



Universidade de São Paulo
Instituto de Física de São Carlos
Programa Interunidades de Pós-Graduação em Bioinformática

Juliana Virginio da Silva

Análise e Reconhecimento de Padrões em Feixes Vasculares Vegetais

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

São Carlos - SP

2021

Juliana Virginio da Silva

Análise e Reconhecimento de Padrões em Feixes Vasculares Vegetais

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Bioinformática da USP, como parte dos requisitos necessários para a obtenção do Título de “Mestre em Ciências”, no Programa: Bioinformática.

Versão Corrigida

Universidade de São Paulo – USP

Instituto de Física de São Carlos – IFSC

Programa Interunidades de Pós-Graduação em Bioinformática

Orientador: Prof. Dr. Luciano da Fontoura Costa

São Carlos - SP

2021

Juliana Virginio da Silva

Análise e Reconhecimento de Padrões em Feixes Vasculares Vegetais/ Juliana
Virginio da Silva. – São Carlos - SP, 2021-
206 p. : il.

Orientador: Prof. Dr. Luciano da Fontoura Costa

Dissertação (Mestrado) – **Versão Corrigida**

Universidade de São Paulo – USP

Instituto de Física de São Carlos – IFSC

Programa Interunidades de Pós-Graduação em Bioinformática, 2021.

1. Visão computacional 2. Feixes vasculares vegetais 3. Histologia vegetal 4.
Atributos I. Orientador: Prof. Dr. Luciano da Fontoura Costa II. Universidade de
São Paulo. IV. Título

*Aos meus pais Donizete e Sueli
e ao meu irmão Rafael,
pelo apoio incondicional
e por todos os valores transmitidos.
Gratidão à vocês.*

Agradecimentos

Agradeço a Deus pelo dom da vida.

Agradeço aos meus pais, Donizete e Sueli, por todo amor, carinho e valores transmitidos. Ao meu irmão Rafael por ser minha força e motivação. Agradeço a vocês pelo apoio incondicional e pela compreensão dos momentos de ausência. Obrigada por compartilharem comigo esse sonho, por acreditarem que ele era possível e por não medirem esforços para a realização. Vocês são minha base!

Meus sinceros agradecimentos a toda minha família pelo apoio, pela torcida, e pelo incentivo. Obrigada pelos momentos de união, afeto, carinho e amor. Agradeço por sempre acreditarem em mim e não me deixarem desistir dos meus sonhos.

Agradeço ao meu orientador, Prof. Dr. Luciano da Fontoura Costa, por todos os conselhos, pela paciência e ajuda nesse período. Obrigada pela oportunidade de desenvolver esse trabalho e por todo ensinamento compartilhado. Sem vocês esse trabalho não teria saído do papel.

Minha gratidão ao Prof. Dr. Cesar Henrique Comin por toda ajuda tanto na idealização quanto na execução desse trabalho.

Agradeço ao Prof. Dr. Marcos Arduin por ter aceitado fazer parte desse trabalho. Seus conhecimentos contribuíram de maneira significativa para o desenvolvimento desse trabalho. Obrigada pela oportunidade e por acreditar em mim.

Aos professores Prof. Dr. Francis Nunes, *Prof^a.Dr^a*. Karina Nogueira, *Prof^a.Dr^a*. Fernanda Anibal, *Prof^a.Dr^a*. Ana Afonso e *Prof^a.Dr^a*. Luciana Regitano: muito obrigada por abrirem as portas para mim, por acreditarem no meu trabalho e, principalmente, por compartilharem comigo o amor pela ciência, pela docência e pela educação. Vocês são meus exemplos e minhas inspirações.

Agradeço também a todos os professores que cruzaram o meu caminho e que com amor e dedicação ministraram ensinamentos que vão além da grade curricular. Vocês foram não só a base para minha formação profissional como também me impulsionaram a sonhar e me deram asas para voar.

Minha gratidão a Carla Moreira. Obrigada por me acompanhar nessa jornada. Por me incentivar e não me deixar desistir. Obrigada pelos conselhos, pelo apoio e por todo carinho.

Meu agradecimento aos presentes que a Biotec 012 me deu: Karen Carvalho, Rebeca Cordellini e Rodrigo Odone. Vocês são parte de mim, parte da minha história e parte de

quem eu me tornei. Obrigada por tudo ao longo desses anos: os conselhos, as conversas, as risadas, as noites de estudos e tudo que vivemos. Eu não chegaria aqui sem vocês.

Aos amigos que cruzaram meu caminho durante toda a vida e sempre estarão presentes: Ana Cordeiro, Ariadne Marino, Julia Matias, Juliana Afonso, Matheus Souza e Samanta Malvino. Obrigada pela amizade, pelo incentivo e apoio.

Meu agradecimento aos meus amigos de laboratório do Grupo de Pesquisa em Visão Cibernética. Obrigada pelas ajudas, discussões e apoio. Eu aprendi muito com vocês.

Minha gratidão a todo corpo de funcionários do Instituto de Física de São Carlos, principalmente, a Livia Ricci Costa e a Isabel de Cássia Sertori. Obrigada por todo auxílio com as documentações, informações, apoio e torcida. De maneira análoga, gostaria de agradecer ao corpo de funcionários do Instituto de Matemática e Estatística especialmente a secretaria do programa de Bioinformática por todo apoio e informações.

Agradeço aos representantes discentes do programa de Bioinformática, Paulo Burke e Raquel Miranda. Obrigada por serem a nossa voz e a ponte entre os alunos e a coordenação. Obrigada por lutarem pelos nossos direitos e pela qualidade da educação. Obrigada também pela disponibilidade, disposição e pelo auxílio com todas as normas e diretrizes do programa.

Agradeço a todos os funcionários, profissionais e alunos que passaram pela minha vida acadêmica e contribuíram de maneira direta ou indireta para minha formação profissional e pessoal. Todos os conselhos e ensinamentos transmitidos a mim me auxiliaram a crescer. Sou extremamente grata a cada um de vocês.

À Capes pelo apoio financeiro para realização deste trabalho de pesquisa.

Obrigada ao universo pelos reencontros, desencontros e encontros desta vida. Obrigada pelas lições valiosas ao longo desse período.

Meu muito obrigada e minha eterna gratidão a todos vocês!!

*“Por que tens medo de piar?
Precisa fazer o mundo
ouvir a sua voz.”
(Rodrigo Oliveira Odone)*

Resumo

Os tecidos vasculares ou tecidos condutores são os responsáveis pelo transporte de água, sais minerais e nutrientes ao longo do vegetal. A literatura evidencia a existência de cinco tipos de feixes vasculares vegetais sendo eles: colateral aberto, colateral fechado, bicolateral, anficrival e anfivasal. Tradicionalmente, a identificação desses tipos de feixes é feita por especialistas em anatomia vegetal. Contudo, os avanços na área de análise de imagens digitais vem abrindo caminho para que estudo morfológico vegetal ganhem novas ferramentas de análise. Os avanços na ciência também têm promovido a interdisciplinaridade em pesquisas e isso tem gerado novas aplicações e abordagens. Partindo dessa premissa, este trabalho analisou imagens microscópicas dos diversos feixes vasculares com o objetivo de promover um estudo capaz de 1) prover subsídios para segmentação em imagens microscópicas de vegetais e 2) contribuir para trabalhos de identificação tecidual. A aquisição das imagens foi feita por meio de microscopia óptica. As lâminas histológicas preparadas com cortes transversais de caules frescos, de espécies representativas de cada tipo de feixe vascular foram visualizadas e digitalizadas. As imagens foram segmentadas manualmente e foram extraídos atributos de quatro categorias: Forma, densidade, regularidade estrutural e multiescala. As matrizes contendo os valores obtidos foram submetidas a análise individual, análise par a par e análise de componentes principais, com o objetivo de selecionar atributos relevantes para a caracterização de tecidos vasculares. Todas as categorias de atributos foram avaliadas pelas abordagens *k-Nearest Neighbors* e Perceptron Multicamada, tanto em suas composições originais quanto após aplicação da análise de componente principal. Ao final, um conjunto de atributos foi selecionados. Esse conjunto obteve 91,67% de acertos em ambos classificadores ou seja, foi o que obteve melhor desempenho classificatório dos feixes vasculares vegetais. Um outro resultado relevante foi que o uso de atributos multiescala levou a uma porcentagem de acerto maior (81,66%) do que atributos tradicionalmente usados para descrever feixes vasculares, como por exemplo descritores de forma (70,84%). Ademais, nosso trabalho também estudou as possíveis explicações biológicas dos descritores selecionados. Isso indica que o presente trabalho foi capaz, não só, de elencar atributos capazes de classificar com bom desempenho os feixes vasculares vegetais, como também apresentar interpretações biológicas que potencialmente justifiquem os resultados aqui apresentados.

Palavras-chaves: feixes vasculares; atributos; histologia vegetal; visão computacional.

Abstract

Vascular tissues or conducting tissues are responsible for transporting water, mineral salts and nutrients throughout the plant. The literature shows the existence of five types of vegetable vascular bundles, namely: open collateral, closed collateral, bicollateral, amphicrival and amphivasal. The science that studies plants tissues is known as plant histology. Traditionally, the identification of these types of bundles is done by specialists in plant anatomy. However, advances in the area of digital image analysis are paving the way for plant morphological studies to gain new analysis tools. Advances in science have also promoted interdisciplinarity in research and this has generated new applications and approaches. Based on this premise, this work analyzed microscopic images of the various vascular bundles in order to promote a study capable of 1) providing subsidies for segmentation in microscopic images of plants and 2) contributing to the work of tissue identification. The acquisition of the images was made by means of optical microscopy. We visualized and scanned the histological slides prepared with cross sections of fresh stems, of species representative of each type of vascular bundle. The images were manually segmented and attributes from four categories were extracted: Shape, density, structural regularity and multiscale. The matrices containing the attribute values were subjected to individual analysis, peer analysis and principal component analysis in order to select the attributes that were most relevant throughout the work. The four categories of attributes, the selected attributes and a set containing all the extracted descriptors were evaluated by the k-Nearest Neighbors and Multilayer Perceptron approaches, both in their original dimensions and after application of the main component analysis. In the end, the category that contained the selected attributes obtained the best classificatory performance of the vascular bundles, reaching 91.67% of correct answers in both classifiers. Another relevant result was that the multiscale category obtained a higher percentage of correct answers (81.66%) than categories traditionally used to describe vascular bundles, such as shape (70.84%). In addition, our work also studied the possible biological explanations of the selected descriptors. This indicates that the present work was able, not only, to list attributes capable of classifying plant vascular bundles with good performance, but also to present biological interpretations that potentially justify the results presented here.

Keywords: vascular bundles; attributes; plant histology; computer vision.

Lista de ilustrações

Figura 1 – Visão geral do trabalho apresentado.	6
Figura 2 – Divisão dos tecidos vegetais em dois grupos: Tecidos meristemáticos e Tecidos permanentes	10
Figura 3 – Cladograma ilustrando os marcos evolutivos vegetais	12
Figura 4 – Representação da direção da condução dos tecidos vasculares vegetais ao longo do organismo.	13
Figura 5 – Feixes vasculares vegetais do tipo colateral	18
Figura 6 – Feixes vasculares vegetais do tipo bicolateral	18
Figura 7 – Feixes vasculares vegetais do tipo concêntrico	19
Figura 8 – Microscópio óptico evidenciando suas partes constituintes e o caminho percorrido pela luz.	22
Figura 9 – Sistema de digitalização de imagens acoplado ao microscópio óptico. . .	23
Figura 10 – Imagem colorida e em tons de cinza.	25
Figura 11 – Imagem colorida com os histogramas das bandas R, G e B.	25
Figura 12 – Histograma da imagem em tons de cinza	26
Figura 13 – Matriz de <i>pixel</i> de uma imagem segmentada manualmente.	27
Figura 14 – Visão geral de todos os passos que compõe a metodologia desse trabalho.	35
Figura 15 – Feixes vasculares vegetais obtidos por meio da captura e digitalização das lâminas histológicas vegetais.	38
Figura 16 – Divisão das imagens digitalizadas dos feixes vasculares vegetais entre as etapas de análise (laranja) e validação (verde).	39
Figura 17 – Exemplo da segmentação manual em um tecido xilemático.	40
Figura 18 – Exemplo da segmentação manual do floema.	41
Figura 19 – Segmentação completa dos tecidos vasculares vegetais em um feixe do tipo anfigasal.	41
Figura 20 – Categoria de atributos extraídas.	42
Figura 21 – Contagem de <i>pixel</i> em uma imagem segmentada manualmente	43
Figura 22 – Casco convexo e resíduo celular	44
Figura 23 – Diâmetro celular	45
Figura 24 – Diâmetro celular equivalente	46
Figura 25 – Elongação celular	47
Figura 26 – Centro de massa ou centroide	48
Figura 27 – Densidade radial do centroide	49
Figura 28 – Lacunaridade	50
Figura 29 – Poligonalidade.	52
Figura 30 – <i>Flowchart</i> com o passo a passo genérico da etapa de análises de atributos.	53

Figura 31 – Gráfico de densidade de probabilidade para três espécies vegetais. . . .	55
Figura 32 – Gráficos de dispersão referente as variáveis ilustrativas A, B, C e D, tomadas par a par.	57
Figura 33 – Visualização das mudanças espaciais antes e depois da aplicação da <i>ACP</i> .	62
Figura 34 – <i>Flowchart</i> com o passo a passo genérico da etapa de validação.	64
Figura 35 – Exemplo da representação do <i>k-NN</i> para classificar três indivíduos: Dispersão dos indivíduos já classificados nas espécies 1,2 e 3 e dos novos indivíduos <i>A, B e C</i>	65
Figura 36 – Exemplo da representação do <i>k-NN</i> para classificar três indivíduos: Distância dos pontos <i>A, B e C</i> em relação aos seus respectivos três vizinhos mais próximos.	66
Figura 37 – Exemplo da topologia de uma <i>PMC</i> com duas camadas intermediárias.	69
Figura 38 – Esquema de uma topologia geral para a classificação do <i>dataset</i> Iris. . .	72
Figura 39 – Matrizes de confusão após aplicação do método <i>k-fold cross-validation</i> com <i>k = 10</i> para escolha da melhor topologia.	73
Figura 40 – Matriz de confusão da etapa de teste para validação da acurácia da rede <i>PMC</i>	74
Figura 41 – Fluxograma indicando a organização dos resultados.	75
Figura 42 – Fluxograma com as divisões de classes e as etapas da análise individual.	76
Figura 43 – Função densidade probabilidade, para os atributos de forma, relativa aos feixes vasculares estudados.	77
Figura 44 – Função densidade probabilidade relativa aos feixes concêntricos com ênfase no floema.	79
Figura 45 – Função densidade probabilidade relativa aos feixes concêntricos com ênfase no xilema.	81
Figura 46 – Função densidade probabilidade relativa aos feixes colaterais com ênfase no floema.	83
Figura 47 – Função densidade probabilidade relativa aos feixes colaterais com ênfase no floema.	86
Figura 48 – <i>Colormap</i> representando a importância de cada atributo considerando a análise individual.	87
Figura 49 – Fluxograma com as divisões das classes estudadas e as etapas da análise Par a Par.	88
Figura 50 – Mapas da correlação de Pearson referente a todas as classes teciduais estudadas.	90
Figura 51 – Mapas da correlação de Pearson referente ao floema dos feixes do tipo concêntrico.	91
Figura 52 – Mapas da correlação de Pearson referente ao xilema dos feixes do tipo concêntrico.	93

Figura 53 – Mapas da correlação de Pearson referente ao floema dos feixes do tipo colateral.	94
Figura 54 – Mapas da correlação de Pearson referente ao xilema dos feixes do tipo colateral.	96
Figura 55 – Mapas da correlação de Pearson referente aos tecidos vasculares dos feixes do tipo bicolateral.	97
Figura 56 – Gráfico de barras com os atributos mais relevantes quando foram estudados todos os feixes vasculares.	101
Figura 57 – Gráfico de barras com os atributos mais relevantes quando foram estudados os floemas dos feixes vasculares do tipo concêntrico.	102
Figura 58 – Gráfico de barras com os atributos mais relevantes quando foram estudados os xilemas dos feixes vasculares do tipo concêntrico.	104
Figura 59 – Gráfico de barras com os atributos mais relevantes quando foram estudados os floemas dos feixes vasculares do tipo colateral.	105
Figura 60 – Gráfico de barras com os atributos mais relevantes quando foram estudados os xilemas dos feixes vasculares do tipo colateral.	107
Figura 61 – Gráfico de barras com os atributos mais relevantes quando foram estudados os feixes vasculares do tipo bicolateral.	108
Figura 62 – Gráfico de barras com os atributos mais relevantes considerando os resultados obtidos na análise par a par.	109
Figura 63 – Fluxograma com as divisões das classes estudadas e as etapas da ACP.	110
Figura 64 – Gráfico de dispersão obtido através da projeção ACP para todos os feixes vasculares estudados.	113
Figura 65 – Gráfico de dispersão obtido através da projeção ACP para os floemas dos feixes vasculares do tipo concêntrico.	116
Figura 66 – Gráfico de dispersão obtido através da projeção ACP para os xilemas dos feixes vasculares do tipo concêntrico.	119
Figura 67 – Gráfico de dispersão obtido através da projeção ACP para os floemas dos feixes vasculares do tipo colateral.	122
Figura 68 – Gráfico de dispersão obtido através da projeção ACP para os xilemas dos feixes vasculares do tipo colateral.	125
Figura 69 – Gráfico de dispersão obtido através da projeção ACP para os feixes vasculares do tipo bicolateral.	128
Figura 70 – <i>Colormap</i> representando a importância de cada atributo considerando a análise ACP.	129
Figura 71 – Detalhamento da de filtragem dos resultados obtidos na etapa de análise desse trabalho.	130
Figura 72 – <i>Colormap</i> representando a importância de cada atributo considerando as três abordagens aplicadas.	131

Figura 73 – Gráfico de barras com os acertos da abordagem <i>k-NN</i> para as categorias estudadas.	133
Figura 74 – Gráfico de barras com os acertos da abordagem <i>PMC</i> para as categorias estudadas.	136
Figura 75 – Gráfico de barras com os acertos médios das abordagens <i>k-NN</i> e <i>PMC</i> para as categorias estudadas.	137
Figura 76 – Função densidade probabilidade, para os atributos de forma, relativa aos feixes concêntricos com ênfase no floema.	156
Figura 77 – Função densidade probabilidade, para os atributos de densidade, relativa aos feixes concêntricos com ênfase no floema.	157
Figura 78 – Função densidade probabilidade do atributo poligonalidade média, relativa aos feixes concêntricos com ênfase no floema.	157
Figura 79 – Função densidade probabilidade, para os atributos de forma, relativa aos feixes concêntricos com ênfase no xilema.	158
Figura 80 – Função densidade probabilidade, para os atributos de densidade, relativa aos feixes concêntricos com ênfase no xilema.	159
Figura 81 – Função densidade probabilidade, para os atributos de regularidade, relativa aos feixes concêntricos com ênfase no xilema.	160
Figura 82 – Função densidade probabilidade, para os atributos de forma, relativa aos feixes colaterais com ênfase no floema.	161
Figura 83 – Função densidade probabilidade, para os atributos de densidade, relativa aos feixes colaterais com ênfase no floema.	161
Figura 84 – Função densidade probabilidade do atributo Lacunaridade com raio 6, relativa aos feixes colaterais com ênfase no floema.	162
Figura 85 – Função densidade probabilidade, para os atributos de forma, relativa aos feixes colaterais com ênfase no xilema.	163
Figura 86 – Função densidade probabilidade, para os atributos de densidade, relativa aos feixes colaterais com ênfase no xilema.	164
Figura 87 – Função densidade probabilidade, para os atributos de regularidade, relativa aos feixes colaterais com ênfase no xilema.	165
Figura 88 – Função densidade probabilidade, para os atributos de forma, relativa aos feixes bicolaterais	166
Figura 89 – Função densidade probabilidade, para o atributo Uniformidade espacial com sigma 30.	166
Figura 90 – Função densidade probabilidade, para os atributos de regularidade, relativa aos feixes bicolaterais.	167
Figura 91 – Gráficos de dispersão com os melhores atributos capazes de segregar os feixes vasculares vegetais estudados.	169

Figura 92 – Gráficos de dispersão com os melhores atributos capazes de segregar os feixes vasculares vegetais estudados (continuação).	170
Figura 93 – Gráficos de dispersão com os melhores atributos capazes de segregar os floemas dos feixes vasculares vegetais do tipo concêntrico.	171
Figura 94 – Gráficos de dispersão com os melhores atributos capazes de segregar os floemas dos feixes vasculares vegetais do tipo concêntrico (continuação).	172
Figura 95 – Gráficos de dispersão com os melhores atributos capazes de segregar os xilemas dos feixes vasculares vegetais do tipo concêntrico.	173
Figura 96 – Gráficos de dispersão com os melhores atributos capazes de segregar os xilemas dos feixes vasculares vegetais do tipo concêntrico (continuação).	174
Figura 97 – Gráficos de dispersão com os melhores atributos capazes de segregar os floemas dos feixes vasculares vegetais do tipo colateral.	175
Figura 98 – Gráficos de dispersão com os melhores atributos capazes de segregar os floemas dos feixes vasculares vegetais do tipo colateral (continuação).	176
Figura 99 – Gráficos de dispersão com os melhores atributos capazes de segregar os xilemas dos feixes vasculares vegetais do tipo colateral.	177
Figura 100 – Gráficos de dispersão com os melhores atributos capazes de segregar os xilemas dos feixes vasculares vegetais do tipo colateral (continuação).	178
Figura 101 – Gráficos de dispersão com os melhores atributos capazes de segregar os feixes vasculares vegetais do tipo bicolateral.	179
Figura 102 – Gráficos de dispersão com os melhores atributos capazes de segregar os feixes vasculares vegetais do tipo bicolateral (continuação).	180
Figura 103 – Matriz de confusão da abordagem k - NN para a categoria Forma no espaço original.	181
Figura 104 – Matriz de confusão da abordagem k - NN para a categoria Forma após aplicação da ACP com dois componentes principais.	182
Figura 105 – Matriz de confusão da abordagem k - NN para a categoria Forma após aplicação da ACP com três componentes principais.	182
Figura 106 – Matriz de confusão da abordagem k - NN para a categoria Densidade no espaço original.	183
Figura 107 – Matriz de confusão da abordagem k - NN para a categoria Densidade após aplicação da ACP com dois componentes principais.	184
Figura 108 – Matriz de confusão da abordagem k - NN para a categoria Densidade após aplicação da ACP com três componentes principais.	184
Figura 109 – Matriz de confusão da abordagem k - NN para a categoria Regularidade estrutural no espaço original.	185
Figura 110 – Matriz de confusão da abordagem k - NN para a categoria Regularidade estrutural após aplicação da ACP com dois componentes principais.	186

Figura 111 – Matriz de confusão da abordagem k -NN para a categoria Regularidade estrutural após aplicação da ACP com três componentes principais. . .	186
Figura 112 – Matriz de confusão da abordagem k -NN para a categoria Multiescala no espaço original.	187
Figura 113 – Matriz de confusão da abordagem k -NN para a categoria Multiescala após aplicação da ACP com dois componentes principais.	188
Figura 114 – Matriz de confusão da abordagem k -NN para a categoria Multiescala após aplicação da ACP com três componentes principais.	188
Figura 115 – Matriz de confusão da abordagem k -NN para a categoria <i>Todos</i> os atributos no espaço original.	189
Figura 116 – Matriz de confusão da abordagem k -NN para a categoria <i>Todos</i> os atributos após aplicação da ACP com dois componentes principais. . .	190
Figura 117 – Matriz de confusão da abordagem k -NN para a categoria <i>Todos</i> os atributos após aplicação da ACP com três componentes principais. . .	190
Figura 118 – Matriz de confusão da abordagem k -NN para a categoria <i>Selecionados</i> no espaço original.	191
Figura 119 – Matriz de confusão da abordagem k -NN para a categoria <i>Selecionados</i> após aplicação da ACP com dois componentes principais.	192
Figura 120 – Matriz de confusão da abordagem k -NN para a categoria <i>Selecionados</i> após aplicação da ACP com três componentes principais.	192
Figura 121 – Matriz de confusão da abordagem PMC para a categoria Forma no espaço original.	194
Figura 122 – Matriz de confusão da abordagem PMC para a categoria Forma após aplicação da ACP com dois componentes principais.	195
Figura 123 – Matriz de confusão da abordagem PMC para a categoria Forma após aplicação da ACP com três componentes principais.	195
Figura 124 – Matriz de confusão da abordagem PMC para a categoria Densidade no espaço original.	196
Figura 125 – Matriz de confusão da abordagem PMC para a categoria Densidade após aplicação da ACP com dois componentes principais.	197
Figura 126 – Matriz de confusão da abordagem PMC para a categoria Densidade após aplicação da ACP com três componentes principais.	197
Figura 127 – Matriz de confusão da abordagem PMC para a categoria Regularidade estrutural no espaço original.	198
Figura 128 – Matriz de confusão da abordagem PMC para a categoria Regularidade estrutural após aplicação da ACP com dois componentes principais. . .	199
Figura 129 – Matriz de confusão da abordagem PMC para a categoria Regularidade estrutural após aplicação da ACP com três componentes principais. . .	199

Figura 130–Matriz de confusão da abordagem PMC para a categoria Multiescala no espaço original.	200
Figura 131–Matriz de confusão da abordagem PMC para a categoria Multiescala após aplicação da ACP com dois componentes principais.	201
Figura 132–Matriz de confusão da abordagem PMC para a categoria Multiescala após aplicação da ACP com três componentes principais.	201
Figura 133–Matriz de confusão da abordagem PMC para a categoria <i>Todos</i> os atributos no espaço original.	202
Figura 134–Matriz de confusão da abordagem PMC para a categoria <i>Todos</i> os atributos após aplicação da ACP com dois componentes principais. . .	203
Figura 135–Matriz de confusão da abordagem PMC para a categoria <i>Todos</i> os atributos após aplicação da ACP com três componentes principais. . .	203
Figura 136–Matriz de confusão da abordagem PMC para a categoria <i>Selecionados</i> no espaço original.	204
Figura 137–Matriz de confusão da abordagem PMC para a categoria <i>Selecionados</i> após aplicação da ACP com dois componentes principais.	205
Figura 138–Matriz de confusão da abordagem PMC para a categoria <i>Selecionados</i> após aplicação da ACP com três componentes principais.	205

Lista de tabelas

Tabela 1 – Quadro resumido relacionando as espécies vegetais com seus feixes vasculares correspondentes.	36
Tabela 2 – Divisão dos tecidos vasculares com base no tipo do feixe vascular. . . .	54
Tabela 3 – Valores dos diâmetros caulinares (x) dos espécimes vegetais (esp1, esp2 e esp3).	55
Tabela 4 – Valores do traço das matrizes R dos gráficos selecionados	58
Tabela 5 – Valores da distância entre os pontos a serem classificados e os respectivos três vizinhos mais próximos	66
Tabela 6 – Exemplo de matriz de confusão para a classificação dos indivíduos A , B e C por meio do algoritmo k - NN	67
Tabela 7 – Topologias testadas para a classificação do <i>dataset</i> Iris	72
Tabela 8 – Atributos que foram relevantes para segregar tecidos do floema em feixes vasculares concêntricos.	78
Tabela 9 – Atributos que foram relevantes para segregar tecidos do xilema em feixes vasculares concêntricos.	80
Tabela 10 – Atributos que foram relevantes para segregar tecidos do floema em feixes vasculares colaterais.	82
Tabela 11 – Atributos que foram relevantes para segregar tecidos do xilema em feixes vasculares colaterais.	84
Tabela 12 – Atributos que foram relevantes para segregar tecidos dos feixes vasculares bicolaterais.	85
Tabela 13 – Correspondência dos atributos com os índices.	99
Tabela 14 – Tabela com os valores de $t(R)$ para todos os feixes vasculares vegetais estudados.	100
Tabela 15 – Tabela com os valores de $t(R)$ para gráficos de dispersão referentes ao floema dos feixes vasculares do tipo concêntrico.	102
Tabela 16 – Tabela com os valores de $t(R)$ para gráficos de dispersão referentes ao xilema dos feixes vasculares do tipo concêntrico.	103
Tabela 17 – Tabela com os valores de $t(R)$ para gráficos de dispersão referentes ao floema dos feixes vasculares do tipo colateral.	105
Tabela 18 – Tabela com os valores de $t(R)$ para gráficos de dispersão referentes ao xilema dos feixes vasculares do tipo colateral.	106
Tabela 19 – Tabela com os valores de $t(R)$ para gráficos de dispersão referentes aos feixes vasculares do tipo bicolateral.	108
Tabela 20 – Pesos absolutos para a projeção de ACP referente a todas as classes. .	112

Tabela 21 – Pesos absolutos para a projeção de ACP referente aos floemas dos feixes concêntricos.	115
Tabela 22 – Pesos absolutos para a projeção de ACP referente aos xilemas dos feixes concêntricos.	118
Tabela 23 – Pesos absolutos para a projeção de ACP referente aos floemas dos feixes colateral.	121
Tabela 24 – Pesos absolutos para a projeção de ACP referente aos xilemas dos feixes colateral.	124
Tabela 25 – Pesos absolutos para a projeção de ACP referente aos feixes bicolateral.	127
Tabela 26 – Acurácia obtida por meio da aplicação da abordagem k -NN frente as categorias estudadas.	132
Tabela 27 – Topologias selecionadas	134
Tabela 28 – Acurácia obtida por meio da aplicação da abordagem PMC frente as categorias estudadas.	135

Lista de abreviaturas e siglas

ACP	Análise de componentes principais
ATP	<i>Adenosine TriPhosphate</i>
CCD	<i>Charge-Coupled Device</i>
Desv. P.	Desvio Padrão
Diâmetro eq.	Diâmetro equivalente
Esp1	Espécie 1
Esp2	Espécie 2
Esp3	Espécie 3
FNT	Formalina tamponada
FPA	Formalina, ácido propiônico e álcool etílico
k-NN	<i>K-Nearest Neighbors</i>
PCA	<i>Principal Component Analysis</i>
PMC	Perceptron Multicamada

Lista de símbolos

Ac	Acurácia
p	Grau de correlação de Pearson
α	Letra grega alfa
β	Letra grega beta
ε	Letra grega epsilon
Γ	Letra grega gama
λ	Letra grega lambda
σ	Letra grega sigma
S_{inter}	Matriz de espalhamento <i>entreclusters</i>
S_{intra}	Matriz de espalhamento <i>intraclusters</i>
μ	micro
\in	Pertence
T	Traço

Sumário

1	INTRODUÇÃO	1
1.1	Motivação e objetivos	4
1.2	Visão geral do trabalho apresentado	5
1.3	Organização da dissertação	7
2	O REINO VEGETAL	9
2.1	Histologia vegetal	10
2.2	Tecido vascular vegetal	12
2.2.1	Xilema	14
2.2.2	Floema	15
2.3	Feixes vasculares vegetais	17
3	ANÁLISE DE IMAGENS E RECONHECIMENTO DE PADRÕES	20
3.1	Biomimética e a visão artificial	20
3.2	Aquisição de imagens	21
3.2.1	O microscópio óptico	21
3.2.2	Digitalização de imagens microscópicas	23
3.3	A imagem digital	24
3.4	Processamento de imagens digitais	27
3.4.1	Segmentação de imagens digitais	28
3.4.2	Extração de atributos	29
3.4.3	Análise dos atributos	31
3.5	Reconhecimento e classificação de padrões	32
4	METODOLOGIA	34
4.1	Montagem das lâminas histológicas	36
4.2	Aquisição de imagens	37
4.3	Segmentação manual das imagens	39
4.4	Extração de atributos	42
4.4.1	Medidas de forma celular	43
4.4.1.1	Área celular	43
4.4.1.2	Resíduo celular e casco convexo	44
4.4.1.3	Diâmetro celular	44
4.4.1.4	Diâmetro celular equivalente	45
4.4.1.5	Elongação celular	46
4.4.2	Medidas de densidade	47

4.4.2.1	Uniformidade espacial da densidade dos centroides	48
4.4.2.2	Circularidade <i>isoline</i>	49
4.4.2.3	Densidade radial do centroide	49
4.4.3	Medidas de regularidade estrutural	50
4.4.3.1	Lacunaridade	50
4.4.3.2	Poligonalidade	51
4.5	Análise dos atributos	53
4.5.1	Análise individual	54
4.5.2	Análise par a par	56
4.5.3	Análise de componentes principais	59
4.6	Validação dos resultados	63
4.6.1	<i>k-Nearest Neighbor(k-NN)</i>	65
4.6.2	Perceptron Multicamada (<i>PMC</i>)	68
5	RESULTADOS	75
5.1	Etapa de análise	75
5.1.1	Análise individual	75
5.1.1.1	Todos os feixes vasculares	76
5.1.1.2	Feixes vasculares concêntricos	77
5.1.1.2.1	Floema nos feixes vasculares concêntricos	77
5.1.1.2.2	Xilema nos feixes vasculares concêntricos	79
5.1.1.3	Feixes vasculares colaterais	81
5.1.1.3.1	Floema nos feixes vasculares colaterais	81
5.1.1.3.2	Xilema nos feixes vasculares colaterais	83
5.1.1.4	Feixe vascular bicolateral	85
5.1.1.5	Considerações gerais	86
5.1.2	Análise Par a Par	88
5.1.2.1	Correlação de Pearson	88
5.1.2.1.1	Todos os feixes vasculares	89
5.1.2.1.2	Feixes vasculares do tipo concêntrico: Floema	90
5.1.2.1.3	Feixes vasculares do tipo concêntrico: Xilema	92
5.1.2.1.4	Feixes vasculares do tipo colateral: Floema	93
5.1.2.1.5	Feixes vasculares do tipo colateral: Xilema	95
5.1.2.1.6	Feixes vasculares do tipo bicolateral	96
5.1.2.2	Análise de dispersão: Matriz de espalhamento	98
5.1.2.2.1	Todos os feixes vasculares	100
5.1.2.2.2	Feixes vasculares do tipo concêntricos: Floema	101
5.1.2.2.3	Feixes vasculares do tipo concêntricos: Xilema	103
5.1.2.2.4	Feixes vasculares do tipo colateral: Floema	104
5.1.2.2.5	Feixes vasculares do tipo colateral: Xilema	106

5.1.2.2.6	Feixes vasculares do tipo bicolateral	107
5.1.2.3	Considerações gerais	109
5.1.3	Análise de componentes principais	110
5.1.3.1	Todos os feixes vasculares	111
5.1.3.2	Feixes vasculares do tipo concêntricos: Floema	114
5.1.3.3	Feixes vasculares do tipo concêntricos: Xilema	117
5.1.3.4	Feixes vasculares do tipo colateral: Floema	119
5.1.3.5	Feixes vasculares do tipo colateral: Xilema	122
5.1.3.6	Feixes vasculares do tipo bicolateral	126
5.1.3.7	Considerações gerais	129
5.2	Filtragem dos resultados	130
5.3	Validação	132
5.3.1	K-Nearest Neighbor(k-NN)	132
5.3.2	Perceptron Multicamada (<i>PMC</i>)	134
5.3.3	Considerações finais	137
6	DISCUSSÃO	139
7	CONCLUSÕES E TRABALHOS FUTUROS	145
7.1	Conclusões	145
7.1.1	Atributos selecionados	145
7.1.2	Atributos de forma	146
7.1.3	Abordagens pouco utilizadas	146
7.1.4	Relevância biológica	147
7.2	Trabalhos Futuros	147
	REFERÊNCIAS	148
	APÊNDICES	155
	APÊNDICE A – ANÁLISE INDIVIDUAL: FUNÇÃO DENSIDADE	
	PROBABILIDADE SELECIONADAS	156
A.1	Feixes vasculares concêntricos	156
A.1.1	Floema nos feixes vasculares concêntricos	156
A.1.2	Xilema nos feixes vasculares concêntricos	158
A.2	Feixes vasculares colaterais	160
A.2.1	Floema nos feixes vasculares colaterais	160
A.2.2	Xilema nos feixes vasculares colaterais	162
A.3	Feixe vascular bicolateral	165

	APÊNDICE B – GRÁFICOS DE DISPERSÃO SELECIONADOS	168
B.1	Todos os tecidos vasculares	169
B.2	Feixes vasculares do tipo concêntrico: Floema	171
B.3	Feixes vasculares do tipo concêntrico: Xilema	173
B.4	Feixes vasculares do tipo colateral: Floema	175
B.5	Feixes vasculares do tipo colateral: Xilema	177
B.6	Feixes vasculares do tipo bicolateral	179
	 APÊNDICE C – K-NN: MATRIZES DE CONFUSÃO	 181
C.1	Forma	181
C.1.1	Espaço original	181
C.1.2	Dimensionalidade reduzida: 2D	181
C.1.3	Dimensionalidade reduzida: 3D	182
C.2	Densidade	183
C.2.1	Espaço original	183
C.2.2	Dimensionalidade reduzida: 2D	183
C.2.3	Dimensionalidade reduzida: 3D	184
C.3	Regularidade estrutural	185
C.3.1	Espaço original	185
C.3.2	Dimensionalidade reduzida: 2D	185
C.3.3	Dimensionalidade reduzida: 3D	186
C.4	Multiescala	187
C.4.1	Espaço original	187
C.4.2	Dimensionalidade reduzida: 2D	187
C.4.3	Dimensionalidade reduzida: 3D	188
C.5	Todos	189
C.5.1	Espaço original	189
C.5.2	Dimensionalidade reduzida: 2D	189
C.5.3	Dimensionalidade reduzida: 3D	190
C.6	Selecionados	191
C.6.1	Espaço original	191
C.6.2	Dimensionalidade reduzida: 2D	191
C.6.3	Dimensionalidade reduzida: 3D	192
	 APÊNDICE D – PMC: MATRIZES DE CONFUSÃO	 194
D.1	Forma	194
D.1.1	Espaço original	194
D.1.2	Dimensionalidade reduzida: 2D	194
D.1.3	Dimensionalidade reduzida: 3D	195
D.2	Densidade	196

D.2.1	Espaço original	196
D.2.2	Dimensionalidade reduzida: 2D	196
D.2.3	Dimensionalidade reduzida: 3D	197
D.3	Regularidade estrutural	198
D.3.1	Espaço original	198
D.3.2	Dimensionalidade reduzida: 2D	198
D.3.3	Dimensionalidade reduzida: 3D	199
D.4	Multiescala	200
D.4.1	Espaço original	200
D.4.2	Dimensionalidade reduzida: 2D	200
D.4.3	Dimensionalidade reduzida: 3D	201
D.5	Todos	202
D.5.1	Espaço original	202
D.5.2	Dimensionalidade reduzida: 2D	202
D.5.3	Dimensionalidade reduzida: 3D	203
D.6	Selecionados	204
D.6.1	Espaço original	204
D.6.2	Dimensionalidade reduzida: 2D	204
D.6.3	Dimensionalidade reduzida: 3D	205

1 Introdução

Ao longo da história da humanidade, cientistas e filósofos tentaram organizar os indivíduos que apresentavam características em comum em grupos. Muitas divisões surgiram ao longo dos anos e, todas elas, foram influenciadas pela tecnologia que era vigente naquele contexto (CATTAL; MENEZES, 2010).

Um grande marco para a ciência ocorreu no século XVI com a chamada "*Revolução Científica*" a qual desenvolveu uma nova forma de olhar a natureza (HENRY, 1998). Um dos feitos mais notórios dessa época foi a invenção do microscópio pelos holandeses Hans Janssen e Zacharias Janssen. Esse instrumento foi utilizado por Robert Hooke em 1663 quando o inglês descobriu, ao analisar cortes finos de cortiça, pequenos espaços vazios os quais ele nomeou de célula (AMABIS; RODRIGUES, 2006).

Com a invenção do microscópio, a histologia passou a ser utilizada na busca de novos atributos que auxiliassem na elucidação dos problemas taxonômicos (STACE, 1991), uma vez que, ela estuda os tecidos e suas formas de organização (JUNQUEIRA; JUNQUEIRA, 1983).

No reino vegetal, dois tecidos que apresentam importância evolutiva significativa são o xilema e o floema, comumente conhecidos como tecidos condutores ou vasculares. Esses tecidos se organizam em forma de feixes vasculares e, foi por meio do surgimento desse sistema organizado de condução de água, sais minerais e nutrientes que os vegetais puderam alcançar definitivamente o meio terrestre resultando na possibilidade do crescimento sem prejuízo hídrico e energético para a planta (AMABIS; RODRIGUES, 2006; GLÓRIA; GUERREIRO, 2006; MODESTO; SIQUEIRA, 2002; RAVEN; EVERT; EICHHORN, 2014).

A identificação de espécies e tecidos vegetais é baseada na morfologia vegetal que se utiliza dos olhos treinados de botânicos para diferenciar as células, seus formatos e posições. Muitas vezes, a subjetividade do olho humano pode ocasionar erros ao classificar a morfologia das células.

É notório que desde a "*Revolução Científica*" a ciência está em constantes transformações e novas descobertas são feitas diariamente, sejam elas ferramentas, novos produtos, novas abordagens ou simplesmente novas aplicações para abordagens já existentes. Nesse cenário, o avanço da tecnologia na área computacional possibilitou a utilização da visão computacional para caracterizar tecidos vegetais.

Os avanços contínuos nos conceitos e métodos de análise de imagens, aliados às abordagens de reconhecimento de padrões e à análise de formas, abriram caminho para

a obtenção de caracterizações cada vez mais abrangentes e precisas dos tecidos vegetais (MAYOR; SILVA; SERENO, 2005; MAYOR; PISSARA; SERENO, 2008; RAMOS et al., 2004; SANYAL; BHATTACHARYA; BANDYOPAHYAY, 2008). Embora a busca por descritores microhistológicos pareça estar bem estabelecida na área de identificação e caracterização de tecidos e espécies vegetais e que os desenvolvimentos da ciência contribuam para aperfeiçoar os trabalhos nessa área, os estudos publicados baseiam-se principalmente em uma classe de descritores: medidas de forma (MAYOR; SILVA; SERENO, 2005; MAYOR; PISSARA; SERENO, 2008; MIAO; GANDELIN; YUAN, 2006; MURRAY; O'MALLEY, 1993; RAMOS et al., 2004; SANYAL; BHATTACHARYA; BANDYOPAHYAY, 2008). Esse viés metodológico pode ser entendido como uma herança visual advinda da histologia tradicional que é fundamentada nas diferenças de formato dos tecidos e de seus constituintes (MINERVINI; SCHARR; TSAFTARIS, 2015).

Abordar um assunto pelas metodologias padrões e bem estabelecidas de uma determinada área é uma prática muito comum no meio científico. Porém, segundo o filósofo e poeta Henry David Thoreau (1817 - 1862), "*A questão não é o que você olha, mas de que maneira você vê.*", em tradução livre. Na ciência, esta sentença nos faz refletir sobre o quão importante é a perspectiva sob a qual analisamos um problema e quais são as ferramentas ou abordagens que estamos utilizando para explorar de maneira satisfatória o assunto em questão.

Também é importante ressaltar que a caracterização precisa estar fundamentada na funcionalidade e nas particularidades do objeto a que se deseja estudar, ou seja, a escolha dos atributos e das abordagens aplicadas precisam estar alinhadas com as interpretações biológicas desses tecidos. Dessa maneira, a extração de descritores pode ser entendida como uma via de mão dupla pois, da mesma forma que ela se baseia nas individualidades de cada tecido ela também promove subsídios para novas interpretações tanto fisiológicas quanto anatômicas.

Existe uma parábola Hindu intitulada "Os cegos e o elefante" que retrata bem a importância da perspectiva para solucionar um problema (BAQUIÃO et al., 2019). De maneira resumida, a parábola ilustra a história de sete sábios cegos que têm um dilema em comum: conseguir descrever um animal desconhecido que havia chegado a cidade. O primeiro sábio apalpou a barriga e declarou que o animal era forte como uma parede. O segundo não concordou e disse, tocando a presa, que ele se parecia mais com uma lança. O terceiro, ao tocar a tromba, retrucou e disse que o animal era parecido com uma serpente. O quarto achou o animal semelhante a uma árvore enquanto apalpava o joelho. O quinto, por sua vez, tocou as orelhas e afirmou que o animal era parecido com uma cortina. O sexto, descontente com todos os outros, tocou o rabo do animal e declarou que ele era semelhante a uma corda. Os seis sábios ficaram discutindo sem entrar em um acordo até que o sétimo cego chegou e pediu para que uma criança desenhasse no chão a figura do

animal e representasse todas as perspectivas sentidas pelos outros sábios. Quando ele tateou o contorno do desenho, declarou: “É assim que os homens se comportam perante a verdade. Pegam apenas uma parte, pensam que é o todo, e continuam tolos!”.

A parábola apresentada nos faz refletir sobre a importância da união das diversas teorias para que o todo possa ser mais compreendido e estudado. Essa reflexão pode ser interpretada de diversas maneiras e por diferentes vieses. Na ciência, por exemplo, ela pode ser compreendida como se os cegos fossem áreas distintas e o elefante o objeto em estudo. Sendo assim, se as pesquisas estiverem focadas em apenas um campo do conhecimento, ela provavelmente chegará a menos conclusões do que outras que estiverem fundamentadas em mais de uma esfera.

Atualmente, a interdisciplinaridade vem abrindo caminho para o desenvolvimento de novos olhares e percepções. A literatura mostra que trabalhos interdisciplinares que unem a histologia vegetal com a análise de imagens e o reconhecimento de padrões têm apresentados resultados significativos e contribuído para: identificação de espécies (ROSITO; MARCHESAN, 2003), comparação entre tecidos (SÁNCHEZ-GUTIÉRREZ et al., 2016), análise na forma de plantas (MIAO; GANDELIN; YUAN, 2006) e até mesmo no estudo do movimento de células dentro dos tecidos vegetais (LIU et al., 2017). Além disso, segundo Richmond e Roehner (2019), um olhar físico é incomum em trabalhos escritos por profissionais da área de biologia porém, ele pode trazer conclusões não antes obtidas.

Uma das abordagens física bastante utilizada em problemas biológicos é a variação da escala que é fundamental para entender e descrever objetos tanto na percepção natural quanto na artificial (LINDEBERG, 2015). Esse processo, chamado de análise multiescala, tem sido amplamente utilizado na literatura física e biológica (COSTA; GOLDBERGER; PENG, 2005; LIU et al., 2017; SHANG et al., 2019; SONG et al., 2018; XING et al., 2017; YIN et al., 2019). No entanto, a aplicação de métodos multiescalares para caracterização de tecidos histológicos de plantas tem sido pouco discutida na literatura e geralmente está mais relacionada aos descritores de textura foliar (PLOTZE; BRUNO, 2007; SILVA; FLORINDO, 2019). Assim, há uma lacuna na exploração dos descritores originados da análise multiescalar para a caracterização de tecidos vegetais, como, por exemplo, os feixes vasculares vegetais.

Um olhar mais centrado nas particularidades dos feixes vegetais revela que uma das diferenças entre eles está na posição de um tecido vascular em relação ao outro (MODESTO; SIQUEIRA, 2002). Existem dois tipos de categorias de atributos que visam extrair informações relacionadas às posições das células de um tecido, são eles: atributos relacionados a densidade dos centroides e atributos relacionados a regularidade estrutural. A literatura evidencia, novamente, uma escassez de trabalhos que caracterizem tecidos vegetais por meio dessas abordagens.

Partindo desses cenários, nota-se uma oportunidade de desenvolvimento na ciência

das plantas no âmbito da análise de imagens e de reconhecimento de padrões. Uma vez que, pode ser explorado o efeito em se estudar atributos pouco utilizados na literatura para caracterizar feixes vasculares vegetais.

1.1 Motivação e objetivos

A caracterização de tecidos vegetais é uma prática que se repete desde a invenção do microscópio até o dias de hoje. Esse processo é importante para a classificação de um tecido, para classificação de uma espécie ou para o controle de qualidade de um vegetal.

Nota-se que os trabalhos que envolvem tecidos vegetais ainda estão fundamentados na histologia clássica e, portanto, caracterizam esses tecidos por meio de descritores de forma.

Os avanços tecnológicos nas abordagens de análise de imagens digitais e métodos de reconhecimento têm possibilitado uma maior abrangência na extração de atributos de imagens. Esses avanços têm contribuído positivamente para que particularidades possam ser melhores estudadas, ou seja, a caracterização anda junto com a funcionalidade de cada tecido e suas individualidades. Isso abre caminho para que novas interpretações possam ser obtidas.

Dentro desse cenário, e levando em consideração as particularidades de cada feixe vascular, o presente trabalho buscou explorar essas lacunas propondo estudar tecidos vegetais não apenas por atributos referente a forma mas também extrair e estudar descritores referente a densidade de centroide e a regularidade estrutural. Alguns atributos foram extraídos segundo a abordagem multiescala.

Essas novas abordagens, atrelada à tradicional, poderão contribuir para novas conclusões, ou seja, o presente trabalho se aproveita dos benefícios da análise digital de imagens e reconhecimento de padrões para melhor compreender os feixes vasculares vegetais, especialmente no sentido de alcançar resultados mais quantitativos e objetivos, gerando novas interpretações na caracterização desses tecidos.

Em particular, pretendemos encontrar descritores que contribuam para a diferenciação efetiva entre feixes vasculares vegetais provenientes do caule de cinco espécies vegetais: *Cyperus rotundus* (Cyperaceae), *Baccharis dracunculifolia* (Asteraceae), *Sechium edule* (Cucurbitaceae), *Dracaena marginata* (Convallariaceae), *Microgramma vacciniifolia* (Polypodiaceae). Cada espécie listada tem um tipo de feixe vascular e são eles, respectivamente, colateral aberto, colateral fechado, bicolateral, anfigasal e anficrival.

As imagens digitais de cada feixe vascular vegetal foram segmentadas manualmente para que os tecidos floema e xilema pudessem ser segregados. A segmentação manual foi a estratégia escolhida uma vez que tínhamos como objetivo construir um banco de dados

"Padrão Ouro" no quesito segmentação de tecidos vegetais em feixes vasculares.

Os descritores foram extraídos por meio de quatro abordagens: forma, densidade, regularidade estrutural e multiescala. Foi objetivo desse trabalho selecionar um conjunto de atributo que melhor caracterizasse os feixes vasculares vegetais. Ao final, todas as categoriais de atributos foram testadas com o objetivo de avaliar a acurácia de cada uma delas por meio de métodos de reconhecimento de padrões supervisionados. Assim, pode-se determinar qual categoria descreveu melhor os tecidos em estudo.

Ademais, o presente trabalho também objetivou encontrar a relevância biológica dos descritores que melhor caracterizassem os feixes vasculares vegetais. Assim, podemos auxiliar na compreensão da arquitetura desses tecidos e futuramente motivar estudos de cunho fisiológico e anatômico.

Sendo assim, esse trabalho foi motivado pelas lacunas encontradas na caracterização em tecidos vegetais e tem como objetivos específicos:

- a. Caracterizar os diferentes tipos de feixes vasculares vegetais;
- b. Diferenciar os diferentes tipos de feixes vasculares vegetais;
- c. Contribuir como base para trabalhos de identificação tecidual vascular vegetal;
- d. Prover subsídios para segmentação em imagens microscópicas de tecido vascular vegetal;
- e. Mostrar a relevância biológica dos atributos que melhor caracterizarem os feixes vasculares vegetais.

1.2 Visão geral do trabalho apresentado

A Figura 1 ilustra a visão geral de todas as partes que constituem esse trabalho. O trabalho se inicia com a montagem das lâminas histológicas e a aquisição das imagens digitais das amostras dos 3 tipos de feixes vasculares vegetais: feixes concêntricos, feixes colaterais e feixes bicolaterais. Os feixes concêntricos se subdividem em anfigival e anfigival. Os feixes colaterais se subdividem em abertos e fechados. Em nosso trabalho, também abordamos os anfigivais primários e secundários.

Ao todo, foram digitalizadas 25 imagens de cada um dos seis feixes estudados. Por meio da segmentação manual, cada figura foi dividida em seus tecidos constituintes, xilema e floema. Portanto, as 150 imagens iniciais deram origem a 300 imagens. Dessas 300, 240 foram destinadas para a etapa de análise (representada em laranja na Figura 1) e 60 para a etapa de validação (representada em verde na Figura 1). Todas as classes de tecidos possuíram o mesmo número de imagens nas duas etapas.

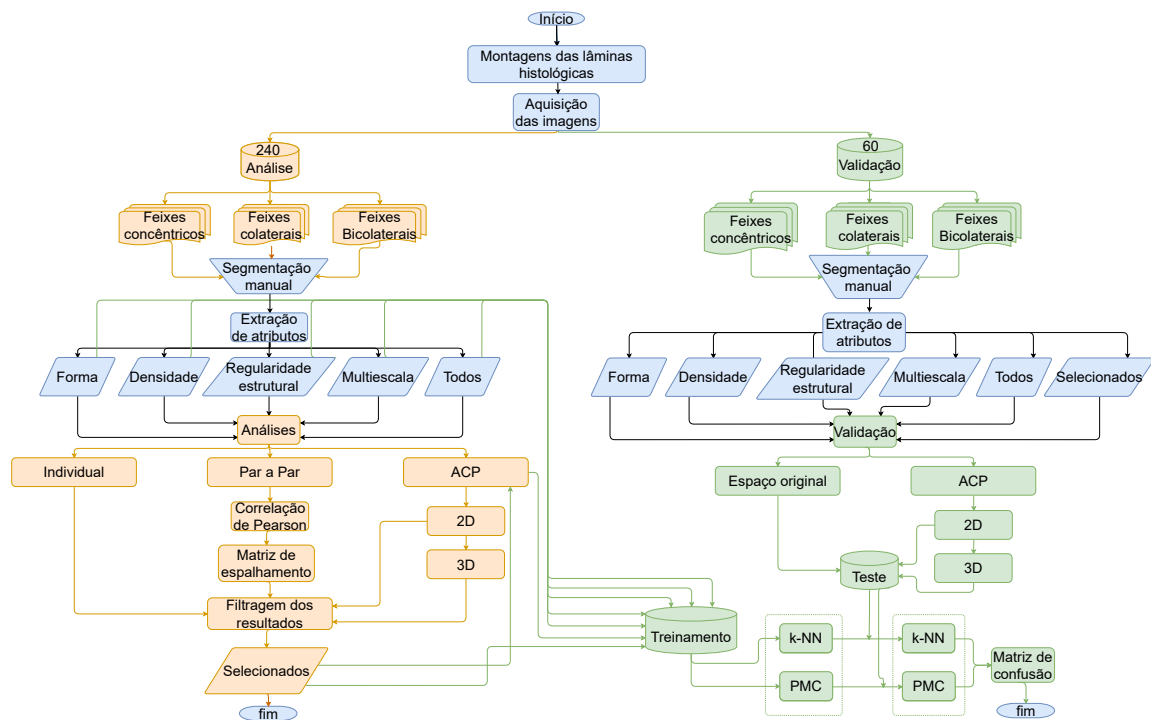


Figura 1 – Visão geral do trabalho apresentado. Os processos em laranja correspondem a etapa de análise, os em verde a etapa de validação e os azuis são referentes aos processos que são comuns as duas etapas.

Fonte: Autor

Na etapa de análise, foram extraídos cinco categorias de atributos: *Forma*, *Densidade*, *Regularidade estrutural*, *Multiescala* e *Todos* (sendo que essa última é composta por todos os atributos extraídos em todas as outras categorias). Os atributos foram extraídos com base não só na literatura como também em consideração às particularidades dos feixes em estudo.

O objetivo da etapa de análise foi estudar se os atributos extraídos eram capazes de segregar e caracterizar os tecidos vegetais. Por isso, os atributos extraídos foram analisados por três abordagens distintas: individual, par a par e análise de componentes principais (ACP). Ao final dessa etapa, os atributos foram selecionados com a finalidade de elencar quais foram os descritores que apresentaram maior influência na caracterização e separação dos feixes vasculares vegetais.

A partir dos resultados obtidos na etapa de análise, foi possível realizar a etapa de validação. Nela, a extração de atributos contou com a inserção da categoria *Selecionados* a qual foi resultado da etapa de análise. Ou seja, aqui os atributos extraídos foram organizados em seis categorias: *Forma*, *Densidade*, *Regularidade estrutural*, *Multiescala*, *Todos* e *Selecionados*.

A etapa de validação foi responsável por medir e comparar a acurácia das categorias de atributos frente ao objetivo geral desse trabalho que é a caracterização e a separação dos

feixes vasculares vegetais. A acurácia foi medida por meio de duas abordagens classificatórias: *k-Nearest Neighbors* e *Perceptron Multicamadas*. Os resultados foram apresentados na forma de matrizes de confusão.

Ao final do trabalho, foi possível analisar qual categoria de atributos foi capaz de caracterizar e separar os tecidos vegetais dos feixes vasculares vegetais com melhor média de acertos. A categoria que obteve a maior média de acertos foi a categoria *Selecionados*. Também foi observado que a categoria *Multiescala* obteve maior número de acertos do que a categoria *Forma*.

Ademais, os atributos que constituíram a categoria *Selecionados* foram estudados buscando encontrar interpretações biológicas que justifiquem o motivo pelo qual eles obtiveram o melhor desempenho computacional.

1.3 Organização da dissertação

O documento inicia-se com dois Capítulos (2 e 3) destinados à apresentar conceitos básicos que foram utilizados ao longo dessa dissertação: O primeiro evidencia a importância da classificação para segregar os vegetais como um grupo de organismos únicos e traz informações acerca da biologia vegetal como por exemplo, a histologia. O segundo é composto por conceitos básicos das áreas de análise de imagem e reconhecimento de padrões e traz informações sobre imagens digitais, processamento de imagens, extração e análise de atributos e, por fim, classificação de padrões.

O próximo Capítulo 4 apresenta toda metodologia realizada ao longo dessa dissertação. Nesse Capítulo foram apresentadas abordagens de extração e análise de atributos bem como algumas abordagens de reconhecimento de padrões. A primeira atividade realizada foi a montagem das lâminas histológicas, seguida da aquisição e da segmentação das imagens dos feixes vasculares vegetais. Posteriormente, os atributos foram extraídos. Após essa etapa, a dissertação é subdividida em duas etapas: a primeira é análise dos atributos e a segunda é a validação dos resultados. Na primeira, os atributos são avaliados por três abordagens (individual, par a par e análise de componente principal) com o objetivo de elencar um subconjunto de atributos que fosse mais relevante para a caracterização dos feixes vasculares vegetais. A segunda etapa consistiu em avaliar a acurácia por meio de classificadores supervisionados (*k-Nearest Neighbor*) e *Perceptron Multicamada*) de todos os conjuntos de atributos estudados nessa dissertação.

O Capítulo 5 apresenta os resultados ao longo dessa dissertação, ou seja, são apresentados os resultados referentes a etapa de análise e de validação.

O Capítulo 6 apresenta a discussão sobre as categorias de atributos estudadas ao longo desse trabalho. Esse capítulo também traz informações referentes a relevância

biológica dos atributos que melhor caracterizaram os feixes vasculares vegetais.

Por fim, a dissertação é encerrada com o Capítulo 7 que traz as conclusões sobre o trabalho desenvolvido e as possibilidades para trabalhos futuros.

2 O Reino Vegetal

O Reino Vegetal também conhecido como *Plantae* reúne as plantas que são seres eucariontes, multicelulares e autotróficos fotossintetizantes. São representantes desse grupo os musgos, as samambaias, os pinheiros e as árvores frutíferas (AMABIS; RODRIGUES, 2006).

Atualmente, são conhecidas aproximadamente 404 mil espécies de plantas (IUCN, 2019) que variam quanto ao tamanho, forma e formação corporal. Devido a essa variação estrutural, as plantas podem ser divididas em grupos, conforme descrito a seguir.

As briófitas são plantas de pequeno porte que vivem em ambientes úmidos e sombreados. São avasculares, ou seja, não possuem sistema de transporte de substância ao longo do corpo vegetal. O exemplo mais comum dessa classe são os musgos (AMABIS; RODRIGUES, 2006; RAVEN; EVERT; EICHHORN, 2014).

Com exceção das briófitas, todos os demais grupos apresentam sistema vascular e, por esse motivo, podem ser classificadas como plantas vasculares ou traqueófitas (AMABIS; RODRIGUES, 2006). O primeiro grupo que apresentou um sistema de transporte de água e substâncias orgânicas e inorgânicas foi o das pteridófitas. Elas são representadas pelas samambaias e avencas e podem atingir mais de 4m de altura (AMABIS; RODRIGUES, 2006; RAVEN; EVERT; EICHHORN, 2014).

Os outros dois grupos de plantas vasculares são fanerógamas, ou seja, apresentam sementes. Dentro desse grupo temos as gimnospermas que são representadas pelos pinheiros e as coníferas. O outro grupo pertencente às fanerógamas são as angiospermas que se diferenciam as gimnospermas por apresentarem flores e frutos (AMABIS; RODRIGUES, 2006; RAVEN; EVERT; EICHHORN, 2014).

As angiospermas podem ser subdivididas em monocotiledôneas e dicotiledôneas dependendo do número de cotilédone no embrião, da direção das nervuras foliares, do tipo de raiz, da estrutura das flores e do sistema vascular (AMABIS; RODRIGUES, 2006; RAVEN; EVERT; EICHHORN, 2014).

Ao observarmos o ambiente ao nosso redor, nos deparamos com dezenas de indivíduos representantes do Reino Vegetal. Esse reino tem importância alimentar, energética, ornamental e ecológica. E, foi por meio dele que o ecossistema se organizou da forma como vemos hoje.

Das quase 404 mil espécies de plantas descritas, aproximadamente 382 mil possuem sistema vascular (IUCN, 2019). Isso indica a importância evolutiva desse sistema de transporte e o quanto isso impulsionou a colonização do ambiente terrestre e, desse modo,

contribuiu para o espalhamento do reino vegetal por toda a Terra.

