UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUACAO EM ENGENHARIA ELÉTRICA

AVALIAÇÃO DE IMAGENS DE TOMOGRAFIA COMPUTADORIZADA SUBMETIDAS A CODIFICAÇÃO COM PERDAS USANDO REDES BAYESIANAS

LUANA GONÇALVES

DM:35/2019

UFPA/ITEC/PPGEE
CAMPUS UNIVERSITÁRIO DO GUAMÁ
BELÉM - PARÁ-BRASIL
2019

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUACAO EM ENGENHARIA ELÉTRICA

LUANA GONÇALVES

AVALIAÇÃO DE IMAGENS DE TOMOGRAFIA COMPUTADORIZADA SUBMETIDAS A CODIFICAÇÃO COM PERDAS USANDO REDES BAYESIANAS

DM:35/2019

UFPA/ITEC/PPGEE
CAMPUS UNIVERSITÁRIO DO GUAMÁ
BELÉM - PARÁ-BRASIL
2019

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUACAO EM ENGENHARIA ELÉTRICA

LUANA GONÇALVES

AVALIAÇÃO DE IMAGENS DE TOMOGRAFIA COMPUTADORIZADA SUBMETIDAS A CODIFICAÇÃO COM PERDAS USANDO REDES BAYESIANAS

Dissertação de mestrado submetida à Banca Examinadora do Programa de Pós-Graduação em Engenharia Elétrica para obtenção do Grau de Mestre em Engenharia Elétrica na área de Telecomunicações.

UFPA/ITEC/PPGEE
CAMPUS UNIVERSITÁRIO DO GUAMÁ
BELÉM - PARÁ-BRASIL
2019

Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD
Sistema de Bibliotecas da Universidade Federal do Pará
Gerada automaticamente pelo módulo Ficat, mediante os dados fornecidos pelo(a) autor(a)

G635a Goncalves, Luana
Avaliação de imagens de tomografia computadorizada submetidas a codificação com perdas usando redes Bayesianas / Luana Goncalves. — 2019.
57 f. : il. color.

Orientador(a): Prof. Dr. Aldebaro Barreto da Rocha Klautau Júnior
Dissertação (Mestrado) - Programa de Pós-Graduação em Engenharia Elétrica, Instituto de Tecnologia, Universidade Federal do Pará, Belém, 2019.

1. Avaliação de qualidade visual. 2. Imagens médicas. 3. Compressão de imagens. 4. Redes Bayesianas. I. Título.

CDD 621.367

AVALIAÇÃO DE IMAGENS DE TOMOGRAFIA COMPUTADORIZADA SUBMETIDAS A CODIFICAÇÃO COM PERDAS USANDO REDES BAYESIANAS

Dissertação de mestrado submetida à avaliação da banca examinadora aprovada pelo colegiado do Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal do Pará e julgada adequada para obtenção do Grau de Mestre em Engenharia Elétrica na área de Telecomunicações.

Aprovada em _____/ _____/ _____.

_____

Prof. Dr. Aldebaro Barreto da Rocha Klautau Júnior

ORIENTADOR

_____

Prof. Dr. Ronaldo de Freitas Zampolo

MEMBRO DA BANCA EXAMINADORA

_____

Prof. Dr. Antônio Pereira Júnior

MEMBRO DA BANCA EXAMINADORA

_____

Prof^{a}. Dr^{a}. Maria Emília de Lima Tostes

DIRETORA DA PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

À minha família e amigos.

Agradecimentos

Agradeço, acima de tudo, aos meus pais, Rosilene e Ribamar, pela dedicação, compreensão e paciência, por terem sempre investido em minha educação e sido meus exemplos de integridade e honestidade. Pelos ensinamentos diretos e indiretos, onde pude construir meus princípios e personalidade.

Ao meu orientador, Prof. Aldebaro, por ter permitido a realização desse sonho, pela instrução nos momentos mais críticos e, principalmente, por confiar no meu trabalho.

Ao Prof. Zampolo, pelo conhecimento fornecido, pelas palavras de apoio nos momentos difíceis, conselhos e por acreditar no meu potencial quando nem eu mesma acreditava. Agradeço pela confiança, paciência e preocupação.

À Profª. Valquíria, por traduzir o sentido de maternidade acadêmica, pela dedicação integral aos alunos e por ser o alicerce da Engenharia Biomédica.

Ao Prof. Silveira, por ser um incentivador de pessoas, por não ter me deixado desistir de entrar no programa de pós-graduação, pelas palavras de apoio, por ensinar que a Universidade pode ser dinâmica e o berço de inovações.

À minha irmã Danielle, pelo apoio, paciência e compreensão nessa trajetória.

Aos amigos Wictor e Ewerton, pelo companheirismo, amizade, pelos muitos momentos de diversão, de estudos e por terem me incentivado durante os momentos difíceis do mestrado.

Às amigas Jaqueline e Rafaela, por mostrarem que nem a erosão do tempo é capaz de agir sobre amizades que se fundaram em bases sólidas.

A todos os colaboradores do LINSE, os quais me acolheram de forma muito especial durante minha estadia na UFSC. Em especial, o Hermeson, por ter me apresentado ferramentas essenciais no desenvolvimento dessa dissertação.

Em especial, ao meu companheiro Hugo, pela paciência e compreensão, por representar minha família, quando longe de casa e por me dar suporte para finalizar essa jornada.

"Everything in life is important, important things are simple, simple things are never easy"
Murphy's Law.

Resumo

Dado que imagens médicas subsidiam o diagnóstico clínico, o desenvolvimento de métricas para avaliação automática de qualidade de tais imagens se tornou de grande importância na atualidade para preservar a informação original após a realização de operações de compressão, realce e transmissão digital. Os dados de experimentais de referência, usados no desenvolvimento de métricas, devem revelar os aspectos de interesse para uma dada aplicação. A maioria das métricas de qualidade visual são desenvolvidas a partir de experimentos que refletem a opinião de usuários comuns sobre a qualidade de cenas naturais. Experimentos especificamente concebidos para avaliação da qualidade clínica de imagens médicas em que especialistas tomam parte são raros. As contribuições deste trabalho são: a proposta de uma metodologia para aquisição de dados subjetivos para avaliação de imagens, voltada para a caracterização dos critérios de qualidade utilizados por especialistas na fase de pré-laudo; e o desenvolvimento de uma métrica baseada em redes Bayesianas para avaliação da qualidade clínica. Ambas contribuições consideram a avaliação de imagens de tomografia computadorizada de crânio. Os resultados apresentados atestam que a abordagem proposta apresenta boa correlação entre a qualidade predita e a subjetiva.

Palavras-chave: Qualidade visual. DICOM. PACS. Imagens médicas. Compressão de imagens.

Abstract

Currently, the medical images have a significant impact on the diagnosis of pathologies, consequently, the development of metrics for the evaluation of such images has been increasingin in order to preserve or improve its original features when subjected to compression, enhancement and digital transmission operations. The complexity in order to performance analysis of a metric includes the correlation of metrics with the subjective data obtained experimentally and also the difficulty of obtaining such data. Therefore, development of new metrics and comparative evaluation of metrics applied to diagnostic images are relevant. In this work, a methodology for the acquisition of subjective data and visual quality assessment metrics Bayesian network based developed specifically for the context of diagnostic images are proposed.
Keywords: Visual quality. DICOM. PACS. Medical images. Image compression.

Lista de Abreviaturas

ACR/NEMA American College of Radiology and National Equipment Manufacturers Association

ANOVA Analysis of Variance

AUC Area Under Curve

CEP Comitê de Ética em Pesquisa

CT Computed Tomography

DAIC Diagnostically Acceptable Irreversible Compression

DCT Discrete Cosine Transform

DICOM Digital Imaging and Communication in Medicine

DWT Discrete Wavelet Transform

EPM Edge Preservation Measure

FDA Food and Drug Administration

GMW-GDSIM Gradient Map Weighting-Gradient Diretion Similarity

HU Hounsfield Unit

HVS Human Visual System

ICS/UFPA Instituto de Ciências da Saúde da Universidade Federal do Pará

ITU International Telecommunication Union

MOS Mean Opinion Score

MSSIM Mean Structural Similarity Index

OS Opinion Score

PACS Picture Archiving and Communications Systems

ROC Receiver Operating Characteristic

SSIM Structural Similarity Index

TCLE Termo de Consentimento Livre e Esclarecido

VGC Visual Grading Characteristics

VQEG Video Quality Experts Group

VQM Visual Quality Metric

Lista de Símbolos

M Largura da imagem

N Altura da imagem

c Coeficiente Wavelet

Δ Passo de quantização

$G_{abs,s,t}$ Taxa em relação à aparência de uma estrutura na imagem

S Estrutura na imagem

T Número de avaliadores

S_x Filtro de Sobel horizontal

S_y Filtro de Sobel vertical

G_x Gradiente da imagem na direção horizontal

G_y Gradiente da imagem na direção vertical

GD_r Direção do gradiente da imagem original

GD_d Direção do gradiente da imagem distorcida

GN_r Magnitude do gradiente da imagem original

GM_d Magnitude do gradiente da imagem distorcida

g Kernel de visualização

σ^2 Variância

Δ_g Variação da magnitude entre imagens

Δ_α Variação da direção entre imagens

μ Média

σ Desvio padrão

SE Sensibilidade

SP Sensitividade

Lista de Figuras

2.1 Etapas dos sistemas de gestão de imagens médicas desde sua aquisição até seu armazenamento. 7

4.1 Exemplo de uma função de janelamento: p [HU], e O denota o valores de pixels antes (em Unidades de Hounsfield) e depois do processo de janelamento, respectivamente; o intervalo [0, L-1] representa o range de exibição de preto e branco. 17

4.2 Slice com diferentes janelamentos: (a) osso; (b) tecidos moles. 18

4.3 Algoritmo para compressão de imagens. 19

4.4 Matriz de normalização JPEG utilizada no processo de quantização. 21

4.5 Exemplo de imagem codificada por JPEG: (a) imagem original; (b) imagem reconstruída. 22

4.6 Sub-bandas da decomposição da imagem através da Transformada de Wavelet Discreta. 23

4.7 Exemplo de imagem codificada por JPEG 2000: (a) imagem original; (b) imagem reconstruída. 23

4.8 Etapas do procedimento para coleta de dados. 26

4.9 Interface gráfica para coleta de dados de imagens. 27

5.1 Diagrama de fluxo do cálculo do GMW-GDSIM, onde $r(i,j)$, $d(i,j)$ representam a imagem de referência e distorcida, respectivamente; $GD_r(i,j)$ e $GD_d(i,j)$ são dados pela Eq. 5.2; $GM_r(i,j)$ e $GM_d(i,j)$ calculados a partir da Eq. 5.4; $GDSIM(i,j)$ é dado pela Eq. 5.3; $GM_c(i,j)$ pela Eq. 5.5; Filtro baixa magnitude pela Eq. 5.6; Filtro Gaussiano pela Eq. 5.8; $GMW(i,j)$ pela Eq. 5.7 e $GMW - GDSIM$a partir da Eq. 5.9. 30

5.2 (a) Similaridade entre as direções de gradiente das imagens de referência e distorcida (GDSIM); (b) Similaridade ponderada entre as magnitudes dos gradientes das imagens de referência e distorcida (GMW). 33

5.3 Diagrama de blocos do EPM, onde $r(i,j)$, $d(i,j)$ representam a imagem de referência e distorcida, respectivamente; $GD_r(i,j)$ e $GD_d(i,j)$ são obtidos através da Eq. 5.2; $GM_r(i,j)$ e $GM_d(i,j)$ são dados pela Eq. 5.4; $\Delta_g(i,j)$ e $\Delta_\alpha(i,j)$ são computados pelas Eq. 5.10 e Eq. 5.11, respectivamente; $Q_{GM}(i,j)$ e $Q_{GD}(i,j)$ pela Eq. 5.12; $w(i,j)$ dado pela Eq. 5.14 e $Q(i,j)$ e Q_w pelas Eq. 5.13 e Eq. 5.15, respectivamente. 34

5.4 (a) Avaliação de qualidade de preservação (Q); (b) Avaliação de importância perceptual (w). 35

5.5 Mapa de similaridade estrutural (SSIM). 37

5.6 Gráfico de dispersão entre o MOS e VQMs. (a) Imagens comprimidas por JPEG; (b) Imagens comprimidas por JPEG 2000. 39

6.1 Matriz de confusão da Rede Bayesiana em relação ao MOS. (a) Imagens comprimidas por JPEG; (b) Imagens comprimidas por JPEG 2000. 43

6.2 Curva ROC de VQMs em relação ao MOS. (a) Imagens codificadas por JPEG; (b) Imagens codificadas por JPEG 2000. 44

6.3 Matriz de contingência da rede Bayesiana em relação ao MOS. (a) Imagens comprimidas por JPEG; (b) Imagens comprimidas por JPEG 2000. 44

6.4 Gráfico do teste de Kolmogorov-Smirnov, onde SE é a sensibilidade e SP representa a sensitividade dos dados de MOS em relação à VQM; KS é dado pela Equação 6.1. 45

6.5 Gráfico do teste de Kolmogorov-Smirnov das VQMs em relação às imagens JPEG. 46

6.6 Gráfico do teste de Kolmogorov-Smirnov das VQMs em relação às imagens JPEG 2000. 46

6.7 Matriz de contingência da VQM em relação ao MOS. Baseada em 47

6.8 Matrizes de contingência das VQMs em relação ao MOS para as imagens JPEG. 48

6.9 Matrizes de contingência das VQMs em relação ao MOS para as imagens JPEG. 48

Lista de Tabelas

3.1 Estruturas para avaliação qualidade para exames de tomografia computadorizada de crânio 14
3.2 Critérios de reprodução visual de qualidade para exames de tomografia computadorizada de crânio 14

4.1 Exemplos de *Hounsfield units* de alguns tipos de materiais e tecidos. 18
4.2 Escala para avaliação de imagens clínicas 26

5.1 Saída da rede Bayesiana em relação às aparências das estruturas 38

6.1 Resultados da avaliação de desempenho das VQM's em relação aos mapeamentos estatísticos. 42
6.2 Teste de Kolmogorov-Smirnov. 47

Sumário

Resumo **VII**

Abstract **VIII**

Lista de Abreviaturas **XI**

Lista de Abreviaturas **XI**

Lista de Figuras **IX**

Lista de Tabelas **X**

Sumário **XI**

1 Introdução **1**

1.1 Motivação 1

1.2 Justificativa 2

1.3 Objetivos 2

1.4 Visão geral do trabalho 3

1.5 Organização do texto 3

2 Compressão de imagens médicas **5**

2.1 Introdução 5

2.2 Regulamentação sobre padrões de compressão 6

2.3 Armazenamento e transmissão de imagens médicas 7

3 Avaliação de qualidade visual de imagens médicas **9**

3.1 Introdução 9

3.2 Tipos de avaliação de qualidade visual 10

3.3 Benefícios de novos bancos de imagens médicas 11
3.4 Critérios de qualidade para imagens médicas 12

4 Banco de imagens médicas avaliadas 15
4.1 Introdução 15
4.2 Processamento das imagens experimentais 16
4.2.1 Janelamento 16
4.2.2 Codificação 17
4.2.3 Interpolação 23
4.2.4 Conjunto de teste 24
4.3 Comitê de ética 24
4.4 Perfil dos avaliadores 25
4.5 Sistema de Avaliação 25

5 Redes Bayesianas para avaliação de qualidade visual 28
5.1 Introdução 28
5.2 Métricas de avaliação de qualidade visual 29
5.2.1 *Gradient map weighting-gradient diretion similarity* 29
5.2.2 *Edge preservation measure* 32
5.2.3 *Mean structural similarity index* 36
5.3 Redes Bayesianas 37

6 Resultados 41
6.1 Introdução 41
6.2 Experimento de validação 41
6.3 Discussão 48

7 Considerações Finais 49
7.1 Conclusões 49
7.2 Trabalhos Futuros 50
7.3 Artigos Publicados 50

Referências Bibliográficas 51

Capítulo 1

Introdução

1.1 Motivação

O desenvolvimento de algoritmos para para avaliação automática de qualidade visual é explicado pela redução de custo de sistemas computacionais e pelo crescente número de aplicações em que imagens digitais desempenham um papel central. Atualmente na área médica, imagens possuem função indispensável no processo de diagnóstico. A grande quantidade de exames produzidos em hospitais e clínicas pode limitar a implementação de PACS (*Picture Archiving and Communication Systems*), levando à necessidade de otimização dos sistemas de gestão de imagens médicas. A compressão de imagens médicas pode reduzir os custos de transmissão e armazenamento de dados, sendo estratégico dispor de técnicas automáticas de ajuste de codificadores para maximizar a compressão, sem prejuízo da qualidade clínica. A versão atual do DICOM (*Digital Imaging and Communication in Medicine*), em decorrência da necessidade de economia de unidades de armazenamento, bem como da melhor utilização da banda disponível para transmissão em PACS, permite o uso de estratégias para compressão de imagens, dentre as quais os padrões JPEG e JPEG 2000 [1]. Os maiores ganhos em termos de compressão advêm do uso de codificadores com perdas, havendo um compromisso entre qualidade e fator de compressão. Como a precisão diagnóstica depende intrinsecamente da qualidade da imagem apresentada, há grande restrição ao uso de codificadores com perdas na área médica.

Nesse cenário, a seleção de parâmetros de codificadores requer uma métrica de qualidade visual (VQM, *Visual Quality Metric*) concebida para otimizar o compromisso entre qualidade clínica e compressão. Contudo, a grande maioria das VQMs apresentadas na literatura tem por objetivo representar a opinião média de usuários comuns (não especialistas) a respeito de

cenas naturais (paisagens, pessoas, rostos, etc.), onde não há preocupação em relação ao valor diagnóstico face às características de imagens clínicas. A ausência de bancos de imagens médicas com avaliações subjetivas emitidas por especialistas dificulta o desenvolvimento de VQMs voltadas para a finalidade em pauta. Nesse cenário, a definição de um protocolo de avaliação subjetiva de imagens médicas incentivaria mais amplas pesquisas na área, suscitando a construção padronizada de bases públicas de imagens médicas avaliadas, e consequentemente, o desenvolvimento de VQMs e de novos algoritmos.

1.2 Justificativa

Considerando a Resolução nº 1821/07, do Conselho Federal de Medicina, que regulamenta que é responsabilidade da unidade médico-hospitalar geradora de prontuários estabelecer o armazenamento permanente das imagens para exames clínicos, gerir toda essa quantidade de informação que cresce continuamente com o tempo é um desafio uma vez que os arquivos tendem a ocupar grande volume nos dispositivos de armazenamento de clínicas e hospitais. [2].

Adicionalmente, em aplicações de telemedicina, a transmissão de exames contendo imagens requer infraestrutura de comunicação compatível, o que raramente ocorre em regiões brasileiras rurais ou remotas, especialmente carentes da presença de especialistas em saúde [3].

