itemsets*. Em seguida, ocorre a identificação e geração das regras de associação, sendo a saída do processo.

As saídas do processo de mineração de texto e do processo de geração das regras de associação são os recursos utilizados para a construção das regras semânticas embutidas na estrutura de conhecimento que possibilita a fusão das multimodalidades de informação.

4.4.4 Geração de Regras Semânticas

O processo de geração de regras visa construir as regras semânticas relacionadas ao contexto compreendido no domínio em discussão. O contexto neste trabalho está relacionado ao conteúdo presente nos laudos. Os resultados dos processos anteriores possibilitam a junção das informações obtidas na forma de regras semânticas. As regras são construídas com os resultados produzidos no processo de mineração de texto e no processo de geração de regras de associação. Após a definição da estrutura das regras elas são convertidas para o formato SWRL e inseridas em AnotaMammo que é a ontologia desenvolvida neste trabalho descrita mais adiante neste capítulo (4.6.1).

Figura 28 – Processo de geração de Regras Semânticas



Fonte: Elaborada pela autora.

Também foram adicionados na abordagem regras semânticas geradas a partir das informações conceituais coletadas da ontologia de domínio desenvolvida neste trabalho de doutorado e apresentada nas próximas seções.

4.5 Processo de Mineração de Imagens

O processo de mineração de imagens consiste em analisar e classificar as ROIs de imagens de mamografia. Para isso, o processo engloba uma estratégia *end-to-end deep learning*. A abordagem inclui o procedimento de aprendizado por transferência (*transfer learning*), por meio do ajuste fino (*fine-tuning*) e definição de hiperparâmetros (*hyperparameter*).

O *transfer learning* permite usar o conhecimento descrito por meio de características, pesos ou parâmetros de modelos previamente treinados para treinar novos modelos. Várias arquiteturas CNNs foram propostas nos últimos anos, essas arquiteturas são treinadas com grandes conjuntos de dados como, o ImageNet (DENG et al., 2009). Além disso, permitem sua reutilização por meio de *transfer learning*. Essa técnica é amplamente utilizada para lidar com imagens médicas (ISLAM; WIJEWICKREMA; O'LEARY, 2019).

Com a técnica de *fine-tuning* é possível realizar o congelamento de camadas, que é um processo de congelamento do peso de algumas camadas treináveis da rede. As camadas inferiores das redes contêm informações sobre as características simples, ou de mais baixo nível, dos *pixels* da imagem como, bordas, cantos e contornos, entre outros. Essas características ficam mais complexas, ou de mais alto nível, nas camadas mais avançadas da rede. Conforme descrito em Alom et al. (2019), para um pequeno conjunto de dados com baixa similaridade com os dados da rede pré-treinada, congela-se camadas e treina o classificador com características de baixo nível.

A hiperparametrização é o ajuste de parâmetros fornecidos como argumento para a arquitetura CNN. Naturalmente, os resultados diferem para diferentes valores de hiperparâmetros, e um ajuste adequado pode melhorar o desempenho do treinamento e da classificação (TAJBAKSHI et al., 2016). Para otimização dos hiperparâmetros nesta

pesquisa de doutorado, investigou-se qual o método otimizador baseado em gradiente descendente, e a qual a taxa de aprendizado, são mais adequados para utilização no treinamento do novo modelo, buscando melhorar as taxas de classificação do ROI. Existem vários tipos de hiperparâmetros descritos na literatura. Neste trabalho de doutorado, as técnicas de *transfer learning*, *fine-tuning* e hiperparametrização são aplicadas sobre as três arquiteturas CNN: *DenseNet*, *EfficientNet* e *Xception*. A abordagem inclui ajuste fino e hiperparâmetros para treinar e construir os modelos usando as arquiteturas investigadas. Além disso, os melhores modelos foram combinados como um *ensemble* para melhorar ainda mais o desempenho da classificação.

4.5.1 Arquiteturas CNN pré-treinadas

Dentre as diversas arquiteturas CNN existentes, a seguir são destacadas as arquiteturas utilizadas neste trabalho.

1) DenseNet: Dense Convolutional Network (DenseNet) é uma CNN que conecta cada camada L com todas as outras através de *feed-forward*, então a rede possui $L(L+1)2$ conexões. Nas redes convolucionais tradicionais, cada camada se conecta à sua camada subsequente, tendo assim L conexões (HUANG et al., 2017). Os mapas de características de cada camada são usados como entrada para a próxima camada e assim por diante. Os autores em Huang et al. (2017) destacam algumas vantagens relevantes para DenseNets, como diminuir o número de parâmetros, reforçar a propagação de características e sua reutilização e lidar com o problema do gradiente de fuga. O número de camadas distingue cada versão do Densenet como, DenseNet121, DenseNet169 e DenseNet201.

2) EfficientNet: EfficientNetV2 é uma evolução da EfficientNet, uma família de redes convolucionais (TAN; LE, 2019). Essa versão é uma família de redes convolucionais que otimiza velocidade de treinamento e eficiência de parâmetros usando uma combinação de busca e dimensionamento de arquitetura neural com reconhecimento de treinamento. Os experimentos relatados em Tan e Le (2021) mostram que os modelos EfficientNetV2 treinam mais rápido que outros modelos. Além disso, um método de aprendizado progressivo é proposto para acelerar o treinamento, no qual a regularização é ajustada dinamicamente à medida que as imagens aumentam gradativamente (TAN; LE, 2021).

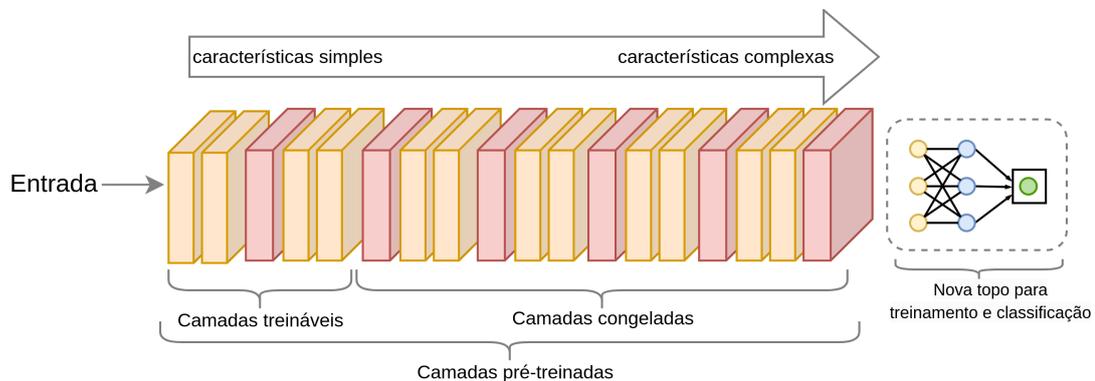
3) Xception: Xception é uma arquitetura CNN inspirada em InceptionV3 (SZEGEDY et al., 2016), uma versão “extrema”. Na Xception, os módulos Inception são substituídos por convoluções separáveis em profundidade. Esta nova versão é mais eficiente que o InceptionV3, especialmente para um grande conjunto de dados de imagem. Além disso, essa arquitetura efetua o uso mais eficiente dos parâmetros do modelo (CHOLLET, 2017).

4.5.2 *Fine-Tuning*

Nesta pesquisa de doutorado, um pequeno conjunto de dados é analisado, considerando o tamanho das grandes bases utilizadas para o treinamento das redes pré-treinadas. Conforme descrito na literatura, para pequenos conjuntos de dados e com baixa similaridade em relação aos dados da rede pré-treinada, é possível manter algumas camadas congeladas, aproveitando os pesos da rede pré-treinada, e treinar novamente outras camadas para a tarefa específica. As técnicas de congelamento são utilizadas para fixar determinadas camadas da rede e mantê-las inalteradas durante o treinamento (CHU et al., 2016; ALOM et al., 2019).

Uma representação do congelamento de camadas pode ser observada na figura 29. As camadas inferiores são retreinadas enquanto as camadas superiores são congeladas. Além disso, é possível observar, um novo topo da rede customizado e adaptado para treinar o novo conjunto de dados e prover a classificação. Por fim, a última camada do novo topo é formada por uma camada densa com o número de nós de acordo o número de classes esperado como saída.

Figura 29 – Representação da arquitetura com o novo topo treinável. Camadas pré-treinadas da arquitetura CNN, com camadas treináveis e camadas congeladas



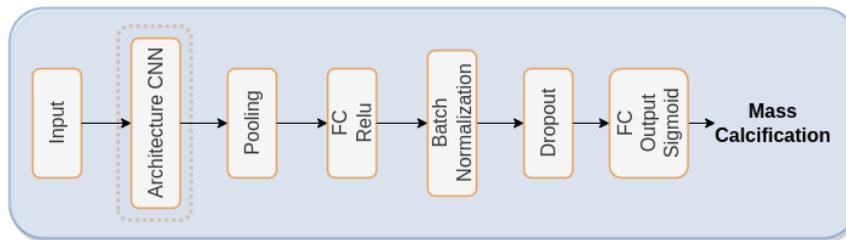
Fonte: Elaborada pela autora.

Ao adicionar o novo topo, a arquitetura resultante do processo de *fine-tuning* é representado na figura 30. Como pode ser visualizado, o novo topo é formado pela adição de uma camada de *pooling*, uma camada totalmente conectada com a função de ativação *Relu*, uma camada *batch normalization*, uma camada de *dropout*. Por fim, uma camada de saída, totalmente conectada, com a função de ativação *Sigmoid*. A saída de rede atribui uma classe ao ROI. Que pode ser, *mass/calcification*, ou *benign/malign*.

4.5.3 *Hiperparâmetros*

A hiperparametrização é o ajuste de parâmetros fornecidos como argumento para a arquitetura CNN. Os resultados diferem para diferentes valores de hiperparâmetros, e um ajuste adequado pode melhorar o desempenho do treinamento e da classificação (TAJ-

Figura 30 – Arquitetura após o processo *fine-tuning* com adição de novo topo



Fonte: Elaborada pela autora.

BAKSH et al., 2016). Existem vários tipos de hiperparâmetros descritos na literatura. E para otimização dos *hiperparameters* nesta pesquisa, a abordagem visa investigar qual o método otimizador baseado em gradiente descendente e qual taxa de aprendizado são mais adequados para utilização no treinamento do novo modelo, buscando melhorar as taxas de classificação do ROI.

O ajuste de hiperparâmetros, ou *hyperparameter tuning*, permite definir qual o melhor otimizador e a taxa de aprendizado (*Learning Rate* (LR)) mais adequada. Os dois hiperparâmetros correspondem a parâmetro de treinamento da rede. Vale ressaltar que os hiperparâmetros da estrutura de rede não foram modificados porque foram utilizadas arquiteturas CNN já estabelecidas.

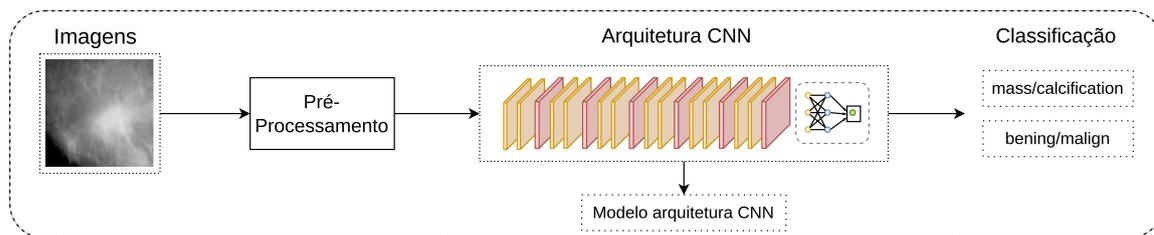
A definição do otimizador mais adequado é realizada separadamente para cada arquitetura de CNN (DenseNet121, EfficientNet2VS, Xception). Cada um dos otimizadores disponíveis no TensorFlow foi avaliado: Adadelta, Adagrad, Adam, Adamax, Ftrl, Nadam e RMSprop. Cada otimizador foi testado com diferentes taxas de aprendizado, começando com um valor baixo de LR, sendo incrementado ao final de cada época. Inicialmente o LR é $2e - 4$ e seus valores aumentam ligeiramente em 10% a cada época. Ao todo foram testadas 150 épocas para cada otimizador. O processo é realizado para diferentes combinações de camadas de treinamento congeladas, sendo escolhido o de melhor desempenho.

Após definição do melhor otimizador, ocorre a busca pela melhor taxa de aprendizagem. Para isso, são testados quatro diferentes valores para LR, são eles: $LR = [1e - 5, 1e - 4, 1e - 3, 1e - 2]$ durante 150 épocas. E aquele que apresentar melhores resultados é selecionado. O ajuste adequado de LR é essencial para obter melhores resultados de precisão.

4.5.4 Construção do Modelo e Classificação

Para a tarefa de classificação é utilizada a estrutura ilustrada na figura 31. Primeiramente, os ROIs são submetidos a uma etapa de pré-processamento. Em seguida, ocorre o treinamento para a construção do modelo e, por fim, a classificação do conjunto de teste com o modelo induzido. Esse processo ocorre para cada uma das arquiteturas CNN separadamente.

Figura 31 – Construção do modelo e classificação binária de um ROI



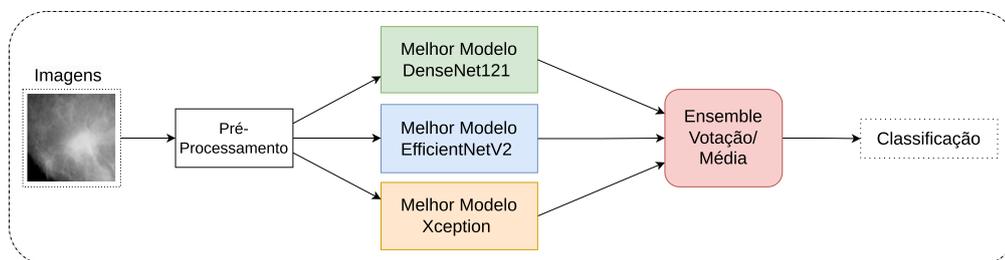
Fonte: Elaborada pela autora.

Para a construção do modelo, cada uma das arquiteturas CNN é incorporada na estrutura proposta pelo processo de ajuste fino, ilustrada na figura 30. Com isso, é construído o modelo de cada uma das arquiteturas. Após, ao instanciar a rede são definidos o otimizador e a taxa de aprendizado selecionados pelo processo de ajuste de hiperparâmetros. O partir do conjunto de treinamento é gerado o modelo instanciado. Por meio do modelo produzido por cada arquitetura é realizada a classificação binária sobre o conjunto de teste. Cada imagem de teste (ROI) é classificada como *mass* ou *calcification*; e, como *benign* ou *malign*.

4.5.5 Ensemble

Ensembles são métodos que combinam diferentes modelos para melhorar a classificação e geralmente produzir melhores resultados do que um único modelo. O *ensemble* de votação classifica cada instância por votação, sendo o resultado aquele que receber mais votos. O *ensemble* de médias classifica cada instância pela predição média. Foram combinados na forma de um *ensemble* os melhores modelos produzidos por cada uma das três CNNs.

Figura 32 – *Ensemble* produzido com os melhores modelos de cada uma das três CNNs, DenseNet121, EfficientNetV2, Xception



Fonte: Elaborada pela autora.

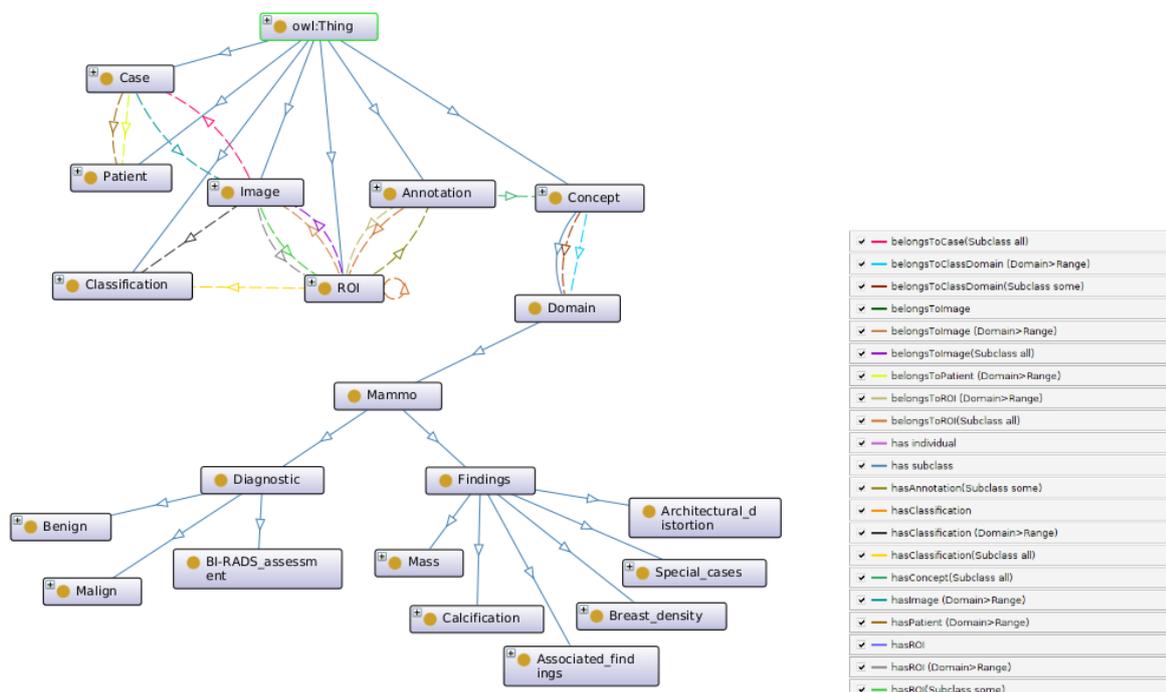
4.6 Engenharia de Ontologia

Para a construção da ontologia Anotamammo é utilizada a metodologia *Ontology Development 101* (NOY; MCGUINNESS et al., 2001) e os princípios de melhores práticas para a construção de ontologias de *Basic Formal Ontology* (ARP; SMITH; SPEAR, 2015).

4.6.1 Ontologia AnotaMammo

A ontologia AnotaMammo, ou modelo estruturado de conhecimento, é composta pela ontologia de anotação, a ontologia de domínio e os artefatos semânticos responsáveis pela fusão multimodal de informação. Essa estrutura pode ser observada na figura 33. Na parte superior da figura 33 estão as classes e propriedades da ontologia de anotação. E, na parte inferior, relacionado com a classe *Domain*, as classes e propriedades da ontologia de domínio, a Mammo.

Figura 33 – Ontologia AnotaMammo



Fonte: Elaborada pela autora.