2.1 Histologia vegetal

Se na antiguidade uma prática que facilitava a cognição e o estudo era dividir o organismo em suas partes constituintes, com o surgimento da histologia essa abordagem ganhou ainda mais importância e auxiliou na classificação dos seres vivos e dos seus tecidos. Tecidos são, por definição, conjuntos de células, iguais ou diferentes entre si, que desempenham função específica (JUNQUEIRA; JUNQUEIRA, 1983). Sendo assim, a busca por atributos histológicos passou a auxiliar na elucidação dos problemas taxonômicos e a promover uma classificação mais objetiva desses organismos (STACE, 1991).

No reino vegetal, os tecidos foram divididos em dois grupos (vide Figura 2): tecidos meristemáticos que são constituídos por células em constantes divisões, e os tecidos permanentes formados por células diferenciadas com formato, função e tamanho permanentes (TAIZ; AL., 2017).

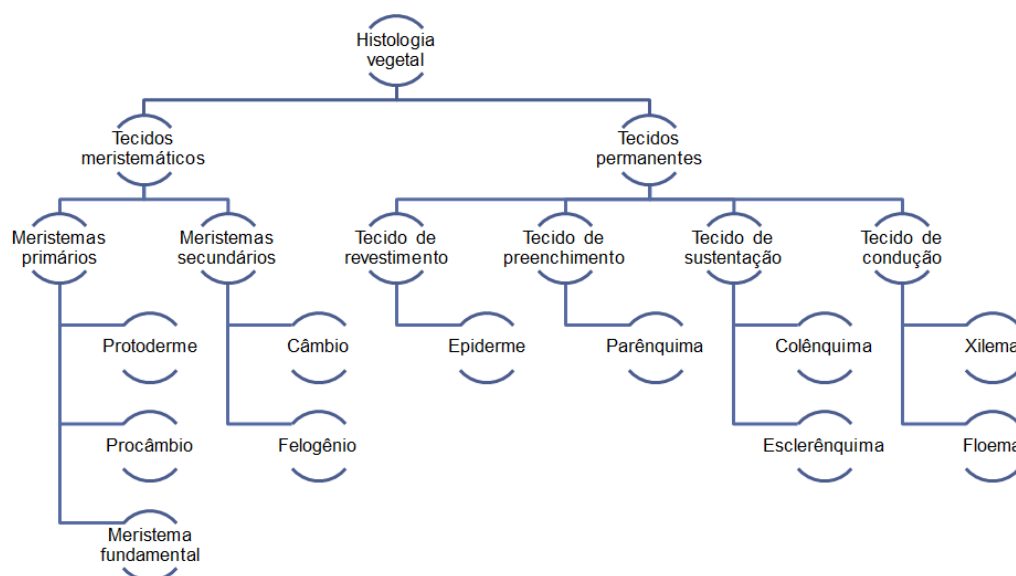


Figura 2 – Divisão dos tecidos vegetais em dois grupos: Tecidos meristemáticos e Tecidos permanentes

Fonte: Autor

Para entendermos melhor a organização tecidual do corpo vegetal, precisamos observar e compreender as fases do seu desenvolvimento. Segundo Glória e Guerreiro (2006) o desenvolvimento morfológico das plantas inicia-se com a germinação da semente, estrutura que contém o embrião. O embrião apresenta dois tipos de meristema apical, um localizado no caule (ou caulinar) e o localizado na raiz (ou radicular). Esses tecidos, após sucessivas divisões celulares, são responsáveis pela formação, respectivamente, do caule primário e da raiz primária. Esses tecidos meristemáticos primários são formados pela

protoderme, pelo meristema fundamental e pelo procâmbio. Eles irão formar os tecidos primários do vegetal como a epiderme, córtex, periciclo, xilema e floema. Ou seja, os meristemas apicais produzem as células que se diferenciam em novas partes tanto do caule quanto da raiz e proporcionam o crescimento em extensão (longitudinal), também conhecido como crescimento primário do vegetal.

Para muitas plantas, como, por exemplo, a maioria das espécies de monocotiledôneas e poucas de dicotiledôneas herbáceas, apenas o crescimento primário é necessário para o desenvolvimento do corpo tecidual do vegetal adulto (GLÓRIA; GUERREIRO, 2006; MODESTO; SIQUEIRA, 2002). Em contrapartida, na maioria das dicotiledôneas e nas gimnospermas é necessário o crescimento em espessura ou secundário. Esse crescimento é oriundo da atividade do câmbio e do felogênio. Esses dois tecidos, originados do procâmbio, periciclo e da epiderme, são denominados meristemáticos secundários ou laterais e seu desenvolvimento dá origem aos tecidos secundários (súber, feloderme, floema e xilema secundários) (GLÓRIA; GUERREIRO, 2006).

De maneira geral, podemos resumir o corpo celular em três sistemas de tecidos: Dérmicos que são responsáveis pelo revestimento; Fundamentais, responsáveis pelo preenchimento e pela sustentação; e Vasculares, responsáveis pelo transporte de nutriente ao longo do corpo vegetal. (AMABIS; RODRIGUES, 2006; GLÓRIA; GUERREIRO, 2006; RAVEN; EVERT; EICHHORN, 2014). A distribuição desses sistemas ao longo do corpo vegetal pode variar de acordo com o órgão e o grupo vegetal, mas basicamente os padrões seguem a orientação: o sistema vascular é envolto pelo sistema fundamental e o sistema dérmico reveste a planta (GLÓRIA; GUERREIRO, 2006).

Desde modo, então, utilizar os tecidos para entender a fisiologia e as características dos vegetais passou a ser uma abordagem muito comum. Saber que aquele organismo era integrante do Reino *Plantae* já não era o suficiente. A qual subgrupo aquele indivíduo pertencia? Quais suas principais características histológicas? E foi nesse cenário que a classificação dos vegetais surgiu.

A Figura 3 ilustra o cladograma que evidencia a importância da histologia para a divisão dos grupos vegetais. O primeiro grande marco evolucionário das plantas foi o surgimento dos tecidos vasculares e, com o advento da microscopia e da histologia, o xilema e o floema puderam ser caracterizados e estudados tornando-se assim, tecidos com valor taxonômico (CHEADLE, 1955). Sendo assim, as plantas puderam ser separadas em vasculares ou avasculares, ou seja, puderam ser classificadas em relação à presença ou ausência desses tecidos.

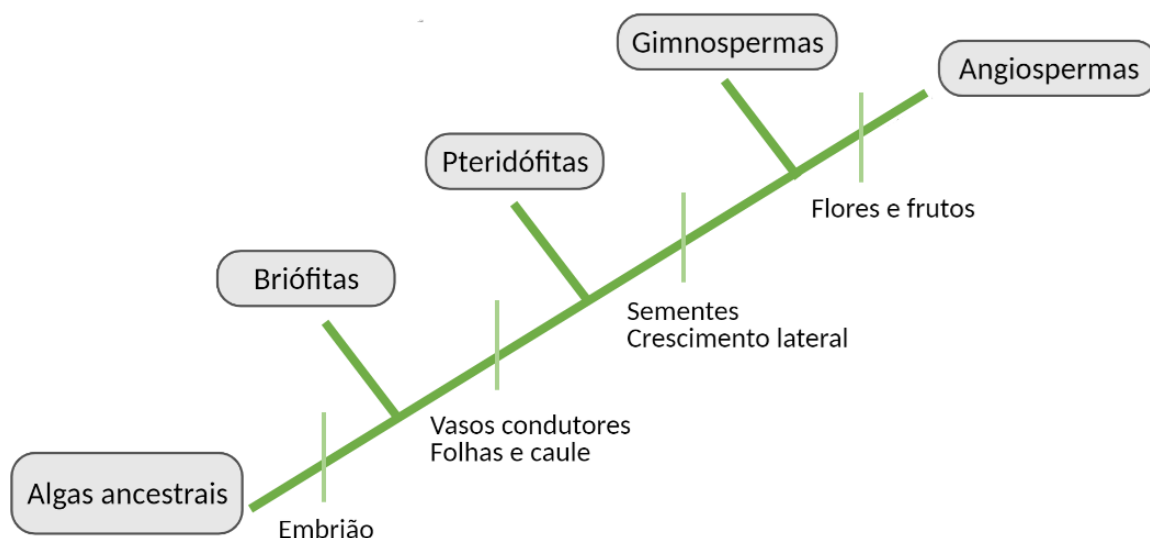


Figura 3 – Cladograma ilustrando os marcos evolutivos vegetais

Fonte: Autor.

Os outros dois marcos evolutivos estão relacionados com a presença de sementes e a presença de flores e frutos. Apesar de todos os órgãos e tecidos serem importantes e possuírem funções específicas que contribuem para o desenvolvimento e manutenção do vegetal, o presente trabalho irá focar apenas nos tecidos vasculares.

2.2 Tecido vascular vegetal

Os tecidos vasculares, também conhecidos como tecidos condutores, têm como principal função transportar água, nutrientes, sais minerais e metabólitos ao longo do corpo vegetal. Essas substâncias são conduzidas em duas direções: das raízes para as folhas e das folhas para as raízes. (AMABIS; RODRIGUES, 2006; FERRI, 1983; MODESTO; SIQUEIRA, 2002; RAVEN; EVERT; EICHHORN, 2014). Os tecidos vasculares, são contínuos ao longo de todos os órgãos das plantas vasculares e formam um sistema vascular capaz de suprir toda demanda hidráulica e energética desses organismos (GLÓRIA; GUERREIRO, 2006).

Como discutido brevemente nas seções anteriores, os tecidos vasculares apresentam grande importância evolutiva para os vegetais. Anteriormente ao surgimento evolutivo desse sistema condutor, o transporte de substâncias era feito por difusão, ou seja, célula a célula por meio da parede celular e isso resultava em duas limitações para a planta: a) necessidade de contato com o meio aquático; b) pequeno porte do organismo (MODESTO; SIQUEIRA, 2002; RAVEN; EVERT; EICHHORN, 2014). Foi por meio do surgimento de um sistema organizado de condução de água, sais minerais e nutrientes que os vegetais puderam alcançar definitivamente o meio terrestre e houve a possibilidade de crescimento

sem prejuízo hídrico e energético para a planta.

Os tecidos vasculares são tecidos permanentes, ou seja, formados por células diferenciadas com formato, função e tamanho permanentes (TAIZ; AL., 2017) e podem ser classificados em relação ao tipo de tecido (xilema ou floema) ou sua origem (câmbio ou procâmbio). O xilema é responsável pelo transporte de água e nutrientes inorgânicos (como sais minerais) das raízes até as folhas da planta. Em contrapartida, o floema é um tecido encarregado de transportar substâncias orgânicas, como, por exemplo, os metabólitos formados pela fotossíntese nas folhas até as demais partes do vegetal pelo floema (AMABIS; RODRIGUES, 2006). A Figura 4 ilustra a direção da condução do xilema (vide Figura 4a) e do floema (vide Figura 4b).

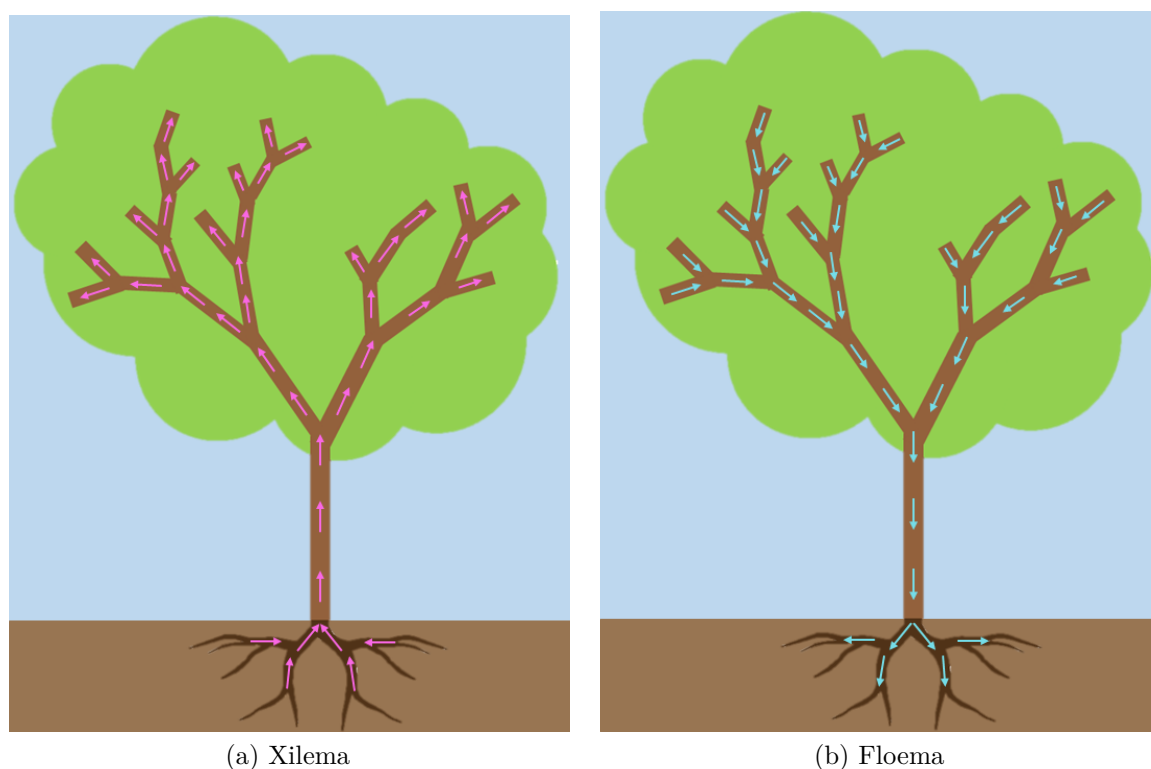


Figura 4 – Representação da direção da condução dos tecidos vasculares vegetais ao longo do organismo. (a) Ilustra o xilema tecido responsável pelo transporte de água e nutrientes inorgânicos das raízes até as folhas; (b) Ilustra o floema tecido responsável pelo transporte de água e substâncias orgânicas das folhas até as raízes.

Fonte: Autor.

O xilema e o floema primário derivam do procâmbio, que está localizado na parte inferior do meristema apical e, por esse motivo, à medida que a planta cresce longitudinalmente, novos elementos de xilema e floema primários são formados. Por outro lado, quando o vegetal cresce em espessura, há a formação do tecido vascular secundário a partir do meristema secundário mais especificamente do câmbio (FERRI, 1983).

Normalmente, as monocotiledôneas e pteridófitas apresentam somente tecidos condutores primários. Já as dicotiledôneas e as gimnospermas formam os tecidos secundários graças à atividade cambial (MODESTO; SIQUEIRA, 2002).

Outra classificação desses tecidos é com relação à disposição do xilema em relação ao floema e vice e versa. Eles podem se organizar de três tipos diferentes de feixes: colaterais (aberto e fechado), bicolateral e concêntrico (anfivasal e anficrival) (MENEZES, 2016; METCALFE; CHALK, 1950; MODESTO; SIQUEIRA, 2002; OGURA, 1972).

2.2.1 Xilema

O xilema é o tecido vegetal responsável pela condução de água e sais minerais desde as raízes até as folhas das plantas (GLÓRIA; GUERREIRO, 2006). Essa solução de água e soluto é comumente chamada de seiva bruta. Por essa razão, alguns livros falam que a principal função do xilema é o transporte de seiva bruta ao longo do corpo vegetal (AMABIS; RODRIGUES, 2006; MENEZES, 2016).

Esse tecido vegetal também apresenta mais duas funções: armazenamento de nutriente e sustentação (suporte mecânico) (GLÓRIA; GUERREIRO, 2006; MODESTO; SIQUEIRA, 2002; RAVEN; EVERT; EICHHORN, 2014).

O xilema, também conhecido como lenho (MODESTO; SIQUEIRA, 2002), é constituído principalmente por três tipos celulares: elementos traqueais, células parenquimáticas e as fibras (GLÓRIA; GUERREIRO, 2006; MODESTO; SIQUEIRA, 2002). Essa formação vale tanto para o xilema primário quanto para o secundário. A diferença deles, além da origem embrionária está no sistema de organização, pois no primário as células organizam-se apenas no sistema radial enquanto que no secundário elas organizam-se tanto no radial quanto no axial (GLÓRIA; GUERREIRO, 2006).

Os elementos traqueais são assim chamados porque são estruturas reforçadas por anéis de lignina e se assemelham a traqueia dos insetos (MODESTO; SIQUEIRA, 2002). Existem dois tipos básicos de elementos traqueais: as traqueídes e os elementos de vaso (AMABIS; RODRIGUES, 2006; GLÓRIA; GUERREIRO, 2006; MODESTO; SIQUEIRA, 2002; RAVEN; EVERT; EICHHORN, 2014). Ambas são células alongadas, com paredes secundárias lignificadas, portanto, mortas e sem protoplastos. A principal função desses elementos traqueais é a condução de água ao longo do vegetal (GLÓRIA; GUERREIRO, 2006; MODESTO; SIQUEIRA, 2002).

Em termos evolutivos, as traqueídes representam os elementos de condução de água mais primitivo. Essas células são imperfuradas e, portanto, a comunicação entre as células é feita por pares de pontuações, que são pequenos orifícios, em suas paredes.

Em contrapartida, os elementos de vaso são característicos das angiospermas e das ordens mais complexas das gimnospermas (GLÓRIA; GUERREIRO, 2006; MODESTO;

SIQUEIRA, 2002). Essas células são perfuradas e essas perfurações podem ser simples ou múltiplas. A comunicação entre as células é feita não só pelos pares de pontuações, mas também pelas placas de perfurações. Como os elementos de vasos unem-se uns aos outros formando longas colunas contínuas por onde a água consegue passar livremente através das perfurações (MODESTO; SIQUEIRA, 2002).

O outro tipo de célula que constitui o xilema são as células parenquimáticas. Elas são células vivas que envolvem os vasos. As suas principais funções são: armazenamento de substâncias de reservas como amido, óleos e cristais; translocação de água e soluto á curta distância (GLÓRIA; GUERREIRO, 2006; MODESTO; SIQUEIRA, 2002). Elas estão mais presentes em angiospermas e raramente são encontradas em gimnospermas (GLÓRIA; GUERREIRO, 2006). Morfologicamente, elas são alongadas e possuem paredes mais delgadas quando comparadas com as paredes dos elementos de vaso e das fibras (GLÓRIA; GUERREIRO, 2006).

O último tipo de célula que compõe o xilema são as fibras. Elas são responsáveis pela sustentação do vegetal. São alongadas e possuem extremidades afiladas. Suas paredes variam em espessura, mas normalmente, são mais espessas que os demais tipos celulares do xilema (GLÓRIA; GUERREIRO, 2006; MODESTO; SIQUEIRA, 2002).

2.2.2 Floema

O tecido vegetal vascular responsável pelo transporte de materiais orgânicos e inorgânicos em solução nas plantas é conhecido por floema ou líber (AMABIS; RODRIGUES, 2006). Entre os materiais orgânicos transportados pelo floema destacam-se os carboidratos na forma de sacarose, as substâncias nitrogenadas como aminoácidos e amidas, os lipídios, os ácidos orgânicos, os ácidos nucleicos, substâncias reguladoras de crescimentos e as vitaminas. As substâncias inorgânicas resumem-se aos íons inorgânicos como os sais minerais (AMABIS; RODRIGUES, 2006; GLÓRIA; GUERREIRO, 2006; RAVEN; EVERT; EICHHORN, 2014). Dessa maneira, podemos inferir com o floema é o tecido responsável pelo transporte do alimento, ou seiva elaborada, desde o sítio de produção (folhas) até um sítio consumidor (meristemas) (GLÓRIA; GUERREIRO, 2006; RAVEN; EVERT; EICHHORN, 2014).

Análogo ao xilema, esse tecido está presente em todos os órgãos da planta vascular e, além de ser responsável pela condução de substâncias, também possui as funções de armazenamento de nutriente e sustentação (GLÓRIA; GUERREIRO, 2006; MODESTO; SIQUEIRA, 2002).

O floema é composto por quatro tipos celulares: os elementos crivados; as células parenquimáticas; as fibras; e as esclereídes (GLÓRIA; GUERREIRO, 2006; MODESTO; SIQUEIRA, 2002; RAVEN; EVERT; EICHHORN, 2014). Os elementos crivados são as

células especializadas em condução e se apresentam de duas maneiras: as células crivadas e os elementos de tubo crivados (GLÓRIA; GUERREIRO, 2006; MODESTO; SIQUEIRA, 2002).

As células crivadas são células longas e mais primitivas que estão presentes em pteridófitas e gimnospermas (GLÓRIA; GUERREIRO, 2006). São caracterizadas por possuírem grupo de poros estreitos na parede celular denominados áreas crivadas (GLÓRIA; GUERREIRO, 2006; MODESTO; SIQUEIRA, 2002). Em contrapartida, os elementos dos tubos crivados são células mais curtas que se caracterizam por apresentarem placas crivadas nas paredes laterais (GLÓRIA; GUERREIRO, 2006). Placas crivadas são áreas crivadas que apresentam poros maiores e altamente especializadas (MODESTO; SIQUEIRA, 2002). Uma placa pode ser composta por uma ou mais áreas crivadas e os poros podem variar de 1 a aproximadamente $15\mu\text{m}$ (GLÓRIA; GUERREIRO, 2006). Essas estruturas são exclusivas das angiospermas (GLÓRIA; GUERREIRO, 2006; MODESTO; SIQUEIRA, 2002).

Vale ressaltar que tanto os elementos do tubo crivado e as células crivadas apresentam protoplasto, com exceção do núcleo, e não são lignificadas, portanto, são consideradas células vivas. E a condução de nutrientes pelos elementos crivados é feita por meio dos orifícios da placa crivada. Cada poro é atravessado por um plasmodesmo, uma fina ponte citoplasmática, que comunica uma célula a outra (AMABIS; RODRIGUES, 2006; MODESTO; SIQUEIRA, 2002).

Os elementos dos tubos crivados são associados às células parenquimáticas (MODESTO; SIQUEIRA, 2002). Estas se diferenciam uma das outras e se especializam em uma das categorias a seguir: células companheiras; células albuminosas e células intermediárias (GLÓRIA; GUERREIRO, 2006). Também existem células parenquimáticas que não se especializam, conhecidas como parenquimáticas não-especializadas e estas são responsáveis por armazenar substâncias como amidos e cristais (GLÓRIA; GUERREIRO, 2006; MENEZES, 2016; MODESTO; SIQUEIRA, 2002).

As células companheiras são as mais relacionadas com os elementos do tubo crivado e se associam a eles por numerosas conexões citoplasmáticas (GLÓRIA; GUERREIRO, 2006; MODESTO; SIQUEIRA, 2002). Elas têm importante papel na distribuição dos assimilados dos elementos de tubo crivado e fazem o controle dos elementos do tubo crivado por meio da transferência de moléculas informacionais e moléculas como ATP (GLÓRIA; GUERREIRO, 2006).

As células albuminosas, ou células de Strasburger, são análogas às células companheiras, mas estão associadas às células crivadas e não aos elementos do tubo crivado como as células companheiras (GLÓRIA; GUERREIRO, 2006; MENEZES, 2016).

A outra categoria de células parenquimáticas associadas ao floema são as células intermediárias. Elas estão presentes nas nervuras de menor calibre de folhas adultas e são

responsáveis por mediar o acúmulo e carregamento de solutos orgânicos, principalmente carboidratos (GLÓRIA; GUERREIRO, 2006; MENEZES, 2016).

Os outros tipos celulares encontrados comumente no floema são as fibras e as esclereídes. As fibras possuem além da função de sustentação (MODESTO; SIQUEIRA, 2002), a de reserva de substâncias assim como as parenquimáticas (GLÓRIA; GUERREIRO, 2006). As esclereídes são encontradas nas partes mais velhas do floema, são ramificadas e resultantes da esclerificação (lignificação) de células do parênquima. Assim como as fibras, atuam na sustentação do vegetal com as esclereídes são lignificadas e, portanto, mortas (GLÓRIA; GUERREIRO, 2006).

2.3 Feixes vasculares vegetais

Feixes vasculares vegetais são as associações do xilema e do floema nos órgãos vegetais, ou seja, é a disposição do xilema em relação ao floema e vice e versa em um órgão específico (MENEZES, 2016; MODESTO; SIQUEIRA, 2002).

No caule, esses tecidos podem se organizar de três tipos diferentes de feixes: colaterais (aberto e fechado), biclateral e concêntrico (anfivasal e anficrival) (MENEZES, 2016; METCALFE; CHALK, 1950; MODESTO; SIQUEIRA, 2002; OGURA, 1972). Os tipos de feixes estão ilustrados, respectivamente, nas Figuras 5, 6, 7.

Os colaterais são caracterizados por apresentarem cordões de xilema e floema situados lado a lado, conforme ilustrado na Figura 5. Normalmente, o floema fica localizado do lado de externo ao xilema tendo o centro do cilindro como referência (MODESTO; SIQUEIRA, 2002). Esse tipo de feixe é subdividido em dois grupos: colateral aberto (vide Figura 5a) e colateral fechado (vide 5b). A diferença entre eles é que no aberto existe a presença de células vivas entre os tecidos vasculares como, por exemplo, o procâmbio ou células potencialmente meristemáticas, ou seja, o feixe pode produzir mais elementos vasculares e crescer. Em contrapartida, no colateral fechado não há praticamente células vivas entre o xilema e o floema, devido a isso, ele não poderá mais produzir novos elementos vasculares (MENEZES, 2016).

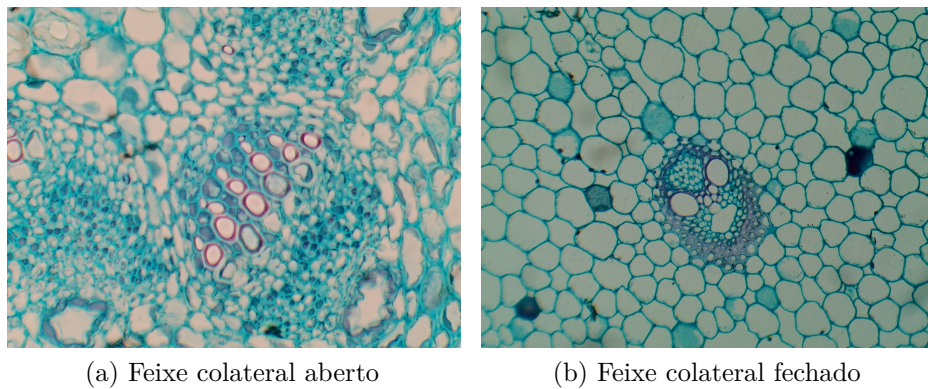


Figura 5 – Feixes vasculares vegetais do tipo colateral. (a) Representa um feixe colateral aberto no qual existe uma camada de células vivas (procâmbio) entre o floema e o xilema; (b) Representa um feixe colateral fechado no qual não há células intermediárias entre os tecidos vasculares. Células do floema estão coradas em azul e do xilema em rosa.

Fonte: Autor

O feixe é considerado biclateral quando o xilema é envolvido externa e internamente pelo floema e é possível diferenciar o floema externo e o interno (MENEZES, 2016; MODESTO; SIQUEIRA, 2002). Um esquema representativo da organização dos tecidos vasculares no feixe biclateral está representado na Figura 6.

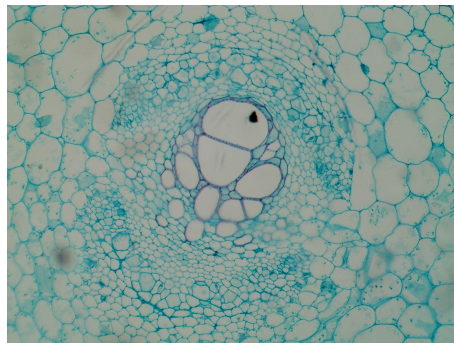


Figura 6 – Feixes vasculares vegetais do tipo biclateral o apresenta um floema externo e um interno mediado por um cordão de xilema. Células do floema estão coradas em azul e do xilema em rosa.

Fonte: Autor

Já os feixes concêntricos são feixes que um tecido vascular circunda o outro (MODESTO; SIQUEIRA, 2002), conforme representado na Figura 7. Quando o floema circunda o xilema o feixe é denominado anficrival (vide 7a). Em contrapartida, quando o xilema circunda o floema, ele é denominado anfivasal (vide Figura 7b) (MENEZES, 2016; MODESTO; SIQUEIRA, 2002).

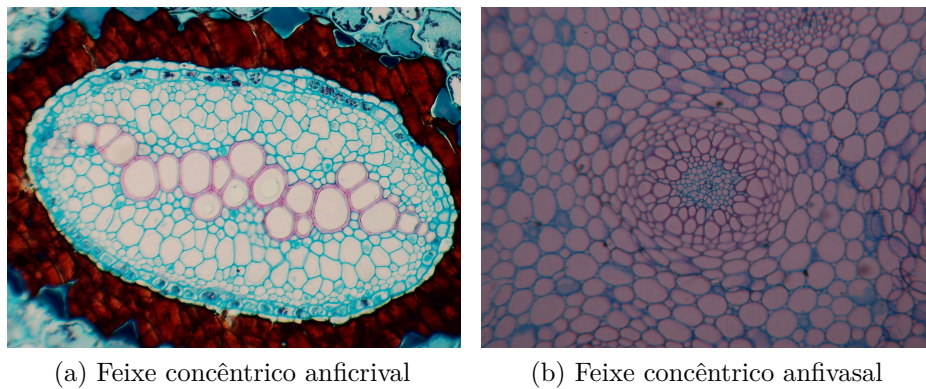


Figura 7 – Feixes vasculares vegetais do tipo concêntrico. (a) Representa um feixe concêntrico anicrival no qual o xilema é o vaso central e o floema o circunda; (b) Representa um feixe concêntrico anivasal no qual o floema é o vaso central e o xilema o circunda. Células do floema estão coradas em azul e do xilema em rosa.

Fonte: Autor

Esses feixes estão distribuídos entre as 382 mil espécies vegetais vasculares descritas (IUCN, 2019). Por exemplo, as monocotiledôneas podem ter feixes do tipo colateral ou concêntrico (CATTAI; MENEZES, 2010; JURA-MORAWIEC, 2017; TOMLINSON; ZIMMERMANN, 1969). Dentro das monocotiledôneas, temos a espécie *Dracaena marginata*, por exemplo, que possui exclusivamente feixe do tipo concêntrico anivasal (JURA-MORAWIEC, 2015) e a espécie *Cyperus rotundus*, uma monocotiledônea herbácea, que apresenta o feixe do tipo colateral fechado (GLÓRIA; GUERREIRO, 2006; MENEZES, 2016).

As dicotiledôneas podem ter feixes do tipo colateral ou bicolateral (GLÓRIA; GUERREIRO, 2006; MENEZES, 2016; MODESTO; SIQUEIRA, 2002). Dentro das dicotiledôneas, temos as espécies *Baccharis dracunculifolia* e a *Pimpinella anisum* que apresentam feixes vasculares do tipo colateral aberto (GLÓRIA; GUERREIRO, 2006; MENEZES, 2016). Em contrapartida, as espécies *Sechium edule* e a *Cucurbita moschata* são dicotiledôneas que apresentam feixes vasculares do tipo bicolateral (GLÓRIA; GUERREIRO, 2006; MENEZES, 2016; MODESTO; SIQUEIRA, 2002).

Já as pteridófitas podem ter feixes do tipo concêntrico (MODESTO; SIQUEIRA, 2002) como, por exemplo, a *Selaginella spp* que apresenta o feixe concêntrico do tipo anicrival (MENEZES, 2016).

É notório que os feixes vasculares apresentem particularidades e estejam distribuídos nas mais variadas espécies vegetais. Portanto, estudos que contribuam para descrever e caracterizar esses tecidos podem promover subsídios para estudos que busquem aprimorar a classificação vegetal.

3 Análise de imagens e Reconhecimento de padrões

3.1 Biomimética e a visão artificial

Em 1941, o engenheiro suíço George Mestral enquanto caminhava pelo campo notou que as sementes do *Arctium* colavam constantemente em sua roupa. Esse fato deixou o cientista curioso e, então, ele resolveu observá-las no microscópio. Ao fazer isso, ele notou a presença de filamentos entrelaçados que apresentavam pequenos ganchos em suas pontas e, por esse motivo, se aderiam bem aos tecidos. Essa característica encontrada e observada na Natureza foi o que o inspirou para uma de suas grandes invenções: o velcro (NUNES, 2014).

A história da invenção do velcro ilustra a essência da Biomimética a qual visa tentar aprender com a Natureza e reproduzir suas estratégias para fins específicos. Essa prática pode ser aplicada aos mais diversos nichos da ciência e, atualmente, o design, a arquitetura, a química, as engenharias e a computação são algumas das áreas que mais aplicam os conhecimentos aprendidos e observados na Natureza.

Em computação, a terminologia "Computação Natural" descreve todos os sistemas computacionais desenvolvidos com inspiração ou utilização de algum mecanismo natural ou biológico no processamento de informação (CASTRO, 2006). Um exemplo disso é a Visão Computacional ou *Machine Vision* a qual se baseia nos processos biológicos da formação da imagem para tentar mimetizar a visão natural na denominada visão artificial (MARR, 1982).

A visão artificial está difundida em diversas áreas tais quais: medicina, agricultura, industrial e robótica. Ela vem aprimorando a classificação de produtos, exames clínicos e frutos pois, a inspeção visual desses elementos era tradicionalmente realizada por operadores treinados, mas como as decisões tomadas pelos operadores são subjetivas, ou seja, são afetadas por diversos fatores como por exemplo a fadiga ou hábitos adquiridos, existe um alto risco de ocorrer erro humano no processo classificatório (GONZALEZ; WOODS, 2000). Por isso, atualmente, a classificação é estimada com sistemas de visão artificial por meio da análise e do reconhecimento de padrão (DU; SUN, 2006).

Um estudo que exemplifica bem esse aprimoramento do sistema de classificação foi realizado por Paulus, Busscher e Schrevens (1997) com diferentes variedades de maçãs. Os operadores treinados avaliaram vários formatos, tamanhos e parâmetros de cores e, o resultado demonstrou uma limitada capacidade humana de reproduzir a estimativa

da qualidade se comparado a um sistema de classificação artificial. Outra conclusão interessante deste estudo é que o aumento do número de parâmetros considerados em uma tomada de decisão está diretamente relacionada com o aumento do erro de classificação.

Como pode ser observado no exemplo acima, a extração de atributos realizada no experimento teve base nas observações já realizadas pelos operadores. É natural que isso ocorra pois a visão artificial está intrinsecamente relacionada com a visão natural. A ampla distribuição dessa ferramenta no reconhecimento de padrão é refletida na rapidez e na capacidade de quantificação que os algoritmos de reconhecimento possuem.

Vale ressaltar que a visão artificial envolve as mesmas etapas da visão natural: aquisição, processamento e interpretação da imagem. Ou seja, para gerar o produto final é necessário um dispositivo que irá capturar a luz refletida do objeto, fotosensores que irão transformar o estímulo luminoso em elétrico, um centro de processamento e, por último, software de interpretação e análise da imagem formada (GONZALEZ; WOODS, 2000). Essas etapas fizeram parte da metodologia deste trabalho e, por isso, serão melhores descritas nas próximas seções.

3.2 Aquisição de imagens

Existem vários processos que resultam na aquisição das imagens, porém, todos eles partem do mesmo princípio: Um objeto é iluminado por uma fonte luminosa que pode ser a luz natural, luz artificial raio-X, radar, infravermelho, entre outras e, dependendo da natureza da fonte, a energia da iluminação é refletida dos objetos ou transmitida através deles (GONZALEZ; WOODS, 2000). Por fim, essa luz é captada por um dispositivo que possui sensores que transformam a energia da iluminação em imagens digitais.

Neste trabalho, focaremos no processo de aquisição de imagem obtida a partir de um microscópio óptico acoplado a uma câmara fotográfica digital, uma vez que, esses foram os instrumentos utilizados na metodologia deste estudo.

3.2.1 O microscópio óptico

A invenção do microscópio revolucionou a forma como os organismos eram vistos e estudados pelos cientistas e, por isso, é um grande marco na história da ciência. Com o passar do anos, a microscopia evoluiu e, os microscópios rudimentares como aquele inventado por Antonie van Leeuwenhoek em 1674 (ANTONIE. . . , 2009) deram lugar a microscópios mais complexos e tecnológicos.

Dentro dessa caminhada evolutiva, encontra-se o microscópio óptico também conhecido como microscópio de luz. Seu funcionamento é bem simples se comparado aos microscópios mais atuais. Essa simplicidade mecânica o torna fácil de manusear e, com

isso, sua aplicabilidade ganha grandes proporções.

A Figura 8 representa um microscópio óptico e suas principais partes constituintes. A lâmina montada da amostra de interesse é colocada sobre a platina que é uma base de suporte e fixação que tem uma abertura central que deixa passar a luz. Essa luz é proveniente de uma lâmpada e, por meio de um conjunto de espelhos, é direcionada para o condensador cuja função é concentrar a luz e projeta-la através do espécime (transiluminação) (JUNQUEIRA; CARNEIRO, 2013). A intensidade da luz pode ser regulada por meio do reóstato que atua na própria fonte luminosa ou indiretamente através do condensador e do diafragma que é uma abertura que regula a quantidade de luz pelo tamanho do seu diâmetro (MOREIRA, 2013). A luz que atravessa o espécime é recebida pela objetiva que é a lente que fica mais próxima do objeto e projeta uma imagem real, ampliada e invertida do mesmo em direção à ocular que é um cilindro com duas ou mais lentes que permitem ampliar a imagem real fornecida pela objetiva, formando uma imagem virtual que será projetada na retina do observador (JUNQUEIRA; CARNEIRO, 2013; MOREIRA, 2013).

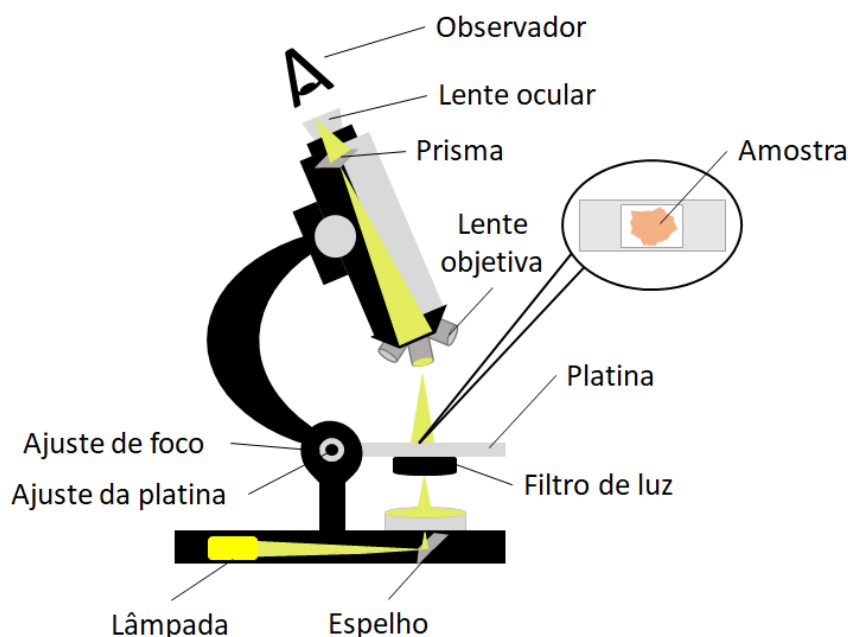


Figura 8 – Microscópio óptico evidenciando suas partes constituintes e o caminho percorrido pela luz.

Fonte: Autor.

O avanço da tecnologia possibilitou o uso de câmeras digitais acopladas aos microscópios ópticos. Dessa maneira, a imagem fornecida pelas lentes oculares do microscópio não precisa ser apenas projetada na retina do observador. Isso tornou possível a digitali-

zação das imagens e sua quantificação por meio de abordagens de análise de imagem e reconhecimento de padrões (JUNQUEIRA; CARNEIRO, 2013)

3.2.2 Digitalização de imagens microscópicas

A Figura 9 ilustra o microscópio óptico Axiostar Plus da Carl Zeiss o qual está acoplado a um dispositivo para digitalização das imagens microscópicas. Esses instrumentos foram utilizados ao longo deste trabalho com o objetivo de obter imagens microscópicas digitais dos feixes vasculares vegetais em estudo.

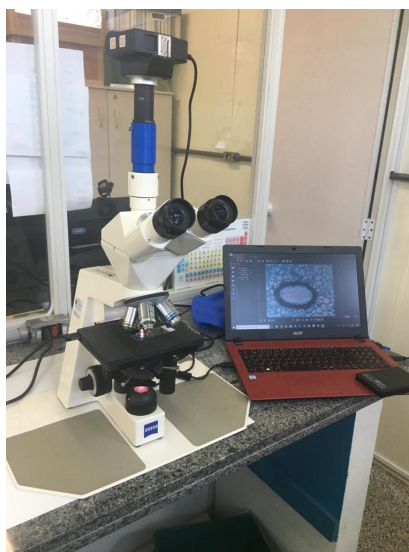


Figura 9 – Sistema de digitalização de imagens acoplado ao microscópio óptico.

Fonte: Autor.

No microscópio óptico Axiostar Plus da Carl Zeiss a luz advinda da lente objetiva pode seguir por dois caminhos: 1) em direção as lentes oculares e projetar a imagem final na retina do observador; 2) seguir em direção a um adaptador, similar a uma lente ocular, inserido acima do microscópio no qual é acoplado um dispositivo para a aquisição da imagem. A escolha de qual caminho a luz irá percorrer é definida pelo operador. Neste trabalho, as imagens foram adquiridas pelo segundo caminho, ou seja, a luz advinda da lente objetiva seguiu para o adaptador e posteriormente para a câmera fotográfica digital onde irá ocorrer a digitalização das imagens.

Dentro das câmeras fotográficas há sensores de luz (fotossensores) individuais dispostos em forma de uma matriz bidimensional de 4000 x 4000 elementos ou mais. Um desses dispositivos é o Dispositivo de carga acoplada (ou do inglês, *Charge-Coupled Device* (CCD)) cuja função é coletar a energia de entrada e projetá-la em um plano de imagem. Os sensores que forem estimulados pela luz advinda inicialmente do espécime irão produzir saídas proporcionais a intensidade de luz recebida. Essas saídas são convertidas

em um sinal analógico que será digitalizado dentro do sistema de aquisição de imagem (GONZALEZ; WOODS, 2000).

Para criar uma imagem digital é necessário então, converter os dados contínuos gerados pelos CCDs para formato digital. Essa transformação envolve dois processos: a amostragem e a quantização (GONZALEZ; WOODS, 2000), sendo que a primeira irá digitalizar os valores de coordenadas e a segunda os valores de amplitude (nível de intensidade).

Tanto a amostragem quanto a quantização ocorrem automaticamente nos dispositivos de aquisição e a imagem digital é mostrada em tempo real para o observador. Sendo assim, é possível armazenar a imagem para posteriormente realizar sua quantificação por meio de abordagens de análise de imagem e reconhecimento de padrões. Sabemos que estas abordagens tem trazido descobertas interessantes e promovendo trabalhos cada vez mais interdisciplinares e inovadores. Sendo assim, se faz necessário entender um pouco mais sobre esses assuntos e este será o foco das próximas seções deste trabalho.

3.3 A imagem digital

Até o presente momento abordamos os mecanismos necessários para a formação de uma imagem digital desde o caminho percorrido pela luz dentro do microscópio óptico até a formação da imagem digital dentro dos dispositivos de aquisição. Porém, o que é uma imagem digital e como podemos quantificá-las por meio das abordagens de análise de imagem e reconhecimento de padrões?

Uma imagem digital pode ser definida com uma função bidimensional $I(x,y)$ onde x e y são coordenadas espaciais, e a amplitude de I é chamado de intensidade (COSTA; CESAR, 2009; GONZALEZ; WOODS, 2000). Como a imagem é digital, os valores de x , y e I são quantidades finitas e discretas. Em linhas gerais, uma imagem digital é representada por duas dimensões de um conjunto de pontos finitos, formando uma matriz, onde cada ponto da matriz é denominado *pixel* (COSTA; CESAR, 2009; GONZALEZ; WOODS, 2000).

Quando pensamos em imagem digital logo vem a nossa mente dois tipos de imagens: as coloridas e as em tons de cinza. A Figura 10 ilustra uma mesma imagem sendo apresentada em cores (vide 10a) e em tons de cinza (vide 10b). A diferença entre elas está no número de matrizes ou bandas que cada imagem tem. Para entendermos melhor essa diferenças, partiremos da premissa de que grande parte das cores visíveis pelo olho humano podem ser obtidas por meio de uma combinação de três cores primárias: vermelho (do inglês *Red*, R), verde (do inglês *Green*, G) e azul (do inglês *Blue*, B). Sendo assim, as imagens coloridas precisam combinar essas três cores para formarem quaisquer outras tonalidades. Dessa maneira, um *pixel* irá possuir 3 bandas espectrais, ou três matrizes

de intensidades: a R, a G e a B e, por isso, são conhecidas com imagens multibandas (GONZALEZ; WOODS, 2000).

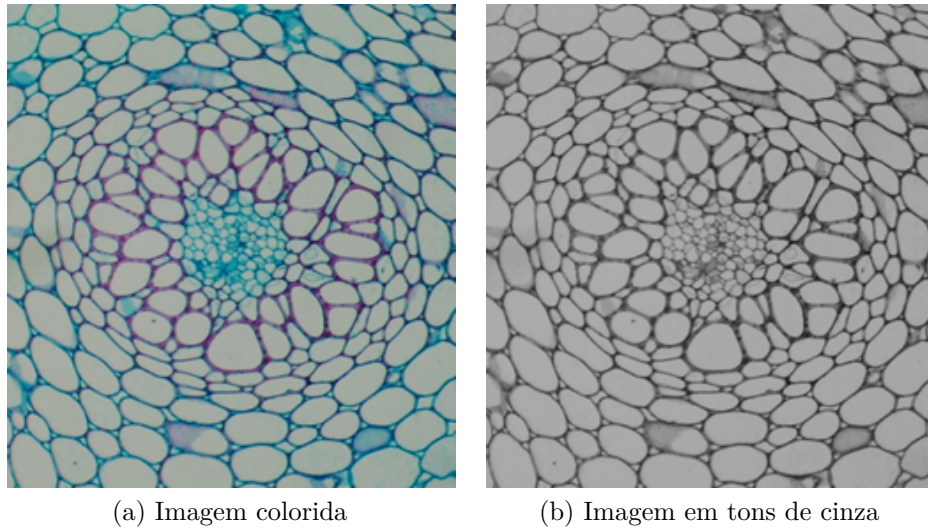


Figura 10 – Imagem colorida e em tons de cinza. (a) Imagem colorida; (b). Imagem em tons de cinza.

Fonte: Autor

Na Figura 11 temos a imagem original e colorida de um feixe vascular vegetal do tipo anfigasal sendo dividida em suas 3 camadas constituintes e seus respectivos histogramas com as intensidades dos *pixels* em cada banda espectral.

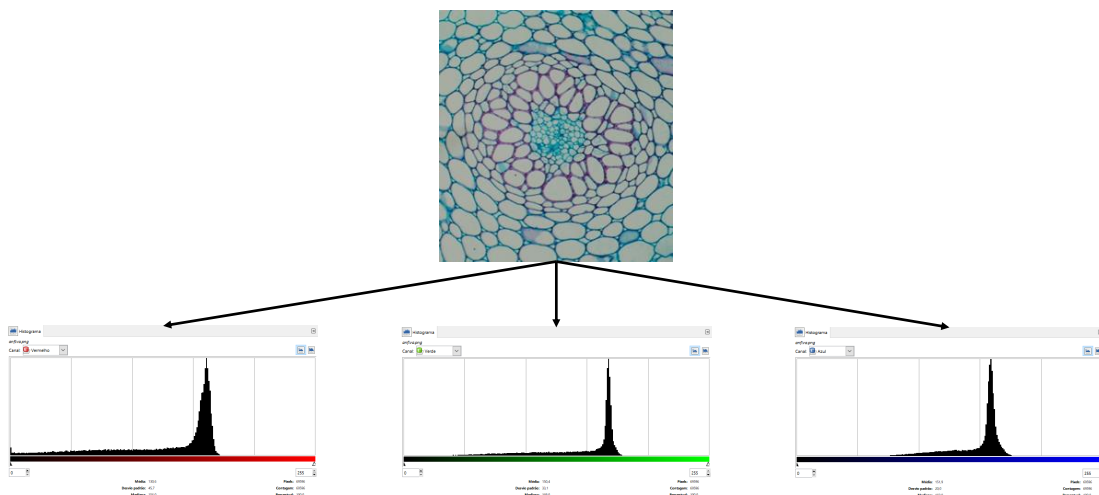


Figura 11 – Imagem colorida com os histogramas das bandas R, G e B.

Fonte: Autor.

Em contrapartida, uma imagem em tons de cinza terá apenas uma banda e, portanto, apenas uma matriz. Nesses casos, o valor de cada *pixel* pode ser chamado de tons de cinza, e é uma única amostra de um espaço de cores, variando entre o preto como a menor intensidade e o branco como maior intensidade, passando pelo cinza, conforme mostrado na Figura 12.

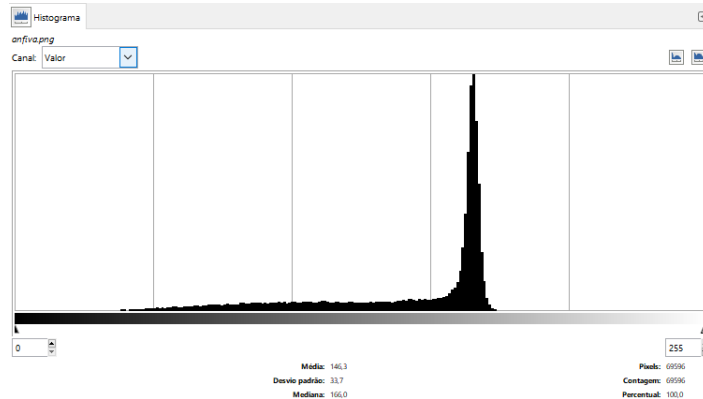


Figura 12 – Histograma da imagem em tons de cinza

Fonte: Autor.

Deste ponto em diante, focaremos apenas nas imagens em tons de cinzas, uma vez que todo trabalho foi feito a partir de imagens digitais com 256 tons de cinza.

Entender que toda e qualquer imagem é uma matriz bidimensional é o primeiro pré-requisito para compreender como são aplicadas as abordagens de análise de imagem e reconhecimento de padrões. Essa característica das imagens é o que torna possível trabalharmos com todas as operações matriciais e, com isso, quantificar, analisar e reconhecer objetos e/ou imagens completas. Normalmente, os valores dessa matriz estão dentro do intervalo de 1 a 255 e são valores inteiros, onde 1 é preto e 255 branco. Os demais níveis são tons discretos ao longo da faixa de intensidade.

A Figura 13 ilustra uma imagem em tons de cinza de um feixe vascular anfvival e sua respectiva matriz I , na qual os valores estiveram dentro do intervalo $1 \leq I(i, j) \leq 255$, onde os índices i e j indicam as coordenadas de um *pixel* na imagem. Nela podemos observar que a célula em destaque tem intensidade $I(i, j) = 67$ enquanto que o fundo apresenta $I(i, j) = 0$. Nesse exemplo apenas uma parte da matriz foi mostrada pois, imprimi-la totalmente traria pouca informação didática.

que abriu o leque de aplicações dessa área. Atualmente, o processamento de imagens é amplamente utilizado na medicina, na geologia, na geografia, nas ciências biológicas entre outras.

Mesmo com a grande dispersão dessa área em toda a ciência Gonzalez e Woods (2000) relatam que não há uma definição bem estabelecida para o termo processamento de imagens. Muitas vezes, autores entendem como processamento a etapa em que a entrada e a saída são imagens. Porém, Gonzalez e Woods (2000) acreditam que essa definição faça com que o processamento fique limitado apenas as operações primitivas o que seria superficial.

Dessa maneira, Gonzalez e Woods (2000) sugerem que a melhor forma de se entender o processamento de imagens é subdividir os processos em níveis: baixo, médio e alto. Sendo que os processos de nível baixo englobam as técnicas onde a entrada e a saída são imagens, como por exemplo, redução de ruídos, realce de contraste e aguçamento de imagens. Os processos médios são constituídos por etapas nas quais as entradas, em geral, são imagens, mas as saídas são atributos extraídos dessas imagens. Os exemplos mais comuns de processamento a nível médio são: segmentação e extração de atributos de objetos. Já o processamento de nível alto consiste em realizar as funções cognitivas da visão. Muitas vezes essa etapa é sobreposta pela etapa de análise das imagens.

Dentro desse contexto, os tópicos seguintes irão explorar temas como segmentação de imagens digitais e a extração de atributos uma vez que essas abordagens foram utilizadas no decorrer desse trabalho.

3.4.1 Segmentação de imagens digitais

A segmentação é a etapa do processamento de imagens que consiste em segregar os elementos de uma imagem. Ela é importante pois promove a seleção das partes ou dos objetos alvos em uma imagem (GONZALEZ; WOODS, 2000). Em geral, a segmentação se faz necessária quando o problema a ser resolvido requer a individualização de uma parte ou de um objeto da imagem, sendo que, o grau de detalhamento em que a divisão da imagem é realizada depende do problema a ser resolvido.

Esse processo pode ser dividido em três categorias que variam de acordo com a participação humana na segmentação: automática, semiautomática e manual. A segmentação automatizada é caracterizada pelo uso de algoritmos que sejam capazes de segregar os objetos alvos em larga escala e de maneira não subjetiva, ou seja, sem que ocorra a intervenção de um usuário (BARBEDO, 2017). Há duas abordagens tipicamente usadas para segmentar automaticamente as imagens, são elas: detecção de bordas e limiarização (GONZALEZ; WOODS, 2000).

De maneira geral, a detecção de bordas é fundamentada nas mudanças bruscas de

intensidade entre os objetos e o fundo da imagem, ou seja, seus diferentes tons de cinza. Enquanto que a limiarização se baseia na divisão de uma imagem em regiões que sejam semelhantes, por exemplo, segmentar objetos dispostos em um fundo que possui nível de cinza dentro de uma faixa pré estabelecida. Ao realizar a limiarização é possível retirar determinados tons de cinza da imagem (aqueles que caracterizam o fundo, por exemplo) e manter apenas tons com as intensidades dos objetos alvos (COSTA; CESAR, 2009).

Em contrapartida, a segmentação semiautomática é caracterizada pela interferência do usuário em alguma etapa do processo sendo que as demais etapas ocorrem de maneira automatizadas (BARBEDO, 2017; BARBEDO, 2016). Nesses casos, a segmentação é um pouco mais demorada e mais subjetiva quando comparada com a automática.

A última categoria de segmentação é a manual. Nela, o usuário interfere em todas as etapas da segmentação e é o controlador de toda a operação (BARBEDO, 2017). Das três categorias a mais subjetiva e com maior demanda de tempo é a segmentação manual. Porém, ela é fundamental para estabelecer um padrão comparativo entre o ideal e o alcançado por meio dos algoritmos. Em biologia, essa comparação é muito comum pois, muitas vezes, a segmentação é a primeira etapa de uma série de outros processos que irão culminar na identificação de: uma doença (tanto animal (CHIU et al., 2010; BERGOUIGNAN et al., 2009) quanto vegetal (BARBEDO, 2017)), de espécies vegetais (CERUTTI et al., 2013), de um tecido em específico (HERREMANS et al., 2015; VU et al., 2019), ou do posicionamento de células vegetais (LIU et al., 2010; MKRTCHYAN et al., 2011).

O processo de segmentação não é algo trivial e está diretamente relacionado com a resolução do problema proposto. Dessa maneira, uma segmentação bem feita é muito importante para o decorrer do trabalho (GONZALEZ; WOODS, 2000; COSTA; CESAR, 2009). Como exposto acima, muitas vezes a segmentação manual pode ser entendida como um *padrão ouro* de comparação. Essa característica se deve ao fato dela ser feita sob os olhares atentos e treinados de profissionais qualificados. Em nosso trabalho, optamos por segmentar manualmente as imagens dos feixes vasculares vegetais com o objetivo de disponibilizarmos esse material como um banco de dados para trabalhos futuros na área. As segmentações manuais foram supervisionadas pelo anatomista vegetal Prof. Dr. Marcos Arduin responsável pelo Laboratório de Anatomia Vegetal do Departamento de Botânica da Universidade Federal de São Carlos, campus São Carlos.

3.4.2 Extração de atributos

Após a etapa de segmentação, inicia-se a etapa de extração de atributos. Segundo a língua portuguesa, a palavra "atributo" significa o que é próprio, característico de algo ou de alguém, uma particularidade. Ou seja, atributo pode ser interpretado como uma característica de um objeto. Em visão computacional, extrair atributos de uma imagem

consiste em extrair as características mais importantes de uma imagem com o objetivo de evidenciar as diferenças e similaridades entre os objetos alvos (OLIVEIRA; DUTRA; RENNÓ, 2005; DUDA; HART; STORK, 2012). Sendo assim, a principal função da extração de atributos é caracterizar os objetos para estes possam ser reconhecidos (DUDA; HART; STORK, 2012).

Dentro desse contexto, deve-se levar em consideração as particularidades de cada imagem e as particularidades do problema a ser solucionado, ou seja, os atributos extraídos devem conter informações pertinentes sobre as classes a serem reconhecidas (OLIVEIRA; DUTRA; RENNÓ, 2005; OLIVEIRA et al., 2009). Por isso, é muito importante estudar e avaliar o objeto em estudo antes de definir quais serão os atributos que serão extraídos.

Neste trabalho, foram avaliadas imagens histológicas de feixes vasculares vegetais. Portanto, a escolha dos atributos extraídos foi feita respeitando as particularidades de cada feixes e a literatura vigente.

A literatura mostra que muitos trabalhos envolvendo tecidos vegetais são fundamentados na histologia tradicional que é baseada nas diferenças de formato dos tecidos e de seus constituintes (MINERVINI; SCHARR; TSAFTARIS, 2015). Isso faz com que a maioria dos trabalhos com tecidos vegetais publicados extraiam principalmente uma classe de descritores: medidas de forma (MAYOR; SILVA; SERENO, 2005; MAYOR; PISSARA; SERENO, 2008; MIAO; GANDELIN; YUAN, 2006; MURRAY; O'MALLEY, 1993; RAMOS et al., 2004; SANYAL; BHATTACHARYA; BANDYOPAHYAY, 2008).

Em nosso trabalho, todos os tecidos estudados são vasculares, ou seja, todos são formados por xilema e floema. Porém os formatos e tamanho das células constituintes desses tecidos podem apresentar particularidades entre os feixes. Por esse motivo, estudar e entender a forma e o tamanho das células pode trazer informações cruciais sobre o tecido em questão. Este fato atesta a necessidade de extrairmos atributos relacionados a forma das células.

Por outro lado, foi observado que existem outras classes de atributos que descrevem tecidos vegetais mas são poucos recorrentes na literatura, são eles: a multiescala (PLOTZE; BRUNO, 2007; SILVA; FLORINDO, 2019), a densidade celular (MAYOR; SILVA; SERENO, 2005; MAYOR; PISSARA; SERENO, 2008) e a regularidade estrutural (COSTA; ROCHA; LIMA, 2006; JURA-MORAWIEC, 2017).

Tomando como premissa que cada feixe possui uma disposição organizacional tecidual específica (MODESTO; SIQUEIRA, 2002) e que as relações de fronteira entre as células é variável (JURA-MORAWIEC, 2017), estudos que extraiam atributos referentes a multiescala, densidade celular e regularidade estrutural podem ser potencialmente descritivos dos feixes vasculares vegetais. Assim, nota-se que há uma lacuna na exploração desses descritores para a caracterização de tecidos vegetais, como, por exemplo, os feixes

vasculares vegetais.

Nesse cenário, o presente trabalho escolheu extrair medidas referentes a forma, densidade celular, regularidade estrutural e multiescala com o objetivo de caracterizar os feixes vasculares vegetais.

3.4.3 Análise dos atributos

Podemos entender a análise dos atributos como a etapa que irá selecionar quais descritores representam a informação primordial contida nos dados originais (LIU; MOTODA, 2012). Sendo assim, essa etapa busca pelo subconjunto com o menor número de atributos e com melhor acurácia no processo de classificação (OLIVEIRA; DUTRA; RENNÓ, 2005).

Nem sempre um grande número de atributos garante maior acurácia na classificação. Muitas vezes, dentro do conjunto original de descritores existem atributos irrelevantes ou redundantes e isso afeta a precisão dos classificadores e, portanto, precisam ser removidos (BRAMER, 2007).

Os métodos descritos para seleção de atributos selecionam os atributos pela avaliação individual ou pela avaliação de subconjuntos de descritores (LEE, 2005).

A análise que visa avaliar os atributos individualmente está atrelada a descrever a relevância dos atributos. E, portanto, eliminam apenas os descritores que são irrelevantes para a classificação (LEE, 2005). Um exemplo de uma análise individual é a inferência Bayesiana que avalia a sobreposição entre duas curvas de densidade de probabilidade, ou seja, a região de incerteza (GONZALEZ; WOODS, 2000). Quanto mais sobrepostas estiverem as curvas, menos relevante será o atributo analisado.

Já os métodos que avaliam subconjuntos, sejam eles mínimos (como por exemplo par a par) ou total (como o uso do próprio conjunto original de atributos) podem remover tanto atributos irrelevantes quanto redundantes (LEE, 2005). Um exemplo que analisa os atributos par a par é a correlação de Pearson (BENESTY et al., 2009). Esse método indica o grau de correlação entre dois atributos e, portanto, a redundância desses descritores. Esse subconjunto de atributos não correlacionados podem ser analisados por meio da dispersão e, conseqüentemente, pela matriz de espalhamento (S) que irá indicar o quão espalhado estão os clusteres no plano $2D$ (COSTA; CESAR, 2009). Dessa maneira, conseguimos remover tanto os atributos redundantes quanto os irrelevantes.