A natureza do problema é otimizar a demanda por espaço, minimizar o custo, maximizar a resiliência e obter alta performance. Logo, se a avaliação de qualidade de imagens para diagnóstico for automatizada, o sistema poderá reduzir a quantidade de dados que será estocada, diminuindo o tempo e custo dos processos de armazenamento e comunicação [4].

1.3 Objetivos

Os objetivos do trabalho são:

1. **Objetivo primário:** Desenvolver modelos de avaliação de qualidade visual voltados para imagens de tomografia computadorizada que caracterizem a inspeção visual da qualidade clínica realizada por especialistas em fase de pré-laudo, ou seja, antes das imagens produzidas pelo tomógrafo serem enviadas ao médico que redigirá o laudo.

2. **Objetivos secundário:**

a) Propor procedimentos experimentais para avaliar a qualidade clínica de imagens de tomografia computadorizada submetidas a compressão JPEG e JPEG2000 que siga as recomendações do International Telecommunication Union (ITU) e considere os critérios definidos no European Guidelines on Quality Criteria for Computed Tomography;

b) Implementação e realização do procedimento para obtenção de dados de avaliação de qualidade clínica de imagens de tomografia computadorizada;

c) Avaliar e comparar o desempenho de métricas de qualidade visual já propostas na literatura especializada à luz dos dados obtidos neste projeto;

d) Propor métricas de qualidade visual que possam ser aplicadas na otimização de sistemas de codificação de imagem, de maneira a maximizar a compressão sem que haja prejuízo à qualidade clínica.

1.4 Visão geral do trabalho

Este trabalho apresenta uma metodologia para aquisição de dados subjetivos a partir de imagens de tomografia computadorizada de crânio, voltada para a caracterização dos critérios de avaliação utilizados por especialistas na fase pré-laudo. Foi implementado em linguagem Python um sistema para visualização de imagens de teste e coleta da opinião de especialistas que segue as recomendações do *International Telecommunication Union* (ITU) para avaliação de imagens [5] e também incorpora elementos do *European Guidelines on Quality Criteria for Computed Tomography* [6]. Adicionalmente, é proposta uma estratégia de avaliação de qualidade visual baseada em redes Bayesianas, especificamente para o contexto de imagens de tomografia de crânio.

1.5 Organização do texto

O restante deste documento está organizado como segue. O Capítulo 2 discute a compressão de imagens para imagens de diagnóstico. O Capítulo 3 apresenta um estudo sobre a avaliação de qualidade visual com foco em imagens médicas, com ênfase no compromisso entre compressão e qualidade clínica das imagens. Os Capítulos 2 e 3 constituem a revisão bibliográfica do documento. Em seguida, o Capítulo 4 aborda a definição do protocolo de avaliação subjetiva, assim como, a implementação de um sistema para coleta de dados. No

Capítulo 5, a métrica de qualidade visual proposta é detalhada. O Capítulo 6 apresenta e discute os resultados. As conclusões e propostas de trabalhos futuros são expostas no Capítulo 7.

Capítulo 2

Compressão de imagens médicas

2.1 Introdução

O principal objetivo do uso de algoritmos de compressão de imagens é minimizar o número de bits necessários para a reprodução de uma determinada cena, mantendo aceitável a qualidade visual. Para imagens médicas, em caso de transmissão ou armazenamento, técnicas de compressão representam uma alternativa atraente, desde que sua qualidade clínica seja preservada. Uma vez que a presença de artefatos visuais pode comprometer a conclusão diagnóstica [7], o uso de codificação sem perda (ou reversível) é recomendado. Nesta abordagem, a imagem reconstruída é igual à imagem original, ou seja, à imagem antes de realizada a operação de codificação.

A codificação sem perda, no entanto, apresenta reduzidas taxas de compressão em comparação às abordagens com perda (ou irreversíveis). Assumindo que, além das redundâncias espacial e de codificação, existem artefatos imperceptíveis para o sistema visual humano (HVS, *Human Visual System*), é plausível a obtenção de um codificador com perdas dirigido segundo um modelo perceptual inspirado em experimentos subjetivos [8], neste caso, visando a preservação dos elementos da imagem responsáveis pelo seu valor clínico. Atualmente, a compressão irreversível diagnosticamente aceitável, (DAIC, *diagnostically acceptable irreversible compression*) pressupõe supervisão de um especialista para garantir que não haja redução da qualidade clínica da imagem. A taxa de compressão tolerada depende do tecido analisado, da modalidade do exame e principalmente do objetivo do estudo [6].

De uma perspectiva prática, isso significa que nenhum artefato gerado por esquemas de compressão deva ser perceptível ao visualizador humano ou interferir na interpretação da ima-

gem. Logo, garantir a qualidade aceitável da imagem reconstruída é o foco principal no domínio da compressão de imagens médicas. Este capítulo discute algoritmos e a regulamentação de técnicas de compressão para o caso específico de imagens para diagnóstico.

2.2 Regulamentação sobre padrões de compressão

A fim de melhorar a precisão, clareza e utilidade de documentos clínicos, em 1993 o *American College of Radiology* (ACR) e o *National Electrical Manufacturers Association* (NEMA) desenvolveram o padrão *Digital Imaging and Communication in Medicine* (DICOM) que viabiliza a armazenagem, transmissão, impressão, e exibição de imagens médicas, em uma estrutura universalmente compreendida [4].

A versão atual do DICOM, em decorrência da necessidade de economia de unidades de armazenamento, bem como melhor utilização da banda disponível para transmissão de dados em sistemas PACS, admite o uso de estratégias para compressão de imagens. Adicionalmente, o ACR/NEMA define que se esquemas de compressão reversíveis ou irreversíveis forem adotados, somente algoritmos como *Joint Photographic Experts Group* devam ser utilizados, entre eles JPEG-LS, JPEG-2000, ou MPEG [1].

O FDA/U.S. (*Food and Drug Administration*) também não restringe o uso de algoritmos de compressão para imagens médicas, desde que a qualidade da imagem reconstruída seja aprovada por um especialista e que os sistemas que adotam compressão com perdas (equipamentos em estações de trabalho, por exemplo) submeta-se à avaliação dos impactos do esquema de compressão adotado, utilizando métricas de qualidade visual (VQM, *visual quality metrics*). Todavia, o FDA/U.S não recomenda a adoção de algoritmos de compressão para imagens digitais de mamografia porque estudos anteriores mostraram que as imagens decodificadas não são pertinentes para diagnóstico. No Brasil, a situação regulatória não é definida [9].

Nem o ACR/NEMA, FDA/U.S. ou a documentação do DICOM estabelecem recomendação, aprovação ou declaração geral sobre as taxas de compressão apropriadas para qualquer modalidade, doença ou aplicação clínica específica que assegure a manutenção da qualidade para fins de diagnóstico [6, 9].

O padrão JPEG apresenta limitações se comparadas a outras técnicas, tais como: pobre eficiência de compressão em esquemas reversíveis, e presença de artefatos em formato de blocos provenientes da selção de altas taxas de compressão em esquemas irreversíveis. Além

disso, o JPEG básico aceita como entrada imagens com no máximo 12 bits/pixel de profundidade de bits, o que pode representar uma grande limitação visto que imagens médicas podem frequentemente ultrapassar esse grau de profundidade.

O padrão JPEG2000 [10], por sua vez, representa uma alternativa em relação ao JPEG, considerando o compromisso entre taxas de compressão e distorções, apresentando ainda várias funcionalidades convenientes a aplicações na área médica, como por exemplo, compressão com ou sem perda integradas, codificação de região de interesse, e decodificação progressiva [1].

2.3 Armazenamento e transmissão de imagens médicas

Um sistema PACS é definido por todas as tecnologias que integram os processos de gestão de imagens radiológicas, ou de suporte a diagnóstico, de determinada unidade medico-hospitalar. A administração de tais imagens vai desde a coleta de dados demográficos durante o cadastro do paciente até a emissão do laudo final pelo médico radiologista com o subsequente armazenamento da imagem.

O fluxo de trabalho se inicia com a análise da qualidade de uma imagem médica pelo técnico ou médico radiologista, logo após o procedimento de aquisição. Se a imagem não apresentar qualidade suficiente para diagnóstico, repete-se o exame; caso contrário, a imagem é armazenada nos servidores PACS até a requisição de um profissional para emissão do laudo. Em seguida, a imagem deve ser armazenada para futuras consultas de confirmação ou análise de progressão do diagnóstico. Na Figura 2.1 é apresentado um diagrama representado o fluxo de trabalho que uma imagem médica percorre desde sua aquisição até ser armazenada [4].

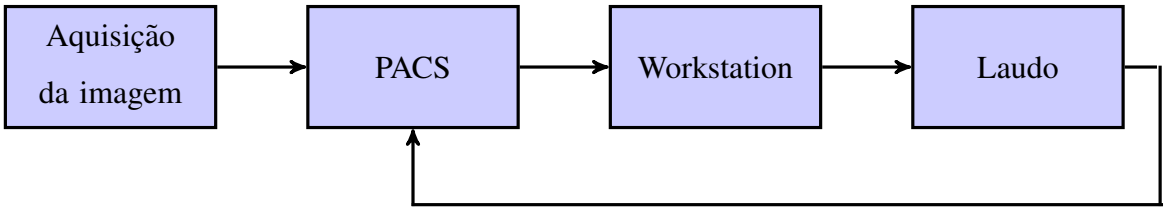


Figura 2.1: Etapas dos sistemas de gestão de imagens médicas desde sua aquisição até seu armazenamento.

Considerando os avanços da tecnologia da informação e de telecomunicações, que oferecem novos métodos de armazenamento e transmissão de dados, a Resolução nº 1821/07, do Conselho Federal de Medicina, regulamenta que é responsabilidade da unidade médico-

hospitalar geradora do arquivo estabelecer o armazenamento permanente dos prontuários dos pacientes arquivados eletronicamente em meio óptico, microfilmado ou digitalizado [2].

Logo, os arquivos de prontuário tendem a ocupar grande volume nos dispositivos de armazenamento de clínicas e hospitais. Com o advento de *scanners* por *slices*, como no caso da tomografia computadorizada ou ressonância magnética, é possível examinar estruturas de interesse milimetricamente espassadas, resultando em exames contendo alguns milhares de *slices*. Por dia, um hospital regular produz imagens em volume tal que sejam necessários dezenas de gigabytes para armazená-las. Gerir toda essa quantidade de informação que cresce continuamente com o tempo é um desafio.

Em aplicações de telemedicina, a transmissão de exames contendo imagens requer infraestrutura de comunicação compatível, o que raramente ocorre em regiões brasileiras rurais ou remotas, especialmente carentes da presença de especialistas em saúde [3].

A natureza do problema é otimizar a demanda por espaço, minimizar o custo, maximizar a resiliência e obter alta performance. Logo, se a informação não-crítica for removida, o sistema poderá reduzir a quantidade de dados que será estocada, diminuindo o tempo e custo dos processos de armazenamento e comunicação [4].

Atenta ao problema do tamanho dos arquivos de imagens médicas, a comunidade científica investiga estratégias de compressão com perda para o caso específico de imagens para diagnóstico e a avaliação da qualidade visual das imagens a fim de evitar a degradação da informação clínica.

Capítulo 3

Avaliação de qualidade visual de imagens médicas

3.1 Introdução

A eclosão de novas tecnologias tem facilitado o aperfeiçoamento de métodos de aquisição de imagens, assim como, o surgimento de novas modalidades de sinais visuais para fins diagnósticos [11], levantando questões a respeito da aquisição, armazenamento, transferência, compressão, e exibição de imagens médicas [12]. Entre elas, a qualidade visual, que pode sofrer perdas durante tais processos ao ponto de inviabilizar seu uso [13–16].

Em imagens médicas, artefatos visuais podem levar o especialista ao erro.Portanto, preservar a qualidade clínica de uma imagem médica ao longo da cadeia de processamento é uma tarefa crítica. A avaliação de qualidade visual por si, mesmo em um contexto mais geral, é um processo complexo em que diversos fatores são levados em consideração: percepção de contraste; resolução da imagem; presença de variados tipos de ruído e artefatos na imagem; iluminação ambiente; e distância de visualização. A complexidade é ainda maior em imagens médicas, pois modalidade tem suas especificidades correspondentes aos fenômenos físicos do sistema de aquisição e fisiológicos da região analisada [12].

Métricas de qualidade para imagens médicas desempenham relevante papel no apoio à minimização da radiação emitida a pacientes por ocasião de exames de raio-X ou tomografia. O objetivo, nesse caso, é obter uma imagem viável à tarefa de diagnóstico (imagem contém informação suficiente para a melhor decisão médica) com a menor dose de radiação possível, preservando o paciente e reduzindo os custos de operação [17].

Mas, para além da definição de parâmetros de tomógrafos e outros equipamentos, como avaliar a qualidade visual de imagens médicas? A resposta considera que é a avaliação da condição patológica que determina a informação que deve ser preservada nessa imagem.Ou seja, o objetivo não deve ser julgar a mera aparência da imagem, mas sua utilidade para as tarefas de detecção e classificação de doenças (qualidade clínica) [11].

Neste capítulo, é apresentada uma revisão de como a avaliação de qualidade visual tem sido estudada para o caso específico de imagens médicas. Serão abordados os tipos de avaliação de qualidade, assim como, suas respectivas especificidades para dada aplicação. Em seguida, são discutidos os benefícios da criação de novos bancos de imagens médicas. E finalmente, quais os critérios definidos para a avaliação de qualidade visual de imagens para diagnóstico.

3.2 Tipos de avaliação de qualidade visual

De maneira geral, há duas formas de avaliação de qualidade visual: a avaliação subjetiva, onde a qualidade é determinada a partir da opinião de pessoas sobre imagens que lhes são apresentadas em situações controlados; e a avaliação objetiva, que corresponde ao resultado obtido pelo uso algoritmos ou funções matemáticas na estimação de qualidade de uma imagem [18].

Considerando o disposto na Resolução do Conselho Federal de Medicina, nº 1.643/2002, é responsabilidade do médico avaliar cuidadosamente a informação que recebe e somente emitir opiniões e recomendações ou tomar decisões médicas se a qualidade da informação recebida for suficiente e pertinente para o cerne da questão. Isso implica que a qualidade de uma imagem que foi submetida a algum tipo de processamento, em qualquer campo médico, deve ser avaliada por especialistas antes que a mesma chegue a etapa de laudo, ou sirva de suporte a diagnóstico [19].

A maneira mais confiável de avaliação de qualidade clínica é a subjetiva com a participação de médicos especialistas. Além de pouco prática e dispendiosa, ao se pensar em grande quantidade de imagens, as respostas podem variar significativamente entre profissionais. Nesse contexto, métricas de qualidade visual representam uma alternativa à avaliação subjetiva, principalmente no tratamento automático de grande volume de imagens.A comunidade radiológica ainda não adotou um método padrão ou protocolo de avaliação objetiva para aferição de qualidade clínica [20].

A maioria das métricas de qualidade visual existentes foram desenvolvidas a partir de experimentos cujas imagens de teste são convencionais (paisagens naturais, pessoas, rostos,

animais, etc.) e os avaliadores, usuários comuns (não médicos ou técnicos radiologistas). Para imagens médicas, o avaliador deve ser um especialista e o critério de qualidade está especificamente relacionado ao valor diagnóstico da imagem.

Na maioria dos estudos, a performance de observadores humanos é o teste final para a validação de determinada métrica e uma vez que dada métrica, ou um conjunto de métricas forem validados para o caso específico de imagens médicas o número de testes de avaliações de qualidade sem radiologistas pode aumentar, otimizando o processo de avaliação [21].

Em [22, 23], são apresentados estudos baseados em experimentos de avaliação de qualidade visual subjetiva de imagens codificadas aplicados à caracterização de desempenho de métodos objetivos para avaliação de qualidade diagnóstica de tais imagens.

3.3 Benefícios de novos bancos de imagens médicas

Apesar de haver vários bancos de imagens de cenas naturais avaliadas subjetivamente disponíveis em domínio público, no campo da avaliação da qualidade de imagens médicas a situação é diferente. A ausência de tais bancos dificulta pesquisas na área e mesmo no caso de imagens médicas não avaliadas é difícil encontrar um banco de imagens disponível publicamente com imagens com características uniformes, ou seja, a amplitude, orientação, tamanho, posição e número de lesões das imagens sejam semelhantes.

Isso porque implicações éticas impedem que imagens de pacientes sejam expostas, uma vez que o disposto na Resolução do Conselho Federal de Medicina nº 1.643/2002 considera que as informações sobre o paciente identificado só podem ser transmitidas, até mesmo a outro profissional, com prévia permissão do paciente, mediante seu consentimento livre e esclarecido e sob rígidas normas de segurança capazes de garantir a confidencialidade e integridade de tais informações [11, 24].

Além disso, para a criação de novos bancos, os avaliadores devem ser submetidos a um ambiente controlado e condições adaptadas similares às aplicadas durante a rotina clínica de exames, como luminosidade da sala, distância e ângulo de visualização, e resolução da tela de exibição. É recomendado que o conjunto de imagens deva ser escolhido de forma que as características das imagens sejam uniformes. Testes psicovisuais requerem protocolos estritos e difíceis de implementar, são demorados, dispendiosos e dependem da especialidade médica e do tipo de exame, além do fato que os métodos de avaliação dependem da informação que deve

ser extraída das imagens, o que os tornam pouco práticos [12, 19, 21].

Logo, apesar de métricas que representem a percepção humana de qualidade visual serem objeto de numerosos estudos em processamento de sinais, a aplicação do conhecimento decorrente de tais estudos na caracterização da qualidade de imagens com objetivo diagnóstico ainda é bastante incipiente, visto a ausência de banco de imagens específicas avaliadas e dada a complexidade de tal tarefa, havendo poucos trabalhos abordando esse assunto, dentre eles [12].

Pois, apesar de muitas das métricas propostas na literatura serem baseadas em propriedades do sistema visual humano, no entanto, até agora nenhuma é capaz de levar em consideração todas as peculiaridades do HVS. Isso porque o sistema ainda não é totalmente compreendido pela medicina e ainda não é claro como modelar todas os possíveis tipos e níveis de distorções em uma imagem.

Além disso, as condições para validar os experimentos, metodologia para processar os resultados desses experimentos, o número necessário de participantes, são outras questões a serem respondidas. Todos esses fatores tornam a avaliação de qualidade de imagens médicas um desafio.

A disponibilidade de bancos de imagens médicas avaliadas subjetivamente facilitaria a caracterização do desempenho de métricas para o caso específico de imagens de diagnóstico, assim como o desenvolvimento de novos algoritmos específicos para esse contexto.

