



Universidade Estadual de Feira de Santana
Programa de Pós-Graduação em Ciência da Computação

Um Método Semiautomático para Estimação de Fibrose Intersticial

Marcos Morais de Sousa

Feira de Santana

2023



Universidade Estadual de Feira de Santana
Programa de Pós-Graduação em Ciência da Computação

Marcos Morais de Sousa

Um Método Semiautomático para Estimação de Fibrose Intersticial

Dissertação apresentada à Universidade Estadual de Feira de Santana como parte dos requisitos para a obtenção do título de Mestre em Ciência da Computação.

Orientador: Angelo Amâncio Duarte

Coorientador: Washington Luis Conrado dos Santos

Feira de Santana

2023

Ficha Catalográfica – Biblioteca Central Julieta Carteado

S696m Sousa, Marcos Morais de
Um método semiautomático para estimação de fibrose
intersticial / Marcos Morais de Sousa.- Feira de Santana, 2023.

97.: il

Orientador: Angelo Amâncio Duarte

Coorientador: Washington Luis Conrado dos Santos

Dissertação (Mestrado) – Universidade Estadual de Feira de Santana,
Programa de Pós-graduação em Ciência da Computação , 2023.

1. Patologia computacional 2. Processamento de imagem. 3. Fibrose
intersticial - Computação. 4. Seguimento de cores - Computação. I.
Duarte, Angelo Amâncio, orient. II. Santos, Washington Luis Conrado
dos, coorient. III. Universidade Estadual de Feira de Santana. IV. Título.

CDU: 616.681.3

Tatiane Souza Santos - Bibliotecária CRB5/1634

Marcos Morais de Sousa

Um Método Semiautomático para Estimação de Fibrose Intersticial

Dissertação apresentada à Universidade Estadual de Feira de Santana como parte dos requisitos para a obtenção do título de Mestre em Ciência da Computação.

Feira de Santana, 8 de fevereiro de 2023

BANCA EXAMINADORA

Documento assinado digitalmente
 ANGELO AMANCIO DUARTE
Data: 04/04/2023 08:55:03-0300
Verifique em <https://validar.iti.gov.br>

Angelo Amâncio Duarte (Orientador(a))
Universidade Estadual de Feira de Santana



Karla Lais Pêgas

Irmandade Santa Casa de Misericórdia de Porto Alegre / Laboratório de Patologia e Micologia do Hospital Santa Rita

Assinado de forma digital por MICHELE FULVIA ANGELO:26554093885
Dados: 2023.04.03 11:06:18 -03'00'

Michele Fúlvia Angelo
Universidade Estadual de Feira de Santana

Abstract

The estimation of the fibrosis content in the interstitium, suggested that it may accompany immunological and infectious complications, which may be associated with a process of rejection of a transplanted kidney, is possible with the use of the Banff Classification, analyzing slides of renal biopsies stained with Dye Masson's Trichrome. The evaluation of Interstitial Fibrosis is semi-quantitative and causes subjectivity, which can be excessive, since the diagnosis and correct analysis of images is a complex task and highly dependent on the experience of the pathologists who perform it. In this scenario of reflection regarding the analysis of images of kidney biopsies, automatic methods to aid diagnosis using computational techniques can contribute to an increase in the quality of medical diagnosis. Considering that the Banff Classification is a relevant criterion for evaluating patients with chronic kidney disease modified by kidney transplantation, it is important to find ways to help mitigate subjectivity and intra- and inter-examiner variability in the evaluation of indicators used in this classification, especially in the inherited degree of interstitial fibrosis. In this context, this work presents a computational method to search semiautomatically for interstitial fibrosis, from digital images of renal biopsies, from the perspective of the Banff Classification. A *dataset* of 309 images was annotated by 5 pathologists and the results compared with the measurements made by the developed system. The results showed little agreement in the evaluation between the physicians, in addition to a large variance when they performed two consecutive measurements of the same image set. In relation to the quality of the system, it carried out the measurement invariably, but its precision could not be definitively evaluated, due to the variability obtained in the analysis of the pathologists. Due to the variable degree of reproducibility and agreement presented by the pathologists, it was not possible to assess the quality of the computational method, since a gold standard was not found in the medical team. Even so, this project built a semi-automatic model for evaluating interstitial fibrosis that is available as part of the functionality of the PathoSpotter project (<https://pathospotter.bahia.fiocruz.br>). The system can be used as an estimator of the degree of interstitial fibrosis in biopsy slide sections that are selected by the pathologist, and can be adjusted in the future by obtaining a *dataset* with consistent notes of interstitial fibrosis.

Keywords: image processing; core segmentation; interstitial fibrosis; computational

pathology.

Resumo

A estimativa do teor de fibrose no interstício, critério que pode apontar complicações imunológicas e infecciosas, que podem estar associadas a um processo de rejeição de um rim transplantado, é possível com o uso da Classificação de Banff, analisando lâminas de biópsias renais coradas com Corante Tricromo de Masson. A avaliação da Fibrose Intersticial é semi-quantitativa e ocasiona subjetividade, que pode ser excessiva, uma vez que o diagnóstico e a análise correta de imagens são uma tarefa complexa e altamente dependente da experiência dos patologistas que a realizam. Neste cenário de incertezas quanto às análises de imagens de biópsias renais, métodos automáticos para auxílio ao diagnóstico utilizando técnicas computacionais podem contribuir para um aumento da qualidade do diagnóstico médico. Considerando que a Classificação de Banff é um relevante critério para avaliação do paciente com doença renal crônica submetido a transplante renal, é importante encontrar formas de auxiliar na mitigação da subjetividade e na variabilidade intra e interexaminador na avaliação de indicadores usados nesta classificação, especialmente na determinação do grau de fibrose intersticial. Neste contexto, este trabalho apresenta um método computacional para estimação semiautomática de fibrose intersticial, a partir de imagens digitais de biópsias renais, sob a perspectiva da Classificação de Banff. Um *dataset* de 309 imagens foi anotado por 5 patologistas e os resultados comparados com as medições feitas pelo sistema desenvolvido. Os resultados apontaram pouca concordância na avaliação entre os médicos, além de grande variância quando estes realizaram duas medidas consecutivas de um mesmo conjunto de imagem. Já em relação à qualidade do sistema, o mesmo realizou a medição de forma invariável, porém sua precisão não pôde ser avaliada de forma definitiva, em virtude da variabilidade obtida na análise dos patologistas. Em decorrência do grau variável de reprodutibilidade e concordância apresentado pelos médicos patologistas, a avaliação do desempenho do método computacional não trouxe resultados conclusivos, uma vez que não foi encontrado um padrão-ouro na equipe médica. Ainda assim, este projeto construiu um modelo semi-automático de avaliação de fibrose intersticial que está disponibilizado como parte das funcionalidades do projeto PathoSpotter (<https://pathospotter.bahia.fiocruz.br>). O sistema pode ser usado como estimador do grau de fibrose intersticial de seções da lâmina de biópsia que sejam selecionadas pelo patologista, e pode ser futuramente ajustado a partir da obtenção de um *dataset* com anotações consistentes de fibrose intersticial.

Palavras-chave: processamento de imagem; segmentação de cores; fibrose intersticial; patologia computacional.

Prefácio

Esta dissertação de mestrado foi submetida à Universidade Estadual de Feira de Santana (UEFS) como requisito parcial para obtenção do grau de Mestre em Ciência da Computação.

A dissertação foi desenvolvida no Programa de Pós-Graduação em Ciência da Computação (PGCC), tendo como orientador o Prof. Dr. **Angelo Amâncio Duarte**. O Prof. Dr. **Washington Luis Conrado dos Santos** foi coorientador(a) deste trabalho.

Agradecimentos

Agradeço primeiramente a Deus, por me presentear com tudo que foi necessário para eu sobrepujar todos os desafios inerentes enfrentados neste mestrado.

Agradeço de forma muito especial à minha esposa Carla e aos meus filhos, Dimitri, Dianna e David, por todo amor e compreensão, bem como pelas inúmeras horas furtadas de seu convívio.

Agradeço a minha mãe e ao meu pai, postumamente, por terem me dado os ensinamentos estruturais que ajudaram na formação dos meus valores morais e éticos.

Agradeço especialmente ao amigo-irmão Pedro Alan e a minha querida irmã Michelle que sempre me incentivaram, acreditando em meu potencial, muitas vezes bem mais que eu mesmo.

Agradeço aos gestores do Grupo Valdir Barreto, se não fosse pela compreensão e fé em meu profissionalismo, não seria possível o meu ingresso e a conclusão deste mestrado.

Agradeço ao Prof. Dr. Angelo Amâncio Duarte e ao Prof. Dr. Washington L. C. dos Santos que aceitaram o desafio da orientação e coorientação, respectivamente, e que, com muito entusiasmo, me auxiliaram tanto na germinação das ideias durante todo o processo de desenvolvimento desta pesquisa.

Não posso esquecer-me de agradecer especialmente ao Dr. José Nathan Andrade que, com sua paciência e conhecimento, tanto contribuiu para a concretização deste trabalho, bem como à equipe de médicos patologistas do projeto PathoSpotter.

Agradeço aos professores e colegas do Programa de Pós-Graduação em Ciência da Computação (PGCC) da Universidade Estadual de Feira de Santana (UEFS), aos colegas do Laboratório de Computação de Alto Desempenho (LaCAD), bem como aos demais amigos cujo convívio me proporcionaram momentos de relaxamento necessários para a continuidade desta pesquisa.

A todos que não foram citados, mas que tiveram participação nesta etapa da minha vida, meus sinceros e verdadeiros agradecimentos.

*“Lembrem disto: quem planta pouco
colhe pouco; quem planta muito
colhe muito.”*

(Bíblia Sagrada, 2 Coríntios 9:6)

Sumário

Abstract	i
Resumo	iii
Prefácio	v
Agradecimentos	vi
Sumário	x
Lista de Tabelas	xi
Lista de Figuras	xiv
Lista de Abreviações	xv
Lista de Símbolos	xvii
1 Introdução	1
1.1 Objetivos	3
1.1.1 Objetivo Geral	3
1.1.2 Objetivos Específicos	3
1.2 Contribuições	4
1.3 Organização do Trabalho	4
2 Revisão Bibliográfica	5
2.1 Conceitos elementares de nefropatologia usados neste trabalho	5
2.1.1 Os rins	5
2.1.2 Fibrose Renal	7
2.1.3 Como é avaliada a Fibrose Renal	7
2.1.4 O Corante tricrômico de Masson	9
2.1.5 Classificação de Banff	10
2.2 Conceitos de computação usados neste trabalho	13
2.2.1 Processamento digital de imagens	13
2.2.2 Visão Computacional	15

2.2.3	Cores do Espectro Visível	19
2.2.4	Modelos de cores	20
2.2.5	Segmentação de imagens baseada na cor	23
2.2.6	O problema da Segmentação de Imagens por Cor	28
2.3	Conceitos complementares	30
2.3.1	Coefficiente Kappa	30
2.4	Trabalhos Relacionados	33
3	Metodologia	37
3.1	Fase 1: Montagem da equipe de patologistas.	38
3.2	Fase 2: Pesquisa e escolha do <i>software</i> para anotação médica.	38
3.3	Fase 3: Escolha das imagens digitalizadas de biópsia.	41
3.4	Fase 4: Preparação do <i>Dataset</i>	43
3.5	Fase 5: Curadoria do <i>Dataset</i>	43
3.6	Fase 6: Avaliação do percentual de fibrose intersticial pelo método desenvolvido.	44
3.7	Fase 7: 2 ^a avaliação do percentual de fibrose intersticial pelos médicos.	47
3.8	Fase 8: Comparação dos resultados obtidos.	47
4	Resultados	51
4.1	Resultados de concordância intraexaminador	51
4.1.1	Concordância intraexaminador do médico A	51
4.1.2	Concordância intraexaminador do médico B	52
4.1.3	Concordância intraexaminador do médico C	52
4.1.4	Concordância intraexaminador do médico D	53
4.1.5	Concordância intraexaminador do médico E	53
4.1.6	Conclusões sobre análise interexaminador	54
4.2	Resultados de concordância interaexaminador	54
4.2.1	Resultados de concordância entre médicos A e B	54
4.2.2	Resultados de concordância entre médicos A e C	55
4.2.3	Resultados de concordância entre médicos A e D	55
4.2.4	Resultados de concordância entre médicos A e E	55
4.2.5	Conclusões sobre análise interexaminador do médico A	56
4.2.6	Resultados de concordância entre médicos B e C	56
4.2.7	Resultados de concordância entre médicos B e D	56
4.2.8	Resultados de concordância entre médicos B e E	57
4.2.9	Conclusões sobre análise interexaminador do médico B	57
4.2.10	Resultados de concordância entre médicos C e D	57
4.2.11	Resultados de concordância entre médicos C e E	58
4.2.12	Conclusões sobre análise interexaminador do médico C	58
4.2.13	Resultados de concordância entre médicos D e E	58
4.2.14	Conclusões sobre análise interexaminador D	59
4.3	Resultados de concordância interexaminador entre o método computacional e os médicos A,B,C,D,E.	59

4.3.1	Resultados de concordância interexaminador: $[\Sigma,A]$	59
4.3.2	Resultados de concordância interexaminador: $[\Sigma,B]$	59
4.3.3	Resultados de concordância interexaminador: $[\Sigma,C]$	60
4.3.4	Resultados de concordância interexaminador: $[\Sigma,D]$	60
4.3.5	Resultados de concordância interexaminador: $[\Sigma,E]$	61
4.3.6	Conclusões sobre análise de concordância entre médicos e o método computacional	61
5	Considerações Finais	65
5.1	Trabalhos Futuros	66
	Referências	68

Lista de Tabelas

2.1	Escore de Banff para lesão por fibrose intersticial (<i>score Banff ci</i>) . . .	11
2.2	Exemplo de tabela de contigência r x r	31
2.3	Exemplo de aplicação do Coeficiente de Concordância de Kappa . . .	33
3.1	Equipe de Médicos Patologistas.	38
3.2	<i>Softwares</i> para auxílio à anotação de imagens médicas que foram encontrados e avaliados.	39
3.3	Exemplo de comprimentos de ondas no espectro visível para a cor azul encontrado em algumas literaturas	46
3.4	Resultados da Concordância Intraexaminador	49
3.5	Tabela baseada na interpretação Kappa sugerida por Landis e Koch (1977)	50
4.1	Resultados da Concordância Intraexaminador	52
4.2	Resultados da Concordância Intraexaminador	52
4.3	Resultados da Concordância Intraexaminador	53
4.4	Resultados da Concordância Intraexaminador	53
4.5	Resultados da Concordância Intraexaminador	54
4.6	Resultados da Concordância Interexaminador	54
4.7	Resultados da Concordância Interexaminador	55
4.8	Resultados da Concordância Interexaminador	55
4.9	Resultados da Concordância Interexaminador	56
4.10	Resultados da Concordância Interexaminador	56
4.11	Resultados da Concordância Interexaminador	57
4.12	Resultados da Concordância Interexaminador	57
4.13	Resultados da Concordância Interexaminador	58
4.14	Resultados da Concordância Interexaminador	58
4.15	Resultados da Concordância Interexaminador	59
4.16	Resultados da Concordância Interexaminador	59
4.17	Resultados da Concordância Interexaminador	60
4.18	Resultados da Concordância Interexaminador	60
4.19	Resultados da Concordância Interexaminador	61
4.20	Resultados da Concordância Interexaminador	61
4.21	Resumo Comparativo de Concordância	63

Lista de Figuras

2.1	Estruturas que formam o rim humano visto em seção transversal (University Of Utah, 2022).	6
2.2	Esquemas de rim vistos por cortes longitudinais do órgão (Junqueira e Carneiro, 2017).	6
2.3	Interstício com fibrose. O azul “intenso” é a fibrose no interstício. Elaboração do Autor (2022).	8
2.4	Rim: Glomerulonefrite Membranoproliferativa, Tipo II. Glomérulos com leve aumento da matriz mesangial. Corante tricrômico de Masson, 200x. Adaptado de Rosman (2012).	9
2.5	O conteúdo da Classificação de Banff de Transplante Renal pela qual a patologia pode ser inventariada com escores de lesão de Banff e Parâmetros de diagnóstico exigidos pelos algoritmos por trás das Categorias de diagnóstico Banff e chegar a um diagnóstico (Roufosse et al., 2018).	11
2.6	Imagem Conceitual, retratando de forma similar como o patologista enxerga a fibrose intersticial. A coloração da imagem é simbólica e não se relaciona com as cores da lâmina de biópsia renal, coradas com tricrômico de Masson (Roufosse et al., 2018).	12
2.7	Etapas fundamentais em processamento digital de imagens (Gonzalez e Woods, 2009).	14
2.8	Campos de estudo interligados a Visão Computacional (Barelli, 2018).	16
2.9	Exemplo de um Sistema de Visão Computacional e suas principais etapas (Marques Filho e Neto, 1999)	17
2.10	Detecção e segmentação de objetos (Granaty, 2022).	18
2.11	(A) Esquema de cor RGB e (B) Esquema de cor HSV (Gonzalez e Woods, 1999; Gomes J.; Velho, 1994).	21
2.12	(A) Espaço das Cores RGB aditivas e (B) espaço das cores CMYK subtrativas (George e Beaird, 2012).	21
2.13	Esquema de cor Matiz, Saturação e Valor (George e Beaird, 2012).	22
2.14	Esquema do algoritmo JSEG (Deng et al., 1999).	23
2.15	(A) Imagem original (B) e Mapa de classes. Adaptado de Deng et al. (1999).	25

2.16	Segmentação de imagem no espaço HSV. (a) Original. (b) Matiz. (c) Saturação. (d) Intensidade. (e) Máscara de saturação binária (preto = 0). (f) Produto de (b) por (e). (g) Histograma de (f). (h) Segmentação dos componentes vermelhos em (a) (Gonzalez e Woods, 2009).	26
2.17	Segmentação no espaço de cores RGB. (a) Imagem original com as cores de interesse delimitadas por um retângulo. (b) Resultado da segmentação no espaço de cores RGB. Compare com a Figura 2.16 (h) (Gonzalez e Woods, 2009).	28
2.18	Espaço de cores idealizado por Pitágoras (Colorsystem, 2020).	29
2.19	Exemplo de medidas de acordo para dados categóricos (Landis e Koch, 1977)	32
2.20	Quadro-Resumo dos trabalhos considerados relevantes quando comparados à proposta deste estudo, e dos datasets utilizados. Elaboração do Autor (2022).	36
3.1	Metodologia. Elaboração do Autor (2022)	37
3.2	Arquitetura do <i>software</i> Cytomine (Montefiore Institute, 2010).	40
3.3	Aparelhos usados para obtenção de Imagens Médicas: (a) Pathobin USB C-Mount Camera; (b) Olympus BX40 Clinical Microscope (Pathobin, 2022) e (Microscope, 2022).	41
3.4	Imagens completas de biópsia renal (do inglês (<i>Whole Slide Imaging-WSI</i>) com visão panorâmica do fragmento (coloração TM). Elaboração do Autor (2022)	42
3.5	Ilustração do processo de envio de uma imagem digitalizada de lâmina de biópsia ao Cytomine, e da identificação pelo médico da área de interesse e da quantificação da fibrose intersticial, gerando uma nova imagem (recorte). Elaboração do Autor(2022).	43
3.6	Exemplo de recortes de imagens removidos do <i>dataset</i> porque são áreas da medula renal, que não devem ser consideradas para a avaliação de fibrose intersticial. Elaboração do Autor(2022).	44
3.7	Exemplo de recortes de imagens que compõem o <i>dataset</i> : (a) a (d) Representação histológica da medula renal. Elaboração do Autor(2022).	45
3.8	Estrutura do método computacional. Elaboração do Autor (2022)	46
3.9	Imagem de Lâmina Inteira (<i>Whole Slide Imaging, WSI</i>) com regiões afetadas por fibrose intersticial. A Figura é um exemplo em que patologistas distintos podem olhar para regiões diferentes de uma mesma imagem de Lâmina Inteira (<i>Whole Slide Imaging, WSI</i>). Elaboração do Autor (2022).	47
3.10	Exemplo da normalização de valores percentuais usando critério de Banff. Elaboração do Autor (2022).	48

4.1	Resultado Gráfico da Comparação entre Anotadores. Elaboração do Autor (2022).	64
-----	---	----

Lista de Abreviações

Abreviação	Descrição
ABTO	Associação Brasileira de Transplante de Órgãos
ADI	Análise Digital de Imagens
API	<i>Application Programming Interface</i>
BMP	<i>Bitmap</i>
COCO	<i>Common Objects in Context</i>
CIE	<i>Comission Internationale de l'Eclairage</i>
CMYK	<i>Cyan, Magenta, Yellow, black</i>
CNN	<i>Convolutional Neural Networks</i>
DAS	Anticorpos Específicos Anti-HLA
DRC	Doença Renal Crônica
EPS	<i>Encapsulated PostScript</i>
FI	Fibrose Instersticial
FIOCRUZ	Fundação Oswaldo Cruz
GUI	<i>Graphical User Interface</i>
HSV	<i>Hue, Saturation, Value</i>
Lab	Luminosidade, coordenada vermelho/verde e amarelo/azul
LPBI	Laboratório de Patologia e Biointervenção do Centro de Pesquisas Gonçalo Moniz
HLA	<i>Human Leukocyte Antige</i>
MRXS	Extensão de Arquivos de Imagem Raster
PDI	Processamento Digital de Imagens
PNG	<i>Portable Network Graphic</i>
PGCC	Programa de Pós-Graduação em Ciência da Computação
RGB	<i>Red, Green, Blue</i>
HSV	<i>Hue, Saturation, Value</i>
JPEG	<i>Joint Photographic Experts Groups</i>
TM	Tricrômico de Masson
TIFF ou TIF	<i>Tagged Image File Format</i>
TRS	Terapia Renal Substitutiva
UEFS	Universidade Estadual de Feira de Santana
VC	Visão Computacional
VMS	<i>Video Management System</i>
XML	<i>Extensible Markup Language</i>

WSI

Whole Slide Imaging

Lista de Símbolos

Símbolo	Descrição
λ	Lambda
Σ	Sigma

Capítulo 1

Introdução

Estudos apontam que o Brasil possui o maior programa público de saúde no mundo, sendo o segundo maior transplantador de órgãos em números absolutos, atrás apenas dos EUA (Brasil, 2019, 2020; Maior, 2021). Em 2019, ocorreram cerca de 6417 transplantes renais, denominados por alguns especialistas como Terapia Renal Substitutiva (TRS), posicionando o país em segundo lugar em um universo de 35 países (ABTO, 2020).

A TRS é um tratamento complexo, podendo envolver complicações imunológicas e infecciosas que se devem, principalmente, à rejeição do rim (Torres et al., 2021). Para acompanhamento do paciente submetido a transplante renal, o exame de biópsia renal é fundamental (Colvin e Chang, 2019).

A biópsia renal é um procedimento invasivo, no qual se remove um fragmento do rim para análise médica, feita com auxílio de um microscópio (Colvin e Chang, 2019; Kasvi, 2020).

No exame anatomopatológico da biópsia renal, é possível avaliar diversos parâmetros histopatológicos, o que é de grande importância para o diagnóstico de doenças, formulação de planos de tratamento e julgamento do prognóstico (Zheng et al., 2021). Para a TRS um marcador de grande importância é a fibrose renal.

A fibrose renal é acumulada após a perda de tecido por alguma lesão ocasionada no rim (Sette et al., 2010; Menn-Josephy et al., 2016; Brasil, 2020), que pode ocorrer devido à progressão da Doença Renal Crônica (DRC)¹ ou após transplantes renais (Boor et al., 2010; Farris e Alpers, 2014; Santos, 2021). Quando essa fibrose renal ocorre no interstício renal (espaço entre os néfrons e vasos sanguíneos e linfáticos), é conhecida como fibrose intersticial (FI) (Farris e Alpers, 2014).

A estimativa da quantidade de fibrose intersticial serve para avaliar o grau de lesão renal (Farris e Alpers, 2014). Essa estimativa é possível com o uso da técnica

¹Doença Renal Crônica (DRC), que é uma doença altamente prevalente (da população), doença irreversível, progressiva e associada a maior risco cardiovascular.

histológica de coloração tricrômica, ou técnica de Masson. Essa técnica é usada rotineiramente nos laboratórios, porque apresenta pouca variação da intensidade de cor e efeito artefactual (objetos falsos que aparecem por conta de defeitos nas imagens e que não são reais) (Van De Vlekkert et al., 2020; Farris e Alpers, 2014).

A estimativa da FI é uma avaliação semiquantitativa da extensão de uma cicatriz no tecido renal ou parênquima (Menn-Josephy et al., 2016). Nesta avaliação, associa-se valores às escalas qualitativas, não sendo necessário que o número atribuído a cada descrição corresponda exatamente à magnitude real do fenômeno analisado, o que pode ocasionar subjetividade na avaliação (Farris et al., 2014; Berchtold et al., 2017).

A subjetividade da avaliação da FI pode ser expressiva, uma vez que o diagnóstico de imagens é uma tarefa complexa e altamente dependente da experiência dos patologistas que realizam a análise, podendo comprometer a qualidade da análise da FI (Pinheiro, 2020). Além da subjetividade da avaliação da FI, patologistas apresentam grau variável de reprodutibilidade em diferentes técnicas de histoquímicas, conforme relatado por Berchtold et al. (2017) e Farris et al. (2014).

Em um esforço para melhorar a qualidade da análise da FI, o método classificatório de Banff, criado em 1991, busca tornar o processo de avaliação da FI mais objetivo e consistente, minimizando a subjetividade e a variabilidade interobservador na quantificação da fibrose intersticial (Loupy et al., 2020).

Resumidamente, esse método de classificação padroniza a estimativa de FI, tendo como premissa fundamentos histológicos. Desde então, padrões avaliativos da FI vêm sendo estabelecidos e revisitados, reduzindo a subjetividade e a heterogeneidade nos diagnósticos clínicos (Shashiprakash et al., 2021).

Mesmo com o uso da classificação de Banff para uma melhor avaliação do grau de FI, esta avaliação ainda está sujeita a subjetividades por parte do patologista que analisa a imagem. Uma abordagem que pode ser usada para melhorar o processo, sempre que possível, é avaliar o grau de concordância da avaliação da FI entre dois ou mais patologistas, procedimento que pode ser realizado via métodos estatísticos, como o método Kappa de Cohen (Glen, 2014), tipicamente usado em ciências médicas ou sociais para obter uma avaliação quantitativa do grau de concordância entre diferentes métodos ou analistas (Vergni et al., 2021).

Outra estratégia para enfrentar o problema de subjetividade e a variabilidade interobservador é desenvolver técnicas computacionais para auxílio ao diagnóstico, as quais analisam a imagem de maneira determinística, não sofrendo influência das variações inerentes à análise da imagem por um patologista (Cerqueira et al., 2021).

Os autores Farris et al. (2021) fizeram uma revisão da literatura discutindo algoritmos de Inteligência Artificial, com ênfase em transplante renal, e relataram que este procedimento é frequentemente classificado pelo uso da classificação de Banff. No trabalho de Ginley et al. (2021), os autores sugerem que uso de métodos de apren-

dizagem de máquina podem realizar a tarefa de classificação da fibrose intersticial com desempenho comparável ao de um patologista.