Em contrapartida, um método que consegue avaliar todos os atributos conjuntamente e, reduzir a dimensionalidade de maneira a priorizar os descritores que possuem maior peso para segregar as classes em questão é a análise de componente principal (COSTA; CESAR, 2009; JOLLIFFE, 1986). Sendo assim, ela também consegue remover atributos redundantes e irrelevantes.

Dentro desse cenário, o presente trabalho escolheu avaliar seus atributos por três vieses diferentes: individualmente, par a par e todo conjunto de atributos. As abordagens escolhidas foram capazes de remover descritores irrelevantes, redundantes e reduzir a dimensionalidade dos dados originais.

3.5 Reconhecimento e classificação de padrões

Após os atributos serem extraídos, eles são organizados em vetores de características (F) e assumem a seguinte forma:

$$F = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix}$$

na qual cada elemento, f_i , representa o i -ésimo descritor ou atributo e n é o número total desses descritores.

Um padrão é um conjunto de atributos que consegue caracterizar um objeto (GONZALEZ; WOODS, 2000; COSTA; CESAR, 2009). A classificação de imagens consiste em associar uma região da imagem a uma classe que descreve um objeto real. Sendo assim, o classificador usa vetor característica para atribuir o objeto a uma dada classe (DUDA; HART; STORK, 2012; OLIVEIRA; DUTRA; RENNÓ, 2005).

A classificação de um padrão pode ser feita de duas maneiras: supervisionada e não supervisionada. A diferença entre elas está que na primeira, as classes dos elementos são conhecidas pelo usuário. Em contrapartida, no segundo tipo as classes são definidas pelo próprio algoritmo de segregação (COSTA; CESAR, 2009; SILVA; SPATTI; FLAUZINO, 2010; DUDA; HART; STORK, 2012).

Em nosso trabalho, o objetivo foi classificar os feixes vasculares vegetais por meio de imagens provenientes da histologia. Portanto, aqui, as classes já eram conhecidas *a priori*. Por esse motivo, os classificadores utilizados são supervisionados.

Dentro do nicho dos classificadores supervisionados existem dois que são amplamente empregados em problemas envolvendo reconhecimento de padrões: K-Nearest Neighbor (k - NN) (COVER; HART, 1967) e Perceptron Multicamadas (PMC) (SILVA; SPATTI; FLAUZINO, 2010). O k - NN possui uma abordagem mais simples e que envolve a distância euclidiana entre uma amostra que se deseja descobrir a classe e as demais amostras previamente classificadas. A amostra será classificada de acordo com a classe mais próxima (COSTA; CESAR, 2009). Já o PMC é uma rede neural artificial e seu algoritmo é mais complexo e envolve uma série de equações que culminam no ajuste dos pesos sinápticos

até que a rede esteja convergida e, conseqüentemente, possa classificar uma amostra (GONZALEZ; WOODS, 2000).

Ambas as abordagens podem ser usadas para validar se uma classificação tem alta taxa de acurácia ou não. Para isso, basta utilizar um novo conjunto de dados previamente classificado após a etapa de treinamento tanto do *k-NN* quanto do *PMC*. Assim, ao final do algoritmo, basta comparar se as classes determinadas pelo algoritmo são as mesmas desejadas pelo usuário. Dentro desse cenário, um método muito utilizado para determinar a acurácia de um classificador é a matriz de confusão. Nela, a acurácia é obtida por meio da razão entre a soma dos acertos e o número de indivíduos testados.

O reconhecimento de padrões é a última etapa de todo processo que envolve a visão computacional, uma vez que, ao final desta etapa os objetos são atribuídos aos seus rótulos correspondentes.

4 Metodologia

A Figura 14 ilustra a visão geral de todas as partes que constituem a metodologia desse trabalho. As primeiras fases desse trabalho consistiram em fazer a montagem das lâminas histológicas e a aquisição das imagens digitais das amostras de seis tipos de feixes vasculares vegetais.

Posteriormente, as imagens foram divididas de maneira que 80% foram utilizadas na etapa de análise (ilustrada em laranja na Figura 14) e 20% na etapa de validação (em verde). Todas as classes de tecidos foram quantitativamente igualitárias nas duas etapas. Ambas as etapas possuem em comum os processos de segmentação manual e extração de atributos. Os atributos foram extraídos com base não só na literatura como também em consideração às particularidades dos feixes em estudos.

O objetivo da etapa de análise foi estudar se os atributos extraídos eram capazes de segregar e caracterizar os tecidos vegetais. Também, buscou-se nessa etapa selecionar quais foram os atributos que aprestaram maior influência na caracterização e separação desses tecidos.

A partir do resultado obtidos na etapa de análise, foi possível realizar a etapa de validação a qual foi responsável por atribuir valores capazes de medir e comparar a acurácia das categorias de atributos frente ao objetivo geral desse trabalho que é a caracterização e a separação dos tecidos vegetais dos feixes vasculares vegetais.

Os detalhes de cada processo serão esmiuçados nos itens a seguir.

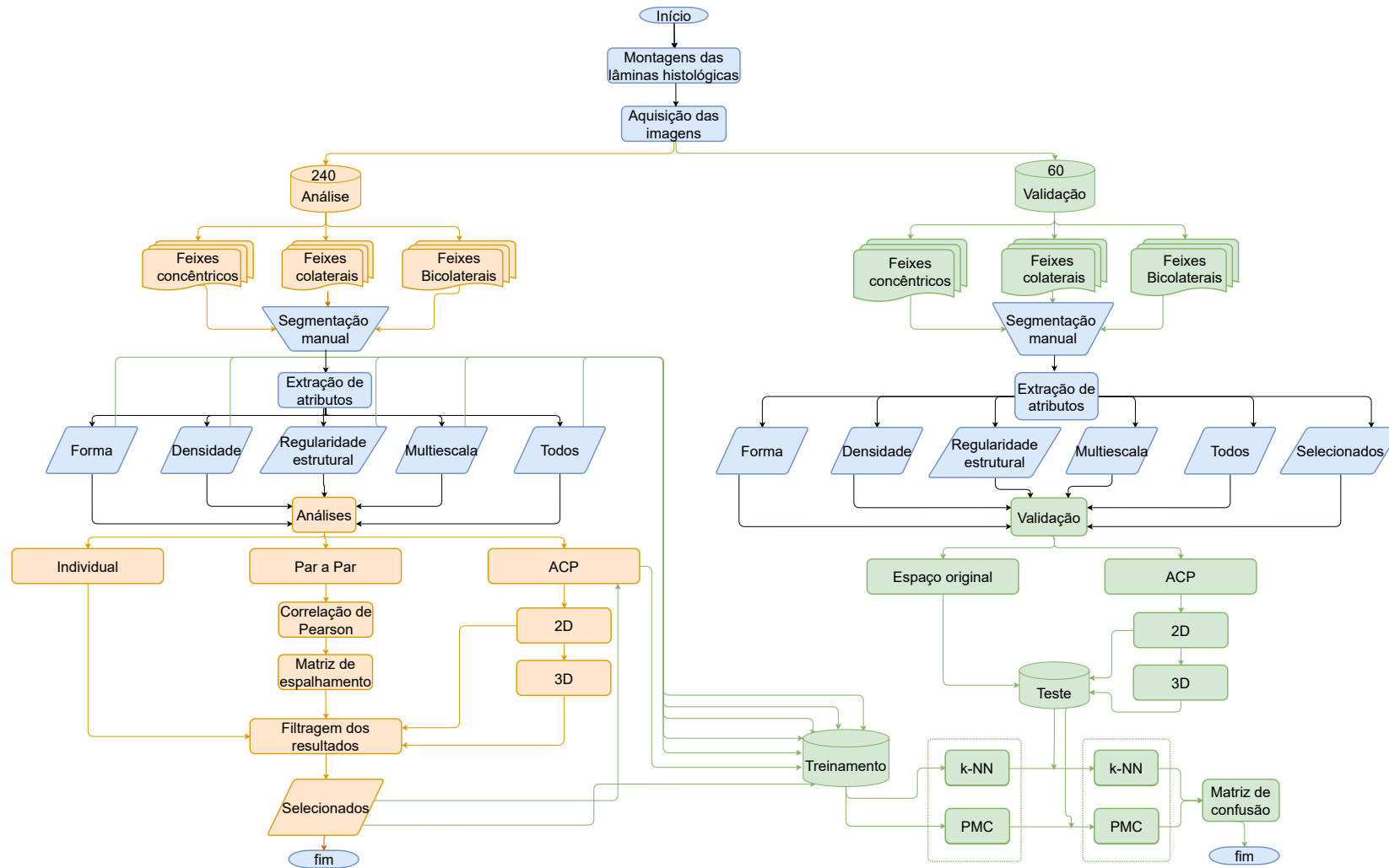


Figura 14 – Visão geral de todos os passos que compõe a metodologia desse trabalho. Os processos em laranja correspondem a etapa de análise, os em verde a etapa de validação e os azuis são referentes aos processos que são comuns às duas etapas.

Fonte: Autor

4.1 Montagem das lâminas histológicas

Os feixes vasculares são encontrados no caule e na raiz dos espécimes vegetais. Um mesmo indivíduo poderá apresentar diferentes tipos de feixes vasculares nesses órgãos (MODESTO; SIQUEIRA, 2002). Independente do órgão originário desses feixes, eles podem se organizar em três tipos diferentes: colaterais (aberto e fechado), biclateral e concêntrico (anfivasal e anficrival) (MENEZES, 2016; METCALFE; CHALK, 1950; MODESTO; SIQUEIRA, 2002; OGURA, 1972). Alguns tipos, como por exemplo o anfivasal, podem ser composto por tecidos primários e secundários. Para que todos os tipos de tecidos vasculares fossem investigados, esse trabalho explorou histologicamente o caule de algumas espécies vegetais conforme descrito na Tabela 1.

Tabela 1 – Quadro resumido relacionando as espécies vegetais com seus feixes vasculares correspondentes.

Espécies	Tipo do feixe vascular
<i>Cyperus rotundus</i> (Cyperaceae)	Colateral Fechado
<i>Baccharis dracunculifolia</i> (Asteraceae)	Colateral aberto
<i>Sechium edule</i> (Cucurbitaceae)	Biclateral
<i>Dracaena marginata</i> (Asparagaceae)	Anfivasal primário
<i>Dracaena marginata</i> (Asparagaceae)	Anfivasal secundário
<i>Microgramma vacciniifolia</i> (Polypodiaceae)	Anficrival

A primeira etapa da análise histológica vegetal consiste em confeccionar cortes histológicos. Eles foram feitos com lâminas de barbear novas e eles tinham uma espessura de aproximadamente 0,5cm. Posteriormente, foram mantidos por 15 dias em fixadores com o objetivo de manter sua morfologia o mais semelhante possível ao do vegetal vivo. Cada espécie foi submetida a um dos três fixadores mais comumente utilizados: FPA (mistura de formalina, ácido propiônico e álcool etílico) (PUIGDEFABREGAS; BENASET, 2000), FNT (formalina tamponada) (CLARK, 1981) e Zirkle (ZIRKLE, 1928).

Após ser fixado, os materiais biológicos foram desidratado por meio de série etanólica com concentração crescente de etanol (50, 70, 90 e 100%). Posteriormente, passaram por série etanol-xilólica (3:1, 1:1, 1:3, xilol puro), com a finalidade de prepará-los para receber a parafina e serem emblocados. Os blocos resultantes foram seccionados em micrótomo manual com lâmina de aço resultando em cortes individuais com 5mm de espessura que foram aderidos nas lâminas histológicas. As lâminas foram levadas a uma placa aquecedora para distensão dos cortes e, posteriormente, levadas à estufa para secar (MOLINARO; AL., 2009). A etapa seguinte consistiu em corar os cortes histológicos com hematoxilina-eosina seguindo o protocolo descrito em Junqueira e CARNEIRO (2013).

O preparo das lâminas foi feito sob a supervisão do Prof. Dr. Marcos Arduin no Laboratório de Anatomia e Morfologia Vegetal do Departamento de Botânica do Centro de

Ciências Biológicas e da Saúde (CCBS) da Universidade Federal de São Carlos (UFSCar), campus São Carlos.

4.2 Aquisição de imagens

As imagens dos cortes histológicos transversais dos caules vegetais foram adquiridas por meio de um microscópio óptico (*Carl Zeiss - Axiostar Plus Transmitted Light Microscope*). Nele existe um sistema de captura e digitalização de imagens histológicas que é composto por uma câmera inserida na objetiva do microscópio e um computador ligado diretamente a câmera, conforme evidenciado na Figura 9.

A captura e a digitalização das imagens histológicas foram feitas por meio do *software MicroCapture Versão 6.9.7*. Nele, foi possível fazer a leitura das lâminas histológicas, encontrar os feixes vasculares vegetais e realizar a digitalização desses tecidos.

O equipamento descrito acima foi utilizado no Laboratório de Imunologia Aplicada, cujo responsável é o Prof. Dr. Ricardo Carneiro Borra. Esse laboratório está localizado no Departamento de Genética e Evolução (DGE) da UFSCar, campus São Carlos.

Foram capturadas e digitalizadas imagens de todos os tipos de feixes vasculares vegetais indicados na Tabela 1. A Figura 15 ilustra um exemplo de cada feixe vascular estudado neste trabalho. As Figuras 15a e 15b ilustram os feixes vasculares do tipo colaterais fechado e aberto, respectivamente. A Figura 15c evidencia o feixe vascular do tipo bicolateral. As Figuras 15d, 15e e 15f ilustram os feixes vasculares do tipo concêntrico, sendo eles, respectivamente, anfigasal primário, anfigasal secundário e anfigival.

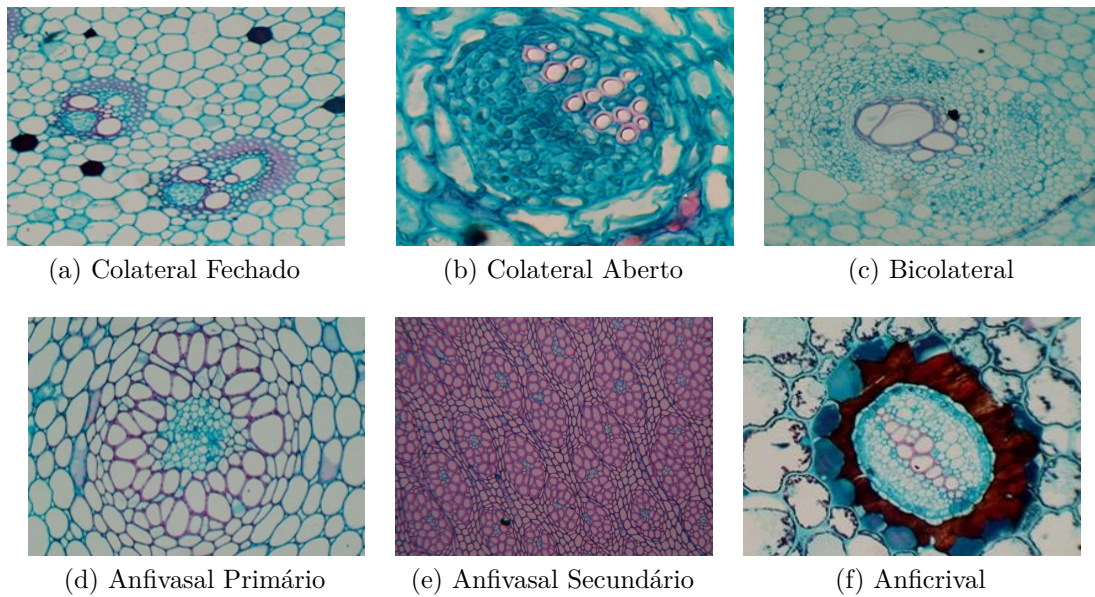


Figura 15 – Feixes vasculares vegetais obtidos por meio da captura e digitalização das lâminas histológicas vegetais. (a) Colateral Fechado (20x) extraído de *Cyperus rotundus*. (b) Colateral Aberto (40x) extraído de *Baccharis dracunculifolia*. (c) Bicolateral (10x) extraído de *Sechium edule*. (d) Anfivasal Primário (10x) extraído de *Dracaena marginata*. (e) Anfivasal Secundário (10x) extraído de *Dracaena marginata*. (f) Anficrival (20x) extraído de *Microgramma vacciniifolia*.

Fonte: Autor

Vale ressaltar que cada feixe é composto pelo floema e pelo xilema. Portanto, a mesma imagem foi contabilizada para o floema e para o xilema. Ao final, foram escolhidas 25 imagens de cada tipo organizacional, sendo que 20 imagens de cada tipo de feixe foram usadas para a etapa de análise (representada em laranja) e 5 para a fase de validação (representada em verde), conforme ilustrado na Figura 16.

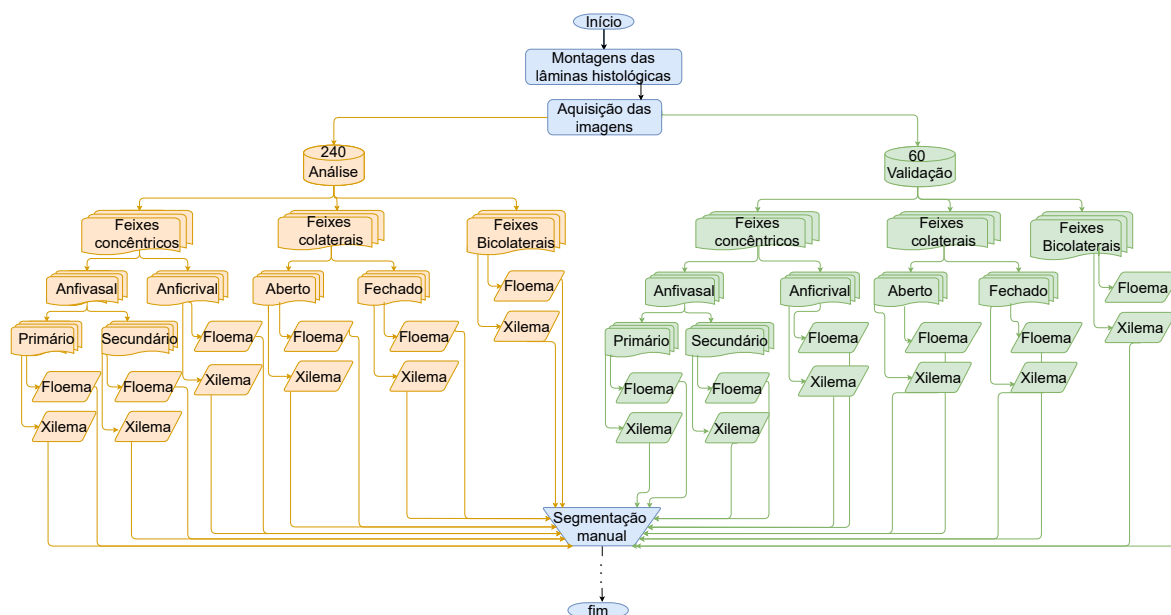


Figura 16 – Divisão das imagens digitalizadas dos feixes vasculares vegetais entre as etapas de análise (laranja) e validação (verde).

Fonte: Autor

Todas as imagens selecionadas foram segmentadas e analisadas conforme descrito a seguir.

4.3 Segmentação manual das imagens

As imagens de cada tipo de tecido vascular foram segmentadas manualmente com auxílio do *software Paint.net (versão 3.08)* com o objetivo de mapear os tecidos separando o xilema do floema. Essa etapa foi realizada no Grupo de Pesquisa em Visão Cibernética sob a supervisão do Prof. Dr. Luciano da Fontoura Costa do Instituto de Física de São Carlos (IFSC) localizado na Universidade de São Paulo (USP), campus São Carlos e foi orientada pelo Prof. Dr. Marcos Arduin do Departamento de Botânica do Centro de Ciências Biológicas e da Saúde (CCBS) da Universidade Federal de São Carlos (UFSCar), campus São Carlos.

Considerando a imagem como sendo uma matriz I , cada elemento de interesse foi separado por um tom de cinza que variou dentro do intervalo $1 \leq I(i, j) \leq 255$, onde os índices i e j indicam as coordenadas de um *pixel* na imagem. O fundo e os objetos não relevantes para o estudo receberam a intensidade zero ($I(i, j) = 0$).

A Figura 17, por exemplo, ilustra a segmentação manual do xilema em um feixe vascular do tipo anfivasal primário (vide 17a). A Figura 17b mostra o início da segmentação manual na qual, cada célula do xilema foi segmentada individualmente com um tom de cinza diferente. A Figura 17c ilustra a segmentação completa na qual fica evidente que cada

célula tem uma intensidade diferente. Nesse exemplo em específico, foram segmentadas 125 células do xilema, portanto, os tons de cinza ficaram dentro do intervalo $1 \leq I(i, j) \leq 125$. Para finalizar essa etapa, os tecidos não relevantes e o fundo receberam intensidade igual a zero ($I(i, j) = 0$), conforme ilustrado na Figura 17d.

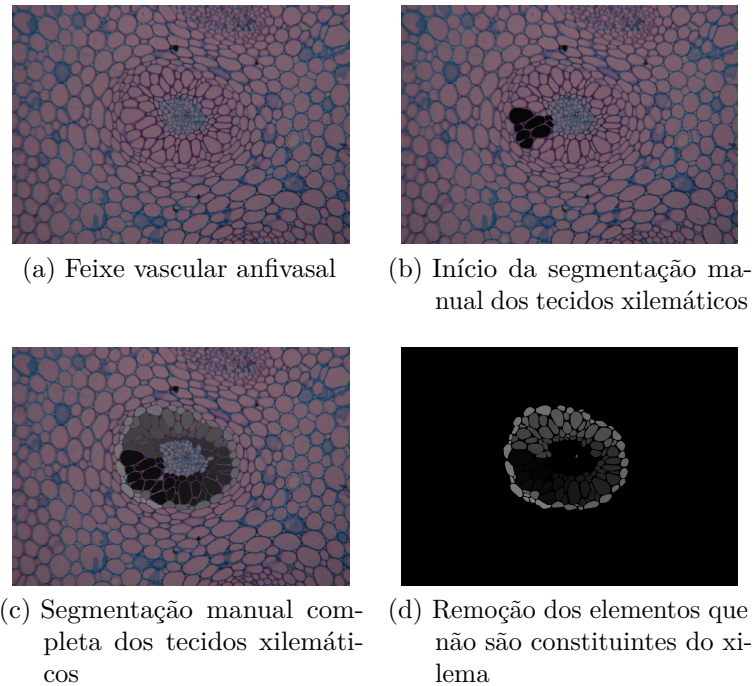


Figura 17 – Exemplo da segmentação manual em um tecido xilemático. (a) Imagem original com todos os tecidos. (b) Início da segmentação manual do xilema. (c) Todas as células do xilemas foram segmentadas com um tom de cinza diferente dentro do intervalo $1 \leq I(i, j) \leq 125$. (d) Os tecidos não relevantes e o fundo receberam intensidade igual a zero ($I(i, j) = 0$).

Fonte: Autor

Os feixes vasculares são formados pelo xilema e pelo floema. A segmentação de um feixe consiste em separar não só as células de um tecido individualmente, mas também diferenciar os dois tecidos. A Figura 18, por exemplo, ilustra a segmentação manual do floema no mesmo feixe vascular do tipo anfigvasal da Figura 17a. A Figura 18a mostra o início da segmentação manual e nota-se que os níveis de cinza são subsequentes ao do xilema. Como a última célula do xilema recebeu $I(i, j) = 125$, a primeira do floema tem $I(i, j) = 126$. A Figura 18b ilustra a segmentação completa com os tons de cinza dentro do intervalo $126 \leq I(i, j) \leq 244$. Para finalizar essa etapa, os tecidos que não são relevantes e o fundo receberam intensidade igual a zero ($I(i, j) = 0$), conforme ilustrado na Figura 18c.

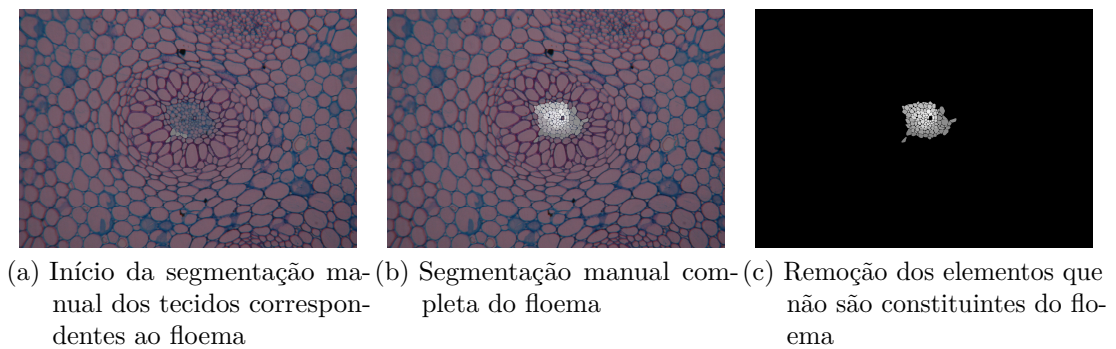


Figura 18 – Exemplo da segmentação manual do floema. (a) Imagem original com todos os tecidos. (b) Início da segmentação manual do floema. (c) Todas as células do floema foram segmentadas com um tom de cinza diferente dentro do intervalo $126 \leq I(i, j) \leq 244$. (d) Os tecidos não relevantes e o fundo receberam intensidade igual a zero ($I(i, j) = 0$).

Fonte: Autor

Como resultado final da segmentação da Figura 17a, teremos a segregação do xilema e do floema. As células correspondentes ao xilema possuem nível de cinza dentro do intervalo $1 \leq I(i, j) \leq 125$ e as do floema $126 \leq I(i, j) \leq 244$. O fundo e os tecidos que não são relevantes para o nosso estudo apresentaram $I(i, j) = 0$. Essa separação completa dos tecidos pode ser observada na Figura 19.

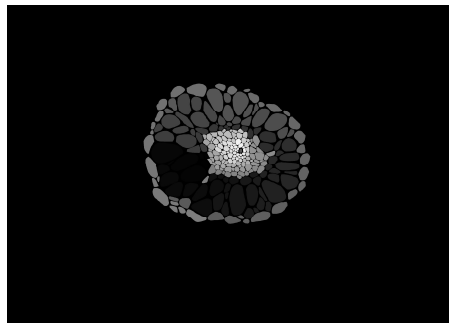


Figura 19 – Segmentação completa dos tecidos vasculares vegetais em um feixe do tipo anfibasal. As células correspondentes ao xilema possuem nível de cinza dentro do intervalo $1 \leq I(i, j) \leq 125$ e as do floema $126 \leq I(i, j) \leq 244$. O fundo e os tecidos não relevantes para o nosso estudo apresentaram $I(i, j) = 0$.

Fonte: Autor

Esse procedimento foi repetido para todas as imagens de todos os feixes vasculares vegetais. Ao final obtivemos 300 imagens, segmentadas e com a separação total entre xilema e floema. Dessas, 240 foram usadas na etapa de análise e as 60 restantes, para a etapa de validação.

4.4 Extração de atributos

Antes da etapa de extração, as imagens segmentadas foram normalizadas levando em consideração as magnitudes com as quais elas foram extraídas. Esse processamento foi importante para que as diferenças de valores extraídos de cada feixes estivessem relacionados apenas as características dos feixes e não ao processo de aquisição das imagens.

Todas as medidas foram extraídas a partir das imagens segmentadas na etapa anterior. Nessas imagens cada célula recebeu um tom de cinza específico. Dessa maneira, podemos trabalhar tanto com medidas que exploram as diferenças entre as células quanto as medidas que exploram as diferenças entre os tecidos.

Visando elucidar as particularidades teciduais dos feixes apresentados acima, foram elencadas quatro categorias de medidas sendo elas: i) medidas de forma; ii) medidas de densidade, iii) medidas de regularidade estrutural e iv) medidas multiescala (esta última categoria é formada pelas medidas de regularidade estrutural e de densidade que foram extraídas por meio da abordagem de multiescala). As categorias i, ii e iii unidas formam uma quinta que é composta por todos os atributos extraídos, por esse motivo, essa categoria foi nomeada como *Todos*. A Figura 20 ilustra quais atributos fazem parte de cada categoria.

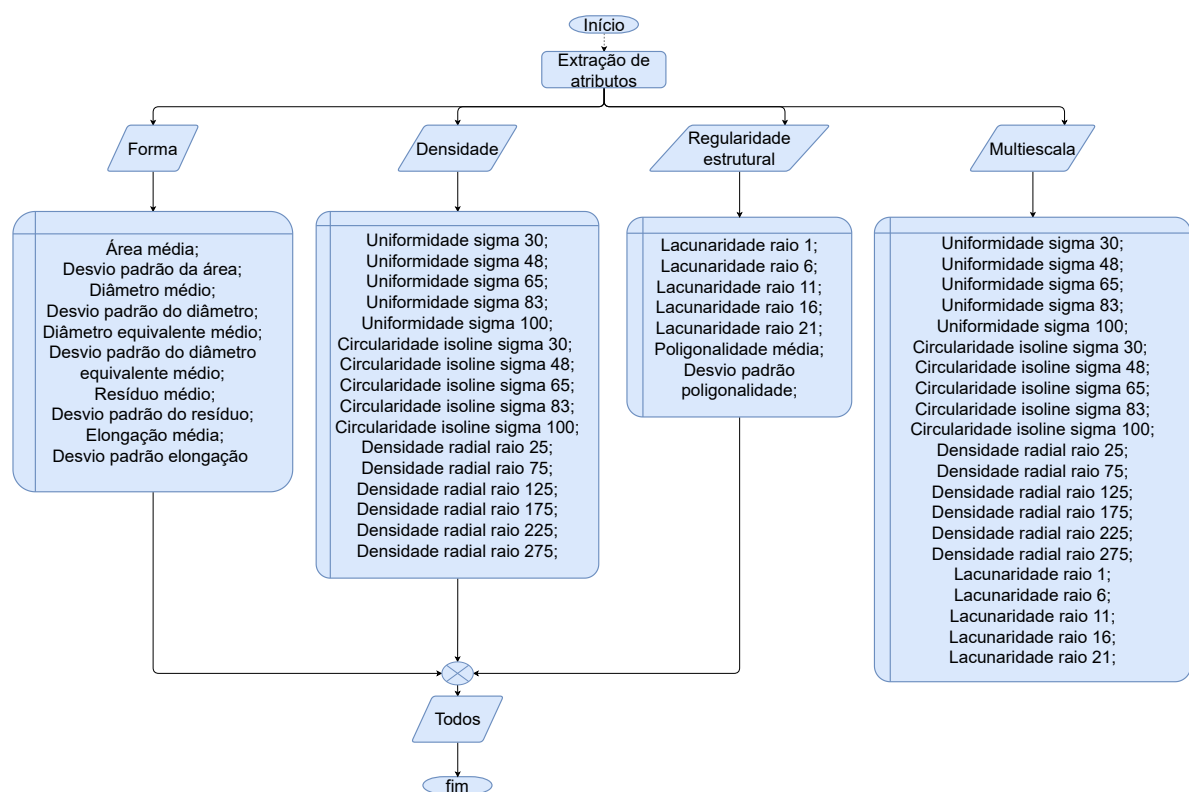


Figura 20 – Categoria de atributos extraídas: i) medidas de forma; ii) medidas de densidade, iii) medidas de regularidade estrutural; iv) medidas multiescala e v) todos

Fonte: Autor

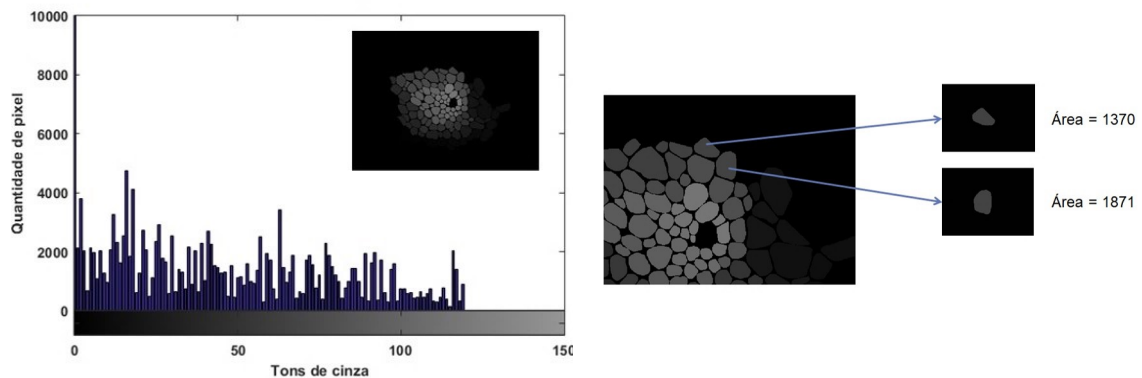
4.4.1 Medidas de forma celular

Conforme evidenciado na Figura 20 foram realizadas as seguintes medidas de forma: área celular, resíduo celular (casco convexo), diâmetro celular, diâmetro celular equivalente e alongação celular.

Como os nomes indicam, os cálculos foram feitos para cada célula, por isso, ao final de cada análise foram calculados a média e o desvio padrão para cada tecido.

4.4.1.1 Área celular

A medida da área foi realizada pela contagem dos *pixels* (COSTA; CESAR, 2009) de cada célula que compõe o xilema e o floema. A Figura 21 ilustra a contagem de *pixel* em uma imagem segmentada manualmente. De maneira didática, foi plotado um histograma no qual cada pico representa um tom de cinza e, conseqüentemente, uma célula. A quantidade de *pixels* correspondente a cada barra é a área da célula segmentada com aquela intensidade (vide Figura 21a). Por exemplo, na Figura 21b ilustramos a contagem de *pixel* de duas células individualmente, a primeira tem 1370 *pixels* e, portanto sua área é 1370. Analogamente, a segunda é constituída por 1871 *pixels* e, concomitantemente, sua área é 1871.



(a) Histograma de uma imagem segmentada manualmente
(b) Exemplo do cálculo da área de duas células do floema de uma imagem segmentada manualmente

Figura 21 – Contagem de *pixel* em uma imagem segmentada manualmente. (a) Contagem de *pixel* por meio de histograma. Cada pico representa um tom de cinza e, conseqüentemente, uma célula. A quantidade de *pixel* correspondente a cada barra é a área celular. (b) Exemplo do cálculo da área de duas células do floema de uma imagem segmentada manualmente.

Fonte: Autor

4.4.1.2 Resíduo celular e casco convexo

Por definição, casco convexo é calculado como sendo o menor polígono capaz de conter o objeto de interesse (COSTA; CESAR, 2009) e (DUDA; AL., 1973). O resíduo celular é definido como a área do casco convexo menos a área da célula.

Para fins didáticos, selecionamos uma célula do floema de um feixe anficrival (vide Figura 22). O menor polígono capaz de circundar essa estrutura é o casco convexo, conforme ilustrado na Figura 22a. Em contrapartida, a área do casco convexo menos a área da célula é o resíduo celular (mostrado em cinza na Figura 22b).

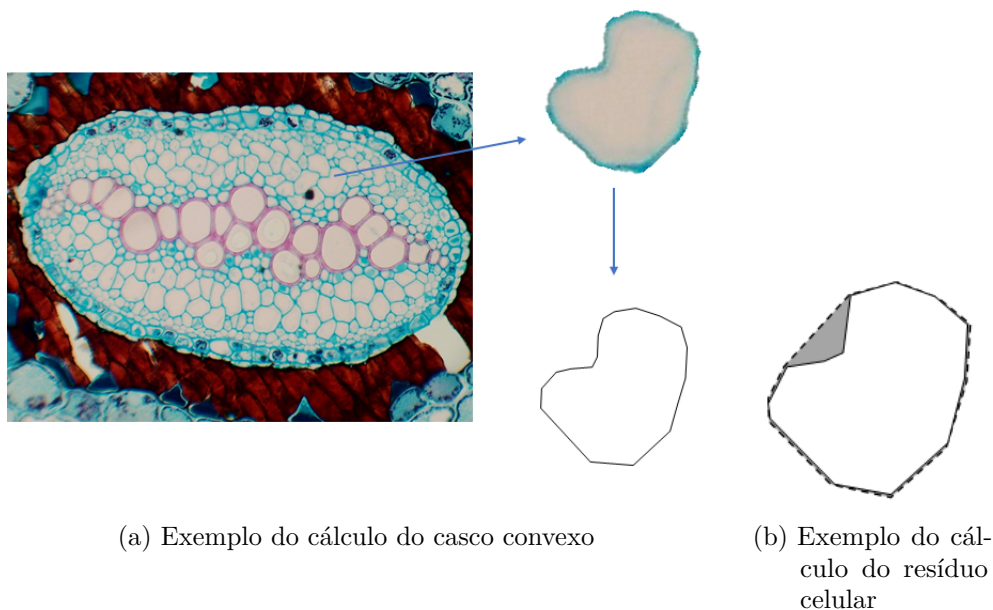


Figura 22 – Casco convexo e resíduo celular. (a) O casco convexo é o menor polígono capaz de conter o elemento de interesse. (b) O resíduo celular está representado em cinza e é definido como a área do casco convexo menos a área da célula.

Fonte: Autor

Nesse trabalho, foi calculado o resíduo de cada célula que compõe cada tipo organizacional dos feixes vasculares vegetais.

4.4.1.3 Diâmetro celular

O diâmetro foi medido como sendo a maior distância entre dois pontos quaisquer da borda da célula (COSTA; CESAR, 2009), conforme ilustrado na Figura 23.

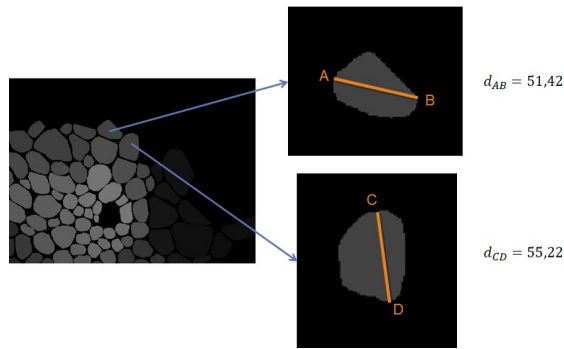


Figura 23 – Exemplo do cálculo do diâmetro celular calculado para duas células individualmente.

Fonte: Autor

A Equação 4.1 descreve o cálculo da distância euclidiana entre pontos A e B . Essa equação foi utilizada para realizar o cálculo do diâmetro celular.

$$D_{AB} = \sqrt{(x_A - x_B)^2 + (y_A - y_B)^2} \quad (4.1)$$

4.4.1.4 Diâmetro celular equivalente

Essa medida foi calculada com base no raio equivalente de cada célula. Para isso, células foram representadas como sendo circunferências e o diâmetro equivalente (COSTA; CESAR, 2009) foi calculado segundo a conforme a Equação 4.2:

$$D_{eq} = 2\sqrt{\frac{A}{\pi}} \quad (4.2)$$

onde A é a área celular calculada anteriormente.

A Figura 24 ilustra o cálculo do diâmetro equivalente celular por meio da Equação 4.2 para duas células cuja área celular foi calculada e mostrada na Figura 21.



Figura 24 – Exemplo do cálculo do diâmetro equivalente celular calculado para duas células individualmente.

Fonte: Autor

4.4.1.5 Elongação celular

Sabemos que uma imagem em tons de cinza pode ser representada por uma matriz I , na qual os valores estão dentro do intervalo $1 \leq I(i, j) \leq 255$, onde os índices i e j indicam as coordenadas de um *pixel* na imagem. Considerando uma célula representada por N *pixels*, as coordenadas dos seus *pixels* podem ser armazenadas em uma matriz X de tamanho $N \times 2$, onde cada linha possui a coordenada de um *pixel*.

Ao calcularmos a matriz de covariância sobre as colunas da matriz X , obtemos a matriz C de tamanho 2×2 . A partir disso, podemos encontrar o eixo principal (λ_1) e o secundário (λ_2) da célula, uma vez que, eles estão relacionados com os autovetores de C . O autovetor relacionado com o maior autovalor indica o eixo principal da célula enquanto que o segundo autovetor define o eixo secundário (COSTA; CESAR, 2009).

Didaticamente, podemos inferir que a direção na qual a célula é mais alongada é chamada de eixo principal (λ_1), enquanto que perpendicularmente a essa direção temos o eixo secundário (λ_2). A Figura 25 ilustra os dois eixos de uma célula.

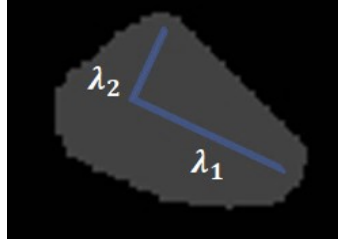


Figura 25 – Elongação celular: a direção na qual a célula é mais alongada é chamada de eixo principal (λ_1), enquanto que perpendicularmente a essa direção temos o eixo secundário (λ_2).

Fonte: Autor

Com os valores de λ_1 e λ_2 obtidos, calculamos a elongação celular de acordo com a Equação 4.3:

$$Elong = \frac{\lambda_1}{\lambda_2} \quad (4.3)$$

onde λ_1 e λ_2 são, respectivamente, o maior e o menor autovalor.

4.4.2 Medidas de densidade

As medidas que foram calculadas nesse item foram: uniformidade espacial da densidade dos centroides; circularidade *isoline*; densidade radial do centroide.

Todas elas estão relacionadas ao centro de massa (centroide) de cada célula e, por isso, se faz necessário entender como o centroide é calculado.

Centro de massa ou centroide

Como as imagens foram segmentadas, o centroide $C_m = (\bar{x}, \bar{y})$ foi estimado como sendo a média dos valores das coordenadas dos *pixels*, conforme descrito nas Equações 4.4 e 4.5.

Seja (x_i, y_i) a coordenada do i -ésimo *pixel* de uma célula, as coordenadas \bar{x} e \bar{y} do centroide são estimadas como

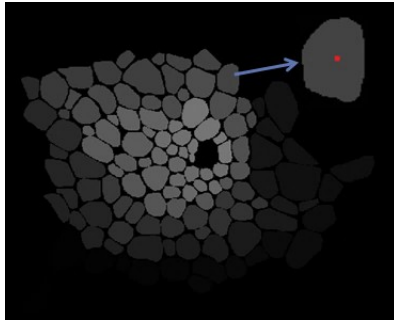
$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad (4.4)$$

$$\bar{y} = \frac{1}{N} \sum_{i=1}^N y_i \quad (4.5)$$

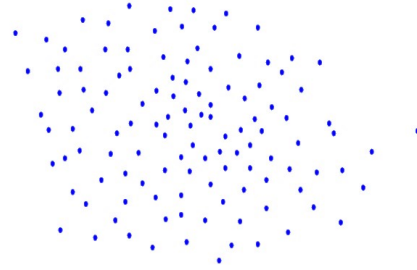
onde N é o número de *pixels* de cada célula.

Na Figura 26 temos o exemplo visual de onde estão localizados os centro de massas de um floema. A Figura 26a pontua a posição do centroide de apenas uma célula desse

tecido. Enquanto que, a Figura 26b evidencia o mapa dos centros de massa de todas as células. Nela o fundo foi removido para enfatizar apenas o posicionamento dos centroides.



(a) Ilustração centro de massa para uma célula



(b) Ilustração dos centros de massa para todas as células de um tecido vegetal

Figura 26 – Centro de massa ou centroíde. (a) O centroíde foi calculado apenas para uma célula. (b) O centro de massa foi calculado para todas as células do tecido vegetal.

Fonte: Autor

Método de Janela de Parzen de estimação de densidades

A uniformidade espacial e a circularidade, descritas a seguir, foram calculadas a partir da estimativa não-paramétrica da distribuição de probabilidades dos centroides celulares. Para isso, foi aplicado o método de Janela de Parzen de estimação de densidades (DUDA; HART; STORK, 2012), que consiste em tratar cada centroíde como um impulso unitário e realizar a convolução tais impulsos com um kernel Gaussiano (IZIDORO, 2005; PARZEN, 1962).

A função gaussiana possui um parâmetro σ (sigma) que determina a largura do pico e está relacionado com a ideia de escala. Portanto, um sigma pequeno leva a uma função bem detalhada, que é influenciada por qualquer variação na posição das células, enquanto um sigma grande gera funções suaves, que não são influenciadas por pequenas variações nas posições das células.

Os sigmas utilizados foram: 30, 48, 65, 83 e 100, portanto, as medidas uniformidade espacial e circularidade foram calculadas para os cinco sigmas listados.

4.4.2.1 Uniformidade espacial da densidade dos centroides

Após a realização dos cálculos do centro de massa e do método de Janela de Parzen de estimação de densidades, descritos acima e, com o objetivo de estimar a uniformidade espacial da densidade dos centros de massa das células, foi aplicado a entropia de Shannon

(SHANNON, 1948) descrita na Equação 4.6.

$$H(x) = - \sum_i p_i \ln p_i \quad (4.6)$$

onde p_i é o valor de probabilidade estimado para o *pixel* i .

4.4.2.2 Circularidade *isoline*

Para calcular essa medida, a distribuição de probabilidades dos centroides foi limiarizada (com limiar sendo o 95º percentil de todos os valores de densidade calculados) obtendo-se então uma imagem binária. Um imagem binária é aquela que só possui apenas duas intensidades I .

O cálculo da circularidade (COSTA; CESAR, 2009) é definido segundo a Equação 4.7.

$$C = \frac{4\pi A}{P^2} \quad (4.7)$$

onde A e P são, respectivamente, a área e o perímetro da imagem binária.

4.4.2.3 Densidade radial do centroide

Tomando como referência o mapa de centro de massa obtido na Figura 26b e partindo do pressuposto que o centroide global (C_g) de uma tecido é o centro de massa de um objeto formado por todas as células da imagem, calculamos o C_g , que está indicado por um ponto vermelho na Figura 27, do floema apresentado na Figura 26a.

Dois círculos concêntricos em C_g foram construídos, sendo que o círculo interno tem raio r_i e o externo, R_e que é definido como $R_e = r_i + \Delta r$, vide Figura 27.

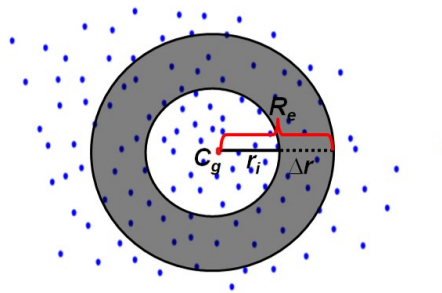


Figura 27 – Densidade radial do centroide: C_g é o centro de massa global da imagem; r_i o raio do círculo interno; R_e o raio do círculo externo e Δr a variação entre os raios.

A densidade radial dos centroides é calculada como sendo a relação do número de centroides celulares (n_i) dentro do anel (hachurado em cinza na Figura 27) e a área desse anel (A_i), conforme descrito em (COSTA; CESAR, 2009) e apresentado na Equação 4.8.

$$D_i = \frac{n_i}{A_i} \quad (4.8)$$

Dessa maneira, um descritor \vec{D}_i foi definido, contendo a densidade de centroides para valores de raio distintos, sendo eles: 25, 75, 125, 175, 225 e 275.

4.4.3 Medidas de regularidade estrutural

Foram analisadas medidas referentes a regularidade estrutural dos tecidos, foram elas: lacunaridade e poligonalidade.

4.4.3.1 Lacunaridade

Essa medida está relacionada com a distribuição de lacunas em uma imagem. Mais especificamente, ela mede como os *pixels* estão distribuídos e organizados na região de interesse, quantificando assim, o espaço preenchido (MANDEBROT, 1982).

Nesse trabalho usamos um método que calcula a lacunaridade de forma auto-referenciada à forma do objeto (RODRIGUES; BARBOSA; COSTA, 2005; ALLAIN; CLOITRE, 1991; LUCENA; STOSIC, 2014; BACKES; BRUNO, 2005; JR; LANGE; MARKS, 1996; EINSTEIN; WU; GIL, 1998).

A lacunaridade foi aplicada aos tecidos vasculares considerando o plano de fundo da imagem como um objeto e as células suas lacunas, conforme ilustrado na Figura 28.

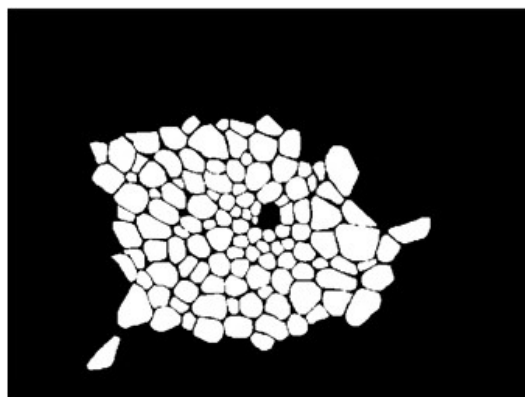


Figura 28 – Lacunaridade: as células do tecido vegetal são consideradas as lacunas da imagem.

Fonte: Autor

Esta propriedade indica a homogeneidade do tecido e foi calculada dentro de uma região limitada por um raio R em torno de um dado *pixel* da imagem. O parâmetro R foi variado entre os seguintes valores: 1, 6, 11, 16 e 21.

Valores baixos de lacunaridade (próximos de um) estão relacionados com formas invariantes à translação (RODRIGUES, 2014). Desse modo, imagens que possuem lacunas de tamanhos irregulares ou distribuídas de maneira heterogênea apresentam alta lacunaridade (RODRIGUES; BARBOSA; COSTA, 2005).

4.4.3.2 Poligonalidade

A poligonalidade (COSTA; ROCHA; LIMA, 2006) está relacionada com a regularidade de ângulos entre pontos vizinhos em um diagrama de Voronoi (AURENHAMMER, 1991).

A Figura 29 mostra as etapas necessárias para que a poligonalidade possa ser medida. Esse atributo foi calculado baseando-se nos centros de massa das células, representados por um conjunto de pontos (p_1, p_2, \dots, p_k) no plano (COSTA; ROCHA; LIMA, 2006). A Figura 29a ilustra os centroides das células correspondentes a um tecido vegetal. Cada ponto p_k é o centro de massa de uma célula desse tecido.

Cada ponto (p_k) está relacionado com uma célula de Voronoi (R_k) . O conjunto de células de Voronoi formam o Diagrama de Voronoi, conforme ilustra a Figura 29b.

Células de Voronoi adjacentes à uma célula R_k específica indicam quais pontos compõe a vizinhança Γ_k do ponto p_k . Cada célula vizinha à (R_k) é denominada (N_k) (COSTA; ROCHA; LIMA, 2006). Na Figura 29c, podemos observar que a célula de Voronoi R_k está indicada pela cor vermelha, enquanto que as que compõe a vizinhança Γ_k estão em amarelo.

A medida de poligonalidade consiste em ordenar os pontos em Γ_k no sentido horário ou anti-horário e medir os ângulos entre pontos adjacentes (α_i) . A Figura 29d ilustra esse processo. Os pontos foram ordenados no sentido horário e ângulos α_i calculados.

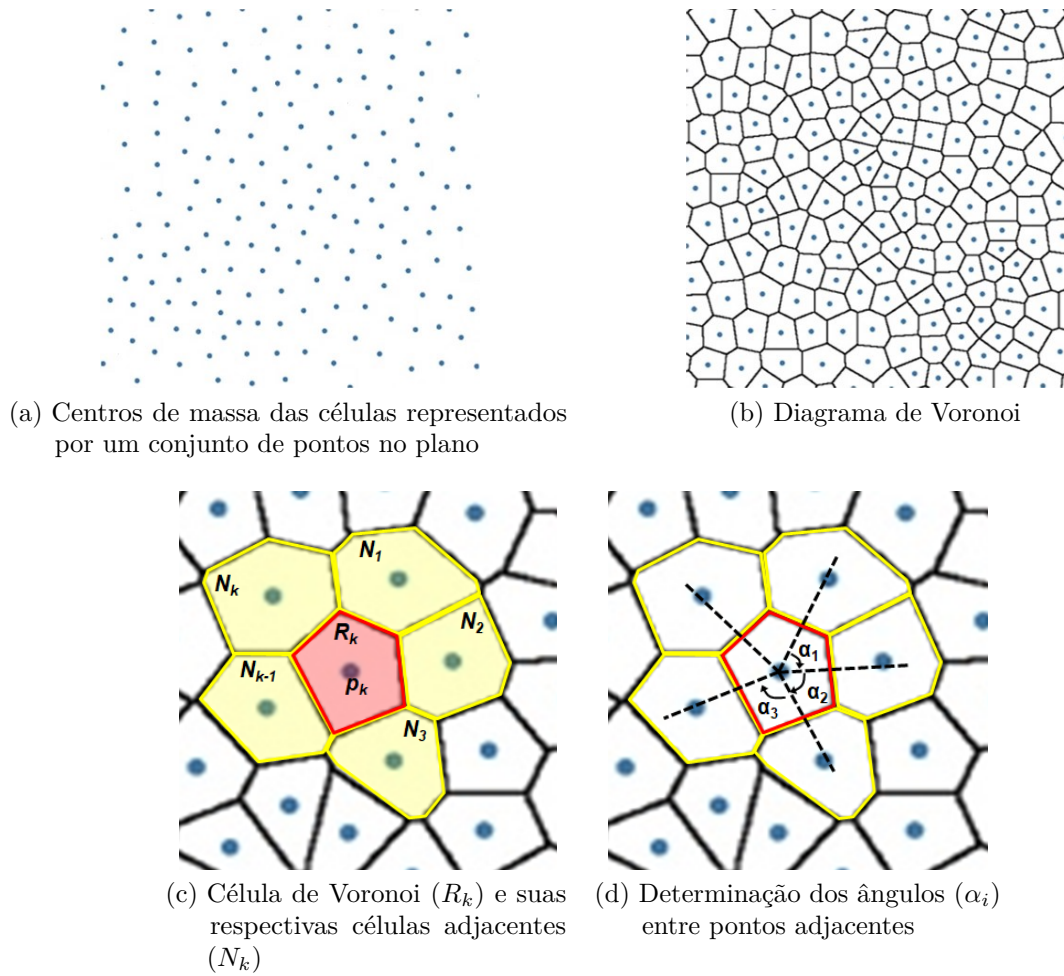


Figura 29 – Etapas para o cálculo da poligonalidade. (a) Ilustra os centros de massa das células de um tecido vegetal. (b) Diagrama de Voronoi. (c) Enfatiza uma célula de Voronoi R_k e suas respectivas vizinhas N_k . (d) Ilustra os ângulos entre pontos adjacentes.

Fonte: Autor

Com os α_i em mãos, podemos aplicar a equação da poligonalidade conforme descrita na Equação 4.9.

$$L_k = \frac{1}{\sum_i |\alpha_i - \beta| + 1} \quad (4.9)$$

onde β é o ângulo de referência a ser utilizado. Neste trabalho $\beta = 60^\circ$.

Dessa maneira, podemos inferir que o cálculo da poligonalidade é feito para todas as células de Voronoi que compõe o diagrama. Devido a isso, a medida final de cada tecido foi obtida com base na média dos valores de poligonalidade de cada célula. Foi também calculado o valor do desvio padrão para cada tecido.

4.5 Análise dos atributos

Uma vez extraídos os atributos, é importante analisar quais atributos possuem maior relevância para descrever os feixes vasculares. Essa análise também foi importante para selecionar um conjunto de atributos capaz de segregar com maior eficiência os tecidos vegetais estudados. A Figura 30 ilustra a visão geral dos processos realizados nessa etapa do trabalho.

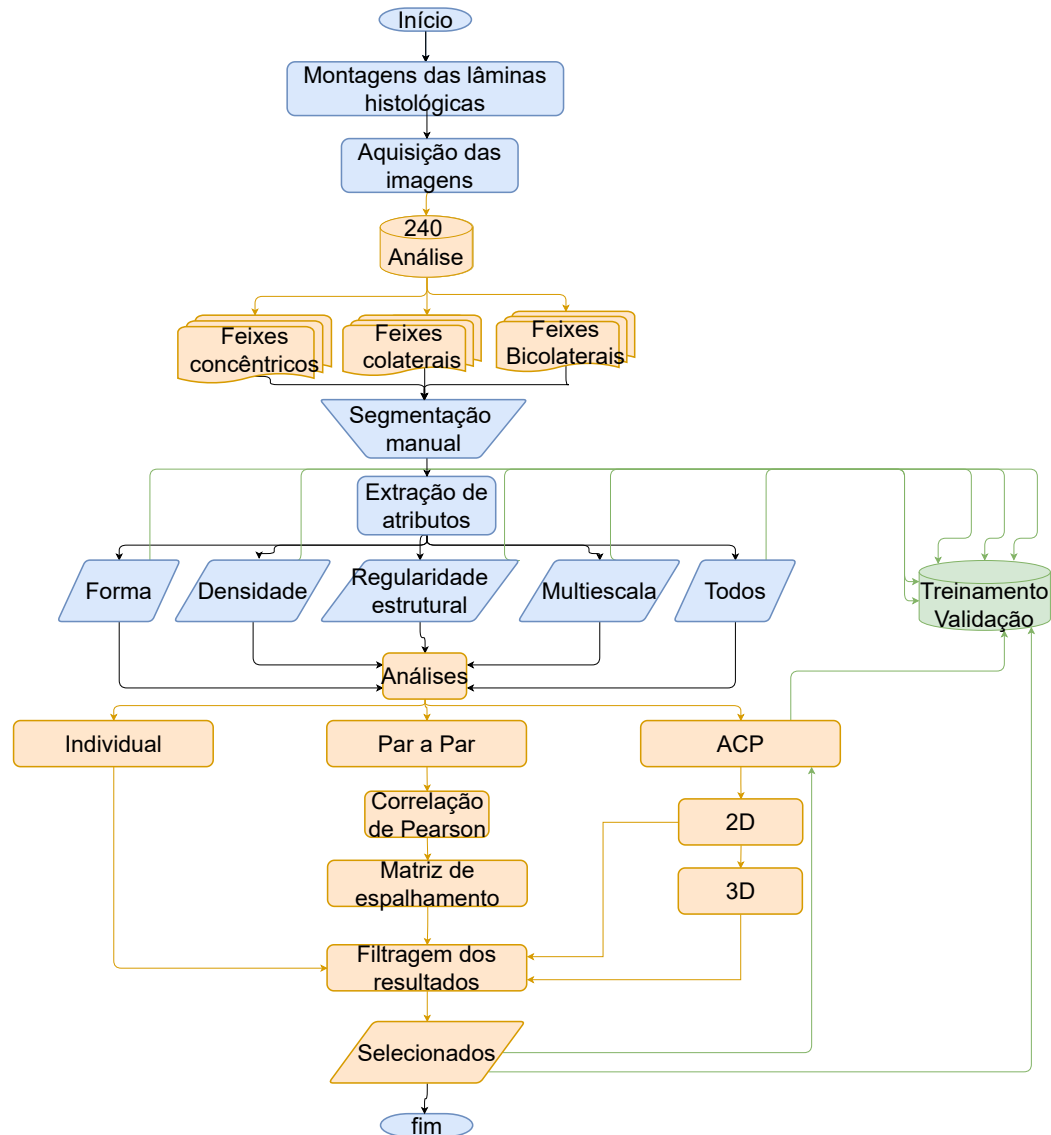


Figura 30 – *Flowchart* com o passo a passo genérico da etapa de análises de atributos. Os processos em laranja são exclusivos dessa etapa, os em azul são gerais ao trabalho e os em verde se relacionam com a etapa de validação.

Fonte: Autor

Com o objetivo de tentar encontrar quais são os melhores atributos para a classificação e identificação dos seis feixes vasculares vegetais estudados, as 240 imagens provenientes dos feixes concêntricos (Anfivascular primário, Anfivascular secundário, Anfivascular),

feixes colaterais (Aberto e Fechado) e feixes bicolaterais (as classes estão detalhadas na Tabela 2), foram analisadas para os 33 atributos extraídos e ilustrados na Figura 20.

As abordagens analíticas utilizadas foram: análise individual (função densidade de probabilidade); análise par a par (coeficiente de correlação de Pearson e matriz de espalhamento) e análise de componente principal (*ACP*).

Tabela 2 – Divisão dos tecidos vasculares com base no tipo do feixe vascular.

	Concêntrico	Bicolateral	Colateral
Composto por	Anfivasal Primário Xilema Anfivasal Primário Floema Anfivasal Secundário Xilema Anfivasal Secundário Floema Anficrival Xilema Anficrival Floema	Bicolateral Xilema Bicolateral Floema	Colateral Aberto Xilema Colateral Aberto Floema Colateral Fechado Xilema Colateral Fechado Floema

Ao final dessa etapa, os atributos mais recorrente considerando as três abordagens aplicadas foram reunidos em um novo grupo denominado *Selecionados*.

Também é importante ressaltar que as matrizes de atributos oriundas das 240 amostras analisadas nessa etapa foram usadas para o treinamento da etapa de validação.

4.5.1 Análise individual

Buscando inferir sobre o potencial de cada medida em caracterizar e segregar os seis feixes vasculares estudados, foram plotados os gráficos individuais de cada medida. Os gráficos foram feitos baseando-se na função densidade de probabilidade. Essa função descreve a probabilidade relativa de uma variável aleatória ter um valor dado, conforme indicado na Equação 4.10.

$$P[a \leq x \leq b] = \int_a^b f(x)dx \quad (4.10)$$

A equação 4.10 indica que a probabilidade P de uma medida (x) ser atribuída a uma faixa particular é dada pela integral da densidade de x , ($f(x)$), sobre um intervalo $a \leq x \leq b$ (GONZALEZ; WOODS, 2000). Por exemplo, vamos imaginar que a variável x é o diâmetro caulinar medido em centímetros e que queremos plotar a função densidade de probabilidade dessa variável para três espécies diferentes de flores. Para isso, foram medidos os diâmetros caulinares desses espécimes vegetais e observou-se que os valores para as espécie 1 (esp1), espécie 2 (esp2) e espécie 3 (esp3) estavam contidos em intervalos distintos conforme descrito na Tabela 3.