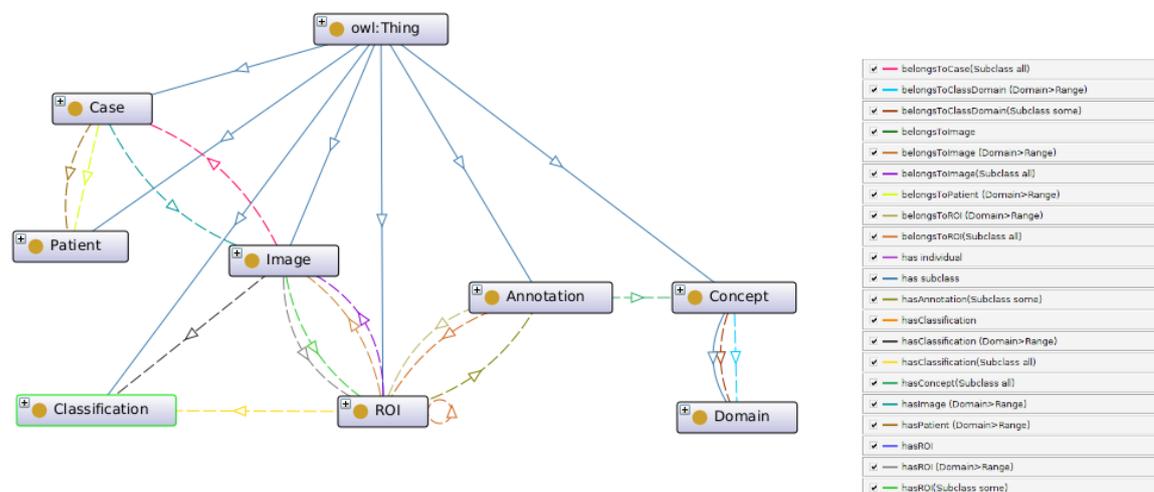
4.6.2 Ontologia de Anotação

A ontologia de anotação foi implementada para representar as informações relacionadas com o conteúdo e identificação do caso, do paciente e das imagens, e com o processo de anotação semântica. A ontologia de anotação da AnotaMammo é baseada no trabalho de Bannour e Hudelot (2014). Na figura 34 pode ser observado um extrato da ontologia de anotação composta pelas seguintes classes:

- ❑ **Case** — corresponde ao caso médico, vinculado a um paciente e a uma ou mais imagens.
- ❑ **Patient** — corresponde ao paciente e abrange suas informações, tais como, idade, histórico de câncer de mama, histórico de câncer de mama na família, terapia hormonal, entre outros.

- ❑ **Image** — corresponde à imagem em si, cada instância possui um *label* identificador e o tipo daquela imagem, se é CC ou MLO.
- ❑ **ROI** — corresponde a cada ROI da imagem, cada instância representa um único ROI e uma imagem pode ter um ou mais ROIs.
- ❑ **Classification** — corresponde às informações de classificação de cada ROI, informações obtidas pelo processo de mineração de imagem.
- ❑ **Annotation** — corresponde às anotações de cada ROI ou imagem. Um ROI possui uma ou mais anotações. Essas anotações interligam cada ROI aos seus respectivos conceitos simples e conceitos complexos.
- ❑ **Concept** — corresponde à classe que realiza a conexão da parte de anotação com a parte do domínio na ontologia. Permite relacionar os conceitos usados na anotação com a respectiva classe na ontologia de domínio.
- ❑ **Domain** — classe que engloba a ontologia de domínio.

Figura 34 – Extrato da ontologia de anotação com as classes e propriedades para a anotação semântica



Fonte: Elaborada pela autora.

4.6.3 Artefatos Semânticos

Os artefatos semânticos são os axiomas e as regras *SWRL* que compõem a Anota-Mammo, especialmente aqueles que permitem a integração das informações multimodais e a anotação semântica.

Alguns exemplos de propriedades de objeto podem ser observados abaixo:

- ❑ **hasImage** — propriedade de objeto que relaciona uma instância da classe caso a uma imagem.

- ❑ **belongsToCase** — propriedade de objeto que relaciona uma instância da classe imagem a um caso.
- ❑ **belongsToPatient** — propriedade de objeto que relaciona o caso com o paciente.
- ❑ **belongsToImage** — propriedade de objeto que relaciona um ROI com uma imagem.
- ❑ **hasROI** — propriedade de objeto que relaciona uma imagem com um ROI.
- ❑ **hasClassification** — propriedade de objeto que relaciona um ROI com a sua classificação, procedente da mineração de imagem aplicada ao ROI.
- ❑ **hasAnnotation** — propriedade de objeto que relaciona uma anotação com um determinado ROI ou imagem, pode existir uma ou mais anotações.
- ❑ **belongsToROI** — propriedade de objeto que relaciona uma anotação com um ROI.
- ❑ **hasConcept** — propriedade de objeto que relaciona uma anotação com um conceito pertencente ao domínio.
- ❑ **belongsToClassDomain** — propriedade de objeto que relaciona um conceito com a classe correspondente na ontologia de domínio.

Alguns exemplos de propriedades de dados podem ser observados abaixo:

- ❑ **Image_Type** — propriedade de dado que relaciona uma imagem com o seu tipo; pode ter os seguintes valores: L-CC, L-MLO, R-CC, R-MLO.
- ❑ **Annotation_ID** — propriedade de dado que identifica a anotação do ROI ou imagem com um valor inteiro.
- ❑ **ROI_ID** — propriedade de dado que identifica o ROI com um valor inteiro.
- ❑ **Similarity_WordConcept** — propriedade de dado que atribui a similaridade entre uma palavra e um conceito, do tipo *float*.

4.6.4 Regras Semânticas

Alguns exemplos de regras podem ser observados abaixo:

- ❑ Regra para recuperar os ROIs que possuem a classificação predita para determinado conceito simples. Neste exemplo, recupera os ROIs cuja classificação é *Mass*.

```
ROI(?r) ^ hasClassification(?r, Mass_classification) -> sqwrl:select(?r)
```

- Regra para anotar um ROI classificado com o respectivo conceito simples, no exemplo o conceito *Mass*

```
ROI(?r) ^ hasClassification(?r, Mass_classification) -> hasAnnotation(?r,
    Mass_anoatation)
```

- Nesse outro exemplo, o ROI recebe a anotação que é vinculada ao conceito simples *Benign*

```
ROI(?r) ^ hasClassification(?r, Benign_classification) ->
    hasAnnotation(?r, Benign_anoatation)
```

- Regra para anotação do ROI com um conceito complexo. No exemplo uma anotação do tipo *Circumscribed*

```
ROI(?r) ^ hasAnnotation(?r, Benign_annotation) ^ hasAnnotation(?r,
    Mass_annotation) -> hasAnnotation(?r, Circumscribed_annotation)
```

- Regra para anotação do ROI com um conceito complexo. No exemplo uma anotação do tipo *Birads3*

```
ROI(?r) ^ hasAnnotation(?r, Benign_annotation) ^ hasAnnotation(?r,
    Mass_annotation) -> hasAnnotation(?r, Birads3_annotation)
```

- Regra para seleccionar os contextos relacionados com o conceito *Bening*, com similaridade maior que 0.5.

```
Annotation(?a) ^ hasConcept(?a, ?c) ^ hasSimilarity_WordContext(?c,
Benign_concept) ^ Similarity_WordContext(?c, ?s) ^ swrlb:greaterThan(?s,
    0.5) -> sqwrl:select(?c, ?s)
```

- Regra para anotação do ROI com um conceito complexo. No exemplo, o contexto fibroadenoma relacionado com o conceito *Circumscribed*.

```
ROI(?r) ^ hasAnnotation(?r, Circumscribed_annotation) ->
    hasAnnotation(?r, Fibroadenoma_annotation)
```

- Regra para seleccionar os itemsets relacionados com o conceito *Bening*, com *confidence* maior que 0.5.

```
Annotation(?a) ^ hasConcept(?a, ?c) ^ hasConfidence(?c, Benign_concept) ^
    Confidence_Itemset(?c, ?s) ^ swrlb:greaterThan(?s, 0.5) ->
    sqwrl:select(?c, ?s)
```

- Regra para recuperar as imagens com anotação do conceito simples *Circumscribed*.

```
ROI(?r) ^ hasAnnotation(?r, Circumscribed_annotation) -> sqwrl:select(?r)
```

4.6.5 Ontologia de Domínio

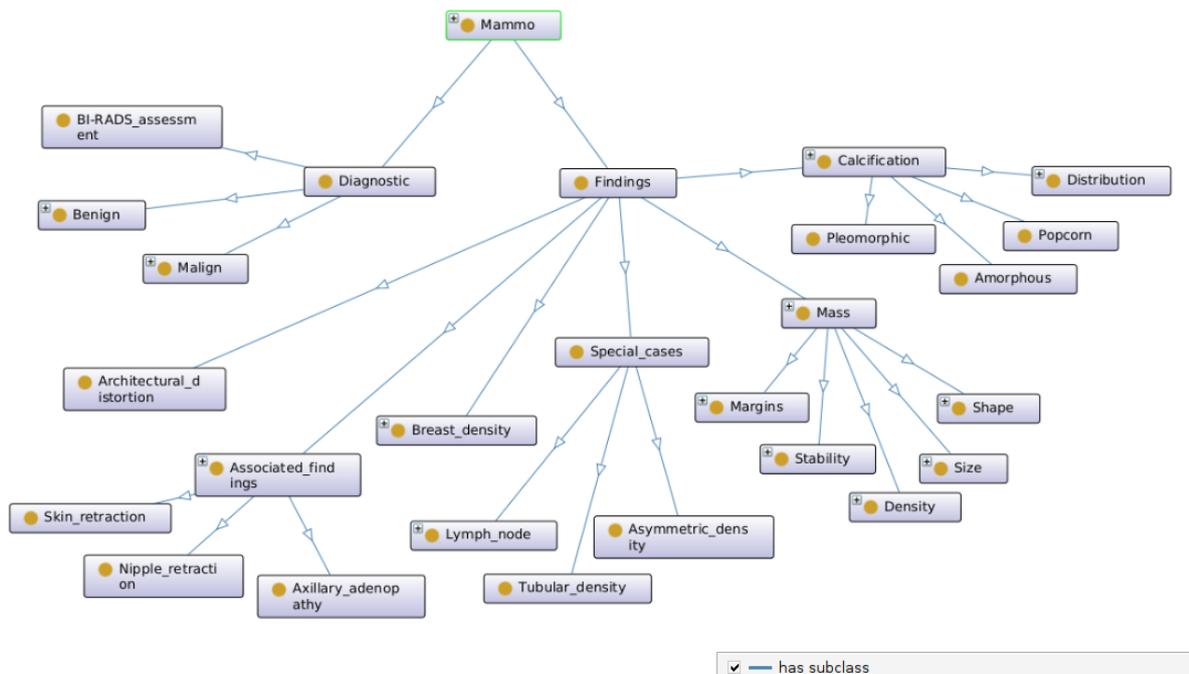
A ontologia de domínio Mammo é construída com base em ontologias e publicações científicas já convalidadas pela comunidade científica. São reutilizadas terminologias, vocabulário médico e relacionamentos. Na figura 35 é possível visualizar um extrato da Mammo.

A ontologia Mammo é baseada na ontologia de domínio de anotação de mamografia *Mammography Annotation Ontology* (MAO) (BULU; ALPKOCAK; BALCI, 2012). Além de parte do vocabulário controlado que descreve as características observadas nas mamografias pelos radiologistas, por meio de algumas das variáveis encontradas em *American College of Radiology National Mammography Database*, (OSUCH et al., 1995; BURNSIDE et al., 2009; CHHATWAL et al., 2009).

Algumas informações adicionais, como comentários explicativos com a definição de determinado conceito, são extraídas da ontologia *Radlex*, via API do BioPortal¹ e também do guia de mamografia (LATTANZIO; SIMONETTI, 2010).

Adicionalmente, foram incorporados na ontologia Mammo alguns termos relacionados com o diagnóstico patológico de biópsias (WU et al., 2019). E, também, termos relacionados aos casos especiais envolvendo *intramammary lymph node*, obtidos em Bitencourt et al. (2019).

Figura 35 – Extrato da ontologia de domínio com classes e propriedades relacionados ao domínio — Ontologia Mammo



Fonte: Elaborada pela autora.

A ontologia Mammo tem duas subclasses principais, a classe *Diagnostic* e a classe

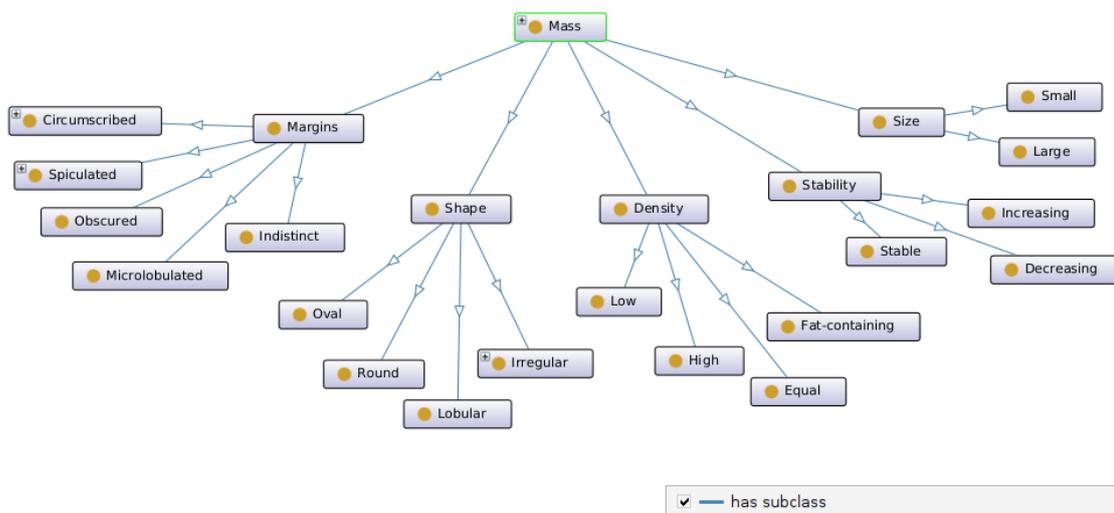
¹ bioportal.bioontology.org/ontologies/RADLEX/

Findings. Na classe **Diagnostic** tem como subclasse **BI-RADS** que inclui as diferentes categorias de classificação, e a subclasse **Pathologic_diagnostic** que inclui a classificação proveniente do diagnóstico patológico realizado na análise de biópsias **Benign** e **Malign**.

A classe **Findings** tem as seguintes subclasses: a classe **Architectural_distortion**, a classe **Associated_findings**, a classe **Breast_density**, a classe **Calcification**, a classe **Mass**, e a classe **Special_cases**.

Na figura 36 é apresentado o diagrama da classe **Mass**. Composta por: a classe **Margins** e suas subclasses **Circumscribed**, **Indistinct**, **Microlobulated**, **Obscured**, **Spiculated**. Classe **Shape** e subclasses **Oval**, **Round**, **Lobular**, **Irregular**. Classe **Density** e subclasses **Low**, **High**, **Equal**, **Fat-containing**. Classe **Stability** e subclasses **Stable**, **Decreasing**, **Increasing**. Classe **Size** e subclasses **Small**, **Large**.

Figura 36 – Extrato da ontologia de domínio — Classe **Mass**

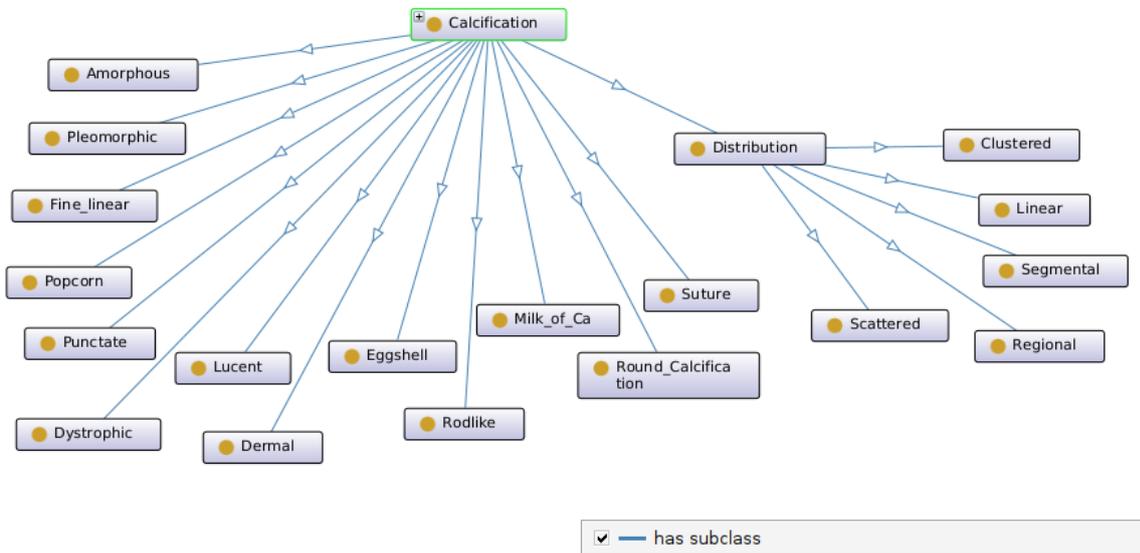


Fonte: Elaborada pela autora.

Na figura 37 é apresentado o diagrama da classe **Calcification**. Composta pelas classes **Amorphous**, **Pleomorphic**, **Dermal**, **Dystrophic**, **Eggshell**, **Fine_linear**, **Lucent**, **Milk_of_Ca**, **Popcorn**, **Punctate**, **Rodlike**, **Round_Calcification**, **Suture**. E, pela classe **Distribution** e subclasses **Clustered**, **Linear**, **Regional**, **Scattered**, **Segmental**.