Apenas muito recentemente, no mês de maio de 2022, foi publicado um trabalho Liu et al. (2022b) no qual os autores apresentam um modelo baseado em aprendizagem de máquina que realiza todo o processo de avaliação da fibrose intersticial. Segundo os autores, o estudo apresenta o primeiro processo totalmente automático baseado em aprendizado de máquina para avaliação de fibrose intersticial em biópsias renais. Na Seção 2.4 do Capítulo 2, comentar-se-á mais sobre este e alguns trabalhos correlacionados à proposta deste estudo.

É importante registrar que no início de 2021, quando este trabalho foi iniciado, não haviam na literatura trabalhos que apresentassem um sistema automático para replicar a avaliação da FI por patologistas na perspectiva da classificação de Banff. Mesmo hoje com esses trabalhos publicados, ainda não existe a disposição para avaliação um sistema que faça a estimativa automática de fibrose, já que nenhum dos estudos oferece um sistema para uso público.

Dentro deste cenário, esse trabalho foi iniciado em 2021 com propósito de desenvolver uma ferramenta para estimação semiautomática da FI como parte das funcionalidades do projeto PathoSpotter (<http://pathospotter.uefs.br/>), o qual desenvolve ferramentas computacionais para auxílio ao diagnóstico de lesões renais a partir de imagens de lâminas completas de biópsias renais (do inglês *Whole Slide Imaging – WSI*).

O método indicado neste trabalho propõe uma solução para diminuir a subjetividade através da avaliação semiautomática da fibrose intersticial (FI), sob a perspectiva da classificação de Banff.

Sabendo que o corante tricrômico de Masson é usado rotineiramente na estimação da FI e que este, ao entrar em contato com fibrose intersticial, reage em cor azul, nossa premissa foi que um algoritmo baseado em segmentação por cor, especificamente a azul, seria capaz de realizar uma quantificação da fibrose intersticial e que apresente desempenho comparável ao de um patologista.

1.1 Objetivos

1.1.1 Objetivo Geral

Desenvolver um método semiautomático para estimação de fibrose intersticial que apresente desempenho similar ao de um médico patologista, a partir de imagens de lâminas completas de biópsias renais, sob a perspectiva da classificação de Banff.

1.1.2 Objetivos Específicos

- Criar um banco de dados de imagens anotadas de fibrose intersticial, disponibilizando-o à comunidade interessada.

- Incluir a solução desenvolvida ao portfólio de ferramentas computacionais do projeto PathoSpotter (<http://pathospotter.uefs.br/>).

1.2 Contribuições

Considerando que a classificação de Banff orienta o manejo do paciente com Doença Renal Crônica (DRC) submetido a transplantes renais, é importante encontrar formas de reduzir a subjetividade e a variabilidade interobservador na avaliação de indicadores usados nessa classificação, principalmente na determinação do grau de fibrose intersticial, requerida para avaliação de Dano Crônico Túbulo Intersticial² (Frank O'brien, 2020), o qual frequentemente aparece na ocorrência de rejeição de transplantes renais.

Nessa perspectiva, o presente estudo pode cooperar de forma prática com os patologistas na análise de imagens de lâminas completas de biópsias renais, apresentando contribuição científica e tecnológica, uma vez que, até o momento da escrita deste trabalho, foi encontrado apenas um estudo publicado que trata da estimativa automática de fibrose intersticial renal, o qual não se encontra disponível como serviço de fácil acesso para patologistas, e ainda precisa ser validado, usando mais casos de análise.

Além disso, este estudo pode contribuir para o avanço da Nefropatologia Computacional, área de pesquisa recente que apresenta problemas em aberto, fomentando a interação entre patologistas e cientistas da computação – um dos objetivos do projeto PathoSpotter.

1.3 Organização do Trabalho

Esse texto foi dividido em cinco capítulos, sendo o Capítulo 1 a presente Introdução, na qual o leitor pode obter um prévio relato daquilo que será abordado nos demais capítulos.

O Capítulo 2 apresenta os temas que foram importantes para a construção do conhecimento utilizado neste trabalho, bem como expõe alguns estudos correlacionados a esta proposta de modo a nortear a presente pesquisa.

O Capítulo 3 descreve os processos para implementação do método computacional, bem como os instrumentos empregados e a avaliação das capacidades e limitações do método computacional.

No Capítulo 4 são apresentados e discutidos os resultados alcançados.

No Capítulo 5 são feitas as considerações finais, os objetivos propostos e os resultados alcançados, com sugestões que possam viabilizar a continuidade desta pesquisa.

²Dano Crônico Túbulo Intersticial é a lesão primária de túbulos renais e interstício resultante na diminuição da função renal.

Capítulo 2

Revisão Bibliográfica

Neste capítulo serão tratados os principais conceitos relacionados a este trabalho, classificados em três sessões: Conceitos Teóricos da Medicina, Conceitos Teóricos da Computação e Trabalhos Relacionados.

2.1 Conceitos elementares de nefropatologia usados neste trabalho

A seguir, são apresentados assuntos da medicina necessários à compreensão dos leitores desse estudo.

2.1.1 Os rins

Responsáveis por filtrar o sangue e produzir a urina por onde são eliminados água, eletrólitos, hormônios, ureia e diversas toxinas do organismo, os rins atuam como órgãos produtores de algumas substâncias como eritropoetina; uma glicoproteína (formada por 165 aminoácidos), que estimula a produção de eritrócitos; renina, que participa da regulação da pressão sanguínea e que, junto ao fígado e pele, é responsável pela ativação da vitamina D – hormônio que atua na saúde óssea, crescimento, imunidade, musculatura, metabolismo e em diversos órgãos e sistemas, como o cardiovascular e o sistema nervoso central (Junqueira e Carneiro, 2017; São Paulo, 2021; Pfizer, 2021; Charles Patrick Davis, 2021; Kidshealth, 2021).

Os rins são dois órgãos localizados em ambos os lados da coluna vertebral à altura da 12^a vértebra torácica e da 3^a vértebra lombar, atrás das últimas costelas, medindo aproximadamente doze centímetros de comprimento e cinco centímetros de largura. Pesando cerca de 150 gramas cada rim, seu formato assemelha-se ao de um grão de feijão, com uma borda convexa e outra côncava onde fica o hilo, região composta por tecido adiposo por onde entram e saem vasos sanguíneos e linfáticos, cálices, nervos e o ureter (Junqueira e Carneiro, 2017; Montanari, 2016). A Figura 2.1 apresenta

em seção transversal, um rim adulto normal, demonstrando o córtex externo mais claro, e a medula mais escura, com as pirâmides renais em que os dutos coletores se fundem e drenam para os cálices e pelve centrais.



Figura 2.1: Estruturas que formam o rim humano visto em seção transversal (University Of Utah, 2022).

Na parte central do rim, localizada ao lado do hilo e da região côncava, delimitada pela medula renal, fica o seio renal, ocupado por espaços conhecidos como cálices menores e cálices maiores. Os cálices fazem parte da parede de um tubo em forma de funil, conhecido como cálice renal, do qual surge o ureter (Junqueira e Carneiro, 2017; Montanari, 2016). A Figura 2.2 apresenta esquemas de um rim vistos por cortes longitudinais do órgão.

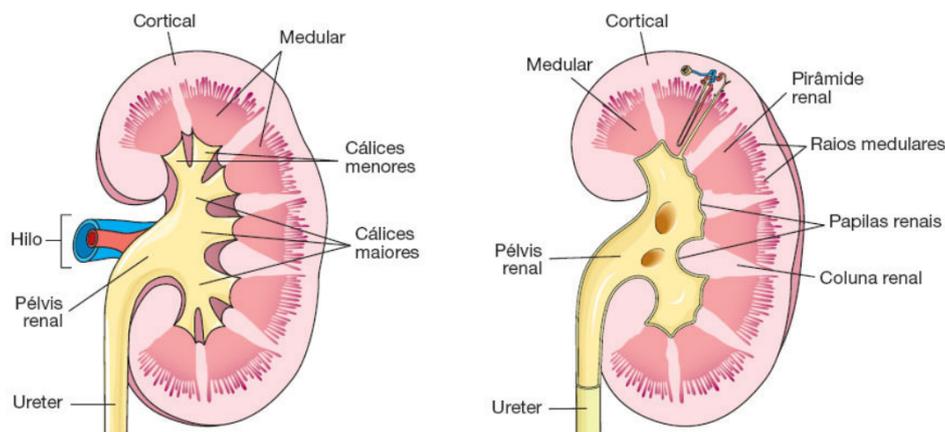


Figura 2.2: Esquemas de rim vistos por cortes longitudinais do órgão (Junqueira e Carneiro, 2017).

À direita destas estruturas, estão as pirâmides que, em sua extremidade mais larga, conectam-se ao córtex renal. Juntos, o córtex e as pirâmides da medula renal constituem a parte funcional do rim, região comumente conhecida como parênquima renal (Junqueira e Carneiro, 2017; Montanari, 2016).

Os rins são órgãos de importância vital que podem adoecer de forma crônica, acarretando uma série de alterações patogênicas, que podem ocasionar o surgimento de fibrose renal e, conseqüentemente, a perda progressiva da função renal (Sette et al., 2010; Brasil, 2020).

2.1.2 Fibrose Renal

Processos inflamatórios crônicos podem ser ocasionados por várias doenças que acometem o parênquima renal (tecido renal) e progridem para várias alterações estruturais do tecido. Essas alterações levam a lesões de natureza hemodinâmica¹, resultante de perda significativa da massa renal, seja por motivo de doença, acidentes ou intervenções cirúrgicas. Esse tipo de lesão, pode ainda acontecer em decorrência de processos inflamatórios crônicos. A inflamação crônica pode resultar na formação de fibras de colágeno produzidas por células miofibroblásticas² (Boor et al., 2010).

O acúmulo progressivo dessas fibras é chamado de fibrose e, quando ocorre nos túbulos renais e nos tecidos que circundam os rins, denomina-se fibrose intersticial (FI) (Sette et al., 2010; Brasil, 2020). A fibrose intersticial, em conjunto de outras alterações, acarreta a perda da função renal. Logo, é imprescindível a avaliação do grau desta alteração durante análise histopatológica³ de um fragmento do tecido renal (Nikonovas et al., 2020).

2.1.3 Como é avaliada a Fibrose Renal

A avaliação da fibrose renal é um processo histológico. Em sentido amplo, um processo histológico ou histologia é sinônimo de anatomia microscópica, não estando limitado ao estudo das células, dos tecidos e de como essas estruturas se organizam para formar os órgãos, mas também tratando das suas funções (Gartber, 2007; Junqueira e Carneiro, 2017).

Para o estudo de tecidos, vasos, células e outras estruturas microscópicas, a histologia analisa amostras processadas de tecido, preparadas em lâminas de vidro e que são estudadas utilizando-se um microscópio. Esse procedimento é realizado por meio do exame de biópsia que consiste na análise microscópica dos tecidos (Nikonovas et al., 2020). De acordo com Belsare e Mushrif (2012), quando essa análise é realizada pelo médico patologista, procurando detectar a presença de doenças como, por exemplo, a fibrose renal, é chamada histopatologia.

¹As lesões hemodinâmicas são os principais mecanismos patogênicos da doença renal crônica, que ocorrem não apenas por processos mecânicos, mas também pela ativação contínua de uma variedade de mediadores inflamatórios.

²As células miofibroblásticas reagem para anticorpos para proteínas contráteis, principalmente actina (1A4, HHHF-35), e/ou associadas ao processo de contração, como proteínas reguladores (calponina, caldesmon)

³Avaliação histopatológica consiste na análise microscópica dos tecidos para a detecção de possíveis lesões existentes

Por meio do exame de biópsia é possível avaliar a fibrose renal. Quando, na avaliação da fibrose renal, leva-se em consideração apenas a região do interstício renal, tem-se a fibrose intersticial (FI). Para a realização desse exame de biópsia, uma agulha específica ou um procedimento cirúrgico é realizado para extrair um fragmento do tecido renal. Esse fragmento é preparado para a lâmina de vidro com substâncias químicas que preservam suas características, a fim de que seja enviado a um laboratório de patologia humana (Kelly e Landman, 2014).

No laboratório de patologia humana, o médico patologista submete o fragmento a diversos cortes por micrótomo. Micrótomo é o aparelho que faz cortes microscópicos, variando geralmente de 1 a 10 micrômetros de espessura, em pequenas amostras de material biológico emblocadas em resinas específicas para análise em microscópio óptico; e realiza sua coloração usando técnicas histoquímicas como, por exemplo, a técnica histoquímica do corante tricrômio de Masson (Junqueira e Carneiro, 2017). A Figura 2.3 mostra uma imagem histológica com a aplicação da técnica de Masson, onde a FI está destacada em cor azul “intenso”.



Figura 2.3: Interstício com fibrose. O azul “intenso” é a fibrose no interstício. Elaboração do Autor (2022).

Um dos motivos da disseminação do uso da técnica de Masson na análise de biópsias renais para avaliação da FI é que essa técnica pode apresentar pouca variação na intensidade de cor e minimizar o aparecimento de objetos falsos que aparecem devido a defeitos nas imagens, também conhecido como efeito artefactual (Menn-Josephy et al., 2016).

Depois de aplicado o corante, a análise é feita com o uso de um microscópio, através do qual um examinador humano deve estimar a FI, em porcentagem, de todo o fragmento (Berchtold et al., 2017).

A análise da FI depende da interpretabilidade humana e pode apresentar variabilidade no grau de reprodutibilidade, atribuído às peculiaridades e complexidades do processo visual de interpretação de imagens (Farris et al., 2014). Como a análise humana da FI não é precisa, é comum certos procedimentos serem usados para melhorar a precisão do diagnóstico como, por exemplo, o uso da Classificação de Banff (Farris et al., 2014) e a aplicação do Coeficiente Kappa de Cohen (Cohen, 1960b; Mchugh, 2012; Glen, 2014).

2.1.4 O Corante tricrômico de Masson

O nome “tricromo” identifica técnicas de coloração em que dois ou mais corantes aniônicos⁴ são usados em conjunção com um heteropoliácido: ácido fosfomolibdico ou fosfotúngstico. Esses ácidos são compostos cristalinos solúveis em água e álcool. Eles podem ser incluídos em soluções de corante ou aplicados a seções sequencialmente, entre os tratamentos com diferentes corantes. A técnica tricrômico de Masson, desenvolvida por C. L. Pierre Masson (1880-1959), é um método no qual as estruturas são representadas em três cores: o núcleo em roxo (hematoxilina); o citoplasma é vermelho (fucsina ácida) e o colágeno é azul (azul de anilina).

Na avaliação da Fibrose Intersticial, são excluídas as áreas de vasos sanguíneos de pequeno e médio calibres, glomérulos, cápsula de Bowman e zona pericapsular por possuir colágeno fisiológico. Esse processo é complexo e, para se ter controle da intensidade da cor no citoplasma e no colágeno, a técnica é realizada em etapas (Masson, 1929; Kiernan, 2015; Rosman, 2012). A ilustração da Figura 2.4 exemplifica o resultado da aplicação da técnica na qual: (A) Interstício; (B) Túbulo; (C) Célula mesangial; (D) Glomérulo; (E) Célula tubular; (F) Capilar do glomérulo. O Citoplasma está exibido em cor vermelho e o colágeno em cor azul.

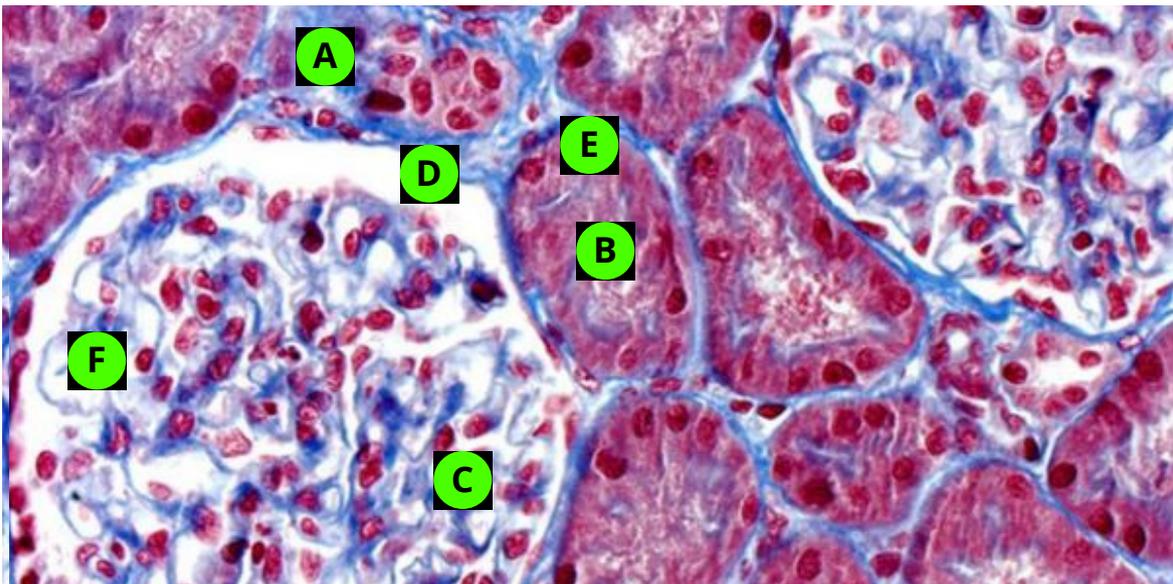


Figura 2.4: Rim: Glomerulonefrite Membranoproliferativa, Tipo II. Glomérulos com leve aumento da matriz mesangial. Corante tricrômico de Masson, 200x. Adaptado de Rosman (2012).

⁴aniônicos refere-se a agentes tensioativos que possuem grupamentos funcionais e que fornecem íons negativos

2.1.5 Classificação de Banff

A Classificação de Banff é organizada internacionalmente pela *Banff Foundation*, e tem como objetivo liderar o desenvolvimento e a disseminação da Classificação Internacional de Patologia de Transplante Renal e facilitar a pesquisa multidisciplinar e colaborativa para aprimorar sua base científica e utilidade clínica. A Classificação Banff é um consenso para a notificação de biópsias de transplantes, tendo sua concepção em 1991. Desde então, passou por revisões a cada 2 anos, proporcionando uma rápida expansão do conhecimento com inúmeras revisões científicas, limitando a subjetividade e heterogeneidade do diagnóstico (Roufosse et al., 2018).

A Classificação de Banff padroniza os diagnósticos de rejeição ao transplante renal com base em critérios histológicos (Loupy et al., 2020). As decisões clínicas, geralmente, são feitas após a integração dos diagnósticos de Banff no contexto clínico.

No entanto, a interpretação dos casos de biópsia ainda é heterogênea entre patologistas e clínicos. Os parâmetros usados na Classificação de Banff podem ser refinados e, certamente, os seriam na Convenção de Banff 2019, que foi postergada para outubro de 2022 em decorrência da pandemia da covid-19 (Roufosse et al., 2018).

Atualmente, a versão dessa classificação mais amplamente utilizada é a Classificação de Banff de 1997 (Racusen et al., 1999; Pascual et al., 2002; Nankivell e Alexander, 2010), que foi significativamente revisada em 2007 na *9th Banff Conference in Allograft Pathology* (Solez et al., 2008).

Escore de Banff para Lesão

Importante destacar que o Escore de Banff avalia a presença e o grau de alterações histopatológicas nos diferentes compartimentos das biópsias de transplante renal, focalizando principalmente, mas não exclusivamente, nas características diagnósticas vistas na rejeição. Entretanto, uso desse escore não é, por si só, suficiente para atingir as várias categorias de diagnóstico de Banff (Loupy et al., 2020; Haas et al., 2018; Loupy et al., 2017; Roufosse et al., 2018).

Parâmetros de diagnóstico adicionais, como os histopatológicos, moleculares, sorológicos e ou clínicos, podem ser necessários para determinar o diagnóstico correto. De acordo com Roufosse et al. (2018), a Classificação de Banff nunca deu uma definição precisa para áreas individuais de fibrose intersticial.

A razão para isso é que o Escore de Banff para estimar a fibrose intersticial, do inglês *Banff Lesion Score (Cortical Interstitial)*, destinava-se a refletir puramente o córtex composto de tecido fibroso, o que não corresponde necessariamente a áreas que um patologista identificaria como uma mancha de fibrose patológica. A imagem 2.5 ilustra um guia de referência para a classificação de Banff Haas et al. (2018).

A Classificação de Banff pode apresentar fatores que ocasionam discrepâncias na análise, podendo levar a divergências no diagnóstico como, por exemplo, a graduação da FI. Essa graduação é dividida em quatro categorias, baseadas no valor percentual

Conteúdo da Classificação de Banff para Patologia de Aloenxertos Renais

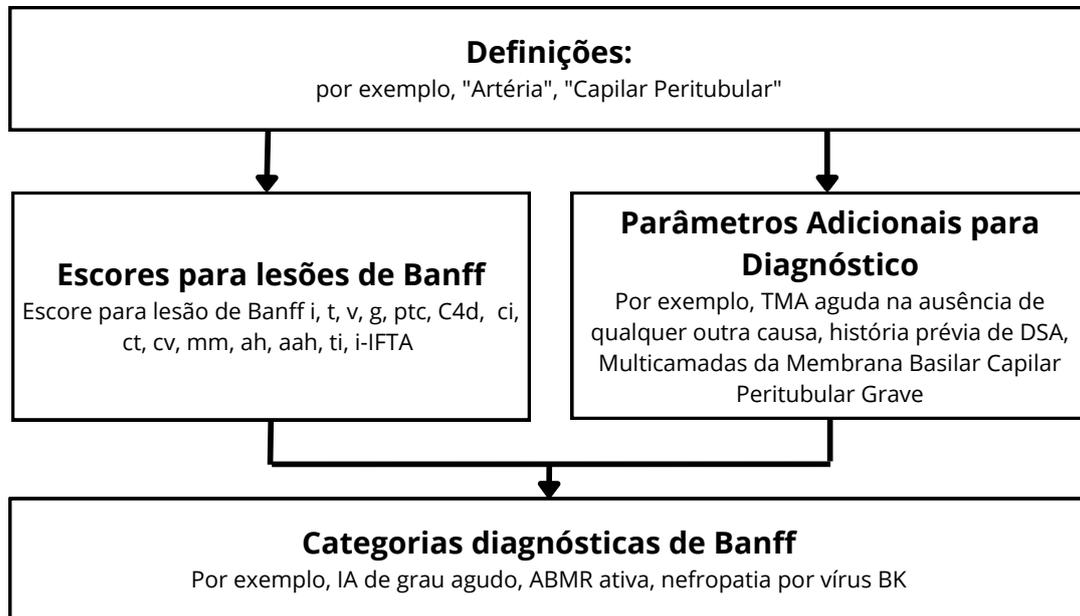


Figura 2.5: O conteúdo da Classificação de Banff de Transplante Renal pela qual a patologia pode ser inventariada com escores de lesão de Banff e Parâmetros de diagnóstico exigidos pelos algoritmos por trás das Categorias de diagnóstico Banff e chegar a um diagnóstico (Roufosse et al., 2018).

de tecido alterado pela FI em relação ao fragmento inteiro, conforme apresentado na 2.1.

Tabela 2.1: Escore de Banff para lesão por fibrose intersticial (*score Banff ci*)

Categoria	Faixa de Valor
ci0	Menor que 5%
ci1	De 6 a 25%
ci2	De 26 a 50%
ci3	Maior que 51%

Fonte: Recuperado de Roufosse et al. (2018)

Este trabalho de pesquisa se limita a olhar a Classificação de Banff para lesão por fibrose intersticial (*Banff Lesion Scores ci*), conforme apresentado na Tabela 2.1.

Para facilitar e padronizar a pontuação da lesão por fibrose intersticial (*Banff Lesion Scores ci*), a imagem conceitual na Figura 2.6 retrata, de forma similar, como o patologista deve enxergar a fibrose intersticial sem atrofia tubular (Haas et al., 2018).

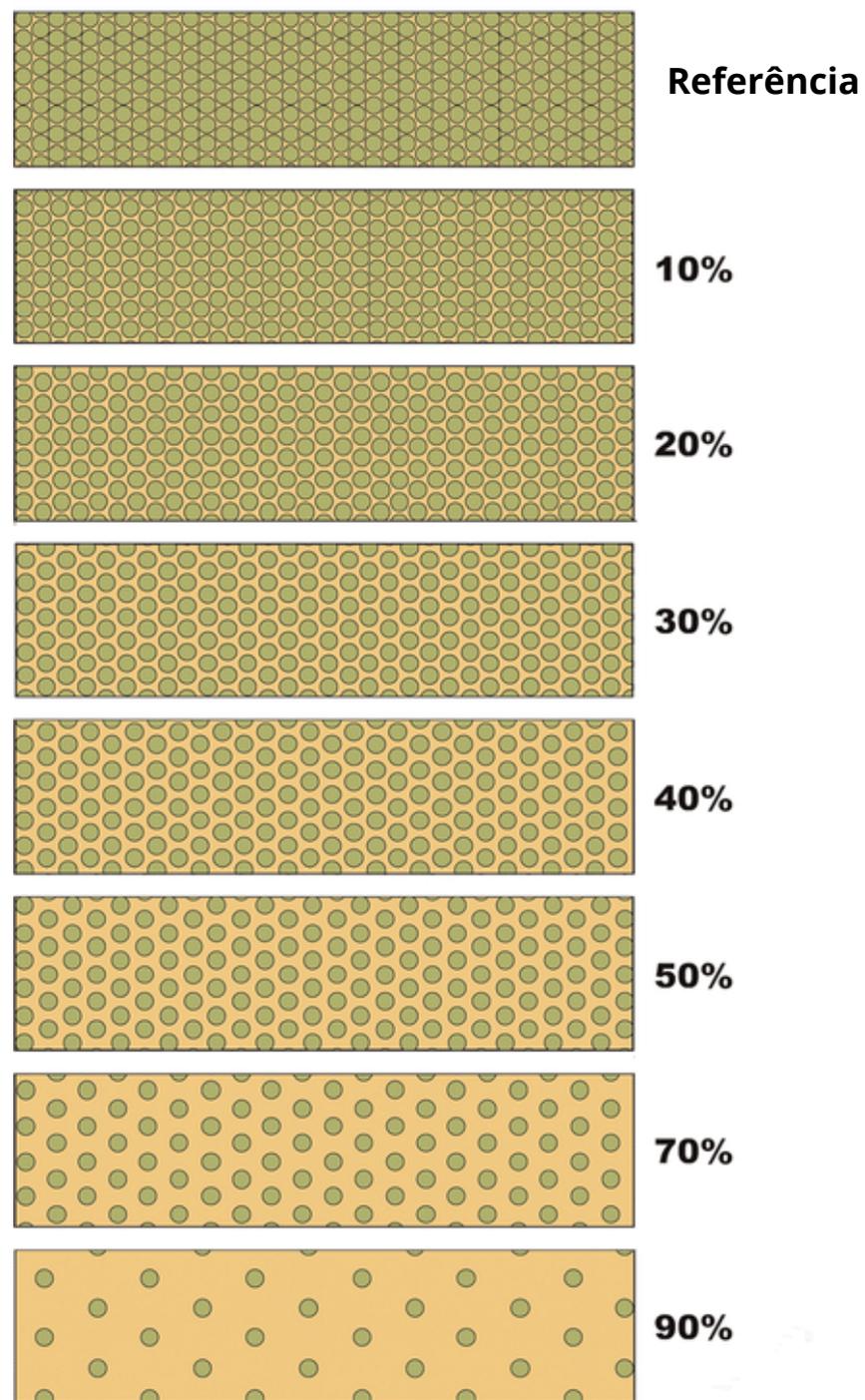


Figura 2.6: Imagem Conceitual, retratando de forma similar como o patologista enxerga a fibrose intersticial. A coloração da imagem é simbólica e não se relaciona com as cores da lâmina de biópsia renal, coradas com tricrômico de Masson (Roufousse et al., 2018).