Tabela 3 – Valores dos diâmetros caulinares (x) dos espécimes vegetais (esp1, esp2 e esp3).

	Diâmetro caulinar (x)
Esp1	$11,2 \leq x \leq 13,1$
Esp2	$13,2 \leq x \leq 13,9$
Esp3	$13,5 \leq x \leq 14,4$

A Figura 31 ilustra o gráfico da densidade de probabilidade do exemplo acima. O gráfico tem no eixo x os valores de diâmetro caulinar e no eixo y a probabilidade da variável ser atribuída a uma faixa particular.

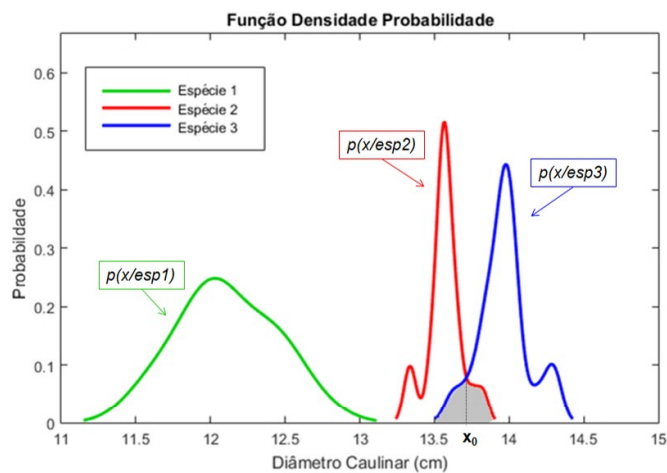


Figura 31 – Exemplo de um gráfico de densidade de probabilidade para três espécies vegetais. Cada curva representa a densidade de probabilidade de uma espécie. A curva verde não possui nenhuma sobreposição, em contrapartida, as curvas vermelha e azul apresentam uma região de incerteza, hachurada em cinza.

Fonte: Autor

Segundo a inferência Bayesiana, quando há uma região de sobreposição entre duas curvas de densidade de probabilidade, temos uma região de incerteza, ou seja, se uma amostra cair dentro dessa região, haverá uma certa indecisão sobre qual é a classe correspondente daquela amostra. Dessa maneira, quanto menos sobrepostas estiverem as curvas, melhor os grupos estarão separados e, por consequência, as classes estarão melhores definidas. Por exemplo, na Figura 31 as curvas referente as espécies 1 e 2 apresentam uma pequena área de incerteza (hachurada no gráfico) enquanto que a curva para a espécie 3 não possui sobreposição. Portanto, a medida x (diâmetro caulinar) foi uma boa escolha para classificar essas três espécies vegetais.

O ponto x_0 é a fronteira de decisão, ou seja, se uma amostra a ser classificada tiver um valor inferior a x_0 ela terá maior probabilidade de pertencer a espécie 2 enquanto que se o valor for superior ao x_0 será da espécie 3.

Essa etapa é importante pois é por meio dela que saberemos se a escolha de um atributo foi eficaz para conseguir extrair informações precisas que podem culminar na segregação e caracterização de um determinado feixe vascular. Nela, selecionamos qualitativamente os gráficos que possuíram nenhuma ou pequena área de sobreposição entre as curvas de densidade de probabilidade.

4.5.2 Análise par a par

Nessa etapa, as medidas foram avaliadas aos pares. Essa abordagem é baseada no coeficiente de correlação de Pearson (p) que indica o grau da correlação entre as variáveis X e Y (BENESTY et al., 2009), sendo que $-1 \leq p \leq 1$. Esse coeficiente foi calculado conforme descrito na Equação (4.11):

$$p = \frac{cov(X, Y)}{\sqrt{var(X).var(Y)}} \Leftrightarrow \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{[\sum_{i=1}^n (x_i - \bar{x})^2][\sum_{i=1}^n (y_i - \bar{y})^2]}} \quad (4.11)$$

na qual cov é a covariância calculada entre as variáveis X e Y e var é a variância de cada uma delas.

Para fins didáticos apresentaremos um exemplo ilustrativo considerando quatro variáveis (A, B, C e D) extraídos de 3 espécies vegetais (1, 2 e 3). A Figura 32 ilustra os gráficos de dispersão entre as variáveis estudadas nesse exemplo, tomadas par a par. Vale ressaltar que uma vez plotado o gráfico cujo eixo y corresponde a variável A e o eixo x a variável B , não é necessário estudar o gráfico cujo eixo y representa a variável B e o eixo x a variável A . A inversão dos eixos não traz mudança no coeficiente de Pearson e nem influencia no comportamento da disposição dos grupos estudados.

Para cada gráfico foi calculado o coeficiente de correlação de Pearson. Os valores estão indicados na parte superior de cada *plot* da Figura 32. O índice p além de indicar o grau de correlação entre as duas variáveis em estudo, fornece informações sobre a direção dessa correlação. Valores de p que estão contidos no $0,9 \leq p \leq 1$ indicam uma relação linear positiva, ou seja, as variáveis são diretamente proporcionais. Valores de p que fazem parte do intervalo $-0,9 \leq p \leq -1$ representam uma relação linear negativa, ou seja, as variáveis são inversamente proporcionais. Seguindo o mesmo raciocínio, valores de p contidos no intervalo $-0,9 < p < 0,9$ indicam que as variáveis em estudo não possuem uma relação bem estabelecida.

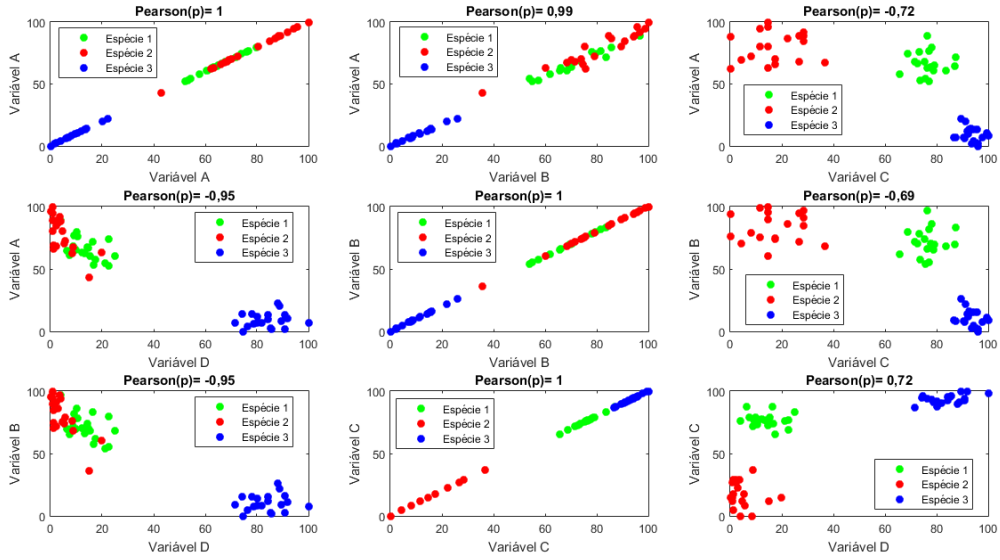


Figura 32 – Gráficos de dispersão referente as variáveis ilustrativas A, B, C e D, tomadas par a par.

Fonte: Autor

A última coluna da Figura 32 é composta por gráficos em que as variáveis que compõem o eixo x e o eixo y não possuem uma correlação bem estabelecida. E, como pode ser observado, os *clusters* das espécies ficam mais distribuídos por todo espaço $2D$. Para atingir o objetivo desse trabalho, ou seja, conseguir encontrar atributos capazes de segregar bem os tipos de feixes vasculares em estudo, foram levados em consideração apenas os pares que possuem um valor de p dentro do intervalo $-0,9 < p < 0,9$.

Com a finalidade de avaliar o quão espalhado estão os *clusters* em um gráfico de dispersão, utilizamos os traços das matrizes de espalhamento (S) *interclusters* (S_{inter}) e *intraclusters* (S_{intra}), assim, poderemos ter uma melhor quantificação sobre a dispersão entre os grupos e dentro dos grupos, respectivamente.

A matriz de espalhamento (S) indica a dispersão total de um vetor de característica (COSTA; CESAR, 2009). Ela pode ser definida conforme indicado na Equação 4.12.

$$S = \sum_{i=1}^N (\vec{f}_i - \vec{M})(\vec{f}_i - \vec{M})^T \quad (4.12)$$

A matriz de espalhamento *intraclusters* (S_{intra}) é a somatória da dispersão em cada *cluster*, ou seja, em cada classe estudada (C_i). A dispersão de cada classe S_i pode ser calculada conforme demonstrado na Equação 4.13.

$$S_i = \sum_{i \in C_i} (\vec{f}_i - \vec{\mu}_i)(\vec{f}_i - \vec{\mu}_i)^T \quad (4.13)$$

Dessa maneira então, conseguimos calcular a S_{intra} , conforme descrito na Equação 4.14.

$$S_{intra} = \sum_{i=1}^K S_i = \sum_{i \in C_i} \sum_{i=1}^K (\vec{f}_i - \vec{\mu}_i)(\vec{f}_i - \vec{\mu}_i)^T \quad (4.14)$$

A matriz de espalhamento *interclusters* (S_{inter}) expressa como uma classe está dispersa em relação a outra. As classes aqui são representadas em termos de seus centroides. A S_{inter} pode ser definida conforme descrito na Equação 4.15.

$$S_{inter} = \sum_{i=1}^K N_i (\vec{\mu}_i - \vec{M})(\vec{\mu}_i - \vec{M})^T \quad (4.15)$$

Para o nosso estudo, é necessário que as classes estejam poucos dispersas internamente e bastante dispersas entre si. Por esse motivo, foi feita a relação (R) entre S_{inter} e S_{intra} e, posteriormente, calculado o traço dessa matriz R . O traço (t) é definido como sendo a somatória dos elementos da diagonal principal de uma matriz.

Retomando o exemplo ilustrativo da Figura 32, calculamos o traço da matriz R dos três gráficos de dispersão contidos na última coluna da Figura 32. Os valores do traço desses gráficos estão mostrados na Tabela 4.

Tabela 4 – Valores do traço das matrizes R dos gráficos selecionados

Eixo Y	Eixo X	$T(S_{intra})$	$T(S_{inter})$	$T(R)$
A	C	9047,25	119454,96	35,02
B	C	10441,64	123190,63	32,85
C	D	4851,20	139893,89	59,07

A Tabela 4 evidencia que o gráfico que contém no eixo y a variável C e no eixo x a variável D é o que possui o maior $T(R)$ e, portanto, é o gráfico que apresenta os *clusters* relacionados as três espécies menos dispersos internamente e mais distantes entre eles.

De maneira análoga, a mesma metodologia apresentada no exemplo ilustrativo foi seguida para os diferentes feixes vasculares estudados nesse trabalho. Os gráficos de dispersão foram plotados respeitando as classes de tecidos e com os 33 atributos foram tomados par a par. Os gráficos que apresentaram coeficiente de correlação de Pearson dentro do intervalo $-0,9 < p < 0,9$ foram selecionados e tiveram os traços das matrizes *intra* e *interclusters* calculados. Posteriormente, os gráficos foram filtrados por ordem decrescente de $T(R)$, sendo que os 10 gráficos que obtiveram maiores valores de $T(R)$ foram escolhidos.

A partir desses 10 gráficos escolhidos, pode-se elencar quais os 20 atributos foram mais representativos para cada determinada classe tecidual. Para melhor visualização da

ocorrência dos atributos, foram plotados gráfico de barras com base na frequência onde cada barra indica um dos 33 atributos extraídos e a altura dessas barras ilustra quantas vezes esse atributo foi recorrente dentro daquela classe.

O objetivo principal dessa etapa foi encontrar pares de atributos que apresentem um alto valor para S_{inter} e baixo para S_{intra} . Ademais, também objetivou estudar a frequência desses atributos para cada classe tecidual.

4.5.3 Análise de componentes principais

A Análise de Componentes Principais (*ACP* ou do inglês *PCA - Principal Component Analysis*) (JOLLIFFE, 1986) é uma técnica utilizada para eliminar redundâncias e reduzir a dimensionalidade dos dados conservando ao máximo a variância dos dados originais (VASCONCELOS, 2007). Dessa maneira, a *ACP* está relacionada com a ideia de redução do número de atributos, com a menor perda possível de informação (VARELLA, 2008).

De maneira geral, a *ACP* consiste em uma transformação linear em que cada componente principal é uma combinação linear das variáveis originais (VASCONCELOS, 2007; COSTA; CESAR, 2009). Em outras palavras, a matriz de covariância da matriz de atributos é calculada e os autovetores são usados para que a transformação linear seja efetuada. A principal propriedade dessa técnica é minimizar a covariância entre os atributos, maximizando, assim, a variância ao longo dos novos eixos não correlacionados (COSTA; CESAR, 2009).

Por outra forma, a partir de uma matriz F , de tamanho $n \times m$, preenchida com os dados obtidos na etapa de extração de atributos, foi gerada uma matriz de covariância simétrica A de tamanho $m \times m$.

Os componentes principais serão os autovetores com os maiores autovalores associados à matriz A . Os autovalores são as raízes do polinômio de grau n em λ , (BOLDRINI; AL., 1980) resultante da Equação 4.16.

$$\det[A - \lambda I] = 0 \quad (4.16)$$

onde I é a matriz identidade.

Para encontrar os autovetores basta substituir cada autovalor na Equação 4.17:

$$(A - \lambda_i I)\vec{v}_i = 0 \quad (4.17)$$

sendo que cada \vec{v}_i terá o λ_i correspondente (BOLDRINI; AL., 1980).

Após a determinação dos componentes principais, calcula-se a contribuição de cada componente principal λ_i na projeção dos dados. Essa contribuição é dada em porcentagem

e representa a proporção de variância total explicada pelo componente principal \vec{v}_i . Em nosso trabalho, os três primeiros componentes principais foram escolhidos. Essa escolha foi feita para cada classe tecidual estudada.

Para fins didáticos, usaremos o *dataset* nomeado Iris para ilustrar um exemplo sobre o uso e a aplicação da análise de componentes principais. Esse *dataset* é bastante utilizado em livros didáticos sobre processamento de imagem, como por exemplo Gonzalez e Woods (2000). Ele foi criado pelo biólogo Ronald Fisher e é constituído por um vetor de características composto por 4 atributos sendo eles: comprimento das sépalas, largura das sépalas, comprimento das pétalas e largura das pétalas. Esse atributos foram extraídos de três espécies vegetais: *Iris setosa*, *Iris virginica* e *Iris versicolor* (FISHER, 1936). Cada grupo de espécie foi analisado com base em 50 indivíduos. Para efeitos didáticos, apresentaremos esse *dataset* apenas com os atributos comprimento sépala, largura sépala e comprimento pétala.

A matriz F representa uma parte da matriz completa $F_{150 \times 3}$ e nela as colunas 1, 2 e 3 ilustram, respectivamente, os valores obtidos por meio dos atributos comprimento sépala, largura sépala e comprimento pétala. A matriz $A_{3 \times 3}$ representa a matriz de covariância obtida com os valores associados a matriz F .

$$F = \begin{bmatrix} \vdots & \vdots & \vdots \\ 5,1 & 3,5 & 1,4 \\ 4,9 & 3,0 & 1,4 \\ 4,7 & 3,2 & 1,3 \\ 4,6 & 3,1 & 1,5 \\ 5,0 & 3,6 & 1,4 \\ 5,4 & 3,9 & 1,7 \\ 4,6 & 3,4 & 1,4 \\ \vdots & \vdots & \vdots \end{bmatrix} \rightarrow \text{MaCov}(F) = A = \begin{bmatrix} 0,6857 & -0,0393 & 1,2737 \\ -0,0393 & 0,1880 & -0,3217 \\ 1,2737 & -0,3217 & 3,1132 \end{bmatrix}$$

O próximo passo para a realização da *ACP* consiste em calcular os autovalores λ_i da matriz de covariância A . Os autovalores λ_i são reais e positivos e podem ser ordenados. A matriz $\lambda(A)$ ilustra o cálculo dos autovalores para a matriz A .

$$\lambda(A) = \begin{bmatrix} 0,0594 & 0 & 0 \\ 0 & 0,2410 & 0 \\ 0 & 0 & 3,6865 \end{bmatrix} \rightarrow \begin{cases} \lambda_1(A) = 0,0594 \\ \lambda_2(A) = 0,2410 \\ \lambda_3(A) = 3,6265 \end{cases}$$

$$\lambda_3(A) \geq \lambda_2(A) \geq \lambda_1(A) \geq 0$$

Cada $\lambda_i(A)$ terá associado a ele um autovetor $\vec{v}_i(A)$, sendo que $\vec{v}_1(A) \rightarrow \lambda_1(A)$, $\vec{v}_2(A) \rightarrow \lambda_2(A)$ e $\vec{v}_3(A) \rightarrow \lambda_3(A)$. A matriz $\vec{v}(A)$ representa os autovetores associados

aos autovalores da matriz $\lambda(A)$. Ao final, a matriz $\vec{v}(A)$ foi ordenada de acordo com os autovalores ordenados da matriz A . Esses vetores ordenados estão na matriz $v_{ordem}(A)$, sendo esses chamados de componentes principais.

$$\vec{v}(A) = \begin{bmatrix} -0,6627 & 0,6392 & 0,3902 \\ 0,6640 & 0,7425 & -0,0887 \\ 0,3464 & -0,2003 & 0,9165 \end{bmatrix}$$

$$\xrightarrow{\text{Ordenado}} v_{ordem}(A) = \begin{bmatrix} 0,3902 & 0,6392 & -0,6627 \\ -0,0887 & 0,7425 & 0,6640 \\ 0,9165 & -0,2003 & 0,3464 \end{bmatrix}$$

O cálculo da contribuição de cada componente principal ($v_{ordem(i)}(A)$) na projeção dos dados foi feito com base na variância total ($VarTotal$) dos dados. Esse valor foi obtido como sendo a soma da diagonal principal da matriz A . Para esse exemplo, a $VarTotal = 3,9869$. A partir disso, foi possível calcular a variância de cada componente principal ($V_{(v_{ordem(i)}(A))}$).

$$V_{(v_{ordem(i)}(A))} = \begin{cases} V_{(v_{ordem(1)}(A))} = \left(\frac{\lambda_{(3)}(A)}{VarTotal}\right)x100 = \left(\frac{3,6865}{3,9869}\right)x100 = 92,47\% \\ V_{(v_{ordem(2)}(A))} = \left(\frac{\lambda_{(2)}(A)}{VarTotal}\right)x100 = \left(\frac{0,2410}{3,9869}\right)x100 = 6,04\% \\ V_{(v_{ordem(3)}(A))} = \left(\frac{\lambda_{(1)}(A)}{VarTotal}\right)x100 = \left(\frac{0,0594}{3,9869}\right)x100 = 1,49\% \end{cases}$$

Dessa forma, podemos inferir que o primeiro componente principal detém 92,47% da variação total dos dados originais, o segundo 6,04% e o terceiro 1,49%. Conseqüentemente, os dois primeiros eixos principais correspondem a 98,51% da variação total dos dados originais. Foram escolhidos os componentes principais que acumularam aproximadamente 70% ou mais da proporção da variância total (VARELLA, 2008). Portanto, o primeiro e o segundo componente principal foram escolhidos.

Partindo dessa premissa, podemos construir a matriz de transformação (L) que é composta pelos dois primeiros autovetores da matriz $v_{ordem}(A)$.

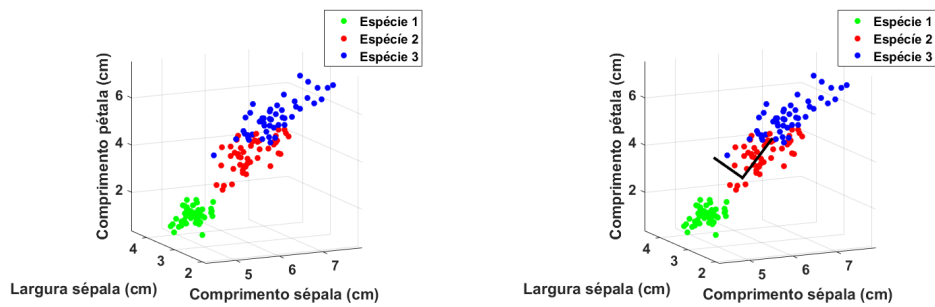
$$L = \begin{bmatrix} 0,3902 & 0,6392 \\ -0,0887 & 0,7425 \\ 0,9165 & -0,2003 \end{bmatrix}$$

Por último, calculamos a transformação linear dos novos eixos (\tilde{F}_i), conforme descrito na Equação 4.18.

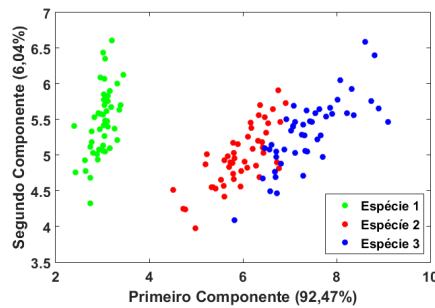
$$\tilde{F}_i = L\vec{F}_i \quad (4.18)$$

Com a finalidade de ilustrar as mudanças espaciais obtidas no exemplo citado acima, representamos graficamente os efeitos da *ACP* sobre os dados contidos na matriz F , vide Figura 33. O gráfico de dispersão original dos dados contidos na matriz F foi plotado e pode ser observado na Figura 33a. Nela os dados estão plotados em um gráfico 3D onde os eixos x , y e z correspondem, respectivamente, ao comprimento das sépalas, largura das sépalas e comprimento de pétalas.

A Figura 33b ilustra a orientação do $\lambda_{(3)}(A)$ e do $\lambda_{(2)}(A)$, sendo que o $\lambda_{(3)}(A)$ representa a orientação da maior variância dos dados. Ambos estão representados pelas linhas sólidas de cor preta.



(a) Representação no espaço original dos dados contidos na matriz F (b) *Plot* 3D dos dados originais da matriz F com as principais orientações de variância dos dados apresentadas pelas linhas sólidas pretas.



(c) Nova distribuição espacial dos dados após a aplicação da *ACP*.

Figura 33 – Visualização das mudanças espaciais antes e depois da aplicação da *ACP*.(a) Representação no espaço original dos dados contidos na matriz F ; (b) *Plot* 3D dos dados originais da matriz F com as principais orientações de variância dos dados apresentadas pelas linhas sólidas pretas; (c) Nova distribuição espacial dos dados após a aplicação da *ACP* onde é possível notar que houve uma redução espacial a pequena perda de informação.

Fonte: Autor

Após calculados os $\lambda_{(i)}(A)$ são obtidos os autovetores ordenados (matriz $v_{ordem}(A)$) e, conseqüentemente, a matriz de transformação L pode ser obtida. A partir dela, foi feita a combinação linear dos valores originais de \vec{F}_i , resultando em novos eixos \vec{F}_i , conforme

descrito na Equação 4.18. A Figura 33c ilustra a distribuição dos dados após a aplicação da *ACP*. O gráfico *2D* foi plotado sendo que o eixo x e y são, respectivamente, o primeiro e o segundo componente principal. Juntos, eles totalizam 98,51% da variação total dos dados originais. Esse exemplo ilustra como a *ACP* é importante para reduzir a dimensionalidade dos dados, minimizar a covariância entre os atributos e maximizar a variância ao longo dos novos eixos não correlacionados.

Para o trabalho em questão, essa abordagem foi aplicada seguindo todas as etapas exemplificadas nessa seção e respeitando as classes teciduais em avaliação.

4.6 Validação dos resultados

O objetivo dessa etapa foi avaliar de forma quantitativa qual o melhor conjunto de atributos capaz de segregar com maior eficiência os tecidos vegetais estudados. Para isso, as categorias de atributos i) Forma. ii) Densidade; iii) Regularidade estrutural; iv) Multiescala; v) Todos e vi) Selecionados foram avaliadas por meio de duas abordagens: *k-Nearest Neighbor* (*k-NN*) e por redes neurais, mais especificamente, Perceptron Multicamada (*PMC*). A Figura 34 ilustra a visão geral dessa etapa.

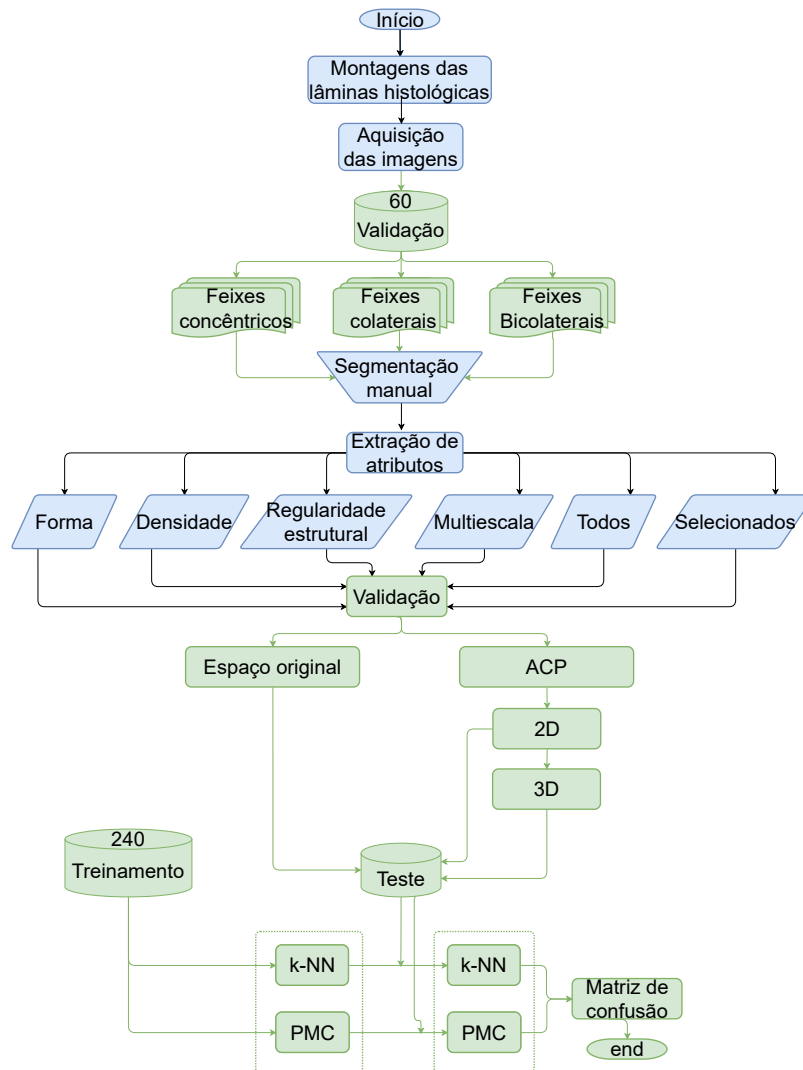


Figura 34 – *Flowchart* com o passo a passo genérico da etapa de validação. Os processos em verde são exclusivos dessa etapa, os em azul são gerais ao trabalho e os em laranja se relacionam com a etapa de análise dos atributos.

Fonte: Autor

Esse trabalho utilizou ao total 300 imagens. Ambas as abordagens classificadoras foram treinadas com as 240 amostras da etapa de análise dos atributos. Sendo assim, as 60 amostras remanescentes foram utilizadas na fase de teste tanto do *K-NN* quanto do *PMC*. Vale ressaltar que todas as categorias de atributos foram avaliadas tanto nas dimensões originais quanto após a diminuição de dimensionalidade oriunda da análise de componente principal (*ACP*).

A avaliação da acurácia foi obtida por meio da matriz de confusão.

4.6.1 k -Nearest Neighbor(k -NN)

k -Nearest Neighbor(k -NN) foi proposto por Cover e Hart (1967) e é uma das abordagens supervisionadas não paramétricas mais simples de classificação, uma vez que, ela é baseada em encontrar os k vizinhos mais próximos de uma determinada amostra. O k é um número inteiro positivo e representa o número de vizinhos que será buscado, por exemplo se $k = 1$ a amostra que se deseja classificar terá sua classe atribuída à mesma do vizinho mais próximo. Em casos onde $k > 1$ a classe é considerada como aquela exibida pela maioria desses vizinhos. Em casos onde k é par pode ocorrer empate, nesse contexto, uma das classes é selecionada arbitrariamente (COSTA; CESAR, 2009).

Por se tratar de uma abordagem supervisionada, se faz necessário conhecer previamente as classes de um conjunto de dados. Essa primeira etapa é conhecida como fase de treinamento e é caracterizada por separar os grupos de maneira que cada classe esteja representada no espaço de dispersão.

Supondo que temos um conjunto \tilde{F}_i de 150 amostras já classificadas em 3 classes (Espécie 1, Espécie 2 e Espécie 3) conforme ilustrado na Figura 33c e que o objetivo agora seja classificar três novos indivíduos A , B e C dispersos pela extensão do plano formado pelas coordenadas x e y que são respectivamente, o primeiro e o segundo componente principal, conforme ilustrada na Figura 35.

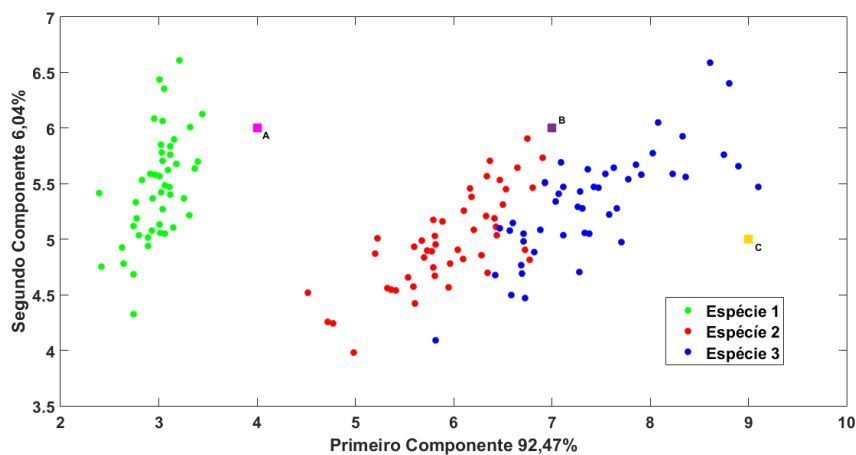


Figura 35 – Exemplo da representação do k -NN para classificar três indivíduos: Dispersão dos indivíduos já classificados nas espécies 1,2 e 3 e dos novos indivíduos A , B e C .

Fonte: Autor

Para classificar os três indivíduos A , B e C basta realizar o cálculo da distância euclidiana, vide Equação 4.19, entre o novo ponto A , B e C e todos os pontos já classificados, sendo que A é um novo indivíduo e p_i os pontos já classificados. Nesse exemplo $1 \leq i \leq 150$. Posteriormente, as distâncias são ordenadas a fim de elencar quais são os vizinhos mais

próximos a esses pontos.

$$D_{Ap_i} = \sqrt{(x_A - x_{p_i})^2 + (y_A - y_{p_i})^2} \quad (4.19)$$

O número de vizinhos selecionados irá depender o valor atribuído a variável k . Nesse exemplo, assim como em toda dissertação, o k escolhido foi 3 ($k = 3$). Sendo assim, para cada ponto A , B e C foram selecionados os três vizinhos mais próximos, conforme ilustrado na Figura 36.

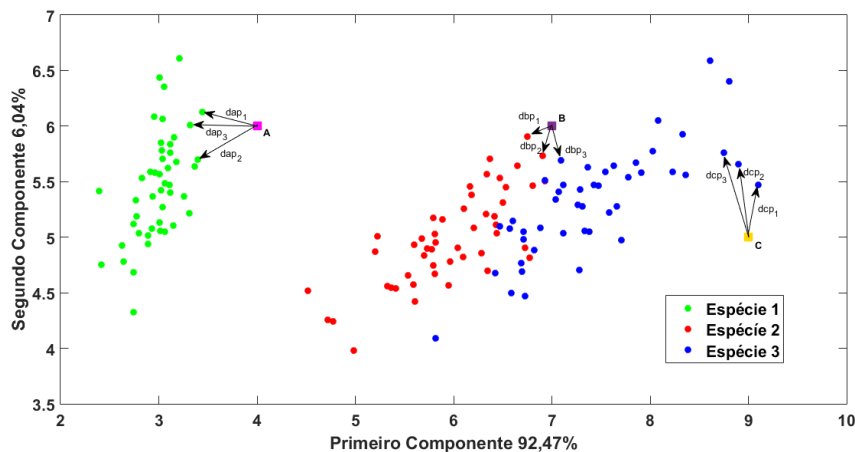


Figura 36 – Exemplo da representação do k -NN para classificar três indivíduos: Distância dos pontos A , B e C em relação aos seus respectivos três vizinhos mais próximos.

Fonte: Autor

Os valores da distância entre os pontos a serem classificados e seus três vizinhos mais próximos estão apresentados na Tabela 5.

Tabela 5 – Valores da distância entre os pontos a serem classificados e os respectivos três vizinhos mais próximos

Indivíduo	Distância	Classe do vizinho	Classe atribuída
A	dap1 = 0,569	Espécie 1	Espécie 1
	dap2 = 0,676	Espécie 1	
	dap3 = 0,681	Espécie 1	
B	dbp1 = 0,262	Espécie 2	Espécie 2
	dbp2 = 0,284	Espécie 2	
	dbp3 = 0,322	Espécie 3	
C	dcp1 = 0,480	Espécie 3	Espécie 3
	dcp2 = 0,667	Espécie 3	
	dcp3 = 0,804	Espécie 3	

Pela definição da abordagem k -NN, a amostra que se deseja classificar terá sua classe atribuída a mesma dos vizinhos mais próximos (COSTA; CESAR, 2009). Nesse exemplo,

o $k = 3$ e, portanto, a classe foi considerada como aquela exibida pela maioria desses vizinhos. Para os indivíduos A e C os três vizinhos mais próximos eram, respectivamente, da espécie 1 e da espécie 3. Devido a isso, a amostra A foi atribuída a espécie 1 e a B a espécie 3. Em contrapartida, a amostra B teve dois de seus vizinhos mais próximos da espécie 2 e um da espécie 3. Por esse motivo, foi atribuída a amostra B a espécie 2, conforme evidenciado na Tabela 5.

Outra abordagem para o uso do k - NN está na validação de dados. A técnica é a mesma com a diferença de que já se sabe as classes dos indivíduos A , B e C . Sendo assim, é possível inferir se as classes atribuídas pelo algoritmo estão certas ou não. Essa etapa também pode ser denominada como etapa de teste.

Vamos supor que as classes dos indivíduos A , B e C são, respectivamente, espécie 1, espécie 3 e espécie 3. Vimos que ao final da aplicação foi atribuído para essas amostras as classes espécie 1, 2 e 3, respectivamente. Uma forma didática de apresentar esse resultado é a matriz de confusão.

A Tabela 6 representa a matriz de confusão para a classificação dos indivíduos A , B e C por meio do algoritmo k - NN . Ela é composta pela classe prevista e pela classe obtida, ou seja, aquela que foi atribuída pelo algoritmo. Quando a classe prevista e a atual são a mesma, a linha e a coluna dessa respectiva classe somam um. Portanto, os acertos estarão acumulados na diagonal principal da matriz de confusão. Em contrapartida, quando a classe prevista é diferente da obtida, a linha e coluna correspondente também irá somar um mas, ela não estará na diagonal principal da matriz de confusão.

Tabela 6 – Exemplo de matriz de confusão para a classificação dos indivíduos A , B e C por meio do algoritmo k - NN .

		Classificação obtida		
		Espécie 1	Espécie 2	Espécie 3
Classificação prevista	Espécie 1	1	0	0
	Espécie 2	0	0	0
	Espécie 3	0	1	1

Em nosso exemplo, era previsto que as classes fossem espécie 1, espécie 3 e espécie 3 porém, as classes obtidas foram espécie 1, 2 e 3. Em outras palavras, quando a classe prevista foi a espécie 1, o algoritmo apresentou como resposta a espécie 1, portanto, a posição *Espécie 1 x Espécie 1* somou 1. Quando a classe prevista era espécie 3, o algoritmo apresentou como resposta a espécie 3 e, nesse caso, a posição *Espécie 3 x Espécie 3* somou 1. E, para outra amostra o algoritmo indicou como resposta a espécie 2 o que resultou na soma de um na posição *Espécie 3 x Espécie 2*.

A acurácia (A_c) de uma matriz de confusão é determinada pela Equação 4.20 e pode ser interpretada como a razão entre a somatória da diagonal principal da matriz de

confusão, ou seja, a soma dos acertos, sobre o número de indivíduos testados.

$$Ac = \frac{\sum DiagonalPrincipal}{PopTotal} \quad (4.20)$$

Em nosso exemplo, a soma dos acertos é igual a 2 e a população total testada é igual a 3. Sendo assim, $Ac_{exemplo} = \frac{2}{3} = 0,667$.

Partindo dessa premissa, podemos inferir que, para esse exemplo ilustrativo, o k -NN teve uma acurácia de 66,7%.

Essa mesma abordagem foi aplicada para os dados obtidos nessa dissertação. Sendo que para a fase de treinamento foram usadas 240 amostras e para a etapa de validação 60. As categorias de atributos i) forma; ii) densidade; iii) regularidade estrutural; iv) multiescala; v) todos e vi) selecionados foram validadas em relação ao espaço original e após aplicação do *ACP* (no espaço 2 e 3D), seguindo a abordagem k -NN com $k = 3$. Os resultados da validação foram apresentados na forma de matriz de confusão.

4.6.2 Perceptron Multicamada (*PMC*)

Perceptron MultiCamadas (*PMC*) é uma rede neural artificial constituída por, pelo menos, uma camada intermediária, também conhecida como escondida, de neurônios, situada entre a camada de entrada e a respectiva camada neural de saída (SILVA; SPATTI; FLAUZINO, 2010). A Figura 37 ilustra um exemplo de um rede do tipo *PMC*. Ela é composta pela camada de entrada que é formada por x_n , duas camadas intermediárias sendo que a primeira é formada por n_1 neurônios, e a segunda por n_2 neurônios e a camada de saída composta por m neurônios que resultam nas y_m categorias de classificação. As variáveis x_n , n_1 , n_2 , m e y_m são números reais e positivos.

Outra característica importante desse tipo de rede é que ela pertence à arquitetura *feedforward*, cujo treinamento é realizado de forma supervisionada, ou seja, quando se conhece *a priori* as classes dos dados de entrada. Conforme observado na Figura 37, o fluxo de informações se inicia na camada de entrada, percorre em seguida as camadas intermediárias, sendo então finalizado na camada neural de saída (GONZALEZ; WOODS, 2000) e (SILVA; SPATTI; FLAUZINO, 2010).

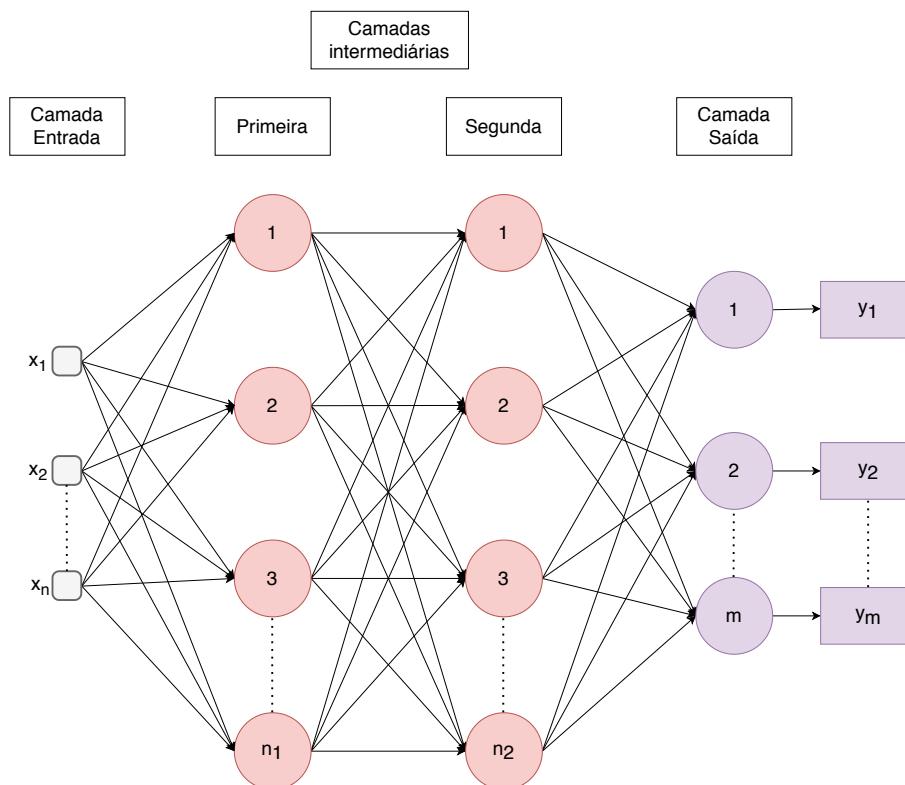


Figura 37 – Exemplo da topologia de uma *PMC* com duas camadas intermediárias. Topologia contendo x_n neurônios na camada de entrada, n_1 neurônios na primeira camada intermediária, n_1 neurônio na segunda camada intermediária e m neurônios na camada de saída.

Fonte: Autor

O processo de treinamento dessa rede é denominado *backpropagation* e consiste nas aplicações sucessivas de duas fases: “propagação adiante” (*forward*) e “propagação reversa” (*backward*) (SILVA; SPATTI; FLAUZINO, 2010). Na primeira fase os dados de entrada do conjunto de treinamento são inseridos nas entradas da rede e são propagados por todas as camadas intermediárias até a produção das respectivas saídas. Portanto, esta fase tem por objetivo obter as respostas da rede, levando-se em consideração apenas valores atuais de pesos sinápticos e limiares de seus neurônios, os quais permanecerão inalterados durante cada execução desta fase.

Em contrapartida, a fase *backward* tem por objetivo ajustar os pesos de todos os neurônios de uma rede. Isso é possível uma vez que a rede *PMC* é supervisionada, ou seja, sabemos qual é a saída desejada. Dessa maneira, o algoritmo de retropropagação do erro irá ajustar os pesos levando-se em consideração a diferença observada entre as respostas produzidas por suas saídas em relação aos seus respectivos valores desejados. Vale ressaltar que as saídas dos neurônios da primeira camada neural serão as próprias entradas daqueles neurônios pertencentes à segunda camada neural escondida e, conseqüentemente, o ajuste de peso e limiares irá ocorrer tanto nas camadas intermediárias quanto na camada de

saída (SILVA; SPATTI; FLAUZINO, 2010).

A camada de saída pode ser composta por diversos neurônios, sendo que cada um representa uma das saídas desejadas. Conforme apresentado na Figura 37 a camada de saída tem m neurônios, portanto, a rede *PMC* tem também y_m classes, ou seja, se assume que cada uma das classes é representada por um neurônio de saída. Esse método para a escolha da representação das classes é denominado *one of c-classes* e é o mais utilizado (SILVA; SPATTI; FLAUZINO, 2010).

Não entraremos em detalhes sobre as equações matemáticas que envolvem o algoritmo de retropropagação do erro, uma vez que, o foco do nosso trabalho não está na implementação das redes do tipo *PMC*. Contudo, para melhor entendimento dessa tão valiosa ferramenta, sugerimos a leitura do capítulo 5 do livro intitulado *Redes Neurais Artificiais para engenharia e ciências aplicadas* escrito por Silva, Spatti e Flauzino (2010). O mais importante neste trabalho é entendermos as principais variáveis como podemos ajusta-las com a finalidade de se obter um bom resultado para nossos dados.

Uma vez ajustados os pesos sinápticos, a rede estará treinada. Cada vez que as amostras do conjunto de treinamento é apresentada é denominada *época*. Ou seja, a rede do tipo *PMC* estará treinada quando a variação do erro entre duas épocas sucessivas for inferior à precisão (ε), que, normalmente é $\varepsilon = 10^{-6}$ (SILVA; SPATTI; FLAUZINO, 2010).

Existem duas variáveis que estão diretamente ligadas ao número de iterações que uma rede do tipo *PMC* irá precisar para estar treinada: a taxa de aprendizagem (η) e o *momentum* (α). A taxa de aprendizagem ilustra o quão rápido o processo de treinamento da rede estará sendo conduzido rumo à sua convergência, ou seja, até o treinamento estar finalizado (SILVA; SPATTI; FLAUZINO, 2010). De maneira análoga, o *momentum* irá exprimir a variação entre as matrizes de pesos de duas iterações subsequentes. De maneira geral, α indica o "tamanho do passo" em direção ao mínimo da função erro (SILVA; SPATTI; FLAUZINO, 2010). Segundo Rumelhart, Hinton e Williams (1986) os valores típicos para os parâmetros η e α para treinamento de redes *PMC* são: $0,05 \leq \eta \leq 0,75$ e $0 \leq \alpha \leq 0,9$ (SILVA; SPATTI; FLAUZINO, 2010).

Outro aspecto que está diretamente relacionado com o desempenho de uma rede *PMC* é a topologia, ou seja, como essa rede foi arquitetada. Achar as quantidades ideais de camadas intermediárias e de neurônios por camada para se estudar um determinado conjunto de dados é uma tarefa não trivial. Silva, Spatti e Flauzino (2010) relata em seu livro que a melhor maneira de se encontrar a topologia que mais adequada aos seus dados é testando as topologias candidatas. Esse cenário, a técnica estatística conhecida por validação cruzada (*cross-validation*), cujo objetivo é avaliar a aptidão de cada estrutural neural quando aplicadas a um conjunto de dados, ganha visibilidade e é fortemente empregada (SILVA; SPATTI; FLAUZINO, 2010), (GAMBÄCK; SIKDAR, 2017) e (SHENFIELD; DAY; AYESH, 2018).

A *cross-validation* possui três métodos de aplicação. Nesse trabalho, focaremos o método denominado *k*-partições (*k-fold cross-validation*). Este método baseia-se em dividir do conjunto total de amostras em *k*-partições, sendo que *k* – 1 delas irão compor o subconjunto de treinamento e a partição restante, o conjunto de teste. Em outras palavras, o subconjunto de treinamento será utilizado para treinar todas as topologias candidatas, sendo que o subconjunto de teste é somente aplicado para selecionar aquela apresentar os melhores resultados de generalização (SILVA; SPATTI; FLAUZINO, 2010). O valor da variável *k* está relacionado a quantidade de dados porém, usualmente $5 \leq k \leq 10$.

A escolha da melhor topologia é feita baseada na topologia que teve a maior acurácia. Porém, em caso de empate, normalmente, a arquitetura com menor custo computacional é escolhida.

Após a escolha da melhor topologia, a rede selecionada poderá ser utilizada para classificar novas amostras cuja classificação é desconhecida. Essa fase é denominada etapa de teste a qual é caracterizada pela entrada de um novo conjunto de dados (conjunto de teste) na rede visando a classificação. Essa etapa também pode ser utilizada para validar qual é a acurácia da rede frente a um novo conjunto de dados. Nesse caso, é necessário conhecer as classes originais desses novos dados.

Para fins didáticos, ilustraremos a utilização da rede do tipo *PMC* para classificar o *dataset* Iris (FISHER, 1936). Para isso, elencamos nove topologias com potencial de classificação desses dados. Essa escolha foi baseada em trabalhos anteriores de nosso grupo de pesquisa porém, ainda não publicados. A Figura 38 mostra a topologia geral da rede *PMC* que será utilizada. Nela podemos observar que a camada de entrada é composta por quatro neurônios x_1, x_2, x_3, x_4 sendo eles, respectivamente, relativos aos atributos comprimento das sépalas, largura das sépalas, comprimento das pétalas e largura das pétalas. As camadas intermediárias possuem n_1 e n_2 neurônios que irão variar de acordo com cada uma das nove topologias elencadas. Assim como a camada de entrada, a camada de saída é composta por um número fixo de neurônios que, nesse caso, é três, sendo eles relativos as classes *Espécie 1*, *Espécie 2* e *Espécie 3*.

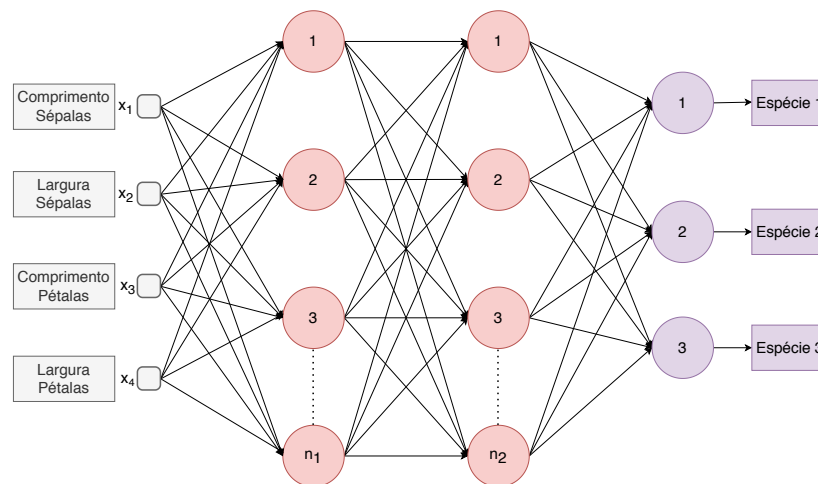


Figura 38 – Esquema de uma topologia geral para a classificação do *dataset* Iris. Topologia contendo 4 neurônios na camada de entrada, n_1 neurônios na primeira camada intermediária, n_1 neurônio na segunda camada intermediária e 3 neurônios na camada de saída.

Fonte: Autor

A Tabela 7 mostra as características de cada topologia elencada. Topologias onde $n_2 = 0$ apresentam apenas uma camada intermediária. Vale ressaltar que para todas as topologias foram utilizados $\eta = 0,3$ e $\alpha = 0,2$. A variável a diz respeito ao número de neurônio da primeira camada intermediária e ela é calculada como sendo: $a = \frac{\text{atributos} + \text{classes}}{2}$. Esse índice a é utilizado como *default* no *software Weka*. Em nosso exemplo, temos 4 atributos (comprimento sépalas, largura sépalas, comprimento pétalas, largura pétalas) e 3 classes (espécie 1, espécie 2 e espécie 3). Portanto, $a = \frac{3+4}{2} = \frac{7}{2} = 3,5$. Sendo assim, $n_1 = 3$.

Tabela 7 – Topologias testadas para a classificação do *dataset* Iris

Topologia	n_1	n_2	Época
1	a	0	500
2	a	0	1000
3	a	0	2000
4	10	20	500
5	10	20	1000
6	10	20	2000
7	50	30	500
8	50	30	1000
9	50	30	2000

As redes do tipo *PMC* foram construídas, treinadas e testadas no *software Weka* 3.8.4. Para a fase de treinamento, foram usados 80% dos dados e para a fase de teste 20%.

Com o objetivo de avaliar qual das nove topologias possui maior rendimento frente ao *dataset* Iris, utilizamos o método *k-fold cross-validation* com $k = 10$. Os resultados foram apresentados por meio de matriz de confusão conforme ilustra a Figura 39. Nela podemos ver que as topologias 1, 2, 4, 5, 6 e 7 apresentaram uma acurácia de 95,83%. Como várias topologias obtiveram o mesmo desempenho, escolhemos a com menor número de neurônios, camadas e épocas, ou seja, a rede escolhida foi a de Topologia 1.

Topologia 1 (95,83%)	Atual			Topologia 2 (95,83%)	Atual			Topologia 3 (95%)	Atual					
	Espécie 1	Espécie 2	Espécie 3		Espécie 1	Espécie 2	Espécie 3		Espécie 1	Espécie 2	Espécie 3			
	Espécie 1	40	0		0	Espécie 1	40		0	0	Espécie 1	40	0	0
	Espécie 2	0	37		3	Espécie 2	0		37	3	Espécie 2	0	37	3
Prevista	Espécie 3	0	2	38	Prevista	Espécie 3	0	2	38	Prevista	Espécie 3	0	3	37
Topologia 4 (95,83%)	Atual			Topologia 5 (95,83%)	Atual			Topologia 6 (95,83%)	Atual					
	Espécie 1	Espécie 2	Espécie 3		Espécie 1	Espécie 2	Espécie 3		Espécie 1	Espécie 2	Espécie 3			
	Espécie 1	40	0		0	Espécie 1	40		0	0	Espécie 1	40	0	0
	Espécie 2	0	36		4	Espécie 2	0		36	4	Espécie 2	0	36	4
Prevista	Espécie 3	0	1	39	Prevista	Espécie 3	0	1	39	Prevista	Espécie 3	0	1	39
Topologia 7 (95,83%)	Atual			Topologia 8 (95%)	Atual			Topologia 9 (95%)	Atual					
	Espécie 1	Espécie 2	Espécie 3		Espécie 1	Espécie 2	Espécie 3		Espécie 1	Espécie 2	Espécie 3			
	Espécie 1	40	0		0	Espécie 1	40		0	0	Espécie 1	40	0	0
	Espécie 2	0	37		3	Espécie 2	0		36	4	Espécie 2	0	36	4
Prevista	Espécie 3	0	2	38	Prevista	Espécie 3	0	2	38	Prevista	Espécie 3	0	2	38

Figura 39 – Matrizes de confusão após aplicação do método *k-fold cross-validation* com $k = 10$ para escolha da melhor topologia. As topologias 1, 2, 4, 5, 6 e 7 apresentaram uma acurácia de 95,83%. Enquanto que as topologias 3, 8 e 9 obtiveram acurácia de 95%.

Fonte: Autor

Com a topologia já escolhida e a rede treinada, podemos classificar um novo conjunto de dados. Essa etapa é denominada "Etapa de teste" e pode ser feita de duas maneiras: com um novo conjunto de dados cujo objetivo é classificar esses indivíduos; ou validando a acurácia de rede frente a um conjunto de dados cujas as classes foram definidas *a priori*.

Para nosso exemplo, assim como foi utilizado ao longo desse trabalho, focaremos na validação da acurácia da rede frente a um conjunto de dados com classes já definidas. Para isso, um subconjunto com 20% dos dados originais foi utilizado, portanto, 30 indivíduos. Eles foram aplicados sobre a rede *PMC* constituída por 4 neurônios na camada de entrada, 3 na camada intermediária e 3 na camada de saída. Com as variáveis $epoca = 500$, $\eta = 0,3$ e $\alpha = 0,2$.

O resultado dessa validação está na matriz de confusão ilustrada na Figura 40. Nela podemos ver que os 10 indivíduos de cada classe foram classificados de maneira correta, ou seja, a acurácia dessa rede foi de 100%.

Topologia 1 (100%)		Atual		
		Espécie 1	Espécie 2	Espécie 3
Espécie 1	10	0	0	
Espécie 2	0	10	0	
Prevista	Espécie 3	0	0	10

Figura 40 – Matriz de confusão da etapa de teste para validação da acurácia da rede *PMC*. A topologia utilizada foi de *PMC* constituída por 4 neurônios na camada de entrada, 3 na camada intermediária e 3 na camada de saída. Com as variáveis $epoca = 500$, $\eta = 0,3$ e $\alpha = 0,2$. A acurácia obtida frente a um novo subconjunto de dados foi de 100%.

Fonte: Autor

Essa mesma abordagem foi aplicada para os dados obtidos nessa dissertação. Sendo que para a fase de treinamento foram usadas 240 amostras e para a etapa de validação 60. As categorias de atributos i) Forma. ii) Densidade; iii) Regularidade estrutural; iv) Multiescala; v) Todos e vi) Seleccionados foram validadas em relação ao espaço original e após aplicação do *ACP* (no espaço 2 e 3D), seguindo a abordagem *PMC* com a escolha entre as nove topologias (descritas na Tabela 7) feita por meio do método *k-fold cross-validation* com $k = 10$ e a validação feita com base nas 60 amostras das quais já se conhecia as classes. Os resultados da validação foram apresentados na forma de matriz de confusão.

5 Resultados

Nesse Capítulo iremos apresentar os resultados obtidos ao longo de etapa de análise e da etapa de validação. A Figura 41 ilustra quais abordagens foram aplicadas em qual etapa e, portanto, como os resultados serão apresentados.

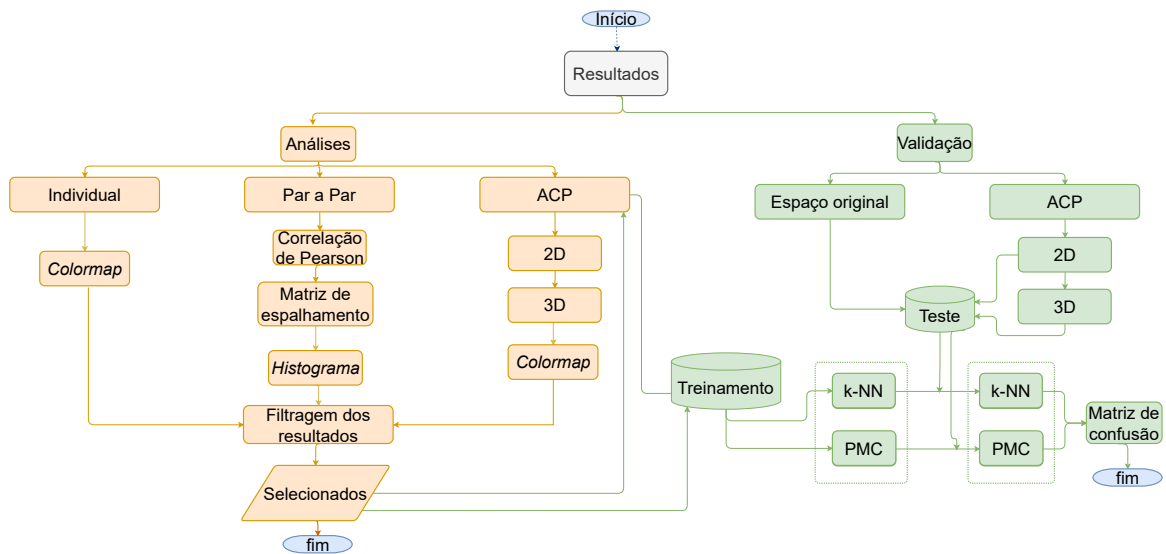


Figura 41 – Fluxograma indicando a organização dos resultados.

Fonte: Autor

Para fins didáticos, inserimos ícones em cada subseção para a melhor visualização do leitor sobre a etapa do trabalho que está sendo apresentada.

5.1 Etapa de análise

5.1.1 Análise individual

Nessa seção, iremos apresentar os resultados obtidos por meio da análise individual. A Figura 42 ilustra as divisões das classes teciduais estudadas nesse etapa: Todas as classes; feixes concêntricos (floema e xilema); feixes colaterais (floema e xilema); e feixes bicolaterais.

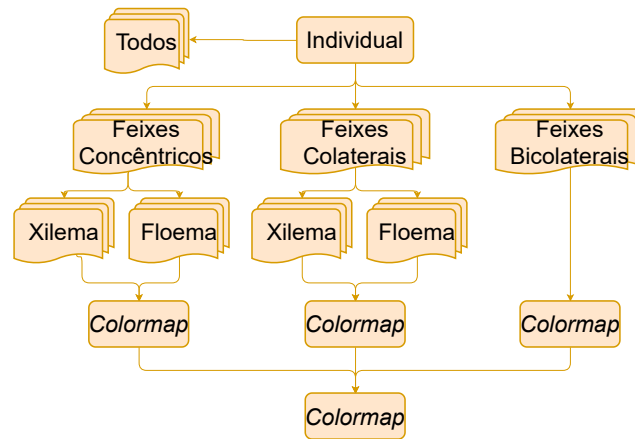
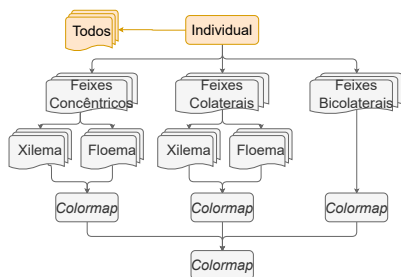


Figura 42 – Fluxograma com as divisões de classes e as etapas da análise individual.

Fonte: Autor

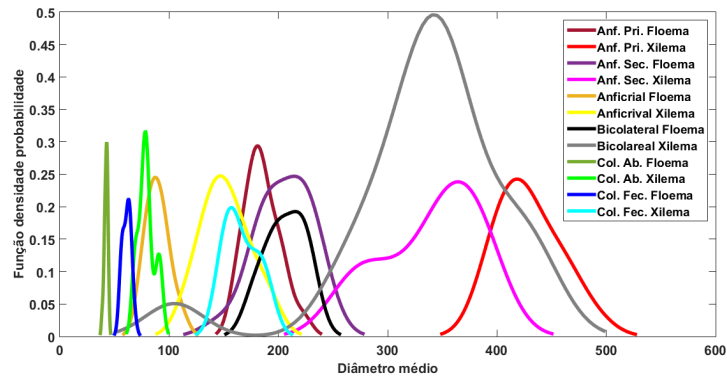
Os atributos que foram mais relevantes para a segregação dos tecidos vasculares de cada classe foram apresentados em *colormaps* específicos. Ao final, apresentamos um *colormap* geral no qual foi apresentado os atributos mais recorrentes durante a aplicação da abordagem individual.