3.4 Critérios de qualidade para imagens médicas

O bom desempenho do diagnóstico exige que a qualidade da imagem seja suficiente para atender ao requisito clínico do exame, mantendo a mínima dose de radiação ao paciente possível. Para garanti-la, deve haver uma seleção cuidadosa dos parâmetros técnicos que controlam a exposição do paciente, assim como, da exibição das imagens e também a verificação do desempenho do equipamento como parte de um programa de garantia de qualidade. Com isso, o uso ótimo de radiação em exames de imagens de Tomografia Computadorizada (CT, *Computed Tomography*) envolve três aspectos: a qualidade da imagem, a dose de radiação e a escolha adequada da técnica ao exame [25].

Nesse cenário, o *European Study Group of Radiologists and Physicists* propôs-se a investigar os princípios associados a boas técnicas de imagem e estabeleceu critérios de qualidade de exames de imagem que garantam a preservação de toda informação clínica necessária

ao diagnóstico, com doses mínimas para o paciente para diferentes modalidades de exames, como: radiologia convencional, tomografia computadorizada, medicina nuclear, tomografia por emissão de pósitron, imagens de ressonância magnética e ultrassom. Tais critérios são utilizados por médicos ou técnicos radiologistas nas rotinas de exames para definir se a imagem possui o padrão de qualidade de diagnóstico.

Os critérios não podem ser os mesmos aceitos em todos os casos, isso porque algumas áreas anatômicas em determinados exames não interferem no diagnóstico e podem ser visualizadas com baixa qualidade em relação a outras. Como a qualidade difere para cada técnica é necessário que haja diferentes critérios para cada tipo de exame. Em relação a imagens de CT, o grupo lista critérios de qualidade através do guia qualitativo *European Guidelines on Quality Criteria for Computed Tomography* para as técnicas de exames: crânio, face e pescoço, coluna vertebral, tórax, abdômen e pélvis, e ossos e articulações. Cada grupo de exames é subdividido em órgãos específicos ou partes do corpo que devem ser analisados de acordo com a área de interesse.

Os critérios de qualidade são baseados nas especificações anatômicas, com foco na avaliação da área anatômica dos órgãos e do contraste entre diferentes tecidos. Dessa forma, o guia define quais as estruturas essenciais e o grau necessário de visibilidade delas para garantir precisão diagnóstica. Atualmente, não existem definições internacionalmente aceitas, para o propósito dessas diretrizes. Como um exemplo prático, o grau de visibilidade de Crânio de CT para diferentes tipos de janelamento (tecido mole e osso) é definido da seguinte forma:

- Visualização: órgãos e estruturas que devem ser perceptíveis. Na Tabela 3.1, são listadas as estruturas definidas em [6].

- Reprodução visual: estruturas particulares da indicação específica que devem ser discriminadas a um nível essencial para o diagnóstico. Na Tabela 3.2, são listados os critérios de reprodução visual definidos em [6]. Isso incluirá os termos:

  - Reprodução: detalhes das estruturas anatômicas devem ser visíveis, mas não necessariamente precisam ser claramente definidos.

  - Reprodução crítica: os detalhes anatômicos devem ser claramente definidos.

Tabela 3.1: Estruturas para avaliação qualidade para exames de tomografia computadorizada de crânio

Estruturas para tecidos moles
Cérebro
Cerebelo
Base do crânio
Vasos após contraste intravenoso
Estruturas para osso
Base do crânio
Cerebelo
Parte basal dos lobos frontais
Vasos após contraste intravenoso

Tabela 3.2: Critérios de reprodução visual de qualidade para exames de tomografia computadorizada de crânio

Critérios para tecidos moles
Reprodução crítica dos limite entre a massa branca e a cinzenta
Reprodução crítica dos gânglios da base
Reprodução crítica do sistema ventricular
Reprodução crítica do espaço do líquido cefalorraquidiano em volta do mesencéfalo
Reprodução crítica do espaço do líquido cefalorraquidiano em volta do cérebro
Reprodução crítica dos grandes vasos e plexo coroide após de contraste intravenoso
Critérios para osso
Reprodução crítica das estruturas corticais e trabeculares dos ossos
Reprodução crítica dos compartimentos gasosos
Reprodução crítica da sela túrcica
Reprodução crítica dos contornos cerebelares
Reprodução crítica dos limites entre substância branca e cinzenta do cerebelo
Reprodução do espaço do líquido cefalorraquidiano em volta do tronco encefálico
Reprodução crítica dos grandes vasos e plexo coroide após de contraste intravenoso

Capítulo 4

Banco de imagens médicas avaliadas

4.1 Introdução

A disponibilidade de um banco de imagens médicas avaliadas subjetivamente permite a avaliação de desempenho de métricas de qualidade visual para o caso específico de imagens de diagnóstico, como discutido no Capítulo 3. Neste capítulo um sistema é proposto – segundo as recomendações do ITU [5] e os critérios definidos no *European Guidelines on Quality Criteria for Computed Tomography* [6] – para aquisição de dados de avaliação subjetiva de qualidade visual de imagens de tomografia computadorizada de crânio codificadas através das técnicas de compressão JPEG e JPEG2000.

Dado que imagens de tomografia computadorizada são representadas por um conjunto de *slices* que reproduzem o objeto escaneado, permitindo variadas manipulações na imagem durante o diagnóstico [12], é delineado adicionalmente um protocolo de processamento a fim de uniformizar a avaliação e codificar as imagens apresentadas. O perfil dos avaliadores participantes do experimento e os critérios éticos da pesquisa também são apresentados. O experimento foi idealizado para que seja realizado em condições similares àquelas aplicadas durante as rotinas clínicas permitindo a observação de todos os detalhes necessários para o diagnóstico.

O principal objetivo é obter uma coleção de imagens distorcidas e a avaliação de tais imagens, as quais podem servir de suporte para o desenvolvimento de novos modelos de observadores numéricos para avaliação de imagens médicas, verificar a eficiência de algoritmos já disponíveis e compará-los para grupos particulares de distorções [26].

4.2 Processamento das imagens experimentais

O conjunto de imagens distorcidas foi obtido a partir de 3 imagens originais de tomografia computadorizada de crânio, com cortes no plano transversal, dentre as quais 2 imagens sem indicação de estruturas patológicas e 1 imagem com patologias inerentes, todas em formato DICOM (dcm). O objetivo da utilização de ferramentas de processamento de imagens é modificar cada *slice* com as características em que os laudos médicos são emitidos em rotinas clínicas garantindo que o radiologista não modificasse a imagem durante o experimento, resultando em um conjunto uniforme para comparação dos resultados. Logo, todas as imagens foram janeladas, codificadas e interpoladas para exibição.

4.2.1 Janelamento

Como certos detalhes ou estruturas anatômicas ocupam uma faixa estreita de níveis de cinza, é desejável reescalonar a faixa de interesse para toda a gama de exibição acessível. Janelamento de imagens consiste em uma mudança artificial na faixa dinâmica dos pixels, onde um intervalo dos valores dos pixels que serão exibidos é delimitado de acordo com a estrutura anatômica de interesse e os valores fora desse intervalo são ajustados para os valores mínimos ou máximos de exibição. Um exemplo de janelamento é a transformação linear dada por:

$$j(m,n) = \begin{cases} 0 & \text{se } x(m,n) \leq f_1 \\ \frac{r(m,n)-f_1}{f_2-f_1} & \text{se } f_1 < r(m,n) < f_2 \\ 1 & \text{se } r(m,n) \geq f_2 \end{cases} \tag{4.1}$$

onde $r(m,n)$ é a imagem original, $j(m,n)$ é a imagem janelada com a faixa dinâmica normalizada entre $[0,1]$, e $[f_1, f_2]$ são os limites da janela. A faixa dinâmica resultante entre $[0,1]$ precisa ser mapeada para uma faixa de exibição disponível, o que pode ser solucionado multiplicando os valores pelo valor máximo de exibição disponível. Detalhes (pixels) abaixo do limite mínimo f_1 são eliminados (renderizados em preto) e aqueles acima do limite máximo f_2 são saturados (renderizados em branco), a imagem resultante exibe os detalhes da faixa $[f_1, f_2]$ com aumento de contraste, utilizando toda a faixa de exibição disponível (Figura 4.1) [27].

São duas as razões que justificam o janelamento como uma operação necessária na visualização de imagens DICOM. A primeira é que imagens médicas apresentam pixels com

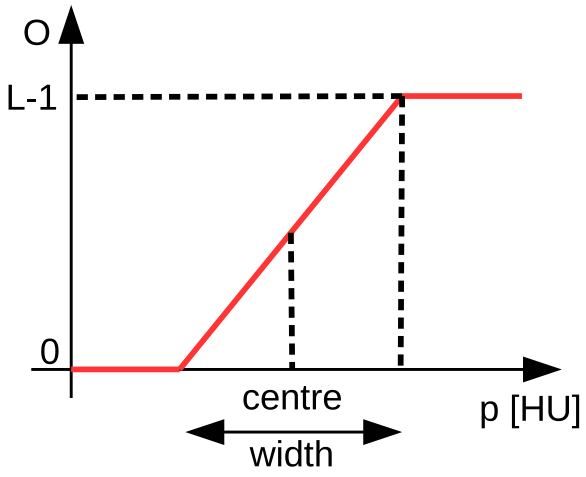


Figura 4.1: Exemplo de uma função de janelamento: p [HU], e O denota o valores de pixels antes (em Unidades de Hounsfield) e depois do processo de janelamento, respectivamente; o intervalo [0, L-1] representa o range de exibição de preto e branco.

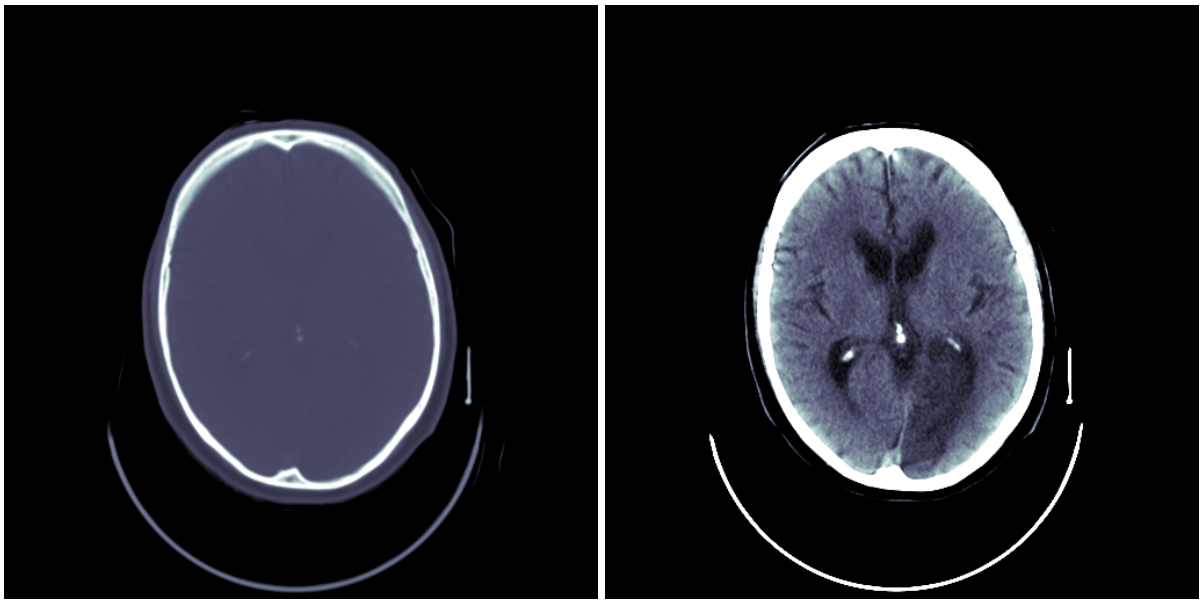
profundidade de bit maior que 8, sendo necessário um reescalamento de valores para adequar a faixa dinâmica das imagens DICOM à capacidade de exibição de dispositivo de visualização como monitores convencionais. A segunda razão é que a escolha de janelas com diferentes valores de centro e largura permite enfatizar diferentes tipos de tecidos. Na Figura 4.2 são apresentado dois resultados aplicando diferentes janelas para a mesma imagem DICOM.

Ar, água, osso e tecidos moles, por exemplo, possuem janelas previamente tabeladas de acordo com a escala *Hounsfield Unit* (HU), a qual associa os valores de pixel referentes a cada tipo de estrutura anatômica. Na Tabela 4.1 são mostrados exemplos de HU referentes a diferentes estruturas anatômicas [28].

4.2.2 Codificação

O mecanismo de algoritmos de compressão de imagens baseada em transformação pode ser sintetizado em dois ou três passos, como descrito na Figura 4.3. Nos casos de compressão sem perda, primeiramente, a imagem é mapeada para outro domínio e em seguida codificada. Se, no entanto, o passo de quantização é introduzido, o método de compressão é dito com perda.

O NEMA define como padrão para codificação de imagens os esquemas de compressão JPEG e JPEG 2000, para imagens de 8-bits e 12-bits, respectivamente, e JPEG-XT para imagens de 16 bits de profundidade [29]. Apesar do uso do JPEG ainda ser dominante no processamento



(a) (b)

Figura 4.2: Slice com diferentes janelamentos: (a) osso; (b) tecidos moles.

Tabela 4.1: Exemplos de *Hounsfield units* de alguns tipos de materiais e tecidos.

Tecido	**HU**
Ar	$-1008... - 1004$
Adiposidade	$-108... - 72$
Pulmão	$-900... - 200$
Água	0
Bile	$8...24$
Rim	$22...42$
Sangue	$24...60$
Pâncreas	$26...54$
Músculo	$30...58$
Necrose	$30...60$
Baço	$34...58$
Fígado	$46...74$
Abcesso	$66...116$
Osso	$902...1108$

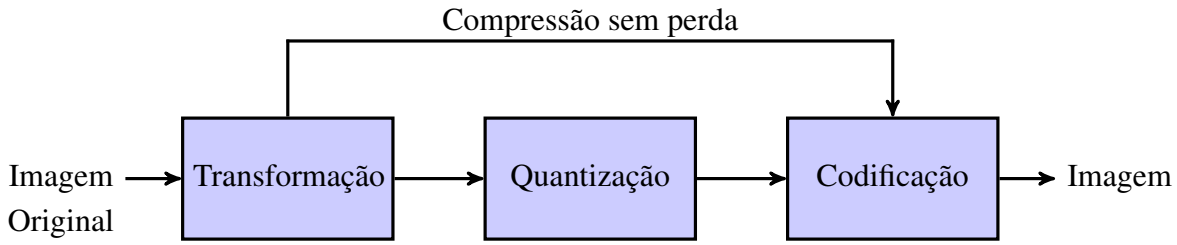


Figura 4.3: Algoritmo para compressão de imagens.

de imagens, o padrão apresenta certas limitações, entre elas, a pobre eficiência de compressão se comparadas a outras técnicas para esquemas reversíveis e a presença de artefatos em blocos. Além disso, para aplicações com imagens médicas, atualmente o estudo de JPEG 2000 em relação a outros algoritmos vem se destacando. Isso porque o JPEG2000 pode fornecer taxas de compressão maiores em relação ao JPEG para o mesmo nível de qualidade, assim como, apresenta várias funcionalidades favoráveis à aplicações em imagens de diagnóstico, como por exemplo, técnicas de compressão com ou sem perda integradas, codificação de região de interesse e decodificação progressiva, entre outros [1].

Embora as imagens originais utilizadas possuam 16 bits de profundidade, os algoritmos JPEG e JPEG 2000 foram escolhidos uma vez que adotam padrões de compressão com perda.

JPEG

O primeiro passo na codificação JPEG de uma imagem em tons de cinza é a decomposição da imagem em blocos de $8X8$ pixels. Tais blocos são processados e armazenados separadamente e somente são agrupados novamente quando a imagem é reconstruída. Essa metodologia, no entanto, apesar de melhorar a eficiência computacional, pode introduzir artefatos em blocos, caracterizados por uma descontinuidade nas bordas dos blocos na imagem reconstruída, principalmente em altas taxas de compressão.

Depois da etapa de decomposição mencionada, cada bloco é mapeado através da Transformada de Cossenos Discreta (DCT, *Discrete Cosine Transform*), a qual é definida como uma representação de uma função por uma combinação linear de funções cossenos. A DCT, $c(i,j)$, de uma imagem denominada $r(x,y)$, de dimensões $M \times N$, é dada por:

$$c(i,j) = a(i)a(j)r(x,y)cos\left[\frac{(2x+1)\pi i}{2M}\right] cos\left[\frac{(2y+1)\pi j}{2N}\right] \tag{4.2}$$

onde

$$a(k) = \begin{cases} \sqrt{\frac{1}{N(\text{ou}M)}} & \text{para } k = 0 \\ \sqrt{\frac{2}{N(\text{ou}M)}} & \text{para } k = 1, 2, \ldots, N(\text{ou}M) - 1 \end{cases} \tag{4.3}$$

Após essa transformação e a etapa de codificação, uma vez que se deseja que a imagem seja reconstruída, a inversa da DCT é aplicada aos blocos codificados, e então o método de compressão é dito sem perda, onde a imagem original pode ser perfeitamente recuperada.

No entanto, se o passo de quantização é introduzido, o método de compressão é com perda. Especificamente, a perda de informação é consequência de dois processos: a normalização psicovisual e o truncamento desses valores para o inteiro mais próximo.

Cada bloco DCT 8×8 corresponde à informação espectral local da imagem. Como uma imagem é um sinal bidimensional não-estacionário, essas assinaturas locais tendem a ser distintas umas das outras. Logo, um bloco com níveis de cinza homogêneos tem um domínio de DCT com poucos coeficientes significativos, enquanto que um bloco com muitas variações tem um domínio de DCT com maiores coeficientes espectrais.

A estratégia da quantização é favorecer os coeficientes espectrais que são mais sensíveis ao HVS, com isso, divide-se os coeficientes da DCT pelos valores correspondentes em uma matriz de normalização.

Estudos psicovisuais permitem definir a sensibilidade do HVS às frequências espaciais. Os valores na matriz de quantização foram escolhidos de maneira a serem inversamente proporcionais à importância visual da componentes espectral na mesma posição. Na figura 4.4 é mostrada a matriz de normalização usada pelo método de compressão JPEG, cujos valores crescem no sentido da esquerda para direita e de cima pra baixo, resultando em uma redução progressiva dos valores de coeficiente de DCT que representam as altas frequências presentes no bloco.

Posteriormente, os valores resultantes da divisão dos coeficientes DCT pela matriz de normalização são arredondados para o inteiro mais próximo, dada que a representação de casas decimais implica em mais consumo de memória. O ajuste do grau de compressão empregado pelo algoritmo é permitido através do escalonamento da lista normalizada pela multiplicação por uma constante. Logo, para altas taxas de compressão, se é realizada alta redução no intervalo dos coeficientes DCT's então muitos valores serão atribuídos a zero. Com certeza, o aumento no número de coeficientes de valor zero também resulta na aumento da degradação do sinal.

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

Figura 4.4: Matriz de normalização JPEG utilizada no processo de quantização [4].