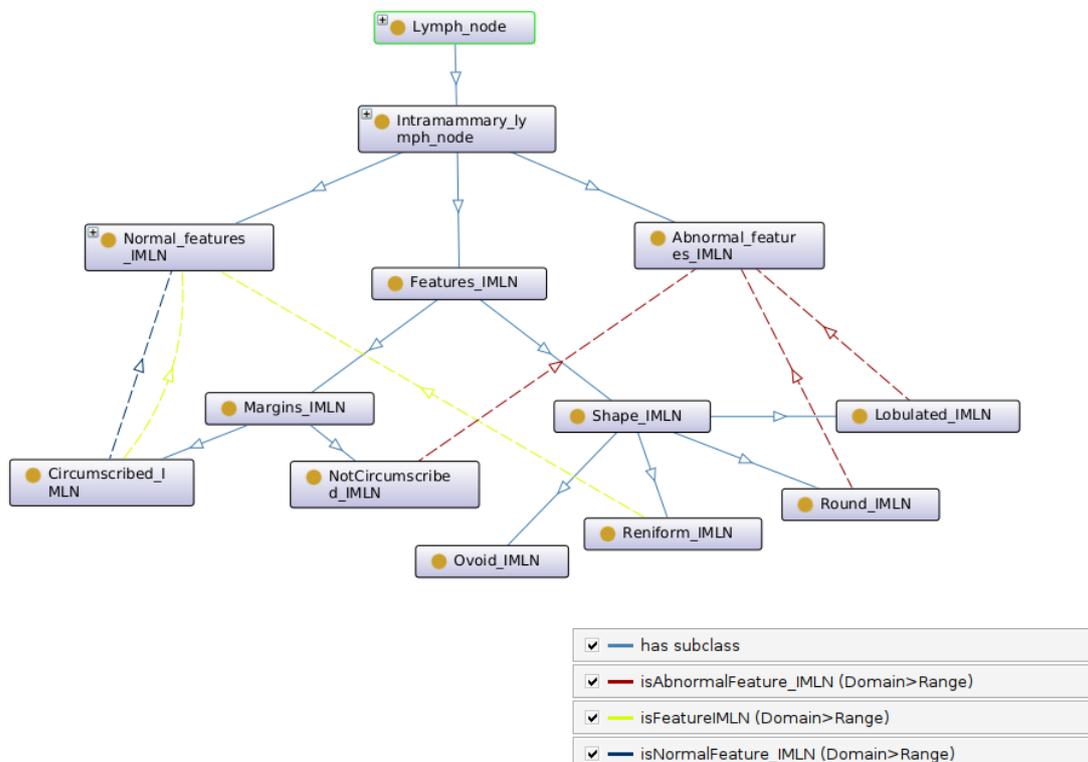
Na figura 38 é apresentado o diagrama da classe **Lymph_node**. Como é possível observar, a classe **Intramammary_lymph_node** representa um tipo especial da classe massa. Esses conceitos e relações foram obtidos em Bitencourt et al. (2019). Por meio dos relacionamentos é possível inferir se dada característica é normal ou anormal.

Figura 37 – Extrato da ontologia de domínio — Classe Calcification



Fonte: Elaborada pela autora.

Figura 38 – Extrato da ontologia de domínio — Classe Lymph_node



Fonte: Elaborada pela autora.

4.7 Considerações Finais

A abordagem desenvolvida neste trabalho permite, a partir de mamografia, classificá-la e ir mais além, adicionando anotações semânticas e expandindo-as, a partir do conhecimento incorporado na Ontologia AnotaMammo e usando o raciocinador para as regras.

O desenvolvimento da abordagem para anotação semântica de imagens é composto pelas seguintes etapas:

- ❑ Trabalho com várias bases de mamografia: para desenvolver estratégia robusta e o mais completa possível para a anotação de mamografias;
- ❑ Mineração de texto de laudos médicos: para identificar relações de similaridade e coocorrência entre pares de palavras presentes em um mesmo laudo. Os resultados são utilizados para geração de regras semânticas;
- ❑ Mineração de regras de associação: para identificação de outras regras e relações semânticas entre os *itemsets* frequentes;
- ❑ Classificação via *deep learning* de imagens com base em suas características visuais: geração de conceitos iniciais para iniciar o processo de anotação semântica de imagens médicas via termo resultante da classificação. *Deep learning* foi usado por ser o método de classificação que apresenta melhor resultado na literatura;
- ❑ Uso de Ontologia de domínio: para estruturar adequadamente o conhecimento sobre o domínio;
- ❑ Uso de Ontologia para anotação: para promover o processo de anotação semântica das imagens.

Os experimentos realizados para o método MUSA são apresentados no próximo capítulo.

Capítulo 5

Experimentos

Neste capítulo são apresentados os experimentos realizados nesta pesquisa de doutorado, detalhando sua implementação, configuração e resultados.

5.1 Implementação

Nesta seção serão apresentadas as técnicas e ferramentas utilizadas para a implementação dos experimentos desenvolvidos neste trabalho de doutorado. Os algoritmos de mineração de dados, tanto para imagens quanto para textos, são implementados na linguagem *Python*¹, no ambiente *Colaboratory*². *Python* que é uma linguagem amplamente utilizada na área de ciência de dados e possui inúmeras bibliotecas para manipulação, análise de dados, aprendizado de máquina e *deep learning*, tais como, *Numpy*³, *Pandas*⁴, *Scikit-Learn*⁵, *Keras*⁶ e *TensorFlow*⁷. O ambiente de trabalho *Weka*⁸ junto ao algoritmo *Apriori* (AGRAWAL; SRIKANT et al., 1994) também é utilizado neste trabalho. A linguagem de desenvolvimento da ontologia é a *OWL2* e para as regras e consultas *SWRL* e *SQWRL*, o ambiente de desenvolvimento utilizado é o *Protégé*⁹ e o como raciocinador o *Pellet*.

¹ python.org

² colab.research.google.com

³ numpy.org

⁴ pandas.pydata.org

⁵ scikit-learn.org

⁶ keras.io

⁷ tensorflow.org

⁸ cs.waikato.ac.nz/ml/weka

⁹ protege.stanford.edu

5.2 Bases de Dados

5.2.1 CBIS-DDSM

Curated Breast Imaging Subset of DDSM (CBIS-DDSM), é uma versão atualizada e padronizada do DDSM. O CBIS-DDSM é um conjunto de dados público composto por casos de calcificação e massa, incluindo imagens completas de mamografia e suas regiões de interesse (ROIs). As imagens estão disponíveis no formato DICOM e são compostas por imagens da mama direita e da mama esquerda na incidência craniocaudal (CC) e na incidência mediolateral oblíqua (MLO). ROIs são segmentações que correspondem a cada lesão mamária encontrada em cada imagem. Além disso, o conjunto de dados contém subconjuntos de treinamento e teste para achados clínicos de massa e calcificação (LEE et al., 2017; LEE et al., 2016).

5.2.2 CDD-CESM

A base *Categorized Digital Database for Low energy and Subtracted Contrast Enhanced Spectral Mammography images* (CDD-CESM) (KHALED et al., 2022; KHALED et al., 2021) é composta por imagens de mamografia espectral com contraste. Essa é uma modalidade recente de obtenção de imagens que utiliza contraste iodado, ideal para o rastreamento em mamas densas ou com achados na mamografia convencional de difícil interpretação. A base CDD-CESM inclui 2006 imagens com visualizações CC e MLO, entre elas, 310 imagens de massa e 238 imagens de calcificação. Essa base foi incorporada a este trabalho de doutorado por apresentar relatórios médicos com rico conteúdo em relação ao diagnóstico, recurso essencial para extração de conhecimento que apoia a construção de regras semânticas. Os relatórios médicos foram subdivididos pelo tipo de achado clínico. Foram utilizados os achados referentes às seguintes categorias, *mass_description*, *mass_enhancement_description*, *calcification* e *allFindings*.

5.2.3 MAMMOSET

A base de dados *MAMMOSET* reúne três bases de dados de mamografia: DDSM, Mini-Mias e Vienna, bases de dados reconhecidas pela comunidade acadêmica para validação de métodos para análise de dados de imagens de mamografia. Os autores integraram essas três bases para disponibilizar em um único conjunto, de maneira padronizada e organizada, os ROIs das imagens de mamografia, os vetores de características extraídos dos ROIs e os respectivos metadados (OLIVEIRA et al., 2017).

A base Vienna foi criada pelo *Department of Radiology at University of Vienna* e é composta por mamografias coletadas do BI-RADS. São 447 imagens de ROIs, de 256 x 256 *pixels* de resolução. Os metadados incluem o tipo de lesão (*mass* e *calcification*), seus

subtipos como, por exemplo, *circumscribed* um tipo de margem da massa ou *amorphous* um tipo de calcificação, BI-RADS, entre outros.

5.2.4 VIENNA

A base de dados Vienna também faz parte do acervo do Grupo de Pesquisa de Banco de Dados do PPGCC, nessa versão a base conta com 446 imagens de ROIs de mamografia e seus respectivos laudos médicos. Nesses laudos estão presentes os achados clínicos e o diagnóstico histopatológico relativo ao ROI, as descrições encontradas nesses laudos (L) são compostas por palavras (p_n) que pertencem a um vocabulário controlado. Os laudos são formatados da seguinte maneira:

L(identificador imagem) -> (BI-RADS, $p_1, p_2, p_3, \dots, p_n$)

Conforme pode ser observado na figura 39, cada imagem contém o identificador, um valor de *BI-RADS* e as palavras que compõe o laudo. Por exemplo, a imagem com identificador 4144907, possui *BI-RADS* = 3, e as palavras *fibroadenoma*, *mass* e *circum* que corresponde à palavra *circumscribed*. Algumas palavras aparecem abreviadas e algumas aparecem como siglas. As palavras que compõe o vocabulário controlado serão detalhadas posteriormente nesta tese.

Os dados presentes nos laudos são usados para enriquecer a construção das regras que formam o modelo estruturado de conhecimento. Aqui essas informações são obtidas por meio de técnicas e métodos de processamento de linguagem natural aplicados sobre os laudos.

5.3 Mineração de Texto

O processo de mineração de texto apresentado na seção 4.4, tem por objetivo a representação de palavras, a identificação de coocorrências e a similaridade entre pares de palavras. Para a realização do processo de mineração de texto foram utilizados algoritmos tradicionais de PLN^{10,11}, e adaptados de implementações disponibilizadas pela comunidade acadêmica¹².

5.3.1 Pré-Processamento

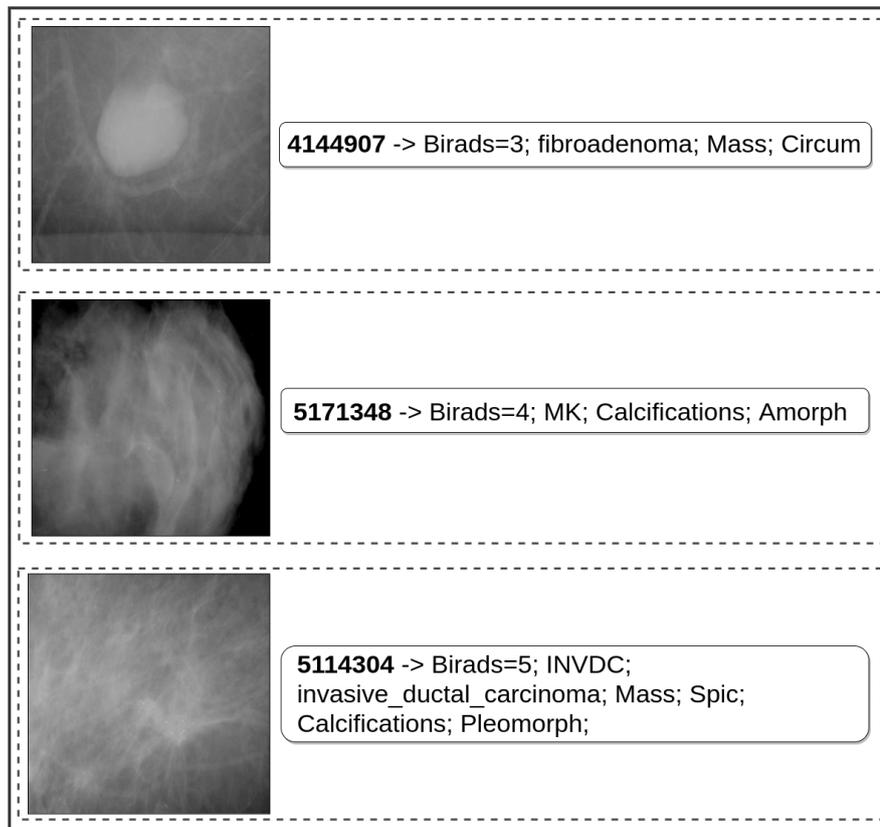
Inicialmente ocorre uma etapa de preparação dos dados. Os laudos selecionados são analisados e ajustados para serem compatíveis com o processamento computacional. As palavras com abreviações ou siglas são substituídas pela sua forma extensa. Isso ocorre

¹⁰ <https://web.stanford.edu/~jurafsky/li15/lec3.vector.pdf>

¹¹ <https://courses.engr.illinois.edu/cs447/fa2018/Slides/Lecture17HO.pdf>

¹² <https://www.kaggle.com/kenshoresearch/kdwd-pmi-word-vectors>

Figura 39 – Exemplos de ROIs de imagens de mamografia e os respectivos laudos



Fonte: Elaborada pela autora.

quando a descrição correta está presente na documentação da base. A sigla é excluída caso não tenha a palavra correspondente na documentação. São acrescentados o símbolo de *underline* (_) entre as palavras compostas, para que as mesmas sejam caracterizadas como um único *token* no próximo passo. Além disso, ocorre a tarefa de normalização das palavras para o formato minúsculo.

Como exemplo, na base Vienna, os laudos originais possuem uma série de simplificações de palavras ou abreviações. As abreviações e o tipo de lesão, podem ser conferidos na tabela 8. Nos laudos aqui analisados, os tipos de lesão podem ser benignas, de alto-risco ou malignas, esses tipos representam a gravidade da lesão. Na tabela 9 podem ser observados as 26 palavras e seus respectivos números de ocorrências, a tabela está ordenada pela palavra mais comum, no caso é *mass* que ocorre 228 vezes.

Para que os laudos se tornem legíveis pelo computador é necessário transformá-los em dados que permitam a leitura e interpretação por algoritmos de mineração de dados. Para isso, é necessário a realização de algumas etapas que compõem o processo de mineração. Entre elas, o pré-processamento que engloba a identificação de unigramas, de *skipgrams*, e a representação no formato de matriz. Além disso, ocorre o cálculo de similaridade sobre

Tabela 8 – Abreviações de diagnóstico histológico

Diagnóstico	Abreviação
<i>Benign Lesions</i>	
<i>breast tissue</i>	BT
<i>cyst</i>	C
<i>calcifications</i>	CALC
<i>ductal hyperplasia</i>	DH
<i>fibrosis</i>	F
<i>fibroadenoma</i>	FA
<i>fatty tissue</i>	FT
<i>haematoma</i>	HAEM
<i>harmartoma</i>	HARM
<i>lymphangioma</i>	LA
<i>mastitis</i>	MAST
<i>mastopathia</i>	MP
<i>papilloma</i>	MP
<i>sclerosing adenosis</i>	SA
<i>scar</i>	SCR
<i>High-risk Lesions</i>	
<i>atypical ductal hyperplasia</i>	ADH
<i>lobular carcinoma in situ</i>	LCIS
<i>phyllodes tumor</i>	PHYL
<i>radial scar</i>	RS
<i>Malignant Lesions</i>	
<i>ductal carcinoma in situ</i>	DCIS
<i>invasive ductal cancer</i>	INVDC
<i>invasive lobular cancer</i>	INVLC
<i>invasive tub. cancer</i>	INVTC
<i>muc. cancer</i>	MC

a matriz esparsa procedente da contagem simples de coocorrência entre as palavras. Bem como, sobre as matrizes resultantes do PMI e suas variantes e do SVD.

No pré-processamento, a primeira tarefa é a tokenização, em que cada palavra é representada no formato de um *token*. Tem-se então listas de *tokens*, conforme pode ser observado abaixo. Após isso, ocorre a tarefa de remoção dos *stop words* que são aquelas palavras comuns que não acrescentam muita informação ao contexto como, por exemplo, (*a, an, are, by, has...*).

```
['fibroadenoma', 'mass', 'circumscribed'],
    ['calcification', 'amorph'],
['invasive_ductal_carcinoma', 'mass', 'speculated', ...
    ... 'calcification', 'pleomorph']
```

O próximo passo consiste na aplicação do algoritmo para identificar os unigramas do vocabulário. Um unigrama é a união de uma palavra e a quantidade de vezes que a mesma aparece no *corpus*. Na etapa seguinte é aplicado o algoritmo *skipgram*. Cada palavra da lista de *tokens* é percorrida (palavra foco ou *word*) e, em seguida, são formados os *skipgrams* com palavras anteriores e posteriores à palavra foco (palavras de contexto ou *context*). Isso é feito por meio de janelas de tamanho pré-determinado pelo usuário. Por exemplo, se o tamanho da janela é "3" as três palavras anteriores e as três posteriores serão percorridas, essas são as palavras de contexto. Nessa etapa é possível determinar a coocorrência entre as palavras. Um exemplo, considerando uma janela de tamanho 3, é ilustrado abaixo.

```
fibroadenoma fibrosis mass circumscribed
```

```
( 'fibroadenoma', 'fibrosis' )
( 'fibroadenoma', 'mass' )
( 'fibroadenoma', 'circumscribed' )
( 'fibrosis', 'fibroadenoma' )
( 'fibrosis', 'mass' )
( 'fibrosis', 'circumscribed' )
( 'mass', 'fibroadenoma' )
( 'mass', 'fibrosis' )
( 'mass', 'circumscribed' )
( 'circumscribed', 'fibroadenoma' )
( 'circumscribed', 'fibrosis' )
( 'circumscribed', 'mass' )
```

Neste trabalho, estabeleceu-se que o tamanho da janela padrão é o tamanho máximo de *tokens* que um laudo do *corpus* pode ter. Assim, assegura-se o relacionamento entre todas as palavras contidas no laudo, ora sendo *word*, ora sendo *context*. Alguns dos achados de unigramas e *skipgrams* são demonstrados a seguir.

Na base Vienna foram identificados 30 unigramas e 312 *skipgrams*, os exemplos mais frequentes de unigramas e *skipgrams* são demonstrados na tabela 9.

Tabela 9 – Base Vienna. Exemplos de ocorrências de unigramas(a) e *skipgrams* (b)

(a) Unigramas		(b) <i>Skipgrams</i>	
Palavra	Ocorrência	<i>Skipgram</i>	Ocorrência
<i>mass</i>	228	(<i>calcification, pleomorph</i>)	105
<i>calcification</i>	224	(<i>calcification, amorph</i>)	94
<i>pleomorph</i>	105	(<i>mass, circumscribed</i>)	90
<i>amorph</i>	94	(<i>mass, indistinct</i>)	69
<i>circumscribed</i>	90	(<i>invasive_ductal_carcinoma, mass</i>)	61
<i>fibroadenoma</i>	89	(<i>fibroadenoma, mass</i>)	59
<i>mastopathia</i>	87	(<i>mass, speculated</i>)	58
<i>invasive ductal carcinoma</i>	86	(<i>mass, calcification</i>)	46
<i>indistinct</i>	69	(<i>mastopathia, calcification</i>)	40
<i>speculated</i>	58	(<i>invasive_ductal_carcinoma, speculated</i>)	40

Na base CDD-CESM, na categoria *Mass description*, foram identificados 105 unigramas e 2902 *skipgrams*. Na tabela 10 é possível observar alguns exemplos de unigramas e *skipgrams* com os respectivos valores de ocorrências. Como pode ser observado, a palavra *mass* ocorre 214 vezes, e o *skipgram* (*malignant, irregular*) ocorre 150 vezes.