5.1.1.1 Todos os feixes vasculares

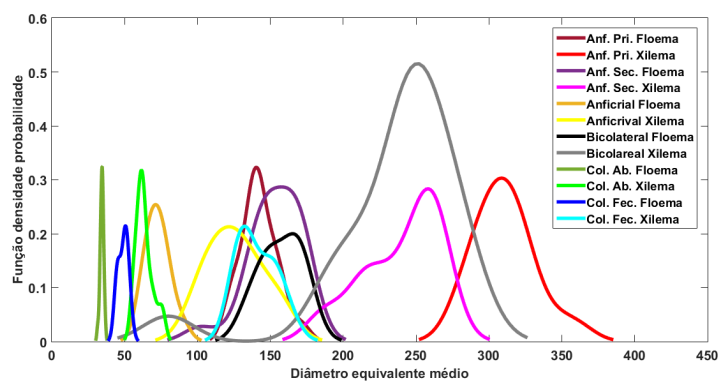


Para essa etapa, foi plotada a função de densidade probabilidade de todos os atributos. As funções foram feitas considerando todos os feixes vasculares estudados: anfigasal primário, anfigasal secundário, anfigicrival, bicolateral, colateral aberto e colateral fechado. Os feixes foram subdivididos em seus tecidos constituintes, ou seja, floema e xilema.

A Figura 43 ilustra quais foram os atributos capazes de segregar pelo menos um dos tecidos vasculares vegetais estudados e, portanto, relevantes nesse tipo de análise.



(a) Diâmetro médio



(b) Diâmetro equivalente médio

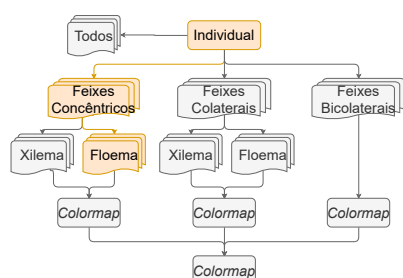
Figura 43 – Função densidade probabilidade para os atributos relacionados a forma capazes de segregar pelo menos um dos tecidos vasculares estudados. (a) Diâmetro médio; (b) Diâmetro equivalente médio.

Fonte: Autor

Dessa maneira, os únicos atributos capazes de segregar pelo menos um dos tecidos vasculares estudados foram da categoria forma, sendo eles: diâmetro médio (Figura 43a) diâmetro equivalente médio (Figura 76d). Ambos os atributos foram capazes de segregar as classes teciduais colateral aberto floema e colateral fechado floema.

5.1.1.2 Feixes vasculares concêntricos

5.1.1.2.1 Floema nos feixes vasculares concêntricos



Para essa etapa, foi plotada a função de densidade probabilidade de todos os atributos. As funções foram feitas considerando os seguintes tecidos floemáticos: Anfivascular Primário, Anfivascular Secundário e Anficrival.

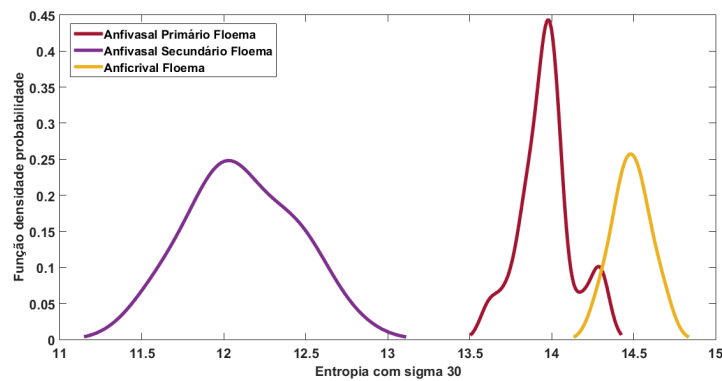
A Tabela 8 ilustra quais atributos, relacionados a forma, a densidade e a regularidade respectivamente, foram capazes de segregar pelo menos uma classe de tecidos estudadas e, portanto, relevantes nesse tipo de análise. Os atributos marcados com asterisco também fazem parte da categoria multiescala.

Tabela 8 – Atributos que foram relevantes para segregar tecidos do floema em feixes vasculares concêntricos.

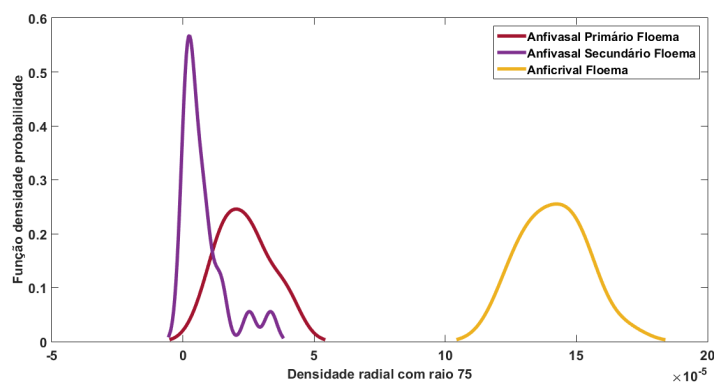
	Atributos relevantes
Medidas forma	Área média
	Resíduo médio
	Diâmetro médio
	Diâmetro equivalente médio
Medidas densidade	Uniformidade espacial 30 *
	Uniformidade espacial 48 *
	Uniformidade espacial 65 *
	Circularidade <i>isoline</i> 48 *
	Circularidade <i>isoline</i> 65 *
	Densidade radial 75 *
Medidas regularidade estrutural	Poligonalidade média

No total, foram selecionados 11 atributos sendo 4 relativos a forma: Área média, Resíduo médio, Diâmetro médio e Diâmetro equivalente médio; 7 relativos a densidade: Uniformidade espacial com sigma 30, Uniformidade espacial com sigma 48, Uniformidade espacial com sigma 65, Circularidade *isoline* 48, Circularidade *isoline* 65 e Densidade radial com raio 75; e 1 relacionada com a regularidade estrutural: Poligonalidade média. Todos esses gráficos estão ilustrados no Apêndice A.1.1 desta dissertação.

De maneira geral, podemos inferir que dos 11 atributos listados, apenas dois foram capazes de segregar mais de um tecido. Os gráficos desses dois atributos estão ilustrados na Figura 44. Nesse sentido, a Figura 44a destaca-se por segregar totalmente o tecido anfigival secundário e possuir uma pequena área de sobreposição entre os tecidos anfigival primário e anfigival, bem como a Figura 44b que apresenta segregação completa do tecido anfigival e pequena área de sobreposição entre os anfigivais primários e secundário.



(a) Uniformidade espacial com sigma 30

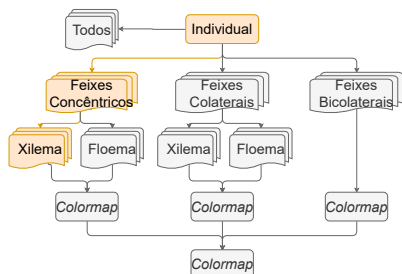


(b) Densidade radial com raio 75

Figura 44 – Função densidade probabilidade para os atributos capazes de segregar pelo menos um dos tecidos floemáticos dos feixes vasculares concêntricos. (a) Uniformidade espacial com sigma 30; (b) Densidade radial com raio 75.

Fonte: Autor

5.1.1.2.2 Xilema nos feixes vasculares concêntricos



Para essa etapa, foi plotada a função de densidade probabilidade de todos os atributos. As funções foram feitas considerando os seguintes tecidos xilemáticos: Anfivasal Primário, Anfivasal Secundário e Anficrival.

A Tabela 9 ilustra quais atributos, relacionados a forma, a densidade e a regularidade respectivamente, foram capazes de segregar pelo menos uma classe dos tecidos estudados e, portanto, relevantes nesse tipo de análise. Os atributos marcados com asterisco também fazem parte da categoria multiescala.

Tabela 9 – Atributos que foram relevantes para segregar tecidos do xilema em feixes vasculares concêntricos.

	Atributos relevantes
Medidas forma	Área média
	Desvio Padrão área
	Resíduo médio
	Desvio padrão resíduo
	Diâmetro médio
	Diâmetro equivalente médio
Medidas densidade	Uniformidade espacial 30 *
	Uniformidade espacial 48 *
	Uniformidade espacial 65 *
	Uniformidade espacial 83 *
	Uniformidade espacial 100 *
	Densidade radial 75 *
	Densidade radial 125 *
Medidas regularidade estrutural	Poligonalidade média
	Lacunaridade 1 *
	Lacunaridade 6 *

Ao total, foram selecionados 16 atributos sendo 6 relativos a forma: Área média, Desvio padrão da área, Resíduo médio, Desvio padrão do resíduo, Diâmetro médio e Diâmetro equivalente médio; 7 relativos a densidade: Uniformidade espacial com sigma 30, Uniformidade espacial com sigma 48, Uniformidade espacial com sigma 65, Uniformidade espacial com sigma 83, Uniformidade espacial com sigma 100, Densidade radial com raio 75 e Densidade radial com raio 125; e 3 relacionados com a regularidade estrutural: Poligonalidade média, Lacunaridade com raio 1 e Lacunaridade com raio 6. Todos esses gráficos estão ilustrados no Apêndice A.1.2 desta dissertação.

De maneira geral, os atributos que apresentaram separação para mais de um tecido são: Área média, Diâmetro equivalente médio e a Uniformidade espacial com sigma 30, 48 e 65. Todos eles estão apresentados na Figura 45. Nesse sentido, as Figuras 45a e 45b destacam-se pois os tecidos apresentam boa distribuição e pouca área de sobreposição. Os atributos Uniformidade espacial com sigma 30, 48 e 65 (respectivamente Figuras 45c, 45d e 45e) apresentam o tecido anficrival totalmente segregado e pouca sobreposição entre os tecidos anfigvais primário e secundário.

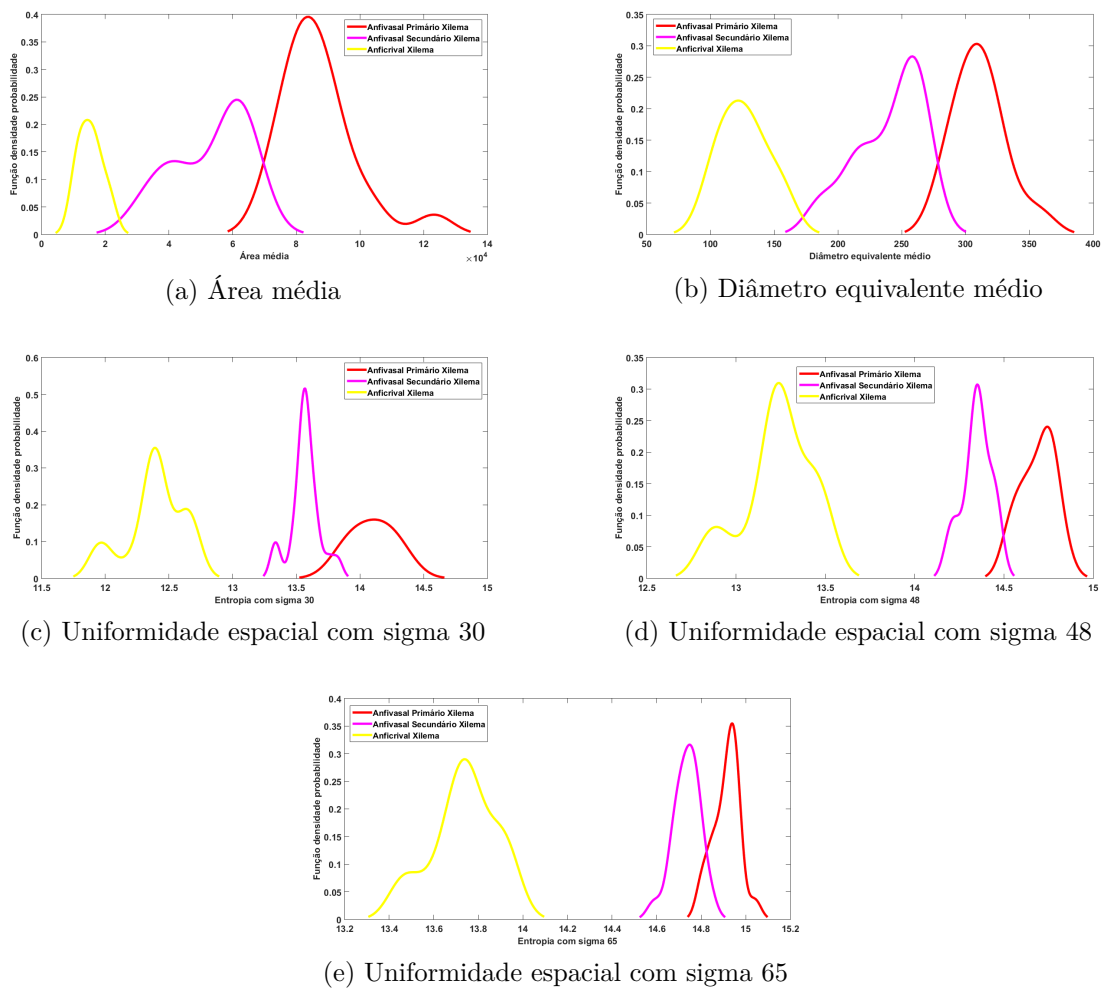
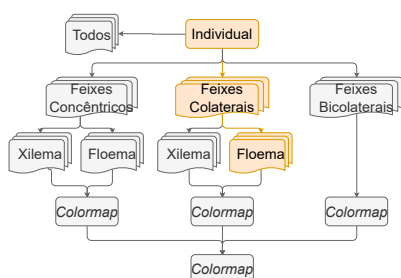


Figura 45 – Função densidade probabilidade para os atributos capazes de segregar pelo menos um dos tecidos xilemáticos dos feixes vasculares concêntricos. (a) Área média; (b) Diâmetro equivalente médio; (c) Uniformidade espacial com sigma 30; (d) Uniformidade espacial com sigma 48 e (e) Uniformidade espacial com sigma 65.

Fonte: Autor

5.1.1.3 Feixes vasculares colaterais

5.1.1.3.1 Floema nos feixes vasculares colaterais



Para essa etapa, foi plotada a função de densidade probabilidade de todos os atributos. As funções foram feitas considerando os seguintes tecidos floemáticos: Colateral aberto e Colateral fechado.

A Tabela 10 ilustra quais atributos, relacionados a forma, a densidade e a regularidade respectivamente, foram capazes de segregar os tecidos estudados. Os atributos marcados com asterisco também fazem parte da categoria multiescala.

Tabela 10 – Atributos que foram relevantes para segregar tecidos do floema em feixes vasculares colaterais.

	Atributos relevantes
Medidas forma	Área média
	Resíduo médio
	Desvio padrão resíduo
	Diâmetro médio
	Diâmetro equivalente médio
Medidas densidade	Densidade radial 75 *
	Densidade radial 275 *
Medidas regularidade estrutural	Lacunaridade 6 *

No total, foram selecionados 8 atributos sendo 5 relativos a forma: Área média, Resíduo médio, Desvio padrão do resíduo, Diâmetro médio e Diâmetro equivalente médio; 2 relativos a densidade: Densidade radial com raio 75 e Densidade radial com raio 275; e 1 relacionada com a regularidade estrutural: Lacunaridade com raio 6. Todos esses gráficos estão ilustrados no Apêndice A.2.1 desta dissertação.

De maneira geral, podemos inferir que dos 8 atributos apresentados, apenas os atributos Área média, Figura 46a, Resíduo médio, Figura 46b, Diâmetro médio, Figura 46c e Diâmetro equivalente médio, Figura 46d, foram capazes de segregar os dois tecidos sem nenhuma área de sobreposição entre eles.

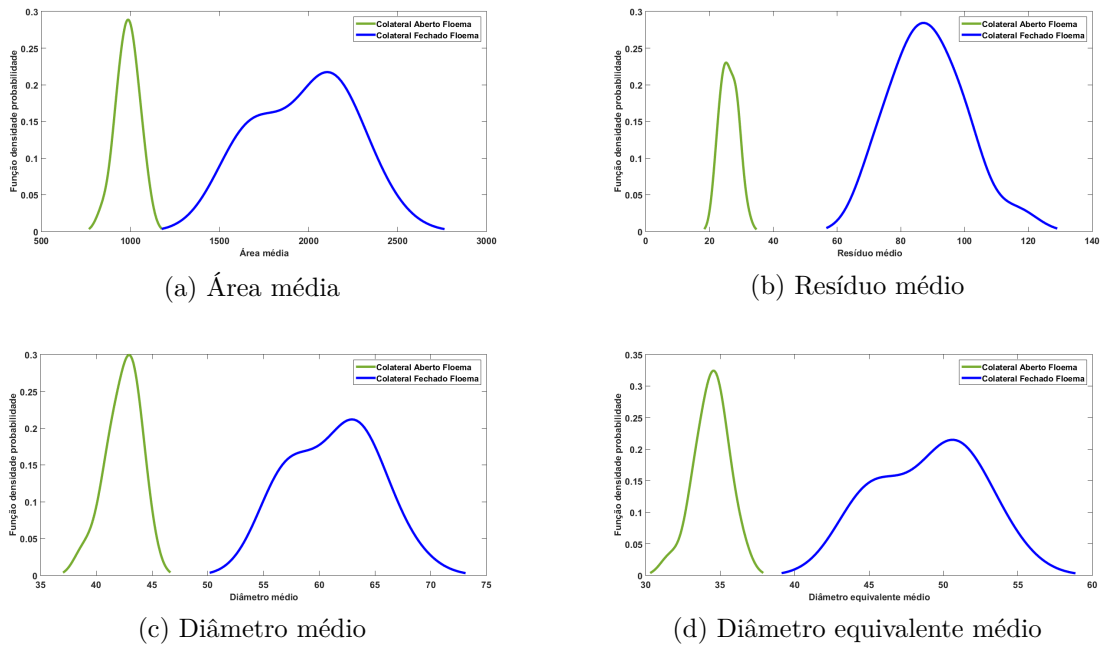
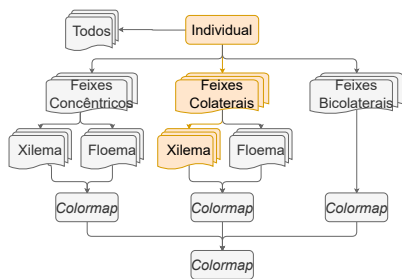


Figura 46 – Função densidade probabilidade para os atributos capazes de segregar pelo menos um dos tecidos floemáticos dos feixes vasculares colaterais. (a) Área média; (b) Resíduo médio; (c) Diâmetro médio; (d) Diâmetro equivalente médio.

Fonte: Autor

5.1.1.3.2 Xilema nos feixes vasculares colaterais



Para essa etapa, foi plotada a função de densidade probabilidade de todos os atributos. As funções foram feitas considerando os seguintes tecidos xilemáticos: Colateral aberto e Colateral fechado.

A Tabela 11 ilustra quais atributos, relacionados a forma, a densidade e a regularidade respectivamente, foram capazes de segregar os tecidos estudados. Os atributos marcados com asterisco também fazem parte da categoria multiescala.

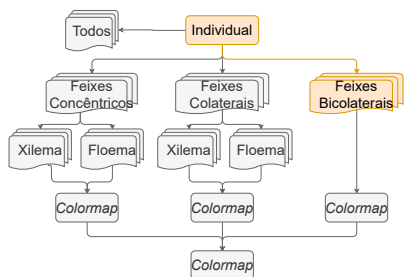
Tabela 11 – Atributos que foram relevantes para segregar tecidos do xilema em feixes vasculares colaterais.

	Atributos relevantes
Medidas forma	Área média
	Desvio Padrão área
	Resíduo médio
	Desvio padrão resíduo
	Diâmetro médio
	Desvio padrão diâmetro
	Diâmetro equivalente médio
	Desvio padrão diâmetro equivalente
Medidas densidade	Uniformidade espacial 30 *
	Uniformidade espacial 48 *
	Uniformidade espacial 65 *
	Uniformidade espacial 83 *
	Uniformidade espacial 100 *
Medidas regularidade estrutural	Poligonalidade média
	Lacunaridade 1 *
	Lacunaridade 6 *
	Lacunaridade 11 *
	Lacunaridade 16 *
	Lacunaridade 21 *

Ao total, foram selecionados 19 atributos sendo 8 relativos a forma: Área média, Desvio padrão da área, Resíduo médio, Desvio padrão do resíduo, Diâmetro médio, Desvio padrão do diâmetro, Diâmetro equivalente médio, e Desvio Padrão do diâmetro equivalente; 5 relativos a densidade: Uniformidade espacial com sigma 30, Uniformidade espacial com sigma 48, Uniformidade espacial com sigma 65, Uniformidade espacial com sigma 83 e Uniformidade espacial com sigma 100; e 6 relacionados com a regularidade estrutural: Poligonalidade média, Lacunaridade com raio 1, Lacunaridade com raio 6, Lacunaridade com raio 11, Lacunaridade com raio 16 e Lacunaridade com raio 21. Todos esses gráficos estão ilustrados no Apêndice A.2.2 desta dissertação.

Dos atributos selecionados, apenas os atributos Lacunaridade com raio 1, Lacunaridade com raio 6 e Lacunaridade com raio 11 possuíram uma pequena área de sobreposição entre os tecidos analisados. Os demais separaram totalmente os tecidos ou seja, sem nenhuma área de sobreposição entre eles. Como 16 atributos apresentam destaque, apresentar todos os eles aqui é inviável. Por isso, todos os gráficos podem ser visualizados no Apêndice A.2.2 desta dissertação.

5.1.1.4 Feixe vascular bicolateral



O feixe bicolateral é composto por xilema e floema, como os outros feixes. Porém, ele não apresenta subtipos, como é o caso dos feixes concêntricos e colaterais. Devido a isso, os resultados referentes ao xilema e ao floema para a análise individual dos feixes bicolaterais serão apresentados juntos. Para essa etapa, foi plotada a função de densidade probabilidade de todos os atributos. As funções foram feitas considerando os tecidos que compõem os feixes bicolaterais: Floema e Xilema.

A Tabela 12 ilustra quais atributos, relacionados a forma, a densidade e a regularidade respectivamente, foram capazes de segregar os tecidos estudados. Os atributos marcados com asterisco também fazem parte da categoria multiescala.

Tabela 12 – Atributos que foram relevantes para segregar tecidos dos feixes vasculares bicolaterais.

	Atributos relevantes
Medidas forma	Desvio Padrão área
	Resíduo médio
	Desvio padrão resíduo
	Desvio padrão diâmetro equivalente
Medidas densidade	Uniformidade espacial 30 *
Medidas regularidade estrutural	Lacunaridade 16 *
	Lacunaridade 21 *

Ao total, foram selecionados 8 atributos sendo 5 relativos a forma: Desvio padrão da área, Resíduo médio, Desvio padrão do resíduo, Desvio padrão do diâmetro e Desvio Padrão do diâmetro equivalente; apenas 1 relativos a densidade: Uniformidade espacial com sigma 30; e 2 relacionados com a regularidade estrutural: Lacunaridade com raio 16 e Lacunaridade com raio 21.

De maneira geral, os atributos que separaram totalmente os tecidos, ou seja, sem nenhuma área de sobreposição estão apresentados na Figura 47 e são eles: Desvio Padrão do resíduo (Figura 47a), Desvio padrão do diâmetro (Figura 47b) e Desvio padrão do diâmetro equivalente 47c.

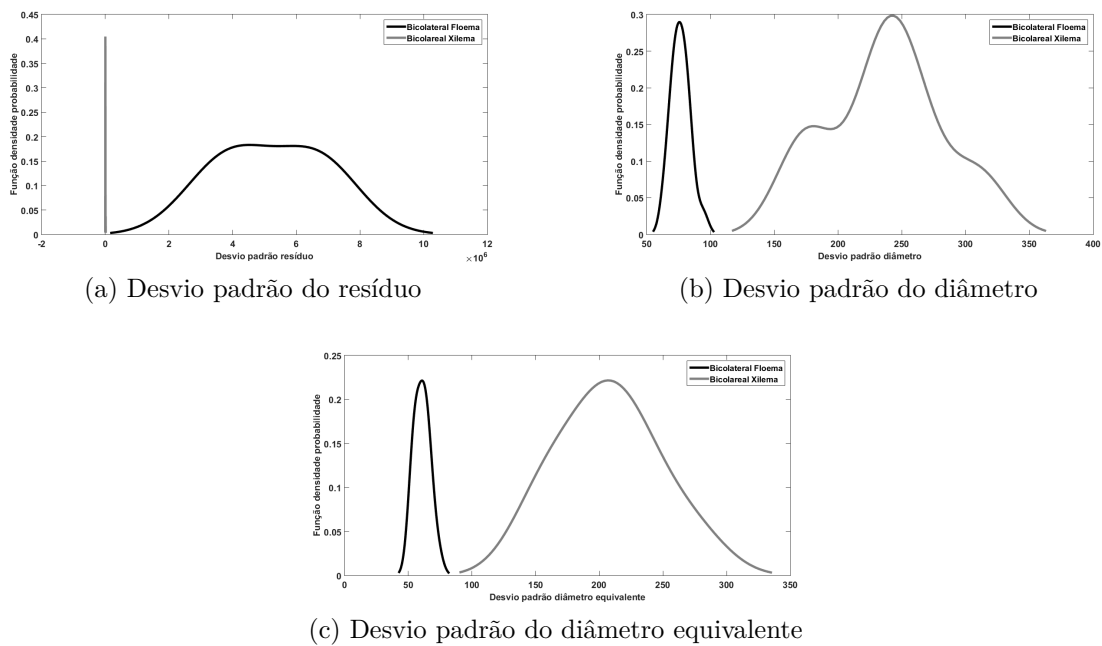
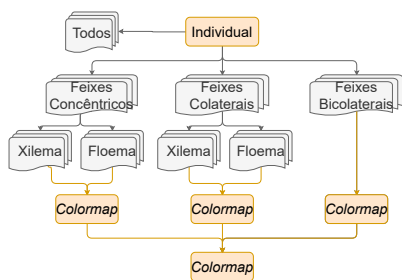


Figura 47 – Função densidade probabilidade para os atributos capazes de segregar pelo menos um dos tecidos floemáticos dos feixes vasculares colaterais. (a) Desvio padrão do resíduo; (b) Desvio padrão do diâmetro; (c) Desvio padrão do diâmetro equivalente.

Fonte: Autor

5.1.1.5 Considerações gerais



Tendo como premissa os resultados apresentados para a análise individual dos feixes vasculares vegetais podemos buscar entender quais medidas foram relevantes ao longo dessa análise. Para isso, os resultados foram concatenados em um *colormap*, Figura 48, no qual a escala de cor varia do vermelho (menos importante) para o verde (mais importante).

A importância de cada atributo foi estabelecida com base na recorrência desse atributo nos resultados apresentados. Essa abordagem foi feita para cada tipo de feixe levando em consideração os dois de tecidos vasculares que compõem esses feixes, o floema e o xilema. Ao final, os resultados foram agrupados em uma escala na tentativa de elencar quais medidas foram importantes na análise individual.

Ao analisar a Figura 48 podemos observar que a distribuição das medidas ao longo do tipo de feixe vascular é diferente. Os feixes colaterais foram melhores caracterizados pelos atributos referente ao formato celular e a regularidade estrutural. Os feixes concêntricos foram melhores representados pelos atributos referente a forma e a densidade. Em

contrapartida, os feixes bicolaterais apresentaram como atributos mais recorrentes os que fazem parte do grupo referente ao formato celular e a regularidade estrutural.

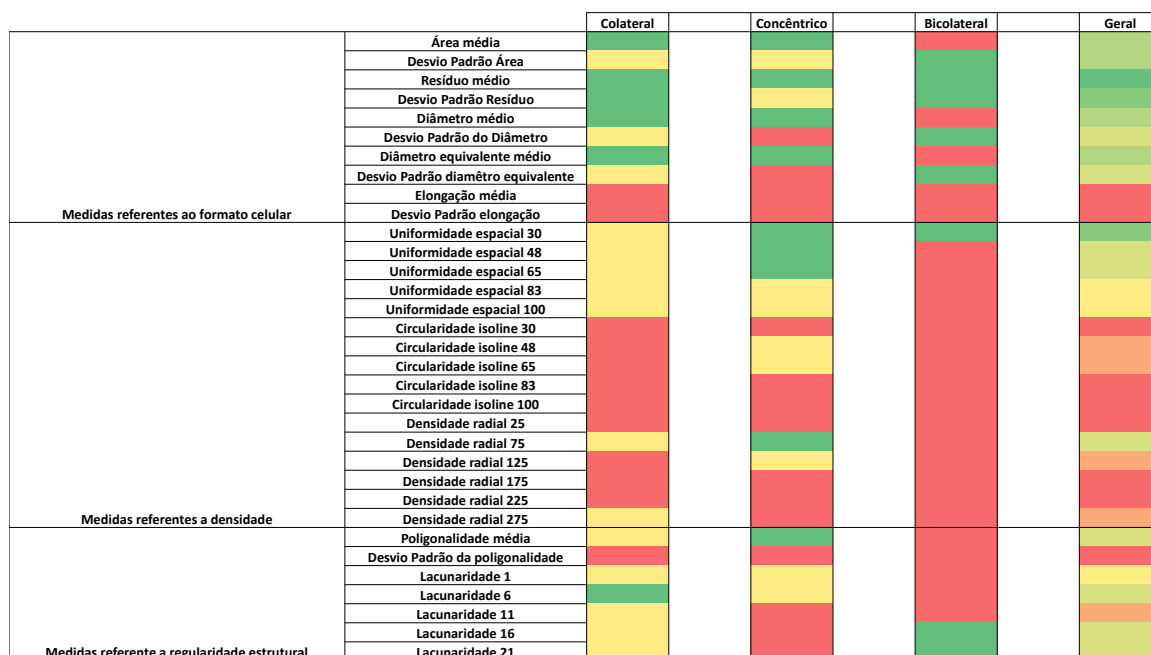


Figura 48 – *Colormap* representando a importância de cada atributo. A escala varia do vermelho (menos recorrente, logo menos importante) ao verde (mais recorrente, logo mais importante).

Fonte: Autor

De maneira geral, podemos notar que a distribuição geral dos atributos ao longo da análise individual engloba as quatro categorias de medidas estudadas (formato celular, densidade, regularidade estrutural e multiescala). Assim, podemos concluir que os atributos que foram importantes para caracterizar e segregar os feixes vasculares vegetais considerando apenas a análise individual são: Área média, Desvio padrão da área, Resíduo médio, Desvio padrão do resíduo, Diâmetro médio, Desvio padrão do diâmetro, Diâmetro equivalente médio, Desvio padrão do diâmetro equivalente, Uniformidade espacial com os sigmas 30, 48 e 65, Densidade radial com raio 75, poligonalidade média e Lacunaridade com raio 6, 16 e 21.

Outro ponto relevante que podemos sugerir com base nos resultados da análise individual tem relação com a multiescala dos atributos. Alguns atributos, como por exemplo, a Uniformidade espacial e a Densidade radial apresentaram recorrência em escalas específicas. Isso indica que o uso da análise multiescala para extração dos atributos é importante e pode contribuir na caracterização e na segregação de feixes vasculares vegetais.

5.1.2 Análise Par a Par

Nessa seção, iremos apresentar os resultados obtidos por meio da análise par a par. A Figura 49 ilustra as classes teciduais estudadas e as etapas que envolveram a análise Par a Par. As seguintes classes teciduais foram analisadas: Todas as classes; Feixes concêntricos (floema e xilema); Feixes colaterais (floema e xilema); e feixes bicolaterais.

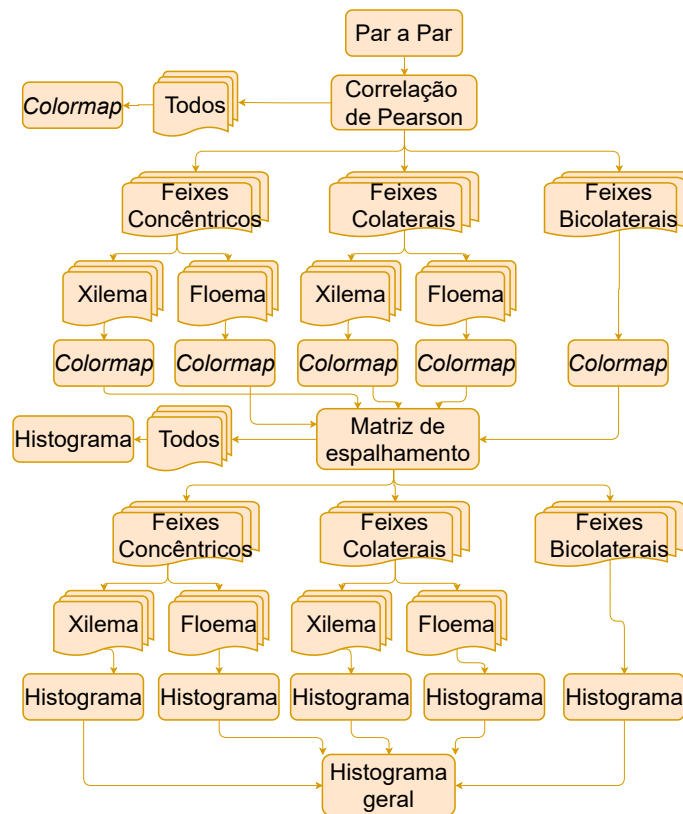


Figura 49 – Fluxograma com as divisões das classes estudadas e as etapas da análise Par a Par.

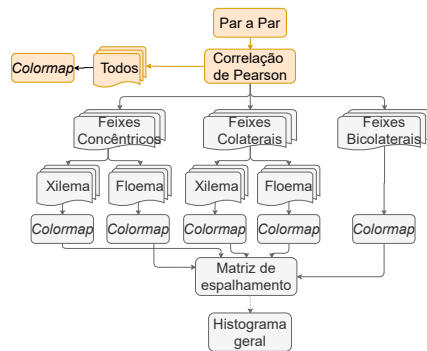
Vale ressaltar que dentro da abordagem par a par, os atributos foram submetidos a duas análises: correlação de Pearson e Matriz de espalhamento.

Ao final de ambas análises, os atributos que foram mais relevantes para cada classe foram apresentados em um gráfico de barras geral no qual elencamos quais atributos foram recorrentes durante a aplicação da abordagem par a par para as classes teciduais estudadas.

5.1.2.1 Correlação de Pearson

Os ícones de cada subseção foram feitos com base na Figura 49, porém omitimos a etapa correspondente à Matriz de espalhamento.

5.1.2.1.1 Todos os feixes vasculares

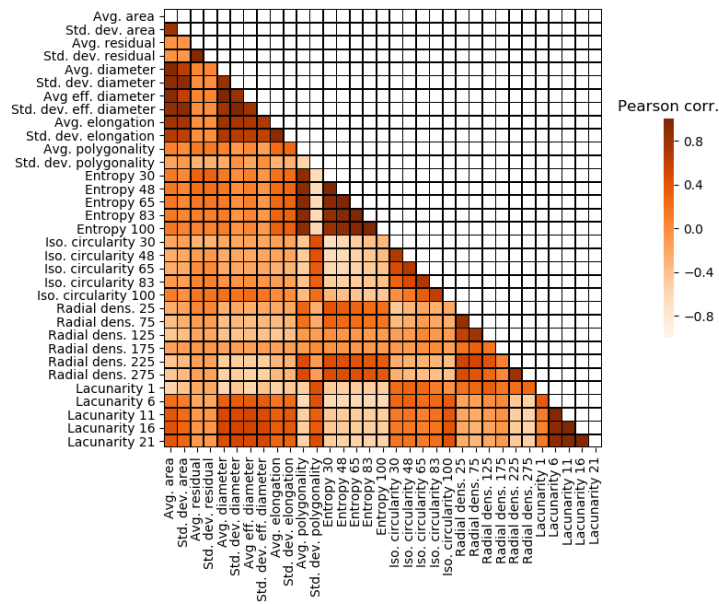


Para essa etapa, foi calculada a correlação de Pearson de todos os atributos tomados em pares. Os cálculos de p foram feitos considerando todos os feixes vasculares estudados: anfigival primário, anfigival secundário, anfigival, bicolateral, colateral aberto e colateral fechado. Os feixes foram subdivididos em seus tecidos constituintes, ou seja, floema e xilema.

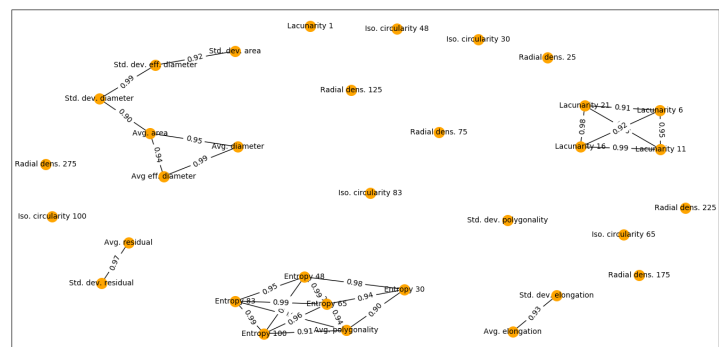
A Figura 50 ilustra os mapas com os valores de p referente ao grupo composto por todos os feixes vasculares estudados nesse trabalho.

A Figura 50a mostra, por meio de um *colormap*, os valores obtidos para a variável p em todos os pares de atributos. Nesse caso, os quadrados com cores mais escuras (laranja escuro) indicam $p \geq 0,9$ e com cores mais claras (transparentes) indicam $p \leq -0,9$. Já a Figura 50b liga somente os atributos que apresentaram $|p| \geq 0,9$. Nesse caso, nota-se correlação $|p| \geq 0,9$ entre os atributos de lacunaridade com exceção do raio 1 e entre os atributos de uniformidade espacial. O atributo poligonalidade média possui $|p| \geq 0,9$ com todos os atributos de uniformidade espacial. Os atributos alongação média e desvio padrão da alongação também estão conectados, bem como os atributos resíduo médio e desvio padrão do resíduo. O atributo desvio padrão da área possui $|p| \geq 0,9$ com desvio padrão do diâmetro equivalente que, por sua vez, possui $|p| \geq 0,9$ com o desvio padrão do diâmetro. O atributo área média tem correlação $|p| \geq 0,9$ em relação aos atributos diâmetro equivalente médio e diâmetro médio que, por sua vez, possuem correlação de $|p| = 0,99$ entre eles.

Os pares de atributos que apresentaram $|p| \geq 0,9$ foram filtrados e não foram selecionados para a próxima etapa do trabalho.



(a) Diagrama da correlação de Pearson

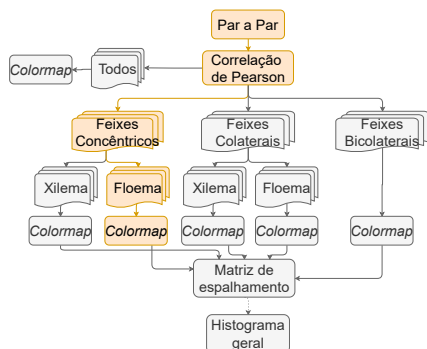


(b) Rede de correlação de Pearson

Figura 50 – Mapas da correlação de Pearson referente a todos as classes teciduais estudadas. (a) Diagrama de Pearson: *Colormap* indica os valores da correlação de Pearson; (b) Rede de correlação de Pearson: Atributos conectados indicam que eles possuem, em módulo, valores de $p \geq 0,9$.

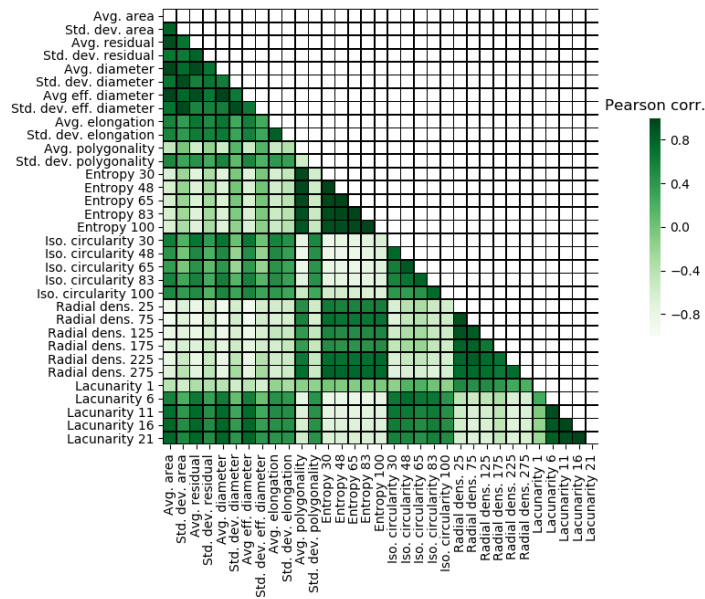
Fonte: Autor

5.1.2.1.2 Feixes vasculares do tipo concêntrico: Floema

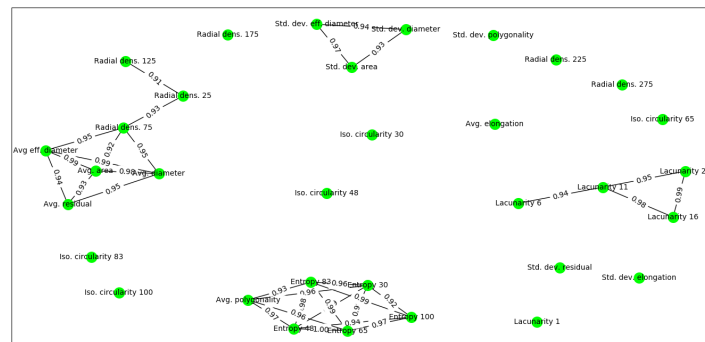


Os cálculos da correlação de Pearson foram feitos considerando os seguintes tecidos floemáticos: Anfivascular Primário, Anfivascular Secundário e Anfivascular.

A Figura 51 ilustra os mapas com os valores de p referente ao floema dos feixes do tipo concêntrico.



(a) Diagrama da correlação de Pearson



(b) Rede de correlação de Pearson

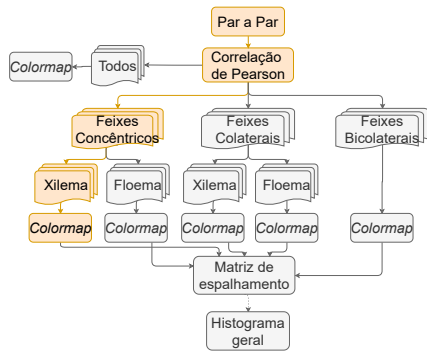
Figura 51 – Mapas da correlação de Pearson referente ao floema dos feixes do tipo concêntrico. (a) Diagrama de Pearson: *Colormap* indica os valores da correlação de Pearson; (b) Rede de correlação de Pearson: Atributos conectados indicam que eles possuem, em módulo, valores de $p \geq 0,9$.

Fonte: Autor

A Figura 51a mostra, por meio de um *colormap*, os valores obtidos para a variável p em todas as duplas de atributos. Nesse caso, os quadrados com cores mais escuras (verde escuro) indicam $p \geq 0,9$ e com cores mais claras (transparentes) indicam $p \leq -0,9$. Já a Figura 51b liga somente os atributos que apresentaram $|p| \geq 0,9$. Nesse caso, os atributos de uniformidade espacial e o atributo poligonalidade média possuem $|p| \geq 0,9$ entre eles. A lacunaridade com raio 6 possui $|p| = 0,94$ com a lacunaridade com raio 11 que, por sua vez, apresenta $|p| \geq 0,9$ com as lacunaridades 21 e 16. Estas últimas também estão conectadas por possuírem $|p| = 0,99$. Os atributos de forma diâmetro equivalente médio, resíduo

médio, diâmetro médio e área média possuem $|p| \geq 0,9$ entre elas e com a densidade radial com raio de 75. Esta, por sua vez, apresenta $|p| = 0,93$ com a densidade radial com raio 25 que apresenta $|p| = 0,91$ com densidade radial com raio 125. Além disso, os atributos de forma desvio padrão do diâmetro equivalente, desvio padrão do diâmetro e desvio padrão da área apresentam $|p| \geq 0,9$. Os pares de atributos que apresentaram $|p| \geq 0,9$ foram filtrados e não foram selecionados para a próxima etapa do trabalho.

5.1.2.1.3 Feixes vasculares do tipo concêntrico: Xilema

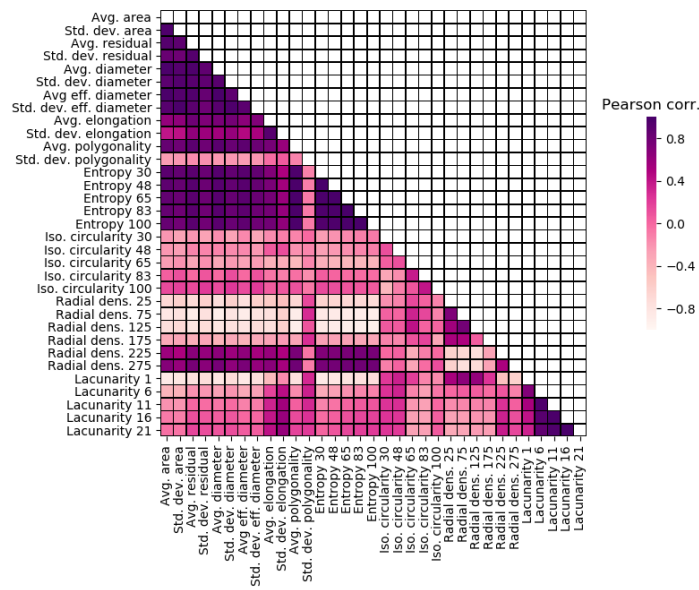


Os cálculos da correlação de Pearson foram feitos considerando os seguintes tecidos do xilema: Anfivascular Primário, Anfivascular Secundário e Anfivascular.

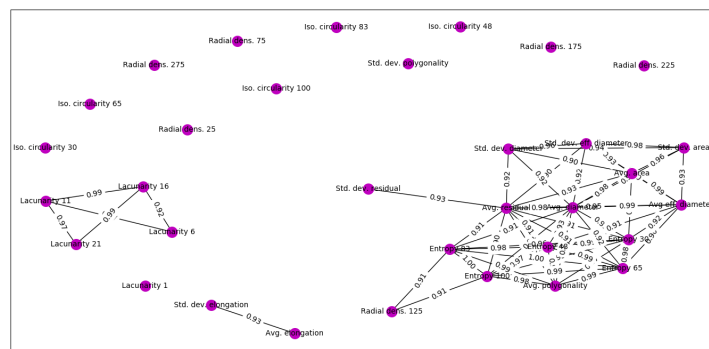
A Figura 52 ilustra os mapas com os valores de p referente ao xilema dos feixes do tipo concêntrico.

A Figura 52a mostra, por meio de um *colormap*, os valores obtidos para a variável p em todas as duplas de atributos. Nesse caso, os quadrados com cores mais escuras (roxo escuro) indicam $p \geq 0,9$ e com cores mais claras (transparentes) indicam $p \leq -0,9$. Já a Figura 52b liga somente os atributos que apresentaram $|p| \geq 0,9$. Nesse caso, as lacunaridades com raios 11 e 16 possuem $|p| \geq 0,9$ entre elas e com as lacunaridades 6 e 21. O atributo desvio padrão da alongação apresenta $|p| = 0,93$ com a alongação média. Os atributos de forma desvio padrão do diâmetro, desvio padrão do diâmetro equivalente, desvio padrão da área, resíduo médio, diâmetro médio, área média e diâmetro equivalente médio possuem $|p| \geq 0,9$ entre eles e os 4 últimos apresentam $|p| \geq 0,9$ com todos os atributos de uniformidade espacial. Estes, apresentam $|p| \geq 0,9$ com a poligonalidade média. Já o atributo desvio padrão do resíduo $|p| = 0,93$ com o resíduo médio. E o atributo densidade radial com raio 125 $|p| \geq 0,9$ com os atributos uniformidade espacial sigma 83 e 100.

Os pares de atributos que apresentaram $|p| \geq 0,9$ foram filtrados e não foram selecionados para a próxima etapa do trabalho.



(a) Diagrama da correlação de Pearson

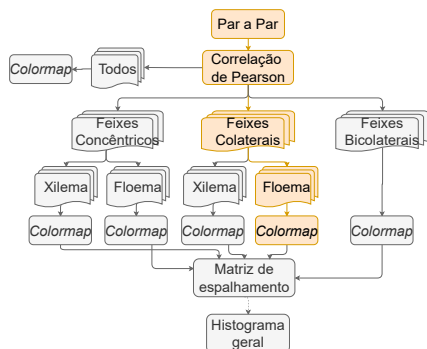


(b) Rede de correlação de Pearson

Figura 52 – Mapas da correlação de Pearson referente ao xilema dos feixes do tipo concêntrico. (a) Diagrama de Pearson: *Colormap* indica os valores da correlação de Pearson; (b) Rede de correlação de Pearson: Atributos conectados indicam que eles possuem, em módulo, valores de $p \geq 0,9$.

Fonte: Autor

5.1.2.1.4 Feixes vasculares do tipo colateral: Floema

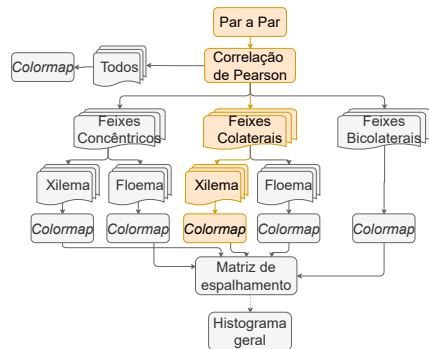


Os cálculos de p foram feitos considerando os seguintes tecidos do floema: Colateral aberto e Colateral fechado.

$|p| = 0,97$ com a uniformidade espacial 83. Já o atributo desvio padrão do diâmetro equivalente apresenta $|p| = 0,93$ com o desvio padrão da área.

Os pares de atributos que apresentaram $|p| \geq 0,9$ foram filtrados e não foram selecionados para a próxima etapa do trabalho.

5.1.2.1.5 Feixes vasculares do tipo colateral: Xilema

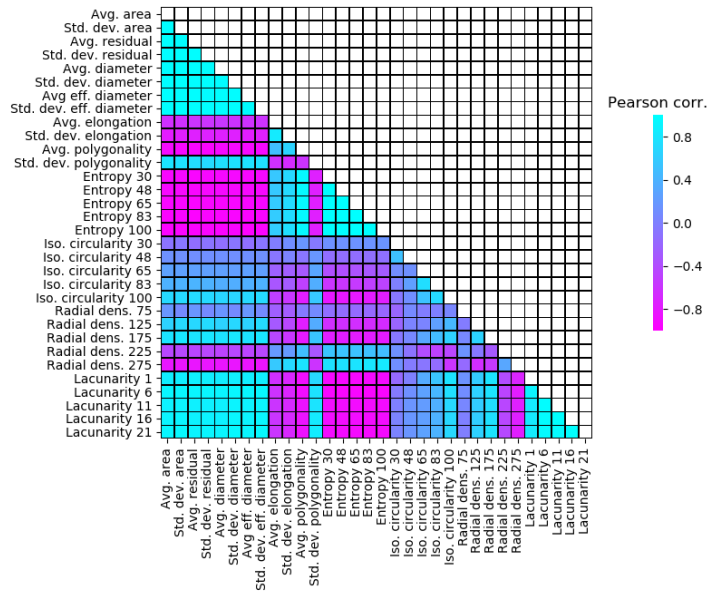


Os cálculos de p foram feitos considerando os seguintes tecidos do xilema: Colateral aberto e Colateral fechado.

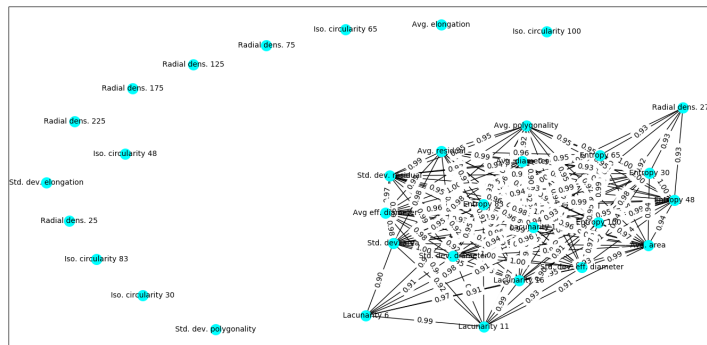
A Figura 54 ilustra os mapas com os valores de p referente ao xilema dos feixes do tipo colateral.

A Figura 54a mostra, por meio de um *colormap*, os valores obtidos para a variável p em todas as duplas de atributos. Nesse caso, os quadrados com tons azuis claro indicam $p \geq 0,9$ em *pink* indicam $p \leq -0,9$. Já a Figura 54b liga somente os atributos que apresentaram $|p| \geq 0,9$. Nesse caso, vemos uma $|p| \geq 0,9$ entre um grande número de atributos. Os atributos de lacunaridade possuem $|p| \geq 0,9$ entre elas e entre os atributos de forma, com exceção da alongação média e do desvio padrão da alongação e entre todos os atributos de Uniformidade espacial. Os atributos de uniformidade espacial, por sua vez, apresentam $|p| \geq 0,9$ entre eles, com os atributos de forma com exceção da alongação média e do desvio padrão da alongação, com a densidade radial com raio 275 e com a poligonalidade média. Este último também apresenta $|p| \geq 0,9$ com os atributos de forma com exceção da alongação média e do desvio padrão da alongação.

Os pares de atributos que apresentaram $|p| \geq 0,9$ foram filtrados e não foram selecionados para a próxima etapa do trabalho.



(a) Diagrama da correlação de Pearson

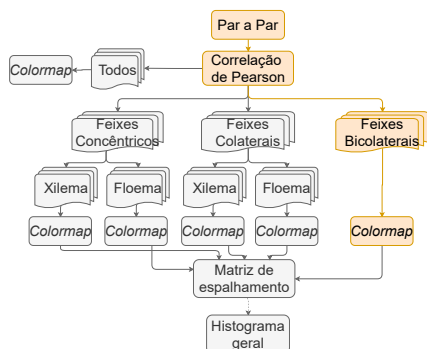


(b) Rede de correlação de Pearson

Figura 54 – Mapas da correlação de Pearson referente ao xilema dos feixes do tipo colateral.
 (a) Diagrama de Pearson: *Colormap* indica os valores da correlação de Pearson;
 (b) Rede de correlação de Pearson: Atributos conectados indicam que eles possuem, em módulo, valores de $p \geq 0,9$.

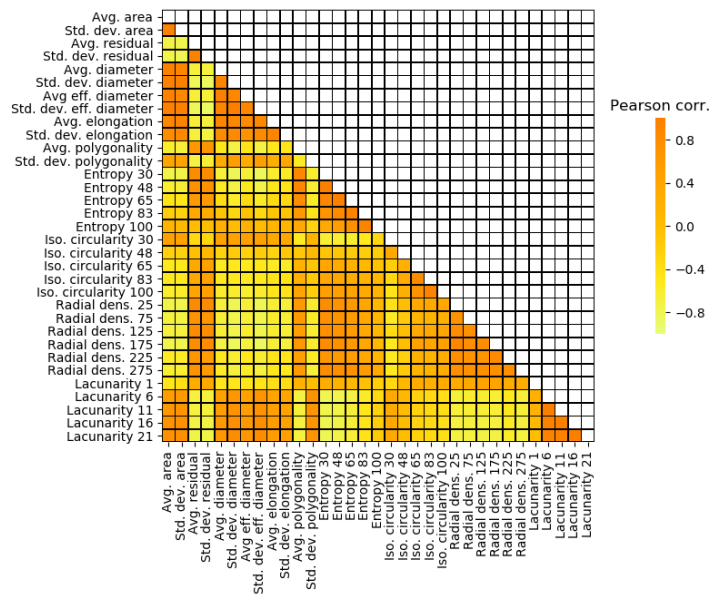
Fonte: Autor

5.1.2.1.6 Feixes vasculares do tipo bicolateral

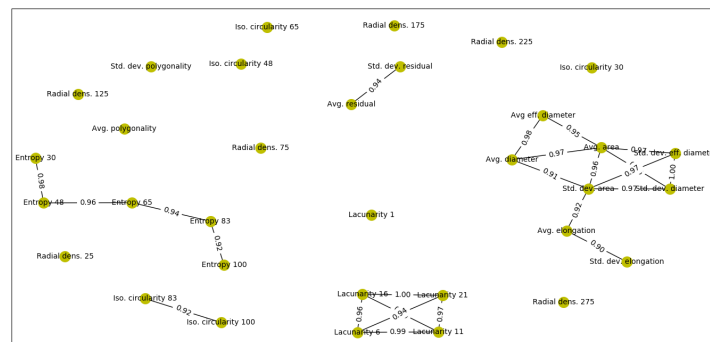


Os cálculos de p foram feitos considerando os tecidos que compõem os feixes bicolaterais: Floema e Xilema.

A Figura 55 ilustra os mapas com os valores de p referente aos tecidos vasculares dos feixes do tipo bicolateral.



(a) Diagrama da correlação de Pearson



(b) Rede de correlação de Pearson

Figura 55 – Mapas da correlação de Pearson referente aos tecidos vasculares dos feixes do tipo bicolateral. (a) Diagrama de Pearson: *Colormap* indica os valores da correlação de Pearson; (b) Rede de correlação de Pearson: Atributos conectados indicam que eles possuem, em módulo, valores de $p \geq 0,9$.

Fonte: Autor

A Figura 55a mostra, por meio de um *colormap*, os valores obtidos para a variável p em todas as duplas de atributos. Nesse caso, os quadrados com cores alaranjadas mais escuras indicam $p \geq 0,9$ e com cores amareladas mais claras indicam $p \leq -0,9$. Já a Figura 55b liga somente os atributos que apresentaram $|p| \geq 0,9$. Nesse caso, os atributos de lacunaridade com raio 6, 11, 16 e 21 apresentam $|p| \geq 0,9$. O atributo circularidade *isoline* 83 apresenta $|p| = 0,92$ com o atributo circularidade *isoline* 100. Já os atributos resíduo médio e o desvio padrão do resíduo apresentam $|p| = 0,94$. O atributo uniformidade espacial com sigma 30 apresenta $|p| = 0,98$ com a uniformidade espacial 48 que, por sua

vez, apresenta $|p| = 0,96$ com a uniformidade espacial 65. Esta apresenta $|p| = 0,94$ com a uniformidade espacial 83 que, por sua vez, apresenta $|p| = 0,92$ com a Uniformidade espacial com sigma 100. Já os atributos de forma diâmetro equivalente médio, diâmetro médio, área média e desvio padrão da área possuem $|p| \geq 0,9$ entre eles. Os dois últimos atributos também apresentam $|p| \geq 0,9$ com os atributos desvio padrão do diâmetro equivalente e desvio padrão do diâmetro. Estes apresentam $|p| = 1$ entre eles. O atributo desvio padrão da alongação apresenta $|p| = 0,9$ com média da alongação que, por sua vez, apresenta $|p| = 0,92$ com o atributo desvio padrão da área.

Os pares de atributos que apresentaram $|p| \geq 0,9$ foram filtrados e não foram selecionados para a próxima etapa do trabalho.

5.1.2.2 Análise de dispersão: Matriz de espalhamento

Ao final da aplicação da abordagem de correlação de Pearson, os pares de atributos selecionados de cada conjunto tecidual foram estudados por meio da técnica matriz de espalhamento.

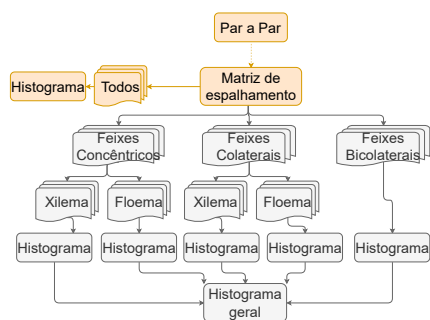
Para facilitar a visualização de quais atributos foram mais relevantes para cada conjunto de tecidos estudados, foi plotado um gráfico de barras onde cada barra representa um atributo. No gráfico de barras os atributos estão numerados de 1-33. A Tabela 13 mostra a relação entre os números e os atributos correspondentes. Ela também mostra qual a categoria de cada atributo sendo que o atributos marcados com * fazem parte da categoria multiescala.

Tabela 13 – Correspondência dos atributos com os índices. Atributos marcados com * fazem parte da categoria multiescala.

Número	Nome do atributo	Categorias dos atributos
1	Área média	Medidas referentes ao formato celular
2	Desvio Padrão Área	
3	Resíduo médio	
4	Desvio Padrão Resíduo	
5	Diâmetro médio	
6	Desvio Padrão do Diâmetro	
7	Diâmetro equivalente médio	
8	Desvio Padrão diâmetro equivalente	
9	Elongação média	
10	Desvio Padrão elongação	
11	Uniformidade espacial 30 *	Medidas referente a densidade
12	Uniformidade espacial 48 *	
13	Uniformidade espacial 65 *	
14	Uniformidade espacial 83 *	
15	Uniformidade espacial 100 *	
16	Circularidade <i>isoline</i> 30 *	
17	Circularidade <i>isoline</i> 48 *	
18	Circularidade <i>isoline</i> 65 *	
19	Circularidade <i>isoline</i> 83 *	
20	Circularidade <i>isoline</i> 100 *	
21	Densidade radial 25 *	Medidas referentes a regularidade estrutural
22	Densidade radial 75 *	
23	Densidade radial 125 *	
24	Densidade radial 175 *	
25	Densidade radial 225 *	
26	Densidade radial 275 *	
27	Poligonalidade média	
28	Desvio Padrão da poligonalidade	
29	Lacunaridade 1 *	
30	Lacunaridade 6 *	
31	Lacunaridade 11 *	
32	Lacunaridade 16 *	
33	Lacunaridade 21 *	

Os ícones de cada subseção foram feitos com base na Figura 49, porém omitimos a etapa correspondente à correlação de Pearson.

5.1.2.2.1 Todos os feixes vasculares



Para essa etapa, foram calculadas as matrizes de espalhamento para os pares de atributos que tinham $|p| \leq 0,9$. Os cálculos foram feitos considerando todos os feixes vasculares estudados: anfigasal primário, anfigasal secundário, anfigival, bicolateral, colateral aberto e colateral fechado. Os feixes foram subdivididos em seus tecidos constituintes, ou seja, floema e xilema.