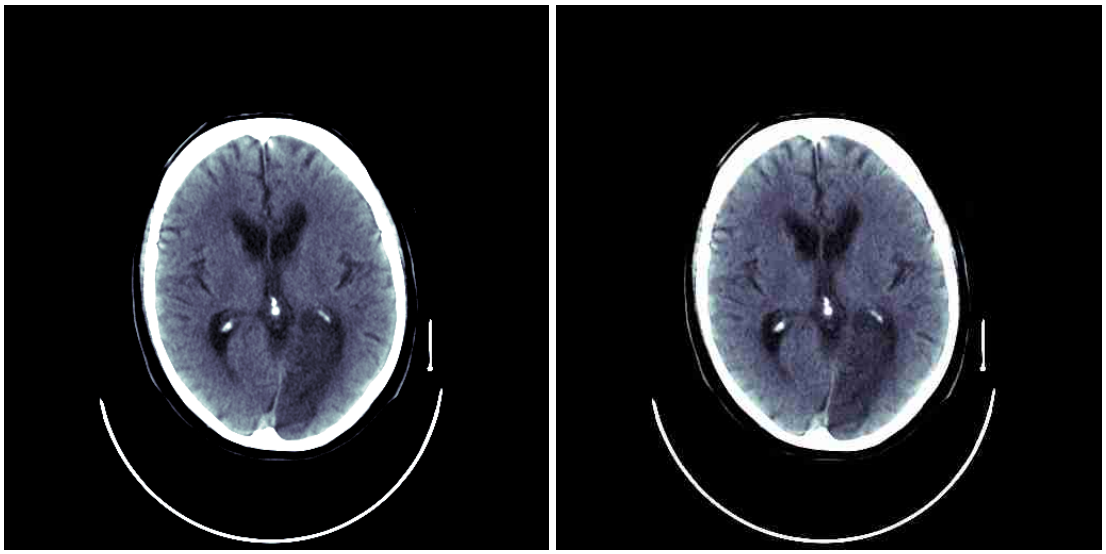
Finalmente,os blocos de coeficientes DCT quantizados são codificados usando uma combinação do método de codificação de Huffman e *run-length*. Antes desses passos, entretanto, os valores na lista são organizados como um vetor maximizando a eficiência da codificação *run-length*, onde o número de coeficientes nulos são colocados no fim do vetor. Para isso, a lista é ordenada usando um esquema *zigzag* começando começando pelos coeficientes DCT's de menor frequência e terminando pelas frequências mais altas.

O vetor resultante é codificado usando um algoritmo similar à codificação de Huffman, entretanto, com uma tabela de codificação pré-determinada. Como é conhecido que cada bloco tem 64 valores, depois que o último coeficiente não nulo é codificado um simples símbolo de fim de bloco é adicionado, o que indica para o algoritmo de decodificação que os valores remanescentes dos coeficientes DCT's normalizados são zeros [4]. Na Figura 4.5 é apresentado um exemplo de imagem codificada utilização o padrão de compressão JPEG.

JPEG2000

Como mostrado na Figura 4.3, algoritmos de compressão usualmente são iniciados por um método de transformação. O JPEG 2000 utiliza a Transformada de Wavelet Discreta (DWT, Discrete Wavelet Transform), que permite melhor compromisso entre compressão e qualidade visual que o JPEG. Um dos motivos, é que no padrão JPEG 2000 a DWT é geralmente aplicada à imagem inteira, evitando os artefatos em blocos que são frequentes no JPEG.

Durante a etapa de transformação, a imagem a ser comprimida é decomposta através da DWT em três sub-bandas de resolução mais baixa da imagem iterativamente, as quais correspondem aos coeficientes wavelet horizontais, verticais e diagonais da imagem processada (Fig. 4.6). Tal processo é repetido até que os valores dos pixels se concentrem próximos a zero.



(a) (b)

Figura 4.5: Exemplo de imagem codificada por JPEG: (a) imagem original; (b) imagem reconstruída.

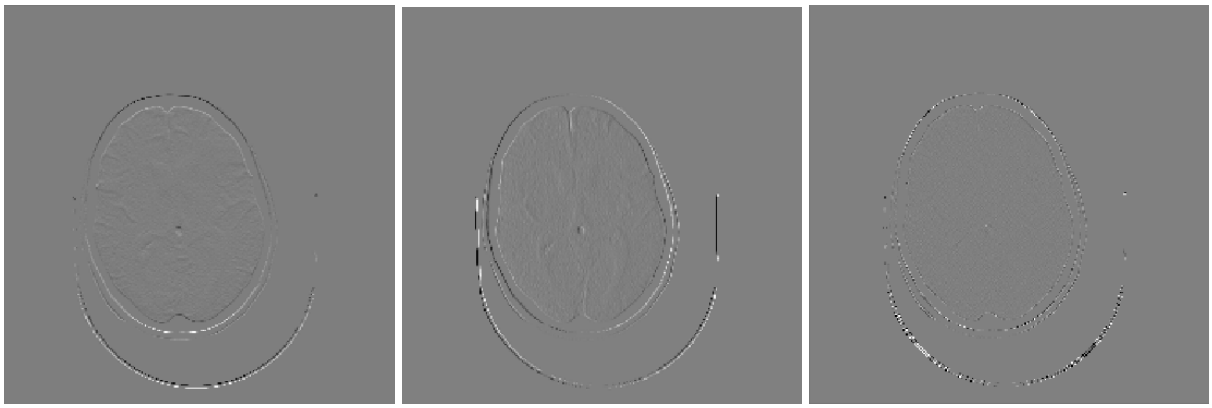
Sendo assim, como na compressão JPEG, o objetivo da DWT no JPEG 2000 é obter imagens correlacionadas com a imagem original, entretanto, com um alto número de pixels com valores pequenos, os quais podem ser facilmente quantizados, o que consequentemente resulta em uma boa taxa de compressão.

Para compressão sem perda a decomposição é calculada através da Transformada de Wavelet Biorthogonal, caso contrário, a Transformada de Wavelet de Daubechies é utilizada. Após a decomposição de DWT, a quantização dos coeficientes é obtida de forma similar à compressão JPEG, em que os coeficientes da transformada são divididos por um escalar, dependendo da taxa de compressão. No JPEG 2000, a quantização é realizada através da equação:

$$q(i,j) = sign(c(i,j))floor\left[\frac{|c(i,j)|}{\Delta}\right] \tag{4.4}$$

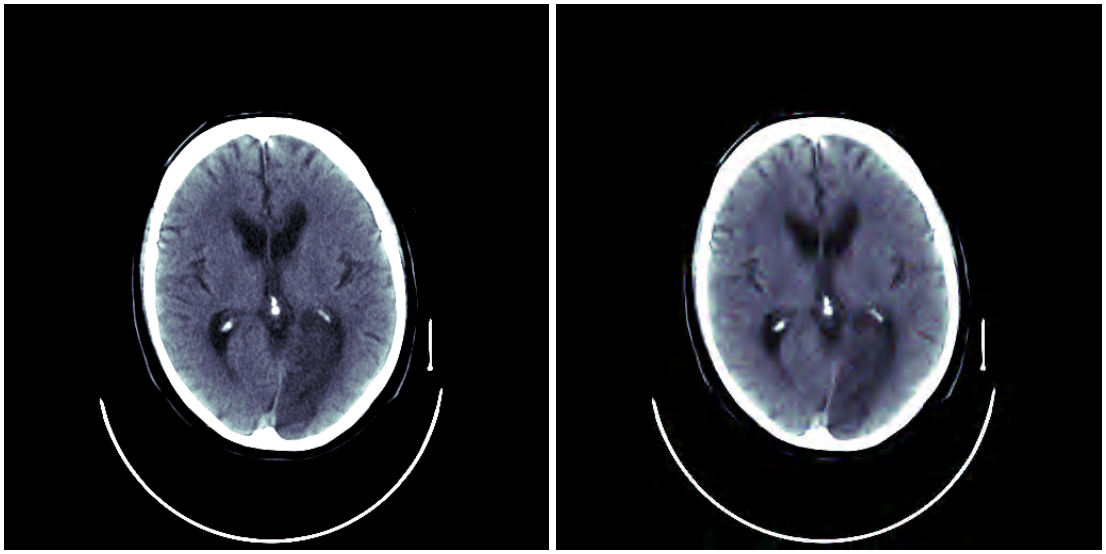
onde $c(i,j)$ é o coeficiente wavelet a ser quantizado, $q(i,j)$ é o valor quantizado e Δ é o tamanho do passo de quantização.

Depois da etapa de quantização, os bits são codificados usando o método de codificação aritmética binária [30]. Na Figura 4.7 é apresentado um exemplo de imagem codificada utilizando o padrão de compressão JPEG 2000.



(a) (b) (c)

Figura 4.6: Sub-bandas da decomposição da imagem através da Transformada de Wavelet Discreta.



(a) b (b)

Figura 4.7: Exemplo de imagem codificada por JPEG 2000: (a) imagem original; (b) imagem reconstruída.

4.2.3 Interpolação

No processo de aquisição da imagem, no caso de *scanners* por *slices*, obtém-se uma sequência de cortes de um órgão com alta resolução, no entanto, o espaçamento entre esse cortes é geralmente muito maior que o tamanho do pixel. Logo, a re-amostragem é necessária para melhorar a qualidade da imagem em dispositivos de exibição.

Além disso, ajustes de zoom ou rotação das imagens após a aquisição são manipulações

frequentes em diagnósticos. Métodos de interpolação são incorporados em sistemas de diagnóstico assistido por computador, cirurgia assistida por computador e arquivamento de imagens e sistemas de comunicação PACS. Em [31], algoritmos simples, como vizinho mais próximo ou interpolação linear, não são recomendados para imagens médicas, em [32], recomenda-se o uso da interpolação bicúbica.

4.2.4 Conjunto de teste

Nesta fase do trabalho, foram usados codificadores JPEG (fator de qualidade = 50, 60, 70 e 100) e JPEG2000 (fator de qualidade = 30, 40, 50 e 100), a partir da compressão dessas imagens em diferente níveis. As imagens foram ampliadas via interpolação bicúbica e janeladas mediante janelamento padrão para osso e tecidos moles. A escolha de janelas para osso e tecidos moles é justificada por serem as opções padrão em análise de imagens de CT de crânio.

Os slices foram originalmente adquiridos em 512 X 512 e 16-bits de profundidade. O resultado foi um conjunto de teste com 48 imagens com 800 $\times$ 800 pixels e 8-bits de profundidade.

4.3 Comitê de ética

Considerando o disposto na Resolução nº 466/12 e na Resolução nº 196/96, do Conselho Nacional de Saúde, do Ministério da Saúde, as pesquisas envolvendo seres humanos devem atender aos fundamentos éticos e científicos pertinentes. Pesquisas em seres humanos são as realizadas em qualquer área do conhecimento que, de modo direto ou indireto, envolvam indivíduos ou coletividades, em sua totalidade ou partes, incluindo o manejo de informações e materiais. Assim, são consideradas como tal pesquisas envolvendo entrevistas, aplicações de questionários, utilização de banco de dados e revisões de prontuários. Logo, a ética da pesquisa implica em que, em qualquer área do conhecimento, tais pesquisas deverão ser apreciadas por um Comitê de Ética em Pesquisa (CEP), e somente deverão ser iniciada interação com seres humanos após análise e aprovação do CEP em questão [33, 34].

Para aquisição das imagens, dados de avaliação e continuidade da pesquisa, o projeto intitulado “Avaliação de qualidade visual aplicada a imagens de tomografia computadorizada” foi submetido ao CEP do Instituto de Ciências da Saúde da Universidade Federal do Pará - ICS/ UFPA, o qual emitiu um parecer consubstanciado definindo que o protocolo encaminhado

dispõe de metodologia e critérios definidos conforme a resolução 466/12 do Conselho Nacional de Saúde do Ministério da Saúde e, portanto, aprovando a continuidade do projeto e consequentemente a coleta de dados. CAAE: 86587118.4.0000.0018.

4.4 Perfil dos avaliadores

Quanto aos avaliadores, a seleção se deu pelos seguintes critérios de inclusão:

- Ser médico de qualquer instituição habilitado ao exercício regular da profissão;

- Concordância em participar da pesquisa, mediante a assinatura do Termo de Consentimento Livre e Esclarecido (TCLE).

4.5 Sistema de Avaliação

O sistema proposto é delineado para coleta de dados de avaliação de qualidade clínica de imagens DICOM codificadas levando em consideração os critérios listados na seção 3.4. No experimento piloto, concebido para automatizar o protocolo, o experimento foi planejado para ser conduzido em um ambiente controlado, então os avaliadores não podem modificar a imagem exibida. Logo, ao invés de disponibilizar todos os *slices* produzidos pelo exame, como no caso de uma avaliação regular, apenas um *slice* é disponibilizado por avaliação onde o especialista analisa na imagem a qualidade dos critérios listados.

O protocolo foi implementado levando em consideração as etapas pelas quais as imagens passam, desde a aquisição até a emissão do laudo diagnóstico. Imediatamente após a aquisição, um técnico ou médico analisa determinados parâmetros da imagem, tais como o grau de precisão dos limites entre estruturas que a compõem, ou o quão regular se mostram seus contornos. A análise desses aspectos teria por finalidade indicar se a imagem apresenta artefatos, ruídos ou distorções. Nesta etapa, a qual o experimento é baseado, o foco é determinar se a imagem em questão apresenta qualidade suficiente para permitir um diagnóstico adequado. Em caso positivo, passa-se então à etapa de laudo, onde serão avaliadas as estruturas anatômicas imageadas. Caso contrário, a fase de aquisição é repetida.

Foram utilizados na implementação a linguagem Python e o aplicativo de código aberto PsychoPy [35, 36], que possui recursos variados para a concepção e execução de experimentos

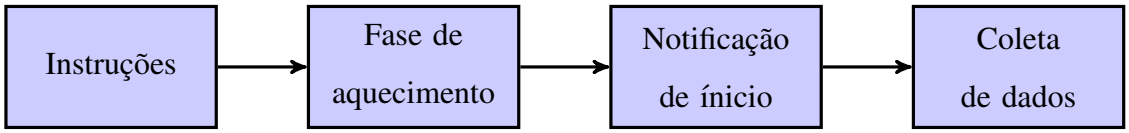


Figura 4.8: Etapas do procedimento para coleta de dados.

em neurociência, psicologia e psicofísica. O experimento pode ser dividido nas etapas ilustradas na Figura 4.8.

Inicialmente são apresentadas orientações a respeito do procedimento e seu objetivo. Em seguida, uma fase de aquecimento familiariza o avaliador com a interface do sistema. Posteriormente, uma notificação é exibida indicando o fim do aquecimento. E por último, a fase de coleta de dados. Nesta, as imagens que compõem o conjunto de teste são apresentadas aleatoriamente, uma de cada vez. Para cada imagem exibida, o avaliador deve responder à pergunta: "Qual o grau de visibilidade das estruturas anatômicas da imagem apresentada?", e pontuar a qualidade da imagem de 1 a 5, de acordo com os critérios listados na Tabela 4.2 [37].

Tabela 4.2: Escala para avaliação de imagens clínicas

Escala	Aparência das Estruturas
1	Qualidade certamente não compromete o diagnóstico
2	Qualidade possivelmente não compromete o diagnóstico
3	Qualidade inconclusiva
4	Qualidade possivelmente compromete o diagnóstico
5	Qualidade certamente compromete o diagnóstico

Apesar de não haver limite de tempo imposto, a duração típica do experimento é em torno de 20 minutos (Fig. 4.9). A qualidade final da imagem (OS, *Opinion Score*) é dada por:

$$OS = \frac{\sum_{s=1}^{S} \sum_{t=1}^{T} G_{abs,s,t}}{S \cdot T}, \tag{4.5}$$

onde $G_{abs,s,t}$ é a escala em relação à aparência de uma estrutura em particular (Tab. 4.2), S é dado pelo número de estruturas apresentadas na imagem (Tab. 3.2) e T representa o número de avaliadores [37]. O range dinâmico do $OS \in [1, 5]$.

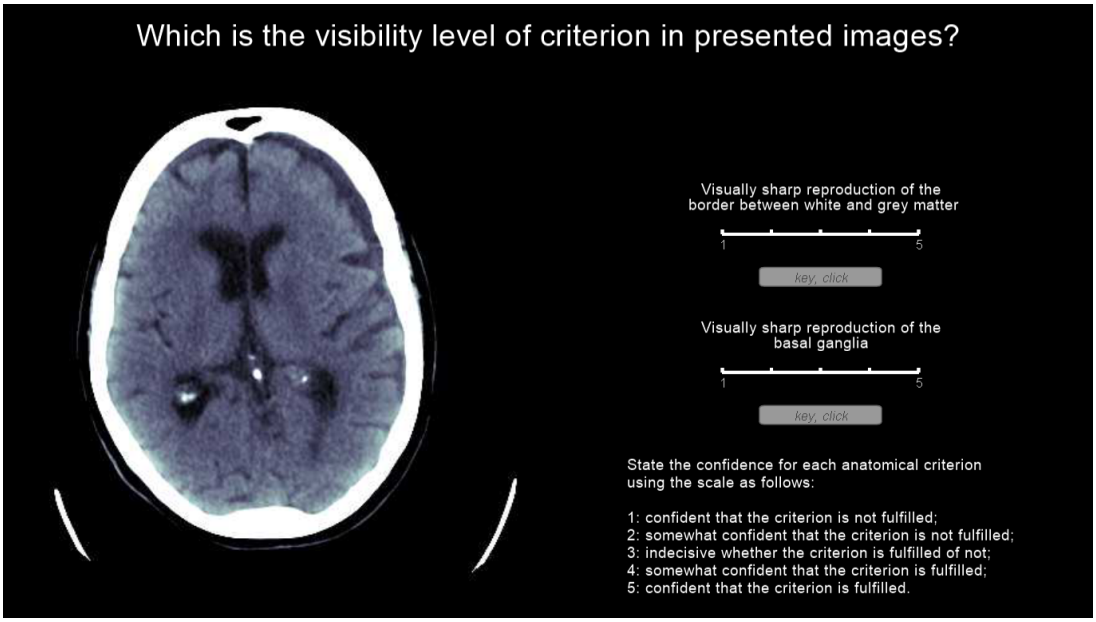


Figura 4.9: Interface gráfica para coleta de dados de imagens.

Capítulo 5

Redes Bayesianas para avaliação de qualidade visual

5.1 Introdução

Uma vez que a maioria das VQMs existentes na literatura foram desenvolvidas a partir de experimentos cujas imagens de teste são convencionais e validadas por ensaios em que os avaliadores são observadores não especializados, transferi-las diretamente para o contexto de imagens médicas pode interferir na performance de tais algoritmos, e com isso não representar bem especificamente a qualidade relacionada ao valor diagnóstico das imagens. Nesse cenário, o desenvolvimento de VQMs específicas para imagens médicas exerce papel fundamental na caracterização de qualidade, por meio da qual algoritmos e sistemas de aquisição e processamento podem ser otimizados.