2.2 Conceitos de computação usados neste trabalho

A seguir, serão apresentados os conceitos computacionais necessários ao entendimento deste trabalho.

2.2.1 Processamento digital de imagens

O processamento digital de imagens é capaz de preparar a imagem para facilitar a aplicação dos métodos de Visão Computacional para realização da análise e interpretação do conteúdo da imagem e obter informações relevantes. O processamento digital de imagens se refere ao processamento por um computador. Uma imagem pode ser definida como uma função bidimensional, $f(x, y)$, em que x e y são coordenadas espaciais (largura x altura), e o valor de f em qualquer ponto dessas coordenadas (x, y) corresponde ao brilho ou intensidade nesse ponto. A imagem é considerada como digital, quando x, y e os valores de intensidade de f são quantidades finitas e discretas (Gonzalez, 2009).

Segundo Marques Filho e Neto (1999), a área de processamento de imagens é classificada em duas categorias distintas: 1) o aprimoramento de informações pictóricas para interpretação humana; e 2) a análise automática por computador de informações extraídas de uma cena. A primeira categoria é classificada como processamento de imagens, enquanto que a segunda categoria é relacionada aos termos análise de imagens, Visão Computacional e reconhecimento de padrões.

Não há consenso entre os autores sobre onde termina o processamento de imagens e começam outras áreas relacionadas à análise de imagem e à Visão Computacional (Gonzalez e Woods, 2009).

Entretanto, um sistema de processamento de imagens pode ser entendido como um conjunto de técnicas de Processamento Digital de Imagens (PDI) que inclui a preparação de imagens para posterior análise e de técnicas de Análise Digital de Imagens (ADI), que inclui análise quantitativa do processo a partir do qual regiões, partículas e objetos na imagem são medidos.

Um sistema de processamento digital de imagens é constituído por um conjunto de etapas capazes de produzir um resultado a partir do domínio do problema. Essas etapas são mostradas na Figura 2.7 que ilustra a organização das duas categorias propostas por Gonzalez e Woods (2009) a respeito dos métodos cujas entradas e saídas são imagens, e métodos cujas entradas podem ser imagens, mas cuja saída é o atributo extraído dessas imagens.

A seguir, serão apresentadas breves descrições de cada uma destas etapas.

- **Aquisição da Imagem** - Etapa responsável pela captura da imagem e conversão em adequada representação para o processamento digital.
- **Filtragem e realce da imagens** - Refere-se a técnicas que permitem acentuar características relevantes de uma imagem, de forma que o resultado seja mais

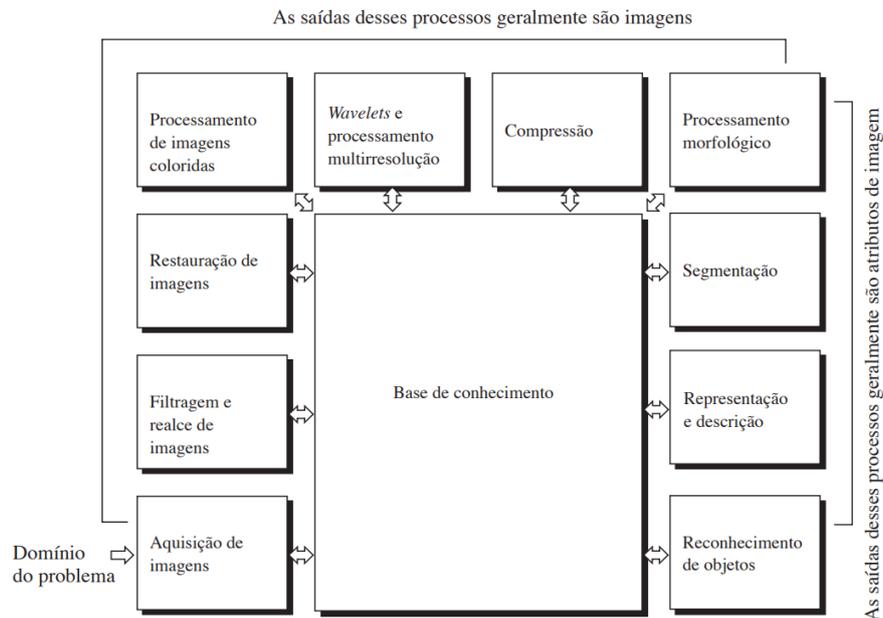


Figura 2.7: Etapas fundamentais em processamento digital de imagens (Gonzalez e Woods, 2009).

adequado do que a imagem original para uma aplicação específica.

- **Restauração de imagens** - Consiste no uso de técnicas matemáticas ou probabilísticas para remoção ou redução de imperfeições ocorridas durante a etapa de aquisição.
- **Processamento de Imagens coloridas** - Etapa que trata de uma área na segmentação de imagens, a qual vem ganhando considerável importância com o aumento significativo de imagens digitais na Internet. Além disso, a cor também pode ser usada como base para a extração de características de interesse em uma imagem.
- **Wavelets e processamento multiresolução** - Nesta etapa, as *wavelets* constituem os fundamentos para a representação de imagem. Gonzalez e Woods usaram as *wavelets* em seu livro *Processamento Digital de Imagens 3a ed.* (2009) para explicar a compressão de dados de imagens e para a representação piramidal, na qual as imagens são subdivididas sucessivamente em regiões menores.
- **compressão de Imagens** - Etapa que lida com técnicas de redução do armazenamento necessário para salvar uma imagem, ou a largura de banda necessária para transmiti-la.
- **Processamento morfológico** - Etapa que trabalha com ferramentas para a extração de componentes de imagens úteis na representação e descrição da forma.

- **Segmentação** - Etapa que extrai e identifica áreas de interesse presentes na imagem. Essa etapa geralmente se baseia na detecção de descontinuidades (bordas) ou semelhanças (regiões) na imagem.
- **Representação e descrição** - Em geral, o resultado de um estágio de segmentação, que normalmente são dados primários em forma de pixels, corresponde tanto à fronteira de uma região (por exemplo, o conjunto de pixels que separa uma região da imagem de outra) quanto a todos os pontos dentro da imagem.
- **Reconhecimento de objetos** - Nesta etapa, um identificador ou etiqueta é atribuído a objetos da imagem, dependendo das características fornecidas por seus descritores. O processo de interpretação é atribuir um significado a todos os objetos reconhecidos.

2.2.2 Visão Computacional

Visão Computacional é área da ciência que estuda a forma de como possibilitar uma máquina enxergar o meio à sua volta e extrair informações importantes a partir de imagens capturadas por dispositivos como *webcams*, *scanners*, *smartphones*, câmeras de vídeo, entre outros (Ballard e Brown, 1982; Backes e Junior, 2019).

Para Barelli (2018), a Visão Computacional pode ser compreendida como um complemento da visão biológica. Alguns cientistas concentram suas pesquisas na compreensão de como o córtex cerebral funciona e, em seguida, incorporam essas ideias à Visão Computacional. É o que pesquisadores do Instituto de Tecnologia de Massachusetts (MIT) definem como ensinar um computador a ver as coisas como seres humanos (Greenemeier, 2008).

Para trazer à luz o entendimento de como uma máquina pode ser capaz de “enxergar”, é importante compreender o processo de funcionamento do olho humano. Este processo inicia-se no par de olhos (cada um tem cerca de 25 milímetros de diâmetro) que capta a luz com comprimento de onda que varia de de 400 nm (violeta) até 700 nm (vermelho) via córnea, um tecido transparente que cobre a parte anterior de cada olho de forma semelhante a um vidro totalmente transparente.

Seguindo seu caminho, a luz atravessa o humor aquoso, alcançando a pupila, formada pela íris e que funciona como um diafragma, atingindo imediatamente o cristalino que funciona como uma lente de focalização, convergindo então os raios luminosos para um ponto sobre a retina.

Na retina, mais de cem milhões de células fotossensíveis (sensíveis à luz) transformam a luz em impulsos conduzidos ao cérebro através do nervo óptico para o córtex visual localizado na parte posterior do cérebro, uma das partes mais complexas do sistema de processamento das imagens recebidas pelo nosso par de olhos.

Após compreender, ainda que superficialmente o processo de funcionamento do olho humano fica mais nítido que mesmo com toda a sofisticação e décadas de pesquisa, os computadores ainda não podem competir com a natureza, onde o cérebro humano

classifica objetos vistos de forma rápida e precisa o suficiente para que as pessoas possam interpretar o que é visto (Greenemeier, 2008).

Vencida as dificuldades de possibilitar uma máquina enxergar os objetos à sua volta, tarefas complexas de classificação e interpretação de imagens podem ser realizadas com maior precisão e velocidade do que os seres humanos seriam capazes de alcançar (Brito et al., 2021).

Segundo Barelli (2018), além da neurobiologia que estuda o sistema biológico de visão, diversas áreas do conhecimento humano vem contribuindo com o avanço da Visão Computacional como é o caso da Inteligência Artificial, do Processamento de Imagens e do Reconhecimento de Padrões. Por exemplo, algumas áreas de estudo que estão diretamente ligadas à Visão Computacional são mostradas na figura 2.8.

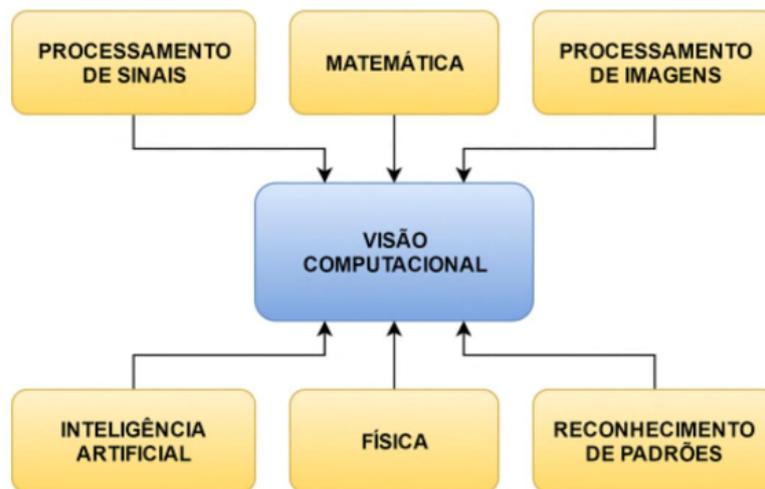


Figura 2.8: Campos de estudo interligados a Visão Computacional (Barelli, 2018).

Com tantos campos de estudos relacionados a Visão Computacional, é importante encontrar maneiras de agrupar os processos envolvidos e facilitar a evolução desses estudos. É possível agrupar esses processos relacionados à Visão Computacional em três níveis de operações: de baixo, de médio e de alto nível (Gonzalez e Woods, 2009).

As operações de baixo nível são processos básicos, como pré-processamento para reduzir o ruído da imagem, melhorar o contraste e o brilho. Os processos de médio nível incluem segmentação (separar regiões de imagens ou objetos) e a descrição do objeto, deduzido-o a uma forma que facilite o processamento computacional. Já os processos de alto nível, tentam reconhecer conjuntos de objetos em imagens (Gonzalez e Woods, 2009).

A seguir, é descrito brevemente as etapas que compõem um sistema de Visão Computacional (Marques Filho e Neto, 1999; Gonzalez e Woods, 2009; Barelli, 2018), conforme retrata a Figura 2.9.

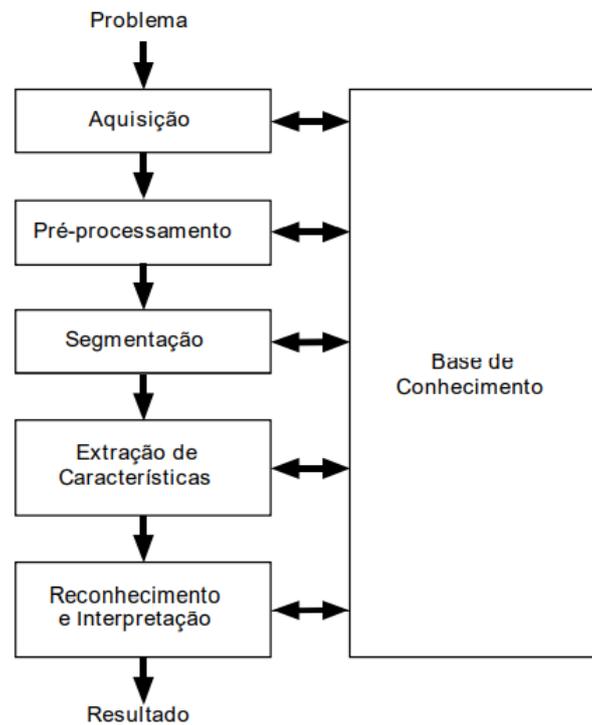


Figura 2.9: Exemplo de um Sistema de Visão Computacional e suas principais etapas (Marques Filho e Neto, 1999)

Aquisição

Nesta etapa, a imagem é capturada e representada de forma computacional para ser interpretada na etapa de Pré-processamento (Marques Filho e Neto, 1999).

Pré-processamento

O objetivo desta etapa é ajustar e otimizar os dados de imagem adquiridos. Para isso, diversas técnicas podem ser aplicadas, como remover ruídos, girar a imagem, aplicar filtros e trocar o espaço de cores (Marques Filho e Neto, 1999).

Segmentação

De acordo com Shapiro et al. (2001) e Marques Filho e Neto (1999), a segmentação de imagens tem basicamente dois objetivos: decompor a imagem em partes, ou seja, separar os objetos de interesse do *background* da imagem original para posterior análise, e realizar uma mudança de representação.

Sendo assim, pode-se dizer que segmentação de imagem se refere ao processo de separação, extração de informações de imagens ou vídeos e de destacar padrões de segmentos, criando sub-regiões que possuem alguma semelhança entre si, dividindo a imagem em objetos e fundo (Gonzalez, 2009).

Esse processo é fundamental em aplicações relacionadas à análise de imagem, uma vez que é responsável por identificar áreas de interesse, contribuindo com a resolução do problema apresentado como, por exemplo, a identificação de doenças renais (Claro et al., 2018).

Embora a segmentação de imagens seja tarefa simples de descrever, sua implementação é muito complexa de realizar, mesmo sendo uma área com grande quantidade de pesquisas em desenvolvimento como nos diz Gonzalez (2009). Dentre algumas das técnicas de segmentações temos a limiarização, crescimento de regiões, filtragem do espaço escalar e métodos estatísticos (Gonzalez e Woods, 2009; Taher et al., 2018; Pandey e Lalchhanhima, 2020). Na segmentação é possível fazer a classificação, localização, classificação + localização, detecção e segmentação de objetos em uma imagem (Parmar, 2018). Na figura 2.10 é mostrado alguns resultados usando técnicas de segmentação de imagens.

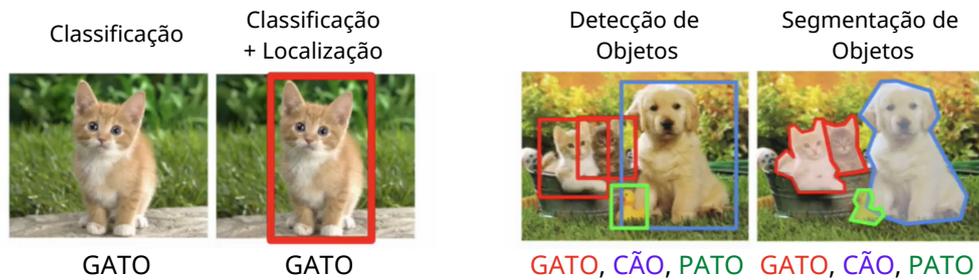


Figura 2.10: Detecção e segmentação de objetos (Granatyr, 2022).

- **Classificação:** A finalidade na classificação é determinar a que categoria um conjunto de dados pertence. O resultado da predição é um valor categórico (discreto e não ordenado) e uma etiqueta que caracteriza um grupo no conjunto de dados (Ham, J., 2011).
- **Classificação + Localização:** Na classificação e localização, cada imagem contém apenas um objeto, enquanto que na localização, além de determinar a que classe pertence uma imagem, também é necessário definir sua localização relativa na imagem, empregando uma caixa delimitadora (*bounding box*). Logo, é possível que haja um ou muitos objetos de classes iguais, ou diferentes (Parmar, 2018).
- **Detecção de Objetos:** Na detecção de objetos, todos os objetos na imagem precisam ser detectados, demarcados com caixas de delimitadoras e etiquetados com a classe a que o objeto pertence. Assim, é possível que haja um ou múltiplos objetos de classes iguais, ou diferentes (Parmar, 2018; Karagiannakos, 2019).
- **Segmentação de Objetos:** Existem dois subtipos de segmentação de objetos: segmentação semântica e segmentação de instância. A segmentação semântica

tem por objetivo classificar todo e qualquer pixel da imagem para sua respectiva classe. Diferente da detecção de objetos, não se usa caixas delimitadoras, todos os pixels do objeto são marcados, não há distinção entre instâncias, dessa forma, em uma imagem que traz duas flores da mesma espécie, elas serão tratadas como o mesmo objeto. Já a segmentação de instância funciona como a segmentação semântica, mas faz distinção entre instâncias, por exemplo, em uma imagem que contém duas flores da mesma espécie, elas serão tratadas como objetos do mesmo tipo, contudo de instâncias diferentes (Parmar, 2018; Karagiannakos, 2019).

Extração de características

Representa, por meio de valores, uma imagem ou partes dela. Esses valores são as características básicas que representam os atributos presentes na imagem. Nesta etapa, a entrada ainda é uma imagem, sendo a saída um conjunto de dados que correspondem a imagem de entrada (Marques Filho e Neto, 1999).

Reconhecimento e Interpretação

Nesta etapa, que também é conhecida por Reconhecimento de padrões, as imagens são classificadas e rotuladas de acordo com suas características semelhantes (Marques Filho e Neto, 1999).

Base de Conhecimento

Segundo Marques Filho e Neto (1999), as etapas descritas acima pressupõem o conhecimento sobre o problema a ser resolvido. Ou seja, para reconhecer objetos em uma imagem, um sistema de visão computacional necessita de uma base de conhecimento desses objetos. Esta base de conhecimento pode ser implementada diretamente no código, através, por exemplo, de um sistema baseado em regras, ou esta base de conhecimento pode ser aprendida a partir de um conjunto de amostras dos objetos a serem reconhecidos utilizando técnicas de aprendizado de máquina (Marengoni e Stringhini, 2009).

2.2.3 Cores do Espectro Visível

Segundo a Comissão Internacional de Iluminação (CIE, 1932; Guild, 1931), todas as cores podem ser formadas com a mistura das cores vermelha, verde e azul, onde cada uma, chamadas de cores primárias, estando no comprimento de onda, podem ser detectadas pelo olho humano. O olho humano é sensível às ondas curtas (420-440 nm), médias (530-540 nm), e longas (560-580 nm). Assim, o comprimento de onda para as cores primárias, segundo a CIE é:

- $R(\lambda)$ luz vermelho com comprimento de onda de 700 nm.
- $G(\lambda)$ luz verde com comprimento de onda de 546 nm.
- $B(\lambda)$ luz azul com comprimento de onda de 435.8 nm.

O modelo ou espaço de cores CIE foi criado em 1931 (CIE, 1932; Guild, 1931). Esse espaço de cores foi resultado de uma série de experimentos feitos no final da década de 1920 por William David Wright, usando dez observadores, e John Guild, usando sete observadores (Smith e Guild, 1931; Wright, 1929).

O espaço de cores CIE é um sistema de mapeamento que usa valores tristimulus (uma combinação de três valores de cores que estão próximas ao vermelho, verde e azul). Quando esses valores são combinados, eles podem reproduzir as cores que o olho humano consegue perceber. Os resultados experimentais foram combinados na especificação do espaço de cores conhecido como CIE RGB, a partir do qual o espaço de cores CIE XYZ foi derivado. O CIE RGB/XYZ são espaços de cores aditivas e servem como base para a definição de cores de forma padronizada (Smith e Guild, 1931; Wright, 1929).

2.2.4 Modelos de cores

Um espaço de cores (também conhecido como modelo de cores ou sistemas de cores) é uma especificação de um sistema de coordenadas e um subespaço dentro desse sistema no qual, cada cor é representada por um único ponto (Gonzalez e Woods, 1999).

Há vários modelos de cor, como por exemplo os modelos de cor são RGB, HSV, LaB, YCbCr, YIQ e CMYK. Segundo Gonzalez e Woods (1999), o espaço de cores RGB (red, green, blue — vermelho, verde, azul) é mais utilizado para monitores e dispositivos de captura de imagem coloridos, como câmeras de vídeo. o espaço de cores CMY (*cyan, magenta, yellow* — ciano, magenta, amarelo) e o espaço de cores CMYK (*cyan, magenta, yellow, black* — ciano, magenta, amarelo, preto) são utilizados em dispositivos de impressão colorida. Já o espaço de cores HSI (*hue, saturation, intensity* — matiz, saturação, intensidade), que também é conhecido como HSV, está relacionado à forma como os seres humanos descrevem e interpretam as cores. Este espaço de cores, tem ainda a vantagem de separar as informações de cor e de escala de cinza da imagem, sendo mais adequado para muitas das técnicas de processamento de imagens. Na Figura 2.11 é possível observar uma ilustração do espaço de cores RGB e HSV.

2.2.4.1 O espaço de cores RGB

Neste espaço, as cores são exibidas em porcentagens de vermelho, verde e azul. Ex: Azul em RGB = $[0, 0, 255]$ = Azul e $[255, 255, 0]$ = Amarelo. Saturando ao máximo essas cores obtém-se o branco $[255, 255, 255]$. Para o inverso se obtém o preto $[0, 0, 0]$. Juntando uma cor com outra consegue-se uma terceira cor.

Na Figura 2.12 é possível observar uma ilustração dos espaço de cores RGB e CMYK. Enquanto o espaço de cores RGB exhibe as cores aditivas, o espaço de cores CMYK

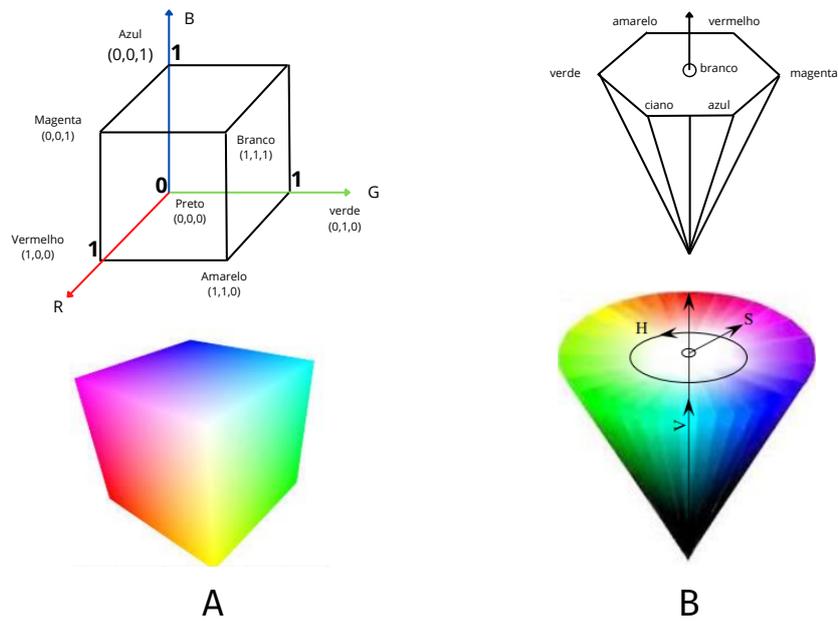


Figura 2.11: (A) Esquema de cor RGB e (B) Esquema de cor HSV (Gonzalez e Woods, 1999; Gomes J.; Velho, 1994).

exibe tanto as cores aditivas (usando a luz) quanto às subtrativas (no papel) (George e Beard, 2012).

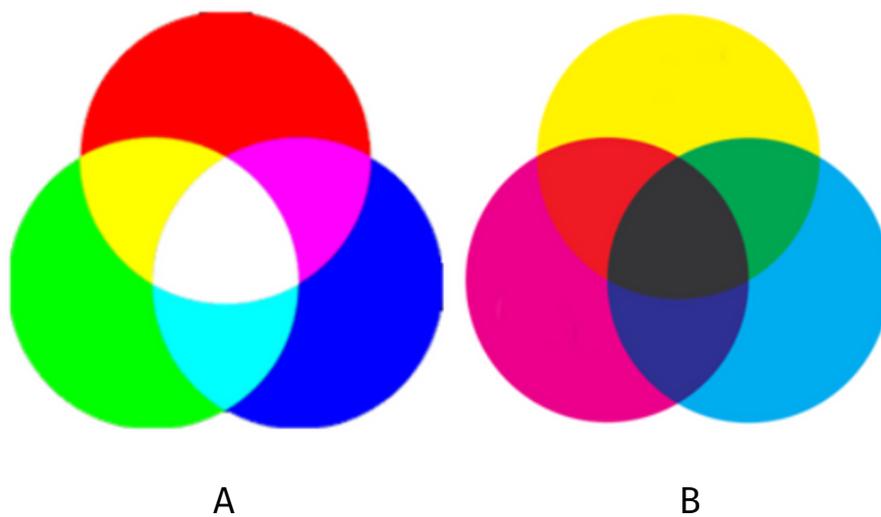


Figura 2.12: (A) Espaço das Cores RGB aditivas e (B) espaço das cores CMYK subtrativas (George e Beard, 2012).

O espaço de cores HSV

O espaço de cores HSV opera de modo a separar o componente de intensidade (*value*) da informação de cor e utilizar os componentes de matiz (*hue*) e saturação (*saturation*) relacionados à percepção humana. Deste modo, é mais indicado para manipulação (Gonzalez e Woods, 2009). No espaço de cores HSV, a tonalidade ou matiz corresponde à medida do comprimento médio de onda da luz que reflete ou emite um objeto. Isto define a cor, atingindo valores de entre 0° e 360° .

Entretanto, em algumas aplicações esse valor é definido em valor escalar de 0 a 180 ou ainda entre percentuais de 0 e 100%. A Saturação é a “pureza” da cor, quanto menor esse valor, mais cinza será a imagem, e quanto maior o valor, mais “pura” será a imagem, ou seja, com as cores mais nítidas e sem a presença do branco. A saturação representa a faixa de comprimentos de onda em torno do comprimento de onda médio em que a energia é refletida ou transmitida, variando de 0 a 100%. O valor ou luminância define o brilho da cor, atingindo valores de 0 a 100% (George e Beard, 2012).