Após aplicação da abordagem matriz de espalhamento, foram selecionados os 10 gráficos de dispersão com maiores valores da variável $t(R)$. Sendo, (R) a relação entre S_{inter} e S_{intra} e, (t) o traço dessa matriz.

A Tabela 14 mostra os valores calculados de $t(R)$ para cada um dos gráficos de dispersão plotados com as variáveis descritas nas colunas *Eixo Y* e *Eixo X*. Os gráficos podemos ser visualizados no Apêndice B.1.

Tabela 14 – Tabela com os valores de $t(R)$ para todos os feixes vasculares vegetais estudados.

Eixo X	Eixo Y	$t(S_{intra})$	$t(S_{entre})$	$t(R)$
Uniformidade espacial 30	Diâmetro eq. médio	15136,73	281934,72	37,33
Uniformidade espacial 48	Diâmetro eq. médio	15886,55	279367,49	35,22
Uniformidade espacial 30	Desv. P. diâmetro eq. médio	13557,89	234457,21	34,96
Poligonalidade média	Diâmetro eq. médio	15921,30	275705,67	34,95
Poligonalidade média	Diâmetro médio	17435,96	290946,71	33,28
Uniformidade espacial 48	Desv. P. diâmetro eq. médio	14307,70	231889,98	32,63
Uniformidade espacial 65	Diâmetro eq. médio	16728,61	269853,19	32,57
Poligonalidade média	Desv. P. diâmetro eq. médio	14342,45	228228,15	32,30
Uniformidade espacial 30	Desv. P. diâmetro médio	15140,53	235596,45	31,84
Uniformidade espacial 65	Diâmetro médio	18243,27	285094,23	31,07

O gráfico de barras evidenciado na Figura 56 mostra quais atributos foram mais relevantes quando foram estudados todos os feixes vasculares juntos. Cada barra representa um atributo, conforme descrito na Tabela 13.

Dessa maneira, vemos que os atributos mais recorrentes no estudo par a par entre todos os feixes vasculares são: diâmetro equivalente médio, desvio padrão do diâmetro equivalente médio, Uniformidade espacial 30 e Poligonalidade média. Eles estão distribuídos entre as quatro categorias de atributos selecionadas o que indica a importância de se estudar mais de uma classe de atributos quanto o objetivo for a caracterização de feixes vasculares vegetais.

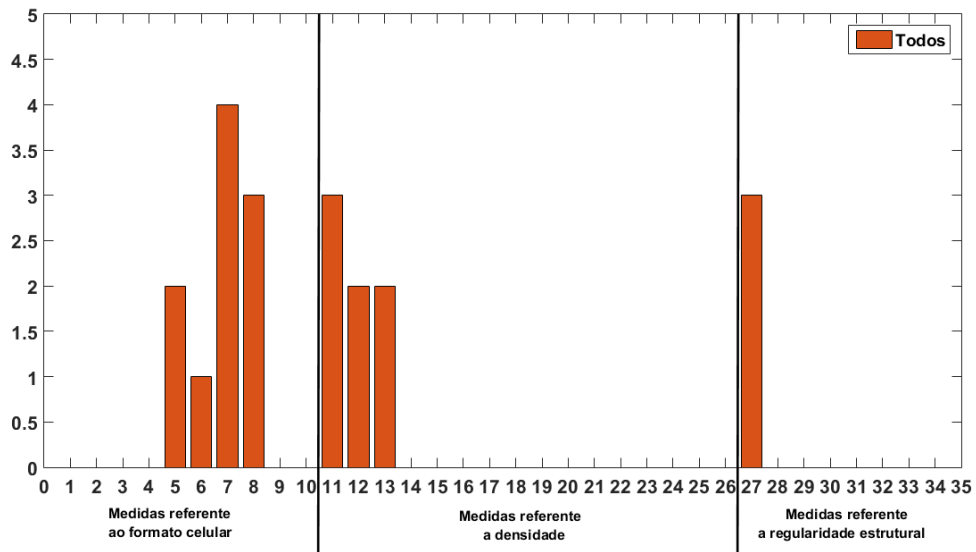
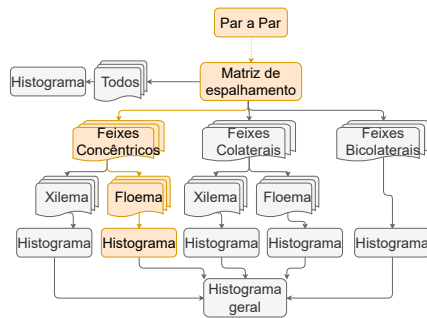


Figura 56 – Gráfico de barras com os atributos mais relevantes quando foram estudados todos os feixes vasculares. Cada barra representa um atributo.

Fonte: Autor

5.1.2.2.2 Feixes vasculares do tipo concêntricos: Floema



Para essa etapa, foram calculadas as matrizes de espalhamento para os pares de atributos que tinham $|p| \leq 0,9$. Os cálculos foram feitos considerando os seguintes tecidos do floema: Anfivascular Primário, Anfivascular Secundário e Anficrival.

Após aplicação da abordagem matriz de espalhamento, foram selecionados os 10 gráficos de dispersão com maiores valores da variável $t(R)$. Sendo, (R) a relação entre S_{inter} e S_{intra} e, (t) o traço dessa matriz.

A Tabela 15 mostra os valores calculados de $t(R)$ para cada um dos gráficos de dispersão plotados com os eixos x e y plotados como descritos nas colunas *Eixo X* e *Eixo Y*. Os gráficos podemos ser visualizados no Apêndice B.2.

Tabela 15 – Tabela com os valores de $t(R)$ para gráficos de dispersão referentes ao floema dos feixes vasculares do tipo concêntrico.

Eixo X	Eixo Y	$t(S_{intra})$	$t(S_{entre})$	$t(R)$
Densidade radial 75	Uniformidade espacial 30	4851,20	139893,89	59,07
Densidade radial 75	Uniformidade espacial 48	6113,49	137278,47	48,91
Densidade radial 75	Poligonalidade média	7136,76	118392,20	43,36
Densidade radial 75	Uniformidade espacial 65	8374,95	128165,12	41,92
Uniformidade espacial 30	Diâmetro médio	9047,25	119454,96	35,02
Uniformidade espacial 30	Diâmetro eq. médio	10441,64	123190,63	32,85
Densidade radial 25	Uniformidade espacial 30	10695,17	119797,09	28,20
Uniformidade espacial 30	Resíduo médio	10444,60	93846,49	26,72
Uniformidade espacial 30	Desv. P. diâmetro eq.	13256,87	80132,64	25,58
Uniformidade espacial 30	Desv. P. resíduo	16601,59	73908,59	25,27

O gráfico de barras evidenciado na Figura 57 mostra quais atributos foram mais relevantes quando foram estudados todos os feixes vasculares juntos. Cada barra representa um atributo.

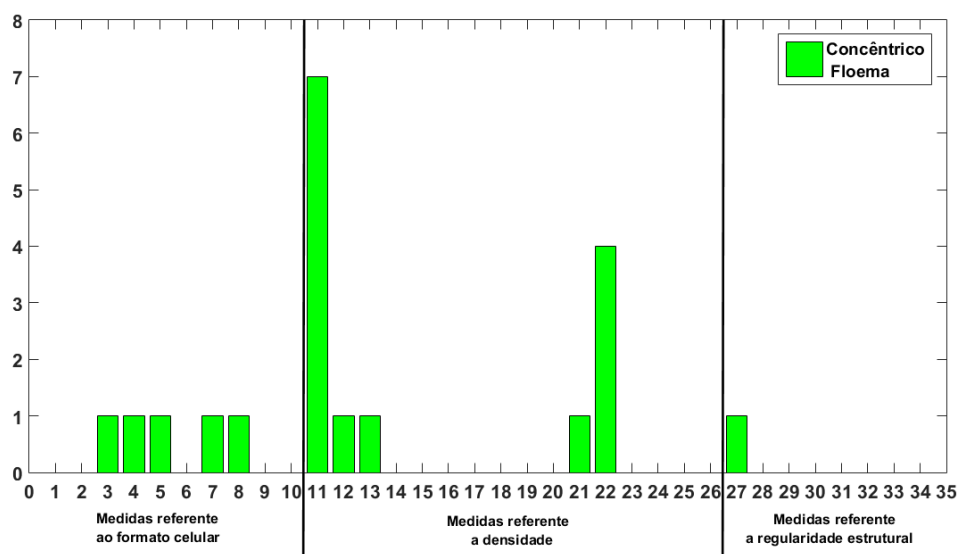
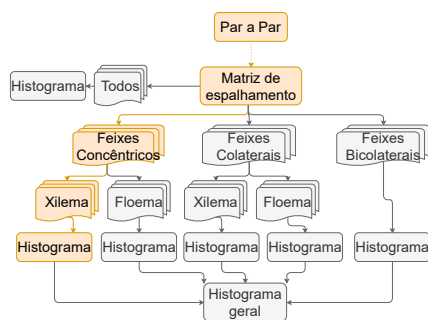


Figura 57 – Gráfico de barras com os atributos mais relevantes quando foram estudados os floemas dos feixes vasculares do tipo concêntrico. Cada barra representa um atributo.

Fonte: Autor

Dessa maneira, vemos que os atributos mais recorrentes no estudo par a par entre os floemas dos feixes vasculares do tipo concêntrico são: Uniformidade espacial 30 e densidade radial com raio 75. Nesse caso, vemos que o atributos com maior relevância fazem parte das categorias densidade e multiescala.

5.1.2.2.3 Feixes vasculares do tipo concêntricos: Xilema



Para essa etapa, foram calculadas as matrizes de espalhamento para os pares de atributos que tinham $|p| \leq 0,9$. Os cálculos foram feitos considerando o xilema nos seguintes feixes: Anfivasal Primário, Anfivasal Secundário e Anficrival.

Após aplicação da abordagem matriz de espalhamento, foram selecionados os 10 gráficos de dispersão com maiores valores da variável $t(R)$. Sendo, (R) a relação entre S_{inter} e S_{intra} e, (t) o traço dessa matriz.

A Tabela 16 mostra os valores calculados de $t(R)$ para cada um dos gráficos de dispersão plotados com os eixos x e y plotados como descritos nas colunas *Eixo X* e *Eixo Y*. Os gráficos podemos ser visualizados no Apêndice B.3.

Tabela 16 – Tabela com os valores de $t(R)$ para gráficos de dispersão referentes ao xilema dos feixes vasculares do tipo concêntrico.

Eixo X	Eixo Y	$t(S_{intra})$	$t(S_{entre})$	$t(R)$
Poligonalidade média	Diâmetro eq. médio	6673,80	111337,63	47,36
Uniformidade espacial 48	Área média	6705,99	98112,08	35,51
Lacunaridade 1	Uniformidade espacial 83	8195,24	107371,15	34,67
Lacunaridade 1	Uniformidade espacial 65	8130,14	102725,98	33,09
Lacunaridade 6	Poligonalidade média	10099,76	85919,51	27,98
Lacunaridade 1	Uniformidade espacial 48	8457,48	97705,43	26,55
Lacunaridade 1	Poligonalidade média	8800,59	100735,69	25,13
Densidade radial 275	Poligonalidade média	20596,03	93797,36	23,19
Lacunaridade 6	Uniformidade espacial 30	10357,73	78711,25	22,04
Lacunaridade 1	Diâmetro eq. médio	10005,37	89625,14	20,96

O gráfico de barras evidenciado na Figura 58 mostra quais atributos foram mais relevantes quando foram estudados os tecidos do xilema dos feixes vasculares do tipo concêntrico. Cada barra representa um atributo, conforme descrito na Tabela 13.

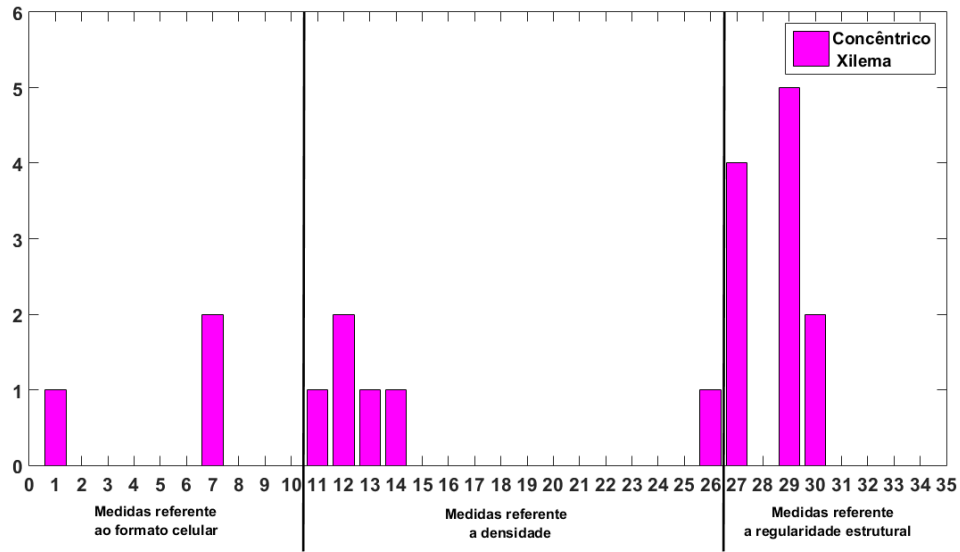
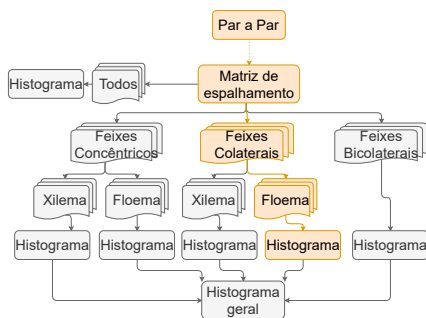


Figura 58 – Gráfico de barras com os atributos mais relevantes quando foram estudados os xilemas dos feixes vasculares do tipo concêntrico. Cada barra representa um atributo.

Fonte: Autor

Dessa maneira, vemos que os atributos mais recorrentes no estudo par a par entre os xilemas dos feixes vasculares do tipo concêntrico são: Diâmetro equivalente médio, uniformidade espacial 48, poligonalidade média, lacunaridade 1 e lacunaridade 6. Nesse caso, os atributos mais relevantes estão distribuídos entre as quatro categorias de atributos selecionadas (formato celular, densidade, regularidade estrutural e multiescala)

5.1.2.2.4 Feixes vasculares do tipo colateral: Floema



Para essa etapa, foram calculadas as matrizes de espalhamento para os pares de atributos que tinham $|p| \leq 0,9$. Os cálculos foram feitos considerando os floemas dos seguintes feixes: Colateral aberto e Colateral fechado.

Após aplicação da abordagem matriz de espalhamento, foram selecionados os 10 gráficos de dispersão com maiores valores da variável $t(R)$. Sendo, (R) a relação entre S_{inter} e S_{intra} e, (t) o traço dessa matriz.

A Tabela 17 mostra os valores calculados de $t(R)$ para cada um dos gráficos de dispersão plotados com os eixos x e y plotados como descritos nas colunas *Eixo X* e *Eixo Y*. Os gráficos podemos ser visualizados no Apêndice B.4.

Tabela 17 – Tabela com os valores de $t(R)$ para gráficos de dispersão referentes ao floema dos feixes vasculares do tipo colateral.

Eixo X	Eixo Y	$t(S_{intra})$	$t(S_{entre})$	$t(R)$
Densidade radial 275	Resíduo médio	9947,42	72853,13	21,66
Densidade radial 275	Diâmetro médio	10376,92	70117,56	17,98
Uniformidade espacial 30	Resíduo médio	11647,57	64123,36	17,35
Densidade radial 275	Diâmetro eq. médio	11257,08	69728,43	15,51
Uniformidade espacial 30	Diâmetro médio	12077,07	61387,79	14,66
Densidade radial 275	Área média	12365,54	68431,26	14,03
Uniformidade espacial 30	Diâmetro eq. médio	12957,24	60998,66	12,07
Uniformidade espacial 30	Área média	14065,70	59701,49	10,22
Lacunaridade 6	Diâmetro eq. médio	15449,01	59085,50	9,53
Densidade radial 275	Densidade radial 25	16495,07	29441,77	5,36

O gráfico de barras evidenciado na Figura 59 mostra quais atributos foram mais relevantes quando foram estudados floema dos feixes vasculares do tipo colateral. Cada barra representa um atributo, conforme descrito na Tabela 13.

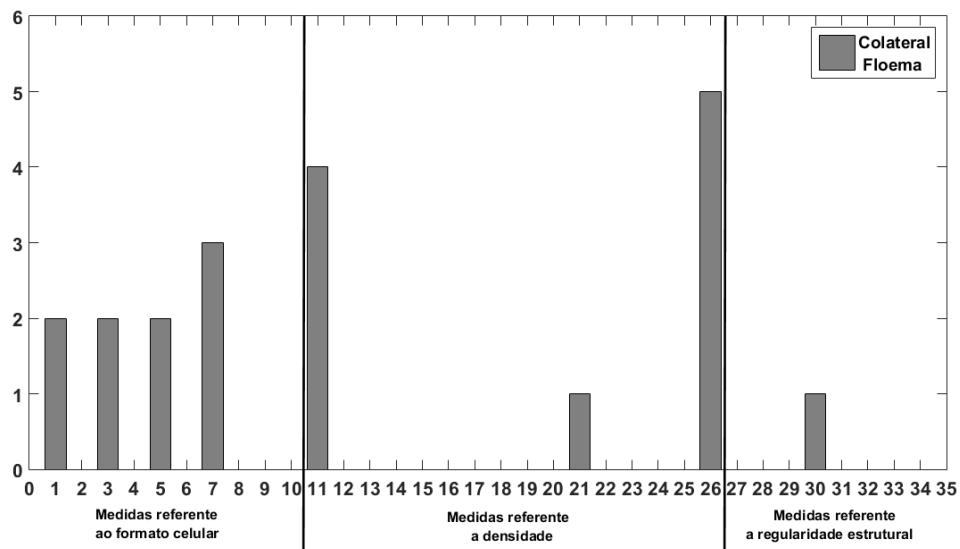
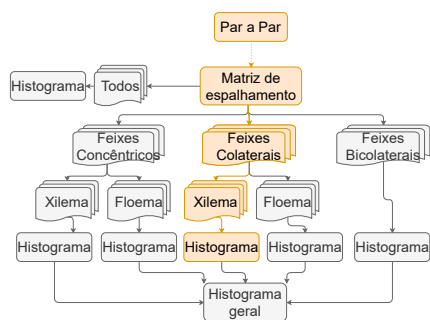


Figura 59 – Gráfico de barras com os atributos mais relevantes quando foram estudados os floemas dos feixes vasculares do tipo colateral. Cada barra representa um atributo.

Fonte: Autor

Dessa maneira, vemos que os atributos mais recorrentes no estudo par a par entre os floemas dos feixes vasculares do tipo colateral são: diâmetro equivalente médio, uniformidade espacial 30 e densidade radial com raio 275. Nesse caso, vemos que a única categoria que não possuiu representantes relevantes foi a regularidade estrutural.

5.1.2.2.5 Feixes vasculares do tipo colateral: Xilema



Para essa etapa, foram calculadas as matrizes de espalhamento para os pares de atributos que tinham $|p| \leq 0,9$. Os cálculos foram feitos considerando o xilema dos seguintes feixes: Colateral aberto e Colateral fechado.

Após aplicação da abordagem matriz de espalhamento, foram selecionados os 10 gráficos de dispersão com maiores valores da variável $t(R)$. Sendo, (R) a relação entre S_{inter} e S_{intra} e, (t) o traço dessa matriz.

A Tabela 18 mostra os valores calculados de $t(R)$ para cada um dos gráficos de dispersão plotados com os eixos x e y plotados como descritos nas colunas *Eixo X* e *Eixo Y*. Os gráficos podemos ser visualizados no Apêndice B.5.

Tabela 18 – Tabela com os valores de $t(R)$ para gráficos de dispersão referentes ao xilema dos feixes vasculares do tipo colateral.

Eixo X	Eixo Y	$t(S_{intra})$	$t(S_{entre})$	$t(R)$
Lacunaridade 16	Poligonalidade média	9594,96	85709,60	36,62
Lacunaridade 1	Desv. P. diâmetro eq.	9117,82	91398,44	36,33
Lacunaridade 11	Poligonalidade média	8908,08	80569,59	31,87
Lacunaridade 1	Resíduo médio	9497,06	84525,17	26,45
Lacunaridade 1	Desv. P. diâmetro	10012,04	84143,98	24,23
Lacunaridade 1	Desv. P. resíduo	9984,14	83453,39	23,17
Lacunaridade 1	Desv. P. área	10328,98	90308,10	22,72
Lacunaridade 1	Diâmetro eq. médio	10015,01	79293,20	21,96
Lacunaridade 1	Diâmetro médio	10403,39	78782,34	21,68
Lacunaridade 16	Uniformidade espacial 65	9227,11	76929,82	18,48

O gráfico de barras evidenciado na Figura 60 mostra quais atributos foram mais relevantes quando foram estudados os tecidos do xilema dos feixes vasculares do tipo colateral. Cada barra representa um atributo, conforme descrito na Tabela 13.

Dessa maneira, vemos que os atributos mais recorrentes no estudo par a par entre os xilemas dos feixes vasculares do tipo colateral são: poligonalidade média, lacunaridade 1 e lacunaridade 16. Nesse caso, os atributos mais relevantes são das classes regularidade estrutural e multiescala.

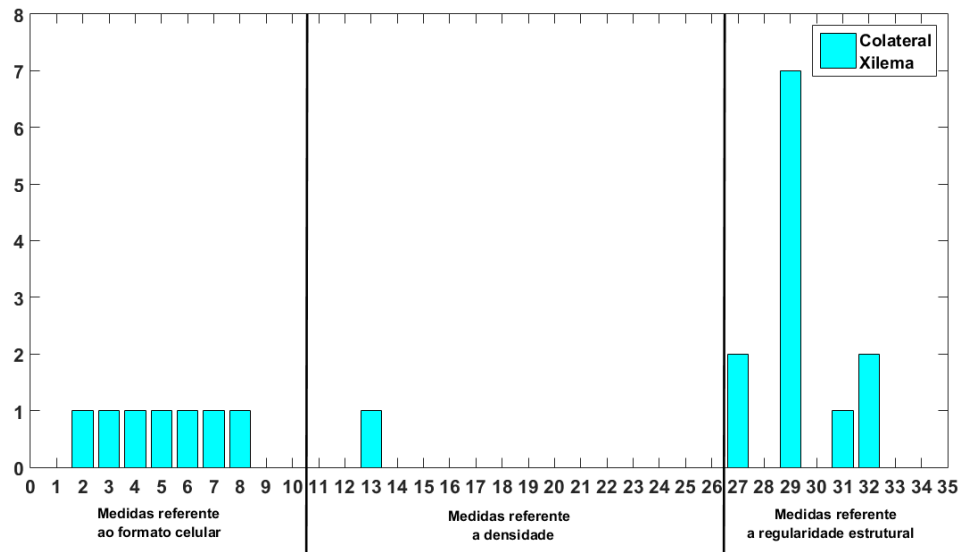
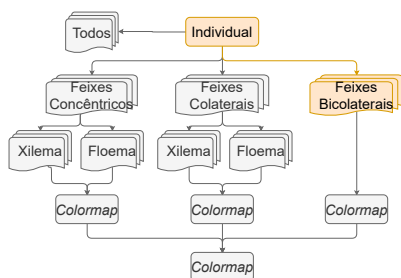


Figura 60 – Gráfico de barras com os atributos mais relevantes quando foram estudados os xilemas dos feixes vasculares do tipo colateral. Cada barra representa um atributo.

Fonte: Autor

5.1.2.2.6 Feixes vasculares do tipo biclateral



Os feixes biclaterais não apresentam subtipos. Por essa razão, os resultados referentes ao xilema e ao floema foram apresentados juntos. Para essa etapa, foram calculadas as matrizes de espalhamento para os pares de atributos que tinham $|p| \leq 0,9$. Os cálculos foram feitos considerando os tecidos que compõem os feixes biclaterais: Floema e Xilema.

Após aplicação da abordagem matriz de espalhamento, foram selecionados os 10 gráficos de dispersão com maiores valores da variável $t(R)$. Sendo, (R) a relação entre S_{inter} e S_{intra} e, (t) o traço dessa matriz.

A Tabela 18 mostra os valores calculados de $t(R)$ para cada um dos gráficos de dispersão plotados com os eixos x e y plotados como descritos nas colunas *Eixo X* e *Eixo Y*. Os gráficos podemos ser visualizados no Apêndice B.6.

Tabela 19 – Tabela com os valores de $t(R)$ para gráficos de dispersão referentes aos feixes vasculares do tipo bicolateral.

Eixo X	Eixo Y	$t(S_{intra})$	$t(S_{entre})$	$t(R)$
Desv. P. diâmetro eq.	Desv. P. resíduo	13162,61	87353,58	14,94
Desv. P. diâmetro	Desv. P. resíduo	13695,51	84196,96	13,47
Densidade radial 175	Desv. P. diâmetro eq.	12588,44	69856,53	11,23
Densidade radial 25	Desv. P. diâmetro eq.	12441,13	62591,19	11,09
Uniformidade espacial 30	Desv. P. resíduo	14992,50	73957,97	10,18
Lacunaridade 21	Desv. P. resíduo	15109,65	61872,17	9,77
Densidade radial 25	Desv. P. diâmetro	12974,04	59434,57	9,70
Lacunaridade 6	Lacunaridade 1	27581,28	14132,51	8,11
Densidade radial 175	Uniformidade espacial 30	14418,34	56460,92	7,46
Lacunaridade 11	Lacunaridade 1	27001,32	14910,61	6,80

O gráfico de barras evidenciado na Figura 61 mostra quais atributos foram mais relevantes quando foram estudados os tecidos vasculares dos feixes do tipo bicolateral. Cada barra representa um atributo, conforme descrito na Tabela 13.

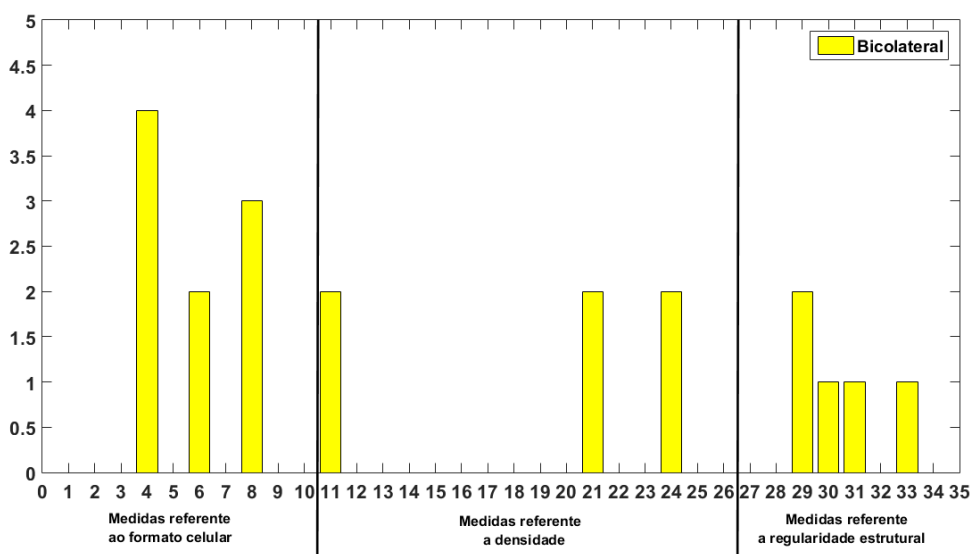
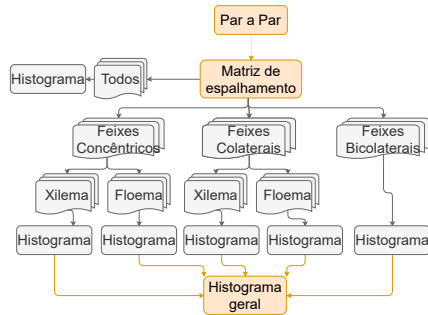


Figura 61 – Gráfico de barras com os atributos mais relevantes quando foram estudados os feixes vasculares do tipo bicolateral. Cada barra representa um atributo.

Fonte: Autor

Dessa maneira, vemos que os atributos mais recorrentes no estudo par a par entre os xilemas dos feixes vasculares do tipo colateral são: desvio padrão resíduo, diâmetro médio, desvio padrão do diâmetro equivalente, uniformidade espacial 30, densidade radial com raios 25 e 175 e lacunaridade 1.

5.1.2.3 Considerações gerais



Partindo como premissa os resultados apresentados para a análise par a par dos feixes vasculares vegetais podemos buscar entender quais medidas foram mais recorrentes ao longo dessa análise. Para isso, os resultados foram concatenados em um gráfico de barras, Figura 62, no qual cada barra corresponde a um atributo conforme descrito na Tabela 13.

A importância de cada atributo foi estabelecida com base na recorrência desse atributo nos resultados apresentados. Essa abordagem foi feita para cada tipo de feixe levando em consideração os dois tipos de tecidos vasculares que compõem esses feixes, o floema e o xilema. Ao final, os resultados foram agrupados e o gráfico de barras construído.

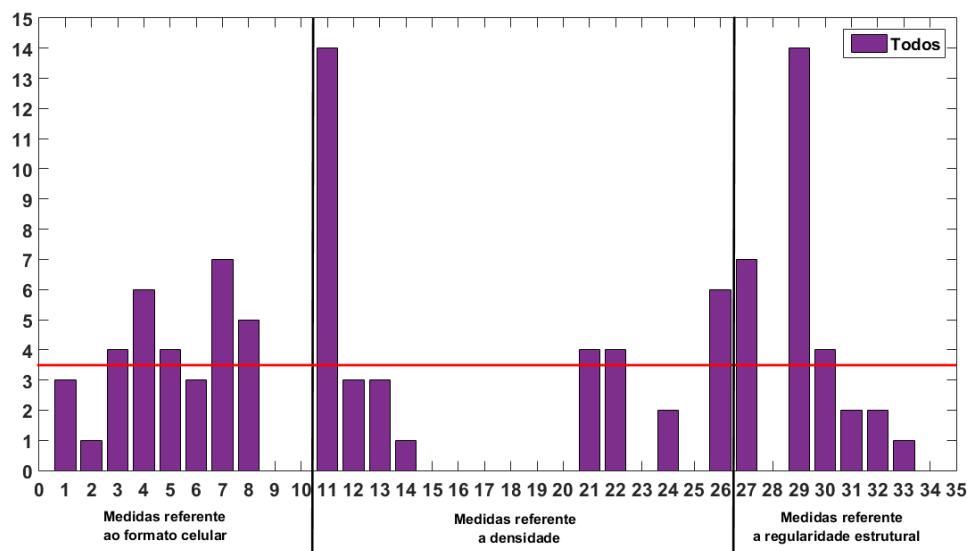


Figura 62 – Gráfico de barras com os atributos mais relevantes considerando os resultados obtidos na análise par a par. Cada barra representa um atributo.

Fonte: Autor

A linha vermelha na Figura 62 corresponde a um *threshold* calculado por meio do valor do terceiro quartil dessa distribuição. Partindo disso, os atributos que possuem uma frequência acima desse *threshold* foram consideradas os mais relevantes para a análise par a par.

De maneira geral, podemos notar que a distribuição geral dos atributos ao longo da análise par a par engloba os quatro grupos de medidas estudados (formato celular, densidade, regularidade estrutural e multiescala). Assim, podemos concluir que os atributos que foram importantes para caracterizar e segregar os feixes vasculares vegetais considerando apenas a par a par são: resíduo médio, desvio padrão do desvio médio, diâmetro médio,

diâmetro equivalente médio, desvio padrão diâmetro equivalente, uniformidade espacial 30, densidade radial 25, densidade radial 75, densidade radial 275, poligonalidade média, lacunaridade 1 e lacunaridade 6.

Outro ponto relevante que podemos sugerir com base nos resultados da análise par a par tem relação com o efeito multiescala dos atributos. Alguns atributos, como por exemplo, a uniformidade espacial, a densidade radial e a lacunaridade apresentaram recorrência em escalas específicas. Isso indica que o uso da análise multiescala para extração dos atributos é importante e pode contribuir na caracterização e na segregação de feixes vasculares vegetais.

5.1.3 Análise de componentes principais

A análise de componentes principais foi feita com todos os atributos, ou seja, inicialmente as classes de tecidos vasculares vegetais estavam distribuídas no espaço original de 33 dimensões. Um dos benefícios em se realizar a análise de componentes principais é a diminuição das dimensões de um conjunto de dados. Em nosso trabalho, optamos por analisar as classes teciduais em duas dimensões (primeiro e segundo componentes principais) e três dimensões (primeiro, segundo e terceiro componentes principais). A fim de elencar as medidas mais relevantes para a segregação e caracterização de cada subclasse, optou-se por selecionar as medidas com maiores pesos do primeiro e do segundo componente principal, uma vez que eles são os responsáveis pela maior variância.

A Figura 63 ilustra a divisão das classes teciduais e das etapas realizadas na análise de componente principal.

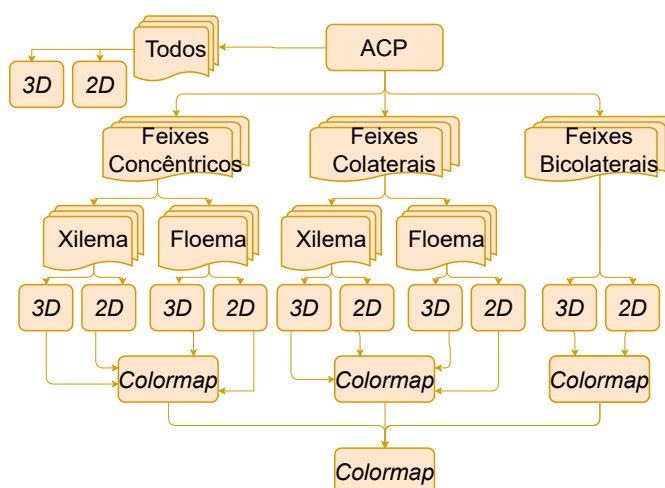


Figura 63 – Fluxograma com as divisões das classes estudadas e as etapas da ACP.

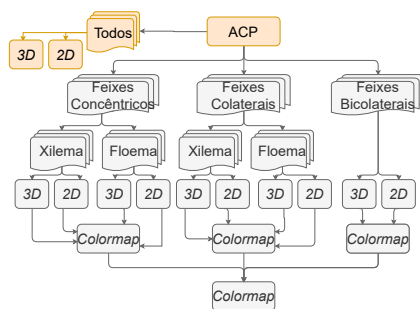
Fonte: Autor

Nessa seção, iremos apresentar os resultados obtidos para as seguintes classes

tecidualais: todas as classes; feixes concêntricos; feixes colaterais; e feixes bicolaterais. Ao final, os atributos mais relevantes para cada classe tecidual serão apresentados em um *colormap* (vide Figura 63) e, assim, podemos elencar quais descritores foram mais recorrentes em todas as classes durante a aplicação da abordagem análise de componentes principais.

Os ícones de cada subseção foram feitos com base na Figura 63.

5.1.3.1 Todos os feixes vasculares



Para essa etapa, foi realizada a ACP considerando todos os feixes vasculares estudados: anfigasal primário, anfigasal secundário, anfigasal crival, biclateral, clateral aberto e clateral fechado. Os feixes foram subdivididos em seus tecidos constituintes, ou seja, floema e xilema.

A Tabela 20 mostra que quando consideramos todos os feixes vasculares vegetais estudados (feixes concêntricos, feixes biclaterais e feixes colaterais) os três primeiros componentes da ACP somam juntos 67,53% da variância dos dados originais, com o primeiro componente principal tendo 31,8%, o segundo 26,6% e o terceiro 9,1%. A Tabela 20 também mostra o peso das 33 medidas analisadas e, em **negrito**, aquelas que tiveram maior relevância no cálculo dos novos eixos. Para o primeiro eixo as medidas com os maiores pesos são: Uniformidade espacial com sigmas 30, 48, 65, 83 e 100; Densidade radial com raios 225 e 275, Poligonalidade média e Lacunaridade com raios 6, 11, 16 e 21. Para o segundo eixo as medidas que apresentaram o maior peso são: Área média, Diâmetro médio, Desvio padrão do diâmetro, Diâmetro equivalente médio, Desvio padrão do diâmetro equivalente; Elongação média e Desvio padrão da elongação média.

Tabela 20 – Todas as classes: Pesos absolutos para a projeção de ACP para os componentes C_1 , C_2 e C_3 . Os valores em negrito indicam os maiores pesos.

Atributos	C_1 (31.8%)	C_2 (26.6%)	C_3 (9.1%)
Área média	0.020	0.059	0.008
Desvio Padrão área	0.017	0.050	0.014
Resíduo média	0.008	0.010	0.060
Desvio Padrão resíduo	0.007	0.010	0.062
Diâmetro médio	0.022	0.058	0.005
Desvio Padrão diâmetro	0.028	0.054	0.019
Diâmetro equivalente médio	0.024	0.056	0.006
Desvio Padrão diâmetro equivalente	0.029	0.051	0.020
Elongação média	0.012	0.057	0.002
Desvio Padrão elongação	0.010	0.054	0.008
Poligonalidade média	0.048	0.029	0.012
Desvio Padrão poligonalidade	0.032	0.029	0.009
Uniformidade espacial 30	0.049	0.033	0.003
Uniformidade espacial 48	0.049	0.035	0.001
Uniformidade espacial 65	0.049	0.034	0.008
Uniformidade espacial 83	0.047	0.033	0.013
Uniformidade espacial 100	0.045	0.031	0.018
Circularidade <i>isoline</i> 30	0.026	0.028	0.036
Circularidade <i>isoline</i> 48	0.028	0.031	0.046
Circularidade <i>isoline</i> 65	0.026	0.033	0.046
Circularidade <i>isoline</i> 83	0.028	0.020	0.048
Circularidade <i>isoline</i> 100	0.030	0.001	0.023
Densidade radial 25	0.030	0.010	0.069
Densidade radial 75	0.029	0.020	0.074
Densidade radial 125	0.015	0.031	0.066
Densidade radial 175	0.004	0.013	0.033
Densidade radial 225	0.042	0.019	0.042
Densidade radial 275	0.043	0.018	0.033
Lacunaridade 1	0.010	0.048	0.044
Lacunaridade 6	0.049	0.001	0.049
Lacunaridade 11	0.047	0.015	0.044
Lacunaridade 16	0.048	0.017	0.041
Lacunaridade 21	0.049	0.013	0.038

Dos 19 atributos com maiores pesos no primeiro e no segundo eixo, 7 são referentes as medidas de forma, 7 estão relacionados com a densidade e 5 são do grupo da regularidade estrutural. Vale ressaltar que o primeiro eixo (eixo com maior variância dos dados originais) teve como atributos de maiores pesos medidas referente a densidade e a regularidade estrutural.

A Figura 64 ilustra o gráfico de dispersão obtido através da projeção ACP. Na Figura 64a temos o primeiro componente principal C_1 no eixo x e o segundo C_2 no eixo y . Juntos, os dois eixos acumulam 58,46% da variância total dos dados originais. Neste

gráfico, há pequenas sobreposição entre as classes teciduais estudadas. Podemos observar também que os *clusters* referentes as classes Colateral Fechado Xilema e Anfisal primário xilema, estão totalmente segregadas nos dois eixos. Na Figura 64b, C_1 é o eixo x , enquanto o terceiro componente principal C_3 é o eixo y . Juntos, os dois eixos acumulam 35,71% da variância total dos dados originais. Neste gráfico, podemos observar a segregação das classe Anficrival floema e Colateral fechado xilema.

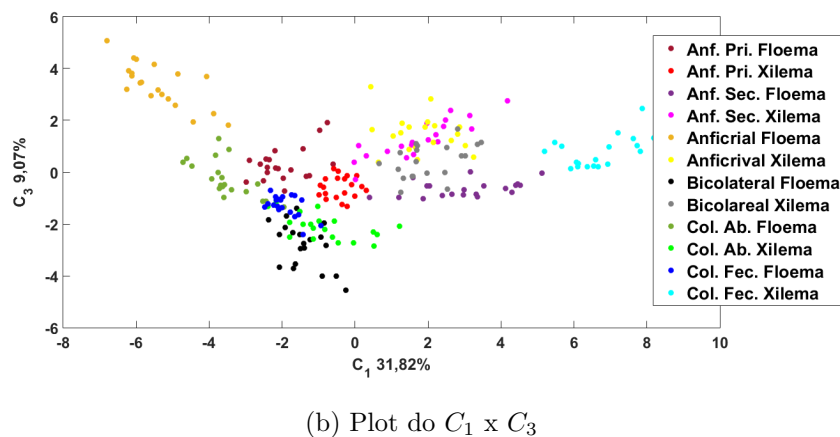
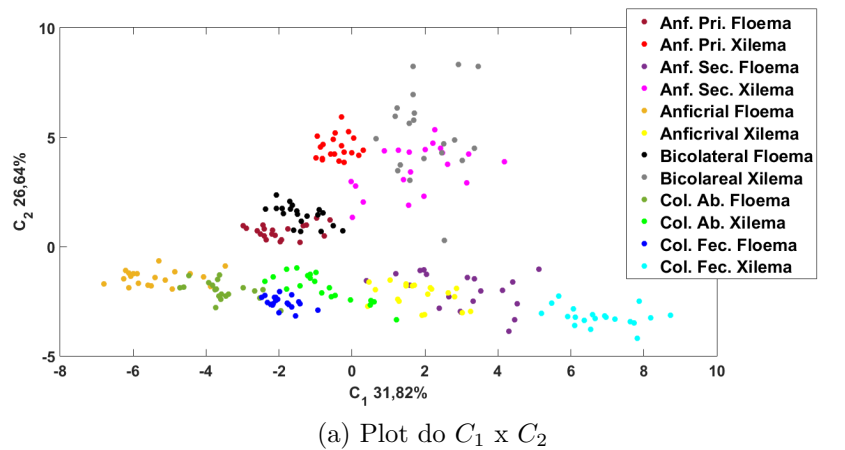
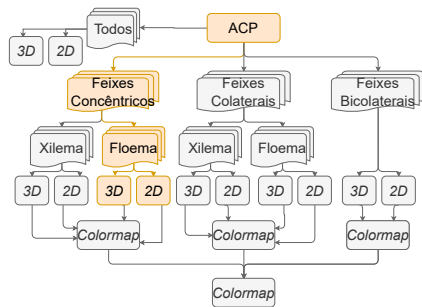


Figura 64 – Gráfico de dispersão obtido através da projeção ACP para todos os feixes vasculares estudados. (a) Ilustra o primeiro componente principal C_1 no eixo x e o segundo C_2 no eixo y ; (b) Ilustra o primeiro componente principal C_1 no eixo x e o terceiro C_3 no eixo y .

Fonte: Autor

5.1.3.2 Feixes vasculares do tipo concêntricos: Floema



Para essa etapa, foi realizada a ACP considerando o floema dos feixes: anfigasal primário, anfigasal secundário e anfigival.

A Tabela 21 mostra que quando consideramos apenas os tecidos do floema dos feixes vasculares concêntricos (Anfigasal primário floema, Anfigasal secundário floema e Anfigival floema) os três primeiros componentes da projeção de ACP somam juntos 81% da variância dos dados originais, com o primeiro componente principal tendo 59,0%, o segundo 17,4% e o terceiro 4,6%. A Tabela 21 também mostra o peso das 33 medidas analisadas e, em negrito, aquelas que tiveram maior relevância no cálculo dos novos eixos. Para o primeiro eixo as medidas com os maiores pesos são: Área média, resíduo médio, diâmetro médio, diâmetro equivalente médio, Uniformidade espacial 30 e densidade radial 75. Para o segundo eixo as medidas que apresentaram o maior peso são: desvio padrão área, desvio padrão diâmetro, desvio padrão diâmetro equivalente, circularidade *isoline* 65 e lacunaridade 1.

Tabela 21 – Floema dos feixes concêntricos: Pesos absolutos para a projeção de ACP para os componentes C_1 , C_2 e C_3 . Os valores em negrito indicam os maiores pesos.

Atributos	C_1 (59.0%)	C_2 (17.4%)	C_3 (4.6%)
Área média	0.036	0.031	0.018
Desvio Padrão área	0.024	0.063	0.025
Resíduo média	0.037	0.020	0.011
Desvio Padrão resíduo	0.024	0.041	0.010
Diâmetro médio	0.038	0.025	0.000
Desvio Padrão diâmetro	0.017	0.071	0.031
Diâmetro equivalente médio	0.037	0.027	0.006
Desvio Padrão diâmetro equivalente	0.017	0.071	0.024
Elongação média	0.028	0.003	0.070
Desvio Padrão elongação	0.025	0.015	0.053
Poligonalidade média	0.031	0.050	0.016
Desvio Padrão poligonalidade	0.025	0.007	0.043
Uniformidade espacial 30	0.036	0.036	0.017
Uniformidade espacial 48	0.035	0.040	0.010
Uniformidade espacial 65	0.035	0.040	0.006
Uniformidade espacial 83	0.035	0.036	0.001
Uniformidade espacial 100	0.035	0.029	0.004
Circularidade <i>isoline</i> 30	0.032	0.023	0.032
Circularidade <i>isoline</i> 48	0.028	0.049	0.008
Circularidade <i>isoline</i> 65	0.026	0.053	0.013
Circularidade <i>isoline</i> 83	0.031	0.025	0.001
Circularidade <i>isoline</i> 100	0.027	0.003	0.018
Densidade radial 25	0.034	0.030	0.043
Densidade radial 75	0.037	0.025	0.033
Densidade radial 125	0.032	0.035	0.007
Densidade radial 175	0.029	0.032	0.056
Densidade radial 225	0.035	0.007	0.008
Densidade radial 275	0.035	0.008	0.037
Lacunaridade 1	0.010	0.058	0.049
Lacunaridade 6	0.031	0.033	0.082
Lacunaridade 11	0.034	0.011	0.085
Lacunaridade 16	0.034	0.001	0.089
Lacunaridade 21	0.033	0.004	0.092

Dos 11 atributos com maiores pesos no primeiro e no segundo eixo, 7 são referentes as medidas de forma, 3 estão relacionados com a densidade e apenas 2 são do grupo da regularidade estrutural. Vale ressaltar que o primeiro eixo (eixo com maior variância dos dados originais) teve como atributos de maiores pesos medidas referente ao formato celular e a densidade.

A Figura 65 ilustra o gráfico de dispersão obtido através da projeção ACP. Na Figura 65a temos o primeiro componente principal C_1 no eixo x e o segundo C_2 no eixo y . Juntos, os dois eixos acumulam 76,4% da variância total dos dados originais. Neste gráfico,

as classes teciduais estão bem separadas em relação aos dois eixos. Podemos observar também que os *clusters* referentes as classes anfigival primário e anfigival estão menos dispersas se comparadas ao *cluster* da classe anfigival secundário. Na Figura 65b, C_1 é o eixo x , enquanto o terceiro componente principal C_3 é o eixo y . Juntos, os dois eixos acumulam 63,6% da variância total dos dados originais. Neste gráfico, podemos observar uma menor distância entre as classes anfigival primário e a anfigival secundário.

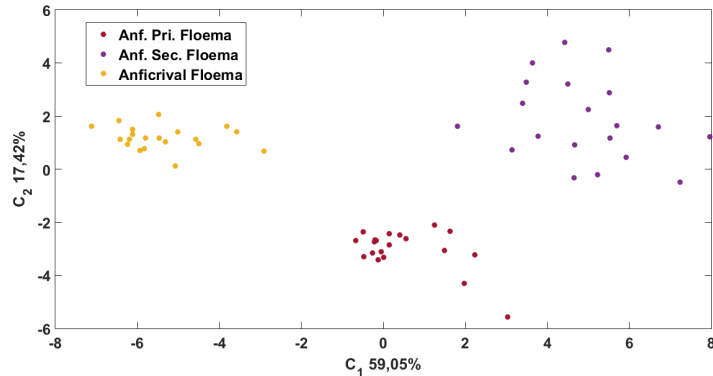
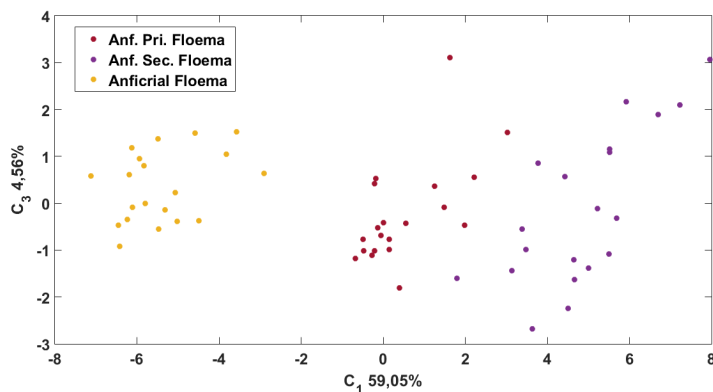
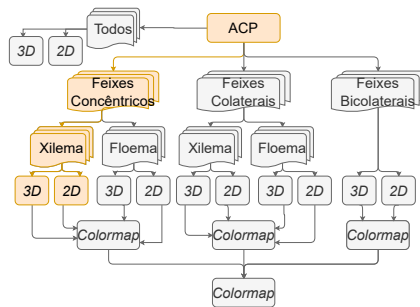
(a) Plot do C_1 x C_2 (b) Plot do C_1 x C_3

Figura 65 – Gráfico de dispersão obtido através da projeção ACP para os floemas dos feixes vasculares do tipo concêntrico. (a) Ilustra o primeiro componente principal C_1 no eixo x e o segundo C_2 no eixo y ; (b) Ilustra o primeiro componente principal C_1 no eixo x e o terceiro C_3 no eixo y .

Fonte: Autor

5.1.3.3 Feixes vasculares do tipo concêntricos: Xilema



Para essa etapa, foi realizada a ACP considerando o xilema dos feixes: anfigasal primário, anfigasal secundário e anfigicrival.

A Tabela 22 mostra que quando consideramos apenas os tecidos do xilema dos feixes vasculares concêntricos (Anfigasal primário xilema, Anfigasal secundário xilema e Anfigicrival xilema) os três primeiros componentes da projeção de ACP somam juntos 76,2% da variância dos dados originais, com o primeiro componente principal tendo 54,5%, o segundo 16,2% e o terceiro 5,5%. A Tabela 22 também mostra o peso das 33 medidas analisadas e, em negrito, aquelas que tiveram maior relevância no cálculo dos novos eixos. Para o primeiro eixo as medidas com os maiores pesos são: Área média, desvio padrão área, resíduo médio, desvio padrão resíduo, diâmetro médio, desvio padrão diâmetro médio, diâmetro equivalente médio, desvio padrão diâmetro equivalente, alongação média, Uniformidade espacial com sigmas 30, 48, 65, 83 e 100, densidade radial com raios de 75, 125, 225 e 275, poligonalidade média e lacunaridade 1. Para o segundo eixo as medidas que apresentaram o maior peso são: desvio padrão alongação e todas as lacunaridade com raios 6, 11, 16 e 21.

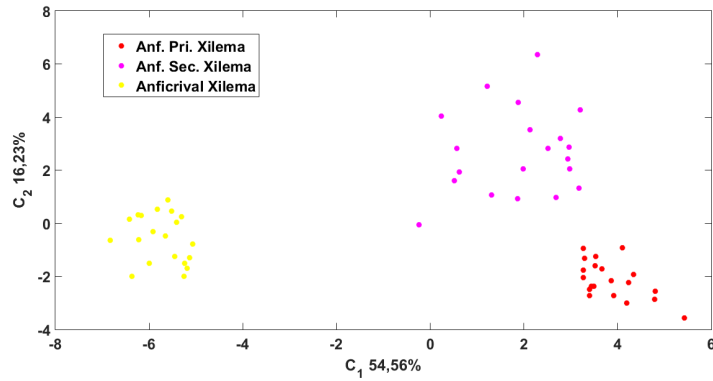
Tabela 22 – Xilema dos feixes concêntricos: Pesos absolutos para a projeção de ACP para os componentes C_1 , C_2 e C_3 . Os valores em negrito indicam os maiores pesos.

Atributos	C_1 (54.5%)	C_2 (16.2%)	C_3 (5.5%)
Área média	0.044	0.025	0.001
Desvio Padrão área	0.042	0.027	0.019
Resíduo média	0.045	0.001	0.010
Desvio Padrão resíduo	0.041	0.006	0.032
Diâmetro médio	0.046	0.009	0.001
Desvio Padrão diâmetro	0.043	0.007	0.038
Diâmetro equivalente médio	0.045	0.014	0.004
Desvio Padrão diâmetro equivalente	0.042	0.015	0.030
Elongação média	0.038	0.038	0.018
Desvio Padrão elongação	0.029	0.061	0.033
Poligonalidade média	0.045	0.004	0.020
Desvio Padrão poligonalidade	0.008	0.035	0.028
Uniformidade espacial 30	0.045	0.015	0.013
Uniformidade espacial 48	0.045	0.007	0.017
Uniformidade espacial 65	0.045	0.001	0.020
Uniformidade espacial 83	0.045	0.007	0.021
Uniformidade espacial 100	0.045	0.009	0.022
Circularidade <i>isoline</i> 30	0.008	0.035	0.089
Circularidade <i>isoline</i> 48	0.007	0.040	0.015
Circularidade <i>isoline</i> 65	0.017	0.029	0.076
Circularidade <i>isoline</i> 83	0.000	0.040	0.114
Circularidade <i>isoline</i> 100	0.007	0.010	0.144
Densidade radial 25	0.034	0.020	0.004
Densidade radial 75	0.043	0.014	0.006
Densidade radial 125	0.041	0.013	0.034
Densidade radial 175	0.017	0.016	0.039
Densidade radial 225	0.034	0.029	0.025
Densidade radial 275	0.037	0.006	0.008
Lacunaridade 1	0.037	0.057	0.018
Lacunaridade 6	0.009	0.102	0.024
Lacunaridade 11	0.001	0.104	0.025
Lacunaridade 16	0.006	0.103	0.026
Lacunaridade 21	0.008	0.100	0.027

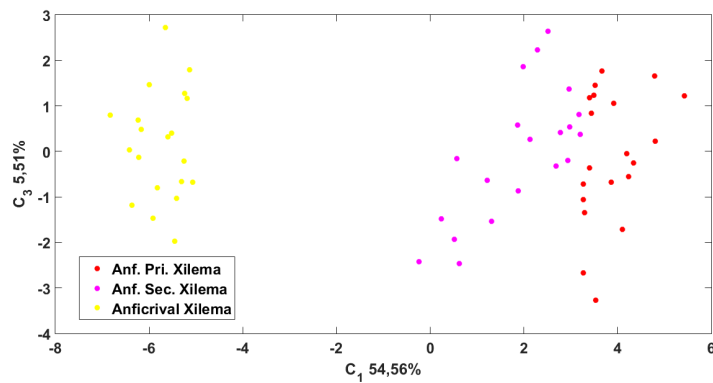
Dos 25 atributos com maiores pesos no primeiro e no segundo eixo, 10 são referentes as medidas de forma, 8 estão relacionados com a densidade e 6 são do grupo da regularidade estrutural.

A Figura 66 ilustra o gráfico de dispersão obtido através da projeção ACP. Na Figura 66a temos o primeiro componente principal C_1 no eixo x e o segundo C_2 no eixo y . Juntos, os dois eixos acumulam 70,7% da variância total dos dados originais. Neste gráfico, as classes teciduais estão bem separadas em relação aos dois eixos. Podemos observar também que os *clusters* referentes as classes anfigasal primário e anficrival estão menos

dispersas se comparadas ao *cluster* da classe anfigvasal secundário. Na Figura 66b, C_1 é o eixo x , enquanto o terceiro componente principal C_3 é o eixo y . Juntos, os dois eixos acumulam 60% da variância total dos dados originais. Neste gráfico, podemos observar uma menor distância entre as classes anfigvasal primário e a anfigvasal secundário.



(a) Plot do C_1 x C_2

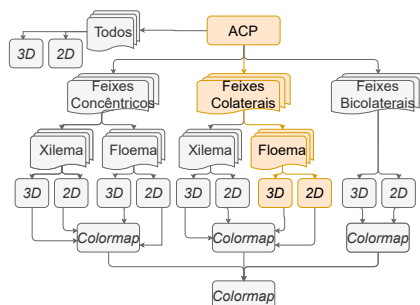


(b) Plot do C_1 x C_3

Figura 66 – Gráfico de dispersão obtido através da projeção ACP para os xilemas dos feixes vasculares do tipo concêntrico. (a) Ilustra o primeiro componente principal C_1 no eixo x e o segundo C_2 no eixo y ; (b) Ilustra o primeiro componente principal C_1 no eixo x e o terceiro C_3 no eixo y .

Fonte: Autor

5.1.3.4 Feixes vasculares do tipo colateral: Floema



Para essa etapa, foi realizada a ACP considerando o floema dos feixes: colateral aberto e colateral fechado.

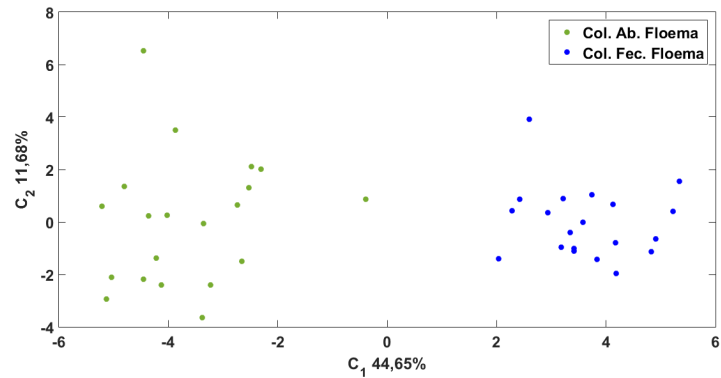
A Tabela 23 mostra que quando consideramos apenas os tecidos do floema dos feixes vasculares colateral (colateral aberto floema, colateral fechado floema) os três primeiros componentes da projeção de ACP somam juntos 64,6% da variância dos dados originais, com o primeiro componente principal tendo 44,7%, o segundo 11,7% e o terceiro 8,2%. A Tabela 23 também mostra o peso das 33 medidas analisadas e, em negrito, aquelas que tiveram maior relevância no cálculo dos novos eixos. Para o primeiro eixo as medidas com os maiores pesos são: Área média, desvio padrão área, resíduo médio, desvio padrão resíduo, diâmetro médio, diâmetro equivalente médio, Uniformidade espacial com sigmas 30, e 48, circularidade *isoline* 30, 48 e 65, densidade radial com raios de 75, 125 e 275 e lacunaridade 1 e 6. Para o segundo eixo as medidas que apresentaram o maior peso são: desvio padrão diâmetro, poligonalidade média, Uniformidade espacial com sigma 65, 83 e 100 e densidade radial com raio 25.

Tabela 23 – Floema dos feixes colaterais: Pesos absolutos para a projeção de ACP para os componentes C_1 , C_2 e C_3 . Os valores em negrito indicam os maiores pesos.

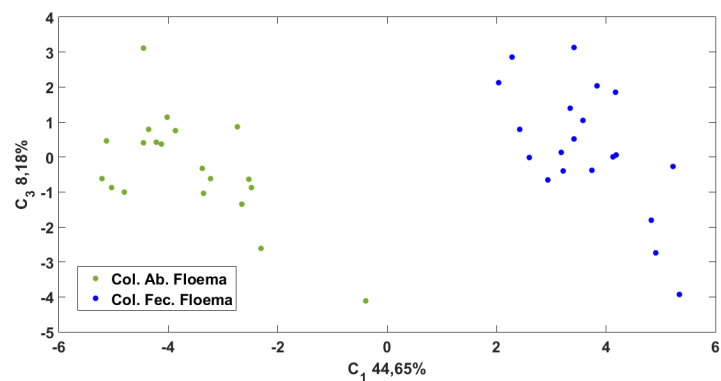
Atributos	C_1 (44.7%)	C_2 (11.7%)	C_3 (8.2%)
Área média	0.050	0.005	0.002
Desvio Padrão área	0.042	0.023	0.024
Resíduo média	0.050	0.003	0.012
Desvio Padrão resíduo	0.043	0.004	0.005
Diâmetro médio	0.050	0.004	0.013
Desvio Padrão diâmetro	0.015	0.064	0.025
Diâmetro equivalente médio	0.050	0.002	0.006
Desvio Padrão diâmetro equivalente	0.032	0.049	0.021
Elongação média	0.003	0.019	0.077
Desvio Padrão elongação	0.003	0.023	0.054
Poligonalidade média	0.005	0.093	0.003
Desvio Padrão poligonalidade	0.011	0.034	0.095
Uniformidade espacial 30	0.046	0.037	0.000
Uniformidade espacial 48	0.041	0.063	0.003
Uniformidade espacial 65	0.006	0.109	0.006
Uniformidade espacial 83	0.029	0.084	0.003
Uniformidade espacial 100	0.038	0.064	0.002
Circularidade <i>isoline</i> 30	0.039	0.036	0.014
Circularidade <i>isoline</i> 48	0.047	0.012	0.005
Circularidade <i>isoline</i> 65	0.041	0.013	0.025
Circularidade <i>isoline</i> 83	0.033	0.005	0.008
Circularidade <i>isoline</i> 100	0.005	0.008	0.023
Densidade radial 25	0.010	0.066	0.048
Densidade radial 75	0.043	0.001	0.061
Densidade radial 125	0.038	0.031	0.056
Densidade radial 175	0.000	0.041	0.039
Densidade radial 225	0.034	0.012	0.008
Densidade radial 275	0.046	0.033	0.031
Lacunaridade 1	0.032	0.017	0.035
Lacunaridade 6	0.045	0.004	0.010
Lacunaridade 11	0.042	0.006	0.051
Lacunaridade 16	0.022	0.017	0.116
Lacunaridade 21	0.009	0.015	0.121

Dos 22 atributos com maiores pesos no primeiro e no segundo eixo, 7 são referentes as medidas de forma, 12 estão relacionados com a densidade e 3 são do grupo da regularidade estrutural.