Tabela 10 – Base CDD-CESM - *Mass description*. Exemplos de ocorrências de unigramas (a). Exemplos de ocorrências de *skipgrams* (b)

(a) Unigramas		(b) <i>Skipgrams</i>	
Palavra	Ocorrência	<i>Skipgram</i>	Ocorrência
<i>mass</i>	214	(<i>malignant, irregular</i>)	150
<i>malignant</i>	186	(<i>benign, circumscribed</i>)	102
<i>benign</i>	177	(<i>speculated, irregular</i>)	101
<i>irregular</i>	168	(<i>malignant, Birads_5</i>)	100
<i>circumscribed</i>	116	(<i>speculated, malignant</i>)	96
<i>speculated</i>	102	(<i>benign, Birads_2</i>)	79
<i>round</i>	101	(<i>oval, benign</i>)	75
<i>oval</i>	82	(<i>malignant, Birads_4</i>)	75
<i>Birads_4</i>	206	(<i>circumscribed, round</i>)	73
<i>Birads_3</i>	74	(<i>benign, Birads_3</i>)	67

Na categoria *Mass Enhancement*, foram identificados 92 unigramas e 1788 *skipgrams*. Na tabela 11 é possível observar alguns exemplos de unigramas e *skipgrams* e seus respectivos valores de ocorrências. Como pode ser observado, a palavra *malignant* ocorre 252 vezes e para a palavra *irregular* são 251 ocorrências. O *skipgram* (*malignant, irregular*) ocorre 241 vezes e (*speculated, malignant*) 110 vezes.

Tabela 11 – Base CDD-CESM - *Mass enhancement*. Exemplos de ocorrências de unigramas(a). Exemplos de ocorrências de *skipgrams* (b)

(a) Unigramas		(b) Skipgrams	
Palavra	Ocorrência	<i>Skipgram</i>	Ocorrência
<i>malignant</i>	252	(<i>malignant, irregular</i>)	241
<i>irregular</i>	251	(<i>irregular, enhancing</i>)	223
<i>heterogenously</i>	206	(<i>malignant, enhancing</i>)	217
<i>mass</i>	200	(<i>irregular, heterogenously</i>)	200
<i>speculated</i>	110	(<i>malignant, heterogenously</i>)	197
<i>benign</i>	82	(<i>irregular, Birads_5</i>)	181
<i>circumscribed</i>	80	(<i>malignant, Birads_5</i>)	179
<i>round</i>	68	(<i>speculated, malignant</i>)	110
<i>oval</i>	36	(<i>circumscribed, oval</i>)	34
<i>nipple_retraction</i>	15	(<i>oval, benign</i>)	32

Na categoria *Calcification*, foram identificados 84 unigramas e 1712 *skipgrams*. Na tabela 12 é possível observar alguns exemplos de unigramas e *skipgrams* e seus respectivos valores de ocorrências. A palavra *macrocalcification* ocorre 127 vezes e a palavra *microcalcification* 77 vezes. O *skipgram* (*malignant, microcalcification*) ocorre 71 vezes e (*malignant, clusters*) 23 vezes.

Tabela 12 – Base CDD-CESM - *Calcification*. Exemplos de ocorrências de unigramas (a). Exemplos de ocorrências de *skipgrams* (b)

(a) Unigramas		(b) Skipgrams	
Palavra	Ocorrência	<i>Skipgram</i>	Ocorrência
<i>benign</i>	145	(<i>benign, macrocalcification</i>)	117
<i>macrocalcification</i>	127	(<i>malignant, microcalcification</i>)	71
<i>malignant</i>	116	(<i>malignant, overlying</i>)	60
<i>microcalcification</i>	77	(<i>overlying, microcalcification</i>)	50
<i>edema_pattern</i>	25	(<i>malignant, clusters</i>)	23
<i>clusters</i>	27	(<i>malignant, diffuse</i>)	22
<i>diffuse</i>	24	(<i>diffuse, edematous</i>)	22
<i>pleomorphic</i>	21	(<i>pleomorphic, clusters</i>)	21
<i>vascular</i>	20	(<i>vascular, benign</i>)	14
<i>nipple_retraction</i>	8	(<i>pleomorphic, microcalcifications</i>)	13

Na categoria *AllFindings* foram identificados 192 unigramas e 5742 *skipgrams*. Na base CBIS-DDSM, para *Mass* foram identificados 43 unigramas e 470 *skipgrams*. Para *calcification* foram identificados 66 unigramas e 472 *skipgrams*.

5.3.2 Cálculo de Similaridade

Após a etapa de formação dos *skipgrams*, são construídas as matrizes simétricas *word*, *context* com a frequência por contagem de pares de palavras. Após isso, essa matriz é transformada com o PMI e suas variantes. Cada uma dessas matrizes é a entrada para o cálculo da similaridade entre os pares de palavras, usando a medida de similaridade do cosseno. Alguns dos resultados dessas etapas são demonstrados adiante.

Na tabela 13 são apresentados os resultados de similaridade de *context* com a palavra foco (*word*), que nesse caso é a palavra *Benign*.

Tabela 13 – Base CDD-CESM - *Mass Enhancement*. Similaridades com a palavra foco (*word*) *Benign*. Em (a) exemplos de palavras *context* e a similaridade (Sim), considerando a matriz de contagem de frequência. Em (b) os resultados de similaridade com a matriz PPMI suavizado. Em (c) os resultados com SVD

(a) Contagem		(b) PPMI suavizado		(c) SVD	
<i>Context</i>	Sim	<i>Context</i>	Sim	<i>Context</i>	Sim
<i>oval</i>	0.80	<i>circumscribed</i>	0.86	<i>circumscribed</i>	0.89
<i>ring</i>	0.76	<i>oval</i>	0.71	<i>focus_enhancement</i>	0.77
<i>homogenously</i>	0.76	<i>Birads_3</i>	0.70	<i>Birads_3</i>	0.68
<i>Birads_3</i>	0.76	<i>round</i>	0.69	<i>imln</i>	0.60
<i>round</i>	0.75	<i>Birads_2</i>	0.56	<i>ring</i>	0.57

Na tabela 14 são apresentados os resultados de similaridade de *context* com a palavra foco (*word*), que nesse exemplo é a palavra *Circumscribed*. Podem ser observadas as palavras *context* e os resultados de similaridade (Sim) utilizando a matriz de contagem de frequência, os resultados de similaridade com a matriz PPMI suavizado e os resultados com SVD.

Tabela 14 – Base Vienna. Similaridades com a palavra foco (*word*) *Circumscribed*. Em (a) exemplos de palavras *context* e a similaridade (Sim), considerando a matriz de contagem de frequência. Em (b) os resultados de similaridade com a matriz PPMI suavizado. Em (c) os resultados com SVD

(a) Contagem		(b) PPMI suavizado		(c) SVD	
<i>Context</i>	Sim	<i>Context</i>	Sim	<i>Context</i>	Sim
<i>fibroadenoma</i>	0.78	<i>fibroadenoma</i>	0.58	<i>Birads_3</i>	0.88
<i>fibrosis</i>	0.72	<i>Birads_3</i>	0.48	<i>fibroadenoma</i>	0.81
<i>harmartomo</i>	0.67	<i>harmartomo</i>	0.37	<i>mass</i>	0.66
<i>Birads_3</i>	0.64	<i>fibrosis</i>	0.36	<i>fibrosis</i>	0.48
<i>papilloma</i>	0.58	<i>mastitis</i>	0.27	<i>harmartomo</i>	0.40

Na tabela 15 são apresentados os resultados de similaridade de *context* com a palavra foco (*word*), que nesse exemplo é a palavra *Malignant*. Podem ser observadas as palavras *context* e os resultados de similaridade (Sim) utilizando a matriz de contagem de frequência, os resultados de similaridade com a matriz PPMI suavizado e os resultados com SVD.

Tabela 15 – Base CDD-CESM — *Mass Enhancement*. Similaridades com a palavra foco (*word*) *Malignant*. Em (a) exemplos de palavras *context* e a similaridade (Sim), considerando a matriz de contagem de frequência. Em (b) os resultados de similaridade com a matriz PPMI suavizado. Em (c) os resultados com SVD.

(a) Contagem		(b) PPMI suavizado		(c) SVD	
<i>Context</i>	Sim	<i>Context</i>	Sim	<i>Context</i>	Sim
<i>nipple_retraction</i>	0.88	<i>heterogenously</i>	0.79	<i>irregular</i>	0.94
<i>speculated</i>	0.87	<i>irregular</i>	0.78	<i>heterogenously</i>	0.75
<i>surrounding_satellites</i>	0.83	<i>Birads_5</i>	0.77	<i>enhancing</i>	0.72
<i>mass</i>	0.82	<i>speculated</i>	0.76	<i>Birads_5</i>	0.69
<i>Birads_5</i>	0.82	<i>retraction</i>	0.37	<i>speculated</i>	0.57

Na tabela 16 são apresentados os resultados de similaridade de *context* com a palavra foco (*word*), que nesse exemplo é a palavra *Speculated*. Podem ser observadas as palavras *context* e os resultados de similaridade (Sim) utilizando a matriz de contagem de frequência, os resultados de similaridade com a matriz PPMI suavizado e os resultados com SVD.

Tabela 16 – Base CBIS-DDSM — *Mass*. Similaridades com a palavra foco (*word*) *Speculated*. Em (a) exemplos de palavras *context* e a similaridade (Sim), considerando a matriz de contagem de frequência. Em (b) os resultados de similaridade com a matriz PPMI suavizado. Em (c) os resultados com SVD.

(a) Contagem		(b) PPMI suavizado		(c) SVD	
<i>Context</i>	Sim	<i>Context</i>	Sim	<i>Context</i>	Sim
<i>obscured-speculated</i>	0.87	<i>obscured-</i>	0.78	<i>architectural-</i>	0.94
<i>ill_defined</i>	0.84	<i>spiculated</i>		<i>distortion</i>	
<i>architectural-</i>	0.76	<i>round-irregular</i>	0.70	<i>round-irregular</i>	0.35
<i>distortion</i>		<i>Birads_5</i>	0.59	<i>Birads_5</i>	0.33
<i>Birads_5</i>	0.74	<i>ill_defined</i>	0.58	<i>irregular</i>	0.26
<i>irregular</i>	0.72	<i>irregular</i>	0.55	<i>lobulated-oval</i>	0.12

Na tabela 17 são apresentados os resultados de similaridade de *context* com a palavra foco (*word*), que nesse exemplo é a palavra *Malignant*. Podem ser observadas as palavras *context* e os resultados de similaridade (Sim) utilizando a matriz de contagem de

frequência, os resultados de similaridade com a matriz PPMI suavizado e os resultados com SVD.

Tabela 17 – Base CBIS-DDSM — *Calcification*. Similaridades com a palavra foco (*word*) *Malignant*. Em (a) exemplos de palavras *context* e a similaridade (Sim), considerando a matriz de contagem de frequência. Em (b) os resultados de similaridade com a matriz PPMI suavizado. Em (c) os resultados com SVD.

(a) Contagem		(b) PPMI suavizado		(c) SVD	
<i>Context</i>	Sim	<i>Context</i>	Sim	<i>Context</i>	Sim
<i>pleomorphic</i>	0.69	<i>clustered-</i>	0.70	<i>punctate-</i>	0.71
<i>punctate-</i>	0.67	<i>linear</i>		<i>amorphous-</i>	
<i>pleomorphic</i>		<i>linear</i>	0.63	<i>pleomorphic</i>	
<i>Birads_4</i>	0.66	<i>fine-linear-</i>	0.53	<i>linear</i>	0.58
<i>pleomorphic-</i>	0.62	<i>branching</i>		<i>pleomorphic</i>	0.40
<i>amorphous</i>		<i>diffusely-</i>	0.50	<i>clustered</i>	0.36
<i>punctate-</i>	0.62	<i>scattered</i>		<i>fine-linear-</i>	0.35
<i>amorphous</i>		<i>Birads_5</i>	0.46	<i>branching</i>	

Os resultados gerados nas etapas descritas anteriormente são os recursos utilizados para a construção das regras semânticas. A seleção do tipo de matriz e das similaridades foi empírica, pois cada uma se comportou melhor em diferentes situações.

5.3.3 Regras de Associação

Além do processo clássico de PLN para cálculo de similaridade entre pares de palavras, neste trabalho de doutorado, também são utilizados os recursos produzidos pelas regras de associação com o algoritmo *Apriori* (AGRAWAL; SRIKANT et al., 1994).

As regras de associação têm por objetivo fornecer recursos adicionais para a construção de regras semânticas e contribuem para confirmar algumas das regras obtidas nos passos anteriores. Foi utilizada a implementação do *Apriori* disponível no Weka. Foi estabelecido como confiança o mínimo de 0.5 (50%), e o suporte mínimo de 0.1 (10%). Alguns exemplos das regras geradas são demonstrados a seguir.

As regras R1 - R3 são obtidas da base Mammoset-Mias. A regra R1, por exemplo, tem como antecedente *circumscribed* e como conseqüente *benign*, com confiança de 0.83, *lift* de 1.47 e convicção de 2.07. Isso significa que em 83% das vezes que ocorrer *circumscribed* também pode ocorrer *benign*.

```
R1- circumscribed ==> benign <conf:(0.83)> lift:(1.47) conv:(2.07)
R2- circumscribed, mass ==> benign <conf:(0.83)> lift:(1.47) conv:(2.07)
R3- mass, benign ==> circumscribed <conf:(0.53)> lift:(2.59) conv:(1.59)
```

As regras R4 - R11 são obtidas da base Mammoset-Vienna. Por exemplo, na regra R5 o antecedente é *malignant* e *amorphous* e o conseqüente *Birads 4*. A confiança é 0.79, o *lift* de 1.52 e convicção de 2.18. Então, o *Birads* terá valor 4 em 79% das vezes que ocorrer o antecedente. Outro exemplo, na R6 o antecedente é *calcification* e *malignant* e o conseqüente *pleomorphic*. A confiança é 0.76, o *lift* de 3.69 e convicção de 3.14. Significa que quando ocorre *calcification* e *malignant* tem 3.69 vezes mais chance de ocorrer também *pleomorphic*.

```
R4- mass, Birads 3 ==> circumscribed <conf:(0.86)> lift:(4.25) conv:(5.01)
R5- amorphous, calcification ==> Birads 4 <conf:(0.79)> lift:(1.52) conv:(2.18)
    R6- calcification, malignant ==> pleomorphic
        <conf:(0.76)> lift:(3.69) conv:(3.14)
    R7- ill-defined, mass ==> Birads 4 <conf:(0.7)> lift:(1.33) conv:(1.49)
        R8- pleomorphic, calcification ==> malignant
            <conf:(0.68)> lift:(2.64) conv:(2.27)
R9- amorphous, calcification ==> benign <conf:(0.67)> lift:(1.31) conv:(1.43)
R10- circumscribed, mass ==> Birads 3 <conf:(0.66)> lift:(2.71) conv:(2.13)
R11- Birads 5 ==> spiculated, mass <conf:(0.51)> lift:(4.03) conv:(1.76)
```

As regras R12 - R15 são obtidas da base CBIS-DDSM - *Calcification*. Como pode ser observado na R12, o antecedente *pleomorphic* e *clustered* tem como conseqüente *Birads 4*, com confiança de 0.82, *lift* de 1.68 e convicção de 2.8. O valor de convicção está dentro do intervalo adequado para uma regra interessante, de acordo com (BRIN et al., 1997).

```
R12- pleomorphic, clustered ==> Birads 4 <conf:(0.82)> lift:(1.68) conv:(2.8)
    R13-clustered, malignant, Birads 4 ==> pleomorphic
        <conf:(0.78)> lift:(1.81) conv:(2.51)
    R14- pleomorphic, clustered, malignant ==> Birads 4
        <conf:(0.72)> lift:(1.47) conv:(1.78)
R15- clustered, malignant ==> Birads 4 <conf:(0.64)> lift:(1.31) conv:(1.4)
```

As regras R16 - R20 são obtidas da base CBIS-DDSM - *Mass*. Como exemplo, a regra R16 tem como antecedente *obscured* e como conseqüente *benign*, com confiança de 0.85, *lift* de 1.95 e convicção de 3.69.

```
R16 - obscured ==> benign <conf:(0.85)> lift:(1.95) conv:(3.69)
R17 - circumscribed ==> benign <conf:(0.74)> lift:(1.69) conv:(2.14)
    R18 - oval ==> benign <conf:(0.69)> lift:(1.57) conv:(1.77)
    R19 - ill-defined ==> malignant <conf:(0.63)> lift:(1.3) conv:(1.38)
R20 - malignant, spiculated ==> irregular <conf:(0.63)> lift:(2.35) conv:(1.95)
```

Os resultados das regras geradas foram utilizados para a construção de regras semânticas, que foram incorporadas como regras SWRL na AnotaMammo.

5.4 Mineração de Imagem

O processo de mineração de imagem apresentado na seção 4.5, tem por objetivo analisar e classificar as ROIs de imagens de mamografia. Num processo que engloba uma estratégia *end-to-end deep learning*. O processo inclui aprendizado por transferência (*transfer learning*), por meio do ajuste fino (*fine-tuning*) e definição de hiperparâmetros (*hyperparameter*), adaptado de (LAI, 2021). Foi utilizado o ambiente *ColabPro* para a realização dos experimentos, já que é necessário acessar *Graphics Processing Unit* (GPU) para o processamento.

5.4.1 Pré-Processamento

Os dados utilizados pertencem à base CBIS-DDSM. Foi utilizada uma versão no formato NumPy contendo um subconjunto dos dados de treinamento e teste. Cada subconjunto é composto por regiões de interesse (ROIs) que correspondem a uma lesão que pode ser do tipo, massa ou calcificação. E, também, do tipo benigno ou maligno.

Cada ROI é classificado individualmente. A classificação determina se a anormalidade detectada nesta ROI é uma massa ou calcificação, ou se é benigna ou maligna. O conjunto de treinamento é composto por 2676 ROIs, sendo 1218 massas e 1458 calcificações. O conjunto de teste é composto por 336 ROIs em que 179 massas e 157 calcificações, sendo 228 benignas e 108 malignas. Os dados de entrada são compostos por ROIs de tamanho 150x150. Os valores de *pixel* de cada imagem são normalizados para corresponder ao esperado para cada modelo.

Além disso, os dados de entrada foram transformados de 16 bits para 8 bits. Uma dimensão é criada para representar o único canal de cor (escala de cinza). Este canal é replicado três vezes para formar um canal de três cores (Red, Green and Blue (RGB)), usado pelas arquiteturas CNN. Os dados de treinamento são subdivididos em 2 para separar os dados de validação, resultando em um subconjunto de treinamento e um subconjunto de validação. O tamanho do *batch* é 128.

Para ajudar a reduzir o *overfitting*, foi utilizada a técnica de aumento de dados. *Data Augmented* é uma técnica para gerar novos dados a partir de transformações nos dados de entrada. É uma técnica eficaz para um problema de classificação (GOODFELLOW; BENGIO; COURVILLE, 2016). As transformações utilizadas nesta pesquisa de doutorado são apresentadas na tabela 18. Também foi executado o método de pré-processamento fornecido por cada arquitetura CNN sobre os dados de teste.