Usar esses atributos ao trabalhar com imagens permite acentuar a diferença entre um objeto e outro, de modo a reforçar a intenção desejada. O oposto é o caso onde há ambiguidade: os elementos se parecem e não é possível identificar a diferença entre eles. Este esquema de cor está exemplificado na Figura 2.13.

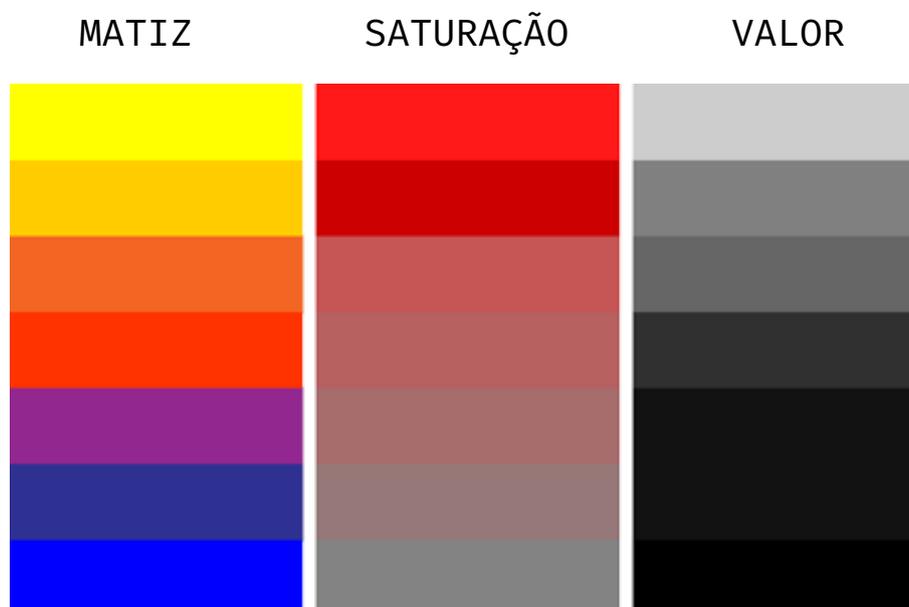


Figura 2.13: Esquema de cor Matiz, Saturação e Valor (George e Beard, 2012).

2.2.5 Segmentação de imagens baseada na cor

A segmentação é um processo que divide uma imagem em regiões. Esse processo tem sido facilitado pelo avanço tecnológico que vem ocorrendo na área de processamento de imagens, desde *Jet Propulsion Laboratory* (Pasadena, California - EUA) em 1964, quando imagens da lua transmitidas por uma sonda Ranger eram processadas por computador para corrigir vários tipos de distorção que vinham da câmera de TV acoplada à sonda (Marques Filho e Neto, 1999).

Muito desse avanço tecnológico deve-se aos esforços empreendidos por pesquisadores que têm favorecido o surgimento e o melhoramento de vários algoritmos e técnicas para segmentação de imagens (Gonzalez e Woods, 2009). Por exemplo, Deng et al. (1999) apresentam uma abordagem para segmentação de imagem. Inclusive, é relatado nesse estudo a complexidade em identificar regiões de imagem contendo padrões de textura de cor. Na Figura 2.14, é mostrado o esquema do algoritmo proposto, chamado JSEG.

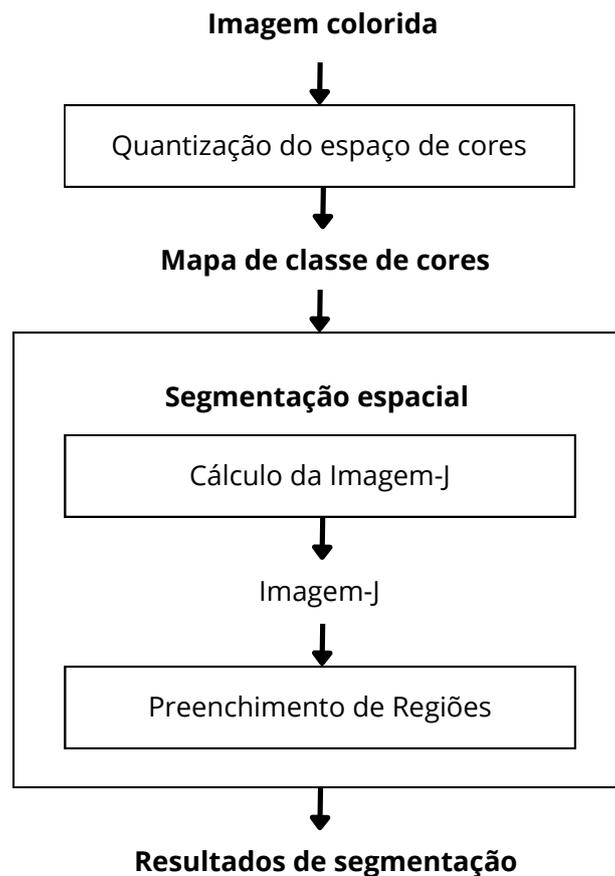


Figura 2.14: Esquema do algoritmo JSEG (Deng et al., 1999).

O trabalho de Cheng et al. (2001) apresenta um levantamento sobre técnicas de segmentação de plantas baseadas em imagens. A pesquisa discute brevemente o pré-processamento de imagens e foca na segmentação. Os autores apresentam três algoritmos primários, a saber: (i) segmentação baseada em índice de cor; (ii) segmentação baseada em limiarização; (iii) segmentação baseada em aprendizado de máquina. Discutem, ainda, o desempenho de segmentação de abordagens baseadas em índices de cores, a partir de estudos da literatura entre 2008 a 2015.

Essas técnicas de segmentação de imagens comumente são empregadas em combinação com o conhecimento do domínio e aplicadas a um determinado problema, colaborando na identificação das diferenças entre os objetos de uma imagem. Essas diferenças permitem que o algoritmo interprete pixels adjacentes e os agrupe em regiões; seu emprego pode determinar o sucesso ou fracasso de toda análise (Gonzalez e Woods, 1992).

Segmentando uma imagem por cor é possível determinar o percentual de contribuição de cada cor nesta imagem. Para determinar o percentual de uma cor na imagem, uma classe é atribuída a um determinado intervalo de valores de *pixel* na imagem, com um rótulo correspondente. Um mapa composto por esses rótulos de camada é chamado de mapa de camadas ou mapa de classes. Por exemplo, a Figura 2.15 mostra uma imagem de 14x14 *pixels* que ilustra como é o mapa de classes do algoritmo JPEG (abreviação de *Joint Photographic Experts Group*, uma organização internacional que padronizou o formato no fim dos anos 80 e no começo dos anos 90). À esquerda, a imagem ampliada, e à direita, a imagem em tamanho real e a ilustração da etiqueta de camada, resultante de uma imagem de 14x14 *pixels*, onde o rótulo são as primeiras letras da imagem que corresponde às cores: verde (G), vermelho (R), amarelo (Y) e azul (B).

As técnicas ou métodos de segmentação de imagens morfométricas dependem principalmente de limites de intensidade de cor predefinidos. Por outro lado, os métodos baseados em aprendizado de máquina incorporam informações estruturais além dos recursos de cor, geralmente resultando em desempenho de segmentação muito mais estável e melhor resistência à variação de coloração Liu et al. (2022b). O que importa a esta dissertação de mestrado é a segmentação de imagem por cor.

Partindo do ponto em que a segmentação é um processo que particiona uma imagem em regiões ou objetos (Gonzalez e Woods, 2000), pode-se compreender a segmentação de imagem por cor como um processo de classificação de cores, tendo como base o atributo cor dos *pixels*. A segmentação por cor pode ser realizada em espaço de cores específicos como por exemplo os espaços de cores HSV e RGB (Gonzalez e Woods, 2000).

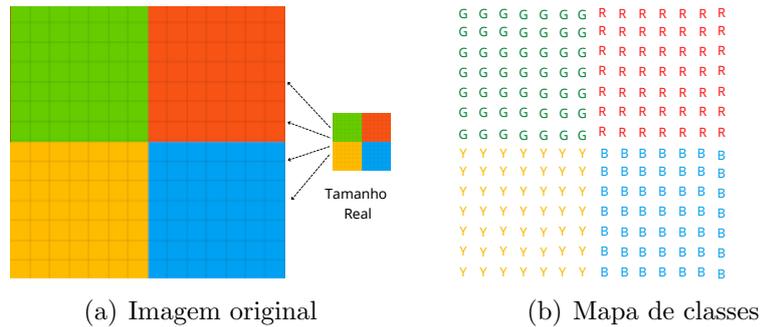


Figura 2.15: (A) Imagem original (B) e Mapa de classes. Adaptado de Deng et al. (1999).

Segmentação no espaço de cores HSV

Para segmentação de uma imagem baseada em cores o ideal é se pensar no espaço de cores HSV, uma vez que a cor neste espaço é representada pela matiz, sendo a saturação usada como máscara para isolar as regiões de interesse (Gonzalez e Woods, 2009). Na Figura 2.16 é demonstrado um exemplo de segmentação realizada no espaço de cores HSV. Nessa Figura, de acordo com Gonzalez e Woods (2009), o interesse é segmentar a região avermelhada. As Figuras 2.16(b) e (d) são as imagens dos componentes HSV.

Ao comparar as Figuras 2.16(a) e (b) (Gonzalez e Woods, 2009) chamam a atenção para os valores relativamente altos de matiz da região de interesse, evidenciando que as cores estão no lado azul-magenta do vermelho. Aplicando uma binarização na imagem de saturação utilizando um limiar igual a 10% é gerado uma máscara mostrada na Figura 2.16(e). Todo pixel cujo valor seja maior que o limiar foi atribuído o valor 1 (branco) e aos demais pixels foi atribuído o valor 0 (preto).

A máscara com a imagem do matiz é mostrada na Figura 2.16(f) e o histograma da imagem resultante é mostrada na Figura 2.16(g). Analisando o histograma é possível notar que os valores de interesse (valores altos) estão agrupados na extremidade superior da escala de cinza, próximo ao valor 1,0. O resultado é mostrado na Figura 2.16(h) (Gonzalez e Woods, 2009).

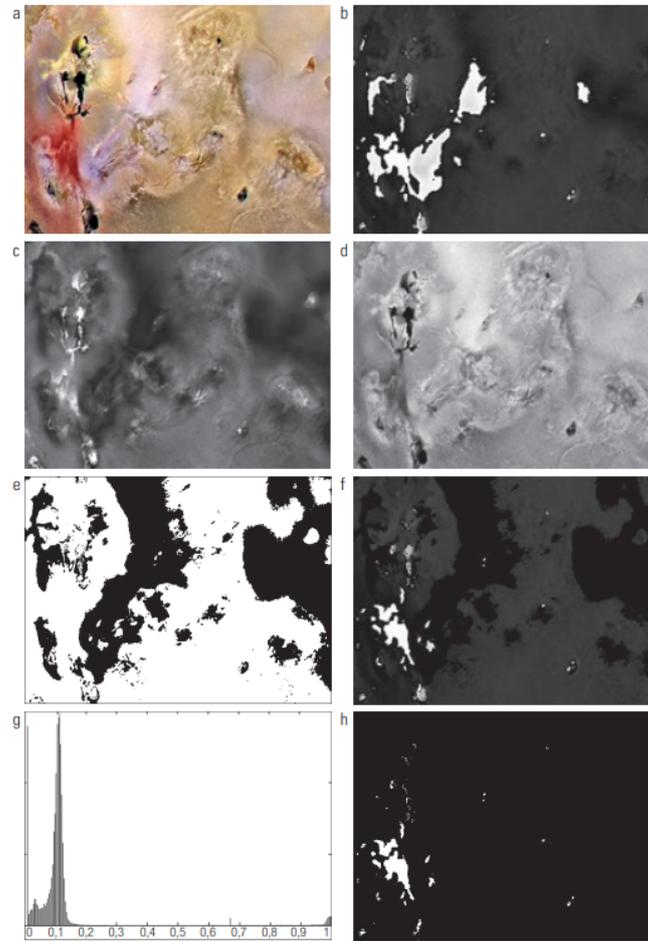


Figura 2.16: Segmentação de imagem no espaço HSV. (a) Original. (b) Matiz. (c) Saturação. (d) Intensidade. (e) Máscara de saturação binária (preto = 0). (f) Produto de (b) por (e). (g) Histograma de (f). (h) Segmentação dos componentes vermelhos em (a) (Gonzalez e Woods, 2009).

Embora essa segmentação exemplo não seja perfeita, Gonzalez e Woods (2009) relata que por experimentação as regiões mostradas em branco na Figura 2.16(h) “são o melhor que esse método pode fazer na identificação dos pontos avermelhados da imagem original”.

Segmentação no espaço de cores RGB

Segundo Gonzalez e Woods (2009) os melhores resultados de segmentação por cor são obtidos utilizando o espaço de cores RGB, embora o próprio autor reconheça que o espaço de cores HSV é mais intuitivo. O autor sustenta essa afirmação afirmando que o método é direto e usa um exemplo no qual o objetivo seria segmentar objetos de uma determinada faixa de cores em uma imagem RGB. Simplificando, o vetor a RGB para a cor média é denominado **a**. Neste exemplo, pretende-se classificar

os *pixels* RGB de uma imagem qualquer como pertencentes ou não a um intervalo de cores. Para a comparação é necessário uma medida de semelhança e Gonzalez e Woods (2009) cita a medida euclidiana como uma das medidas mais simples para este propósito. Para tanto, considera-se \mathbf{z} um ponto arbitrário no espaço de cor RGB, sendo \mathbf{z} similar a \mathbf{a} se a distância entre eles for menor que um limiar determinado em D_0 .

A equação para determinar a distância euclidiana entre \mathbf{z} e \mathbf{a} é demonstrada na Equação 2.1 onde R, G e B indicam os componentes RGB dos vetores \mathbf{a} e \mathbf{z} . Gonzalez e Woods (2009) explica que o conjunto de pontos \mathbf{z} que satisfaz $D(z, a) \leq D_0$ forma uma esfera sólida de raio D_0 . Os pontos na esfera atendem ao critério especificado na cor e os pontos fora da esfera não satisfazem a este critério. Deste modo ao codificar esses dois conjuntos de pontos na imagem, produz a imagem segmentada binária.

$$D(z, a) = \|z - a\| \quad (2.1)$$

Desenvolvendo:

$$D(z, a) = [(z - a)^2 (z - a)]^2$$

$$D(z, a) = [(z_R - a_R)^2 + (z_G - a_G)^2 + (z_B - a_B)^2]^2$$

No entanto, é alertado por Gonzalez e Woods (2009) que este método de segmentação apresenta um alto custo computacional e, sugere como meio termo usar um cubo em vez de esfera para delimitar o conjunto de cores a serem segmentadas. Segundo Gonzalez e Woods (2009), utilizando o método do cubo o centro seria \mathbf{a} e suas dimensões ao longo de cada um dos eixos de cor seriam escolhidos em proporção ao desvio padrão das amostras ao longo de cada um dos eixos.

No exemplo, a segmentação de uma cor arbitrária é segmentada se ela estiver na superfície ou dentro do cubo que em termos computacionais é mais simples determinar se essa cor está dentro ou fora de um cubo do que determinar se essa cor está dentro ou fora de um delimitador esférico ou elíptico (Gonzalez e Woods, 2009).

Vejamos o exemplo da Figura 2.17(a), onde existe uma faixa de cores avermelhadas que deseja-se segmentar. Neste exemplo é calculado o vetor médio de \mathbf{a} usando pontos de cor contidos na figura. Posteriormente, calcula-se o desvio padrão dos componentes RGB (vermelho, verde e azul). Um cubo é centralizado em \mathbf{a} e as dimensões de cada eixo desse cubo foram escolhidas como 1,25 multiplicado pelo desvio padrão dos dados ao longo dos eixos. Na Figura 2.17(b), é apresentado o resultado da codificação sendo o branco quando o ponto está dentro do cubo e preto, quando estiver fora (Gonzalez e Woods, 2009).

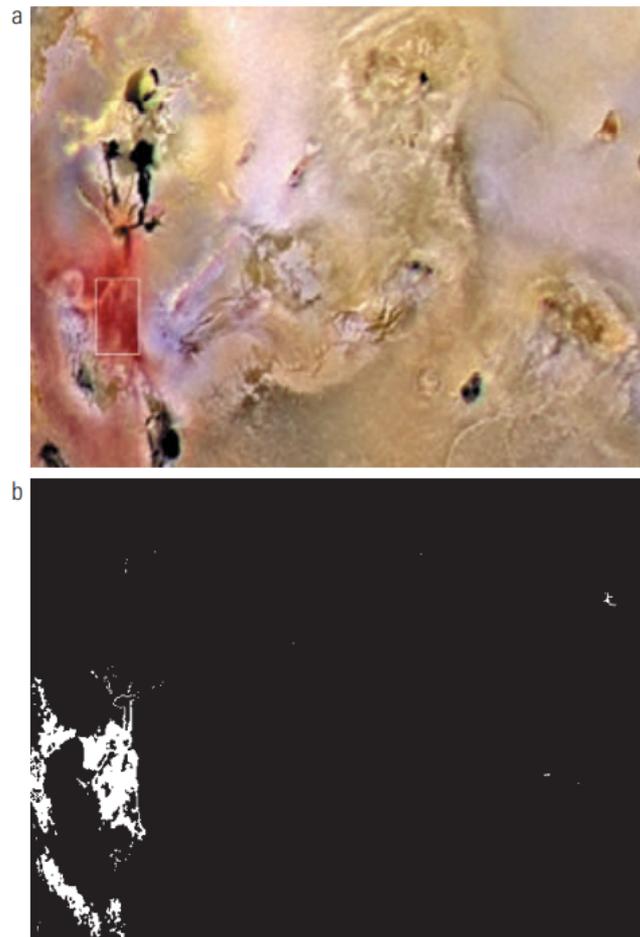


Figura 2.17: Segmentação no espaço de cores RGB. (a) Imagem original com as cores de interesse delimitadas por um retângulo. (b) Resultado da segmentação no espaço de cores RGB. Compare com a Figura 2.16 (h) (Gonzalez e Woods, 2009).

2.2.6 O problema da Segmentação de Imagens por Cor

Ao longo de milhões de anos, o olho humano evoluiu para detectar uma parte do espectro eletromagnético conhecido como luz visível, que percebemos como cor. Entretanto, definir a cor não é algo simples como pode aparentar. Inclusive, desde os primórdios da civilização humana as cores são estudadas.

Nesse sentido, uma das primeiras contribuições foram os estudos de Empédocles⁵ (492 - 431 a.C.), Demócrito⁶ (460 - 370 a.C.) e Platão⁷ (428 - 347 a.C.).

⁵Empédocles - Filósofo e pensador pré-socrático grego. É conhecido por ser o criador da teoria cosmogênica dos quatro elementos clássicos que influenciou o pensamento ocidental até meados do século XVIII.

⁶Demócrito - Filósofo grego do período pré-socrático e agrupado na escola atomista. Julgava que todos os elementos do universo eram compostos de átomos.

⁷Platão - Filósofo e matemático grego da antiguidade, considerado um dos principais pensadores

Entre os primeiros registros de estudos de representação das cores, pode-se destacar o trabalho de Pitágoras (570 - 500 a.C.), que criou um espaço de cor semicircular⁸ conectando as notas da escala musical composta por tons e semitons com planetas representados por certas cores, conforme ilustrado na Figura 2.18.

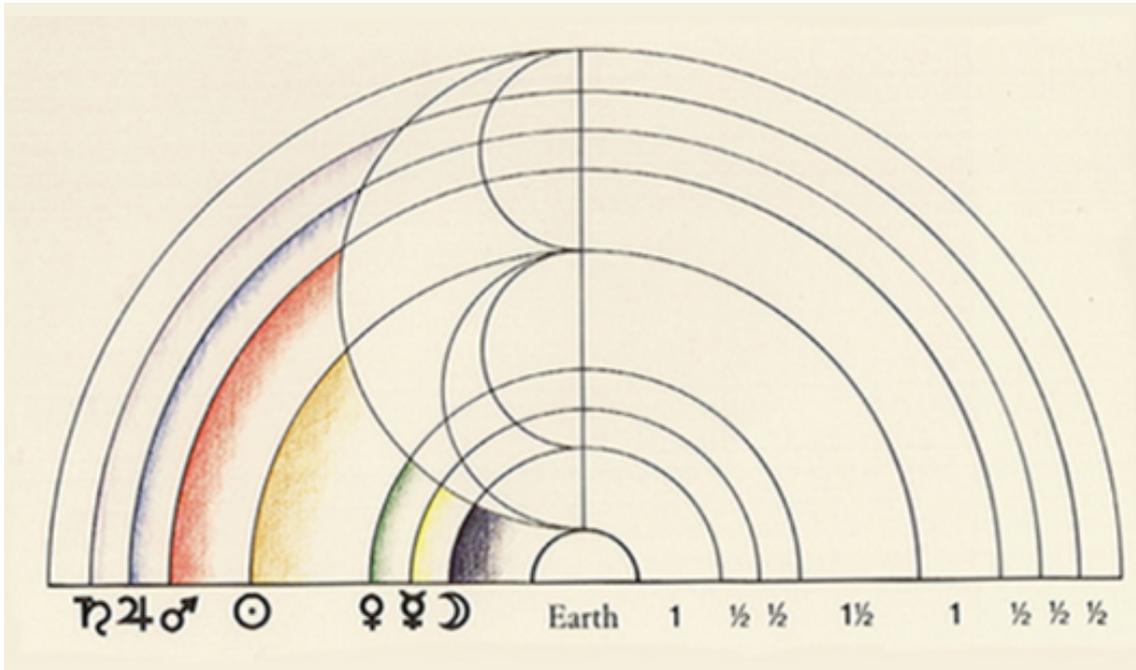


Figura 2.18: Espaço de cores idealizado por Pitágoras (Colorsystem, 2020).

No século XVII, Isaac Newton inseriu o ponto de vista da física no estudo das cores, quando concluiu que para haver cor, é necessária a refração do raio de luz no espectro cromático. Um novo e importante passo ao entendimento das cores só veio no século XVIII com a Doutrina das Cores de Goethe.

Enquanto Newton estudou a cor sob a ótica da física, Goethe adentrou aspectos relacionados ao olho humano abordando a dimensão fisiológica. Essa abordagem contrapõe-se à visão de Newton sobre o fenômeno da luz e das cores, baseando-se em aspectos empíricos que a teoria de Newton não contemplou (Brito e De Oliveira Reis, 2016; Goethe, 2018)

Conforme diz Goethe (2018), percebe-se que a chave para a compreensão dos fenômenos cromáticos está na subjetividade do sentido da visão, representada pelo olho humano. Basicamente, Goethe traduz a formação das cores em um disco cromático de seis cores opostas, enquanto o disco cromático de Newton tem sete cores. Para Goethe, fisiologicamente os fotorreceptores cromáticos presentes nos olhos, enviam pelo nervo ótico os sinais recebidos ao cérebro. Este processo gera além das cores recebidas, as cores opostas.

da história da filosofia

⁸Espaço de cor é um modelo matemático usado para descrever cada cor a partir de equações

Para Brito e De Oliveira Reis (2016) há muita polêmica sobre o trabalho de Goethe com relação às questões físicas da luz, a energia ondulatória e a polaridade das cores. Para entender o fenômeno de formação das cores, não é conveniente isolar o estudo fisiológico de como o olho vê e do comportamento físico da luz, até porque a percepção humana desse fenômeno físico das cores pode variar de pessoa para pessoa.

Goethe e Newton tinham uma relação, não necessariamente oposta, mas de certa forma complementar, pois a teoria de Goethe incluía aspectos que a teoria de Newton não compreendia, como a interpretação de cores a partir de estímulos sensoriais. (Brito e De Oliveira Reis, 2016).

Assim, em primeiro momento, a identificação de uma cor é algo relativamente simples para o campo da computação, mas, ao incluir na equação um humano, a cor de um objeto visto será determinada pela frequência das ondas que esse objeto refletir e da forma como o olho humano é capaz de detectar.

Tendo como base a construção do diálogo sobre aspectos da teoria de Goethe e Newton e entendendo que ainda há muita pesquisa a ser feita, este trabalho limita-se a compreender a cor como uma percepção visual direta, provocada pela ação de um feixe de fótons sobre células especializadas da retina e que transmitem, através de informação pré-processada no nervo óptico, impressões para o sistema nervoso.

2.3 Conceitos complementares

A necessidade de classificar um conjunto de dados em um determinado número de categorias, é algo comum em diversas situações, como por exemplo em contexto clínico, e nesta classificação vários tipos de vieses podem ocorrer (Tinsley e Weiss, 1975). Para minimizar esses vieses, muitas vezes utiliza-se mais de um juiz para categorizar os mesmos dados, analisando-se então seu grau de concordância e, consequentemente, a confiabilidade da classificação. A esse respeito, Hoyle et al. (2002) alerta para a ocorrência de possíveis enviesamentos, nomeadamente por parte de quem realiza essa classificação. Além disso, examinadores diferentes, avaliando o mesmo caso podem chegar a conclusões distintas (Prado, 2004).

Uma forma de constatar a concordância entre as aferições realizadas pelos juízes é o uso de estatísticas de concordância e, uma dessas estatísticas é o coeficiente kappa (Cohen, 1960a; Chen e Krauss, 2004).

2.3.1 Coeficiente Kappa

O coeficiente de kappa criado por Cohen, na década de 60 (Cohen, 1960a) é rotineiramente usado para avaliar a concordância entre examinadores numa escala nominal (variáveis são simplesmente “nomeadas” e não existe uma ordem específica) enquanto que a variação deste coeficiente proposta por Cohen (1968), assume que as categorias são ordenadas e contabilizadas para saber a distância do julgamento

entre os dois examinadores, sendo indicado para os casos em que as categorias da avaliação são ordinais, como por exemplo ruim, médio e ótimo (Prado, 2004).

Um “examinador” neste contexto pode ser um único indivíduo, um conjunto de indivíduos que classificam os itens coletivamente ou um sistema computacional (Vergni et al., 2021; Cantor, 1996). Em outras palavras, o coeficiente de kappa pode ser compreendido como sendo a razão entre duas distâncias, ou seja, quanto mais concordantes forem os examinadores, mais próximo de 1 estará o valor da medida. O coeficiente de kappa varia de -1, em situações em que a amostra é balanceada para duas categorias de resposta e os examinadores discordam em todas as avaliações, até 1, quando os examinadores concordam em todas as avaliações (Prado, 2004; Fonseca et al., 2007).

De forma geral, quando:

- **Kappa** < 0, a concordância é mais fraca que o esperado pelo acaso;
- **Kappa** = 0, a concordância é aquela que seria esperada ao acaso;
- **Kappa** = 1, a concordância perfeita existe.