A Figura 67 ilustra o gráfico de dispersão obtido através da projeção ACP. Na Figura 67a temos o primeiro componente principal C_1 no eixo x e o segundo C_2 no eixo y . Juntos, os dois eixos acumulam 56,4% da variância total dos dados originais. Na Figura 67b, C_1 é o eixo x , enquanto o terceiro componente principal C_3 é o eixo y . Juntos, os dois eixos acumulam 52,9% da variância total dos dados originais. Em ambos os gráficos, as



(a) Plot do C_1 x C_2



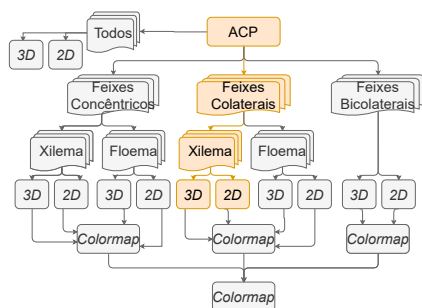
(b) Plot do C_1 x C_3

Figura 67 – Gráfico de dispersão obtido através da projeção ACP para os floemas dos feixes vasculares do tipo colateral. (a) Ilustra o primeiro componente principal C_1 no eixo x e o segundo C_2 no eixo y ; (b) Ilustra o primeiro componente principal C_1 no eixo x e o terceiro C_3 no eixo y .

Fonte: Autor

classes teciduais estão bem separadas em relação aos dois eixos porém, elas apresentam grande dispersão ao longo dos eixos.

5.1.3.5 Feixes vasculares do tipo colateral: Xilema



Para essa etapa, foi realizada a ACP considerando o xilema dos feixes: colateral aberto e colateral fechado.

A Tabela 24 mostra que quando consideramos apenas os tecidos do floema dos feixes vasculares colateral (colateral aberto xilema, colateral fechado xilema) os três primeiros

componentes da projeção de ACP somam juntos 92% da variância dos dados originais, com o primeiro componente principal tendo 82,2%, o segundo 5,5% e o terceiro 4,3%. A Tabela 24 também mostra o peso das 33 medidas analisadas e, em negrito, aquelas que tiveram maior relevância no cálculo dos novos eixos. Para o primeiro eixo as medidas com os maiores pesos são: Área média, desvio padrão área, resíduo médio, desvio padrão resíduo, diâmetro médio, desvio padrão diâmetro, diâmetro equivalente médio, desvio padrão diâmetro equivalente, Uniformidade espacial com sigmas 30, 48, 65, 83 e 100, poligonalidade média e lacunaridade 1, 6, 11, 16 e 21. Para o segundo eixo as medidas que apresentaram o maior peso são: circularidade *isoline* 30, 48, 65, 83 e 100 e densidade radial com raio 125, 225 e 275.

Tabela 24 – Xilema dos feixes colaterais: Pesos absolutos para a projeção de ACP para os componentes C_1 , C_2 e C_3 . Os valores em negrito indicam os maiores pesos.

Atributos	C_1 (82, 2%)	C_2 (5.5%)	C_3 (4.3%)
Área média	0.039	0.016	0.011
Desvio Padrão área	0.039	0.016	0.016
Resíduo média	0.039	0.018	0.000
Desvio Padrão resíduo	0.038	0.017	0.009
Diâmetro médio	0.038	0.021	0.007
Desvio Padrão diâmetro	0.039	0.021	0.017
Diâmetro equivalente médio	0.039	0.019	0.006
Desvio Padrão diâmetro equivalente	0.039	0.018	0.015
Elongação média	0.025	0.007	0.019
Desvio Padrão elongação	0.030	0.030	0.003
Poligonalidade média	0.037	0.003	0.025
Desvio Padrão poligonalidade	0.031	0.008	0.034
Uniformidade espacial 30	0.038	0.019	0.019
Uniformidade espacial 48	0.039	0.017	0.019
Uniformidade espacial 65	0.039	0.010	0.016
Uniformidade espacial 83	0.039	0.001	0.013
Uniformidade espacial 100	0.039	0.007	0.010
Circularidade <i>isoline</i> 30	0.004	0.073	0.180
Circularidade <i>isoline</i> 48	0.005	0.079	0.151
Circularidade <i>isoline</i> 65	0.013	0.142	0.054
Circularidade <i>isoline</i> 83	0.019	0.138	0.060
Circularidade <i>isoline</i> 100	0.031	0.066	0.043
Densidade radial 75	0.005	0.047	0.024
Densidade radial 125	0.027	0.053	0.005
Densidade radial 175	0.032	0.005	0.008
Densidade radial 225	0.019	0.082	0.057
Densidade radial 275	0.032	0.056	0.013
Lacunaridade 1	0.036	0.006	0.034
Lacunaridade 6	0.038	0.002	0.016
Lacunaridade 11	0.037	0.003	0.035
Lacunaridade 16	0.037	0.000	0.041
Lacunaridade 21	0.038	0.000	0.040

Dos 27 atributos com maiores pesos no primeiro e no segundo eixo, 8 são referentes as medidas de forma, 13 estão relacionados com a densidade e 6 são do grupo da regularidade estrutural.

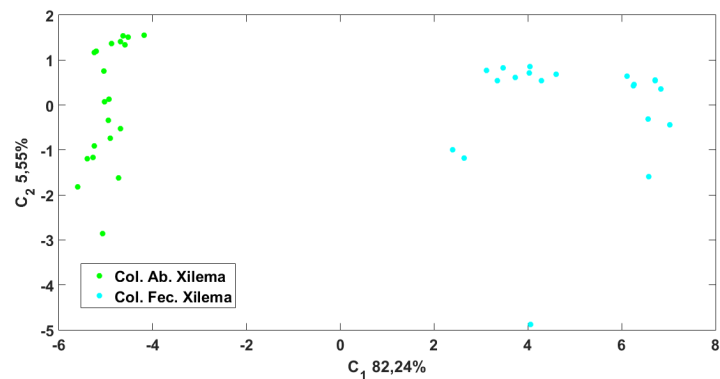
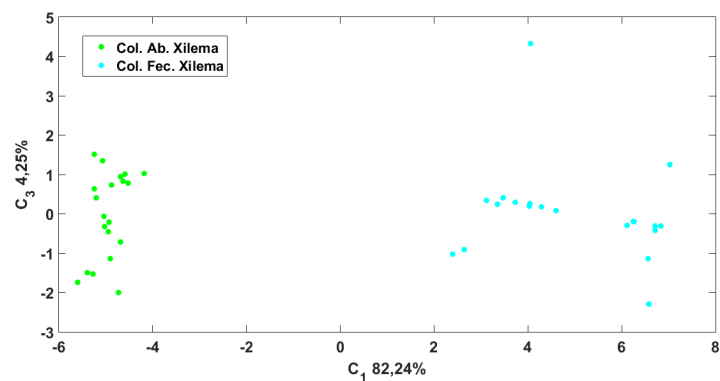
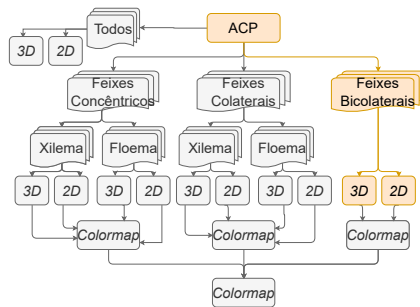
(a) Plot do C_1 x C_2 (b) Plot do C_1 x C_3

Figura 68 – Gráfico de dispersão obtido através da projeção ACP para os xilemas dos feixes vasculares do tipo colateral. (a) Ilustra o primeiro componente principal C_1 no eixo x e o segundo C_2 no eixo y ; (b) Ilustra o primeiro componente principal C_1 no eixo x e o terceiro C_3 no eixo y .

Fonte: Autor

A Figura 68 ilustra o gráfico de dispersão obtido através da projeção ACP. Na Figura 68a temos o primeiro componente principal C_1 no eixo x e o segundo C_2 no eixo y . Juntos, os dois eixos acumulam 87,79% da variância total dos dados originais. Na Figura 68b, C_1 é o eixo x , enquanto o terceiro componente principal C_3 é o eixo y . Juntos, os dois eixos acumulam 86,49% da variância total dos dados originais. Em ambos os gráficos, as classes teciduais estão bem separadas em relação aos dois eixos porém, elas apresentam grande dispersão ao longo dos eixos.

5.1.3.6 Feixes vasculares do tipo biclateral



Para essa etapa, foi realizada a ACP considerando o feixe biclateral. Como não há subdivisão, o feixe biclateral foi estudado apenas em relação ao xilema e o floema.

A Tabela 25 mostra que quando consideramos apenas os tecidos dos feixes vasculares biclateral (biclateral floema, biclateral xilema) os três primeiros componentes da projeção de ACP somam juntos 77,3% da variância dos dados originais, com o primeiro componente principal tendo 57,3%, o segundo 13,2% e o terceiro 6,8%. A Tabela 25 também mostra o peso das 33 medidas analisadas e, em negrito, aquelas que tiveram maior relevância no cálculo dos novos eixos. Para o primeiro eixo as medidas com os maiores pesos são: Área média, desvio padrão área, resíduo médio, desvio padrão resíduo, desvio padrão diâmetro, desvio padrão diâmetro equivalente, Uniformidade espacial com sigma 30, densidade radial com raio 175 e lacunaridade 21. Para o segundo eixo as medidas que apresentaram o maior peso são: Uniformidade espacial com sigma 65, 83 e 100.

Tabela 25 – Feixes bicolaterais: Pesos absolutos para a projeção de ACP para os componentes C_1 , C_2 e C_3 . Os valores em negrito indicam os maiores pesos.

Atributos	C_1 (57.3%)	C_2 (13.2%)	C_3 (6.8%)
Área média	0.038	0.036	0.004
Desvio Padrão área	0.038	0.031	0.002
Resíduo média	0.036	0.000	0.031
Desvio Padrão resíduo	0.037	0.009	0.013
Diâmetro médio	0.034	0.046	0.000
Desvio Padrão diâmetro	0.040	0.024	0.004
Diâmetro equivalente médio	0.034	0.044	0.000
Desvio Padrão diâmetro equivalente	0.040	0.022	0.006
Elongação média	0.032	0.045	0.003
Desvio Padrão elongação	0.032	0.040	0.013
Poligonalidade média	0.032	0.031	0.013
Desvio Padrão poligonalidade	0.027	0.034	0.039
Uniformidade espacial 30	0.039	0.024	0.017
Uniformidade espacial 48	0.036	0.039	0.029
Uniformidade espacial 65	0.031	0.056	0.045
Uniformidade espacial 83	0.021	0.070	0.064
Uniformidade espacial 100	0.007	0.075	0.076
Circularidade <i>isoline</i> 30	0.024	0.016	0.043
Circularidade <i>isoline</i> 48	0.001	0.047	0.034
Circularidade <i>isoline</i> 65	0.020	0.045	0.078
Circularidade <i>isoline</i> 83	0.021	0.036	0.105
Circularidade <i>isoline</i> 100	0.025	0.025	0.104
Densidade radial 25	0.036	0.001	0.028
Densidade radial 75	0.033	0.007	0.050
Densidade radial 125	0.035	0.000	0.028
Densidade radial 175	0.037	0.004	0.016
Densidade radial 225	0.030	0.014	0.041
Densidade radial 275	0.033	0.026	0.031
Lacunaridade 1	0.015	0.047	0.007
Lacunaridade 6	0.033	0.038	0.009
Lacunaridade 11	0.034	0.029	0.017
Lacunaridade 16	0.035	0.021	0.023
Lacunaridade 21	0.036	0.019	0.026

Dos 12 atributos com maiores pesos no primeiro e no segundo eixo, 6 são referentes as medidas de forma, 5 estão relacionados com a densidade e apenas 1 é do grupo da regularidade estrutural.

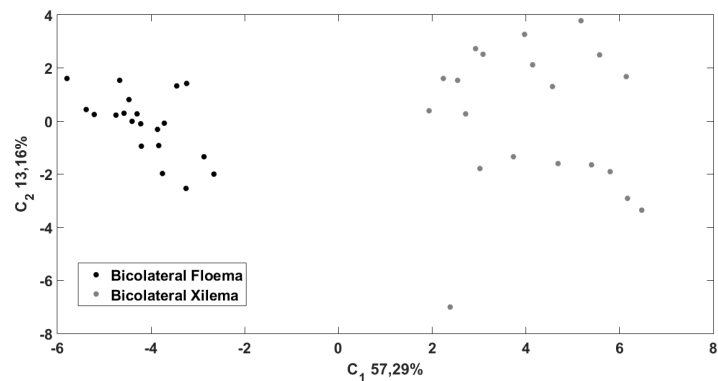
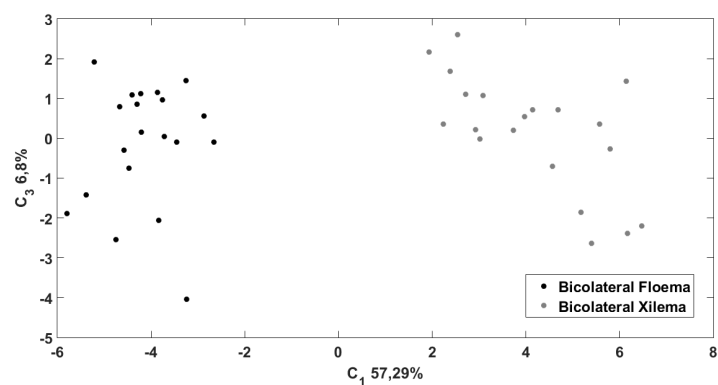
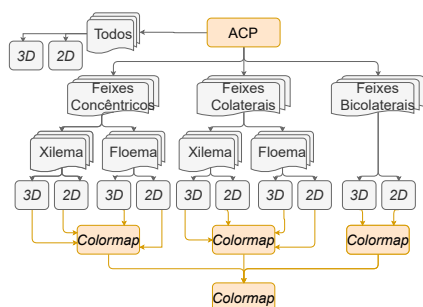
(a) Plot do C_1 x C_2 (b) Plot do C_1 x C_3

Figura 69 – Gráfico de dispersão obtido através da projeção ACP para os feixes vasculares do tipo bicolateral. (a) Ilustra o primeiro componente principal C_1 no eixo x e o segundo C_2 no eixo y ; (b) Ilustra o primeiro componente principal C_1 no eixo x e o terceiro C_3 no eixo y .

Fonte: Autor

A Figura 69 ilustra o gráfico de dispersão obtido através da projeção ACP. Na Figura 69a temos o primeiro componente principal C_1 no eixo x e o segundo C_2 no eixo y . Juntos, os dois eixos acumulam 70,5% da variância total dos dados originais. Na Figura 69b, C_1 é o eixo x , enquanto o terceiro componente principal C_3 é o eixo y . Juntos, os dois eixos acumulam 64,1% da variância total dos dados originais. Em ambos os gráficos, as classes teciduais estão bem separadas em relação aos dois eixos porém, elas apresentam grande dispersão ao longo dos eixos.

5.1.3.7 Considerações gerais



Partindo como premissa os resultados apresentados para a análise de componentes principais dos feixes vasculares vegetais podemos buscar entender quais medidas foram relevantes ao longo dessa análise. Para isso, os resultados foram concatenados em um *colormap*, Figura 70, no qual a escala de cor varia do vermelho (menos importante) para o verde (mais importante).

A importância de cada atributo foi estabelecida com base na recorrência desse atributo nos resultados apresentados. Essa abordagem foi feita para cada tipo de feixe levando em consideração os dois de tecidos vasculares que compõem esses feixes, o floema e o xilema. Ao final, os resultados foram agrupados em uma escala na tentativa de elencar quais medidas foram importantes na análise individual.

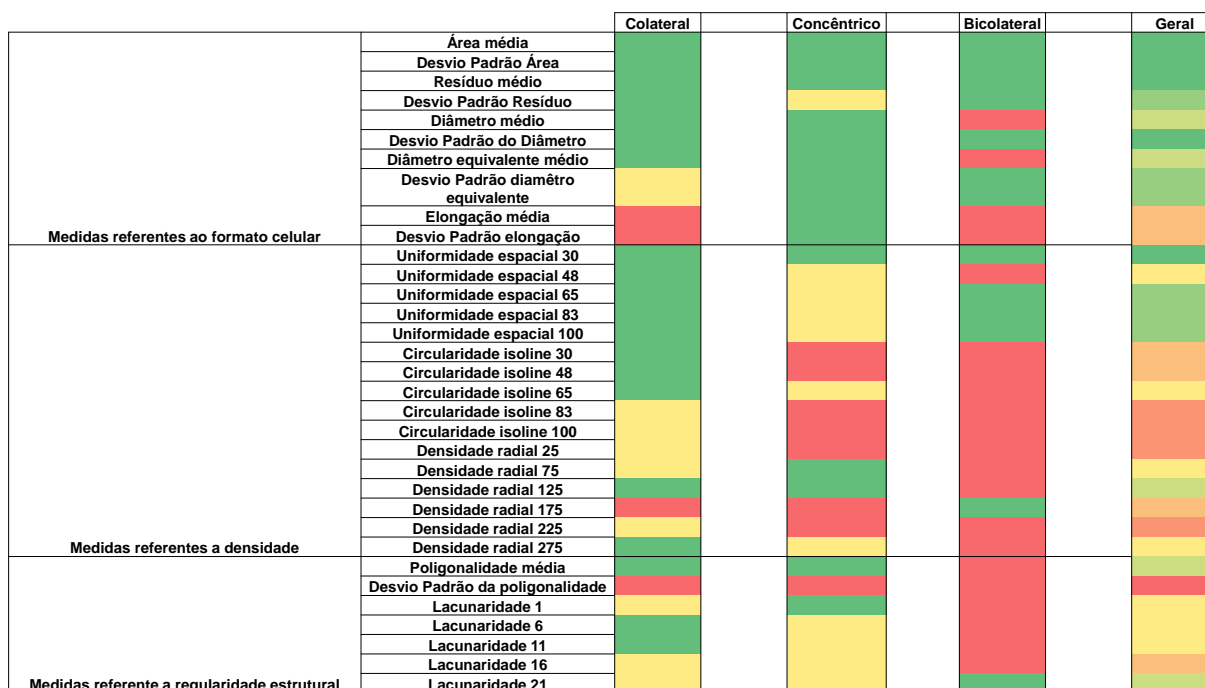


Figura 70 – *Colormap* representando a importância de cada atributo considerando a análise ACP. A escala varia do vermelho (menos recorrente, logo menos importante) ao verde (mais recorrente, logo mais importante).

Fonte: Autor

Ao analisar a Figura 70 podemos observar que a distribuição das medidas ao longo do tipo de feixe vascular é similar pois, todos os feixes vasculares estudados foram bem caracterizados pelos atributos referente aos grupos formato celular e a densidade e regularidade estrutural.

De maneira geral, podemos notar que a distribuição geral dos atributos ao longo da análise de componente principal engloba os quatro grupos de medidas estudados (formato celular, densidade, regularidade estrutural e multiescala). Assim, podemos concluir que os atributos que foram importantes para caracterizar e segregar os feixes vasculares vegetais considerando apenas essa análise são: Área média, Desvio padrão área, resíduo médio, desvio padrão resíduo, diâmetro médio, desvio padrão diâmetro, diâmetro equivalente médio, desvio padrão diâmetro equivalente, Uniformidade espacial com sigmas 30, 65, 83 e 100, densidade radial 125, poligonalidade média, lacunaridade 21.

Outro ponto relevante que podemos sugerir com base nos resultados da análise de componentes principais tem relação com o efeito multiescala dos atributos. Alguns atributos, como por exemplo, a Uniformidade espacial, a densidade radial e a lacunaridade apresentaram recorrência em escalas específicas. Isso indica que o uso da análise multiescala para extração dos atributos é importante e pode contribuir na caracterização e na segregação de feixes vasculares vegetais.

5.2 Filtragem dos resultados

Com base nos resultados obtidos até aqui, podemos inferir que cada tipo de análise resultou em um sub grupo de atributos. Alguns desses atributos são recorrentes em todas as análises, outros foram resultados isolados. A Figura 71 ilustra os detalhes da filtragem dos resultados. Nela é possível verificar que a filtragem foi feita por meio dos resultados gerais obtidos pelas três abordagens utilizadas nesse trabalho.

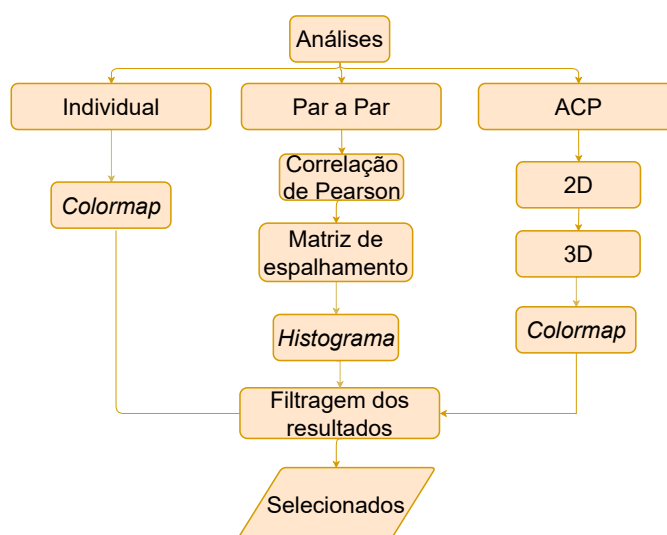


Figura 71 – Detalhamento da de filtragem dos resultados obtidos na etapa de análise desse trabalho.

Fonte: Autor

Buscando inferir quais os atributos mais recorrentes e entender quais medidas foram relevantes ao longo da etapa de análise, os resultados das três abordagens (individual, par a par e análise de componente principal) foram concatenados em um *colormap*, Figura 72, no qual a escala de cor varia do vermelho (menos importante) para o verde (mais importante). A importância de cada atributo foi estabelecida com base na recorrência desse atributo nos resultados apresentados.

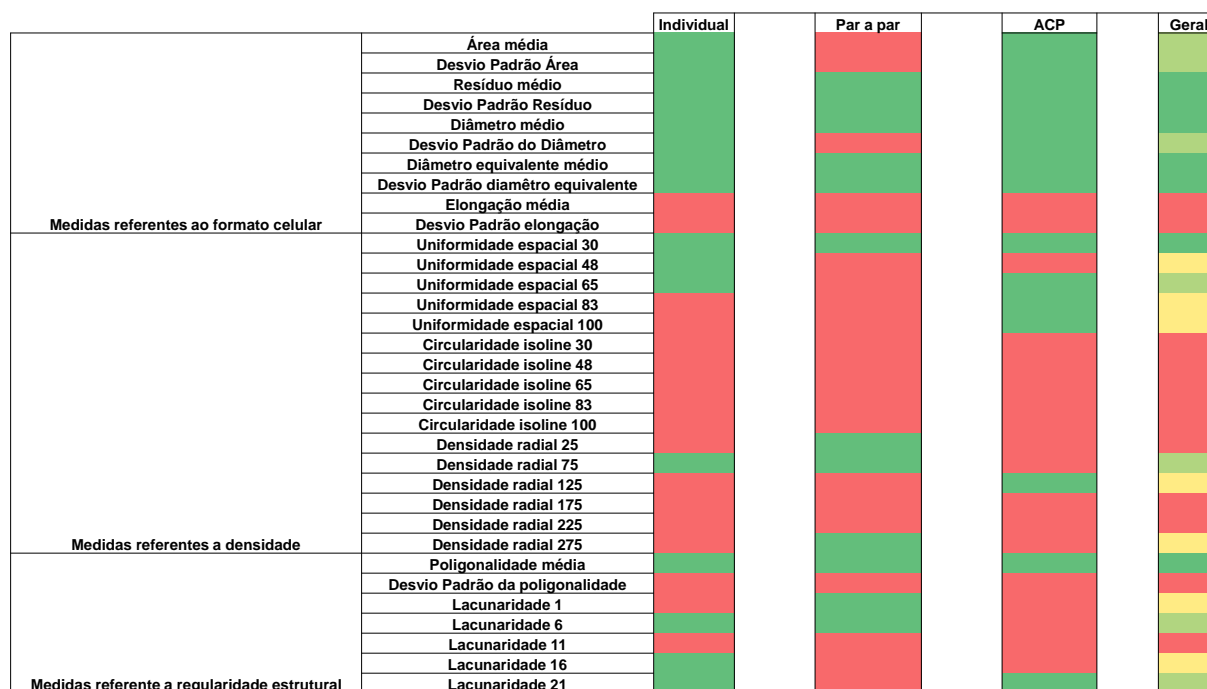


Figura 72 – *Colormap* representando a importância de cada atributo considerando as três abordagens aplicadas. A escala varia do vermelho (menos recorrente, logo menos importante) ao verde (mais recorrente, logo mais importante).

Fonte: Autor

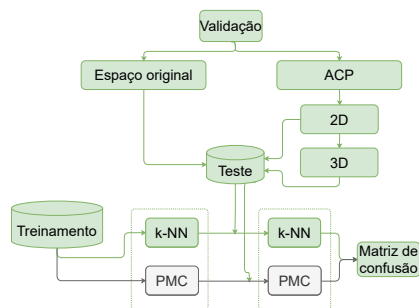
Ao analisar a Figura 72 podemos observar que há 14 atributos que apresentaram relevância em, pelo menos, duas das três abordagens aplicadas ao longo desse trabalho. São eles: área média, desvio padrão área, resíduo médio, desvio padrão resíduo, diâmetro médio, desvio padrão diâmetro, diâmetro equivalente médio, desvio padrão do diâmetro equivalente, uniformidade espacial 30, uniformidade espacial 65, densidade radial com raio 75, poligonalidade média, lacunaridade 6 e 21.

De maneira geral, podemos notar que os atributos selecionados englobam os quatro grupos de medidas estudados (formato celular, densidade, regularidade estrutural e multiescala). Sendo, 8 referente ao formato celular, 3 referentes a densidade e 3 a regularidade estrutural. A multiescala está presente em 5 desses atributos.

5.3 Validação

Nessa seção, iremos apresentar os resultados obtidos na etapa de validação do nosso trabalho. Aqui iremos testar a acurácia de acertos das categorias de atributos extraídos: Forma, Densidade, Regularidade estrutural, Multiescala, *Todos* e *Selecionados*. Vale ressaltar que essa última categoria foi resultado da etapa de análise.

5.3.1 K-Nearest Neighbor(k-NN)



Todos os resultados referentes ao k -NN foram realizados com $k = 3$. A Tabela 26 resume a porcentagem de acerto obtida por meio da abordagem k -NN de todas as categorias estudadas. Vale ressaltar que as categorias foram analisadas tanto no espaço original quanto depois da aplicação do ACP reduzindo a dimensionalidade para $2D$ e $3D$.

Tabela 26 – Acurácia obtida por meio da aplicação da abordagem k -NN frente as categorias estudadas.

Categoria	Acurácia
Forma	66,67%
Forma 2D	45%
Forma 3D	48,33%
Densidade	70%
Densidade 2D	55%
Densidade 3D	65%
Regularidade	63,33%
Regularidade 2D	60%
Regularidade 3D	60%
Multiescala	78,33%
Multiescala 2D	61,67%
Multiescala 3D	75%
Todos	90%
Todos 2D	70%
Todos 3D	78,33%
Selecionados	91,67%
Selecionados 2D	66,67%
Selecionados 3D	73,33%

As matrizes de confusão de cada categoria de atributo estão representadas no Apêndice C desta dissertação.

Para analisar melhor os resultados obtidos pela abordagem k -NN, os valores de acurácia, ou seja, de acerto de cada categoria foi plotado em um gráfico de barras. Esse gráfico de barras está mostrado na Figura 73 e cada barra representa uma categoria em um de seus três espaços amostrais (original, ACP (ou do inglês PCA) 2D, ACP (ou do inglês PCA) 3D).

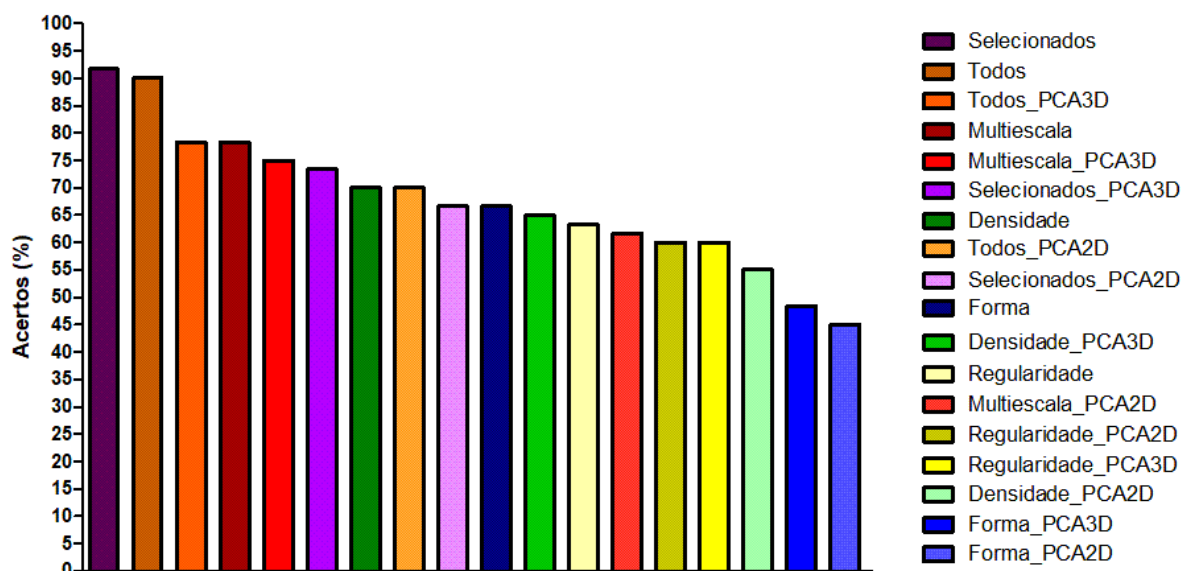


Figura 73 – Gráfico de barras com os acertos da abordagem k -NN para as categorias estudadas. Cada barra representa uma categoria em um de seus três espaços amostrais (original, ACP (ou do inglês PCA) 2D, ACP (ou do inglês PCA) 3D).

Fonte: Autor

Sendo assim, podemos inferir que, por meio da abordagem k -NN, a categoria de atributos que apresentou maior acurácia (A_c) frente a todos os tecidos vegetais estudados foi a categoria *Selecionados* em seu espaço original de 14 dimensões ($A_c = 91, 67\%$). A categoria *Todos* aparece em segundo lugar com uma porcentagem de acerto de 90% quando foi estudada em sua dimensão original, ou seja, 33 dimensões e em terceiro lugar quando foi estudada tomando os três componentes principais ($A_c = 78, 33\%$). Com essa mesma porcentagem de acerto temos a categoria *Multiescala* em seu espaço original, ou seja, 21 dimensões. Essa categoria também ocupa a quarta posição com $A_c = 75\%$ porém, aqui, ela foi estudada tomando os três componentes principais. A categoria *Selecionados* quando estudada tomando os três componentes principais ocupa a quinta posição com $A_c = 73, 33\%$.

Vale ressaltar que quando os feixes vasculares foram validados pela abordagem k -NN, os atributos referente a Forma ocuparam a décima posição quando estudados em seu espaço original de 10 dimensões e obtiveram $A_c = 66, 67\%$. Já quando essa categoria

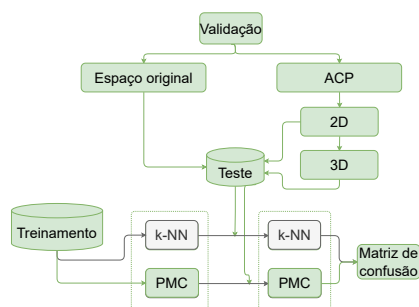
foi estudada após aplicação da análise *ACP*, os erros de classificação ficaram ainda maiores e, conseqüentemente, a acurácia menor. Quando a categoria Forma foi estudada tomando os três componentes principais ela obteve $A_c = 48,33\%$ e ocupou a penúltima posição do *ranking*. Quando estudada tomando os dois componentes principais ela obteve $A_c = 45\%$ e ocupou a última posição do *ranking*.

5.3.2 Perceptron Multicamada (*PMC*)

A primeira etapa dessa abordagem buscou elencar entre as 9 topologias escolhidas qual foi a que melhor classificou os feixes vasculares vegetais. A Tabela 27 mostra qual foi a melhor topologia para cada categoria de atributos.

Tabela 27 – Topologias selecionadas para cada categoria de atributos e suas respectivas acurácias

Categoria	n_1	n_2	epoca	Acurácia
Forma	11	0	2000	88,33%
Forma 2D	50	30	2000	67,08%
Forma 3D	7	0	2000	78,33%
Densidade	14	0	2000	92,08%
Densidade 2D	7	0	500	58,33%
Densidade 3D	7	0	2000	73,33%
Regularidade	10	20	500	82,92%
Regularidade 2D	50	30	1000	56,25%
Regularidade 3D	50	30	2000	69,17%
Multiescala	16	0	500	98,75%
Multiescala 2D	50	30	1000	77,5%
Multiescala 3D	7	0	500	90,42%
Todos	22	0	500	99,58%
Todos 2D	10	20	500	75,83%
Todos 3D	10	20	500	93,33%
Selecionados	13	0	2000	99,17%
Selecionados 2D	7	0	500	81,25%
Selecionados 3D	50	30	2000	97,5%



Cada topologia acima foi re aplicada com os dados de teste e os resultados apresentados na Tabela 28 resumem a porcentagem de acerto obtida por meio da abordagem *PMC* de todas as categorias estudadas. Vale ressaltar que as categorias foram analisadas tanto no espaço original quanto depois da aplicação do *ACP* reduzindo a dimensionalidade para *2D* e *3D*.

Tabela 28 – Acurácia obtida por meio da aplicação da abordagem *PMC* frente as categorias estudadas.

Categoria	Acurácia
Forma	75%
Forma 2D	35%
Forma 3D	40%
Densidade	76,67%
Densidade 2D	66,67%
Densidade 3D	76,33%
Regularidade	61,73%
Regularidade 2D	63,33%
Regularidade 3D	68,33%
Multiescala	85%
Multiescala 2D	63,337%
Multiescala 3D	81,67%
Todos	91,67%
Todos 2D	80%
Todos 3D	86,67%
Selecionados	91,67%
Selecionados 2D	80%
Selecionados 3D	71,67%

As matrizes de confusão de cada categoria de atributo estão representadas no Apêndice D desta dissertação.

Para analisar melhor os resultados obtidos pela abordagem *PMC*, os valores de acurácia, ou seja, de acerto de cada categoria foram plotados em um gráfico de barras. Esse gráfico de barras está mostrado na Figura 74 e cada barra representa uma categoria em um de seus três espaços amostrais (original, *ACP* (ou do inglês *PCA*) 2D, *ACP* (ou do inglês *PCA*) 3D).

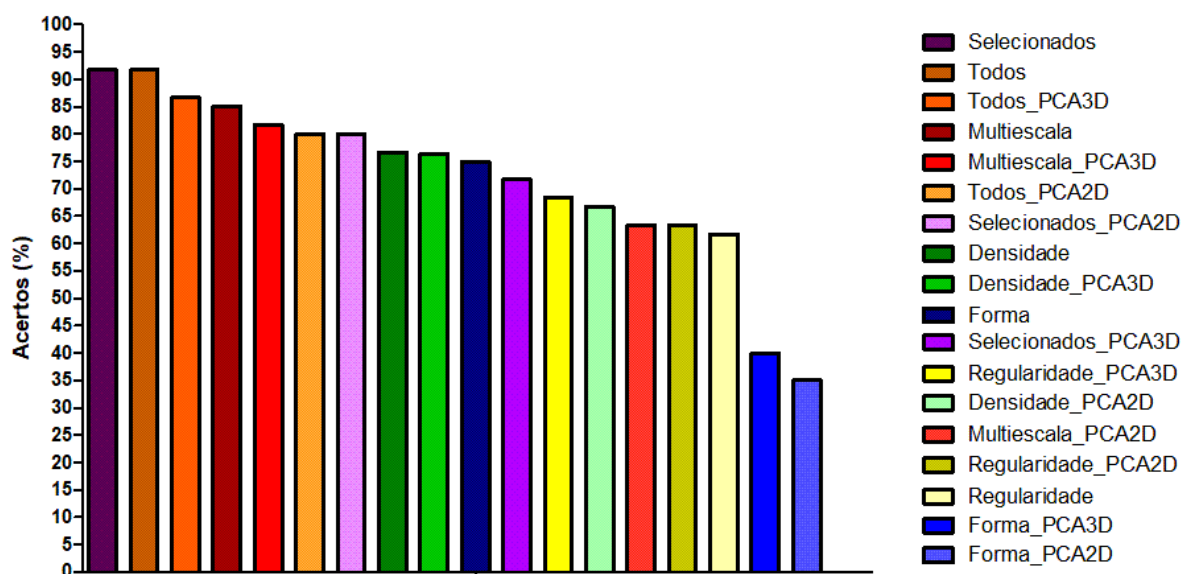


Figura 74 – Gráfico de barras com os acertos da abordagem *PMC* para as categorias estudadas. Cada barra representa uma categoria em um de seus três espaços amostrais (original, *ACP* (ou do inglês *PCA*) 2D, *ACP* (ou do inglês *PCA*) 3D).

Fonte: Autor

Sendo assim, podemos inferir que, na abordagem *PMC*, as categorias de atributos que apresentaram maior acurácia (A_c) frente a todos os tecidos vegetais estudados foram as categoria *Seleccionados* e *Todos* ambas em seus respectivos espaços originais de 14 e 33 dimensões ($A_c = 91,67\%$). A categoria *Todos* aparece em segundo lugar com uma porcentagem de acerto de $86,67\%$ quando foi estudada tomando os três componentes principais. A categoria *Multiescala* em seu espaço original, ou seja, 21 dimensões ocupa a quarta colocação e apresenta $A_c = 85\%$. Essa categoria também ocupa a quinta posição com $A_c = 81,67\%$ porém, aqui, ela foi estudada tomando os três componentes principais. A categoria *Todos* e a *Seleccionados* quando estudadas tomando os dois componentes principais ocupam, respectivamente, a quinta e a sexta posição com a mesma acurácia de $A_c = 80\%$.

Vale ressaltar que quando os feixes vasculares foram validados pela abordagem *PMC*, os atributos referente a *Forma* ocuparam a décima posição quando estudada em seu espaço original de 10 dimensões e obtiveram $A_c = 75\%$. Já quando essa categoria foi estudada após aplicação da análise *ACP*, os erros de classificação ficaram ainda maiores e, conseqüentemente, a acurácia menor. Quando a categoria *Forma* foi estudada tomando os três componentes principais ela obteve $A_c = 40\%$ e ocupou a penúltima posição do *ranking*. Quando estudada tomando os dois componentes principais ela obteve $A_c = 35\%$ e ocupou a última posição do *ranking*.

5.3.3 Considerações finais

Ao final da etapa de validação, nota-se grande semelhança entre as acurácias obtidas por meio da abordagem *k-NN* e *PMC*. Para analisar melhor os resultados obtidos por essas duas abordagens, os valores de acurácia, ou seja, de acerto de cada categoria foram plotados em um gráfico de barras. Esse gráfico de barras está mostrado na Figura 75 e cada barra representa uma categoria em um de seus três espaços amostrais (original, *ACP* (ou do inglês PCA) 2D, *ACP* (ou do inglês PCA) 3D). Esse gráfico de barras foi construído com a acurácia média entre as abordagens *k-NN* e *PMC*.

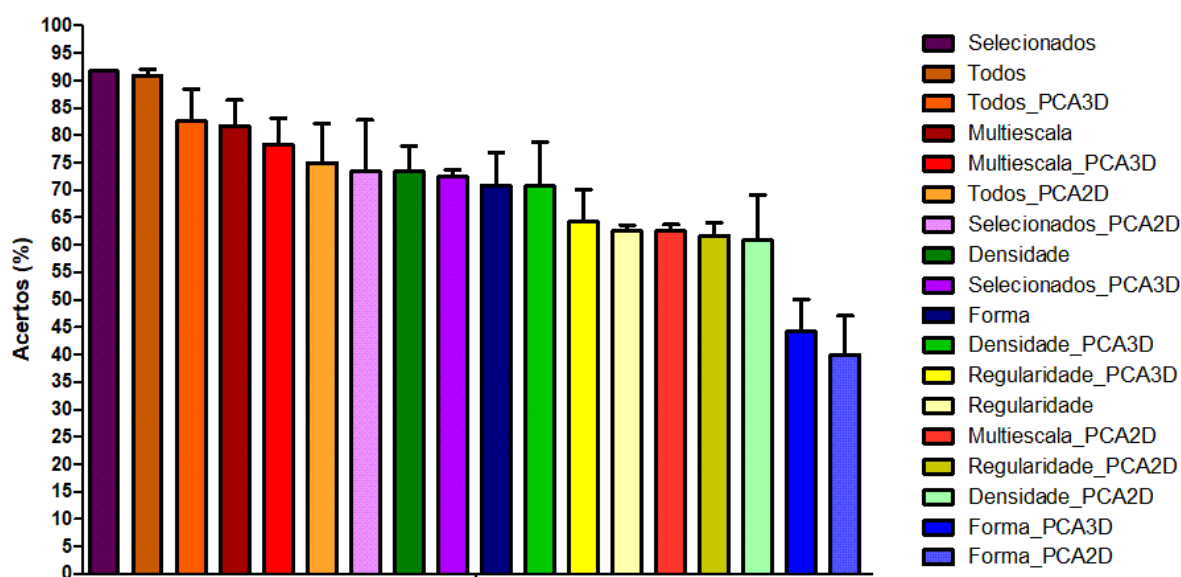


Figura 75 – Gráfico de barras com os acertos médios das abordagens *k-NN* e *PMC* para as categorias estudadas. Cada barra representa uma categoria em um de seus três espaços amostrais (original, *ACP* (ou do inglês PCA) 2D, *ACP* (ou do inglês PCA) 3D).

Fonte: Autor

As cinco melhores categorias foram as mesmas em ambas as abordagens (vide Figura 73 e Figura 74). Consequentemente, podemos inferir que as melhores categorias de atributos capazes de segregar todos os feixes vasculares vegetais são: *Selecionados* (espaço original), *Todos* (espaço original), *Todos* (redução da dimensionalidade para 3 dimensões), *Multiescala* (espaço original) e *Multiescala* (redução da dimensionalidade para 3 dimensões).

Categorias nas quais há mais de um tipo de atributos na sua constituição foram as mais relevantes para nosso estudo e ocuparam da primeira a sétima colocação no *ranking* de acertos médios.

Esse resultado é interessante pois, vemos que um grupo resumido com 14 atributos (categoria *Selecionados*: composta por atributos referentes a Forma, a Densidade, a

Regularidade estrutural e a Multiescala) teve o melhor resultado na classificação de todos os feixes vasculares vegetais. Essa categoria teve um resultado levemente superior quando comparado a categoria composta por todos os 33 atributos extraídos. Isso sugere que categorias formadas por mais de um tipo de atributos conseguem caracterizar melhor os feixes vasculares vegetais e, conseqüentemente, corroboram para uma melhor classificação desses tecidos.

Essa evidência também é reforçada pela quarta posição da categoria Multiescala no *ranking* de acertos médio. Essa categoria é composta por atributos referente a Densidade e a Regularidade estrutural, ou seja, por dois tipos de atributos. Esse resultado também sinaliza a relevância da extração de atributos considerando a análise multiescala. Ao longo de todo trabalho, foi notado que muitos atributos tiveram resultados importantes para uma determinada escala e não apresentaram resultados para outras. Categorias composta por apenas um tipo de atributo, como por exemplo, Densidade, Forma e regularidade estrutural, apresentaram, em seus respectivos espaços originais, a oitava, décima e décima terceira posição no *ranking* de acertos médio.

Outro resultado relevante é o desempenho da categoria Forma na classificação de feixes vasculares vegetais. Quando estudada em seu espaço original de 10 dimensões, ela ocupou o décimo lugar no *ranking* de acertos médio. E quando estudada tomando como base dois e três componentes principais, ela ocupou, respectivamente, a penúltima e a última posição no *ranking* de acertos médio. Como evidenciado ao longo de todo trabalho, atributos de forma são importantes na caracterização dos feixes vasculares vegetais porém, de maneira geral, podemos sugerir que os atributos referentes ao formato celular sozinhos não são promissores para a classificação dos feixes vasculares vegetais.

Também vale ressaltar que as categoriais quando estudadas em seus espaços originais tiveram resultados superiores quando comparadas com as respectivas categorias com redução de dimensionalidade para 2 e 3 dimensões. Essa afirmação só não é verdadeira quando analisamos a categoria regularidade pois, nela o melhor desempenho de acertos foi obtido quando foi ela foi estudada tomando como base os três componentes principais.

Os resultados aqui apresentados podem contribuir para o surgimento de novas metodologias de extração de atributos. Visto que, categorias não muito usuais de atributos (por exemplo, Multiescala) apresentaram melhores rendimentos se comparadas a categorias mais tradicionais (por exemplo, formato celular) dentro da área de classificação de tecidos vegetais.

Os resultados também reforçam a importância em se estudar as particularidades e as características de cada tecido vegetal antes de elencar quais os atributos que serão extraídos, uma vez que, mesmo os tecidos fazendo parte do mesmo grupo eles apresentam individualidades que possibilitaram sua classificação.

6 Discussão

Os avanços recentes em análise de imagens e reconhecimento de padrões abriram caminho para várias aplicações em biologia. Em particular, temos o interessante problema de caracterizar imagens digitais de plantas com o objetivo de caracterizar tecidos de frutas (RAMOS et al., 2004; MAYOR; SILVA; SERENO, 2005; MAYOR; PISSARA; SERENO, 2008; SANYAL; BHATTACHARYA; BANDYOPAHYAY, 2008), análise do formato da flor (MIAO; GANDELIN; YUAN, 2006), extração de elementos agrícolas (BURGOS-ARTIZZU et al., 2010) e rastreamento de células vegetais (LIU et al., 2017). A busca por descritores microhistológicos por meio da análise de imagens digitais de plantas tem contribuído para a identificação das espécies (ROSITO; MARCHESAN, 2003), e para a compreensão da arquitetura do tecido durante o seu desenvolvimento ou quando afetado por doenças (SÁNCHEZ-GUTIÉRREZ et al., 2016).

A caracterização dos tecidos vegetais começa com a busca de descritores microhistológicos que sejam capazes de distinguir as classes em estudo. Os anatomista ao realizarem a caracterização dos tecidos utilizam as características morfofuncionais desses tecidos. Por exemplo, eles analisam qualitativamente as características acessórias dos tecidos tais quais: espessura da parede, presença de compostos como, por exemplo, lignina e a disposição dos tecidos no corte histológico. Nosso trabalho buscou quantificar alguns desses atributos morfofuncionais.

Porém, a literatura atual mostra que as abordagens computacionais de caracterização de tecidos vegetais tendem a extrair descritores microhistológicos de um único grupo: medidas de forma (MAYOR; SILVA; SERENO, 2005; MAYOR; PISSARA; SERENO, 2008; MIAO; GANDELIN; YUAN, 2006; MURRAY; O'MALLEY, 1993; RAMOS et al., 2004; SANYAL; BHATTACHARYA; BANDYOPAHYAY, 2008). Ao analisarmos os trabalhos citados, vemos que, normalmente, os atributos tradicionalmente extraídos incluem a área, perímetro, diâmetro e o alongamento. Observa-se que em todos os trabalhos supracitados somente os atributos de Forma foram suficiente para caracterizar e descrever os tecidos analisados.

Embora o desempenho dos atributos de forma obtido em nosso trabalho pareça contrariar a literatura já publicada, uma vez que, os atributos referentes ao formato celular obtiveram um acerto médio de cerca de 75% quando analisados em seu espaço original de 10 dimensões e 45% e 40%, respectivamente, quando tomados como base três e dois componentes principais. Esses grupos ocuparam, respectivamente, a décima, penúltima e última colocação em nosso *ranking* geral. Esses atributos estão presentes majoritariamente no grupo de atributos *Selecionados* ao longo de todo trabalho e, portanto,

possuem relevância para a caracterização dos tecidos vegetais estudados corroborando com a literatura consultada (MAYOR; SILVA; SERENO, 2005; MAYOR; PISSARA; SERENO, 2008; MIAO; GANDELIN; YUAN, 2006; MURRAY; O'MALLEY, 1993; RAMOS et al., 2004; SANYAL; BHATTACHARYA; BANDYOPAHYAY, 2008). Isso está alinhado com o processo natural do reconhecimento de tecidos vegetais no qual os anatomistas utilizam de diversos outros atributos não mesuráveis pelo processo cognitivo humano para caracterizar e descrever um determinado tecido.

Em nosso trabalho, a categoria que obteve o melhor desempenho na classificação dos tecidos vasculares estudados, com uma média de 91,67% de acertos foi a composta pelos atributos *Seleccionados* ao longo do trabalho. Esse grupo é constituído por 14 atributos, sendo eles: área média, desvio padrão área, resíduo médio, desvio padrão resíduo, diâmetro médio, desvio padrão diâmetro, diâmetro equivalente médio, desvio padrão do diâmetro equivalente, uniformidade espacial com sigma 30, uniformidade espacial com sigma 65, densidade radial com raio 75, poligonalidade média, lacunaridade com raio 6 e lacunaridade com raio 21. Ou seja, 8 atributos são referente ao formato celular, 3 referentes a densidade e 3 a regularidade estrutural. Entre esses 14 descritores selecionados, 5 deles são extraídos por meio da análise multiescala. Isso sugere que para a análise dos feixes vasculares vegetais se faz necessário a extração de mais de uma categoria de atributos, similarmente como é feita na classificação natural desses feixes.

É importante ressaltar que a categoria *Seleccionados* teve um rendimento de acerto levemente superior a categoria *Todos* visto que, essa última teve uma porcentagem de acertos média igual a 90,84% enquanto que a primeira teve uma média de 91,67%. A categoria *Todos* é composta por 33 atributos enquanto que a *Seleccionados* é composta por apenas 14. Ou seja, essa redução de 19 atributos irrelevantes ou redundantes proporcionou aos classificadores o mesmo rendimento de acerto porém necessitando de menos descritores. Isso corrobora com a literatura vigente que indica que a seleção de descritores busca encontrar, por meio da exclusão de atributos irrelevantes ou redundantes, a informação primordial contida nos dados originais (LEE, 2005; LIU; MOTODA, 2012). Assim, temos a vantagem de conseguir os mesmos resultados com menos esforço computacional.

Atualmente, a disponibilidade de recursos computacionais permite que abordagens pouco usuais sejam utilizadas na área de reconhecimento de padrões (JAIN; DUIN; MAO, 2000; PLOTZE et al., 2005). A multiescala tem sido amplamente utilizado na literatura de física e biologia (YIN et al., 2019; SHANG et al., 2019; COSTA; GOLDBERGER; PENG, 2005; LIU et al., 2017; SONG et al., 2018; XING et al., 2017). Ainda assim, a aplicação de métodos multiescala para caracterização de tecidos histológicos de plantas tem sido pouco discutida na literatura, e geralmente está mais relacionada a descritores de textura foliar (PLOTZE; BRUNO, 2007; SILVA; FLORINDO, 2019). Assim, existe uma lacuna na exploração dos descritores oriundos da análise multiescala para a caracterização de tecidos

vegetais, como, por exemplo, tecidos vasculares.

Em nosso trabalho, a categoria Multiescala obteve o quarto lugar no *ranking*, com uma média de acerto de 81,66%. Devido a isso, podemos concluir que os atributos multiescala são potencialmente bons para classificar feixes vasculares vegetais. Vale ressaltar que entre os atributos mais relevantes ao longo de todo trabalho, cinco deles foram extraídos por meio da abordagem multiescala. Isso reforça a importância dessa categoria frente ao objetivo dessa dissertação.

Outro objetivo do nosso trabalho foi tentar entender biologicamente os motivos que justifiquem a relevância computacional apresentadas pelos atributos que melhores descreveram os feixes vasculares vegetais. Como já sabido, em nosso trabalho, 14 descritores foram capazes de classificar os tecidos vegetais com uma pequena taxa de erro. São eles: área média, desvio padrão área, resíduo médio, desvio padrão resíduo, diâmetro médio, desvio padrão diâmetro, diâmetro equivalente médio, desvio padrão do diâmetro equivalente, uniformidade espacial com sigma 30, uniformidade espacial com sigma 65, densidade radial com raio 75, poligonalidade média, lacunaridade com raio 6 e lacunaridade com raio 21. De agora em diante, focaremos nas possíveis relevâncias biológicas de cada um dos descritores que são constituintes da categoria que melhor classificou os feixes vasculares vegetais.

Biologicamente, os feixes vasculares são constituídos pelos tecidos floema e xilema (RAVEN; EVERT; EICHHORN, 2014). O xilema é o tecido responsável pelo transporte de água e minerais, enquanto o floema carrega os compostos orgânicos resultantes da fotossíntese. Assim, a demanda de transporte do xilema excede a do floema, uma vez que a taxa de transporte de fluido é maior no xilema do que no floema. Dessa maneira, a diferença hidráulica entre os tecidos pode justificar a distinção entre o diâmetro do xilema e do floema. A literatura corrobora com essas afirmações, uma vez que há trabalhos que relatam que o xilema, por possuir maior capacidade de fluxo, tende a ser constituído por células maiores quando comparado ao floema (HÖLTTÄ; MENCUCCINI; NIKINMAA, 2009; WU et al., 2011; LUCAS et al., 2013; SANTAROSA et al., 2016). Em nosso trabalho, nota-se que os resultados relacionados à forma celular levaram a uma interpretação biológica relevante, uma vez que, em cada tipo de feixe vascular as células constituintes do xilema são maiores do que as células do floema.

Os atributos área média, diâmetro médio e diâmetro equivalente médio apresentaram coeficiente de correlação de Pearson alto e positivo em todos os feixes estudados. Isso indica que células com grandes áreas também tendem a apresentar valores de diâmetro e diâmetro equivalente altos. Portanto, a relevância biológica desses atributos pode ser justificada de maneira análoga.

Ainda sobre os atributos de forma selecionados ao longo do trabalho, nota-se a presença do desvio padrão da área, desvio padrão do diâmetro e desvio padrão do diâmetro equivalente. Isso sugere que há uma diferença de tamanho entre as células que

constituem o mesmo tecido (xilema ou floema). Biologicamente, tanto no xilema quanto no floema as primeiras células que se diferenciam formam o protoxilema e o protofloema, respectivamente. Posteriormente, novas células do procâmbio se diferenciam formando o metaxilema e o metafloema (MENEZES, 2016; RAVEN; EVERT; EICHHORN, 2014). Existem algumas diferenças entre esses dois tipos celulares mas, para nosso estudo, a diferença primordial é que as células que se originam posteriormente (ou seja, metaxilema e metafloema) são caracterizadas por serem maiores e mais numerosas quando comparadas as células que se diferenciam primeiro (protoxilema e protofloema) (MENEZES, 2016). Sendo assim, é natural que haja uma diferença entre os valores extraídos de área, diâmetro e diâmetro equivalente. Dessa maneira, os atributos relacionados com os desvio padrão dessas medidas ganham relevância no contexto da identificação dos feixes vasculares vegetais.

Embora todas as medidas de forma selecionadas nesse trabalho pareçam se justificar apenas olhando para as características dos tecidos xilema e floema, para o entendimento das outras medidas selecionadas se faz necessário olhar cada tipo de feixe vascular de acordo com suas particularidades em termos de organização. Um desses atributos é a densidade radial, uma vez que, essa medida é fundamentada na relação do número de centroides celulares dentro do anel e a área desse anel (COSTA; CESAR, 2009). Ou seja, ela busca analisar como estão distribuídas as células dentro de um tecido.

Partindo dessa premissa e sabendo que os feixes vasculares concêntricos são caracterizados por apresentarem um tecido vascular circundando o outro (MODESTO; SIQUEIRA, 2002; MENEZES, 2016), podemos inferir que o atributo densidade radial seja potencialmente capaz de caracterizar tantos os xilemas quanto os floemas desses feixes. Em nosso trabalho, esse descritor, quando estudado com raio de 75, apresentou relevância para os feixes do tipo concêntrico nas três abordagens analisadas (individual, par a par e APC). O tamanho do raio foi fundamental pois o atributo conseguiu mapear tanto o tecido interno quanto o externo dos feixes concêntricos. É importante ressaltar que ao longo desse trabalho o descritor densidade radial foi estudado com seis raios (25, 75, 125, 175, 225 e 275). Novamente aqui, fica evidente que a multiescala foi importante pois, dos seis raios analisados, apenas um apresentou relevância geral em nosso trabalho.

Outro descritor que apresentou relevância ao longo do nosso trabalho foi a uniformidade espacial. Essa medida está relacionada com a organização da imagem, ou seja, ela quantifica a uniformidade da densidade dos centroides celular. Biologicamente, alguns feixes vasculares, como por exemplo o feixe anfigasal, apresentam em sua formação uma característica denominada zonação que consiste em dividir em zonas o feixe vascular (JURA-MORAWIEC, 2015). Nesse caso, um feixe anfigasal em desenvolvimento pode conter uma região onde haja células maduras e outro em que hajam células que ainda estão se dividindo (JURA-MORAWIEC, 2015). Esse processo de amadurecimento é rápido e, ao final, as células do xilema estão ordenadas centrifugamente. Essa característica sugere que

alguns feixes sejam mais organizados do que outros e portanto, o atributo uniformidade espacial foi importante para descrever esses tecidos.

Outra medida relevante foi a poligonalidade. Ela está relacionada com a regularidade de ângulos entre pontos vizinhos em um diagrama de Voronoi (AURENHAMMER, 1991; COSTA; ROCHA; LIMA, 2006). Em outras palavras, ela mapeia a vizinhança de uma célula. Os tecidos vasculares podem sofrer um fenômeno conhecido como crescimento intrusivo. Ele ocorre quando uma célula aumenta de tamanho mais rápido do que as vizinhas e se introduz entre elas (GORSHKOV et al., 2019). O trabalho publicado por Jura-Morawiec (2017) evidenciam esse fenômeno ocorrendo em células do xilema de feixes anfigvais da *Dracaena marginata* enquanto que, os trabalhos publicados por Snegireva et al. (2015) e por Raman et al. (2017) descrevem o crescimento intrusivo em células do floema dos feixes vasculares do tipo colateral. Sendo assim, a literatura indica que o crescimento intrusivo altera a vizinhança das células de um tecido vascular. Portanto, estudar os vizinhos de uma célula foi importante para nosso trabalho.

Ademais, a posição do tecido vegetal dentro do feixe pode influenciar sua relação de vizinhança. Por exemplo, nos feixes anfigvais, o xilema está disposto em forma de um cordão no centro do feixe (MODESTO; SIQUEIRA, 2002; MENEZES, 2016). Portanto, a vizinhança de uma célula desse tecido será limitada a poucas células. Em contrapartida, células do floema desse mesmo tecido estão organizadas ao redor do xilema e, portanto, possuem uma vizinhança mais povoada. Sendo assim, além do crescimento intrusivo, a própria organização dos tecidos dentro dos feixes vasculares indicam os motivos pelos quais a poligonalidade foi um dos atributos mais relevantes para descrever os feixes vasculares vegetais.

A lacunaridade também apresentou relevância ao longo do nosso trabalho. Esse atributo indica a homogeneidade de um tecido, ou seja, imagens que possuem lacunas de tamanhos irregulares ou distribuídas de maneira heterogênea apresentam alta lacunaridade (RODRIGUES, 2014; RODRIGUES; BARBOSA; COSTA, 2005). Em nosso trabalho, as células que constituem cada tecido foram entendidas como lacunas. Biologicamente, as células do metaxilema e metafloema são maiores quando comparadas as células protoxilema e protofloema (MENEZES, 2016). Em alguns feixes, como por exemplo, bicolateral e colateral fechado é possível notar variações entre as próprias células do metaxilema (MODESTO; SIQUEIRA, 2002; MENEZES, 2016). Segundo Cheadle e Uhl (1948b), os feixes colateral fechado e bicolateral são caracterizados por apresentarem elementos de metaxilemas de tamanhos grandes e pequenos. Em contrapartida, os metafloemas são mais homogêneos e não apresentam variações de diâmetro dentro desses feixes (CHEADLE; UHL, 1948a). Os resultados apresentados ao longo do nosso trabalho indicam que o xilema foi melhor descrito pela lacunaridade do que o floema. Biologicamente, é sabido que o xilema possui paredes celulares mais espessas. Nossos resultados corroboram com a

literatura descritiva dos feixes vasculares vegetais e sugerem que essa irregularidade de tamanho apresentada pelas células do metaxilema foram determinantes para que esse descritor apresentasse relevância frente aos outros descritores utilizados.

Dos cinco raios estudados para o atributo lacunaridade, esse descritor apresentou relevância apenas em dois: 6 e 21. Essa propriedade está diretamente relacionada com o tamanho dos feixes, uma vez que, espera-se que feixes menores possam ser bem caracterizados por raios grandes e pequenos e feixes maiores sejam caracterizados apenas por raios maiores. Segundo Bhat, Liese e Schmitt (1990), espécies com maior número de feixes vasculares por mm^2 de área apresentam feixes vasculares pequenos, ou seja, há uma redução do tamanho dos feixes vasculares com o aumento do número. A literatura indica que a espécie *Sechium edule* possui cerca de 13 feixes bicolaterais em um corte transversal do caule (SILVA et al., 2020), enquanto que o número de feixes colaterais fechados em um corte transversal do caule de *Cyperus rotundus* é em torno de 30 (JHA; SEN, 1980). Sendo assim, podemos afirmar que os tamanhos dos feixes vasculares avaliados em nosso estudo são variados e que os bicolaterais são maiores do que os colaterais fechados. Portanto, um raio poderá mapear de maneira mais eficiente os feixes bicolaterais.