A percepção visual da imagem, entretanto, não é um processo único, e sim uma tarefa multifacetada, que envolve a combinação de características que engendram a qualidade global da estrutura imageada. Porém, a síntese dessa combinação do ponto de vista perceptual, cognitivo e psicológico não é meramente uma soma dessas características, o que explica a variância entre avaliações obtidas a partir de diferentes observadores. Logo, como as propriedades mensuráveis da imagem e as observações humanas registradas devem ser conectadas? É possível explicar ou simular o fenômeno da qualidade visual? [38].

Nesse sentido, uma abordagem de métrica de qualidade visual específica para imagens médicas é proposta baseada na combinação de VQMs a partir de uma rede Bayesiana, dada a explícita natureza estocástica das avaliações subjetivas. A motivação do uso de redes de decisão

é fundamentada, primeiro pela interpretabilidade, e também porque muito embora a maioria dos modelos para avaliação de qualidade visual seja determinístico, a avaliação subjetiva está sujeita a fatores aleatórios, como fatores ambientais e preferências pessoais do avaliador [39].

Além disso, no caso da necessidade de atualização é possível otimizar os parâmetros da função, ou seja o modelo pode aprender com base em novas experiências apresentadas, e se adaptar a novas perspectivas, caracterizando um paradigma de aprendizado de máquina.

Nessa seção, as VQMs que servem de entrada da rede são detalhadas. Em seguida, a teoria de redes Bayesianas é explanada. Finalmente, é apresentada a estrutura do modelo proposto.

5.2 Métricas de avaliação de qualidade visual

De acordo com *European Guidelines on Quality Criteria for Computed Tomography*, os critérios de qualidade de imagens para diagnóstico são baseados nas especificações anatômicas, com foco na avaliação da área anatômica dos órgãos e do contraste entre diferentes tecidos. Além disso, as degradações mais comuns são distorções orientes de técnicas de compressão, borramentos e ruídos. Dessa forma, as particularidades essenciais para avaliação de qualidade visual para imagens de diagnóstico são os detalhes de borda e textura [6, 40].

Dado que métricas que atuam diretamente a intensidade do pixel da imagem, não correlacionam bem com a qualidade perceptual percebida e modelos baseados no sistema visual humano são limitados em imagens borradas, métricas baseadas nas informações de gradiente da imagem se tornam uma atraente alternativa [40].

Neste cenário, na atual versão do trabalho, das VQMs que foram desenvolvidas ou testadas para o caso específico de avaliação de qualidade clínica e com foco na análise da reprodução das bordas da imagem, destacam-se três VQMs apresentadas a seguir: *Gradient map weighting-gradient diretion similarity* (GMW-GDSIM), *Edge preservation measure* (EPM) e o *Mean structural similarity index* (MSSIM) [40–42].

5.2.1 *Gradient map weighting-gradient diretion similarity*

Em [43–45], métricas baseadas na informação de direção ou magnitude do gradiente da imagem apresentam boa correlação com a avaliação subjetiva. Isso porque a magnitude do gradiente desempenha um papel importante na qualidade da imagem, pois os pixels que representam bordas, textura ou ruído, os quais as diferenças são mais perceptíveis por nossos

olhos, têm maior magnitude. Em contrapartida, a vantagem em métricas baseadas na direção do gradiente é que não são afetadas pelo brilho da imagem. Além disso, a direção do gradiente é propensa a ser mais sensível aos detalhes alterados nas imagens.

Em aplicações práticas, analisar apenas a magnitude do gradiente é inadequado porque as mudanças de intensidade no pixel da imagem podem não afetar criticamente a qualidade global da imagem. Da mesma forma, é impróprio usar apenas a direção para comparação pois não depende do valor dos pixels, com isso, duas imagens que não são semelhantes podem ter a mesma direção de gradiente.

Sendo assim, em [40], uma métrica de qualidade visual é proposta, o GMW-GDSIM, a qual associa as informações de magnitude e direção do gradiente da imagem em um coeficiente, onde a qualidade de pixels de grande magnitude tem prioridade em relação aos pixels que têm baixa magnitude. O foco é priorizar a sensibilidade às mudanças de detalhes nas bordas das imagens. Na Figura 5.1 é apresentado o diagrama de blocos para obtenção do GMW-GDSIM.

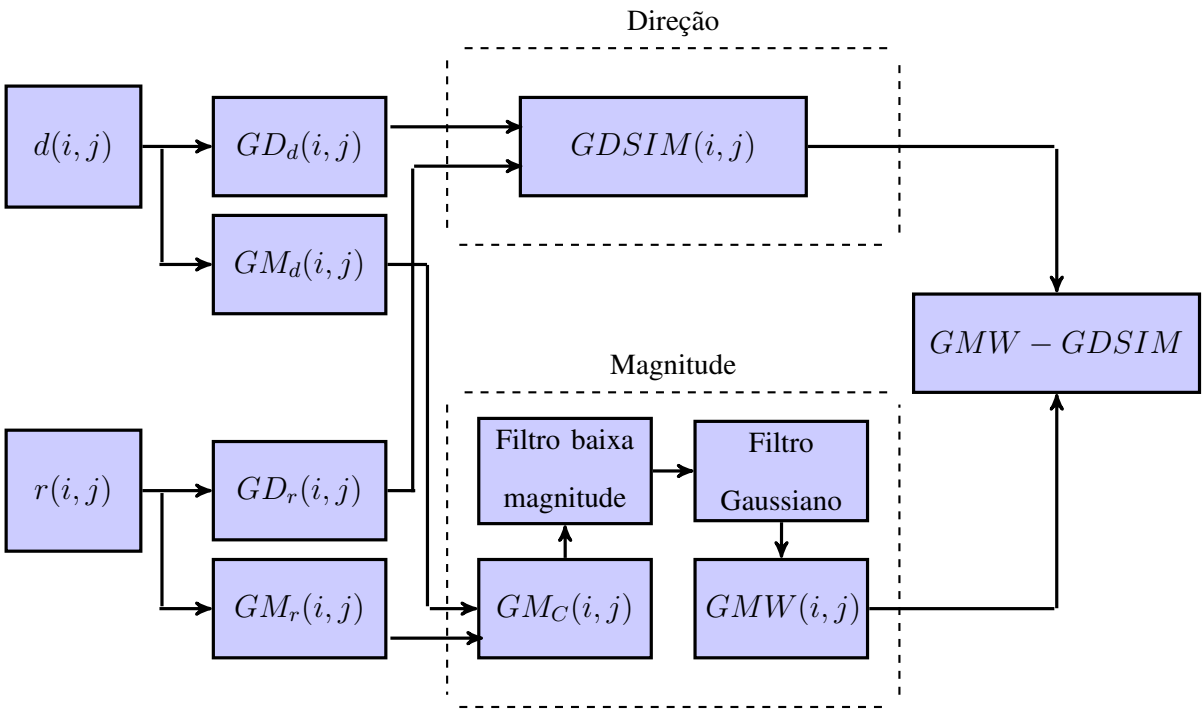


Figura 5.1: Diagrama de fluxo do cálculo do GMW-GDSIM, onde $r(i,j)$, $d(i,j)$ representam a imagem de referência e distorcida, respectivamente; $GD_r(i,j)$ e $GD_d(i,j)$ são dados pela Eq. 5.2; $GM_r(i,j)$ e $GM_d(i,j)$ calculados a partir da Eq. 5.4; $GDSIM(i,j)$ é dado pela Eq. 5.3; $GM_c(i,j)$ pela Eq. 5.5; Filtro baixa magnitude pela Eq. 5.6; Filtro Gaussiano pela Eq. 5.8; $GMW(i,j)$ pela Eq. 5.7 e $GMW - GDSIM$ a partir da Eq. 5.9.

As informações de gradiente de uma imagem nas direções horizontal (G_x) e vertical (G_y) são obtidas através da convolução da imagem com os filtros de Sobel horizontal (S_x) e vertical (S_y), respectivamente, definidos por:

$$S_x = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} \; e \; S_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \tag{5.1}$$

Logo, a direção do gradiente é dada a partir do cálculo da direção do gradiente da imagem computado através de:

$$GD(i,j) = tan^{-1}\left[\frac{G_y(i,j)}{G_x(i,j)}\right]. \tag{5.2}$$

Em seguida, é possível calcular a similaridade entre as direções dos gradientes (GDSIM) dada por:

$$GDSIM(i,j) = \begin{cases} \cos\left(|GD_r(i,j) - GD_d(i,j)|\right) & se\,|GD_r(i,j) - GD_d(i,j)| \leq 180 \\ 360 - \cos\left(|GD_r(i,j) - GD_d(i,j)|\right) & se\,|GD_r(i,j) - GD_d(i,j)| > 180 \end{cases} \tag{5.3}$$

onde as direções de gradiente da imagem de referência e da imagem distorcida são definidas por $GD_r(i,j)$ e $GD_d(i,j)$, respectivamente. Se duas imagens tem a mesma direção de gradiente então $GDSIM(i,j)$ será igual a 1, da mesma forma, se as imagens tiverem direções opostas o coeficiente será igual a 0. Na Figura 5.2(a) é apresentado um exemplo de mapa de similaridade entre as direções dos gradientes de duas imagens (Fig. 4.7(a) e Fig. 4.7(b)).

As informações de magnitude da imagem de referência $(GM_r(i,j))$ e da imagem distorcida $(GM_d(i,j))$ são definidas através de:

$$GM_k(i,j) = \sqrt{{G_x}^2(i,j) + {G_y}^2(i,j)}, k \in \{r,d\}. \tag{5.4}$$

Logo, um mapa de magnitude de gradiente combinado (GM_c) é estruturado usando a média entre a magnitude dos gradientes da imagem de referência e distorcida, definido por:

$$GM_c(i,j) = \frac{GM_r(i,j) + GM_d(i,j)}{2}. \tag{5.5}$$

Por conseguinte, as regiões com magnitude de gradiente extremamente pequena são desprezadas dado que as distorções nessas regiões são raramente notadas. Então, todos os valores de magnitude são atualizados a partir de:

$$GM_c(i,j) = \begin{cases} GM_c(i,j) = 0 & se\, GM_c(i,j) \leq GM_{th} \\ GM_c(i,j) = GMc(i,j) - GM_{th} & se\, GM_c(i,j) > GM_{th} \end{cases} \tag{5.6}$$

onde GM_{th} é um limiar definido em experimentos preliminares e pode variar dada a aplicação. O mapa ponderado de magnitude de gradiente é dado pela convolução do mapa de magnitude do gradiente por um filtro gaussiano, como demonstrado na equação:

$$GMW(x,y) = g(x,y) \cdot GM_c(x,y), \tag{5.7}$$

onde x and y são as distâncias do centro do filtro na direção horizontal e vertical, respectivamente, e $g(x,y)$ é um kernel de suavização gaussiano 2-D com 2 desvios padrões, definido por:

$$g(i,j) = \frac{1}{2\pi\sigma^2} \cdot e^{-\frac{i^2+j^2}{2\sigma^2}}. \tag{5.8}$$

O objetivo da aplicação do filtro gaussiano é aumentar o valor de magnitude do gradiente das regiões de borda, assim como, nas regiões de periferia das bordas. Essa proposta de magnitude ponderada de gradiente realça danos nas bordas causados por compressão JPEG, borramento e ruídos.

Na Figura 5.2(b) é apresentado um exemplo de mapa de similaridade ponderada entre as magnitudes dos gradientes de duas imagens (Fig. 4.7(a) e Fig. 4.7(b)). Finalmente, o mapa $GDSIM(i,j)$ é combinado com o mapa $GMW(i,j)$ definindo o coeficiente GMW-GDSIM, computado por:

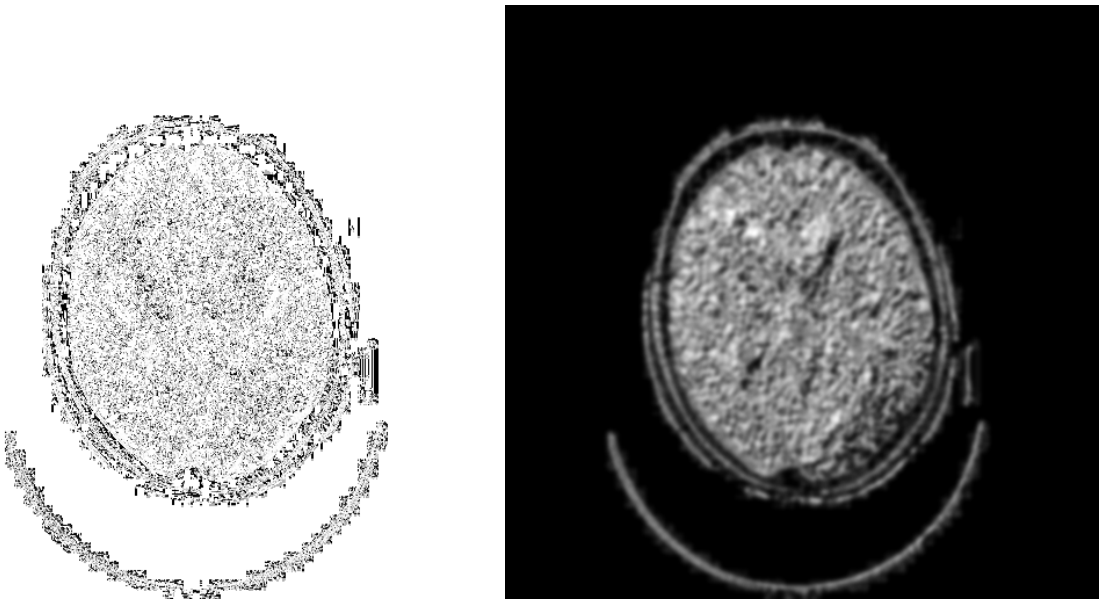
$$\text{GMW-GDSIM} = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M}\{GMW(r,d) \cdot GDSIM(r,d)\}}{\sum_{i=1}^{N}\sum_{j=1}^{M}\{GMW(r,d)\}}, \tag{5.9}$$

onde r e d são as imagens de referência e distorcida, respectivamente.

O range dinâmico do índice de qualidade do GMW-GDSIM $\in [0,1]$, onde 0 indica perda total de informação e 1 indica transmissão ideal com todas as informações da imagem de referência representadas na imagem de teste.

5.2.2 *Edge preservation measure*

Com a finalidade de avaliar a qualidade global e local da imagem, o EPM baseia-se na associação das informações do gradiente com as informações da qualidade das borda da imagem. A ideia é que canais de transmissão ou compressão que transferem informações de gradi-



(a) (b)

Figura 5.2: (a) Similaridade entre as direções de gradiente das imagens de referência e distorcida (GDSIM); (b) Similaridade ponderada entre as magnitudes dos gradientes das imagens de referência e distorcida (GMW).

ente da imagem de origem para a imagem distorcida com mais precisão têm melhor desempenho.

Para isso, parâmetros de borda são extraídos da imagem de referencia e imagem distorcida. Em seguida, a preservação dessas bordas é avaliada, a qual estima quão bem as bordas da imagem distorcida representam as bordas da imagem de referência. Por fim, a qualidade perceptual de cada borda pondera as informações de gradiente. Na Figura 5.3 é desenhado o diagrama de fluxo do cálculo do EPM.

As bordas da imagem são perfeitamente representadas se e somente não houverem mudanças na magnitude ou na direção do gradiente dos pixels correspondentes na imagem de referência. Logo, a estimação da qualidade das bordas é baseada na variação da magnitude $\Delta_g(i,j)$ e da variação da direção $\Delta_\alpha(i,j)$ - entre a imagem de referência, $r(i,j)$, e imagem distorcida, $d(i,j)$. $\Delta_g(i,j)$ é definida por:

$$\Delta_g(i,j) = \begin{cases} \frac{GM_d(i,j)+C}{GM_r(i,j)+C} & GM_r(i,j) > GM_d(i,j) \\ \frac{GM_r(i,j)+C}{GM_d(i,j)+C} & GM_r(i,j) \leq GM_d(i,j) \end{cases} \tag{5.10}$$

onde C é uma constante incluída para evitar a instabilidade quando o denominador está muito

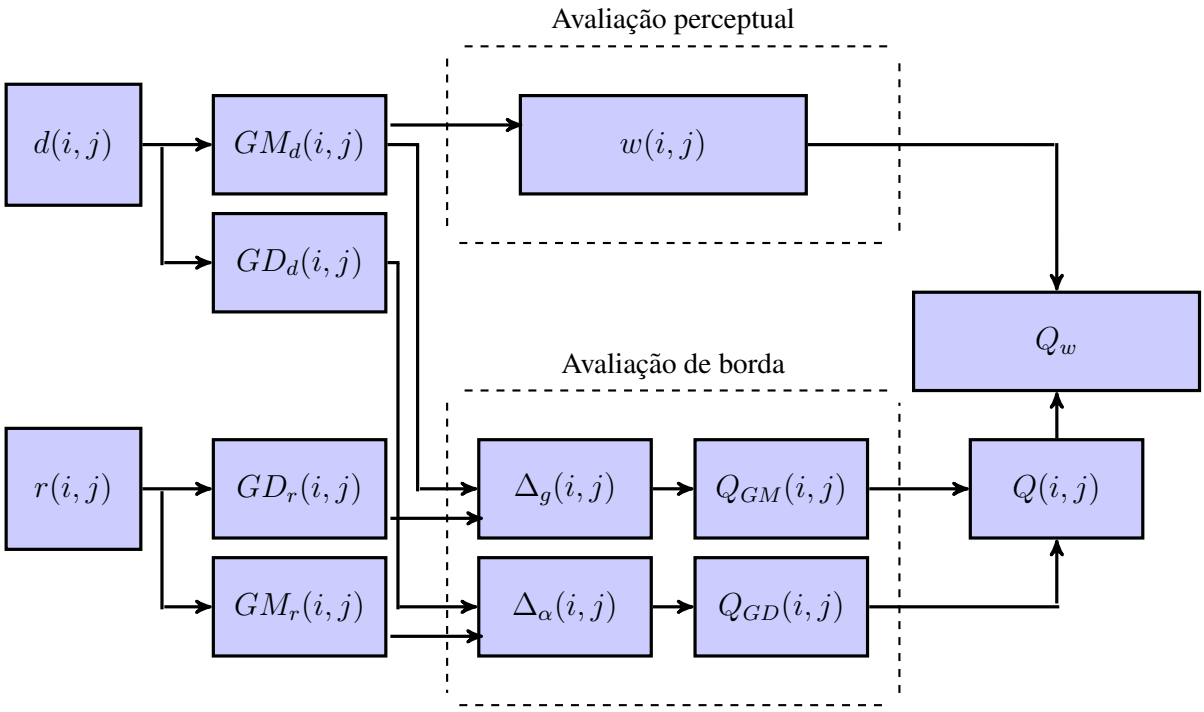


Figura 5.3: Diagrama de blocos do EPM, onde $r(i,j)$, $d(i,j)$ representam a imagem de referência e distorcida, respectivamente; $GD_r(i,j)$ e $GD_d(i,j)$ são obtidos através da Eq. 5.2; $GM_r(i,j)$ e $GM_d(i,j)$ são dados pela Eq. 5.4; $\Delta_g(i,j)$ e $\Delta_\alpha(i,j)$ são computados pelas Eq. 5.10 e Eq. 5.11, respectivamente; $Q_{GM}(i,j)$ e $Q_{GD}(i,j)$ pela Eq. 5.12; $w(i,j)$ dado pela Eq. 5.14 e $Q(i,j)$ e Q_w pelas Eq. 5.13 e Eq. 5.15, respectivamente.

perto de zero. $GM_r(i,j)$ e $GM_d(i,j)$ são definidos pelas Equações 5.4 e 5.2, respectivamente. Como a direção é uma variável cíclica, valores nos dois extremos ($-\pi$ e π) são equivalente. $\Delta_\alpha(i,j)$, portanto, é uma distância relativa normalizada por:

$$\Delta_\alpha(i,j) = \pi^{-1} \left| |GD_r(i,j) - GD_d(i,j)| - \pi \right| . \tag{5.11}$$

Δ_g e Δ_α descrevem mudanças lineares nos parâmetros de informação visual. A modelagem da perda de informação perceptual, entretanto, deve levar em consideração a natureza não linear do HVS. Nesse sentido, as relações não lineares da perda perceptual com a magnitude e direção do gradiente são modeladas por não linearidades sigmodais, dadas por:

$$Q_i(i,j) = \frac{\Gamma_i}{1 + \exp(k_i(\Delta_i(i,j) - \sigma_i))}, i \in \{GM, GD\}. \tag{5.12}$$

Em [41], os parâmetros foram inicializados como $[k_{GM}, \sigma_{GM}, k_{GD}, \sigma_{GD}] = [-11, 0.7, -24, 0.8]$ e o fator Γ_i é otimizado tal que, para $Q_i = 1$, $\Delta_i = 1$.