O funcionamento teórico da estatística kappa, proposto por (Cohen, 1960a) é ilustrado na situação em que dois examinadores avaliam o mesmo grupo de indivíduos e ambos os examinadores avaliem cada individuo de acordo a uma escala de que vai de 1 a r . Os dados relativos às avaliações dos examinadores são apresentados em tabelas de contingência $r \times r$, conforme demonstrado na Tabela 2.2.

Tabela 2.2: Exemplo de tabela de contingência $r \times r$

		Examinador 2				
Examinador 1		1	2		r	Total Linha
1		π_{11}	π_{12}		π_{1r}	$\pi_{1.}$
2		π_{21}	π_{22}	. . .	π_{2r}	$\pi_{2.}$
.	
.	
.	
r		π_{r1}	π_{r2}	. . .	π_{rr}	$\pi_{r.}$
Total Coluna		$\pi_{.1}$	$\pi_{.2}$. . .	$\pi_{.r}$	1

Nesta tabela de contingência, $r = 3$ e π_{ij} representa a parcela de indivíduos que são classificados na i -ésima categoria pelo examinador 1 e na j -ésima categoria pelo examinador 2. Já π_i representa a proporção de indivíduos classificados na i -ésima categoria pelo examinador 1 e π_j representa a proporção de indivíduos classificados na j -ésima categoria pelo examinador 2, para $i, j = 1, 2, \dots, r$.

Neste aspecto, o coeficiente de concordância de Kappa (k), sugerido por Cohen (1960a), pode ser calculado utilizando a equação 2.2 abaixo:

$$k = \frac{(po - pe)}{(1 - pe)} \tag{2.2}$$

Na qual:

$$P_0 = \sum_{i=1}^r \frac{n_{ii}}{n} \quad P_e = \sum_{i=1}^r \frac{n_{i.} * n_{.i}}{n^2}$$

P_0 : Acordo relativo observado entre os examinadores.

P_e : Probabilidade hipotética de acordo casual.

Na literatura pesquisada, há exemplos de tabelas de contigência para interpretação, que ajudam na compreensão da graduação dos valores encontrados ao aplicar a equação do coeficiente Kappa (Cohen, 1960b; Glen, 2014; Landis e Koch, 1977).

<i>Estatística Kappa</i>	<i>Força do acordo</i>
<0,00	Pobre
0,00-0,20	Pouco
0,21-0,40	Justo
0,41-0,60	Moderado
0,61-0,80	Substancial
0,81-1,00	Quase perfeito

Figura 2.19: Exemplo de medidas de acordo para dados categóricos (Landis e Koch, 1977)

Vejamos um exemplo prático em que os examinadores sejam dois médicos, A e B avaliando 100 pacientes diferentes para uma doença particular em que as células na horizontal (linhas) representam, por exemplo, os julgamentos do médico A e as células na vertical (colunas) os julgamentos do médico B.

Assim, para avaliar se o diagnóstico da condição (positiva ou negativa) dos médicos A e B está em concordância, usa-se, neste exemplo, a Tabela 2.3 que mostra uma tabulação cruzada das classificações observadas. Nessa tabela, os valores das células que representam a concordância dos testes são destacadas em vermelho.

Neste exemplo, o valor calculado para a concordância Kappa é 0,492. Na primeira coluna da tabela, é possível ver que 44 resultados foram considerados normais pelos dois médicos, sete foram considerados alterados pelo médico A e normais pelo médico B, e nove foram considerados inconclusivos pelo médico A e normais pelo médico B. Utilizando a interpretação proposta por Landis e Koch (1977) visualizada na Figura 2.19 é possível afirmar que os dois médicos apresentaram uma concordância razoável em relação aos diagnósticos, a partir dos exames de imagem. Estes valores podem ser conseguidos e revisados utilizando-se de softwares específicos, como por exemplo o Simstat, disponível através da Provalis Research (<http://www.provalisresearch.com>).

Tabela 2.3: Exemplo de aplicação do Coeficiente de Concordância de Kappa

Médico A	Médico B			Total Linha
	<i>Normal</i>	<i>Alterado</i>	<i>Inconclusivo</i>	
<i>Normal</i>	44	5	1	50
<i>Alterado</i>	7	20	3	30
<i>Inconclusivo</i>	9	5	6	20
Total Coluna	60	30	10	100
Resultados	Po	Pe	k	Interpretação
Calculados	0,700	0,410	0,492	Moderada

Por meio da página da Internet <https://www.graphpad.com/quickcalcs/kappa1/> ou <http://vassarstats.net/kappa.html> é possível calcular o coeficiente kappa online.

Embora coeficiente kappa de Cohen seja algo amplamente utilizado e recomendado, sendo o mais referido na literatura quando as variáveis são nominais, a sua aplicação, segundo Fonseca et al. (2007) é restrita às situações em que dois juízes avaliam todas as unidades de análise num determinado número de categorias. Entretanto, há situações em que é necessário recorrer-se a medidas de avaliação para múltiplos avaliadores como no exemplo de Fleiss et al. (1969) em que é sugerido com desenvolvimento do coeficiente kappa para múltiplos examinadores e sub-grupos diferentes que avaliam situações, mantendo constante o número de codificações por unidade de análise. Outras propostas vem surgindo para situações mais complexas em que se recorre a mais do que dois juízes ou para abordar melhor casos em que o acordo, ou não acordo, deriva do acaso (Fonseca et al., 2007; Schouten, 1986; Berry e Mielke Jr, 1988; Posner et al., 1990).

2.4 Trabalhos Relacionados

Apresenta-se, nesta seção, um breve resumo de trabalhos considerados relevantes quando comparados à proposta deste estudo.

Inicialmente, foram encontrados poucos estudos norteadores para a proposta deste trabalho que foi desenvolver um método semiautomático para estimação de fibrose intersticial, a partir de imagens digitais de biópsias renais, sob a perspectiva da classificação de Banff.

Como ponto de partida, iniciamos este estudo lendo o trabalho de Farris et al. (2021) no qual os autores apresentam uma revisão da literatura discutindo algoritmos de inteligência artificial, com ênfase na patologia do transplante renal. Nesse estudo é relatado que a patologia do transplante renal é frequentemente classificada pelo uso do escore de Banff. Nos trabalhos de Ginley et al. (2021) e Zheng et al. (2021) é discutida a viabilidade de métodos baseados em aprendizado de máquina para avaliar a fibrose intersticial em imagens de lâmina inteira (do inglês *Whole Slide Image* – WSI).

No trabalho de Ginley et al. (2021), foi relatado que até a data de publicação de seu trabalho em abril de 2021, não havia nenhum estudo concentrado especificamente em um classificador de lâmina inteira para replicar diretamente a avaliação de um patologista de IFTA - fibrose intersticial / atrofia tubular (interstitial fibrosis/ tubular atrophy) e glomerulosclerose em amostras de biópsia renal. Neste trabalho foram estudados meios de automatizar a identificação e quantificação de IFTA e glomerulosclerose em imagens de lâmina inteira com aprendizado de máquina.

Como metodologia usada pelos autores, um patologista renal anotou amostras de biópsia renal de 116 imagens de *slides* inteiros para IFTA e glomerulosclerose. Foram utilizados 79 *slides* de cinco instituições, para o treinamento de diferentes configurações de uma rede neural convolucional, e 17 e 20 WSIs foram utilizados como casos de testes internos e externos, respectivamente. Segundo os autores, o melhor modelo foi comparado com a entrada de quatro patologistas renais em 20 novos slides de teste. Além disso, para 87 amostras de biópsia de teste, as medidas de IFTA e glomerulosclerose feitas por patologistas e a CNN foram correlacionadas.

Como resultado, o melhor desempenho médio em todas as classes de imagem veio de uma rede *deeplab* versão 2 treinada em $40\times$. Os percentuais IFTA e glomerulosclerose derivados desta CNN alcançaram altos níveis de concordância com quatro patologistas renais. As análises patológicas e baseadas na CNN de IFTA e glomerulosclerose mostraram correlação estatisticamente significativa e equivalente com todas as variáveis de desfecho do paciente.

O estudo foi concentrado em áreas fibróticas bem demarcadas e totalmente desenvolvidas e foram restritos a casos de fibrose inicial mínima ou em evolução. Segundo Liu et al. (2022b), embora esse método facilite a anotação e ao mesmo tempo reduza a discordância entre os anotadores, critérios de seleção rigorosos podem dificultar a generalização do modelo.

A fibrose intersticial é um processo progressivo, não sendo uma norma observada na prática clínica a demarcação criteriosa da fibrose. Além disso, embora facilite a anotação e ao mesmo tempo reduza a discordância entre os anotadores, os critérios de seleção rigorosos podem dificultar a generalização do modelo conforme endossa Liu et al. (2022b).

Em resumo, no trabalho de Ginley et al. (2021), foi construído um modelo de segmentação baseado em uma rede neural convolucional para quantificar a fibrose intersticial sugerindo com base em seus resultados que ferramentas de aprendizado de máquina podem possibilitar a identificação automática de estruturas de imagem que podem ser comparáveis com a análise por especialistas humanos.

No estudo de Zheng et al. (2021), foi desenvolvido um modelo associando WSIs ao grau de fibrose intersticial. Os resultados apresentados foram imprecisos (concordância de 0,48-0,63 e kappa 0,31-0,50), conforme relata Liu et al. (2022b).

Apenas muito recentemente, no mês de maio de 2022, foi publicado o trabalho Liu et al. (2022b) no qual é apresentado um modelo baseado em aprendizagem de

máquina que realiza todo o processo de avaliação da fibrose intersticial. Segundo os autores, o estudo apresenta o primeiro processo totalmente automático baseado em aprendizado de máquina para avaliação de fibrose intersticial em biópsias renais.

O estudo aborda a importância do prognóstico preditivo a respeito da fibrose intersticial e reforça a problemática da concordância no diagnóstico dos patologistas. Como objetivo de estudo, os autores propuseram construir um modelo baseado em aprendizado de máquina que permitisse a avaliação automática e confiável da FI em biópsias renais humanas. Para tanto, incorporaram algoritmos de segmentação de córtex, glomérulo e túbulos em um único modelo para avaliar a extensão do dano de FI.

Para construção do modelo, foram usadas imagens de biópsias renais entre 2015 e 2020 coletadas no Linkou Chang Gung Memorial Hospital (LKCGMH), para treinamento/validação e teste. As lâminas foram digitalizadas no LKCGMH com um *scanner* de lâminas NanoZoomer S360 Digital C13220-01 a 400× de ampliação. Anotações foram realizadas por assistentes usando uma ferramenta baseada em Django v0.1.0 e COCO (<https://github.com/jsbroks/coco-annotator>) sendo supervisionado por um patologista renal experiente.

Ademais, os dados apresentados sugeriram que o modelo estudado pode fornecer resultados de desempenho comparável ao dos humanos. Para chegar a essa conclusão, os autores utilizaram casos não seletivos para construção do modelo proposto, considerando padrões de fibrose, e incluindo diferentes doenças renais.

Como limitação, os autores reconhecem que neste estudo os desempenhos dos algoritmos de segmentação não foram perfeitos. Possivelmente, um número maior de casos para treinamento podem alcançar melhores resultados. Outro ponto, é que a fibrose intersticial de biópsias com inflamação e/ou edema acentuados pode ser superestimada pelo modelo adotado. Esses casos apresentam desafios significativos, mesmo para patologistas renais experientes. Portanto, a verificação da inferência do modelo, preferencialmente por consenso, ainda pode ser necessária em casos difíceis. E por último, a correlação observada da função renal com a avaliação do modelo requer validação externa, especialmente a tendência de melhor correlação nos casos com fibrose intersticial moderada merecendo maiores investigações, pois esse grupo de pacientes são os de maior interesse tanto na prática quanto em ensaios clínicos.

Embora os autores relatam que o modelo desenvolvido tem o potencial de fornecer acesso fácil à avaliação de arquivos de imagens de biópsia renal em grande escala e, que também pode ser usado na prática clínica como um segundo especialista em conjunto com patologistas humanos, não conseguimos acesso aos dados para replicar e validar o estudo, mesmo tentando contato inúmeras vezes.

A Figura 2.20 é um quadro-resumo dos trabalhos considerados relevantes quando comparados à proposta deste estudo, e que foram comentados acima.

No Capítulo a seguir, será discutida a metodologia para a implementação da proposta de solução para o problema da estimativa de fibrose intersticial.

Artigo	Métrica de Qualidade https://www.scimagojr.com/	Computacional	Dataset
<p>Autoria: Farris et al. (2021) Data de Publicação: 19-11-2020 Título: Artificial Intelligence and Algorithmic Computational Pathology: Introduction with Renal Allograft Examples. Disponível: https://doi.org/10.1111/his.14304ISSN:1365-2559</p>	<p>Histopathology</p> 	<p>São discutidos métodos de aprendizado profundo, como redes neurais artificiais (ANNs)/redes neurais convolucionais (CNNs).</p>	<p>Não encontrado.</p>
<p>Autoria: Ginley et al. (2021) Data de Publicação: 04-04-2021 Título: Automated Computational Detection of Interstitial Fibrosis, Tubular Atrophy, and Glomerulosclerosis. Disponível: https://doi.org/10.1681/ASN.2020050652ISSN:10466673</p>	<p>Journal of the American Society of Nephrology ...</p> 	<p>Usou-se técnicas de rede neural Convolucional (CNN).</p>	<p>116 imagens de lâminas inteira (WSIs) foram anotadas; 79 WSIs foram usados para treinamento; 37 WSIs foram usadas para testes; 87 amostras de biópsia de teste foram usadas para medições de IFTA e glomeruloesclerose feitas por patologistas e pela CNN foram correlacionadas ao resultado do paciente usando ferramentas estatísticas clássicas.</p>
<p>Autoria: Zheng et al. (2021) Data de Publicação: 21-03-2021 Título: Deep - Learning Driven Quantification of Interstitial Fibrosis in Digitized Kidney Biopsies. Disponível: https://doi.org/10.1016/j.ajpath.2021.05.005ISSN:22132945</p>	<p>Journal of the American Society of Cytopathology</p> 	<p>Foram utilizadas técnicas de deep learning (DL) e o modelo desenvolvido, foi comparado com uma estrutura de modelagem baseada em visão computacional tradicional e aprendizado de máquina que usa descritores de imagem e recursos de textura.</p>	<p>67 WSIs coradas com tricrômico para prever o grau de IFTA. 28 WSIs foram usadas para avaliar o desempenho do modelo proposto. A precisão do modelo proposto foi de 71,8% a 5,3% para as 67 WSIs e 65,0% a 4,2% para as 28 WSIs.</p>
<p>Autoria: Liu et al. (2022) Data de Publicação: 05-05-2022 Título: End-to-end interstitial fibrosis assessment of kidney biopsies with a machine learning-based mode. Disponível: https://doi.org/10.1093/ndt/gfac143ISSN:1460-2385</p>	<p>Nephrology Dialysis Transplantation</p> 	<p>Algoritmos validados de segmentação de córtex, glomérulo e túbulos foram incorporados em um único modelo para avaliar a extensão da fibrose intersticial.</p>	<p>Para segmentação do córtex foram usadas 164 WSIs de 2015 a 2017 sendo 15 WSIs de 2018, usadas para teste; Para segmentação do glomérulo foram usadas 1519 WSIs de 2015 a 2017 sendo 15 WSIs de 2018, usadas para teste; Para segmentação do túbulo foram usadas 22 WSIs de 2015 a 2017 sendo 50 patches (de 10 WSIs) de 2018, usadas para teste.</p>

Figura 2.20: Quadro-Resumo dos trabalhos considerados relevantes quando comparados à proposta deste estudo, e dos datasets utilizados. Elaboração do Autor (2022).

Capítulo 3

Metodologia

O projeto PathoSpotter (<http://pathospotter.uefs.br/>) vêm buscando criar ferramentas computacionais que auxiliem a atividade dos patologistas. Neste caminho, uma destas ferramentas trata da avaliação da Fibrose Intersticial (FI), a partir de imagens de lâminas completas de biópsias renais (do inglês *Whole Slide Imaging–WSI*), sob a perspectiva da Classificação de Banff.

Visto que o objetivo principal desta pesquisa foi desenvolver um método computacional para estimação de Fibrose Intersticial, atendendo em primeira instância a demanda do projeto PathoSpotter, este Capítulo descreve os passos envolvidos na construção da solução proposta.

A ideia geral da metodologia desenvolvida é ilustrada na Figura 3.1.

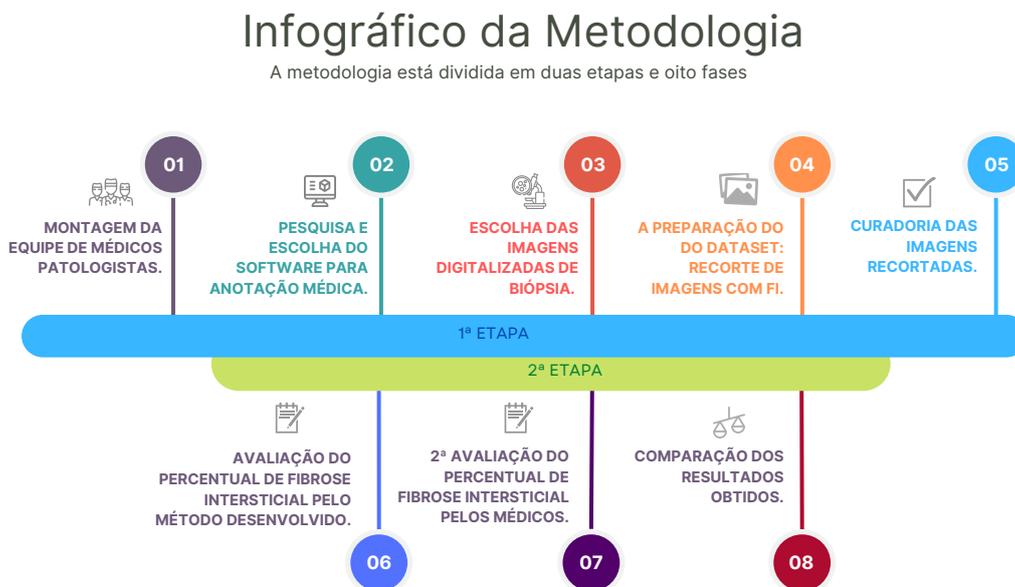


Figura 3.1: Metodologia. Elaboração do Autor (2022)

A seguir, a descrição de cada uma das fases da metodologia.

3.1 Fase 1: Montagem da equipe de patologistas.

A primeira fase compreendeu na montagem de uma equipe de médicos patologistas, demonstrada na Tabela 3.1.

Tabela 3.1: Equipe de Médicos Patologistas.

Nome do Médico (a)	Perfil
Dr. José Nathan Andrade	Mestrando patologia humana pela UFBA, Médico Patologista desde 2022, Infolaudo & Medicina Diagnóstica, Itajaí, SC.
Dr. Mateus Scarabelot Medeiros	Médico Patologista, Fellowship em Anatomia Patológica e Citopatologia DASA, São Paulo, SP.
Dr. ^a Suelen Cunha Santana	Médica Patologista, Laboratório Studart, Salvador, BA.
Dr. ^a Bianca Carla Azevedo de Souza	Mestre em patologia humana pela FIOCRUZ, Médica Patologista, Rede D'Or Salvador e Hospital Aristides Maltez. Salvador, BA.
Dr. ^a Isabela Soares Pimenta	Médica Patologista, Imagepat Sudoeste, Vitória da Conquista, BA.

A equipe de médicos patologistas foi criada com propósito de identificar corretamente o percentual de fibrose intersticial presente em lâminas de biópsia, bem como permitir a comparação do resultado do trabalho do humano e do método proposto.

3.2 Fase 2: Pesquisa e escolha do *software* para anotação médica.

Quando foram inicializados os estudos para a realização desta pesquisa, no início de 2021, o projeto *PathoSpotter* não dispunha de uma ferramenta colaborativa para anotação de imagens médicas de grandes dimensões, que era uma atividade imprescindível para a montagem do *dataset* de imagens. Segundo Bankhead et al. (2017), *scanners* de *slides* inteiros podem gerar rapidamente imagens 2D ultragrandes (por exemplo, imagens com resolução de 2160p, que significa que são 3.840 linhas horizontais e 2.160 linhas verticais de pixels), nas quais cada plano pode conter dados não compactados de até 40 GB ou mais.

Para trabalhar com essas imagens médicas de grandes dimensões (pelo menos 40GB), e partindo do entendimento que *software* colaborativo é um aplicativo que apoia o trabalho em grupo, simultânea e coletivamente, iniciaram-se as buscas por uma ferramenta gratuita, e colaborativa, para anotação de imagens médicas de grandes dimensões, resultando na Tabela 3.2.

Após avaliação dos recursos dos *softwares* levantados, o Cytomine foi adotado pelo projeto PathoSpotter como sistema de anotação e processamento de imagens mé-

Tabela 3.2: *Softwares* para auxílio à anotação de imagens médicas que foram encontrados e avaliados.

Ferramenta	Plataforma	Colaborativo/ Multiusuário	Desenvolvedor
Cytomine	Web	Sim / Sim	Montefiore Institute (2010)
CVAT	Web	Não / Sim	Sekachev et al. (2019)
COCO	Web	Não / Sim	Brooks (2019)
LabelMe	Web	Não / Sim	Science e Laboratory (2008)
QuPath	Desktop	Não / Não	Bankhead (2017)
ImageJ	Desktop	Não / Não	Rasband (2011)

dicas. Assim, uma versão do sistema foi instalada em um servidor proprietário e disponibilizada via *WEB*¹ aos integrantes autorizados do Pathospotter.

O Cytomine tem uma licença que concede direitos de uso, incluindo o direito de relicenciar. Esse sistema utiliza modernas tecnologias *web* e contêineres Docker², bancos de dados geoespaciais e aprendizado de máquina para promover a colaboração ativa e distribuída, facilitando a manipulação de imagens em larga escala (tamanho e quantidade).

Seguindo o padrão de arquitetura REST (*Representational State Transfer*) proposto por Roy Thomas Fielding em sua tese de PhD (Fielding, 2000; Liu et al., 2022a; Silva e Lampkowski, 2022), o Cytomine é considerada uma plataforma web RESTful que interaja com clientes externos desenvolvidos usando diferentes linguagens de programação (Python, Javascript, Java).

A Figura 3.2 é um exemplo de arquitetura do *software* Cytomine.

No exemplo da Figura 3.2 é demonstrado o fluxo de trabalho para segmentação de tumores em lâminas de tecido inteiro de câncer de pulmão de H&E³, a saber: (a) As imagens são carregadas usando *Cytomine-WebUI* ou clientes remotos. (b) As imagens e os dados relacionados são armazenados pelo *Cytomine-Core* e pelo *Cytomine-Image Management System*. (c) Uma vez carregadas, as imagens ficam disponíveis para outros usuários de acordo os direitos de acesso. (d) os usuários anotam de forma colaborativa, regiões de interesse nas imagens. (e) As anotações nas imagens podem ser filtradas e conjuntos de anotações podem ser exibidos ou recuperados por meio de uma API⁴. (f) Algoritmos podem explorar essas anotações.

¹<http://pathospotter-cytomine-core.bahia.fiocruz.br/>

²Docker é um conjunto de produtos de plataforma como serviço que usam virtualização de nível de sistema operacional para entregar *software* em pacotes chamados contêineres.

³Coloração H&E ou coloração HE ou, ainda, coloração hematoxilina-eosina é a principal técnica de coloração de tecidos em histologia. Por meio desta técnica, pode-se diferenciar partes basófilas e acidófilas, ou eosinófilas

⁴*Application Programming Interface (API)*, é um conjunto de rotinas e padrões estabelecidos por um software para a utilização das suas funcionalidades por aplicativos que não pretendem

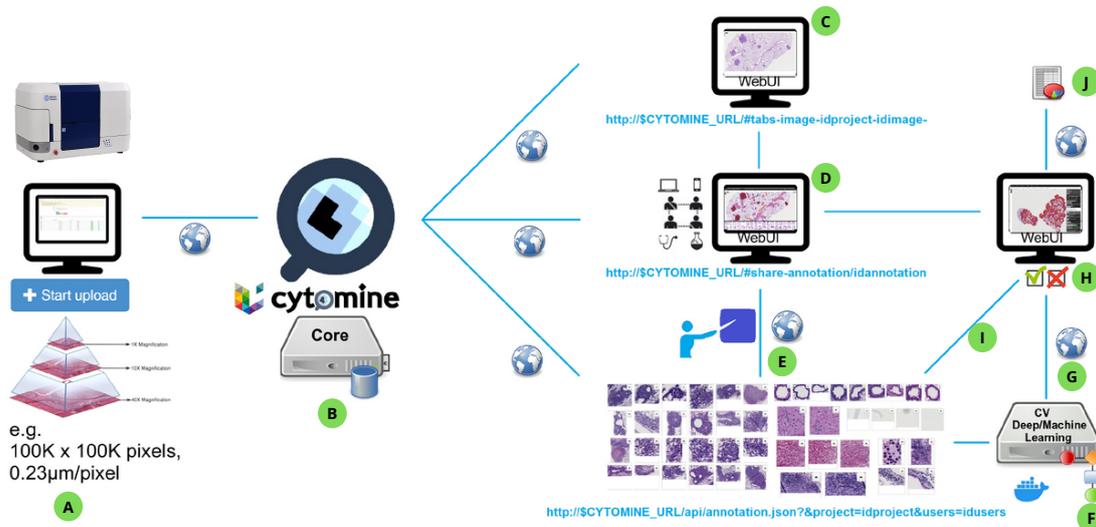


Figura 3.2: Arquitetura do *software* Cytomine (Montefiore Institute, 2010).

(g) É possível o uso de algoritmo ou modelo de reconhecimento externo ser usado remotamente para uma tarefa. (h) O usuário pode revisar anotações automáticas e anotações de outro usuário, usando as ferramentas de revisão *Cytomine-WebUI*. (i) As anotações revisadas podem, eventualmente, ser reutilizadas para refinar e reaplicar o modelo de reconhecimento. (j) Uma vez que as anotações de imagem são validadas por um especialista, os resultados finais podem ser exportados em formatos padrão (Montefiore Institute, 2010; Marée et al., 2016).

No *Cytomine* versão 3.0.3, é permitido *uploads* de arquivos de imagem de até 100 GB. Outros detalhes, como por exemplo o tipo de imagens suportadas, podem ser visualizados na documentação do *Cytomine*: <https://github.com/cytomine/Cytomine-IMS> e <https://doc.cytomine.org/>.

Além de atender as demandas levantadas pelos pesquisadores do PathoSpotter, que no caso era possibilitar o trabalho colaborativo com imagens médicas de grandes dimensões, o Cytomine também foi selecionado por fornecer suporte e toda documentação para que um programa externo interaja com ele através de bibliotecas de clientes de uma API⁵. A API do Cytomine⁶ promove a conexão entre um software proprietário desenvolvido pelo usuário e os serviços do Cytomine.

Todos os arquivos e as orientações necessárias para uso deste projeto estão no repositório do Cytomine no GitHub⁷.

envolver-se em detalhes da implementação do software, mas apenas usar seus serviços.