5.4.2 Hiperparametrização

A hiperparametrização visa definir qual otimizador e qual taxa de aprendizado (LR) são mais adequados. Os dois hiperparâmetros correspondem ao treinamento da rede. Vale

Tabela 18 – Aumento de Dados - transformação dos dados de treinamento em tempo de execução

Transformação	Valor
<i>rotation range</i>	180
<i>rotation shear range</i>	10
<i>zoom range</i>	0.2
<i>horizontal flip</i>	True
<i>fill mode</i>	'reflect'

ressaltar que os hiperparâmetros da estrutura de rede não foram modificados, pois foram utilizadas arquiteturas CNN definidas.

O experimento para definir o otimizador mais adequado foi realizado separadamente para cada arquitetura de CNN (DenseNet121, EfficientNet2VS, Xception). Foram avaliados os otimizadores disponíveis no TensorFlow: Adadelta, Adagrad, Adam, Adamax, Ftrl, Nadam e RMSprop. Cada otimizador foi testado com diferentes taxas de aprendizado, começando com um valor baixo de LR, que foi incrementado ao final de cada época.

Assim, o LR inicial é $2e-4$ e seus valores aumentam ligeiramente em 10% a cada época. Ao todo foram testadas 150 épocas para cada otimizador. Este experimento foi realizado para diferentes combinações de camadas de treinamento congeladas e escolhido a de melhor desempenho.

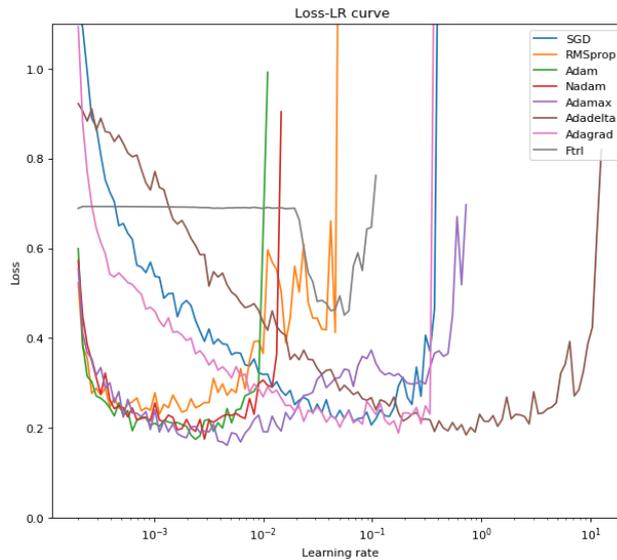
Um exemplo é apresentado na figura 40, que mostra os resultados obtidos em testes com EfficientNetV2S. O otimizador que obteve a menor perda foi o Adamax. Também é possível observar os valores aproximados para a taxa de aprendizado (LR), visto que o aumento da LR pode causar instabilidade após atingir o mínimo global.

Para encontrar o melhor valor de LR, foi testado quatro valores diferentes são eles LR = $[1e-5, 1e-4, 1e-3, 1e-2]$, durante 150 épocas, como pode ser observado na figura 41. A figura mostra que o LR de $LR = 10e - 2$ e $LR = 10e - 3$ têm resultados aproximados. Por esse motivo, esses dois LRs serão investigados na próxima etapa. O ajuste adequado de LR é essencial para obter melhores resultados.

Na figura 42 e na figura 43, é possível comparar os gráficos de acurácia e perda para $LR = 10e - 2$ e $LR = 10e - 3$, os gráficos mostram resultados bastante diferentes. Pode-se observar, no gráfico do lado direito da figura 42, uma grande abertura entre as curvas de *loss* do conjunto de treinamento e do conjunto de validação, caracterizando o *overfitting* ao utilizar o $LR = 10e - 2$. No gráfico do lado direito da figura 43, pode-se observar um *overfitting* mínimo. Portanto, ao utilizar $LR = 10e - 3$ ocorre uma melhora significativa do *overfitting*.

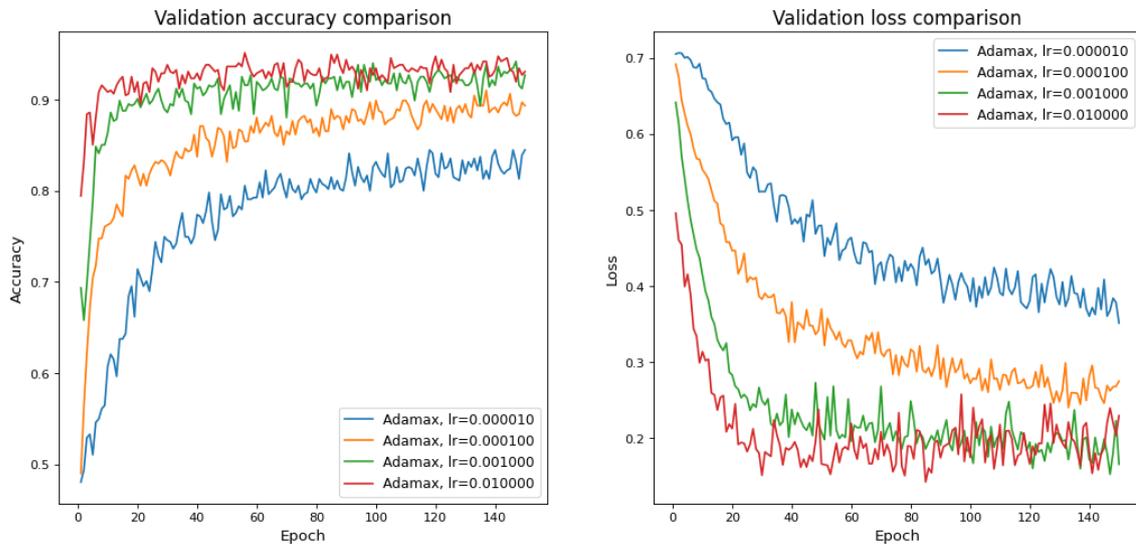
Para analisar a ocorrência de *overfitting*, foi analisado o gráfico de perdas para os experimentos realizados em cada arquitetura CNN. Além disso, para lidar com o *overfitting*,

Figura 40 – Gráfico para escolha do melhor otimizador para EfficientNetV2



Fonte: Elaborada pela autora.

Figura 41 – EfficientNetV2S - Gráfico com quatro valores diferentes de LR para determinar a melhor taxa de aprendizado



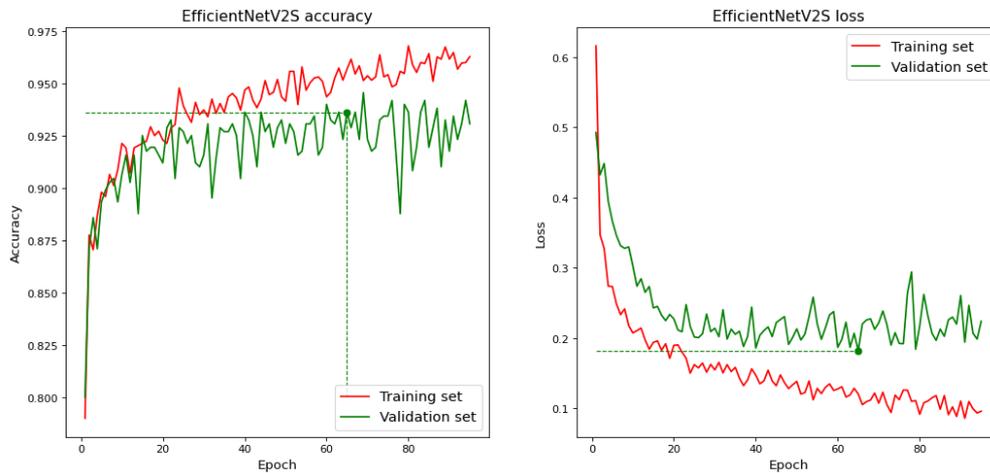
Fonte: Elaborada pela autora.

utilizaram-se alguns métodos de regularização, como aumento de dados, *dropout* e parada antecipada.

Alguns experimentos foram realizados alterando o número de camadas congeladas, foi considerado o congelamento das camadas iniciais e camadas finais. Apesar de ser mais usual o congelamento das camadas iniciais, os experimentos mostraram melhores resultados com o retreino das camadas iniciais e congelamento das posteriores, além do treino do topo da rede.

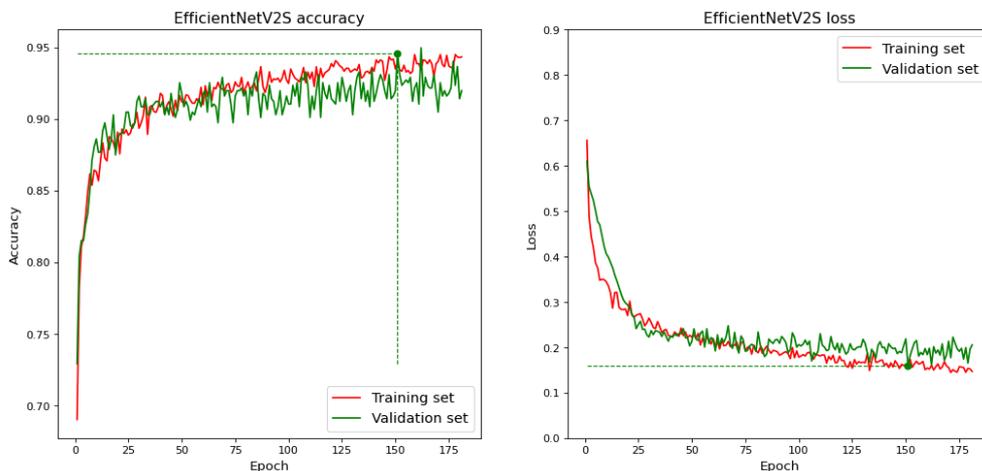
Acredita-se que isso ocorre devido à grande distância entre a similaridade dos dados da

Figura 42 – Gráfico de acurácia e perda para $LR = 1e - 2$. Ao lado esquerdo da figura o gráfico de acurácia e ao lado direito o gráfico de perda



Fonte: Elaborada pela autora.

Figura 43 – Gráfico de acurácia e perda para $LR = 1e - 3$. Ao lado esquerdo da figura o gráfico de acurácia e ao lado direito o gráfico de perda



Fonte: Elaborada pela autora.

base da rede pré-treinada e da base de dados analisada. Assim, ao ajustar as características de mais baixo nível, o modelo se torna mais adequado para a tarefa de classificação dos novos dados analisados.

Além disso, é necessário um menor tempo e custo de processamento para treinar somente poucas camadas e o novo topo, já que menos parâmetros precisam ser treinados. Para cada arquitetura foi identificado empiricamente um valor ideal de camadas iniciais treináveis. Esses ajustes ajudam a lidar com o *overfitting*.

Para as três arquiteturas foi utilizada a taxa de aprendizado $LR = 1e - 3$. Ademais, para a arquitetura *Densenet121* foi utilizado o otimizador *Nadam* e foram treinadas as 15

primeiras camadas. Para a *EfficientNet2VS* foi utilizado o otimizador *Adamax* e foram treinadas as 50 primeiras camadas. Para a *Xception* também foi utilizado o otimizador *Adamax* e foram treinadas as 40 primeiras camadas.

Os parâmetros gerais utilizados para as três arquiteturas são:

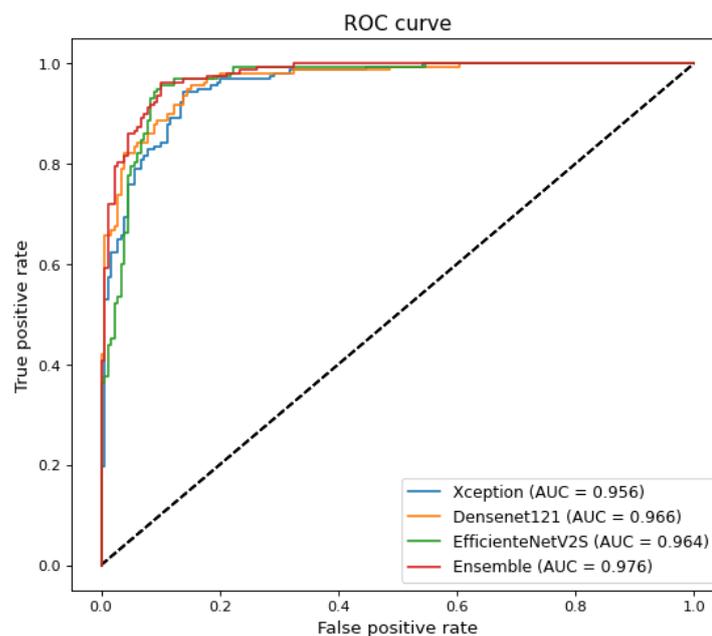
- ❑ *Early Stopping* (Early Stopping (ES)) - Monitora *Validation Loss* pelo número de épocas determinado pela *patience* (P); se não houver melhora, o treinamento é interrompido.
- ❑ *Dropout* - O valor de 0.5 para *dropout* é usado em todos os experimentos.
- ❑ *Function loss* - *Binary Crossentropy*
- ❑ *Checkpoint* - Monitora *Validation Loss* para salvar o melhor modelo.

5.4.3 Ensemble

Ensembles são métodos que combinam diferentes modelos para melhorar a classificação. O *ensemble* por votação classifica cada instância por votos, e o resultado é aquele que recebe mais votos. O *ensemble* de médias classifica cada instância pela previsão média.

Foram combinados os três modelos produzidos na forma de um *ensemble*. A figura 44 mostra a comparação da curva ROC e AUC de cada modelo com o *ensemble* proposto. O valor da acurácia do *ensemble* de média se destacou como o melhor resultado entre todos os modelos.

Figura 44 – ROC Curve - Ensemble



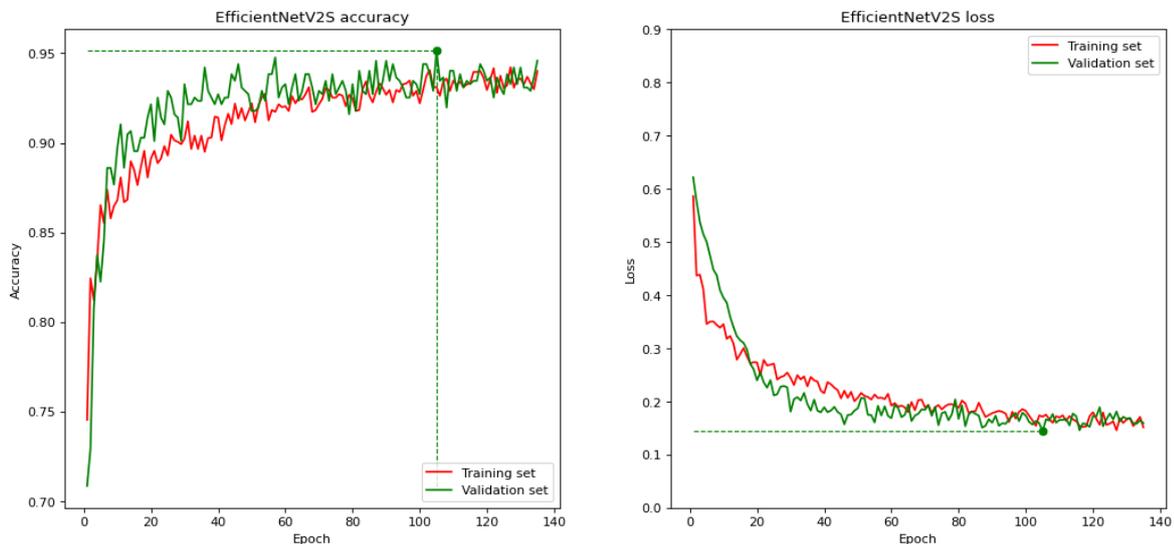
Fonte: Elaborada pela autora.

5.4.4 Resultados

Classificação Massa/Calcificação

Inicialmente, foi definido o número de 250 épocas para cada rede. No entanto, devido à ES, cada arquitetura apresentou um número final de épocas diferente. Para a rede Densenet121, ocorreram 211 épocas em 45 minutos de treinamento. Para o EfficientNetV2S, ocorreram 135 épocas em 42 minutos; para o Xception, ocorreram 138 épocas em 40 minutos. Após o ajuste dos hiperparâmetros, os resultados melhoram significativamente com a escolha do otimizador e a taxa de aprendizado adequada. A Figura 45, mostra as curvas ROC de treinamento e validação para a arquitetura EfficientNetV2. É possível observar um ajuste de curva ideal no gráfico, minimizando a ocorrência de *overfitting*.

Figura 45 – EfficientNetV2 - Depois do ajuste do Otimizador e *Learning Rate*



Fonte: Elaborada pela autora.

Os resultados de classificação, utilizando os modelos obtidos para cada arquitetura, são apresentados na tabela 19. Como pode ser visto, o modelo criado pela arquitetura EfficientNetV2S apresenta os melhores resultados para *Precision*, *Recall* e *F1-Score*.

A acurácia e AUC do conjunto de teste, obtidas por cada modelo e pelo *ensemble* podem ser vistos na tabela 20.

Quando comparado com outras abordagens, o conjunto proposto apresenta resultados consistentes. Como pode ser visto na tabela 21.

Classificação Benigno/Maligno

Para a classificação dos ROIs com a classe benigno ou maligno, optou-se por utilizar somente a arquitetura EfficientNetV2S, tendo em vista que os resultados descritores anteriormente apresentaram melhores resultados para essa arquitetura. A taxa de aprendizado utilizada foi $LR1e - 3$ e o otimizador Adamax. Foi estabelecido 250 épocas, e

Tabela 19 – Resultados de classificação

Modelo	Classe	Precision	Recall	F1-Score
Densenet121	Mass	0.9231	0.8715	0.8966
	Calcification	0.8623	0.9172	0.8889
EfficientNetV2S	Mass	0.9576	0.8827	0.9186
	Calcification	0.8772	0.9554	0.9146
Xception	Mass	0.9023	0.8771	0.8895
	Calcification	0.8642	0.8917	0.8777

Tabela 20 – Acurácia e AUC

Model	Test Accuracy	AUC
Densenet121	0.8929	0.966
EfficientNet2VS	0.9167	0.964
Xception	0.8839	0.956
<i>Ensemble - Voting</i>	0.9196	0.976
<i>Ensemble - Averaging</i>	0.9256	0.976

Tabela 21 – Acurácia - comparação com trabalhos correlatos

Model	Accuracy
<i>Model (KHAN et al., 2019)</i>	0.9373
<i>Model (WANG et al., 2016)</i>	0.8970
<i>Model (AGARWAL; CARSON, 2015)</i>	0.8700
<i>Ensemble - Voting</i>	0.9196
<i>Ensemble - Averaging</i>	0.9256

devido ao *early stopping* somente 108 épocas foram executadas. A GPU utilizada foi a Tesla V100 e o tempo de treinamento foi de 22 minutos. Na tabela 22 é possível observar os valores de precisão, revocação e *F1-Score* para cada uma das classes.