Finalmente, com o objetivo de aproximar a perda de informação perceptual em relação às mudanças nos parâmetros de gradiente, as avaliações não lineares de preservação da magnitude e direção entre as imagens, $Q_{GM}(i,j)$ e $Q_{GD}(i,j)$, são combinadas em uma única avaliação de qualidade de preservação, $Q(i,j)$, dada por:

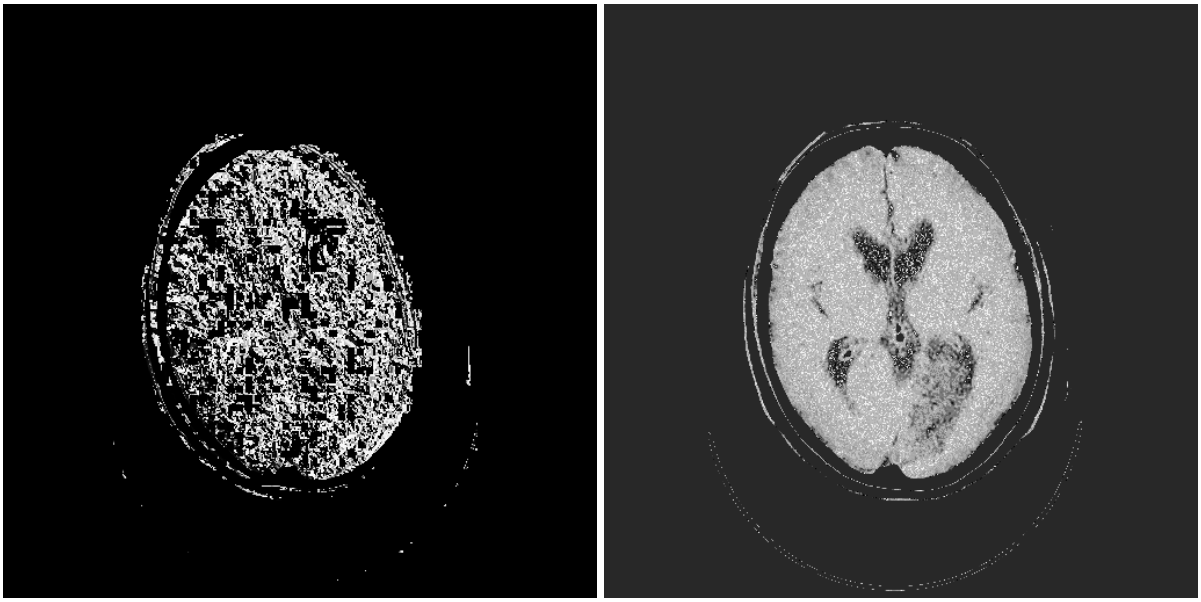
$$Q(i,j) = \sqrt{Q_{GM}(i,j) \cdot Q_{GD}(i,j)}. \tag{5.13}$$

Na Figura 5.4(a) é apresentado um exemplo de mapa de avaliação de preservação de qualidade de duas imagens (Fig. 4.7(a) e Fig. 4.7(b)).

A distribuição da importância perceptual na cena depende de muitos fatores como contraste, orientação, localização, contexto, entre outros [46]. Nessa abordagem, a importância perceptual para cada pixel é representada por $w(i,j)$, o qual é uma função de pesos consistente com a natureza não linear do HVS, dado por:

$$w(i,j) = -\log_2 P(GM_r(i,j)) \tag{5.14}$$

em que $P(GM_r(i,j))$ é a função de densidade de probabilidade da magnitude do gradiente da imagem de referência. Na Figura 5.4(b) é apresentado um exemplo de mapa de avaliação de importância perceptual de duas imagens (Fig. 4.7(a) e Fig. 4.7(b)).



(a) (b)

Figura 5.4: (a) Avaliação de qualidade de preservação (Q); (b) Avaliação de importância perceptual (w).

O índice global de qualidade é dado pela soma normalizada das preservações das bordas ponderadas por sua importância perceptual, definida como:

$$Q_w = \frac{\sum_{i=1}^{i}\sum_{j=1}^{j} Q(i,j)w(i,j)}{\sum_{i=1}^{i}\sum_{j=1}^{j} w(i,j)} \tag{5.15}$$

O range dinâmico do índice de qualidade $Q_w \in [0,1]$, onde 0 indica perda total de informação e 1 indica transmissão ideal com todas as informações da imagem de referência representadas na imagem de teste.

5.2.3 *Mean structural similarity index*

Como imagens naturais são altamente estruturadas, a filosofia por trás do Similaridade Estrutural (SSIM, *structural similarity based image*) é que as degradações da imagem são baseadas nas mudanças das informações estruturais, portanto a avaliação da informação estrutural pode ser uma boa aproximação da distorção da imagem percebida.

A informação da estrutural é definida como os atributos que representam a estrutura dos objetos na cena e independem da iluminação e contraste. Sendo assim, a SSIM se baseia na comparação da luminância, contraste e estrutura na imagem [47]. As três componentes combinadas resultam em:

$$S(x,y) = f(l(x,y), c(x,y), s(x,y)). \tag{5.16}$$

Assim, a componente relacionada a luminância é definida como:

$$l(x,y) = \frac{2\mu_x\mu_y + C_1}{{\mu_x}^2 + {\mu_y}^2 + C_1}. \tag{5.17}$$

Por sua vez, a componente relacionada ao contraste é dada por:

$$c(x,y) = \frac{2\sigma_x\sigma_y + C_2}{{\sigma_x}^2 + {\sigma_y}^2 + C_2}. \tag{5.18}$$

Finalmente, a componente relacionada a estrutura é definida como:

$$s(x,y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}. \tag{5.19}$$

As três funções combinadas resultam em:

$$SSIM(x,y) = [l(x,y)]^\alpha \cdot [c(x,y)]^\beta \cdot [s(x,y)]^\gamma. \tag{5.20}$$

Em [47], considerando $\alpha = \beta = \gamma = 1$ e $C_3 = C_2/2$ para simplificar a expressão, obtém-se:

$$SSIM(i,j) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_x\sigma_y + C_2)}{({\mu_x}^2 + {\mu_y}^2 + C_1)({\sigma_x}^2 + {\sigma_y}^2 + C_2)}, \tag{5.21}$$

onde μ_x denota a média de x, σ_x denota o desvio padrão de x, μ_y representa a media de y, σ_y representa o desvio padrão de y e $\sigma_x y$ a covariância entre x e y. Os parâmetros C_1, C_2 e C_3 são constantes arbitrárias e α, β e γ são parâmetros utilizados para relativizar a importância das três componentes. Na Figura 5.5 é apresentado um exemplo de mapa de avaliação de importância perceptual de duas imagens (Fig. 4.7(a) e Fig. 4.7(b)).

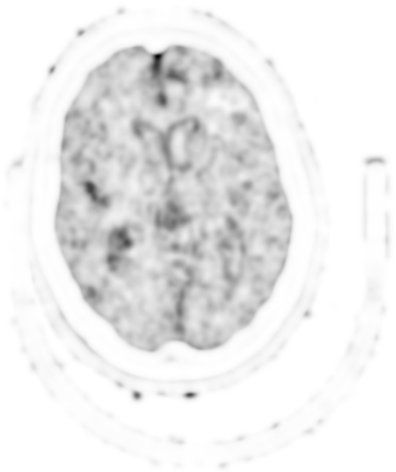


Figura 5.5: Mapa de similaridade estrutural (SSIM).

Para avaliação de qualidade visual, é recomendado aplicar o SSIM localmente. Na prática, o SSIM é aplicado em janelas de $8X8$ pixels que percorrem a imagem pixel a pixel, onde obtém-se o mapa de similaridades, também conhecido como MSSIM. Logo, o MSSIM é dado por:

$$MSSIM = \frac{1}{MN} \sum_{n=1}^{N} SSIM(m, n), \tag{5.22}$$

onde M e N representam as dimensões da imagem e $SSIM(m, n)$ representa o valor de SSIM referente ao elemento (m, n) do mapa de similaridade. O range dinâmico do índice de qualidade $MSSIM \in [0, 1]$, onde 0 indica perda total de informação e 1 indica transmissão ideal com todas as informações da imagem de referência representadas na imagem de teste.

5.3 Redes Bayesianas

Redes Bayesianas são grafos acíclicos orientados capazes de modelar conhecimentos sujeitos à aleatoriedade, sendo a adaptabilidade uma das suas principais características. O objetivo

de uma inferência Bayesiana é produzir um agente mutável que tenha desempenho semelhante às decisões humanas em tarefas de complexidade intelectual, que pode se adaptar a ambientes estocásticos e instáveis, e lidar de maneira sensata com variadas fontes de incertezas [48].

A estrutura de uma rede de Bayes é composta por nós de entrada que são as variáveis causais do evento que será modelado e nós de saída que são dados baseados na decisão de especialistas de casos passados observados. A ideia básica é treinar uma estrutura de árvore de decisão, modelando uma rede multi-conectada que tem por finalidade retornar a probabilidade de cada possibilidade dos nós de saída [38].

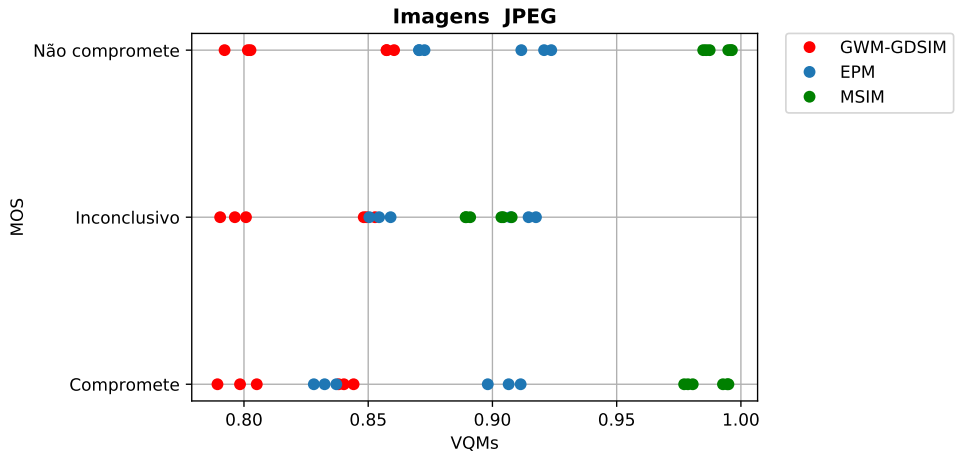
No treinamento de uma rede, a primeira etapa, é identificar as variáveis relevantes que irão compor os nós de entrada da rede, e dispor de uma base de dados fornecidos por especialistas ou proveniente de dados de campo coletados. Dada a diferença perceptual entre imagens codificadas por algoritmos de compressão distintos, a tarefa foi dividida em relação ao algoritmo de compressão da imagem. Com isso, duas redes foram modeladas: uma para imagens codificadas através de JPEG e outra para imagens com padrão de compressão JPEG 2000. No modelo proposto, as entradas são VQMs e a saída é a qualidade percebida da imagem que se quer avaliar. Na Tabela 5.1, é apresentada a associação entre as possíveis saídas das redes de acordo com a qualidade perceptual percebida (Tab. 4.2).

Tabela 5.1: Saída da rede Bayesiana em relação às aparências das estruturas

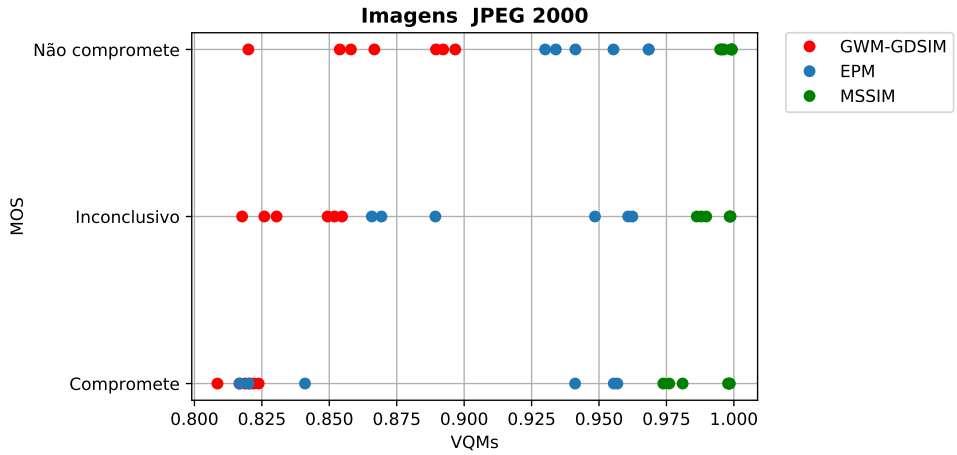
Saída da rede	**Aparência das Estruturas**
Não compromete	Qualidade certamente não compromete o diagnóstico
	Qualidade possivelmente não compromete o diagnóstico
Inconclusivo	Qualidade inconclusiva
Compromete	Qualidade possivelmente compromete o diagnóstico
	Qualidade certamente compromete o diagnóstico

Na fase de treinamento, as métricas que servem como entrada da rede são calculadas e discretizadas para estimação das probabilidades a priori através da frequência relativa. A probabilidade da qualidade também foi estimada a partir da frequência relativa. Como uma ilustração visual da relação entre os dados subjetivos e as VQMs, o gráfico de dispersão entre a média das avaliações subjetivas (MOS, *mean opinion score*) e a predição das métricas é mostrado na Figura 5.6.

Em seguida, a rede calcula as probabilidades a posteriori das saídas, dada cada



(a)



(b)

Figura 5.6: Gráfico de dispersão entre o MOS e VQMs. (a) Imagens comprimidas por JPEG; (b) Imagens comprimidas por JPEG 2000.

configuração dos seus nós de entrada, através do teorema de Bayes definido por:

$$P(y|x_1, x_2, x_3) = \frac{P(y)P(x_1, x_2, x_3|y)}{P(x_1, x_2, x_3)}, \tag{5.23}$$

onde, $P(y)$ é a função de probabilidade das classes de saída e $P(x_1, x_2, x_3)$ é a função de probabilidade conjunta das variáveis aleatórias nós de entrada: GMW-GDSIM, EPM, MSSIM, respectivamente.

Assumindo que as entradas são eventos independentes, temos que:

$$P(x_i|y, x_1, \cdots, x_{i-1}; x_{i+1}, \cdots, x_n) = P(x_i|y), \tag{5.24}$$

então para cada i, a relação é simplificada por:

$$P(y|x_1, x_2, x_3) = \frac{P(y)\prod_{i=1}^{3} P(x_i|y)}{P(x_1, x_2, x_3)}. \tag{5.25}$$

Como $P(x_1, x_2, x_3)$ é constante, a regra de classificação é dada por:

$$\hat{y} = \arg\max_{x} P(y) \prod_{i=1}^{n} P(x_i|y). \tag{5.26}$$

Logo, a classe com maior probabilidade a posteriori é escolhida como saída da rede [49, 50]. O modelo proposto foi implementado em linguagem Python, visando, sobretudo, rapidez na fase de desenvolvimento, facilidade na manutenção do código fonte e portabilidade. Adicionalmente, para implementação das redes foi utilizado o pacote Scikit-Learn, o qual consiste em uma ferramenta destinada à mineração e análise de dados [51].

Capítulo 6

Resultados

6.1 Introdução

Tão importante quanto dispor de VQMs, é fundamental ser capaz de avaliar objetivamente o desempenho dessas métricas nos contextos de aplicação. O procedimento padrão para caracterização do desempenho de métricas de qualidade visual, consolidado atualmente em documentos do VQEG (*Video Quality Experts Group*) e recomendações do ITU, considera três aspectos: acurácia, monotonicidade e consistência, em termos da relação entre os dados MOS e VQMs [5, 52].

Tais atributos são métricas de suporte, que se relacionam com a capacidade de uma VQM prever a avaliação subjetiva em relação a uma sequência de imagens e servem de guia para a seleção de VQMs apropriadas. No caso, o objetivo é encontrar o grau em que uma medida de qualidade pode discriminar os artefatos de codificação e traduzi-la em uma pontuação [53–55].

Para fins de validação, essa seção aborda a avaliação da performance dos modelos propostos, considerando as diretrizes sugeridas pelo relatório VQEG associadas às recomendações em [56], onde é estabelecida uma metodologia estatística para análise de dados categóricos. O propósito é avaliar e, por conseguinte, comparar o desempenho das VQMs - GMW-GDSIM, EPM, MSSIM e as Redes Bayesianas propostas - no contexto de imagens para diagnóstico.

6.2 Experimento de validação

Para examinar a sensibilidade dos modelos de avaliação de qualidade visual implementados, o comportamento estatístico dos algoritmos foi mapeado utilizando a técnica de validação

cruzada *leave one out* e representado na Tabela 6.1. Como recomendação do relatório do VQEG, as predições de acurácia e monotonicidade são apresentadas através do coeficiente de correlação de Spearman e coeficiente de correlação de Pearson, respectivamente. Adicionalmente são apresentadas a análise de variância (ANOVA, *Analysis of Variance*) e a acurácia frente ao conjunto de imagens de teste em relação ao MOS [53, 57].