⁵<https://doc.cytomine.org/>

⁶<https://doc.uliege.cytomine.org/>

⁷<https://github.com/LACAD/mestrado-marcos>

3.3 Fase 3: Escolha das imagens digitalizadas de biópsia.

Para o desenvolvimento da pesquisa, foram utilizadas lâminas (Figura 3.4) extraídas de um acervo criado ao longo de diversos anos pela equipe do Dr. Washington Luis Conrado dos Santos, pesquisador do Laboratório de Patologia e Biointervenção do Centro de Pesquisas Gonçalo Moniz da Fundação Oswaldo Cruz (LPBI/CPqGM/FIOCRUZ).

As lâminas de biópsias renais foram preparadas usando o tricrômio de Masson (TM), usado principalmente para caracterizar e discriminar diferentes tecidos conjuntivos e outros componentes (Masson, 1929). Esse corante histológico reage com músculo liso e queratina em rosa a vermelho, colágeno em azul ou verde, e fibras elásticas em preto. A saber, a FI (Fibrose Intersticial) reage em cor azul neste corante, devendo ser excluídas do diagnóstico áreas de vasos sanguíneos de pequeno e médio calibres, glomérulos, cápsula de Bowman e zona pericapsular (Menn-Josephy et al., 2016; Masson, 1929; Kiernan, 2015; Rosman, 2012).

As lâminas de biópsia foram digitalizadas pelo aparelho *Pathobin USB C-Mount Camera*, acoplado a um microscópio óptico (*Olympus BX40 Clinical Microscope*) e usando o *software* conhecido como *PathobinTools*. O *scanner* apresenta-se na Figura 3.3(a) e o microscópio óptico na Figura 3.3(b).



(a) Pathobin USB C-Mount Camera



(b) Olympus BX40 Clinical Microscope

Figura 3.3: Aparelhos usados para obtenção de Imagens Médicas: (a) Pathobin USB C-Mount Camera; (b) Olympus BX40 Clinical Microscope (Pathobin, 2022) e (Microscope, 2022).

Vale ressaltar que o projeto PathoSpotter, ao qual esse trabalho está vinculado, já recebeu aprovação dos Comitês de Ética para Pesquisa Envolvendo Seres Humanos do Instituto Gonçalo Moniz (FIOCRUZ), protocolos nº 188/09 e Nº 1.817.574, e da UEFS protocolo nº 2.637.620.

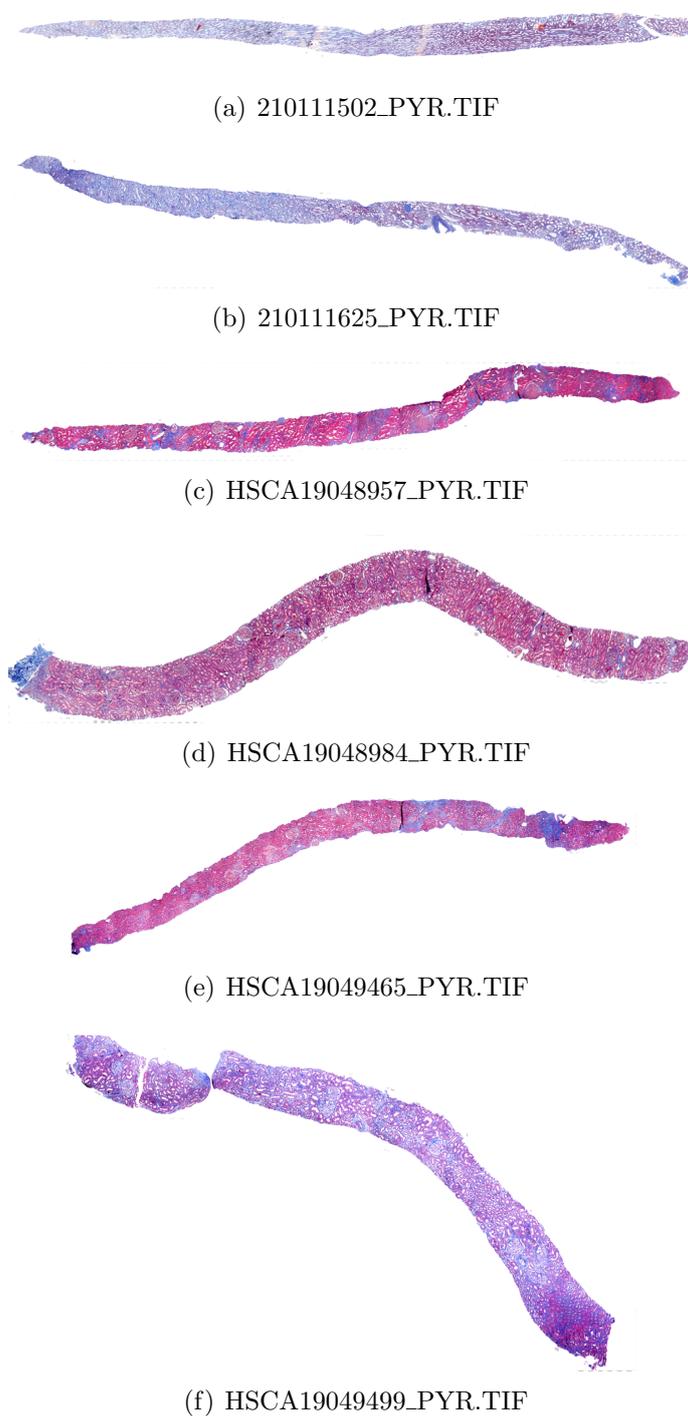


Figura 3.4: Imagens completas de biópsia renal (do inglês (*Whole Slide Imaging-WSI*) com visão panorâmica do fragmento (coloração TM). Elaboração do Autor (2022)

3.4 Fase 4: Preparação do *Dataset*

Esta fase envolveu a preparação do *dataset* de imagens das áreas de interesse (fibrose intersticial) extraídas das WSIs renais mostradas na Figura 3.4.

Utilizando o Cytomine, cada especialista da equipe de médicos Tabela 3.1, identificou regiões de interesse, ou seja, identificou áreas afetadas por fibrose intersticial, resultando em novas imagens com a anotação do valor percentual de fibrose intersticial em relação a toda imagem. Este processo gerou um *dataset* com recortes das lâminas originais. O fluxo deste processo é visualizado na Figura 3.5.

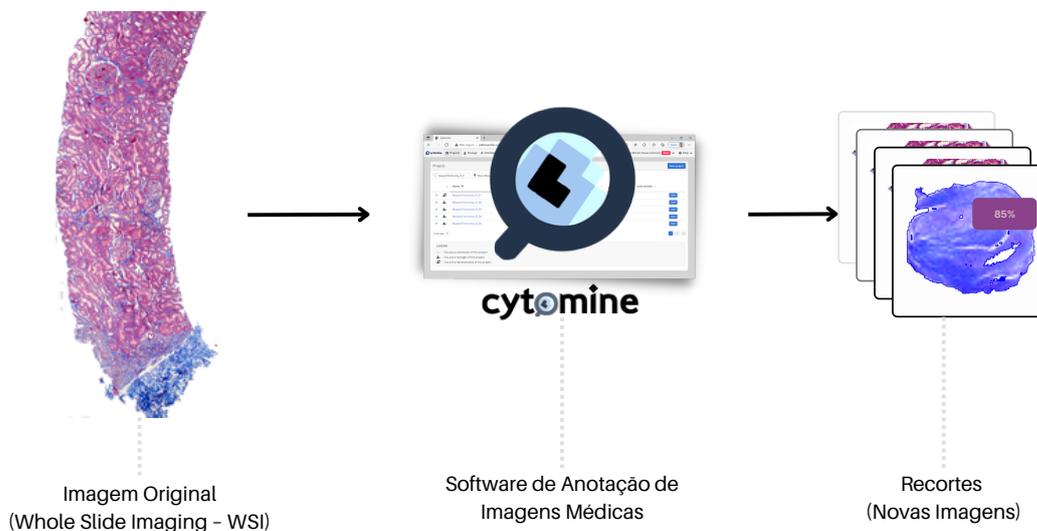


Figura 3.5: Ilustração do processo de envio de uma imagem digitalizada de lâmina de biópsia ao Cytomine, e da identificação pelo médico da área de interesse e da quantificação da fibrose intersticial, gerando uma nova imagem (recorte). Elaboração do Autor(2022).

Cada médico patologista analisou cada uma das seis lâminas preparadas com tri-crômio de Masson, produzindo um *dataset* com 344 recortes de imagens (*crops*) de tamanhos variados gravados em formato PNG, com a informação da quantidade de fibrose intersticial (anotação), não sendo obrigatório que as áreas analisadas e anotadas pelos médicos patologistas fossem as mesmas.

3.5 Fase 5: Curadoria do *Dataset*

O trabalho dos médicos patologistas realizado na quarta fase resultou em um *dataset* com 344 recortes de imagens. Deste *dataset* foram removidos 35 registros que apresentavam algum tipo de inconsistência, como por exemplo, erros artefactuais, erros de anotação ou baixa qualidade na imagem (Figura 3.6).

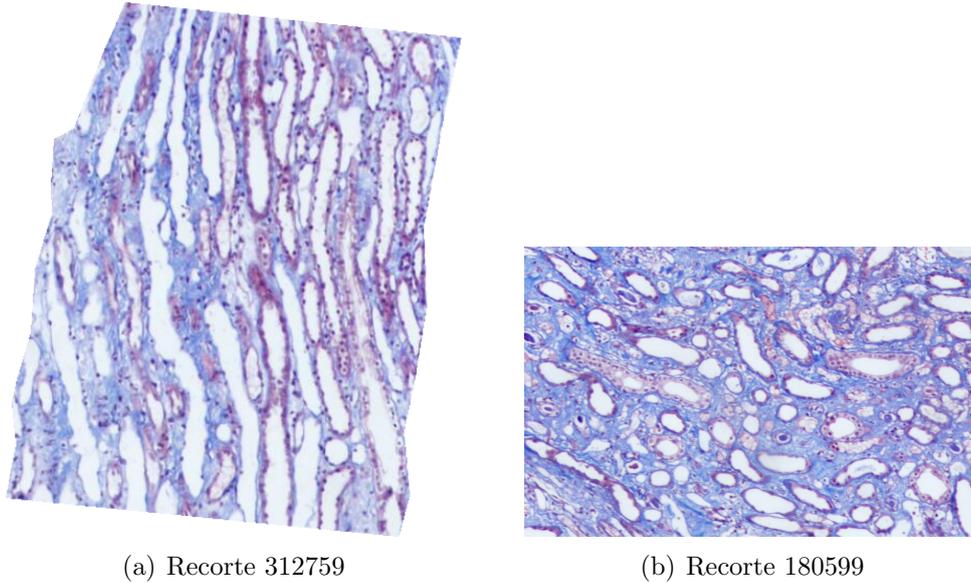


Figura 3.6: Exemplo de recortes de imagens removidos do *dataset* porque são áreas da medula renal, que não devem ser consideradas para a avaliação de fibrose intersticial. Elaboração do Autor(2022).

O conjunto de recortes de imagens final (309 recortes de imagens) compoem o *dataset* disponibilizado ao método computacional na sexta fase e a todos os médicos patologistas na sétima fase (Figura 3.7).

3.6 Fase 6: Avaliação do percentual de fibrose intersticial pelo método desenvolvido.

Nesta etapa, foi desenvolvido o método computacional para a estimação do percentual de fibrose intersticial.

Como o Cytomine tem integração nativa com o Python, optou-se por implementar o método computacional deste trabalho nessa linguagem conjuntamente com o *framework* OpenCV versão 3.9 (<https://pypi.org/project/opencv-python/>), uma vez que a API apresentou incompatibilidade com a versão 3.10 do *OpenCV*.

O *framework* OpenCV é uma biblioteca de funções de programação voltada principalmente para a Visão Computacional. Quanto à plataforma IDE ou *Integrated Development Environment* (Ambiente de Desenvolvimento Integrado) utilizada para o desenvolvimento dos códigos, empregou-se o *Visual Studio Code* versão 1.66.1, por ser uma ferramenta simples, poderosa, customizável por extensões variadas, gratuita e que continua crescendo, sendo comparável em funcionalidades a outras IDEs pagas.

Sendo a proporção de cor azul o marcador para avaliar o percentual de fibrose inters-

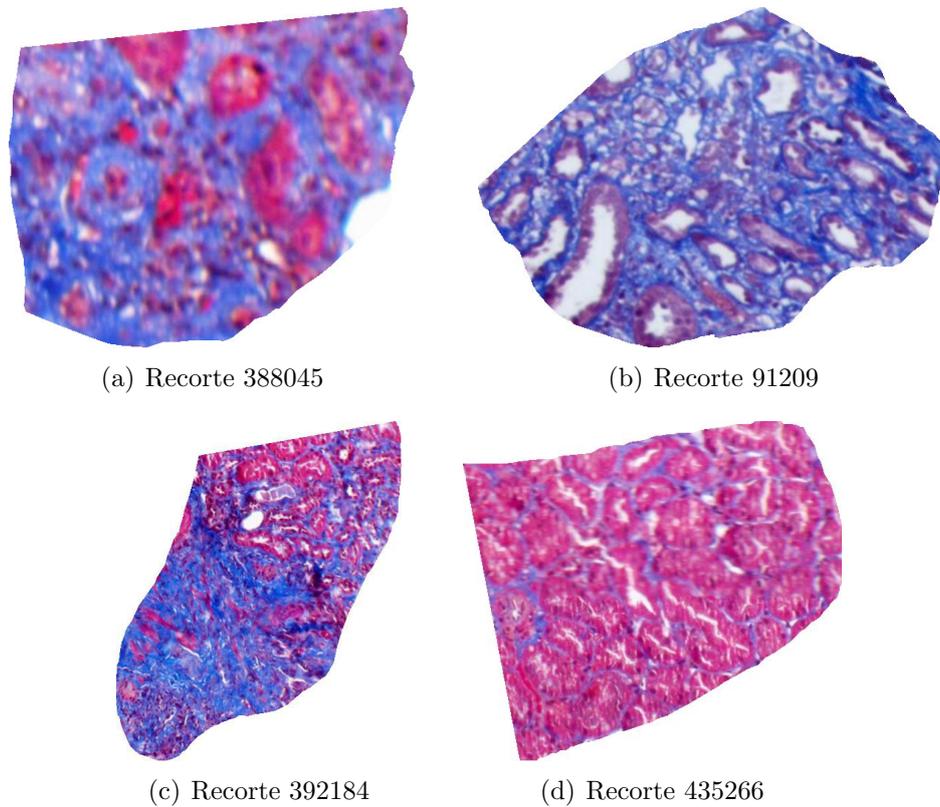


Figura 3.7: Exemplo de recortes de imagens que compõem o *dataset*: (a) a (d) Representação histológica da medula renal. Elaboração do Autor(2022).

ticial (FI) em uma de biópsia corada com trinômio de Masson, na implementação do método computacional foi utilizada a segmentação baseada em cor, que atende ao objetivo principal desta pesquisa: a estimacão de fibrose intersticial, a partir de lâminas de biópsias renais. A escolha da cor azul foi motivada pelo fato de que o tricrômico de Masson reage em azul evidenciando a presença da FI.

Para separar a cor azul que o método computacional deveria analisar, recorreu-se à padronização da Comissão Internacional de Iluminação (abreviada como CIE a partir de seu título francês - *Commission Internationale de l'Eclairage*) que é uma organização dedicada à cooperação internacional e ao intercâmbio de informações nas questões relativas à ciência e à arte da luz. Após diversos estudos, a CIE atribuiu em 1931 (CIE, 1932; Guild, 1931; Smith e Guild, 1931; Wright, 1929), os seguintes comprimentos de onda para cores primárias (que podem ser combinadas para criar outras cores, mas que não podem ser obtidas por mistura de cores): azul = 435,8 nm, verde = 546,1 nm e vermelho = 700 nm.

Na literatura pesquisada encontramos algumas faixas de azul (Tabela 3.3) e optamos por testar três possibilidades para verificar se haveria diferenças entre elas na determinação da FI. Desse modo, foi adotado para uso no método computacional o espectro de cor visível, que corresponde à cor azul com os seguintes comprimentos

de onda: $\lambda = 435$ a 450 nm, $\lambda = 435$ a 470 nm e $\lambda = 435$ a 485 nm.

Tabela 3.3: Exemplo de comprimentos de ondas no espectro visível para a cor azul encontrado em algumas literaturas

Autor	Comprimento de onda (nm)
Cavalheiro (2021)	440 a 485
Silva (2007)	440 a 485
Marques Filho e Neto (1999)	455 a 492
Bruno e Svoronos (2005)	450 a 495
Santos e Scarano (2021)	435.83 a 485.56

O modelo usado para a representação das cores no computador foi o modelo HSV (*hue, saturation e value*) sendo também conhecido como HSI (*hue, saturation e intensity*), que, segundo o CIE, é o modelo que mais se aproxima da maneira como o humano enxerga uma cor. O fluxo de operações do método computacional pode ser visto na Figura 3.8.

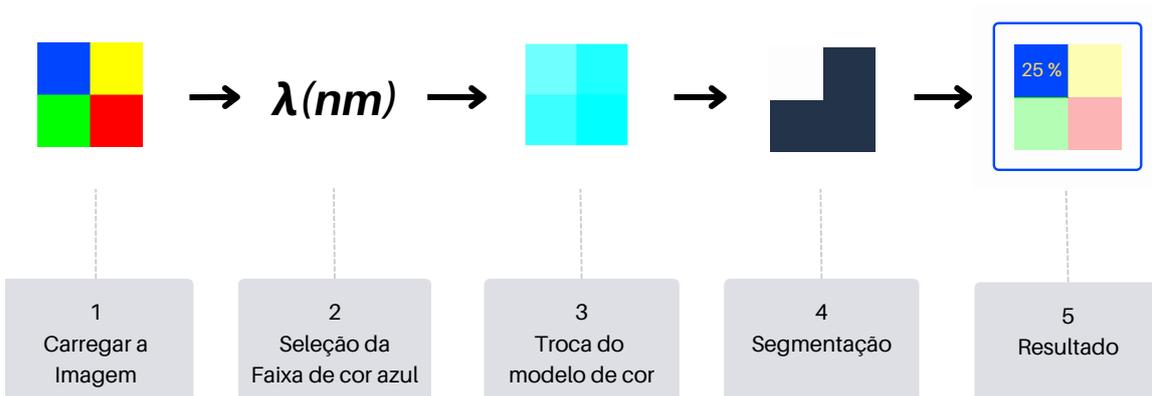


Figura 3.8: Estrutura do método computacional. Elaboração do Autor (2022)

A descrição de cada um dos blocos da Figura 3.8 está elencada a seguir:

1. Carregar imagens com características de interesse no método computacional.
2. Definir a cor azul, de acordo ao comprimento de onda λ visível ao olho humano normal. Para os testes selecionou-se as seguinte faixas de comprimento de onda: $\lambda = 435$ a 450 nm, $\lambda = 435$ a 470 nm e $\lambda = 435$ a 485 nm;
3. Trocar a cor da imagem para o espaço de cores HSV;
4. Segmentar a imagem no espaço de cores HSV, para obter apenas a faixa de cor de interesse, no caso azul;
5. Retorna o percentual de azul encontrado na imagem, obtido pela divisão da quantidade de *pixels* brancos presentes na máscara (representa a cor azul) multiplicados por 100, e dividindo o resultado pelo total de *pixels* de todo o recorte de imagem analisado, conforme Equação 3.1:

$$f_i = \frac{px_w * 100}{px_\Sigma} \quad (3.1)$$

Na qual:

f_i = Fibrose intersticial estimada

px_w = Quantidade de *pixels* brancos

px_Σ = Quantidade de *pixels* total

3.7 Fase 7: 2ª avaliação do percentual de fibrose intersticial pelos médicos.

Nesta fase, todas as imagens resultantes do trabalho individual de cada médico patologista (obtidas na quarta fase) foram reavaliadas por todos os médicos. Assim, cada patologista reavaliou seu trabalho de estimativa da Fibrose Intersticial e o trabalho dos demais membros da equipe de médicos patologistas.

3.8 Fase 8: Comparação dos resultados obtidos.

Sabendo-se que os anotadores humanos normalmente não olham para as mesmas partes afetadas em imagens de lâminas completas de biópsia (do inglês *Whole Slide Imaging*–WSI), ocasionando diferenças interexaminador (Figura 3.9), foi utilizado o escore de Banff e o coeficiente Kappa de Cohen (Cohen, 1960b; Glen, 2014; Landis e Koch, 1977) para medir a concordância dos resultados. Enquanto o escore de Banff foi usado para normalizar faixas de valores percentuais de FI, o coeficiente Kappa de Cohen foi usado para avaliar o grau de concordância inter e intra avaliadores.

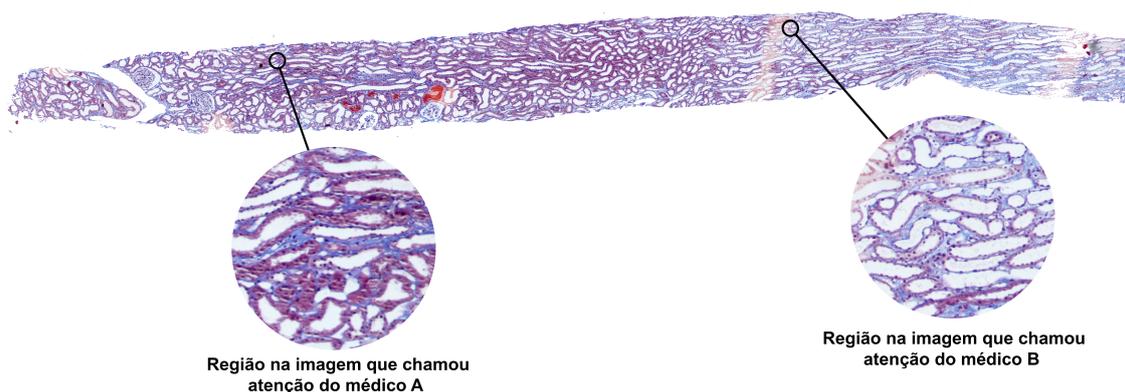


Figura 3.9: Imagem de Lâmina Inteira (*Whole Slide Imaging*, *WSI*) com regiões afetadas por fibrose intersticial. A Figura é um exemplo em que patologistas distintos podem olhar para regiões diferentes de uma mesma imagem de Lâmina Inteira (*Whole Slide Imaging*, *WSI*). Elaboração do Autor (2022).

Na execução dos experimentos, o método computacional é aplicado às imagens anotadas por cada médico patologista da equipe de médicos e retorna o valor percentual de FI. Em seguida, foi aplicado o escore de Banff para Fibrose Intersticial, atribuindo às seguintes faixas de valor (Roufousse et al., 2018; Loupy et al., 2020):

- ci0 – Fibrose Intersticial em até 5% (resultado considerado normal).
- ci1 – Fibrose Intersticial em 6 a 25% (Fibrose Intersticial leve).
- ci2 – Fibrose Intersticial em 26 a 50% (Fibrose Intersticial moderada).
- ci3 – Fibrose Intersticial maior que 50% (Fibrose Intersticial grave).

Para ajudar na compreensão, a tabela com valores percentuais estimados de FI e os valores normalizados após a aplicação do *score de Banff* é demonstrada parcialmente na Figura 3.10.

Identificação da Lâmina	Identificação da Anotação	Valores percentuais originais										Valores normalizados usando critério de Banff						
		Avaliação Automática			Avaliação Humana					Avaliação Automática			Avaliação Humana					
		λ (nm) 435~450	λ (nm) 435~470	λ (nm) 435~485	MÉDICO A	MÉDICO B	MÉDICO C	MÉDICO D	MÉDICO E	λ (nm) 435~450	λ (nm) 435~470	λ (nm) 435~485	MÉDICO A	MÉDICO B	MÉDICO C	MÉDICO D	MÉDICO E	
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	
1	HSCA19048957_P169854_HSCA19048957_309569	8	8	8	60	40	45	10	60	60	ci1	ci1	ci1	ci3	ci2	ci2	ci1	ci3
2	HSCA19048957_P169854_HSCA19048957_309831	4	4	4	35	15	25	10	60	60	ci0	ci0	ci0	ci2	ci1	ci2	ci1	ci3
3	HSCA19048957_P169854_HSCA19048957_309997	2	2	2	40	10	20	15	65	65	ci0	ci0	ci0	ci2	ci1	ci1	ci1	ci3
4	HSCA19048957_P169854_HSCA19048957_310122	1	1	1	5	5	2	0	15	15	ci0	ci0	ci0	ci1	ci1	ci0	ci0	ci1
5	HSCA19048957_P169854_HSCA19048957_310201	8	8	8	55	40	25	10	55	55	ci1	ci1	ci1	ci3	ci2	ci2	ci1	ci3
6	HSCA19048957_P169854_HSCA19048957_310983	1	1	1	20	10	5	5	40	40	ci0	ci0	ci0	ci1	ci1	ci1	ci1	ci2
7	HSCA19048957_P169854_HSCA19048957_311127	1	1	1	10	8	25	5	65	65	ci0	ci0	ci0	ci1	ci1	ci2	ci1	ci3
8	HSCA19048957_P169854_HSCA19048957_311256	1	1	1	10	8	30	5	45	45	ci0	ci0	ci0	ci1	ci1	ci2	ci1	ci2
9	HSCA19048957_P169854_HSCA19048957_311323	6	7	7	60	60	10	20	75	75	ci1	ci1	ci1	ci3	ci3	ci1	ci1	ci3
10	HSCA19048957_P169854_HSCA19048957_311435	3	3	3	30	20	50	30	70	70	ci0	ci0	ci0	ci2	ci1	ci3	ci2	ci3
11	HSCA19048957_P169854_HSCA19048957_311508	10	10	10	70	80	30	70	60	60	ci1	ci1	ci1	ci3	ci3	ci2	ci3	ci3
12	HSCA19048957_P169854_HSCA19048957_311669	7	7	7	30	50	5	5	20	20	ci1	ci1	ci1	ci2	ci3	ci1	ci1	ci1
13	HSCA19048957_P169854_HSCA19048957_311782	5	5	5	30	20	10	15	35	35	ci1	ci1	ci1	ci2	ci1	ci1	ci1	ci2
14	HSCA19048957_P169854_HSCA19048957_311971	2	2	2	10	15	5	20	35	35	ci0	ci0	ci0	ci1	ci1	ci1	ci1	ci2

Figura 3.10: Exemplo da normalização de valores percentuais usando critério de Banff. Elaboração do Autor (2022).

Assim, na Figura 3.10 o valor da linha **1**, coluna **C** é **oito**. Aplicando o critério de Banff passa a ser representado por **ci1** na linha **1**, coluna **K**. Essa interpretação é aplicada para demais linhas e colunas da Figura 3.10.

Em seguida, usando o método kappa de Cohen (Landis e Koch, 1977), foram montadas as tabelas de contigência para realização dos cálculos de concordância interexaminador e intraexaminador. Relembrando que a utilização do coeficiente de concordância Kappa, permite avaliar a medida do grau em que dois juízes ou examinadores, A e B, concordam em suas respectivas classificações de N itens divididos em K categorias mutuamente exclusivas. Um “juiz”ou examinador neste contexto pode ser um único indivíduo, um conjunto de indivíduos que classificam os itens coletivamente ou um sistema computacional Vergni et al. (2021); Cantor (1996), sendo que os cálculos consideram apenas correspondências exatas entre observadores.