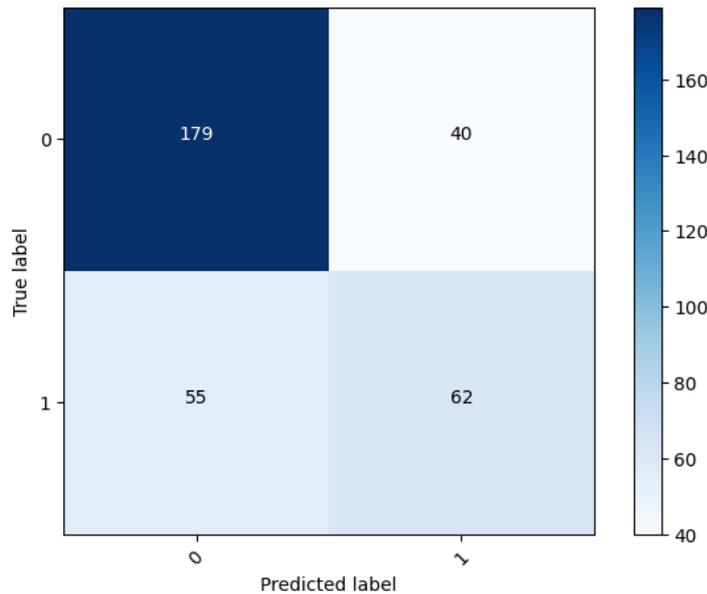
Tabela 22 – Resultados de classificação para benigno e maligno

Modelo	Classe	Precision	Recall	F1-Score
EfficientNetV2S	Benign	0.7650	0.8174	0.7903
	Malign	0.6078	0.5299	0.5662
Test Accuracy	Test Loss	Validation Accuracy	Validation Loss	AUC
0.7173	0.5503	0.7402	0.4978	0.765

Na figura 46 é possível observar a matriz de confusão para as classes benigno e maligno,

onde a classe 0 corresponde a benigno e 1 corresponde a classe maligno. O modelo classificou corretamente 179 ROIs como benigno e 62 como maligno. E atribui a classe incorreta para 40 ROIs que deveriam ter a classe benigno e 55 ROIs que deveriam ter a classe maligno.

Figura 46 – Matriz de Confusão - classificação de benigno e maligno



Fonte: Elaborada pela autora.

Composição dos modelos Massa/Calcificação e Benigno/Maligno

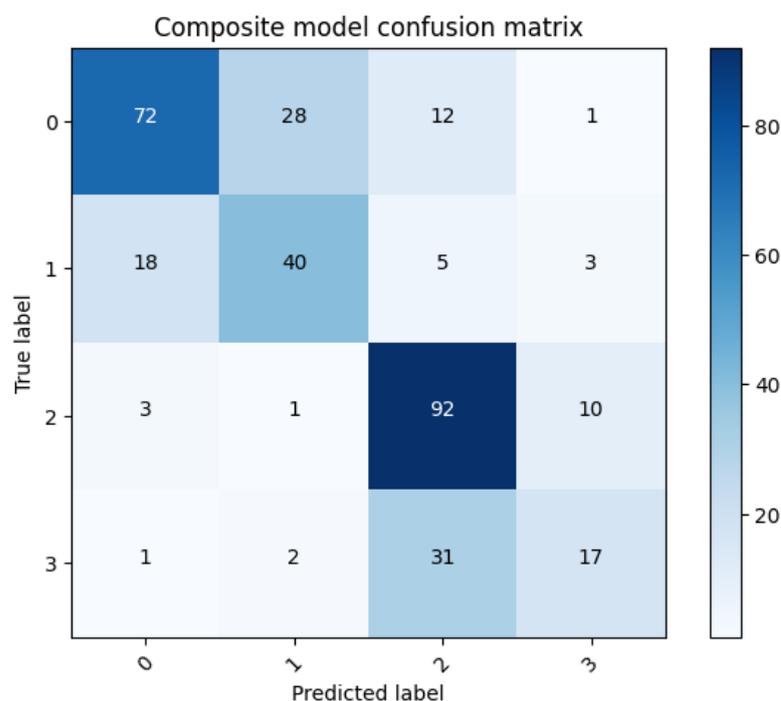
Além dos resultados já apresentados, foi analisado um modelo composto pela combinação dos modelos da arquitetura EfficientNetV2S para classificação de massa/calcificação e benigno/maligno. Essa composição combina as classes como: massa benigna, massa maligna, calcificação benigna e calcificação maligna. Na tabela 23 são apresentados os resultados de precisão, revocação, *F1-Score* e acurácia. A classe massa benigna obteve a melhor precisão entre as quatro classes, no valor de 0.7660. A classe calcificação benigna obteve melhor revocação e *F1-Score* com valores de 0.8679 e 0.7480, respectivamente. O valor da revocação da classe calcificação maligna é de apenas 33%. A acurácia da composição sobre o conjunto de teste foi de 0.6577.

Na figura 47 é apresentada a matriz de confusão do modelo de composição. A diagonal principal da matriz indica os valores corretos da classificação. As classes são: massa benigna (0), massa maligna (1), calcificação benigna (2), calcificação maligna (3). Ao todo, 72 ROIs pertencentes à classe massa benigna são classificados corretamente, enquanto 28 são classificados como massa maligna, 12 como calcificação benigna e 1 como maligna. Para a classe calcificação benigna, 92 ROIs foram classificados corretamente. A classe calcificação maligna é a que apresenta o menor número de acertos na classificação, sendo 17 classificações corretas e 31 classificações incorretas como calcificação benignas.

Tabela 23 – Resultados de classificação

Classe	Precision	Recall	F1-Score
<i>Mass benign</i>	0.7660	0.6372	0.6957
<i>Mass malign</i>	0.5634	0.6061	0.5839
<i>Calcification benign</i>	0.6571	0.8679	0.7480
<i>Calcification malign</i>	0.5484	0.3333	0.4146
Accuracy	0.6577		

Figura 47 – Matriz de Confusão - classificação de massa benigna (0), massa maligna (1), calcificação benigna (2), calcificação maligna (3)



Fonte: Elaborada pela autora.

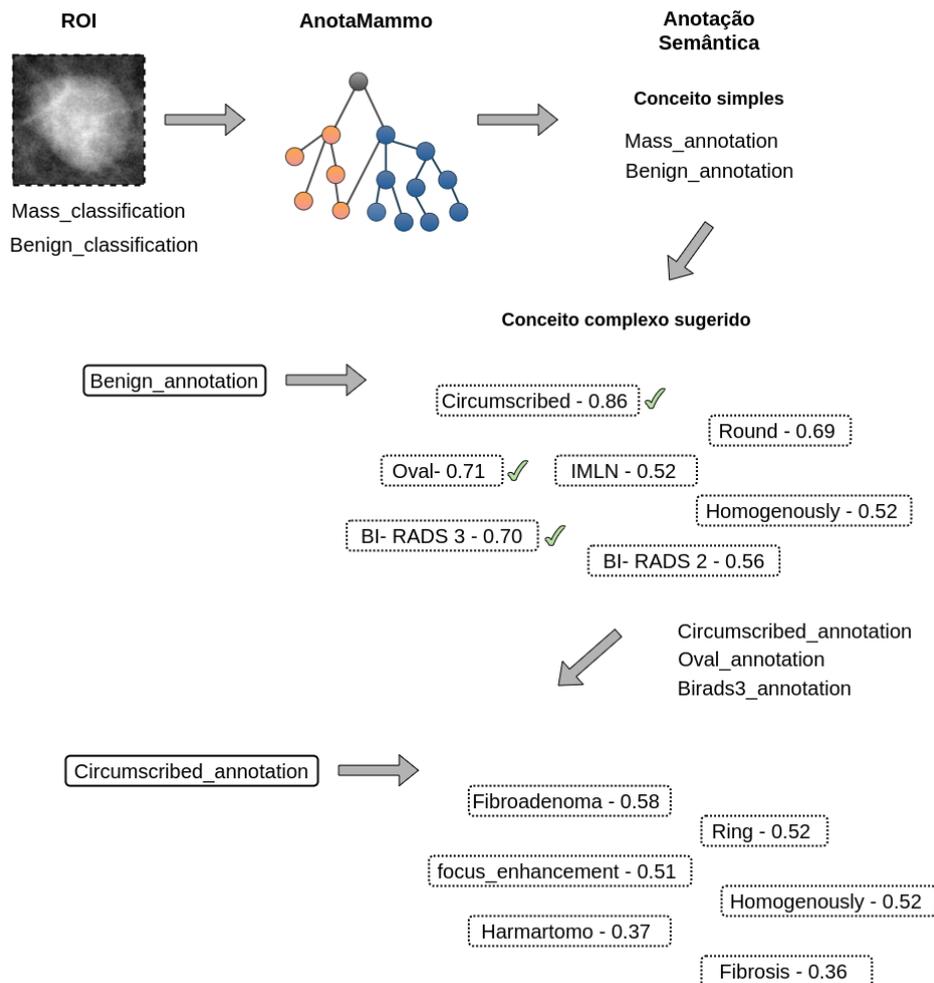
Os resultados gerados pelo processo de mineração de imagem são utilizados como dados iniciais para o processo de anotação semântica da imagem na AnotaMammo, apresentado na próxima seção.

5.5 AnotaMammo

A AnotaMammo é a ontologia que possibilita a anotação semântica das imagens médicas. Após a imagem passar pelo processo de mineração de imagem, os dados resultantes são inseridos no modelo estruturado de conhecimento para prosseguir com a anotação semântica dessas imagens. Uma representação abstrata desse processo é apresentado na figura 48 e na figura 49. As anotações semânticas ocorrem por meio do raciocínio das

regras presentes na AnotaMammo. As regras são construídas com os dados gerados no processo de mineração de texto.

Figura 48 – Processo de Anotação Semântica - ROI classificado no processo de mineração de imagem como massa e benigno, esse dado é entrada na AnotaMammo. O processo de anotação semântica anota o ROI com os conceitos simples obtidos na classificação inicial. Em seguida, conceitos complexos são sugeridos para possível anotação.

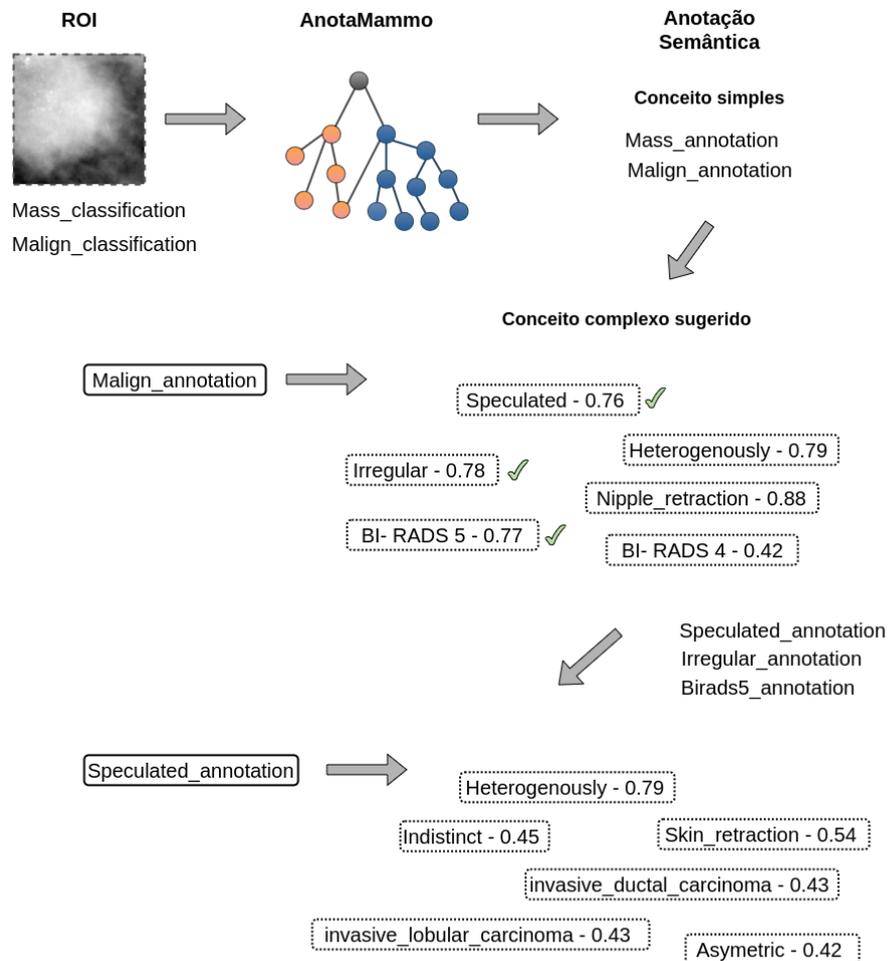


Fonte: Elaborada pela autora.

Na figura 48 um exemplo de uma imagem classificada como massa e benigno. Inicialmente, são inseridos na AnotaMammo a imagem e as classificações obtidas no processo anterior. Em seguida, as primeiras anotações que a imagem recebe são os conceitos simples conforme a classificação. Nesse caso, a imagem recebe a anotação dos conceitos simples *mass_annotation* e *benign_annotation*. Logo após, os conceitos complexos similares à anotação são sugeridos conforme as anotações iniciais. No exemplo ilustrado, são sugeridos os termos complexos para *benign_annotation*, como *circumscribed* com 0.86 de similaridade, *oval* com 0.71 de similaridade e *Birads 3* com 0.70 de similaridade. Ao sele-

cionar uma anotação a ser incluída na imagem, outras sugestões de conceitos complexos são propostas. Por exemplo, ao anotar a imagem com o conceito complexo *circumscribed*, são sugeridos os conceitos complexos *fibroadenoma* com 0.58 de similaridade, *ring* com similaridade de 0.52 e *fibrosis* com 0.36. As sugestões de conceitos complexos que foram selecionadas são de fato anotadas na imagem. Essa seleção pode ser realizada pelo especialista médico. A título de exemplo, os conceitos complexos selecionados foram aqueles que constavam nos relatórios médicos utilizados nesta pesquisa de doutorado.

Figura 49 – Processo de Anotação Semântica - ROI classificado no processo de mineração de imagem como massa e maligno, esse dado é entrada na AnotaMammo. O processo de anotação semântica anota o ROI com os conceitos simples obtidos na classificação inicial. Em seguida, conceitos complexos são sugeridos para possível anotação.



Fonte: Elaborada pela autora.

Na figura 49 é representado o exemplo de uma imagem classificada como massa e maligno. A partir da anotação *malign_annotation*, são sugeridos os conceitos complexos *speculated* com 0.76 de similaridade, *irregular* com similaridade 0.78, Birads 5 com simi-

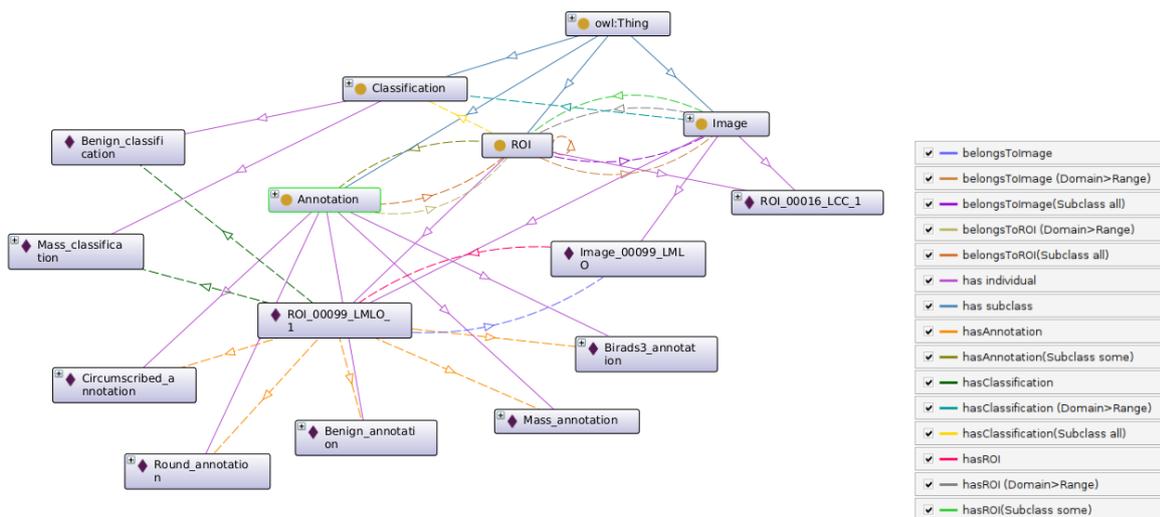
laridade de 0.77 e *nipple_retraction*, entre outros. Com a anotação do conceito complexo *speculated*, outros conceitos complexos são sugeridos. Tais como, *heterogenously* com similaridade de 0.79, *indistinct* 0.45 de similaridade, *skin_retraction* com 0.54 de similaridade, entre outros. Apesar do birads 4 e 5 estarem representados como sugestão de anotação, raramente essas categorias são utilizadas sem outras informações de exames complementares, principalmente porque são categorias que consideram a análise de tecido por meio da biópsia.

Exemplos de anotação de indivíduos na AnotaMammo

Na figura 50 é apresentado o resultado da anotação do indivíduo ROI_00099_LMLO_1 com conceitos simples e conceitos complexos. Também pode-se observar as relações entre classes e indivíduos. Os indivíduos são identificados por um losango roxo. As linhas contínuas indicam que o indivíduo pertence a uma classe, ou entre duas classes que indica a relação de subclasse. As linhas pontilhadas indicam as propriedades. As cores diferenciam as propriedades e podem ser consultadas na legenda (na versão em pdf).

Na parte inferior da figura 50, o ROI_00099_LMLO_1 *belongsToImage* Image_00099. Possui a classificação *mass_classification* e *benign_classification*. Possui as anotações de conceito simples *mass_annotation* e *benign_annotation*. E, anotações de conceitos complexos *circumscribed_annotation*, *round_annotation* e *Birads3_annotation*.

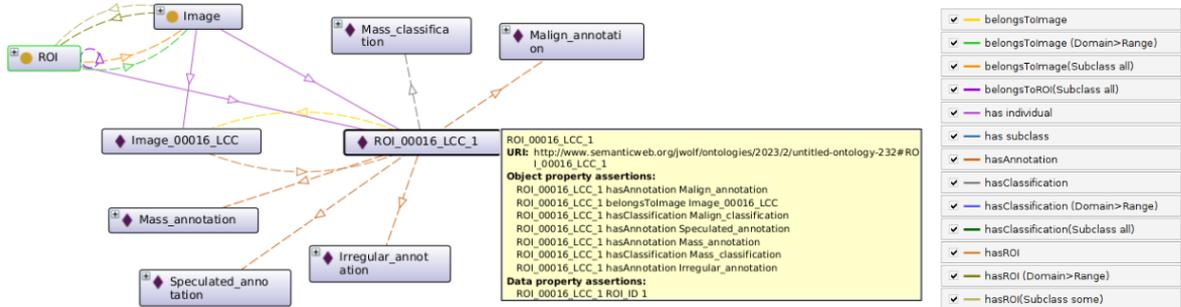
Figura 50 – Anotação Semântica pela AnotaMammo do ROI_00099_LMLO_1. Anotações dos conceitos simples massa e benigno, e anotações dos conceitos complexos *circumscribed*, *round* e *Birads 3*



Fonte: Elaborada pela autora.