O objetivo é estimar a precisão e medir a dispersão entre as previsões das VQMs e o MOS dos espectadores. Além disso, prever o quão monotônica é a VQM em relação ao MOS, uma função monotônica é a qual a derivada não muda de sinal. Uma limitação é que no caso do GMW-GDSIM, EPM e MSSIM, os quais apresentam faixas dinâmicas contínuas, não é possível o calcular a acurácia em relação ao MOS.

Tabela 6.1: Resultados da avaliação de desempenho das VQM's em relação aos mapeamentos estatísticos.

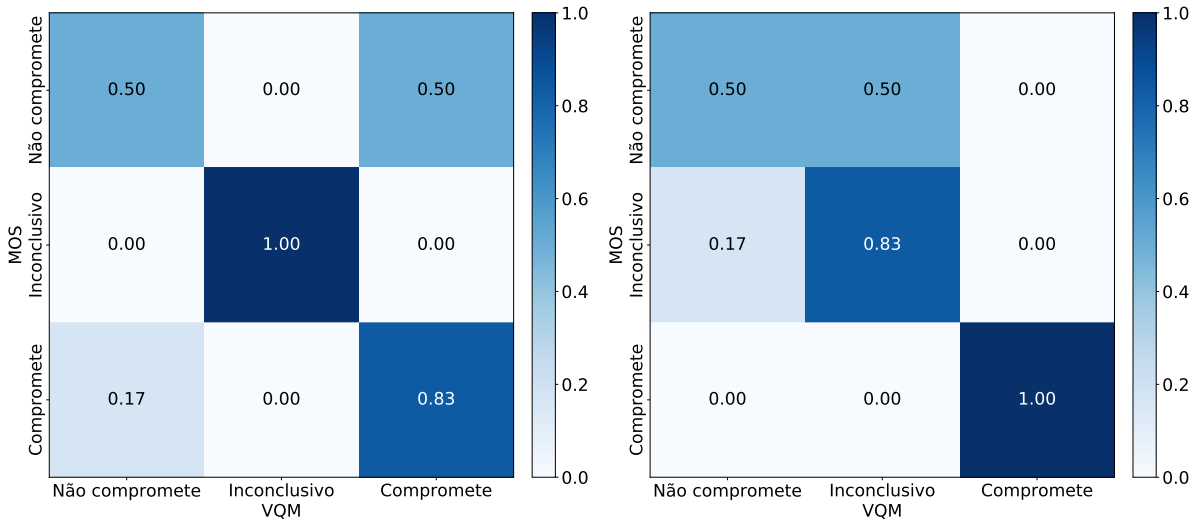
Imagens JPEG

VQM	**Spearman**	**Pearson**	**ANOVA (F score)**	**Acurácia**
GMW-GDSIM	0.142080238157	0.275411316377	35.2513946087	-
EPM	0.328441399114	0.432789211450	31.7896502232	-
MSSIM	0.041772562898	0.209837193430	27.5921907424	-
Rede	0.377216768077	0.414173829464	0.120418848167	0.791666666666

Imagens JPEG 2000

VQM	**Spearman**	**Pearson**	**ANOVA (F score)**	**Acurácia**
GMW-GDSIM	0.862259694785	0.878693247490	34.0207568237	-
EPM	0.486154001423	0.393444737682	29.7337163932	-
MSSIM	0.491317922276	0.459018860629	25.8689607850	-
Rede	0.876459821202	0.931694990624	0.128491620111	0.833333333333

Uma vez que o conjunto de dados esteja desbalanceado, apenas a acurácia e monotonicidade não são suficientes para validar o desempenho de um classificador, isso porque os resultados inferidos podem representar dados tendenciosos. A predição de consistência, por sua vez, relaciona-se com a capacidade da medida de qualidade objetiva fornecer previsões consistentemente precisas para todos os tipos de imagens e não falhar para um subconjunto de imagens. Nesse sentido, na Figura 6.1 são apresentadas as matrizes de confusão das saídas das redes Bayesianas em relação ao MOS.



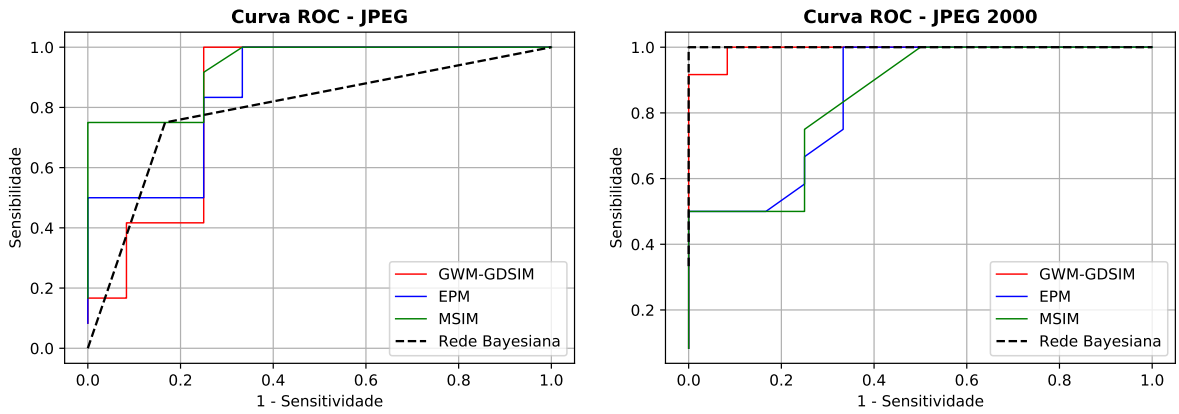
(a) (b)

Figura 6.1: Matriz de confusão da Rede Bayesiana em relação ao MOS. (a) Imagens comprimidas por JPEG; (b) Imagens comprimidas por JPEG 2000.

Outra predição de consistência muito utilizada, a curva *Receiver Operating Characteristic* (ROC) é uma ferramenta que estima quão o modelo é capaz de distinguir as classes precisamente, quanto maior a área abaixo da curva (AUC, *Area Under Curve*) melhor é o desempenho do modelo. A técnica relaciona a sensibilidade (SE), ou seja, a taxa de amostras que o classificador categoriza como positivas e de fato são positivas, à sensitividade (SP) dos dados, que consiste na proporção de amostras que o classificador categoriza como negativas que de fato são negativas.

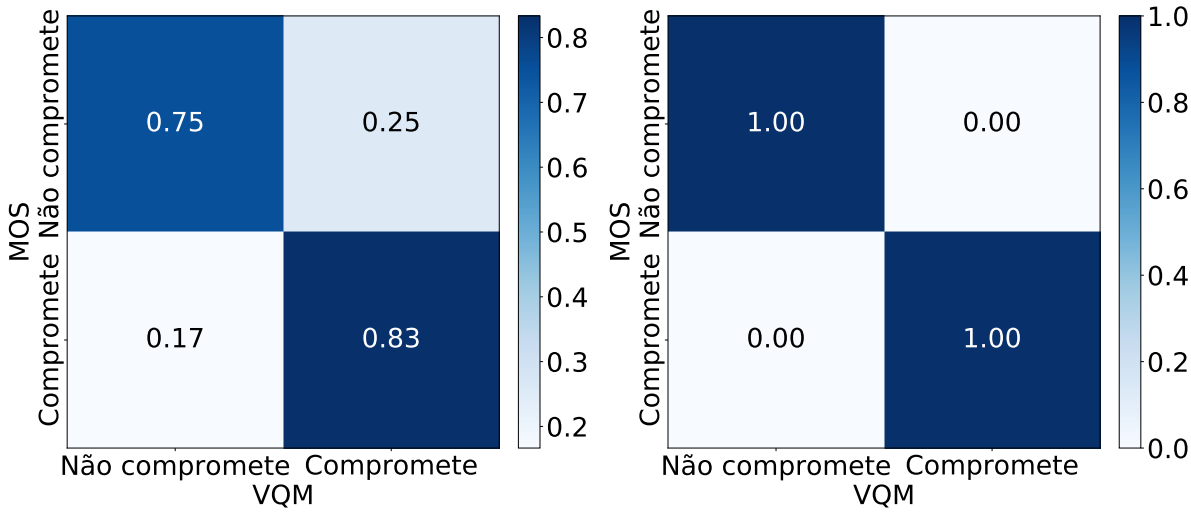
Como a curva ROC só se aplica à classificadores binários, em [58] uma extensão da curva ROC é empregada à tarefas com múltiplas classes, aplicando a abordagem *one vs all*. A técnica consiste em um mapeamento das classes, onde uma delas é definida como positiva e todas as outras classes são unificadas em outra única classe. Na Figura 6.2 é apresentada a curva ROC das métricas GMW-GDSIM, EPM, MSSIM e Rede Bayesiana em relação ao MOS, utilizando *one vs all*. A classe "Não compromete" continuou definida como "Não compromete", enquanto que, "Inconclusivo" e "Compromete" foram unificadas para "Compromete".

Na Figura 6.3 são apresentadas as matrizes de contingência das redes Bayesianas em relação ao MOS utilizando a abordagem *one vs all*.



(a) (b)

Figura 6.2: Curva ROC de VQMs em relação ao MOS. (a) Imagens codificadas por JPEG; (b) Imagens codificadas por JPEG 2000.



(a) (b)

Figura 6.3: Matriz de contingência da rede Bayesiana em relação ao MOS. (a) Imagens comprimidas por JPEG; (b) Imagens comprimidas por JPEG 2000.

Em [20], o teste Kolmogorov–Smirnov é utilizado para estimar a performance de VQMs, onde o desempenho é caracterizado pela distância de Kolmogorov-Smirnov (KS) dada por:

$$KS(P_1, P_2) = \sup_s |P_1(s) - P_2(s)|, \tag{6.1}$$

em que $P_1(s)$ é a função de distribuição acumulada do radiologista aceitar a imagem, e $P_2(s)$

representa a função de distribuição acumulada do radiologista não aceitar a imagem, representadas pelas Equações 6.2 e 6.3, respectivamente.

$$P1(s) = 1 - SE \tag{6.2}$$

$$P2(s) = 1 - SP \tag{6.3}$$

Quanto maior a diferença entre as duas distribuições melhor é a performance do modelo, entretanto, para calcular a distância de Kolmogorov-Sminov de uma VQM que têm a faixa dinâmica é necessário arbitrar um limiar s'. Nesse caso, s' é escolhido tal que obtenha o máximo KS. Uma situação genérica é mostrada na Figura 6.4.

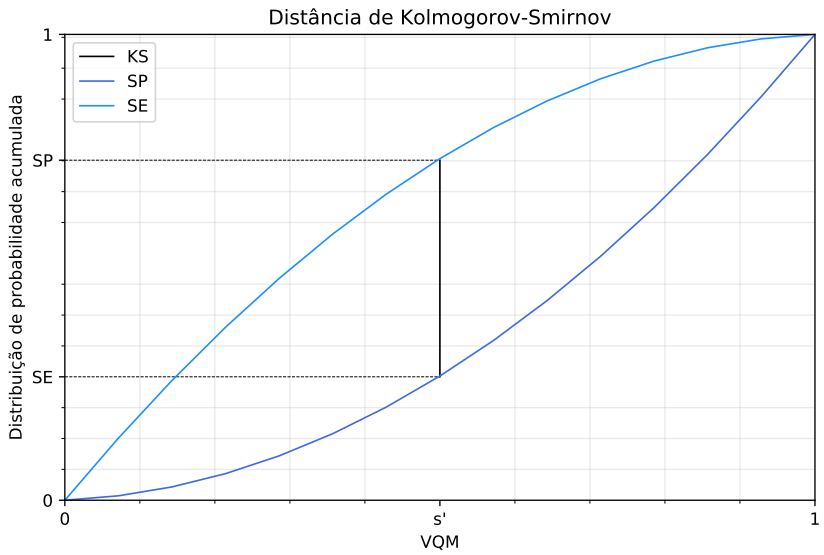


Figura 6.4: Gráfico do teste de Kolmogorov-Smirnov, onde SE é a sensibilidade e SP representa a sensitividade dos dados de MOS em relação à VQM; KS é dado pela Equação 6.1.

Nas Figuras 6.5 e 6.6 são apresentados os gráficos do teste de Kolmogorov-Smirnov em relação às VQMs para as imagens comprimidas pelos algoritmos JPEG e JPEG 2000, respectivamente.

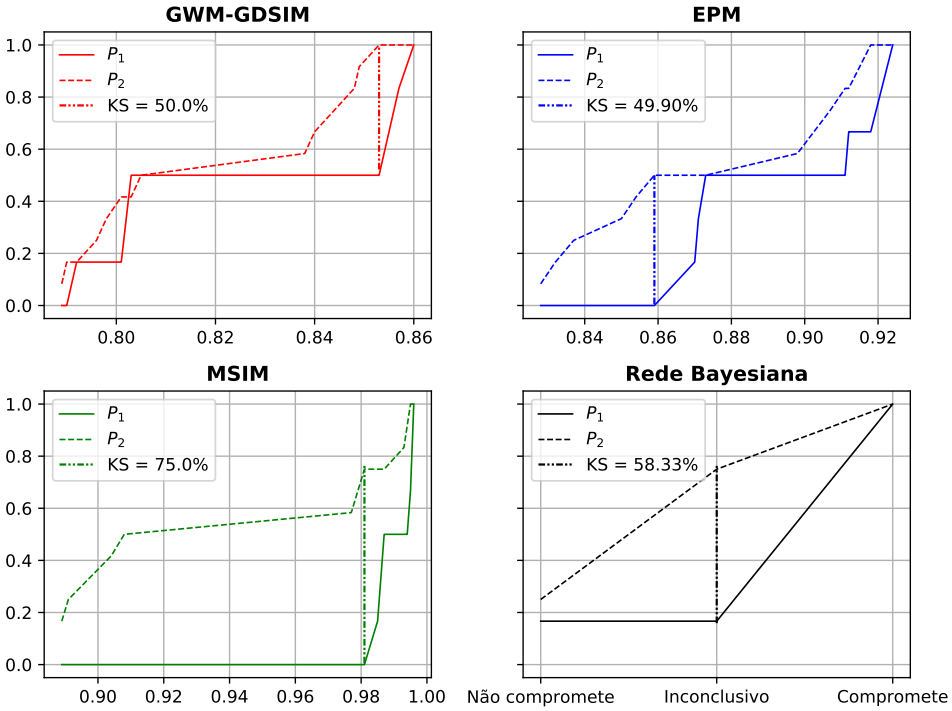


Figura 6.5: Gráfico do teste de Kolmogorov-Smirnov das VQMs em relação às imagens JPEG.

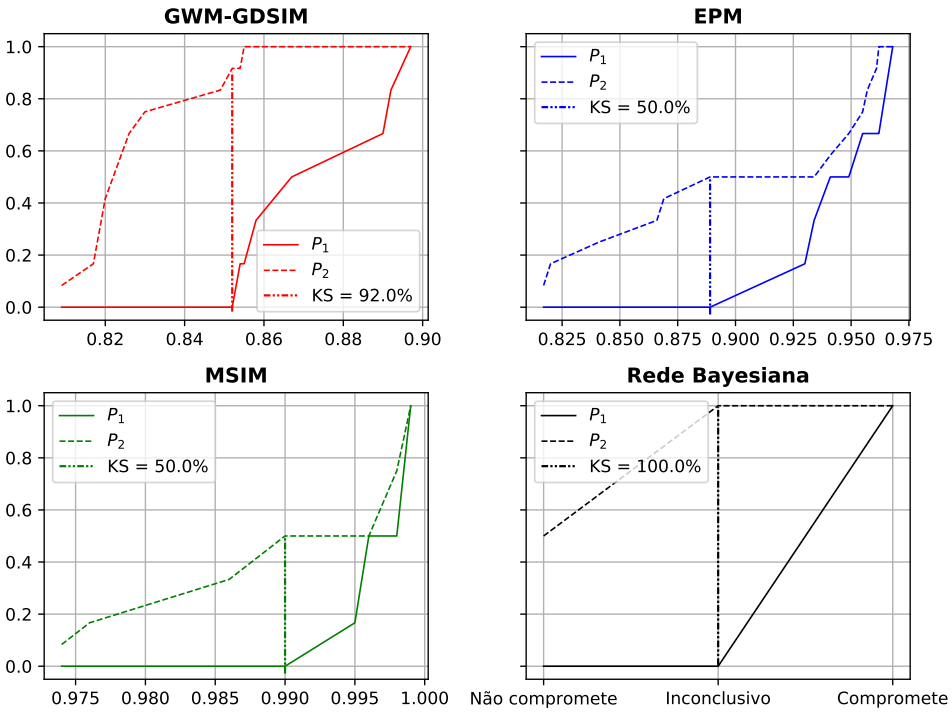


Figura 6.6: Gráfico do teste de Kolmogorov-Smirnov das VQMs em relação às imagens JPEG 2000.

Na Tabela 6.2 a distância de KS; o limiar s' arbitrado e a área abaixo da curva ROC, AUC para cada VQM em relação ao MOS são exibidos para as imagens codificadas por JPEG e JPEG 2000.

Tabela 6.2: Teste de Kolmogorov-Smirnov.

Imagens JPEG			
VQM	**KS**	**s'**	**AUC**
GMW-GDSIM	50.0%	0.859	0.67
EPM	49.9%	0.853	0.72
MSSIM	75.0%	0.981	0.87
Rede	58.33%	-	0.79
Imagens JPEG 2000			
VQM	**KS**	**s'**	**AUC**
GMW-GDSIM	92.0%	0.852	0.99
EPM	50.0%	0.889	0.71
MSSIM	50.0%	0.99	0.69
Rede	100%	-	1.0

Como mostrado na Figura 6.7, a definição de s' permite discretizar as VQMs com faixas dinâmicas contínuas e com isso calcular as tabelas de contingência pertinentes. Logo, nas Figuras 6.8 e 6.9 são apresentadas as matrizes de contingência das VQMs após discretização em relação ao MOS para as imagens JPEG e JPEG 2000, respectivamente.