O coeficiente de kappa criado por Cohen (1960a), conhecidamente é usado para avaliar a concordância entre examinadores numa escala nominal. Enquanto que o coeficiente de kappa ponderado proposto por Cohen (1968) assume que as categorias são ordenadas e contabilizadas para saber a distância entre os dois examinadores, sendo indicado para os casos em que as categorias da avaliação são ordinais, como por exemplo ruim, médio e ótimo. Se as categorias (ci0, ci1,ci2 e ci3 do *score de Banff ci*) forem ordenadas, é possível considerar correspondências próximas. Em outras palavras, se um examinador classifica a fibrose intersticial no grupo ci1 e o outro no grupo ci2, isso é mais próximo do que se um classifica em ci0 e o outro em ci3.

Recaptulando o Capítulo 2, o coeficiente de kappa pode ser compreendido como sendo a razão entre duas distâncias, ou seja, quanto mais concordantes forem os examinadores, mais próximo de 1 estará o valor da medida. O coeficiente de kappa varia de -1, em situações em que a amostra é balanceada para duas categorias de resposta e os examinadores discordam em todas as avaliações, até 1, quando os examinadores concordam em todas as avaliações. Vejamos o exemplo de resultados da concordância intraexaminador da Tabela 3.4, em que cada valor em linha é a classificação da fibrose intersticial pelo Examinador A na primeira aferição e cada valor em coluna é o resultado da segunda aferição pelo mesmo examinador.

Tabela 3.4: Resultados da Concordância Intraexaminador

		Examinador A				
Examinador A	ci0	ci1	ci2	ci3	Total	Linha
<i>ci0</i>			9			9
<i>ci1</i>	1	15	3			19
<i>ci2</i>		1	3	1		5
<i>ci3</i>				1		1
Total Coluna	1	25	6	2		34
	<i>P_o</i>		<i>P_e</i>		<i>kappa</i>	
Resultado	0,559		0,446		0,203	

O coeficiente de concordância de Kappa, sugerido por Cohen (1960a) foi calculado utilizando a Equação 3.2 abaixo:

$$k = \frac{(p_o - p_e)}{(1 - p_e)} \tag{3.2}$$

Na qual:

$$P_o = \sum_{i=1}^r \frac{n_{ii}}{n} \quad P_e = \sum_{i=1}^r \frac{n_{i.} * n_{.i}}{n^2}$$

P_o: Acordo observado entre os examinadores.

P_e: Acordo esperado (probabilidade hipotética).

Na Tabela 3.4 estão apresentados os resultados da comparação entre as duas estimativas de FI feitas pelo médico A. Observa-se que ao cruzar a primeira linha (ci0)

com a segunda coluna (ci1) é possível ver nove resultados que foram considerados (ci0) na primeira avaliação e (ci1) na segunda avaliação.

No cruzamento da segunda linha (ci1) com a primeira coluna (ci1) é possível ver 1 na primeira avaliação e (ci0) na segunda avaliação.

Ocorreram 15 resultados de concordância total quando o cruzamento de informação foi realizado na segunda linha (ci1) e segunda coluna (ci1). E quando a comparação foi feita entre a segunda linha (ci1) e terceira coluna (ci2), três resultados foram considerados: (ci1) na primeira avaliação e (ci2) na segunda avaliação.

No cruzamento da terceira linha (ci2) com a segunda coluna (ci1) é possível ver 1 resultado considerado na primeira avaliação como (ci2) e (ci1) na segunda avaliação. Ainda na terceira linha (ci2), cruzando com a terceira coluna (ci2) vê-se uma concordância de 3 resultados em concordância nas duas avaliações. Para o cruzamento da terceira linha (ci2) com a quarta coluna (ci3), na primeira avaliação ocorreu um resultado considerado (ci2) e (ci3) na segunda avaliação.

Para o cruzamento da quarta linha (ci3) e quarta coluna (ci3) ocorreu 1 resultado concordante para a primeira e segunda avaliação. Logo, observando os resultados de concordância na classificação categórica intraexaminador da Tabela 3.4, obtém-se o valor calculado de Kappa (k) 0.203 conforme demonstrado a seguir.

$$Po = \frac{15}{34} + \frac{3}{34} + \frac{1}{34} = \frac{19}{34} = 0.559$$

$$Pe = \frac{1*9}{34^2} + \frac{25*19}{34^2} + \frac{6*5}{34^2} + \frac{2*1}{34^2} = \frac{516}{1156} = 0.446$$

$$k = \frac{0.559 - 0.446}{0.446} = 0.203$$

Para interpretar o coeficiente Kappa neste trabalho, foi adotado a escala sugerida no artigo de Landis e Koch (1977), exibida na Tabela 3.5. Utilizando a interpretação proposta é possível afirmar que o examinador A apresentou uma concordância “fraca” em relação aos dados analisados na primeira e segunda avaliação. O mesmo princípio de análise aplica-se para os dados apresentados nas Tabelas 4.1 a 4.20, visualizadas no Capítulo 4 o qual apresenta e analisa os resultados obtidos..

Tabela 3.5: Tabela baseada na interpretação Kappa sugerida por Landis e Koch (1977)

(k_n)	Faixa de Valor	Categoria/Concordância
k_1	Menor que zero	Insignificante ou Sem acordo
k_2	De 0,00 a 0,20	Fraca
k_3	De 0,21 a 0,40	Razoável ou Justa
k_4	De 0,41 a 0,60	Moderada
k_5	De 0,61 a 0,80	Forte
k_6	De 0,81 a 1,00	Quase perfeita

Capítulo 4

Resultados

Neste capítulo são apresentados os resultados dos experimentos realizados com o intuito de avaliar o desempenho do método de estimação semi-automática da fibrose intersticial.

4.1 Resultados de concordância intraexaminador

4.1.1 Concordância intraexaminador do médico A

A Tabela 4.1 apresenta a concordância na avaliação que o médico A fez das mesmas imagens em dois momentos diferentes. Para os dados trazidos nesta tabela, são ignoradas informações extras sobre os examinadores ou sobre as imagens avaliadas, configurando assim que o grau de concordância não sofrerá influência em decorrência das condições de avaliação ou características dos examinadores. Este tipo de conjunto de dados é conhecido como população homogênea (Prado, 2004).

Utilizando a Equação 3.2 sugerida por Cohen (1960a), o coeficiente de kappa da Tabela 4.1 é estimado em 0,203. Este kappa não é muito próximo de 1 (concordância total das aferições) e, de acordo com a escala de interpretação kappa adotada neste trabalho (Tabela 3.5), a concordância é considerada fraca.

Tabela 4.1: Resultados da Concordância Intraexaminador

Médico A					
Médico A	ci0	ci1	ci2	ci3	Total Linha
<i>ci0</i>		9			9
<i>ci1</i>	1	15	3		19
<i>ci2</i>		1	3	1	5
<i>ci3</i>				1	1
Total Coluna	1	25	6	2	34
	<i>Po</i>		<i>Pe</i>		<i>kappa</i>
Resultado	0,559		0,446		0,203

4.1.2 Concordância intraexaminador do médico B

A tabela apresenta o resultado e como pode-se ver a concordância do médico com ele mesmo, e em momentos diferentes. O coeficiente de kappa da Tabela 4.2 é estimado em 0,267 sendo considerada razoável pelos critérios adotados neste estudo.

Tabela 4.2: Resultados da Concordância Intraexaminador

Médico B					
Médico B	ci0	ci1	ci2	ci3	Total Linha
<i>ci0</i>	3				3
<i>ci1</i>	6	9	2		17
<i>ci2</i>			2		2
<i>ci3</i>	1	1	10		12
Total Coluna	10	10	14	0	34
	<i>Po</i>		<i>Pe</i>		<i>kappa</i>
Resultado	0,412		0,197		0,267

4.1.3 Concordância intraexaminador do médico C

A tabela apresenta o resultado da concordância do médico C em momentos diferentes. O coeficiente de kappa da Tabela 4.3 foi estimado em 0,030, sendo portanto considerado como concordância fraca.

Tabela 4.3: Resultados da Concordância Intraexaminador

Médico C					
Médico C	ci0	ci1	ci2	ci3	Total Linha
<i>ci0</i>	5				5
<i>ci1</i>	11	3	1		15
<i>ci2</i>	3	3			6
<i>ci3</i>	3	1	4		8
Total Coluna	22	7	5	0	34
	<i>Po</i>		<i>Pe</i>		<i>kappa</i>
Resultado	0,235		0,212		0,030

4.1.4 Concordância intraexaminador do médico D

A tabela apresenta o resultado da concordância em momentos diferentes do médico D com ele mesmo. O coeficiente de kappa da Tabela 4.4 é considerado como concordância fraca, retornando o valor 0,151.

Tabela 4.4: Resultados da Concordância Intraexaminador

Médico D					
Médico D	ci0	ci1	ci2	ci3	Total Linha
<i>ci0</i>	5	9	6		20
<i>ci1</i>		9	2		11
<i>ci2</i>		3			3
<i>ci3</i>					34
Total Coluna	5	21	8	0	68
	<i>Po</i>		<i>Pe</i>		<i>kappa</i>
Resultado	0,412		0,307		0,151

4.1.5 Concordância intraexaminador do médico E

A tabela apresenta o resultado da concordância em momentos diferentes do médico D com ele mesmo. O coeficiente de kappa da Tabela 4.5, com retorno calculado em -0,67 é considerado como sem acordo.

Tabela 4.5: Resultados da Concordância Intraexaminador

Médico E					
Médico E	ci0	ci1	ci2	ci3	Total Linha
<i>ci0</i>					0
<i>ci1</i>	9	3			12
<i>ci2</i>	5	1			6
<i>ci3</i>	1	5	10		16
Total Coluna	15	9	10	0	34
	<i>Po</i>		<i>Pe</i>		<i>kappa</i>
Resultado	0,088		0,145		-0,067

4.1.6 Conclusões sobre análise interexaminador

Baseando-se nos valores dos índices kappa encontrados acima, é possível afirmar que os médicos não apresentam boa concordância com eles mesmos. Como é visto nos resultados apresentados até então, apenas o médico B retornou uma concordância razoável em suas aferições. Enquanto que os médicos A, C e D apresentaram concordância fraca ao comparar a primeira avaliação da fibrose intersticial com a segunda. Já o médico E, discorda de suas próprias avaliações quando comparados os resultados na primeira e segunda etapa de avaliação da fibrose intersticial.

4.2 Resultados de concordância interexaminador

4.2.1 Resultados de concordância entre médicos A e B

A tabela apresenta o resultado e como pode-se ver a concordância do médico A com o médico B, e em momentos diferentes. O coeficiente de kappa da Tabela 4.6 é estimado em 0,228 sendo considerada razoável pelos critérios adotados neste estudo.

Tabela 4.6: Resultados da Concordância Interexaminador

Médico B					
Médico A	ci0	ci1	ci2	ci3	Total Linha
<i>ci0</i>	17	33	8	10	68
<i>ci1</i>	10	97	36	31	174
<i>ci2</i>		10	9	17	36
<i>ci3</i>		1	5	25	31
Total Coluna	27	141	58	83	309
	<i>Po</i>		<i>Pe</i>		<i>kappa</i>
Resultado	0,559		0,446		0,228

4.2.2 Resultados de concordância entre médicos A e C

A tabela apresenta o resultado e como pode-se ver a concordância do médico A com o médico C, e em momentos diferentes. O coeficiente de kappa da Tabela 4.7 é estimado em 0,214 sendo considerada razoável pelos critérios adotados neste estudo.

Tabela 4.7: Resultados da Concordância Interexaminador

		Médico C				
Médico A	ci0	ci1	ci2	ci3	Total	Linha
<i>ci0</i>	27	15	9	17	68	
<i>ci1</i>	19	69	47	39	174	
<i>ci2</i>	1	9	10	16	36	
<i>ci3</i>		1	6	24	31	
Total Coluna	47	94	72	96	309	
	<i>Po</i>		<i>Pe</i>		<i>kappa</i>	
Resultado	0,421		0,263		0,214	

4.2.3 Resultados de concordância entre médicos A e D

A tabela apresenta o resultado e como pode-se ver a concordância do médico A com o médico D, e em momentos diferentes. O coeficiente de kappa da Tabela 4.8 é estimado em 0,260 sendo considerada razoável pelos critérios adotados neste estudo.

Tabela 4.8: Resultados da Concordância Interexaminador

		Médico D				
Médico A	ci0	ci1	ci2	ci3	Total	Linha
<i>ci0</i>	45	19	3	1	68	
<i>ci1</i>	19	69	47	39	174	
<i>ci2</i>	1	9	10	16	36	
<i>ci3</i>		1	6	24	31	
Total Coluna	47	94	72	96	309	
	<i>Po</i>		<i>Pe</i>		<i>kappa</i>	
Resultado	0,528		0,361		0,260	

4.2.4 Resultados de concordância entre médicos A e E

A tabela apresenta o resultado e como pode-se ver a concordância do médico A com o médico E, e em momentos diferentes. O coeficiente de kappa da Tabela 4.9 é estimado em 0,218 sendo considerada razoável pelos critérios adotados neste estudo.

Tabela 4.9: Resultados da Concordância Interexaminador

Médico E					
Médico A	ci0	ci1	ci2	ci3	Total Linha
<i>ci0</i>	26	22	9	11	68
<i>ci1</i>	18	72	42	42	174
<i>ci2</i>		7	7	22	36
<i>ci3</i>			3	28	31
Total Coluna	44	101	61	103	309
	<i>Po</i>		<i>Pe</i>		<i>kappa</i>
Resultado	0,430		0,272		0,218

4.2.5 Conclusões sobre análise interexaminador do médico A

Baseando-se nos valores retornados e analisados, com relação aos resultados de concordância interexaminador, é possível afirmar que a concordância do médico A com os médicos B, C, D e E foi em unanimidade das comparações, considerado como razoável.

4.2.6 Resultados de concordância entre médicos B e C

A tabela apresenta o resultado e como pode-se ver a concordância do médico B com o médico C, e em momentos diferentes. O coeficiente de kappa da Tabela 4.10 é estimando em 0,457 sendo considerada uma concordância moderada.

Tabela 4.10: Resultados da Concordância Interexaminador

Médico C					
Médico B	ci0	ci1	ci2	ci3	Total Linha
<i>ci0</i>	22	5	0	0	27
<i>ci1</i>	25	74	28	14	174
<i>ci2</i>		12	28	18	58
<i>ci3</i>		3	16	64	83
Total Coluna	47	94	72	96	309
	<i>Po</i>		<i>Pe</i>		<i>kappa</i>
Resultado	0,608		0,279		0,457

4.2.7 Resultados de concordância entre médicos B e D

A tabela apresenta o resultado e como pode-se ver a concordância do médico B com o médico D, e em momentos diferentes. O coeficiente de kappa da Tabela 4.11 é estimando em 0,067 sendo considerada uma concordância fraca.

Tabela 4.11: Resultados da Concordância Interexaminador

Médico D					
Médico B	ci0	ci1	ci2	ci3	Total Linha
<i>ci0</i>	11	16	0	0	27
<i>ci1</i>	58	68	14	1	141
<i>ci2</i>	18	28	8	4	58
<i>ci3</i>	25	32	11	15	83
Total Coluna	112	144	33	20	309
	<i>Po</i>		<i>Pe</i>		<i>kappa</i>
Resultado	0,330		0,282		0,067

4.2.8 Resultados de concordância entre médicos B e E

A tabela apresenta o resultado e como pode-se ver a concordância do médico B com o médico D, e em momentos diferentes. O coeficiente de kappa da Tabela 4.12 é estimado em 0,386 sendo considerada uma concordância razoável.

Tabela 4.12: Resultados da Concordância Interexaminador

Médico E					
Médico B	ci0	ci1	ci2	ci3	Total Linha
<i>ci0</i>	22	5	0	0	27
<i>ci1</i>	21	68	35	17	141
<i>ci2</i>	1	21	17	19	58
<i>ci3</i>	0	7	9	67	83
Total Coluna	112	101	61	103	309
	<i>Po</i>		<i>Pe</i>		<i>kappa</i>
Resultado	0,563		0,288		0,386

4.2.9 Conclusões sobre análise interexaminador do médico B

Baseando-se nos valores retornados e analisados, com relação aos resultados de concordância interexaminador, é possível afirmar que a concordância do médico B com os médicos C, D e E apresentou variância de fraca a razoável, no qual razoável está mais perto de 1 (concordância total).

4.2.10 Resultados de concordância entre médicos C e D

A tabela apresenta o resultado da concordância em momentos diferentes do médico C com o médico D. O coeficiente de kappa da Tabela 4.13 foi estimado em 0,099, sendo portanto considerado como concordância fraca.

Tabela 4.13: Resultados da Concordância Interexaminador

Médico D					
Médico C	ci0	ci1	ci2	ci3	Total Linha
<i>ci0</i>	22	24	1	0	47
<i>ci1</i>	30	54	10	0	94
<i>ci2</i>	23	35	8	6	72
<i>ci3</i>	37	31	14	14	96
Total Coluna	112	144	33	20	309
	<i>Po</i>		<i>Pe</i>		<i>kappa</i>
Resultado	0,317		0,242		0,099

4.2.11 Resultados de concordância entre médicos C e E

A tabela apresenta o resultado e como pode-se ver a concordância do médico C com o médico E, e em momentos diferentes. O coeficiente de kappa da Tabela 4.14 é estimando em 0,388 sendo considerada uma concordância razoável.

Tabela 4.14: Resultados da Concordância Interexaminador

Médico E					
Médico C	ci0	ci1	ci2	ci3	Total Linha
<i>ci0</i>	34	12	1	0	47
<i>ci1</i>	9	50	20	15	94
<i>ci2</i>	1	22	24	25	72
<i>ci3</i>	0	17	16	63	96
Total Coluna	44	101	61	103	309
	<i>Po</i>		<i>Pe</i>		<i>kappa</i>
Resultado	0,553		0,271		0,388

4.2.12 Conclusões sobre análise interexaminador do médico C

Baseando-se nos valores retornados e analisados, com relação aos resultados de concordância interexaminador, é possível afirmar que a concordância do médico C com os médicos D e E foram consideradas fraca e razoável sucessivamente. Ou seja, não há uma boa concordância nos resultados anotados pelo médico C e os médicos D e E.

4.2.13 Resultados de concordância entre médicos D e E

A tabela apresenta o resultado da concordância em momentos diferentes do médico D com o médico E. O coeficiente de kappa da Tabela 4.15 foi estimando em 0,098, sendo portanto considerado como concordância fraca.

Tabela 4.15: Resultados da Concordância Interexaminador

Médico E					
Médico D	ci0	ci1	ci2	ci3	Total Linha
<i>ci0</i>	19	38	24	31	112
<i>ci1</i>	25	55	28	36	144
<i>ci2</i>	0	6	8	19	33
<i>ci3</i>	0	2	1	17	20
Total Coluna	44	101	61	103	309
	<i>Po</i>		<i>Pe</i>		<i>kappa</i>
Resultado	0,320		0,247		0,098

4.2.14 Conclusões sobre análise interexaminador D

Com relação a análise de concordância do médico D com o médico E, obviamente é fraca uma vez que o valor retornado de 0,098 está distante de 1 que representa concordância total.

4.3 Resultados de concordância interexaminador entre o método computacional e os médicos A,B,C,D,E.

4.3.1 Resultados de concordância interexaminador: $[\Sigma, A]$

A tabela 4.17 apresenta o resultado comparativo da concordância entre o método computacional e o médico A. O coeficiente de kappa foi estimado em -0,030 que é considerado como sem acordo.

Tabela 4.16: Resultados da Concordância Interexaminador

Médico A					
Método Σ	ci0	ci1	ci2	ci3	Total Linha
<i>ci0</i>	64	154	14	2	234
<i>ci1</i>	4	20	22	22	68
<i>ci2</i>				7	7
<i>ci3</i>					0
Total Coluna	68	174	36	31	309
	<i>Po</i>		<i>Pe</i>		<i>kappa</i>
Resultado	0,272		0,293		-0,030

4.3.2 Resultados de concordância interexaminador: $[\Sigma, B]$

A tabela 4.17 apresenta o resultado comparativo da concordância entre o método computacional e o médico B. O coeficiente de kappa foi estimado em -0,054 que é considerado como sem acordo.

Tabela 4.17: Resultados da Concordância Interexaminador

Médico B					
Método Σ	ci0	ci1	ci2	ci3	Total Linha
<i>ci0</i>	27	130	48	29	234
<i>ci1</i>		11	9	48	68
<i>ci2</i>			1	6	7
<i>ci3</i>					0
Total Coluna	27	141	58	83	309
	<i>Po</i>		<i>Pe</i>		<i>kappa</i>
Resultado	0,126		0,171		-0,054

4.3.3 Resultados de concordância interexaminador: [Σ, C]

A tabela 4.18 apresenta o resultado da concordância comparativo da concordância entre o método computacional e o médico C. O coeficiente de kappa foi estimado em 0,000 que é considerado como uma concordância fraca.

Tabela 4.18: Resultados da Concordância Interexaminador

Médico C					
Método Σ	ci0	ci1	ci2	ci3	Total Linha
<i>ci0</i>	47	84	56	47	234
<i>ci1</i>		10	15	43	68
<i>ci2</i>			1	6	7
<i>ci3</i>					0
Total Coluna	27	141	58	83	309
	<i>Po</i>		<i>Pe</i>		<i>kappa</i>
Resultado	0,188		0,187		0,000

4.3.4 Resultados de concordância interexaminador: [Σ, D]

A tabela 4.19 apresenta o resultado da concordância comparativo da concordância entre o método computacional e o médico D. O coeficiente de kappa foi estimado em 0,098 foi considerado como uma concordância fraca.

Tabela 4.19: Resultados da Concordância Interexaminador

Médico C					
Método Σ	ci0	ci1	ci2	ci3	Total Linha
<i>ci0</i>	102	109	18	5	234
<i>ci1</i>	9	33	14	12	68
<i>ci2</i>	1	2	1	3	7
<i>ci3</i>					0
Total Coluna	112	144	33	20	309
	<i>Po</i>		<i>Pe</i>		<i>kappa</i>
Resultado	0,440		0,379		0,098

4.3.5 Resultados de concordância interexaminador: $[\Sigma, E]$

A tabela 4.20 apresenta o resultado da concordância comparativo da concordância entre o método computacional e o médico E. O coeficiente de kappa estimado em -0,012 denota que não houve concordância entre método computacional e o médico.

Tabela 4.20: Resultados da Concordância Interexaminador

Médico C					
Método Σ	ci0	ci1	ci2	ci3	Total Linha
<i>ci0</i>	44	92	48	50	234
<i>ci1</i>		9	12	47	68
<i>ci2</i>			1	6	7
<i>ci3</i>					0
Total Coluna	44	101	61	103	309
	<i>Po</i>		<i>Pe</i>		<i>kappa</i>
Resultado	0,175		0,184		-0,012

4.3.6 Conclusões sobre análise de concordância entre médicos e o método computacional

A Tabela 4.21 apresenta o resultado de todas as comparações possíveis dos resultados entre patologistas e entre patologistas e método computacional. É necessário desconsiderar-se a comparação do método computacional com ele mesmo uma vez que os resultados são invariáveis. Para chegar a esse resultado, foi usada a equação 4.1:

$$C_k^n = \frac{n!}{k!(n-k)!} \quad (4.1)$$

Na qual:

n: total de elementos no conjunto.

k: total de elementos no subconjunto.

Listando-se as combinações intra e interexaminador possíveis para cinco médicos e o método computacional, têm-se:

[A,A], [A,B], [A,C], [A,D], [A,E],[A, Σ]; [B,B], [B,C], [B,D], [B,E], [B, Σ]; [C,C], [C,D], [C,E], [C, Σ]; [D,D], [DE], [D, Σ] e [E,E], [E, Σ], onde Σ representa o método computacional e as letras A,B,C,D e E, representam os médicos.

Deste modo, cada coluna da Tabela 4.21 apresenta as comparações possíveis entre os patologistas. Para que seja possível a comparação, cada resultado comparativo recebeu um índice Kappa (de k1 a k6, conforme demonstrado na legenda da tabela.

Tabela 4.21: Resumo Comparativo de Concordância

k_n	A						B					C				D			E		TOTAIS
	AA	AB	AC	AD	AE	A Σ	BB	BC	BD	BE	B Σ	CC	CD	CE	C Σ	DD	DE	D Σ	EE	E Σ	
k_1						X					X								X	X	4 (20%)
k_2	X								X			X	X		X	X	X				8 (40%)
k_3		X	X	X	X		X			X				X							7 (35%)
k_4								X													1 (05%)
k_5																					-
k_6																					-

Legenda

k_n	Faixa de Valor	Categoria/Concordância	Anotadores
k_1	Menor que zero	Insignificante	A Médico A
k_2	De 0,00 a 0,20	Fraca	B Médico B
k_3	De 0,21 a 0,40	Razoável	C Médico C
k_4	De 0,41 a 0,60	Moderada	D Médico D
k_5	De 0,61 a 0,80	Forte	E Médico E
k_6	De 0,81 a 1,00	Quase perfeita	Σ Método Computacional

Por exemplo, o médico patologista [A] apresentou concordância Razoável (k3) quando comparado com os médicos [B], [C], [D] e [E]. Já com o método computacional não houve concordância (k1) e quando os resultados da segunda avaliação foram comparados com os resultados da primeira avaliação com ele mesmo, a concordância foi fraca (k2).

O médico patologista [B] retornou uma comparação razoável (k3) com ele mesmo e o médico patologista [E]. Apenas com o médico patologista [C], a concordância foi moderada (k4), sendo fraca (k2) com o médico patologista [D] Sem Acordo com o método computacional.

O médico patologista [C] retornou um acordo moderado (k4) com o médico patologista [D] e acordo fraco com os demais médicos, incluindo ele mesmo.

O médico patologista [D] retornou uma concordância Fraca (k2) com ele mesmo, com o médico patologista [E] e com o método computacional. Quando o resultado do médico patologista [E] foi comparado com ele mesmo e o método computacional, não retornou concordância (k1).

Com relação à comparação de concordância entre os médicos patologistas e o método computacional proposto, um resultado houve concordância Fraca (k2), que foi com o médico patologista [C]. Os demais resultados comparados com os resultados dos patologistas [A], [B], [D] e [E], a concordância foi insignificante.

O resumo desta análise é demonstrado na Figura 4.1 e em seguida, é apresentado o Capítulo 5 que trata das considerações finais.