Na figura 51, o indivíduo ROI_00016_LCC_1 possui as anotações de conceito simples *mass_annotation* e *malign_annotation*. E, anotações de conceitos complexos *irregular_annotation*, *speculated_annotation*. No quadro em destaque constam os detalhes da anotação.

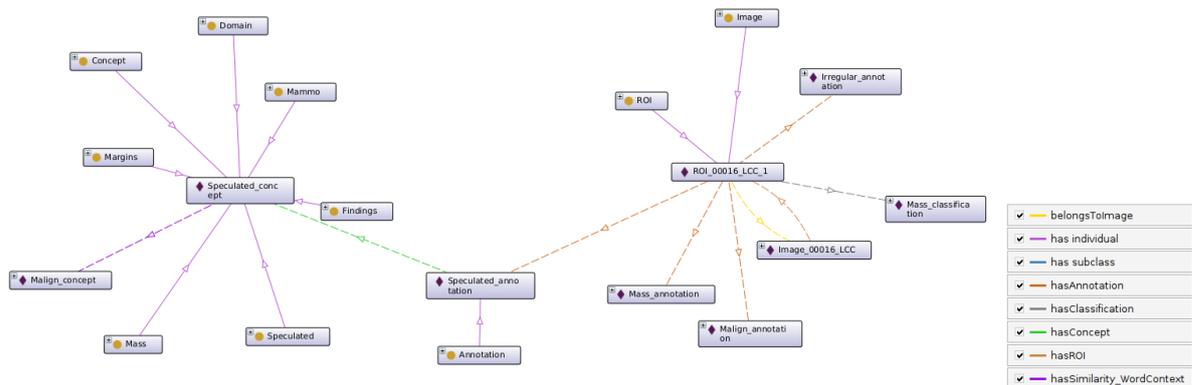
Figura 51 – Anotação Semântica pela AnotaMammo do ROI_00016_LMLO_1. No quadro detalhes da anotação



Fonte: Elaborada pela autora.

Na figura 52, o indivíduo ROI_00016_LCC_1 possui a anotação *speculated_annotation* que se relaciona com o indivíduo *speculated_concept*. Esse relacionamento permite vincular a anotação com a ontologia de domínio Mammo. Ao lado esquerdo da figura as relações do conceito com as classes na ontologia de domínio, além das relações com outros conceitos. No exemplo, o conceito *speculated* se relaciona com o conceito *malign* por meio da propriedade *hasSimilarity_WordContext*, que indica qual a similaridade entre os conceitos.

Figura 52 – Relações entre um indivíduo anotação e um indivíduo conceito. E os relacionamentos com *speculated_concept*



Fonte: Elaborada pela autora.

A AnotaMammo foi avaliada quanto à consistência pelo raciocinador Pellet. A construção da ontologia foi orientada pelos princípios de melhores práticas descritos em Arp, Smith e Spear (2015), Noy, McGuinness et al. (2001).

5.6 Considerações Finais

Neste capítulo foram apresentados os resultados dos experimentos realizados para o método MUSA. Primeiramente foram apresentadas as bases de dados de mamografia utilizadas nos experimentos. Em seguida, foram apresentados os resultados do processo de mineração de texto e regras de associação. Na sequência, foram apresentados os resultados do processo de mineração de imagem. Por último, foram apresentados os resultados produzidos pela AnotaMammo.

No processo de mineração de texto foi apresentada a etapa de pré-processamento dos laudos médicos, incluindo a etapa de formação de unigramas e *skipgrams*. Foram demonstrados alguns dos resultados obtidos de frequência de unigramas e *skipgrams* para diferentes bases de dados. Ademais, foram demonstrados os resultados obtidos pelo cálculo de similaridade entre uma *word* e o *context* para diferentes bases, foi utilizada a medida de similaridade do cosseno. Os recursos obtidos são utilizados para a construção das regras semânticas. Além disso, para as regras de associação foi utilizado o algoritmo *Apriori* (AGRAWAL; SRIKANT et al., 1994). Com as regras geradas foi possível identificar novas relações, que também foram utilizadas para a formação das regras semânticas.

No processo de mineração de imagens foram apresentadas as etapas da estratégia *end-to-end deep learning*. Primeiramente, as imagens de mamografia são transformadas pela etapa de pré-processamento. Em seguida, foram demonstrados os resultados obtidos nos experimentos para definição de otimizador e taxa de aprendizado para cada arquitetura CNN. Na sequência, são apresentados os resultados de classificação.

Para a classificação de *mass/calcification* a arquitetura EfficientNetV2S apresentou os melhores resultados, obtendo mais de 91% de acurácia sobre o conjunto de teste. O *ensemble-averaging* obteve sobre o conjunto de teste, acurácia superior a 92%, resultados similares e até melhores quando comparados com outros trabalhos da literatura.

Para classificação de *benign/malign* os resultados apresentaram valores inferiores quando comparados com à classificação de *mass/calcification*, para o conjunto de teste a taxa de acurácia obtida foi de 71%. Ademais, foram apresentados os resultados obtidos com a combinação dos dois modelos para classificação de quatro classes, a acurácia foi de 65%. As classes inferidas pelo processo de classificação são utilizadas como termos iniciais para o processo de anotação semântica das imagens.

Por fim, foram apresentados os resultados obtidos com o processo de anotação semântica da imagem. A partir dos termos iniciais da classificação são realizadas as primeiras anotações de conceitos simples nas imagens. Em seguida, conceitos complexos são sugeridos e conforme são selecionados como relevantes pelo especialista de domínio, novos conceitos complexos são sugeridos. Os resultados demonstraram que a AnotaMammo permitiu enriquecer semanticamente os termos iniciais provenientes da classificação da imagem, além de propiciar a fusão das informações multimodais.

Capítulo 6

Conclusão

Neste capítulo são apresentadas as conclusões, uma discussão sobre os principais resultados obtidos e contribuições. Além disso, são apresentados os trabalhos futuros que podem ser gerados a partir desta pesquisa de doutorado.

6.1 Principais Contribuições

Nos últimos anos a geração de dados biomédicos tem um crescimento exponencial devido principalmente ao aumento de poder computacional e dos métodos desenvolvidos para lidar com esse tipo de dado. Para o diagnóstico clínico do câncer de mama, e de outras doenças, cada vez mais se faz o uso de exames por imagem. Com isso, a quantidade de dados de imagens gerada é cada vez maior. Dessa forma, deve ser considerado o desenvolvimento de soluções inovadoras que busquem facilitar a rotina clínica, como uma segunda opinião computacional para o diagnóstico clínico, ou para sistemas educacionais para treinamento de novos médicos especialistas.

O processamento e análise de imagens pode gerar resultados pouco interpretáveis para os usuários. Soluções que integram diferentes estratégias para agregar mais semântica a esse resultado, contribuem para diminuição do *gap* semântico entre o entendimento humano em relação àquela imagem e os resultados obtidos pelo processamento do conteúdo ao nível de *pixels*.

Diferentes trabalhos foram propostos para lidar computacionalmente com imagens médicas ou imagens de mamografia. Tanto trabalhos que visam melhorar a tarefa de classificação, quanto os trabalhos que promovem anotação de imagens. No entanto, poucos trabalhos exploram o uso em conjunto de diferentes modalidades de informação para

anotação e melhoria da interpretabilidade dos resultados de classificação, reduzindo o *gap* semântico.

Nesse sentido, o método MUSA foi proposto com o objetivo de classificar e anotar semanticamente imagens médicas, utilizando fusão de informações multimodais. O método integrou diferentes processos para atingir o objetivo. Utilizou várias bases de dados, e diferentes ferramentas e técnicas de mineração de dados complexos e inteligência artificial. Tais como, a similaridade distribucional e regras de associação para lidar com as informações presentes nos laudos médicos. Utilizou também o aprendizado profundo para lidar com a classificação das imagens de mamografia, além do uso de ontologias para representar o domínio de conhecimento e para sustentar o processo de anotação semântica.

O processo de mineração de texto foi responsável por estabelecer as características contextuais, a partir de relações de similaridade e coocorrência entre pares de palavras presentes em um mesmo laudo. Por meio das matrizes de frequência e das variantes de PMI, com o uso da medida de similaridade do cosseno, foi possível determinar valores de similaridade entre palavras. Esses valores foram usados como base para a construção de regras semânticas. Além do mais, com as regras de associação, foi possível identificar novas regras e relações semânticas entre os *termos* frequentes, que também foram utilizadas para a formação das regras semânticas.

O processo de mineração de imagem foi responsável por classificar a imagem de mamografia. A abordagem utilizada é uma estratégia *end-to-end deep learning*, e inclui o aprendizado por transferência, por meio do ajuste fino e definição de hiperparâmetros. As imagens foram classificadas em *mass/calcification* e *benign/malign*. Os resultados foram relevantes e consistentes, e condizem com valores presentes na literatura. Na classificação binária entre massa e calcificação foram obtidos resultados similares e até melhores que os trabalhos encontrados na literatura.

Na classificação de benigno e maligno os resultados tiveram um pior desempenho quando comparado com à classificação de massa/calcificação, mas, mesmo assim, compatível com a literatura. Possivelmente, isso ocorre pela própria natureza das lesões que podem apresentar características semelhantes para patologia diversa, além da pouca quantidade de dados usada no treinamento e o desbalanceamento das classes.

As classes inferidas pelo processo de classificação são utilizadas como termos iniciais para o processo de anotação semântica das imagens.

No processo de engenharia de ontologias foram idealizadas e construídas a ontologia AnotaMammo que engloba a ontologia de anotação e a ontologia de domínio. A ontologia de anotação armazena informações sobre a imagem e comporta o processo de anotação semântica, por meio dos artefatos semânticos incluídos na ontologia.

A ontologia de domínio Mammo representa o conhecimento relacionado com a análise de imagens de mamografia para diagnóstico de câncer de mama.

A separação das ontologias entre anotação e domínio permite que a abordagem seja

adaptada para outros tipos de imagens médicas, em outro domínio. A AnotaMammo permite a anotação semântica das imagens médicas.

Os resultados demonstraram que a ontologia desempenhou apropriadamente o processo de enriquecimento semântico dos termos iniciais provenientes da classificação da imagem, além de ter realizado adequadamente a fusão das informações multimodais.

Por meio das anotações semânticas, o método MUSA possibilita incluir novas informações ao resultado da classificação de uma imagem a partir de suas características visuais. E, como principal vantagem, agrega informações para tornar o resultado mais semântico, mais interpretável, diminuindo o *gap* semântico. Além disso, possibilita uma maior aplicabilidade do que somente a classificação. As anotações semânticas, por exemplo, podem ser incluídas no relatório médico produzido para uma imagem analisada. E, em um sistema educacional, também podem ser utilizadas para apresentar ao aprendiz conceitos relacionados à classificação inicial da imagem, além de informações sobre o significado de cada conceito.

6.2 Publicações

- PEREIRA, Juliana Wolf; RIBEIRO, Marcela Xavier. Semantic annotation and classification of mammography images using ontologies. In: 2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS). IEEE, 2021. p. 378-383. (PEREIRA; RIBEIRO, 2021).
- PEREIRA, Juliana Wolf; RIBEIRO, Marcela Xavier. Hyperparameter for Deep Learning Applied in Mammogram Image Classification. In: 2022 IEEE 35th International Symposium on Computer-Based Medical Systems (CBMS). IEEE, 2022. p. 90-95. (PEREIRA; RIBEIRO, 2022).

6.3 Trabalhos Futuros

Como trabalhos futuros, foram identificadas algumas oportunidades descritas a seguir:

- Projetar e desenvolver e avaliar uma aplicação web para agregar os diferentes processos e permitir a utilização online para usuários especializados. Assim, a ferramenta após testada e validada por órgãos competentes pode contribuir com a prática clínica na identificação de lesões em imagens de mamografia.
- Adquirir, organizar e analisar laudos completos em linguagem natural para elaboração de novas regras e relações semânticas. Essa tarefa não é trivial por ser rara a disponibilidade pública desse tipo de dado.

- Estudar métodos e técnicas para realizar a análise da imagem completa, e não apenas de uma ROI independente, agregando todas as informações referentes ao caso em estudo. Algumas estratégias nesse sentido podem ser adotadas. Por exemplo, classificar individualmente cada ROI e combinar os resultados. Extrair características de cada ROI, e com a técnica de fusão combinar os vetores de características das diferentes visualizações da mamografia em um único vetor para classificação, como realizado em Khan et al. (2019). Além de estudar e adaptar a estratégia apresentada em Cazzolato et al. (2019), que permite detectar anormalidades em regiões, baseada em *superpixel* e visualização de mapa de calor.

Outras oportunidades interessantes dizem respeito ao estudo de técnicas e ferramentas para o aprendizado contínuo de regras semânticas. E, ao estudo de métodos para a incorporação da variável temporal na análise do caso de uma paciente ao longo do tempo. Aqui ocorre novamente a questão de disponibilidade pública desse tipo de dado.

6.4 Considerações Finais

A anotação semântica de imagens médicas a partir de informações multimodais é uma tarefa desafiadora, por envolver o conhecimento multidisciplinar. Diferentes abordagens precisam ser desenvolvidas para obtenção de diferentes modalidades de informação. Por exemplo, processamento de imagens, processamento de texto, classificação, ontologia, entre outros. Para viabilizar a realização de anotação semântica de imagens de mamografia foi proposto o método MUSA.

A abordagem MUSA engloba um processo de mineração de texto e regras de associação para a construção de regras semânticas que compõem o *core* da ontologia AnotaMammo. Inclui ainda um processo de mineração de imagem, a fim de prover informações iniciais quanto à classificação da imagem analisada. Além da ontologia AnotaMammo, composta por uma ontologia de anotação e a ontologia de domínio Mammo. AnotaMammo é responsável por integrar as informações multimodais e realizar a anotação semântica em imagens de mamografia. Os experimentos demonstraram que o método MUSA teve bom desempenho, e contribuiu para adicionar informações semânticas aos resultados de classificação, tornando-os mais completos, interpretáveis e diminuindo o *gap* semântico.

Referências

- AGARWAL, V.; CARSON, C. Using deep convolutional neural networks to predict semantic features of lesions in mammograms. **C231n Course Project Reports**, 2015.
- AGGARWAL, C. C. **Data mining: the textbook**. [S.l.]: Springer, 2015.
- AGRAWAL, R.; IMIELIŃSKI, T.; SWAMI, A. Mining association rules between sets of items in large databases. In: **Proceedings of the 1993 ACM SIGMOD international conference on Management of data**. [S.l.: s.n.], 1993. p. 207–216.
- AGRAWAL, R.; SRIKANT, R. et al. Fast algorithms for mining association rules. In: SANTIAGO, CHILE. **Proc. 20th int. conf. very large data bases, VLDB**. [S.l.], 1994. v. 1215, p. 487–499.
- AKGÜL, C. B. et al. Content-based image retrieval in radiology: current status and future directions. **Journal of digital imaging**, Springer, v. 24, n. 2, p. 208–222, 2011.
- ALLAMPALLI-NAGARAJ, G.; BICHINDARITZ, I. Automatic semantic indexing of medical images using a web ontology language for case-based image retrieval. **Engineering Applications of Artificial Intelligence**, Elsevier, v. 22, n. 1, p. 18–25, 2009.
- ALOM, M. Z. et al. A state-of-the-art survey on deep learning theory and architectures. **electronics**, Multidisciplinary Digital Publishing Institute, v. 8, n. 3, p. 292, 2019.
- ALPAYDIN, E. **Introduction to machine learning**. [S.l.]: MIT press, 2020.
- ARP, R.; SMITH, B.; SPEAR, A. D. **Building ontologies with basic formal ontology**. [S.l.]: Mit Press, 2015.
- AZEVEDO-MARQUES, P. M. d. Diagnóstico auxiliado por computador na radiologia. **Radiologia Brasileira**, SciELO Brasil, v. 34, n. 5, p. 285–293, 2001.
- AZEVEDO-MARQUES, P. M. D. et al. **Medical Image Analysis and Informatics: Computer-Aided Diagnosis and Therapy**. [S.l.]: CRC Press, 2017.
- BANKMAN, I. **Handbook of medical image processing and analysis**. [S.l.]: Elsevier, 2008.
- BANNOUR, H. **Building and Using Knowledge Models for Semantic Image Annotation**. Tese (Doutorado) — Ecole Centrale Paris, 2013.

- BANNOUR, H.; HUDELOT, C. Building and using fuzzy multimedia ontologies for semantic image annotation. **Multimedia tools and applications**, Springer, v. 72, n. 3, p. 2107–2141, 2014.
- BHAGAT, P.; CHOUDHARY, P. Image annotation: Then and now. **Image and Vision Computing**, Elsevier, v. 80, p. 1–23, 2018.
- BIRD, S.; KLEIN, E.; LOPER, E. **Natural language processing with Python: analyzing text with the natural language toolkit**. [S.l.]: "O'Reilly Media, Inc.", 2009.
- BITENCOURT, A. G. et al. Intramammary lymph nodes: normal and abnormal multimodality imaging features. **The British Journal of Radiology**, The British Institute of Radiology., v. 92, n. 1103, p. 20190517, 2019.
- BORST, W. N. **Construction of Engineering Ontologies for Knowledge Sharing and Reuse**. Tese (Doutorado) — Universiteit Twente, 1997.
- BOTTOU, L. Large-scale machine learning with stochastic gradient descent. In: SPRINGER. **Proceedings of COMPSTAT'2010: 19th International Conference on Computational Statistics Paris France, August 22-27, 2010 Keynote, Invited and Contributed Papers**. [S.l.], 2010. p. 177–186.
- BOUCHAKWA, M.; AYADI, Y.; AMOUS, I. A review on visual content-based and users' tags-based image annotation: methods and techniques. **MULTIMEDIA TOOLS AND APPLICATIONS**, Springer, 2020.
- BOUMA, G. Normalized (pointwise) mutual information in collocation extraction. **Proceedings of GSCL**, p. 31–40, 2009.
- BRIN, S. et al. Dynamic itemset counting and implication rules for market basket data. In: **Proceedings of the 1997 ACM SIGMOD international conference on Management of data**. [S.l.: s.n.], 1997. p. 255–264.
- BUENO, R. et al. Time-aware similarity search: a metric-temporal representation for complex data. In: SPRINGER. **International Symposium on Spatial and Temporal Databases**. [S.l.], 2009. p. 302–319.
- BUI, A. A.; TAIRA, R. K. **Medical imaging informatics**. [S.l.]: Springer Science & Business Media, 2009.
- BULU, H. **Ontology-based medical image annotation and retrieval**. Tese (Doutorado) — DEÜ Fen Bilimleri Enstitüsü, 2013.
- BULU, H.; ALPKOCAK, A.; BALCI, P. Ontology-based mammography annotation and case-based retrieval of breast masses. **Expert Systems with Applications**, Elsevier, v. 39, n. 12, p. 11194–11202, 2012.
- _____. Uncertainty modeling for ontology-based mammography annotation with intelligent bi-rads scoring. **Computers in Biology and Medicine**, Elsevier, v. 43, n. 4, p. 301–311, 2013.