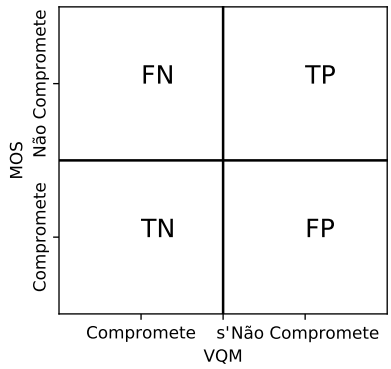
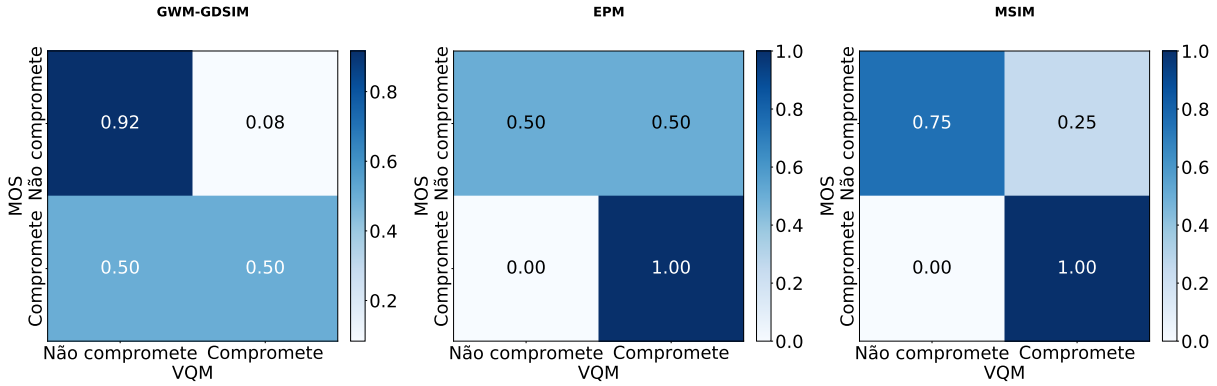
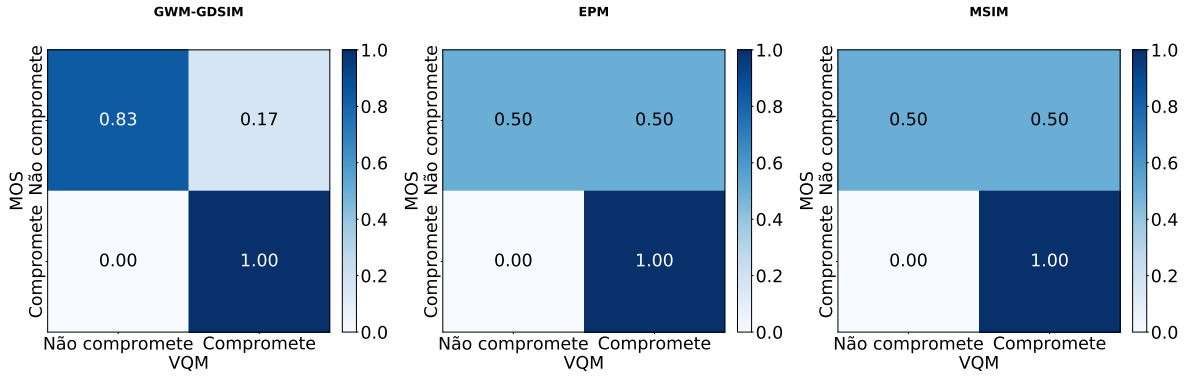


Figura 6.7: Matriz de contingência da VQM em relação ao MOS. Baseada em [20]



(a) (b) (c)

Figura 6.8: Matrizes de contingência das VQMs em relação ao MOS para as imagens JPEG.



(a) (b) (c)

Figura 6.9: Matrizes de contingência das VQMs em relação ao MOS para as imagens JPEG.

6.3 Discussão

Os resultados apresentados atestam que a abordagem proposta - rede Bayesiana para avaliação de qualidade visual - apresenta boa correlação entre a qualidade predita e a subjetiva. Os modelos podem ser utilizados em testes e ajustes de sistemas de processamento de imagens, sendo, contudo, limitados àquelas situações em que uma versão da imagem degradada isenta de ruído seja conhecida.

Capítulo 7

Considerações Finais

7.1 Conclusões

Neste trabalho, foram apresentados um protocolo de avaliação de qualidade clínica, assim como, um sistema experimental para coleta de dados de avaliação de qualidade subjetiva de imagens médicas codificadas. A metodologia foi destinada à construção de um banco de imagens avaliadas que serve de suporte no desenvolvimento de novas métricas de qualidade visual para o caso específico de imagens médicas. O conceito geral é aplicado para o caso de imagens de tomografia dcomputadorizada de crânio e segue as recomendações do *European Guidelines on Quality Criteria for Computed Tomography*.

Adicionalmente, é apresentada uma abordagem de avaliação de qualidade visual para imagens médicas codificadas utilizando redes Bayesianas. Assim como, sua caracterização de desempenho, de acordo com as recomendações do *International Telecommunication Union*.

Os sistemas se complementam, à medida que a coleta de dados subjetivos de imagens médicas, estes ainda incipientes na literatura, viabilizou a confecção de métricas de qualidade visual específica para tais imagens, assim como, a avaliação do desempenho das mesmas. Os resultados mostram que as métricas propostas correlacionam bem com os dados subjetivos. Tal abordagem constitui instrumentos de caracterização de qualidade por meio dos quais algoritmos e sistemas de aquisição e processamento podem ser otimizados. O projeto é mantido em repositório online[1], onde os interessados poderão obter a versão estável mais recente, contribuir com sugestões ou mesmo colaborar no desenvolvimento.

[1]https:/github.com/luaffjk/AQV-RB

7.2 Trabalhos Futuros

A continuação natural deste trabalho consiste na extensão para outros tipos de imagens de CT e modalidades de sistemas de imagens médicas. No mais, é delineada a ampliação do conjunto de dados subjetivos com mais avaliadores e imagens. Além disso, o estudo de reconhecer estruturas na imagem facilitaria a modelagem de algoritmos de compressão adaptativos específicos para imagens de diagnóstico, uma vez que a compressão poderia ser mais agressiva em estruturas não essenciais ao diagnóstico.

7.3 Artigos Publicados

Ressalta-se ainda a publicação dos artigos em conferências: "Aplicativo para caracterização de desempenho de métricas de qualidade visual"publicado nos Anais do XXXIV Simpósio Brasileiro de Telecomunicações (SBrT) 2016, o qual foi premiado com o 2º lugar na categoria de melhor trabalho de Iniciação Científica, *"System and protocol for collecting subjective data on clinical quality of compressed DICOM images"* publicado no XXX *Conference on Graphics, Patterns and Images* (SIBGRAPI) 2017 e "Avaliação de qualidade visual baseada em rede Bayesiana aplicada a imagens médicas"no XIV Congresso Brasileiro de Inteligência Computacional.

Referências Bibliográficas

[1] M.-M. Sung, H.-J. Kim, E.-K. Kim, J.-Y. Kwak, J.-K. Yoo, and H.-S. Yoo, "Clinical evaluation of JPEG2000 compression for digital mammography," *IEEE Transactions on Nuclear Science*, vol. 49, no. 3, pp. 827–832, Jun 2002.

[2] Conselho Federal de Medicina. (2007) Diário oficial da união. [Online]. Available: http://www.portalmedico.org.br/resolucoes/cfm/2007/1821_2007.pdf

[3] M. Kaur and V. Wasson, "ROI based medical image compression for telemedicine application," *Procedia Computer Science*, vol. 70, pp. 579 – 585, 2015, proceedings of the 4th International Conference on Eco-friendly Computing and Communication Systems. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1877050915032019

[4] S. G. L. George C. Kagadis, *Informatics in Medical Imaging*, 2nd ed. Taylor Francis Group, 2012.

[5] ITU-T, *Tutorial – Objective Perceptual Assessment of Video Quality: Full Reference Television*, International Telecommunication Union Geneva, Switzerland, 2004.

[6] European Commission, "European guidelines on quality criteria for computed tomography," http://www.drs.dk/guidelines/ct/quality/htmlindex.htm, 1999, [Online; accessed 08-September-2017].

[7] B. Kumar, S. B. Kumar, and C. Kumar, "Development of improved ssim quality index for compressed medical images," in *2013 IEEE Second International Conference on Image Information Processing (ICIIP-2013)*, Dec 2013, pp. 251–255.

[8] D. S. Taubman and M. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards and Practice*. Springer, 04 2002, vol. 11.

[9] P. Bak, M. Claudon, N. Dugar, D. Harvey, P. H. M. Kämmerer, D. Koff, E. Kotter, T. Kronander, R. Loose, C. Lundström, J. Martin, U.-E. Martin, P. Mildenberger, D. Müller, J. Reponen, A. Rowberg, L. Sutton, D. Taubman, and A. Tzannes, "Usability of irreversible image compression in radiological imaging. a position paper by the european society of radiology (ESR)," *Radiology*, vol. 2, no. 1, p. 103–115, Dec 2011.

[10] D. S. Taubman and M. W. Marcellin, "JPEG2000: standard for interactive imaging," *Proceedings of the IEEE*, vol. 90, no. 8, pp. 1336–1357, Aug 2002.

[11] Y. Ding, Ed., *Visual Quality Assessment for Natural and Medical Image*. Springer, 2008.

[12] C. Cavaro-Menard, L. Zhang, and P. L. Callet, "Diagnostic quality assessment of medical images: Challenges and trends," in *2010 2nd European Workshop on Visual Information Processing (EUVIP)*, July 2010, pp. 277–284.

[13] Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it? a new look at signal fidelity measures," *IEEE Signal Processing Magazine*, vol. 26, no. 1, pp. 98–117, Jan 2009.

[14] L. J. Karam, T. Ebrahimi, S. S. Hemami, T. N. Pappas, R. J. Safranek, Z. Wang, and A. B. Watson, "Introduction to the issue on visual media quality assessment," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 2, pp. 189–192, April 2009.

[15] A. Saha and Q. J. Wu, "Full-reference image quality assessment by combining global and local distortion measures," *Signal Processing*, vol. 128, pp. 186 – 197, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0165168416300196

[16] H. Chang, Q. Zhang, Q. Wu, and Y. Gan, "Perceptual image quality assessment by independent feature detector," *Neurocomputing*, vol. 151, pp. 1142 – 1152, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0925231214013344

[17] C. Martin, D. Sutton, and P. Sharp, "Balancing patient dose and image quality," *Applied Radiation and Isotopes*, vol. 50, no. 1, pp. 1 – 19, 1999. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0969804398000219

[18] Y. Zhou, D. Chen, C. Li, X. Li, and H. Feng, "A practice of medical image quality evaluation," in *International Conference on Neural Networks and Signal Processing, 2003. Proceedings of the 2003*, vol. 1, Dec 2003, pp. 204–207 Vol.1.

[19] C. Delgorge, C. Rosenberger, G. Poisson, and P. Vieyres, "Towards a new tool for the evaluation of the quality of ultrasound compressed images," *IEEE Transactions on Medical Imaging*, vol. 25, no. 11, pp. 1502–1509, Nov 2006.

[20] I. Kowalik-Urbaniaka, D. Bruneta, J. Wangb, D. Koffc, N. Smolarski-Koffc, E. R. Vrscaya, B. Wallaced, and Z. Wangb, "The quest for "diagnostically lossless" medical image compression: A comparative study of objective quality metrics for compressed medical images," *SPIE Medical Images: Image Perception, Observer Performance, and Technology*, vol. 9037, no. 4, Apr. 2014.

[21] L. Zhang, C. Cavaro-Menard, P. L. Callet, and J. Y. Tanguy, "A perceptually relevant channelized joint observer (PCJO) for the detection-localization of parametric signals," *IEEE Transactions on Medical Imaging*, vol. 31, no. 10, pp. 1875–1888, Oct 2012.

[22] D. Koff and H. Shulman, "An overview of digital compression of medical images: can we use lossy image compression in radiology?" *Canadian Association of Radiologists Journal*, vol. 57, no. 4, p. 211–217, Apr. 2006.

[23] D. Koff, P. Bak, P. Brownrig, D. Hosseinzadeh, A. Khademi, A. Kiss, L. Lepanto, T. Michalak, H. Shulman, and A. Volkening, "Pan-canadian evaluation of irreversible compression ratios ("lossy compression") for development of national guidelines," *Journal of Digital Imaging*, vol. 22, no. 6, p. 569–578, Apr. 2009.

[24] Conselho Federal de Medicina. (2012) Diário oficial da união. [Online]. Available: http://www.portalmedico.org.br/resolucoes/cfm/2002/1643_2002.htm

[25] A. Crop, K. Bacher, T. Hoof, P. Smeets, B. Smet, M. Vergauwen, U. Kiendys, p. duyck, K. Verstraete, K. D'Herde, and H. Thierens, "Correlation of contrast-detail analysis and clinical image quality assessment in chest radiography with a human cadaver study," *Radiology*, vol. 262, pp. 298–304, 11 2011.

[26] W. A. J. Jameelah H. Suad, "Subjective quality assessment of new medical image database," in *International Journal of Computer Engineering and Technology (IJCET)*, vol. 4, September 2013, pp. 155–164.

[27] R. Rangayyan, *Biomedical Image Analysis*, 1st ed. CRC PRESS, 2005.

[28] W. Birkfellner, *Applied Medical Image Processing: A Basic Course*. CRC Press, 2016. [Online]. Available: https://books.google.com.br/books?id=iVjOBQAAQBAJ

[29] NEMA. (2019) Data structures and encoding. [Online]. Available: http://dicom.nema.org/medical/dicom/current/output/html/part05.html#sect_8.2.1

[30] D. Taubman, E. Ordentlich, M. Weinberger, G. Seroussi, I. Ueno, and F. Ono, "Embedded block coding in JPEG2000," in *Proceedings 2000 International Conference on Image Processing (Cat. No.00CH37101)*, vol. 2, Sep. 2000, pp. 33–36 vol.2.

[31] T. M. Lehmann, C. Gonner, and K. Spitzer, "Survey: interpolation methods in medical image processing," *IEEE Transactions on Medical Imaging*, vol. 18, no. 11, pp. 1049–1075, Nov 1999.

[32] J. Leng, G. Xu, and Y. Zhang, "Medical image interpolation based on multi-resolution registration," *Computers Mathematics with Applications*, vol. 66, no. 1, pp. 1 – 18, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0898122113002538

[33] Conselho Nacional de Saúde. (2012) Diário oficial da união. [Online]. Available: http://bvsms.saude.gov.br/bvs/saudelegis/cns/2013/res0466_12_12_2012.html

[34] ——. (1996) Diário oficial da união. [Online]. Available: http://bvsms.saude.gov.br/bvs/saudelegis/cns/1996/res0196_10_10_1996.html

[35] J. W. Peirce, "Psychopy—psychophysics software in Python," *Journal of Neuroscience Methods*, vol. 162, no. 1–2, pp. 8 – 13, 2007. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0165027006005772

[36] J. Peirce, "Generating stimuli for neuroscience using PsychoPy," *Frontiers in Neuroinformatics*, vol. 2, p. 10, 2009. [Online]. Available: http://journal.frontiersin.org/article/10.3389/neuro.11.010.2008

[37] A. D. Crop, P. T. V. H. Klaus Bacher, P. V. Smeets, B. S. Smet, M. Vergauwen, U. Kiendys, P. Duyck, K. Verstraete, K. D'Herde, and H. Thierens, "Correlation of contrast-detail analysis and clinical image quality assessment in chest radiography with a human cadaver study," *Radiology*, vol. 262, no. 1, pp. 98–304, Jan. 2012.

[38] T. Eerola, L. Lensu, J.-K. Kamarainen, T. Leisti, R. Ritala, G. Nyman, and H. Kälviäinen, "Bayesian network model of overall print quality: Construction and structural optimisation," *Pattern Recognition Letters*, vol. 32, no. 11, pp. 1558 – 1566, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167865511001061

[39] R. de Freitas Zampolo and R. Seara, "Perceptual image quality assessment based on Bayesian networks," in *Image Processing, 2004. ICIP '04. 2004 International Conference on*, vol. 1, Oct 2004, pp. 329–332 Vol. 1.

[40] W. Preedanan, T. Kondo, P. Bunnun, and I. Kumazawa, "Image quality assessment for medical images based on gradient information," in *2018 5th International Conference on Business and Industrial Research (ICBIR)*, May 2018, pp. 189–194.

[41] B. P. Bondzulic and V. S. Petrovic, "Edge-based objective evaluation of image quality," in *2011 18th IEEE International Conference on Image Processing*, Sept 2011, pp. 3305–3308.

[42] B. Kumar, S. B. Kumar, and C. Kumar, "Development of improved ssim quality index for compressed medical images," in *2013 IEEE Second International Conference on Image Information Processing (ICIIP-2013)*, Dec 2013, pp. 251–255.

[43] Z. Ni, L. Ma, H. Zeng, C. Cai, and K. Ma, "Gradient direction for screen content image quality assessment," *IEEE Signal Processing Letters*, vol. 23, no. 10, pp. 1394–1398, Oct 2016.

[44] G. Chen, C. Yang, and S. Xie, "Gradient-based structural similarity for image quality assessment," in *2006 International Conference on Image Processing*, Oct 2006, pp. 2929–2932.

[45] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "Fsim: A feature similarity index for image quality assessment," *IEEE Transactions on Image Processing*, vol. 20, no. 8, pp. 2378–2386, Aug 2011.

[46] Z. Wang and X. Shang, "Spatial pooling strategies for perceptual image quality assessment," in *2006 International Conference on Image Processing*, Oct 2006, pp. 2945–2948.

[47] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, April 2004.

[48] K. B. Korb and A. E. Nicholson, *Bayesian Artificial Intelligence*, 2nd ed. CHAPMAN HALL/CRC, 2003.

[49] A. Papoulis, *Probability, Ramdom Variables, and Stochastic Processes*, 3rd ed. McGraw-Hill, 1991.

[50] Z. Harry, "The optimality of naive bayes," vol. 2, 01 2004.

[51] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[52] VQEG, "VQEG Final Report of FR-TV Phase II Validation Test," on-line, agosto 2003. [Online]. Available: http://www.its.bldrdoc.gov/vqeg/projects/frtv-phase-ii/frtv-phase-ii.aspx

[53] I. Avcibas, B. Sankur, and K. Sayood, "Statistical evaluation of image quality measures," *Journal of Electronic Imaging*, vol. 15, no. 2, p. 206–223, Nov. 2002.

[54] P. Corriveau and A. Webster, "Vqeg evaluation of objective methods of video quality assessment," *SMPTE Journal*, vol. 108, no. 9, pp. 645–648, Sept 1999.

[55] A. M. Rohaly, "Video quality experts group: Current results and future directions," *Visual Communications and Image Processing*, vol. 4067, no. 9, p. 742–753, Sept 2000.

[56] A. Agresti, *An introduction to categorical data analysis*. New York: Wiley, 1996. [Online]. Available: http://www.worldcat.org/search?qt=worldcat_org_all&q=0471113387

[57] A. Rencher, *Methods of Multivariate Analysis*, ser. Wiley Series in Probability and Statistics. Wiley, 2003. [Online]. Available: https://books.google.com.br/books?id=SpvBd7IUCxkC

[58] T. Landgrebe and R. Duin, "A simplified extension of the area under the roc to the multi-class domain," *Seventeenth Annual Symposium of the Pattern Recognition Association of South Africa*, 01 2006.