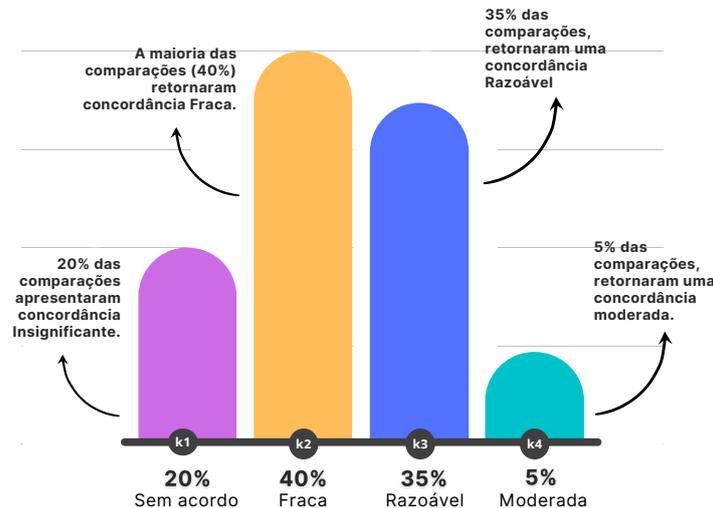


Figura 4.1: Resultado Gráfico da Comparação entre Anotadores. Elaboração do Autor (2022).

Capítulo 5

Considerações Finais

Este trabalho teve como objetivo desenvolver um método semiautomático para estimação de fibrose intersticial, a partir de imagens digitais de biópsias renais, sob a perspectiva da classificação de Banff. Além disso, eram objetivos secundários criar e disponibilizar um banco de dados de imagens anotadas de fibrose intersticial e incluir a solução desenvolvida ao portfólio de ferramentas computacionais fornecidas pelo projeto PathoSpotter.

Para alcançar estes objetivos, foi necessário implantar um sistema de anotação de imagens médicas no projeto PathoSpotter. Este problema foi resolvido com o *software* Cytomine que atende não apenas a demanda desta pesquisa, mas também a demanda de um sistema de anotação de imagens médicas dos patologistas do projeto *PathoSpotter*.

Um conjunto de 309 imagens (*dataset*) foi montado a partir de seis imagens digitais de lâminas inteiras de biópsias renais, coradas com tricrômico de Masson e anotadas por uma equipe de médicos patologistas usando o sistema Cytomine.

Usando técnicas de segmentação por cores, para a identificação de regiões afetadas por fibrose intersticial (FI), um método computacional foi desenvolvido e testado. Os resultados obtidos evidenciam que:

- Destes 309 resultados analisados, aplicando o critério de Banff apenas 2 resultados na faixa λ (nm) 435-450 discordaram (de ci_0 para ci_1) das outras faixas (λ (nm) 435-470 e λ (nm) 435-485). Isto evidencia que neste trabalho, mesmo variando entre as três faixas de cor do espectro azul escolhidas, não houve alterações nos resultados do método computacional.
- Os médicos patologistas apresentaram grau variável de reprodutibilidade e concordância em diferentes cenários de comparação. Tendo como base os resultados analisados é possível afirmar que não houve concordância entre patologistas e o método proposto. Inclusive, essa concordância não ocorreu nem mesmo entre os próprios patologistas. A hipótese é que a complexidade do trabalho

de anotação da fibrose intersticial, o fator da interpretabilidade humana bem como o trabalho repetitivo, contribuíram para a variabilidade na interpretação da fibrose intersticial presente nas imagens médicas. Outros fatores podem ter influenciado nos resultados obtidos, como por exemplo a experiência dos patologistas pode ter influido nos resultados de concordância, o acesso ao laudo da biópsia poderia ajudar na confirmação dos valores estimados e lâminas vistas na tela do computador podem gerar viés de análise uma vez que é comumente os patologistas são treinados para ver no microscópio.

- O método empregado para quantificar a FI apresentou reprodutibilidade e invariância na detecção da cor azul, viabilizando o avanço da pesquisa na construção de cenários distintos de análise. O método permitiu quantificar objetivamente as áreas afetadas por FI, representadas pela cor azul de acordo com as três definições das faixas de comprimentos de onda dessa cor encontradas na literatura.
- Em decorrência do grau variável de reprodutibilidade e concordância apresentado pelos médicos patologistas, não foi possível aferir a qualidade do método computacional, uma vez que não foi encontrado um padrão-ouro na equipe médica.
- O método computacional proposto permite a realização da tarefa de estimativa da fibrose intersticial em grandes volumes de imagens sem apresentar variabilidade de resultados e, atuando muito mais rapidamente do que os patologistas.

É importante destacar que o método proposto não faz anotações, e sim quantifica a fibrose intersticial das áreas delimitadas manualmente pelos patologistas. Por isso, o método proposto neste estudo é um método semiautomático.

Embora a fibrose intersticial seja reconhecidamente o parametro patológico mais relevante associado a função renal dos pacientes, muitos estudos mostram que sua avaliação é pouco reprodutível. Inclusive, em conformidade com trabalhos anteriores, neste estudo foi confirmado que a avaliação humana da fibrose intersticial é variável e demanda de boa confiabilidade inter e intra-avaliadora.

5.1 Trabalhos Futuros

Para o aprimoramento deste trabalho, pretende-se ampliar o *dataset* de imagens de lâmina inteira e de imagens anotadas.

É importante registrar que é nosso intuito refazer o processo de montagem do *dataset* de forma a conseguir imagens que apresentem pelo menos concordância moderada, já que sem isso não se poderá avaliar a qualidade da estimativa do nosso método. Futuramente, pretende-se analisar e considerar outras formas de identificar e quantificar a fibrose intersticial usando outros corantes, bem como empregar outros espaços de cor como, por exemplo, o espaço de cores CIELAB.

Propõe-se, ainda, o aprimoramento do método computacional no intuito de possibilitar a avaliação da fibrose intersticial de modo completamente automático. Para isso, é cogitada a utilização de visão computacional para segmentação automática da fibrose em toda a lâmina, para a partir daí estimar a fibrose das áreas selecionadas. Além disso, tentaremos reproduzir e avaliar métodos publicados de avaliação automática de fibrose intersticial.

Referências

- ABTO, Associação Brasileira De Transplante De Órgãos, D. (2020). Dimensionamento dos transplantes no brasil e em cada estado (2013 - 2020). *RBT - Registro Brasileiro de Transplante, Ano XXVI, n. 4, 2020. 88 p.*
- Backes, A. R. e Junior, J. J. D. M. S. (2019). *Introdução à visão computacional usando Matlab*. Alta Books Editora.
- Ballard, D. e Brown, C. (1982). Computer vision prentice-hall inc. *Englewood Cliffs, New Jersey*.
- Bankhead, P., Loughrey, M. B., FernáNdez, J. A., Dombrowski, Y., Mcart, D. G., Dunne, P. D., Mcquaid, S., Gray, R. T., Murray, L. J., Coleman, H. G., e Others (2017). Qupath: Open source software for digital pathology image analysis. *Scientific reports*, 7(1):1–7.
- Bankhead, P. E. A. (2017). QuPath: Software de código aberto para análise de imagens de patologia digital. Disponível: <<https://doi.org/10.1038/s41598-017-17204-5>>. Acesso: 02 fev 2021.
- Barelli, F. (2018). *Introdução à Visão Computacional: Uma abordagem prática com Python e OpenCV*. Editora Casa do Código.
- Belsare, A. e Mushrif, M. (2012). Histopathological image analysis using image processing techniques: An overview. *Signal & Image Processing*, 3(4):23.
- Berchtold, L., Friedli, I., ValléE, J.-P., Moll, S., Martin, P.-Y., e De Seigneux Matthey, S. (2017). Diagnosis and assessment of renal fibrosis: the state of the art. *Swiss medical weekly*, 147:w14442.
- Berry, K. J. e Mielke Jr, P. W. (1988). A generalization of cohen's kappa agreement measure to interval measurement and multiple raters. *Educational and Psychological Measurement*, 48(4):921–933.
- Boor, P., Ostendorf, T., e Floege, J. (2010). Renal fibrosis: novel insights into mechanisms and therapeutic targets. *Nature Reviews Nephrology*, 6(11):643–656.
- Brasil, M. d. S. (2019). Sistema nacional de transplantes. Disponível: <<https://www.gov.br/saude/pt-br/composicao/saes/snt>>. Acesso: 26 out 2022.

- Brasil, M. d. S. (2020). 12/3 Dia Mundial do Rim. Disponível: <<http://bvsms.saude.gov.br/ultimas-noticias/3138-12-3-dia-mundial-do-rim>>. Acesso: 26 out 2022.
- Brito, N. B. e De Oliveira Reis, J. C. (2016). A teoria das cores de goethe e sua crítica a newton. *Revista Brasileira de História da Ciência*, 9(2):288–298.
- Brito, V. D. C., Santos, P. R. S., e Carvalho Filho, A. O. (2021). Detecção de máscara em python usando opencv e deep learning. *Sociedade Brasileira de Computação*.
- Brooks, J. (2019). COCO Annotator. Disponível: <<https://github.com/jsbroks/coco-annotator/>>. Acesso: 02 fev 2021.
- Bruno, T. J. e Svoronos, P. D. (2005). *CRC handbook of fundamental spectroscopic correlation charts*. CRC Press.
- Cantor, A. B. (1996). Sample-size calculations for cohen's kappa. *Psychological methods*, 1(2):150.
- Cavalheiro, C. A. (2021). Espectro Visível. Disponível: <<https://www.infoescola.com/fisica/espectro-visivel/>>. Acesso: 02 nov 2022.
- Cerqueira, S. P., Aguiar, E. C., Duarte, A. A., Dos Santos, W. L., Oliveira, L. R., e Ângelo, M. F. (2021). Pathospotter classifier: Uma serviço web para auxílio à classificação de lesões em glomérulos renais. In *Anais do XXI Simpósio Brasileiro de Computação Aplicada à Saúde*. SBC.
- Charles Patrick Davis, Md, P. (2021). Medical Definition of Kidney. Disponível: <<https://www.medicinenet.com/kidney/definition.htm>>. Acesso: 21 nov 2021.
- Chen, P. e Krauss, A. (2004). Interrater agreement. *The sage encyclopedia of social science research methods*, 2:511–513.
- Cheng, H.-D., Jiang, X. H., Sun, Y., e Wang, J. (2001). Color image segmentation: advances and prospects. *Pattern recognition*, 34(12):2259–2281.
- CIE, C. (1932). Commission internationale de l'éclairage proceedings, 1931. *Cambridge University, Cambridge*.
- Claro, M. L., Veras, R., Santos, L., FrazÃO, M., Carvalho Filho, A., e Leite, D. (2018). Métodos computacionais para segmentação do disco óptico em imagens de retina: uma revisão. *Revista Brasileira de Computação Aplicada*, 10(2):29–43.
- Cohen, J. (1960a). Chance, skill, and luck: The psychology of guessing and gambling.
- Cohen, J. (1960b). A coefficient of agreement for nominal scales. *Educational and psychological measurement*, 20(1):37–46.

- Cohen, J. (1968). Weighted kappa: nominal scale agreement provision for scaled disagreement or partial credit. *Psychological bulletin*, 70(4):213.
- Colorsystem (2020). Colour order systems in art and science. Disponível: <https://www.colorsystm.com/?page_id=31>. Acesso: 25 mar 2022.
- Colvin, R. B. e Chang, A. (2019). *Diagnostic Pathology: Kidney Diseases E-Book*. Elsevier Health Sciences.
- Deng, Y., Manjunath, B., e Shin, H. (1999). Color image segmentation. In *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*, volume 2, páginas 446–451 Vol. 2.
- Farris, A., Chan, S., Climenhaga, J., Adam, B., Bellamy, C., SeróN, D., Colvin, R., Reeve, J., e Mengel, M. (2014). Banff fibrosis study: multicenter visual assessment and computerized analysis of interstitial fibrosis in kidney biopsies. *American Journal of Transplantation*, 14(4):897–907.
- Farris, A. B. e Alpers, C. E. (2014). What is the best way to measure renal fibrosis?: A pathologist’s perspective. *Kidney international supplements*, 4(1):9–15.
- Farris, A. B., Vizcarra, J., Amgad, M., Cooper, L. A., Gutman, D., e Hogan, J. (2021). Artificial intelligence and algorithmic computational pathology: an introduction with renal allograft examples. *Histopathology*, 78(6):791–804.
- Fielding, R. T. (2000). Architectural styles and the design of network-based software architecture. *Dissertação (PhD) - Curso de Information And Computer Science, University Of California, Irvine, 2010*, página 162. Acesso: 30 mai 2022.
- Fleiss, J. L., Cohen, J., e Everitt, B. S. (1969). Large sample standard errors of kappa and weighted kappa. *Psychological bulletin*, 72(5):323.
- Fonseca, R. J. R. M. d., Silva, P. J. d. S. P. d., e Silva, R. R. d. (2007). Acordo inter-juízes: O caso do coeficiente kappa. *Laboratório de Psicologia*, páginas 81–90.
- Frank O’Brien, M. (2020). Nefrite tubulointersticial. *Ed. Manual Merck de Informação Médica: Versão Para Profissionais de Saúde*.
- Gartber, L. P. (2007). *Tratado de histologia em cores*. Elsevier Brasil.
- George, J. e Beard, J. (2012). *Princípios do web design maravilhoso*.
- Ginley, B., Jen, K.-Y., Han, S. S., Rodrigues, L., Jain, S., Fogo, A. B., Zuckerman, J., Walavalkar, V., Miecznikowski, J. C., Wen, Y., e Others (2021). Automated computational detection of interstitial fibrosis, tubular atrophy, and glomerulosclerosis. *Journal of the American Society of Nephrology*, 32(4):837–850.

- Glen, S. (2014). ‘cohen’s kappa statistic’ de statisticshowto.com: Estatísticas elementares para o resto de nós! Disponível: <<https://www.statisticshowto.com/cohens-Kappa-statistic/>>. Acesso: 08 ago 2021.
- Goethe, J. W. V. (2018). *Doutrina das cores*. Editora Nova Alexandria.
- Gomes J.; Velho, L. (1994). *Computação gráfica: Imagem*. Rio de Janeiro: IMPA/SBM.
- Gonzalez, R. e Woods, R. (1999). *Processamento digital de imagens*. São Paulo: Edgard Blücher.
- Gonzalez, R. e Woods, R. (2000). *Processamento de Imagens Digitais*. Editora Blucher.
- Gonzalez, R. e Woods, R. (2009). *Processamento digital de Imagens*. São Paulo: Pearson Education do Brasil, 3a edição.
- Gonzalez, R. C. (2009). *Digital image processing*. Pearson education india.
- Gonzalez, R. C. e Woods, R. E. (1992). *Digital image processing addison-wesley. Reading, Ma.*
- Granatyr, J. (2022). Segmentação de Imagens com Python de A à Z. Disponível: <<https://iaexpert.academy/2022/02/24/segmentacao-de-imagens-com-python-de-a-a-z/>>. Acesso: 31 mai 2021.
- Greenemeier, L. (2008). Visionary research: Teaching computers to see like a human. Disponível: <<https://www.scientificamerican.com/article/visionary-research>>. Acesso: 29 out 2022.
- Guild, J. (1931). The colorimetric properties of the spectrum. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 230(681-693):149–187.
- Haas, M., Loupy, A., Lefaucheur, C., Roufosse, C., Glotz, D., Seron, D., Nankivell, B., Halloran, P., Colvin, R., Akalin, E., e Others (2018). The banff 2017 kidney meeting report: Revised diagnostic criteria for chronic active t cell-mediated rejection, antibody-mediated rejection, and prospects for integrative endpoints for next-generation clinical trials.
- Ham, J., Kamber, M., P. (2011). *Data Mining :: Concepts and Techniques*, volume 3. Morgan Kaufmann.
- Hoyle, R., Harris, M., e Judd, C. (2002). *Research Methods in Social Relations*. Wadsworth.
- Junqueira, L. C. e Carneiro, J. (2017). *Histologia Básica Texto E Atlas*, volume 13. Ed. Editora Guanabara Koogan Ltda.

- Karagiannakos, S. (2019). Localization and object detection with deep learning. Disponível: <<https://sergioskar.github.io/>>. Acesso: 21 nov 2021.
- Kasvi (2020). Biópsia: o que é e para que serve. Disponível: <<https://kasvi.com.br/biopsia-o-que-e-para-que-serve/>>. Acesso: 27 nov 2021.
- Kelly, C. e Landman, J. (2014). Anatomia do trato urinário. *Coleção Netter de ilustrações médicas. 2a ed. Rio de Janeiro: Saunders-Elsevier*, páginas 24–7.
- Kidshealth (2021). KidsHealth Parents — Definition: Kidney. Disponível: <<https://kidshealth.org/en/parents/kidney.html>>. Acesso: 21 nov 2021.
- Kiernan, J. A. (2015). *Histological and histochemical methods: theory and practice*. © Scion Publishing Ltd, 2015, 5th edition edição.
- Landis, J. R. e Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics*, páginas 159–174.
- Liu, Y., Li, Y., Deng, G., Liu, Y., Wan, R., Wu, R., Ji, D., Xu, S., e Bao, M. (2022a). Morest: Model-based restful api testing with execution feedback. *arXiv preprint arXiv:2204.12148*.
- Liu, Z.-Y., Lin, C.-H., Wang, H.-S., Wen, M.-C., Lin, W.-C., Huang, S.-C., Tu, K.-H., Kuo, C.-F., e Chen, T.-D. (2022b). End-to-end interstitial fibrosis assessment of kidney biopsies with a machine learning-based model. *Nephrology Dialysis Transplantation*.
- Loupy, A., Haas, M., Roufosse, C., Naesens, M., Adam, B., Afrouzian, M., Akalin, E., Alachkar, N., Bagnasco, S., Becker, J. U., e Others (2020). The banff 2019 kidney meeting report (i): Updates on and clarification of criteria for t cell- and antibody-mediated rejection.
- Loupy, A., Haas, M., Solez, K., Racusen, L., Glotz, D., Seron, D., Nankivell, B., Colvin, R., Afrouzian, M., Akalin, E., e Others (2017). The banff 2015 kidney meeting report: current challenges in rejection classification and prospects for adopting molecular pathology.
- Maior, s. p. d. s. d. m. (2021). Sus completa 31 anos. unasus. Disponível: <<https://www.unasus.gov.br/noticia/maior-sistema-publico-de-saude-do-mundo-sus-completa-31-anos>>. Acesso: 26 out 2022.
- Marengoni, M. e Stringhini, S. (2009). Tutorial: Introdução à visão computacional usando opencv. *Revista de Informática Teórica e Aplicada*, 16(1):125–160.
- Marques Filho, O. e Neto, H. V. (1999). *Processamento digital de imagens*. Brasport.

- Marée, R., Rollus, L., Stévens, B., Hoyoux, R., Louppe, G., Vandaele, R., Begon, J.-M., Kainz, P., Geurts, P., e Wehenkel, L. (2016). Collaborative analysis of multi-gigapixel imaging data using Cytomine. *Bioinformatics*, 32(9):1395–1401.
- Masson, P. (1929). Some histological methods: trichrome staining and their preliminary technique. *J Tech Methods*, 12:75–90.
- Mchugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia medica*, 22(3):276–282.
- Menn-Josephy, H., Lee, C. S., Nolin, A., Christov, M., Rybin, D. V., Weinberg, J. M., Henderson, J., Bonegio, R., e Havasi, A. (2016). Renal interstitial fibrosis: an imperfect predictor of kidney disease progression in some patient cohorts. *American journal of nephrology*, 44(4):289–299.
- Microscope, C. (2022). Microscope central. Disponível: <<https://microscopecentral.com/products/olympus-bx40-clinical-microscope>>. Acesso: 27 nov. 2022.
- Montanari, T. (2016). *Histologia: texto, atlas e roteiro de aulas práticas*, volume 4. Ed. do autor.
- Montefiore Institute, D. (2010). Cytomine uliège research&development documentation portal. Disponível: <<https://doc.uliege.cytomine.org/>>. Acesso: 10 fev 2021.
- Nankivell, B. J. e Alexander, S. I. (2010). Rejection of the kidney allograft. *New England Journal of Medicine*, 363(15):1451–1462.
- Nikonovas, T., Spessa, A., Doerr, S. H., Clay, G. D., e Mezbahuddin, S. (2020). Near-complete loss of fire-resistant primary tropical forest cover in sumatra and kalimantan. *Communications Earth & Environment*, 1(1):1–8.
- Pandey, R. e Lalchhanhima, R. (2020). Segmentation techniques for complex image. In *2020 International Conference on Computational Performance Evaluation (ComPE)*, páginas 804–808. IEEE.
- Parmar, R. (2018). Detection and segmentation through convnets. *Towards Data Science*.
- Pascual, M., Theruvath, T., Kawai, T., Tolkoff-Rubin, N., e Cosimi, A. B. (2002). Strategies to improve long-term outcomes after renal transplantation. *New England Journal of Medicine*, 346(8):580–590.
- Pathobin, O. S. (2022). Pathobin on-line store. Disponível: <<https://store.pathobin.com/>>. Acesso: 27 nov. 2022.
- Pfizer (2021). Rins: por que eles são tão importantes e como cuidar deles. Disponível: <<https://cutt.ly/0nvvoPj>>. Acesso: 14 mai 2021.

- Pinheiro, P. (2020). O que o médico nefrologista faz? Disponível: <<https://www.mdsaude.com/nefrologia/quando-procurar-um-nefrologista/>>. Acesso: 18 ago 2021.
- Posner, K. L., Sampson, P. D., Caplan, R. A., Ward, R. J., e Cheney, F. W. (1990). Measuring interrater reliability among multiple raters: an example of methods for nominal data. *Statistics in medicine*, 9(9):1103–1115.
- Prado, R. R. d. (2004). *Modelagem do coeficiente kappa ponderado*. Tese de Doutorado, Universidade de São Paulo.
- Racusen, L. C., Solez, K., Colvin, R. B., Bonsib, S. M., Castro, M. C., Cavallo, T., Croker, B. P., Demetris, A. J., Drachenberg, C. B., Fogo, A. B., e Others (1999). The banff 97 working classification of renal allograft pathology. *Kidney international*, 55(2):713–723.
- Rasband, W. S. (2011). Imagej, us national institutes of health, bethesda, maryland, usa.
- Rosman, F. C. (2012). Evolução histórica no estudo histocitológico “os pioneiros”. *Departamento de Patologia Faculdade de Medicina Universidade Federal do Rio de Janeiro - Serviço de Anatomia Patológica Hospital Municipal Jesus Rio de Janeiro, RJ*, página 172. Acesso: 19 nov 2021.
- Roufosse, C., Simmonds, N., Clahsen-Van Groningen, M., Haas, M., Henriksen, K. J., Horsfield, C., Loupy, A., Mengel, M., Perkowska-Ptasińska, A., Rabant, M., e Others (2018). A 2018 reference guide to the banff classification of renal allograft pathology. *Transplantation*, 102(11):1795.
- Santos, J. P. e Scarano, S. (2021). Explorando quantitativamente observações feitas com espectrógrafos caseiros por meio de recursos digitais. *Revista Brasileira de Ensino de Física*, 43.
- Santos, V. (2021). Rins. Disponível: <<https://brasilecola.uol.com.br/biologia/os-rins.htm>>. Acesso: 21 nov 2021.
- Schouten, H. J. (1986). Nominal scale agreement among observers. *Psychometrika*, 51(3):453–466.
- Science, M. C. e Laboratory, A. I. (2008). LabelMe. Disponível: <<http://labelme.csail.mit.edu/>>. Acesso: 02 fev 2021.
- Sekachev, B., Nikita, M., e Andrey, Z. (2019). Computer vision annotation tool: a universal approach to data annotation. *Intel [Internet]*, 1.
- Sette, L., Titan, S., e Abensur, H. (2010). Doença renal crônica. Disponível: <https://www.medicinanet.com.br/conteudos/revisoes/2518/doenca_renal_cronica.htm>. Acesso: 24 out 2021.

- Shapiro, L. G., Stockman, G. C., e Others (2001). *Computer vision*, volume 3. Prentice Hall New Jersey.
- Shashiprakash, A. K., Lutnick, B., Ginley, B., Govind, D., Lucarelli, N., Jen, K.-Y., Rosenberg, A. Z., Urisman, A., Walavalkar, V., Zuckerman, J. E., e Others (2021). A distributed system improves inter-observer and ai concordance in annotating interstitial fibrosis and tubular atrophy. In *Medical Imaging 2021: Digital Pathology*, volume 11603, páginas 208–213. SPIE.
- Silva, J. H. e Lampkowski, M. (2022). Integração de sistemas de informação: a importância das integrações e como realizá-las com interoperabilidade de forma a produzirem e consumirem informações reutilizáveis.
- Silva, M. F. F. d. (2007). Esclarecendo o significado de “cor” em física. *A Física na Escola*, 8(1):25.
- Smith, T. e Guild, J. (1931). The cie colorimetric standards and their use. *Transactions of the optical society*, 33(3):73.
- Solez, K., Colvin, R., Racusen, L., Haas, M., Sis, B., Mengel, M., Halloran, P., Baldwin, W., Banfi, G., Collins, A., e Others (2008). Banff 07 classification of renal allograft pathology: updates and future directions. *American journal of transplantation*, 8(4):753–760.
- São Paulo, S. D. E. D. S. (2021). Como cuidar dos seus rins e prevenir a doença renal. Disponível: <<http://www.saude.sp.gov.br/ses/perfil/cidadao/orientacoes-gerais-sobre-saude/como-cuidar-dos-seus-rins-e-prevenir-a-doenca-renal>>. Acesso: 16 mai 2021.
- Taher, F., Mahmoud, A., Shalaby, A., e El-Baz, A. (2018). A review on the cerebrovascular segmentation methods. In *2018 IEEE INTERNATIONAL SYMPOSIUM ON SIGNAL PROCESSING AND INFORMATION TECHNOLOGY (ISSPIT)*, páginas 359–364. IEEE.
- Tinsley, H. E. e Weiss, D. J. (1975). Interrater reliability and agreement of subjective judgments. *Journal of Counseling Psychology*, 22(4):358.
- Torres, A. C. O., Dos Santos Neves, B. H., Lembi, B. K., Da Silva Isidoro, C. E., Fernandes, F. P., Oliveira, G. T. Q., De Castro Fischer, I., Batista, M. R. F., De Lanas, P. C. R., e Coutinho, M. N. (2021). Complicações pós-operatórias em pacientes submetidos a transplante renal: uma revisão narrativa. *Revista Eletrônica Acervo Saúde*, 13(4):e6891–e6891.
- University Of Utah (2022). Kidney anatomy. Disponível: <<https://webpath.med.utah.edu/RENAHTML/RENAL116.html>>. Acesso: 25 ago 2022.

- Van De Vlekkert, D., Machado, E., e D'azzo, A. (2020). Analysis of generalized fibrosis in mouse tissue sections with masson's trichrome staining. *Bio-protocol*, 10(10).
- Vergni, L., Todisco, F., e Di Lena, B. (2021). Evaluation of the similarity between drought indices by correlation analysis and cohen's kappa test in a mediterranean area. *Natural Hazards*, páginas 1–23.
- Wright, W. D. (1929). A re-determination of the trichromatic coefficients of the spectral colours. *Transactions of the Optical Society*, 30(4):141.
- Zheng, Y., Cassol, C. A., Jung, S., Veerapaneni, D., Chitalia, V. C., Ren, K. Y., Bellur, S. S., Boor, P., Barisoni, L. M., Waikar, S. S., e Others (2021). Deep-learning-driven quantification of interstitial fibrosis in digitized kidney biopsies. *The American journal of pathology*, 191(8):1442–1453.