- BURNSIDE, E. S. et al. Probabilistic computer model developed from clinical data in national mammography database format to classify mammographic findings. **Radiology**, Radiological Society of North America, v. 251, n. 3, p. 663–672, 2009.
- BUSHBERG, J. T.; BOONE, J. M. **The essential physics of medical imaging**. [S.l.]: Lippincott Williams & Wilkins, 2011.
- CAMILO, C. O. et al. Uma metodologia para mineração de regras de associação usando ontologias para integração de dados estruturados e não-estruturados. Universidade Federal de Goiás, 2010.
- CAZZOLATO, M. T. et al. dp-breath: Heat maps and probabilistic classification assisting the analysis of abnormal lung regions. **Computer Methods and Programs in Biomedicine**, Elsevier, v. 173, p. 27–34, 2019.
- CHHATWAL, J. et al. A logistic regression model based on the national mammography database format to aid breast cancer diagnosis. **American Journal of Roentgenology**, Am Roentgen Ray Soc, v. 192, n. 4, p. 1117–1127, 2009.
- CHOLLET, F. Xception: Deep learning with depthwise separable convolutions. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2017. p. 1251–1258.
- CHU, B. et al. Best practices for fine-tuning visual classifiers to new domains. In: SPRINGER. **Computer Vision–ECCV 2016 Workshops: Amsterdam, The Netherlands, October 8-10 and 15-16, 2016, Proceedings, Part III 14**. [S.l.], 2016. p. 435–442.
- CHURCH, K.; HANKS, P. Word association norms, mutual information, and lexicography. **Computational linguistics**, v. 16, n. 1, p. 22–29, 1990.
- DARBY, M. J.; BARRON, D.; HYLAND, R. **Oxford handbook of medical imaging**. [S.l.]: OUP Oxford, 2011.
- DEERWESTER, S. et al. Indexing by latent semantic analysis. **Journal of the American society for information science**, Wiley Online Library, v. 41, n. 6, p. 391–407, 1990.
- DENG, J. et al. Imagenet: A large-scale hierarchical image database. In: IEEE. **2009 IEEE conference on computer vision and pattern recognition**. [S.l.], 2009. p. 248–255.
- D’ORSI, C. J. et al. **ACR BI-RADS Atlas: Breast Imaging Reporting and Data System; Mammography, Ultrasound, Magnetic Resonance Imaging, Follow-up and Outcome Monitoring, Data Dictionary**. [S.l.]: ACR, American College of Radiology, 2013.
- DOZAT, T. Incorporating nesterov momentum into adam. 2016.
- DUCHI, J.; HAZAN, E.; SINGER, Y. Adaptive subgradient methods for online learning and stochastic optimization. **Journal of machine learning research**, v. 12, n. 7, 2011.

- ESTRELA, V. V.; HERRMANN, A. E. Content-based image retrieval (cbir) in remote clinical diagnosis and healthcare. In: **Encyclopedia of E-Health and Telemedicine**. [S.l.]: IGI Global, 2016. p. 495–520.
- FAYYAD, U. M. et al. Advances in knowledge discovery and data mining. In: AMERICAN ASSOCIATION FOR ARTIFICIAL INTELLIGENCE. [S.l.], 1996.
- FILALI, J.; ZGHAL, H.; MARTINET, J. Ontology and hmax features-based image classification using merged classifiers. In: INSTICC. **International Conference on Computer Vision Theory and Applications**. [S.l.], 2019.
- GALVEIA, J. N.; CRUZ, L. A. da S.; TRAVASSOS, A. Image and ontological information fusion for cataract surgery recommendation. In: IEEE. **2019 27th European Signal Processing Conference (EUSIPCO)**. [S.l.], 2019. p. 1–5.
- GÉRON, A. **Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow**. [S.l.]: "O'Reilly Media, Inc.", 2022.
- GONÇALVES, E. C. Regras de associações e suas medidas de interesse objetivas e subjetivas. **INFOCOMP Journal of Computer Science**, v. 4, n. 1, p. 26–35, 2005.
- GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep learning**. [S.l.]: MIT press, 2016.
- GRUBER, T. Ontology. **Encyclopedia of database systems**, Springer, p. 1963–1965, 2009.
- GRUBER, T. R. A translation approach to portable ontology specifications. **Knowledge acquisition**, Elsevier, v. 5, n. 2, p. 199–220, 1993.
- GUARINO, N. **Formal ontology in information systems: proceedings of the first international conference (FOIS'98), June 6-8, Trento, Italy**. [S.l.]: Ios PressInc, 1998. v. 46.
- GUARINO, N.; OBERLE, D.; STAAB, S. What is an ontology? In: **Handbook on ontologies**. [S.l.]: Springer, 2009. p. 1–17.
- HAN, J.; PEI, J.; KAMBER, M. **Data mining: concepts and techniques**. [S.l.]: Elsevier, 2011.
- HINTON, G.; SRIVASTAVA, N.; SWERSKY, K. Neural networks for machine learning lecture 6a overview of mini-batch gradient descent. **Cited on**, v. 14, n. 8, p. 2, 2012.
- HOCKENMAIER, J. **Distributional similarities, word embeddings**. 2023. Último acesso em 15 de fevereiro de 2023. Disponível em: <<http://courses.grainger.illinois.edu/cs447>>.
- HORROCKS, I. et al. Swrl: A semantic web rule language combining owl and ruleml. **W3C Member submission**, v. 21, n. 79, p. 1–31, 2004.
- HU, B. et al. Ontology-based medical image annotation with description logics. In: IEEE. **Proceedings. 15th IEEE International Conference on Tools with Artificial Intelligence**. [S.l.], 2003. p. 77–82.

- HUANG, G. et al. Densely connected convolutional networks. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2017. p. 4700–4708.
- HWANG, K. H. et al. Building and querying rdf/owl database of semantically annotated nuclear medicine images. **Journal of digital imaging**, Springer, v. 30, n. 1, p. 4–10, 2017.
- IKEDA, D.; MIYAKE, K. K. **Breast imaging: the requisites**. [S.l.]: Elsevier Health Sciences, 2016.
- IM, E.-J.; YELICK, K. Optimization of sparse matrix kernels for data mining. In: **Proceedings of the Workshop on Text Mining**. [S.l.: s.n.], 2001.
- INCA. **Estimativa 2020 : incidência de câncer no Brasil**. [S.l.], 2019.
- _____. **Estimativa 2023 : Incidência de Câncer no Brasil**. [S.l.], 2022.
- ISLAM, K. T.; WIJEWICKREMA, S.; O'LEARY, S. Identifying diabetic retinopathy from oct images using deep transfer learning with artificial neural networks. In: **IEEE 2019 IEEE 32nd international symposium on computer-based medical systems (CBMS)**. [S.l.], 2019. p. 281–286.
- IVANOVIĆ, M.; BUDIMAC, Z. An overview of ontologies and data resources in medical domains. **Expert Systems with Applications**, Elsevier, v. 41, n. 11, p. 5158–5166, 2014.
- JUNIOR, C. R. S. Mineração de regras de associação espaço-temporais temáticas aplicada a imagens de explosões solares. Universidade Federal de São Carlos, 2018.
- KAO, A.; POTEET, S. R. **Natural language processing and text mining**. [S.l.]: Springer Science & Business Media, 2007.
- KHALED, R. et al. Categorized digital database for low energy and subtracted contrast enhanced spectral mammography images [dataset]. The Cancer Imaging Archive, 2021. Último acesso em junho 2023. Disponível em: <<https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=109379611>>.
- _____. Categorized contrast enhanced mammography dataset for diagnostic and artificial intelligence research. **Scientific Data**, Nature Publishing Group UK London, v. 9, n. 1, p. 122, 2022.
- KHAMPARIA, A. et al. Diagnosis of breast cancer based on modern mammography using hybrid transfer learning. **Multidimensional systems and signal processing**, Springer, v. 32, p. 747–765, 2021.
- KHAN, H. N. et al. Multi-view feature fusion based four views model for mammogram classification using convolutional neural network. **IEEE Access**, IEEE, v. 7, p. 165724–165733, 2019.
- KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. **arXiv preprint arXiv:1412.6980**, 2014.

- KIRYAKOV, A. et al. Semantic annotation, indexing, and retrieval. **Web Semantics: Science, Services and Agents on the World Wide Web**, Elsevier, v. 2, n. 1, p. 49–79, 2004.
- KUMAR, A. et al. Adapting content-based image retrieval techniques for the semantic annotation of medical images. **Computerized Medical Imaging and Graphics**, Elsevier, v. 49, p. 37–45, 2016.
- KURTZ, C. et al. A hierarchical knowledge-based approach for retrieving similar medical images described with semantic annotations. **Journal of biomedical informatics**, Elsevier, v. 49, p. 227–244, 2014.
- _____. On combining image-based and ontological semantic dissimilarities for medical image retrieval applications. **Medical image analysis**, Elsevier, v. 18, n. 7, p. 1082–1100, 2014.
- _____. Semantic retrieval of radiological images with relevance feedback. In: SPRINGER. **International Workshop on Multimodal Retrieval in the Medical Domain**. [S.l.], 2015. p. 11–25.
- LAI, L. **leoll2/MedicalCNN: v1.0**. Zenodo, 2021. Disponível em: <<https://doi.org/10.5281/zenodo.4700130>>.
- LANGLOTZ, C. P. **RadLex: a new method for indexing online educational materials**. [S.l.]: Radiological Society of North America, 2006. 1595–1597 p.
- LASSILA, O.; SWICK, R. R. et al. Resource description framework (rdf) model and syntax specification. Citeseer, 1998.
- LATTANZIO, V.; SIMONETTI, G. **Mammography: Guide to Interpreting, Reporting and Auditing Mammographic Images-Re. Co. RM (From Italian Reporting and Codifying the Results of Mammography)**. [S.l.]: Springer Science & Business Media, 2010.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **nature**, Nature Publishing Group UK London, v. 521, n. 7553, p. 436–444, 2015.
- LEE, R. S. et al. Curated breast imaging subset of digital database for screening mammography (cbis-ddsm) [data set]. The Cancer Imaging Archive, 2016. Último acesso em junho 2023. Disponível em: <<https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=22516629>>.
- _____. A curated mammography data set for use in computer-aided detection and diagnosis research. **Scientific data**, Nature Publishing Group, v. 4, n. 1, p. 1–9, 2017.
- LEVY, O.; GOLDBERG, Y.; DAGAN, I. Improving distributional similarity with lessons learned from word embeddings. **Transactions of the association for computational linguistics**, MIT Press, v. 3, p. 211–225, 2015.
- LIMA, L. L.; FERREIRA, J. R.; OLIVEIRA, M. C. Efficient hyperparameter optimization of convolutional neural networks on classification of early pulmonary nodules. In: IEEE. **2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS)**. [S.l.], 2019. p. 144–149.

- LUQUE, E. F. et al. Automatic staging of cancer tumors using aim image annotations and ontologies. **Journal of Digital Imaging**, Springer, p. 1–17, 2019.
- LUQUE, E. F.; RUBIN, D. L.; MOREIRA, D. A. Automatic classification of cancer tumors using image annotations and ontologies. In: IEEE. **2015 IEEE 28th International Symposium on Computer-Based Medical Systems**. [S.l.], 2015. p. 368–369.
- MCMAHAN, H. B. et al. Ad click prediction: a view from the trenches. In: **Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining**. [S.l.: s.n.], 2013. p. 1222–1230.
- MIKOLOV, T. et al. Distributed representations of words and phrases and their compositionality. In: **Advances in neural information processing systems**. [S.l.: s.n.], 2013. v. 26.
- MITCHELL, T. M. et al. **Machine learning**. [S.l.]: McGraw-hill New York, 1997.
- MÖLLER, M.; REGEL, S.; SINTEK, M. Radsem: Semantic annotation and retrieval for medical images. In: SPRINGER. **European Semantic Web Conference**. [S.l.], 2009. p. 21–35.
- MÖLLER, M.; SINTEK, M. A generic framework for semantic medical image retrieval. **KAMC**, Citeseer, v. 253, 2007.
- MOREIRA, D. A. et al. 3d markup of radiological images in epad, a web-based image annotation tool. In: IEEE. **2015 IEEE 28th International Symposium on Computer-Based Medical Systems**. [S.l.], 2015. p. 97–102.
- MURPHY, K. P. **Machine learning: a probabilistic perspective**. [S.l.]: MIT press, 2012.
- NG, K. H.; WONG, J. H. D.; CLARKE, G. D. **Problems and Solutions in Medical Physics: Diagnostic Imaging Physics**. [S.l.]: CRC Press, 2018.
- NOY, N. F.; MCGUINNESS, D. L. **Ontology Development 101: A Guide to Creating Your First Ontology**. [S.l.], 2001.
- NOY, N. F.; MCGUINNESS, D. L. et al. **Ontology development 101: A guide to creating your first ontology**. [S.l.]: Stanford knowledge systems laboratory technical report KSL-01-05 and . . . , 2001.
- O’CONNOR, M. J.; DAS, A. K. Sqwrl: a query language for owl. In: **OWLED**. [S.l.: s.n.], 2009. v. 529, n. 2009.
- OLIVEIRA, P. H. et al. Mammoset: An enhanced dataset of mammograms. In: **Proc. Satellite Events 32nd Brazilian Symp. Databases**. [S.l.: s.n.], 2017. p. 256–266.
- OSUCH, J. et al. A proposal for a national mammography database: content, purpose, and value. **AJR. American journal of roentgenology**, American Public Health Association, v. 164, n. 6, p. 1329–1334, 1995.
- PATTI, J. A. The national radiology data registry: a necessary component of quality health care. **Journal of the American College of Radiology**, Elsevier, v. 8, n. 7, p. 453, 2011.

- PEREIRA, J. W.; RIBEIRO, M. X. Semantic annotation and classification of mammography images using ontologies. In: IEEE. **2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS)**. [S.l.], 2021. p. 378–383.
- _____. Hyperparameter for deep learning applied in mammogram image classification. In: IEEE. **2022 IEEE 35th International Symposium on Computer-Based Medical Systems (CBMS)**. [S.l.], 2022. p. 90–95.
- PONTI, M. A.; COSTA, G. B. P. D. Como funciona o deep learning. **arXiv preprint arXiv:1806.07908**, 2018.
- POWERS, D. M. Evaluation: from precision, recall and f-measure to roc, informedness, markedness and correlation. **arXiv preprint arXiv:2010.16061**, 2020.
- REZENDE, S. O.; MARCACINI, R. M.; MOURA, M. F. O uso da mineração de textos para extração e organização não supervisionada de conhecimento. **Embrapa Informática Agropecuária-Artigo em periódico indexado (ALICE)**, Revista de Sistema de Informação da FSMA, Macaé, n. 7, p. 7-21, 2011., 2011.
- ROSSE, C.; MEJINO, J. L. The foundational model of anatomy ontology. In: **Anatomy Ontologies for Bioinformatics**. [S.l.]: Springer, 2008. p. 59–117.
- RUBIN, D. L. et al. epad: An image annotation and analysis platform for quantitative imaging. **Tomography**, Grapho Publications, v. 5, n. 1, p. 170, 2019.
- _____. Annotation and image markup: accessing and interoperating with the semantic content in medical imaging. **IEEE Intelligent Systems**, IEEE, v. 24, n. 1, p. 57–65, 2009.
- _____. ipad: Semantic annotation and markup of radiological images. In: AMERICAN MEDICAL INFORMATICS ASSOCIATION. **AMIA annual symposium proceedings**. [S.l.], 2008. v. 2008, p. 626.
- SALOMÃO, S. C.; MARQUES, P. M. d. A. Integrando ferramentas de auxílio ao diagnóstico no sistema de arquivamento e comunicação de imagens. **Radiologia Brasileira**, SciELO Brasil, v. 44, n. 6, p. 374–380, 2011.
- SAMET, H. **Foundations of multidimensional and metric data structures**. [S.l.]: Morgan Kaufmann, 2006.
- SCHAEFER, G.; HASSANIEN, A.; JIANG, J. **Computational intelligence in medical imaging: techniques and applications**. [S.l.]: CRC press, 2009.
- SPAK, D. A. et al. Bi-rads® fifth edition: A summary of changes. **Diagnostic and interventional imaging**, Elsevier, v. 98, n. 3, p. 179–190, 2017.
- SRIKANT, R.; AGRAWAL, R. **Mining generalized association rules**. [S.l.]: IBM Research Division Zurich, 1995.
- STANESCU, L. et al. **Creating new medical ontologies for image annotation: a case study**. [S.l.]: Springer Science & Business Media, 2011.

- STUDER, R.; BENJAMINS, V. R.; FENSEL, D. Knowledge engineering: principles and methods. **Data & knowledge engineering**, Elsevier, v. 25, n. 1, p. 161–197, 1998.
- SUETENS, P. **Fundamentals of medical imaging**. [S.l.]: Cambridge university press, 2017.
- SZEGEDY, C. et al. Rethinking the inception architecture for computer vision. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2016. p. 2818–2826.
- TAJBAKHSI, N. et al. Convolutional neural networks for medical image analysis: Full training or fine tuning? **IEEE transactions on medical imaging**, IEEE, v. 35, n. 5, p. 1299–1312, 2016.
- TAN, M.; LE, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In: PMLR. **International conference on machine learning**. [S.l.], 2019. p. 6105–6114.
- _____. Efficientnetv2: Smaller models and faster training. In: PMLR. **International conference on machine learning**. [S.l.], 2021. p. 10096–10106.
- TURNEY, P. D.; PANTEL, P. From frequency to meaning: Vector space models of semantics. **Journal of artificial intelligence research**, v. 37, p. 141–188, 2010.
- UREN, V. et al. Semantic annotation for knowledge management: Requirements and a survey of the state of the art. **Web Semantics: science, services and agents on the World Wide Web**, Elsevier, v. 4, n. 1, p. 14–28, 2006.
- WANG, J. et al. Discrimination of breast cancer with microcalcifications on mammography by deep learning. **Scientific reports**, Nature Publishing Group UK London, v. 6, n. 1, p. 27327, 2016.
- WANI, M. A. et al. **Advances in deep learning**. [S.l.]: Springer, 2020.
- WITTEN, I. H.; EIBE, F.; HALL, M. A. **Data mining : practical machine learning tools and techniques**. [S.l.]: Morgan Kaufmann, 2011.
- WU, N. et al. The nyu breast cancer screening dataset v1. 0. **New York Univ., New York, NY, USA, Tech. Rep**, 2019.
- YAGUINUMA, C. A.; SANTOS, M. T.; BIAJIZ, M. Meta-ontologia difusa para representação de informações imprecisas em ontologias. In: **Workshop on Ontologies and Metamodeling in Software and Data Engineering**. [S.l.: s.n.], 2007. p. 57–67.
- ZEILER, M. D. Adadelta: an adaptive learning rate method. **arXiv preprint arXiv:1212.5701**, 2012.