Sendo assim, nota-se que os atributos selecionados ao longo desse trabalho possuem tanto significância computacional quanto relevância biológica, ou seja, os descritores aqui apresentados foram capazes de refletir a natureza dos diferentes feixes vasculares apresentados.

7 Conclusões e Trabalhos Futuros

7.1 Conclusões

Os constantes avanços da ciência somados ao uso cada vez mais recorrente de abordagens interdisciplinares têm aberto caminho para as mais diferentes aplicações científicas. O presente trabalho trouxe abordagens já bem descritas para a caracterização de tecidos vegetais como também buscou descritores pouco usados mas que possuem uma certa relação com as particularidades de cada tipo de feixe vascular vegetal.

Dentro desse cenário, estudamos seis feixes vasculares vegetais (anfivasal primário, anfivasal secundário, anficrival, colateral aberto, colateral fechado e bicolateral) por meio da digitalização das lâminas histológicas de cortes transversais de caules de espécies de plantas que continham esses feixes.

Ao longo do trabalho, foram apresentados vários descritores. Cada categoria de atributos foi escolhida baseada em características particulares de cada tipo de feixe vascular vegetal. Após a etapa de seleção, esse conjunto de atributos foi reduzido. E, ao final, conseguimos classificar com mais de 90% de acerto os tecidos estudados. Como vimos, todos os atributos selecionados possuem explicações biológicas relevantes. Isso reforça a ideia de que se faz necessário conhecer o objeto de estudo antes de escolher quais descritores serão usados.

7.1.1 Atributos selecionados

Os resultados finais desse trabalho indicam que a categoria de atributos que teve maior média de acertos na classificação dos tecidos foi a *Selecionados*. Vale ressaltar que esse conjunto foi constituído pelos descritores que apresentaram maior recorrência ao longo do trabalho apresentado.

Esse grupo foi constituído por 14 atributos, sendo eles: área média, desvio padrão área, resíduo médio, desvio padrão resíduo, diâmetro médio, desvio padrão diâmetro, diâmetro equivalente médio, desvio padrão do diâmetro equivalente, uniformidade espacial com sigma 30, uniformidade espacial com sigma 65, densidade radial com raio 75, poligonalidade média, lacunaridade com raio 6 e lacunaridade com raio 21.

Sendo assim, podemos concluir que as abordagens utilizadas para elencar os atributos mais relevante e mais recorrentes foram satisfatórias, uma vez que, por meio delas foi possível reduzir o espaço amostral de 33 para 14 atributos e atingir uma média de acertos superior a 90% na classificação dos feixes vasculares vegetais estudados.

Outra conclusão importante é que o conjunto *Selecioneado* é composto por 8 atributos referente ao formato celular, 3 referentes a densidade e 3 a regularidade estrutural. Entre esses 14 descritores selecionados, 5 deles são extraídos por meio da análise multiescala. Portanto, podemos concluir que é importante extrair descritores de várias categorias.

7.1.2 Atributos de forma

Vimos que a literatura indica que um grande número de trabalhos focados na classificação de tecidos vegetais usam como descritores apenas os atributos referentes ao formato celular.

Em nosso trabalho, foi observado que o desempenho classificatório dos feixes vasculares vegetais obtido apenas por meio dos atributos de formato celular foi baixo e, portanto, abaixo do esperado. Porém, é notável que dos 14 atributos selecionados em nosso trabalho, 8 são referentes ao formato celular. Isso sugere a importância que esses descritores possuem frente aos tecidos vegetais visto que, por muito anos, a classificação desse material biológico foi feita pela morfologia desses tecidos.

Portanto, concluímos que os atributos de forma são importantes para a classificação dos feixes vasculares porém, a média de acertos classificatórios é potencializada quando esses atributos são estudados juntamente com atributos de outras categorias.

7.1.3 Abordagens pouco utilizadas

Em nosso estudo, além da categoria de forma, estudamos a eficiência classificatória de outras 3 categorias: Densidade, Regularidade estrutural e Multiescala. Todas essas categorias são pouco utilizadas em trabalhos focadas na classificação e descrição de tecidos vegetais.

Entre essas 3 categorias a que teve maior média classificatória foi a Multiescala. Vale ressaltar que ao longo de todo trabalho, foi comum notarmos que um atributo possuía relevância em uma determinada escala e nas outras não. Dessa maneira, podemos concluir que extrair atributos levando em consideração a abordagem multiescala pode ser uma ferramenta importante na descrição e na classificação dos feixes vasculares vegetais.

Em contrapartida, atributos relacionados a densidade celular e a regularidade estrutural tiveram desempenho classificatórios baixos quando analisados sozinhos. Vale ressaltar que, analogamente aos atributos de forma, os descritores de densidade e de regularidade estrutural também estão presentes no conjunto *Selecioneados*, sendo 3 referentes a densidade e 3 a regularidade estrutural. Portanto, podemos concluir que extrair atributos dessas duas categorias é importante e que as médias de acertos classificatórios é potencializada quando esses atributos são estudados juntamente com atributos de outras categorias.

Em geral, as abordagens poucos descritas na literatura apresentaram média classificatória maior do que a abordagem mais utilizada. Isso reforça dois pontos-chaves de nosso trabalho: o conhecimento *a priori* dos tecidos alvos e o uso da interdisciplinaridade para buscar novas abordagens. Em vista disso, concluímos que os avanços tecnológicos e interdisciplinares podem abrir caminhos para que novas abordagens possam ser testadas e novas interpretações biológicas possam ser obtidas.

7.1.4 Relevância biológica

Em nosso trabalho, foi feito um levantamento biológico de possíveis interpretações que potencialmente justificassem a seleção dos 14 atributos que melhor classificaram os feixes vasculares vegetais.

Ao final, vimos que os descritores selecionados foram capazes de capturar as particularidades de cada tipo tecidual estudado. Isso reforça a necessidade em se conhecer os alvos a serem estudados pois, assim, a extração de atributos pode ser feita de maneira mais apropriada frente as particularidades dos elementos alvos.

7.2 Trabalhos Futuros

Nosso trabalho se mostrou promissor em apresentar a importância do estudo de mais de um tipo de categoria de atributos para descrever feixes vasculares vegetais. Dentro desse cenário, há uma possibilidade imensa de novas aplicações para os resultados aqui apresentados:

Os atributos selecionados podem ser aplicados para um novo grupo de tecidos vasculares obtidos de outras espécies. Isso ajudaria a garantir a generalização da nossa abordagem para outras espécies vegetais.

Os atributos selecionados podem ser aplicados a outros tecidos vegetais para testar se eles são eficazes na classificação de outros tipos de tecidos vegetais.

Ampliar os estudos sobre a categoria Multiescala, extraindo novos atributos por meio dessa abordagem e testando seu poder classificatório frente aos tecidos vasculares bem como outros tecidos vegetais.

Testar o desempenho do algoritmo apresentado sem a inserção da categoria Multiescala.

Nosso trabalho também fornece um banco de dados com imagem segmentadas (*padrão ouro*) dos feixes vasculares vegetais. Esse material pode ser incorporado à trabalhos futuros onde se deseja padronizar abordagens de segmentação automatizada.

Referências

- ALLAIN, C.; CLOITRE, M. Characterizing the lacunarity of random and deterministic fractal sets. *Physical review A*, APS, v. 44, n. 6, p. 3552, 1991.
- AMABIS, J. M.; RODRIGUES, G. M. *Fundamentos da biologia moderna*. [S.l.]: Moderna, 2006.
- ANTONIE van Leeuwenhoek: inventor do microscópio. *Jornal Brasileiro de Patologia e Medicina Laboratorial*, Scielo, v. 45, n. 2, 2009.
- AURENHAMMER, F. Voronoi diagrams—a survey of a fundamental geometric data structure. *ACM Computing Surveys (CSUR)*, ACM, v. 23, n. 3, p. 345–405, 1991.
- BACKES, A.; BRUNO, O. M. Lacunaridade aplicada em análise de textura. *I Workshop de Visão Computacional*, Biblioteca Digital Brasileira de Computação, p. 48–51, 2005.
- BAQUIÃO, A. P. S. S. et al. Percepção de residentes multiprofissionais acerca da interdisciplinaridade. *Saúde e Pesquisa*, Universidade Federal de Juiz de Fora, v. 12, n. 1, p. 187–196, 2019.
- BARBEDO, J. A novel algorithm for semi-automatic segmentation of plant leaf disease symptoms using digital image processing. *Tropical Plant Pathology*, Springer, v. 41, n. 4, p. 210–224, 2016.
- BARBEDO, J. G. A. A new automatic method for disease symptom segmentation in digital photographs of plant leaves. *European journal of plant pathology*, Springer, v. 147, n. 2, p. 349–364, 2017.
- BENESTY, J. et al. Pearson correlation coefficient. In: *Noise reduction in speech processing*. [S.l.]: Springer, 2009.
- BERGOUIGNAN, L. et al. Can voxel based morphometry, manual segmentation and automated segmentation equally detect hippocampal volume differences in acute depression? *Neuroimage*, Elsevier, v. 45, n. 1, p. 29–37, 2009.
- BHAT, K.; LIESE, W.; SCHMITT, U. Structural variability of vascular bundles and cell wall in rattan stem. *Wood Science and Technology*, Springer, v. 24, n. 3, p. 211–224, 1990.
- BOLDRINI, J. L.; AL. et. *Álgebra linear*. [S.l.]: Harper & Row, 1980.
- BRAMER, M. *Principles of data mining*. [S.l.]: Springer, 2007. v. 180.
- BURGOS-ARTIZZU, X. P. et al. Analysis of natural images processing for the extraction of agricultural elements. *Image and Vision Computing*, Elsevier, v. 28, n. 1, p. 138–149, 2010.
- CASTRO, L. N. D. *Fundamentals of natural computing: basic concepts, algorithms, and applications*. [S.l.]: CRC Press, 2006.

- CATTAI, M. B.; MENEZES, N. L. d. Primary and secondary thickening in the stem of *Cordyline fruticosa* (agavaceae). *Anais da Academia Brasileira de Ciências*, SciELO Brasil, v. 82, n. 3, p. 653–662, 2010.
- CERUTTI, G. et al. Understanding leaves in natural images—a model-based approach for tree species identification. *Computer Vision and Image Understanding*, Elsevier, v. 117, n. 10, p. 1482–1501, 2013.
- CHEADLE, V. I. The taxonomic use of specialization of vessels in the metaxylem of gramineae, cyperaceae, juncaceae, and restionaceae. *Journal of the Arnold Arboretum*, JSTOR, v. 36, n. 2/3, p. 141–158, 1955.
- CHEADLE, V. I.; UHL, N. W. The relation of metaphloem to the types of vascular bundles in the monocotyledoneae. *American Journal of Botany*, JSTOR, v. 35, n. 9, p. 578–583, 1948.
- CHEADLE, V. I.; UHL, N. W. Types of vascular bundles in the monocotyledoneae and their relation to the late metaxylem conducting elements. *American Journal of Botany*, JSTOR, v. 35, n. 8, p. 486–496, 1948.
- CHIU, S. J. et al. Automatic segmentation of seven retinal layers in sdoct images congruent with expert manual segmentation. *Optics express*, Optical Society of America, v. 18, n. 18, p. 19413–19428, 2010.
- CLARK, G. Staining procedures. In: *Staining procedures*. [S.l.]: William & Wilkins, 1981.
- COSTA, L. C.; CESAR, R. M. *Shape classification and analysis: theory and practice*. [S.l.]: CRC Press, Inc., 2009.
- COSTA, L. F.; ROCHA, F.; LIMA, S. M. A. de. Characterizing polygonality in biological structures. *Physical Review E*, APS, v. 73, n. 1, p. 011913, 2006.
- COSTA, M.; GOLDBERGER, A. L.; PENG, C.-K. Multiscale entropy analysis of biological signals. *Physical review E*, APS, v. 71, n. 2, p. 021906, 2005.
- COVER, T.; HART, P. Nearest neighbor pattern classification. *IEEE transactions on information theory*, IEEE, v. 13, n. 1, p. 21–27, 1967.
- DU, C.-J.; SUN, D.-W. Learning techniques used in computer vision for food quality evaluation: a review. *Journal of food engineering*, Elsevier, v. 72, n. 1, p. 39–55, 2006.
- DUDA, R. O.; AL. et. *Pattern classification*. [S.l.]: Wiley New York, 1973.
- DUDA, R. O.; HART, P. E.; STORK, D. G. *Pattern classification*. [S.l.]: John Wiley & Sons, 2012.
- EINSTEIN, A. J.; WU, H. S.; GIL, J. Self-affinity and lacunarity of chromatin texture in benign and malignant breast epithelial cell nuclei. *Physical Review Letters*, APS, v. 80, n. 2, p. 397, 1998.
- FERRI, M. G. *Botânica: morfologia interna das plantas: anatomia*. [S.l.]: NBL Editora, 1983.
- FISHER, R. A. The use of multiple measurements in taxonomic problems. *Annals of eugenics*, Wiley Online Library, v. 7, n. 2, p. 179–188, 1936.

- GAMBÄCK, B.; SIKDAR, U. K. Using convolutional neural networks to classify hate-speech. In: *Proceedings of the first workshop on abusive language online*. [S.l.: s.n.], 2017.
- GLÓRIA, B. A.; GUERREIRO, S. M. C. *Anatomia Vegetal*. [S.l.]: UFV, 2006.
- GONZALEZ, R. C.; WOODS, R. E. *Processamento de imagens digitais*. [S.l.]: Editora Blucher, 2000.
- GORSHKOV, O. et al. Intrusive growth of phloem fibers in flax stem: integrated analysis of mirna and mrna expression profiles. *Plants*, Multidisciplinary Digital Publishing Institute, v. 8, n. 2, p. 47, 2019.
- HENRY, J. *A revolução científica*. [S.l.]: Zahar, 1998.
- HERREMANS, E. et al. Automatic analysis of the 3-d microstructure of fruit parenchyma tissue using x-ray micro-ct explains differences in aeration. *BMC plant biology*, BioMed Central, v. 15, n. 1, p. 1–14, 2015.
- HÖLTTÄ, T.; MENCUCCINI, M.; NIKINMAA, E. Linking phloem function to structure: analysis with a coupled xylem–phloem transport model. *Journal of theoretical biology*, Elsevier, v. 259, n. 2, p. 325–337, 2009.
- IUCN. *Number of species evaluated in relation to the overall number of described species, and numbers of threatened species by major groups of organisms*. 2019. IUCN Red List. Disponível em: <<https://www.iucnredlist.org/resources/summary-statistics#Summary%20Tables>>. Acesso em: 10 mar. 2020.
- IZIDORO, S. C. Determinação de número de agrupamentos em conjunto de dados multidimensionais utilizando algoritmos genéticos. *Journal of Computer Science*, INFOCOMP, n. 4, p. 67–72, 2005.
- JAIN, A. K.; DUIN, R. P. W.; MAO, J. Statistical pattern recognition: A review. *Transactions on pattern analysis and machine intelligence*, IEEE, v. 22, n. 1, p. 4–37, 2000.
- JHA, P.; SEN, D. Drought avoidance by cyperus rotundus. *Folia geobotanica & phytotaxonomica*, Springer, v. 15, n. 4, p. 387–394, 1980.
- JOLLIFFE, I. T. Principal component analysis and factor analysis. In: *Principal component analysis*. [S.l.]: Springer, 1986.
- JR, T. G. S.; LANGE, G. D.; MARKS, W. B. Fractal methods and results in cellular morphology—dimensions, lacunarity and multifractals. *Journal of neuroscience methods*, Elsevier, v. 69, n. 2, p. 123–136, 1996.
- JUNQUEIRA, L.; CARNEIRO, J. *Histologia Básica. 12 a Edição*. [S.l.]: Rio de Janeiro: Guanabara Koogan, 2013.
- JUNQUEIRA, L. C. U.; JUNQUEIRA, L. M. M. S. *Técnicas básicas de citologia e histologia*. São Paulo: Santos Editora, 1983.
- JURA-MORAWIEC, J. Formation of amphivasal vascular bundles in dracaena draco stem in relation to rate of cambial activity. *Trees*, Springer, v. 29, n. 5, p. 1493–1499, 2015.

- JURA-MORAWIEC, J. Atypical origin, structure and arrangement of secondary tracheary elements in the stem of the monocotyledonous dragon tree, *dracaena draco*. *Planta*, Springer, v. 245, n. 1, p. 93–99, 2017.
- LEE, H. D. *Seleção de atributos importantes para a extração de conhecimento de bases de dados*. Tese (Doutorado) — Universidade de São Paulo, 2005.
- LINDEBERG, T. Image matching using generalized scale-space interest points. *Journal of Mathematical Imaging and Vision*, Springer, v. 52, n. 1, p. 3–36, 2015.
- LIU, H.; MOTODA, H. *Feature selection for knowledge discovery and data mining*. [S.l.]: Springer Science & Business Media, 2012.
- LIU, M. et al. Plant cell tracking using kalman filter based local graph matching. *Image and Vision Computing*, Elsevier, v. 60, p. 154–161, 2017.
- LIU, M. et al. Automated tracking of stem cell lineages of arabidopsis shoot apex using local graph matching. *The Plant Journal*, Wiley Online Library, v. 62, n. 1, p. 135–147, 2010.
- LIU, T. et al. Multiscale permutation entropy analysis of electrocardiogram. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 471, p. 492–498, 2017.
- LUCAS, W. J. et al. The plant vascular system: evolution, development and functions f. *Journal of integrative plant biology*, Wiley Online Library, v. 55, n. 4, p. 294–388, 2013.
- LUCENA, L.; STOSIC, T. Utilização de lacunaridade para detecção de padrões de imagens de retinas humanas. *Revista da Estatística da Universidade Federal de Ouro Preto*, v. 3, n. 3, p. 789–793, 2014.
- MANDEBROT, B. B. *The fractal geometry of nature*. [S.l.]: Freeman, San Francisco, 1982.
- MARR, D. *Vision: A computational investigation into the human representation and processing of visual information*. [S.l.]: CUMINCAD, 1982.
- MAYOR, L.; PISSARA, J.; SERENO, A. Microstructural changes during osmotic dehydration of parenchymatic pumpkin tissue. *Journal of Food Engineering*, Elsevier, v. 85, n. 3, p. 326–339, 2008.
- MAYOR, L.; SILVA, M.; SERENO, A. Microstructural changes during drying of apple slices. *Drying technology*, Taylor & Francis, v. 23, n. 9-11, p. 2261–2276, 2005.
- MCFARLANE, M. Digital pictures fifty years ago. *Proceedings of the IEEE*, IEEE, v. 60, n. 7, p. 768–770, 1972.
- MENEZES, M. Apontamento básico de morfologia e anatomia vegetal. *IBUSP/DB - UFSCar*, 2016.
- METCALFE, C. R.; CHALK, L. *Anatomy of the Dicotyledons*. [S.l.]: At The Clarendon Press; Oxford, 1950.
- MIAO, Z.; GANDELIN, M.-H.; YUAN, B. A new image shape analysis approach and its application to flower shape analysis. *Image and vision computing*, Elsevier, v. 24, n. 10, p. 1115–1122, 2006.

- MINERVINI, M.; SCHARR, H.; TSAFTARIS, S. A. Image analysis: the new bottleneck in plant phenotyping [applications corner]. *IEEE signal processing magazine*, IEEE, v. 32, n. 4, p. 126–131, 2015.
- MKRTCHYAN, K. et al. Efficient cell segmentation and tracking of developing plant meristem. In: IEEE. *IEEE International Conference on Image Processing*. [S.l.], 2011.
- MODESTO, Z. M. M.; SIQUEIRA, N. J. B. *Botânica*. [S.l.]: EPU, 2002.
- MOLINARO, E. M.; AL. et. *Conceitos e métodos para a formação de profissionais em laboratórios de saúde*. [S.l.]: EPSJV, 2009.
- MOREIRA, C. Microscópio ótico. *Revista de Ciência Elementar*, Casa das Ciências, v. 1, n. 1, p. 1–4, 2013.
- MURRAY, C.; O'MALLEY, M. Segmentation of plant cell pictures. *Image and vision computing*, Elsevier, v. 11, n. 3, p. 155–162, 1993.
- NUNES, J. R. *Biomimética e a moda: a natureza como modelo, medida e mentora na concepção de produtos*. Dissertação (B.S. thesis) — Universidade Tecnológica Federal do Paraná, 2014.
- OGURA, Y. *Comparative anatomy of vegetative organs of the pteridophytes*. [S.l.: s.n.], 1972.
- OLIVEIRA, J.; DUTRA, L.; RENNÓ, C. Aplicação de métodos de extração e seleção de atributos para classificação de regiões. *XII SBSR*, p. 4201–4208, 2005.
- OLIVEIRA, J. de A. et al. Extração de atributos de forma para classificação de imagens de alta resolução do satélite hrc/cbers-2b. *XIV SBSR*, p. 7015–7022, 2009.
- PARZEN, E. On estimation of a probability density function and mode. *The annals of mathematical statistics*, JSTOR, v. 33, n. 3, p. 1065–1076, 1962.
- PAULUS, I.; BUSSCHER, R. D.; SCHREVEN, E. Use of image analysis to investigate human quality classification of apples. *Journal of Agricultural Engineering Research*, Elsevier, v. 68, n. 4, p. 341–353, 1997.
- PLOTZE, R. d. O. et al. Leaf shape analysis using the multiscale minkowski fractal dimension, a new morphometric method: a study with passiflora (passifloraceae). *Canadian Journal of Botany*, NRC Research Press, v. 83, n. 3, p. 287–301, 2005.
- PLOTZE, R. de O.; BRUNO, O. M. Análise de formas e reconhecimento de padrões por meio da assinatura fractal multiescala. *Journal of Computer Science*, INFOCOMP, v. 6, n. 2, p. 91–100, 2007.
- PUIGDEFABREGAS, M. N.; BENASET, A. M. *Pràctiques de fitopatologia*. [S.l.]: Edicions Universitat Barcelona, 2000.
- RAMAN, V. et al. Morpho-anatomy of marijuana (*cannabis sativa* l.). In: *Cannabis sativa L.-Botany and Biotechnology*. [S.l.]: Springer, 2017.
- RAMOS, I. N. et al. Quantification of microstructural changes during first stage air drying of grape tissue. *Journal of Food Engineering*, Elsevier, v. 62, n. 2, p. 159–164, 2004.

- RAVEN, P. H.; EVERT, R. F.; EICHHORN, S. E. *Biologia Vegetal*. [S.l.]: Guanabara Koogan, 2014.
- RICHMOND, P.; ROEHNER, B. M. A physicist's view of the similarities and differences between tuberculosis and cancer. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 534, p. 120761, 2019.
- RODRIGUES, E. P. *Lacunaridade para caracterização de formas de dimensão finita*. Tese (Doutorado) — Universidade de São Paulo, 2014.
- RODRIGUES, E. P.; BARBOSA, M. S.; COSTA, L. F. Self-referred approach to lacunarity. *Physical Review E*, APS, v. 72, n. 1, p. 016707, 2005.
- ROSITO, J. M.; MARCHESAN, E. Determinação de descritores foliares para a identificação micro-histológica de espécies forrageiras. *Acta Scientiarum. Biological Sciences*, v. 25, n. 2, p. 407–413, 2003.
- RUMELHART, D.; HINTON, G.; WILLIAMS, R. Learning internal representation by error propagation. *MIT Press*, Parallel Distributed Processing, 1986.
- SÁNCHEZ-GUTIÉRREZ, D. et al. Fundamental physical cellular constraints drive self-organization of tissues. *The EMBO journal*, EMBO Press, v. 35, n. 1, p. 77–88, 2016.
- SANTAROSA, E. et al. Alterações anatômicas do sistema vascular em porta-enxertos de videira. *Pesquisa Agropecuária Brasileira*, v. 51, n. 4, p. 320–329, 2016.
- SANYAL, P.; BHATTACHARYA, U.; BANDYOPAHYAY, S. K. Analysis of sem images of stomata of different tomato cultivars based on morphological features. In: IEEE. *Modeling & Simulation, 2008. AICMS 08. Second Asia International Conference on*. [S.l.], 2008.
- SHANG, P. et al. Multiscale tsallis permutation entropy analysis for complex physiological time series. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 523, p. 10–20, 2019.
- SHANNON, C. E. A mathematical theory of communication. *The Bell System Technical Journal*, v. 27, n. 1, p. 379–423, 1948.
- SHENFIELD, A.; DAY, D.; AYESH, A. Intelligent intrusion detection systems using artificial neural networks. *ICT Express*, Elsevier, v. 4, n. 2, p. 95–99, 2018.
- SILVA, F. K. A. et al. Anatomical and histochemical study of *Sechium edule* (jacq.) sw. *Anales de Biología*, v. 42, n. 19, p. 173–181, 2020.
- SILVA, I. N. D.; SPATTI, D. H.; FLAUZINO, R. A. *Redes Neurais Artificiais para engenharia e ciências aplicadas curso prático*. [S.l.: s.n.], 2010.
- SILVA, P. M.; FLORINDO, J. B. A statistical descriptor for texture images based on the box counting fractal dimension. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 528, p. 121469, 2019.
- SNEGIREVA, A. et al. Intrusive growth of primary and secondary phloem fibres in hemp stem determines fibre-bundle formation and structure. *AoB Plants*, Oxford Academic, v. 7, p. 1–14, 2015.

- SONG, Z. et al. Scale-specific effects: A report on multiscale analysis of acupunctured eeg in entropy and power. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 492, p. 2260–2272, 2018.
- STACE, C. A. *Plant taxonomy and biosystematics*. [S.l.]: Cambridge University Press, 1991.
- TAIZ, L.; AL. et. *Fisiologia e desenvolvimento vegetal*. [S.l.]: Artmed Editora, 2017.
- TOMLINSON, P. B.; ZIMMERMANN, M. H. Vascular anatomy of monocotyledons with secondary growth—an introduction. *Journal of the Arnold Arboretum*, JSTOR, v. 50, n. 2, p. 159–179, 1969.
- VARELLA, C. A. A. Análise de componentes principais. *Seropédica: Universidade Federal Rural do Rio de Janeiro*, 2008.
- VASCONCELOS, S. Análise de componentes principais (pca). *Instituto de Computação da Universidade Federal Fluminense*, 2007.
- VU, Q. D. et al. Methods for segmentation and classification of digital microscopy tissue images. *Frontiers in bioengineering and biotechnology*, Frontiers, v. 7, p. 53, 2019.
- WU, H. et al. Three-dimensional distribution of vessels, passage cells and lateral roots along the root axis of winter wheat (*triticum aestivum*). *Annals of Botany*, Oxford University Press, v. 107, n. 5, p. 843–853, 2011.
- XING, Q. et al. Using a multiscale image processing method to characterize the periodic growth patterns on scallop shells. *Ecology and evolution*, Wiley Online Library, v. 7, n. 5, p. 1616–1626, 2017.
- YIN, Y. et al. Multiscale joint permutation entropy for complex time series. *Physica A: Statistical Mechanics and its Applications*, Elsevier, v. 515, p. 388–402, 2019.
- ZIRKLE, C. Nucleolus in root tip mitosis in *Zea mays*. *International Journal of Plant Sciences*, Botanical Gazette, v. 86, n. 4, p. 402–418, 1928.

Apêndices

APÊNDICE A – Análise individual: Função densidade probabilidade selecionadas

Essa seção é composta pelos gráficos de função densidade probabilidade que foram selecionados na análise individual dos atributos.

A.1 Feixes vasculares concêntricos

A.1.1 Floema nos feixes vasculares concêntricos

As Figuras 76, 77 e 78 ilustram quais atributos, relacionados a forma, a densidade e a regularidade respectivamente, foram capazes de segregar pelo menos uma classe dos tecidos estudadas e, portanto, relevantes nesse tipo de análise.

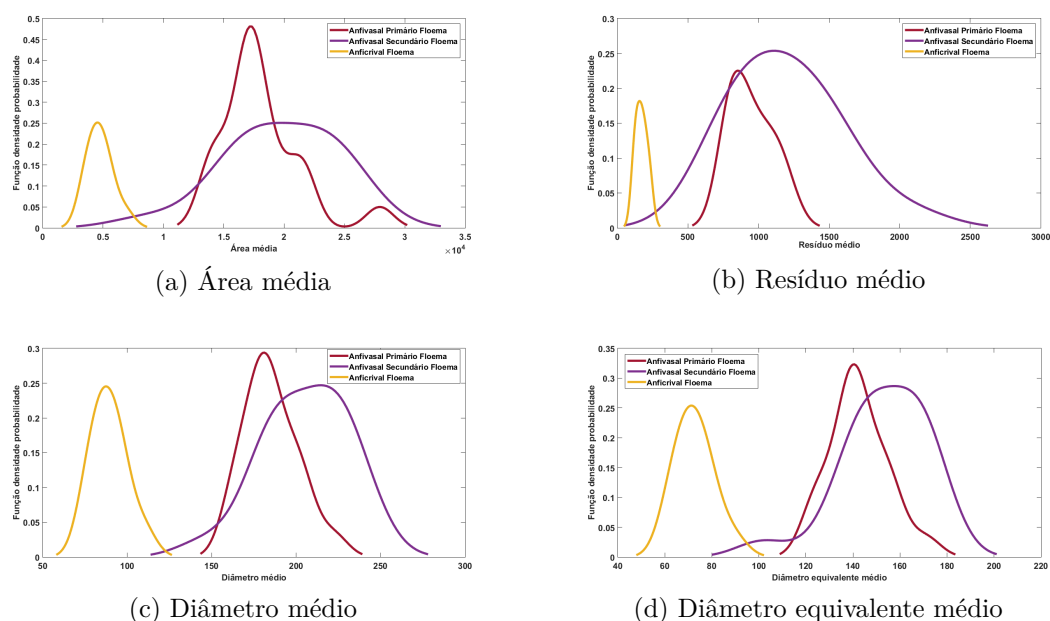


Figura 76 – Função densidade probabilidade, para os atributos de forma, relativa aos feixes concêntricos com ênfase no floema. (a) Área média; (b) Resíduo médio; (c) Diâmetro médio; (d) Diâmetro equivalente médio.

Fonte: Autor

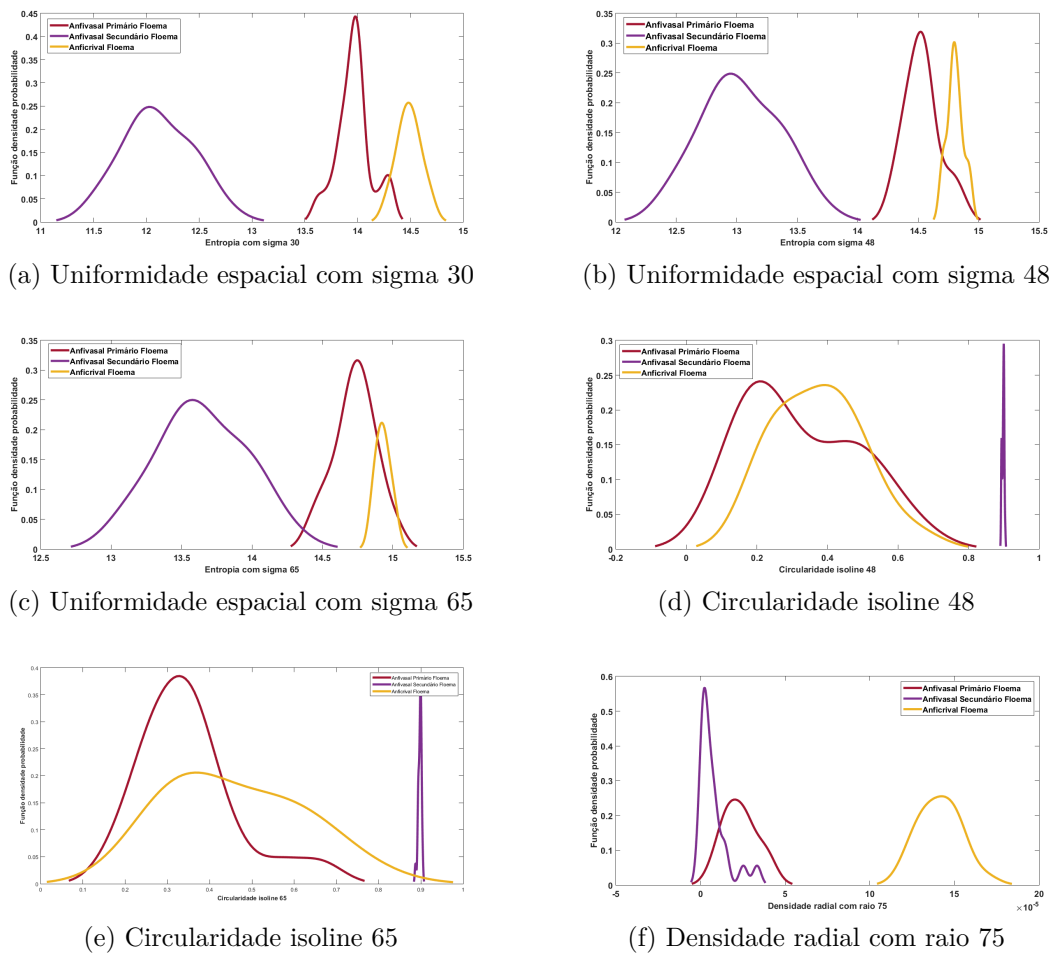


Figura 77 – Função densidade probabilidade, para os atributos de densidade, relativa aos feixes concêntricos com ênfase no floema. (a) Uniformidade espacial com sigma 30; (b) Uniformidade espacial com sigma 48; (c) Uniformidade espacial com sigma 65; (d) Circularidade isoline 48; (e) Circularidade isolone 63; (f) Densidade radial com raio 75.

Fonte: Autor

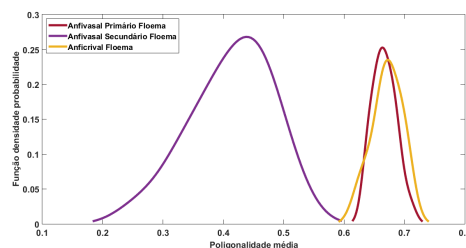


Figura 78 – Função densidade probabilidade do atributo poligonalidade média, relativa aos feixes concêntricos com ênfase no floema.

Fonte: Autor

A.1.2 Xilema nos feixes vasculares concêntricos

As Figuras 79, 80 e 81 ilustram quais atributos foram capazes de segregar pelo menos uma classe dos tecidos estudadas e, portanto, relevantes nesse tipo de análise, sendo que a Figura 79 mostra os atributos referente a forma, a Figura 80 ilustra os atributos relacionados a densidade e a Figura 81 evidencia os atributos relacionados a regularidade estrutural.

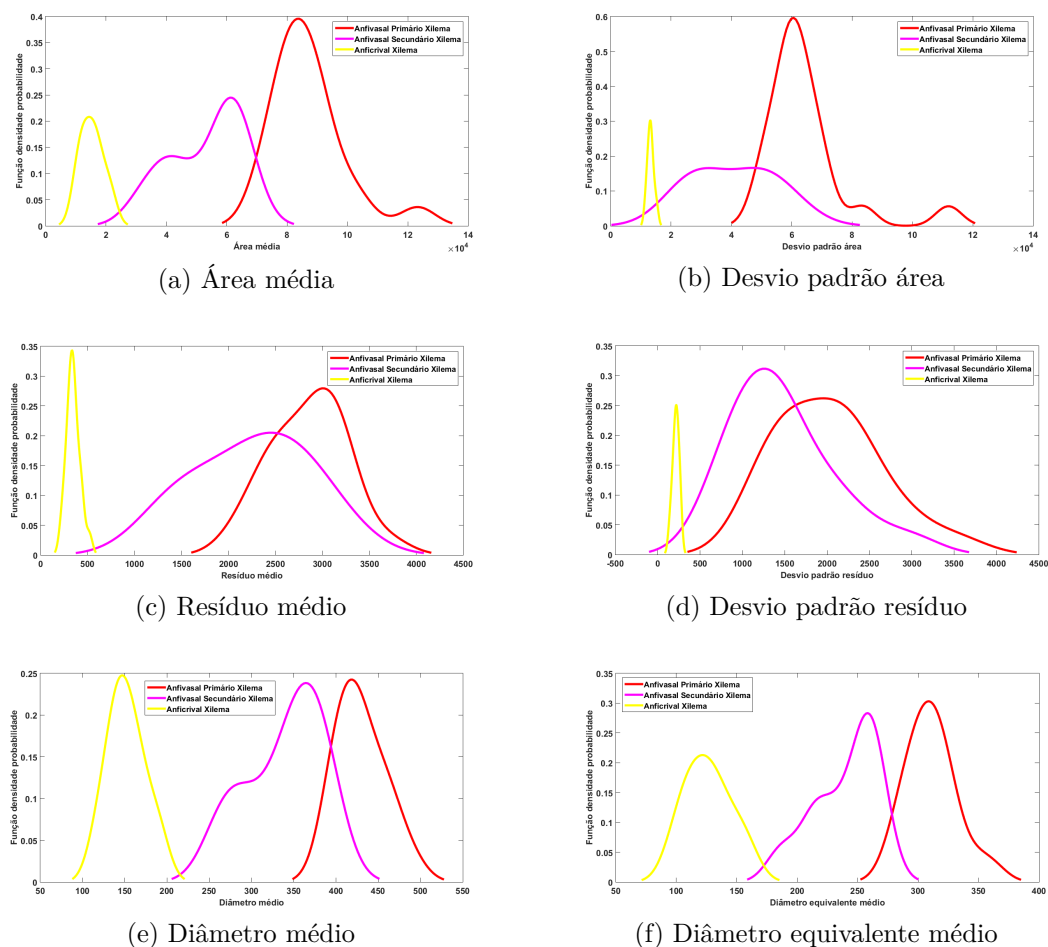


Figura 79 – Função densidade probabilidade, para os atributos de forma, relativa aos feixes concêntricos com ênfase no xilema. (a) Área média; (b) Desvio padrão da área; (c) Resíduo médio; (d) Desvio padrão do resíduo; (e) Diâmetro médio; (f) Diâmetro equivalente médio.

Fonte: Autor

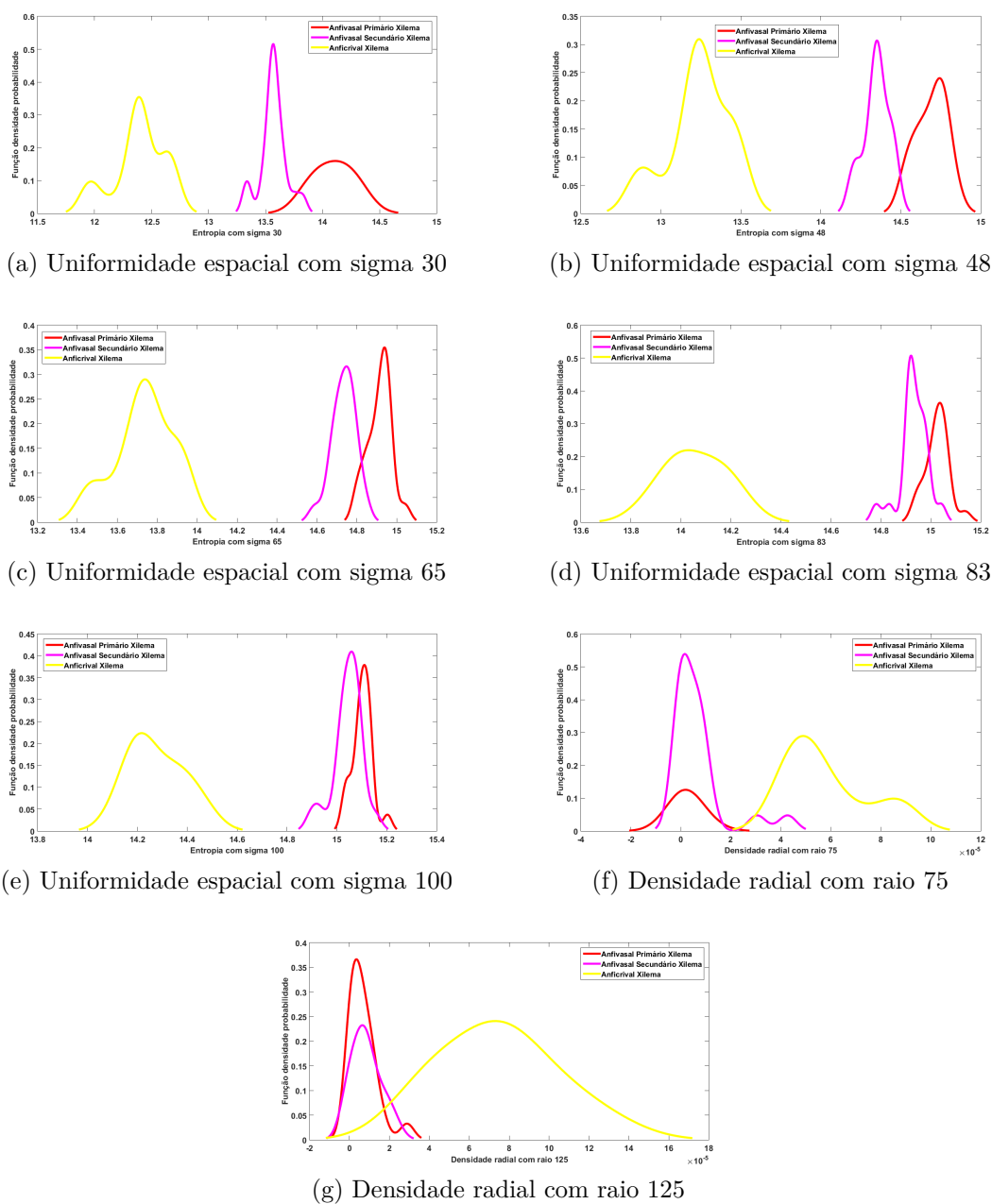


Figura 80 – Função densidade probabilidade, para os atributos de densidade, relativa aos feixes concêntricos com ênfase no xilema. (a) Uniformidade espacial com sigma 30; (b) Uniformidade espacial com sigma 48; (c) Uniformidade espacial com sigma 65; (d) Uniformidade espacial com sigma 83; (e) Uniformidade espacial com sigma 100; (f) Densidade radial com raio 75; (g) Densidade radial com raio 125.

Fonte: Autor

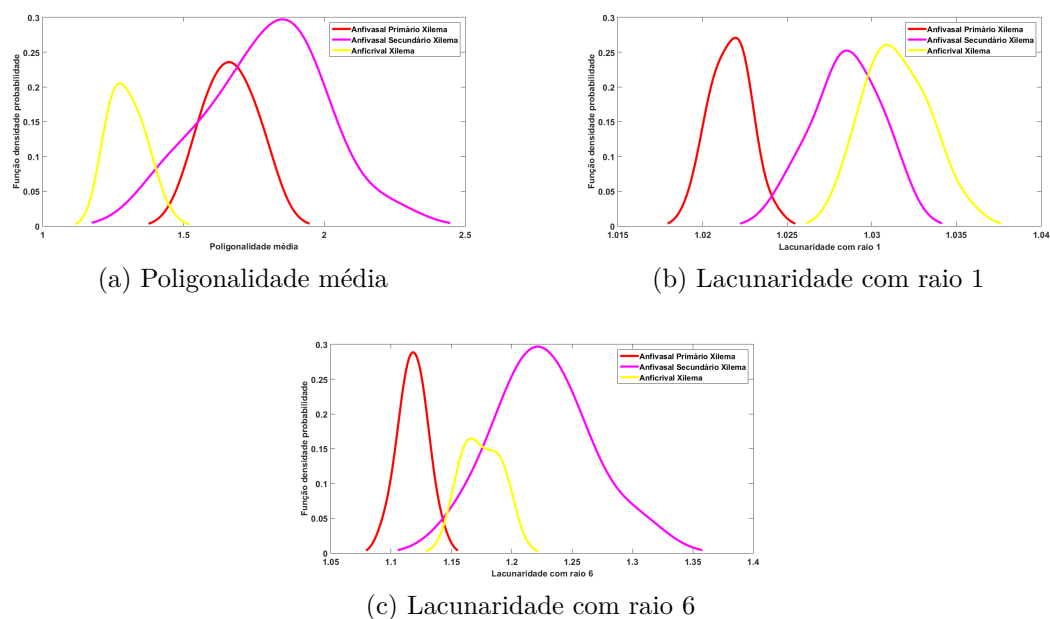


Figura 81 – Função densidade probabilidade, para os atributos de regularidade, relativa aos feixes concêntricos com ênfase no xilema. (a) Média poligonalidade; (b) Lacunaridade com raio 1; (c) Lacunaridade com raio 6.

Fonte: Autor

A.2 Feixes vasculares colaterais

A.2.1 Floema nos feixes vasculares colaterais

As Figuras 82, 83 e 84 ilustram quais atributos, relacionados a forma, a densidade e a regularidade respectivamente, foram capazes de segregar pelo menos uma classe dos tecidos estudadas e, portanto, relevantes nesse tipo de análise.

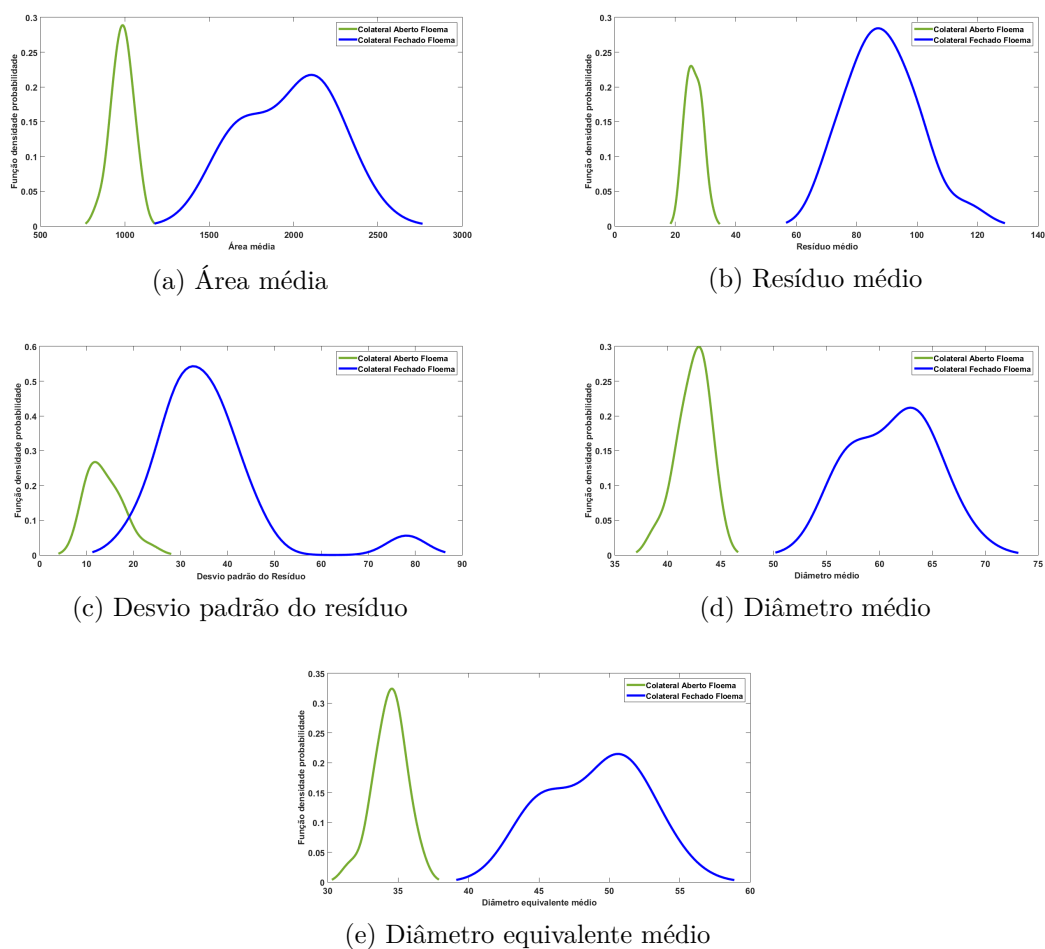


Figura 82 – Função densidade probabilidade, para os atributos de forma, relativa aos feixes colaterais com ênfase no floema. (a) Área média; (b) Resíduo médio; C Desvio padrão do resíduo; (d) Diâmetro médio; (e) Diâmetro equivalente médio.

Fonte: Autor

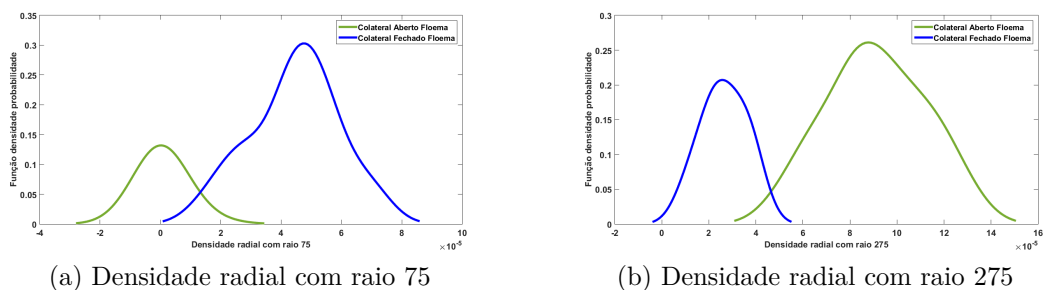


Figura 83 – Função densidade probabilidade, para os atributos de densidade, relativa aos feixes colaterais com ênfase no floema. (a) Densidade radial com raio 75, Densidade radial com raio 275.

Fonte: Autor

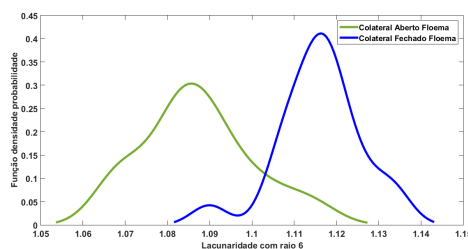


Figura 84 – Função densidade probabilidade do atributo Lacunaridade com raio 6, relativa aos feixes colaterais com ênfase no floema.

Fonte: Autor

A.2.2 Xilema nos feixes vasculares colaterais

As Figuras 85, 86 e 87 ilustram quais atributos foram capazes de segregar pelo menos uma classe dos tecidos estudadas e, portanto, relevantes nesse tipo de análise, sendo que a Figura 85 mostra os atributos referente a forma, a Figura 86 ilustra os atributos relacionados a densidade e a Figura 87 evidencia os atributos relacionados a regularidade estrutural.

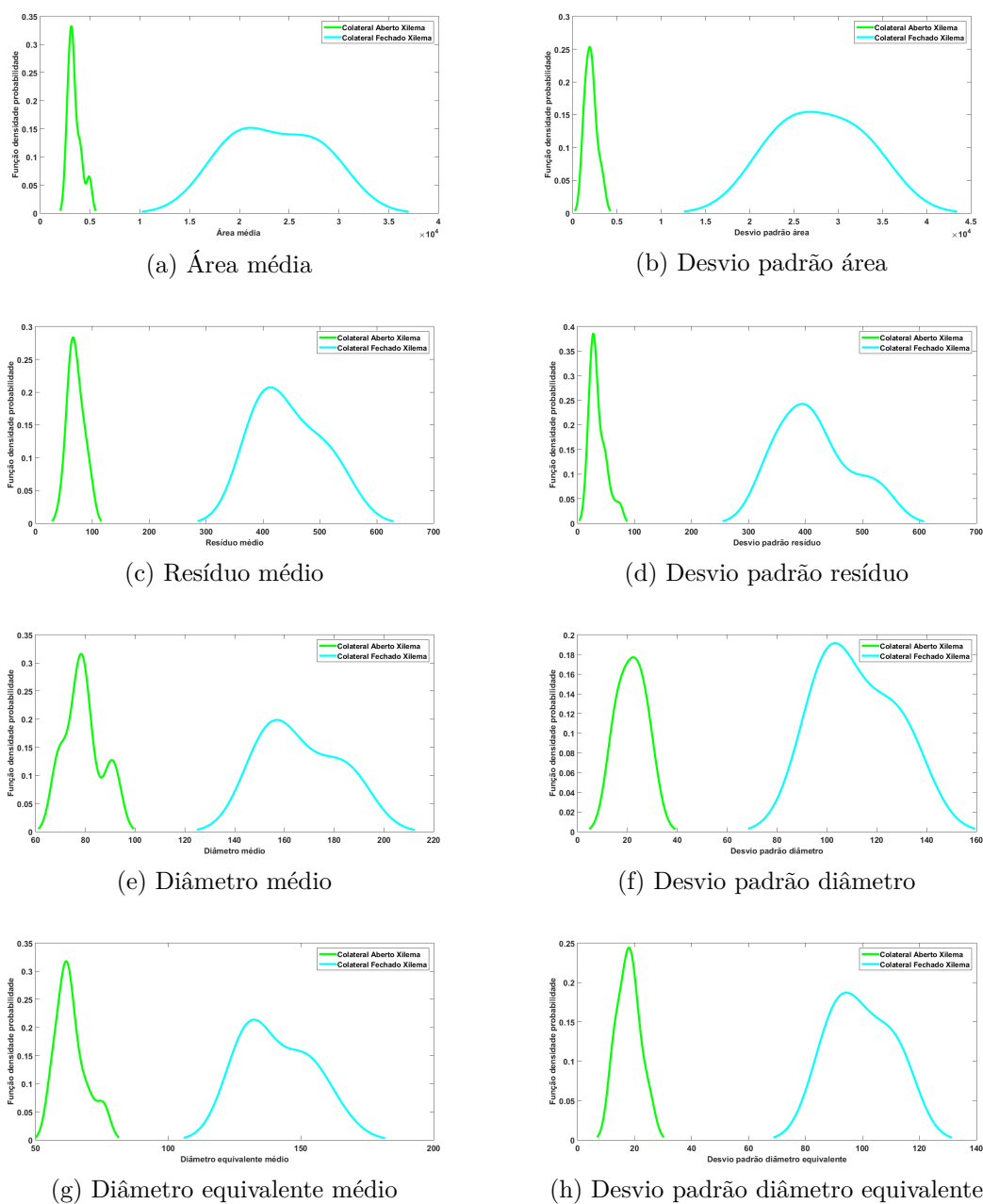


Figura 85 – Função densidade probabilidade, para os atributos de forma, relativa aos feixes colaterais com ênfase no xilema. (a) Área média; (b) Desvio padrão da área; (c) Resíduo médio; (d) Desvio padrão do resíduo; (e) Diâmetro médio; (f) Desvio padrão diâmetro; (g) Diâmetro equivalente médio; (h) Desvio padrão do diâmetro equivalente.

Fonte: Autor

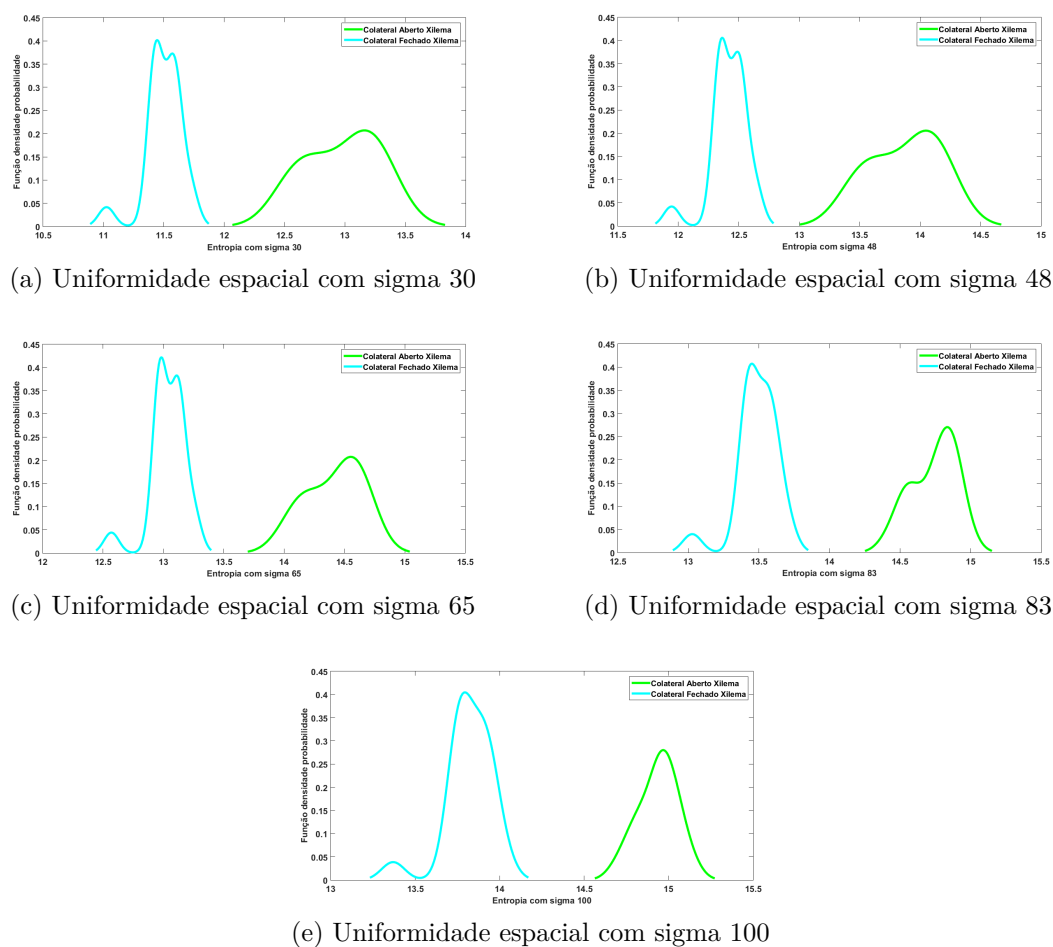


Figura 86 – Função densidade probabilidade, para os atributos de densidade, relativa aos feixes colaterais com ênfase no xilema. (a) Uniformidade espacial com sigma 30; (b) Uniformidade espacial com sigma 48; (c) Uniformidade espacial com sigma 65; (d) Uniformidade espacial com sigma 83 e (e) Uniformidade espacial com sigma 100.

Fonte: Autor

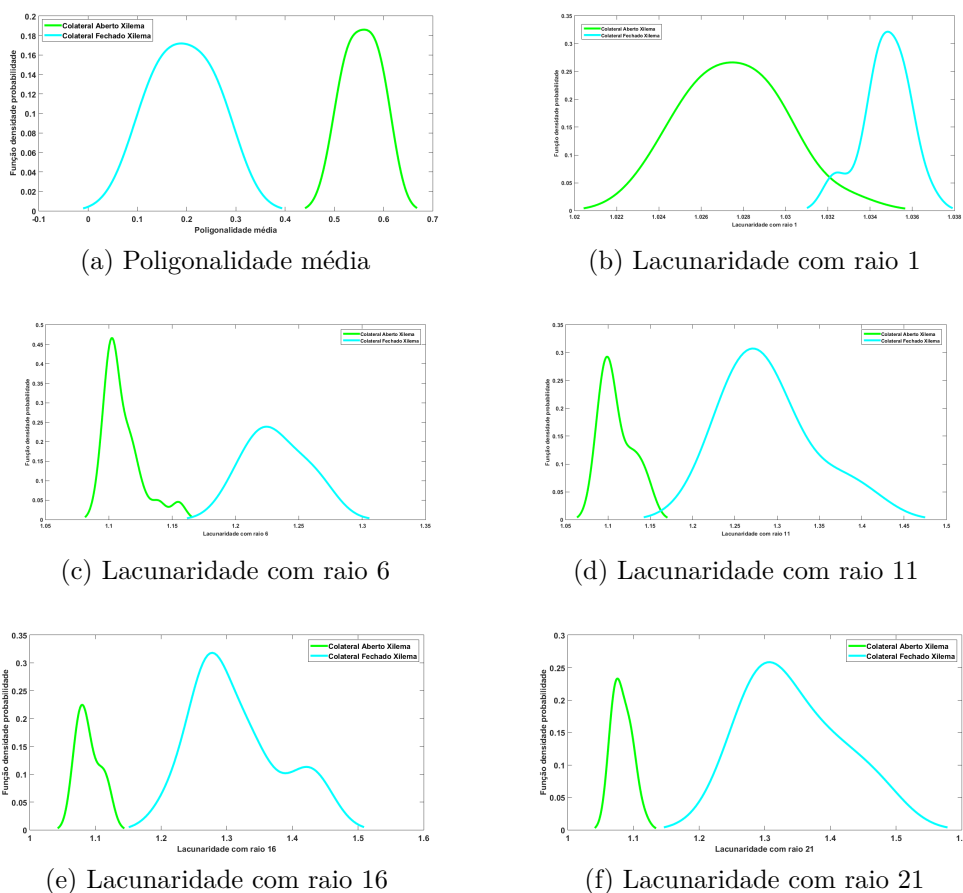


Figura 87 – Função densidade probabilidade, para os atributos de regularidade, relativa aos feixes colaterais com ênfase no xilema. (a) Poligonalidade média, (b) Lacunaridade com raio 1; (c) Lacunaridade com raio 6; (d) Lacunaridade com raio 11; (e) Lacunaridade com raio 16; (f) Lacunaridade com raio 21.

Fonte: Autor

A.3 Feixe vascular biclateral

As Figuras 88, 89 e 90 ilustram quais atributos foram capazes de segregar os tecidos e, portanto, relevantes nesse tipo de análise, sendo que a Figura 88 mostra os atributos referente a forma, a Figura 89 ilustra os atributos relacionados a densidade e a Figura 90 evidencia os atributos relacionados a regularidade estrutural.

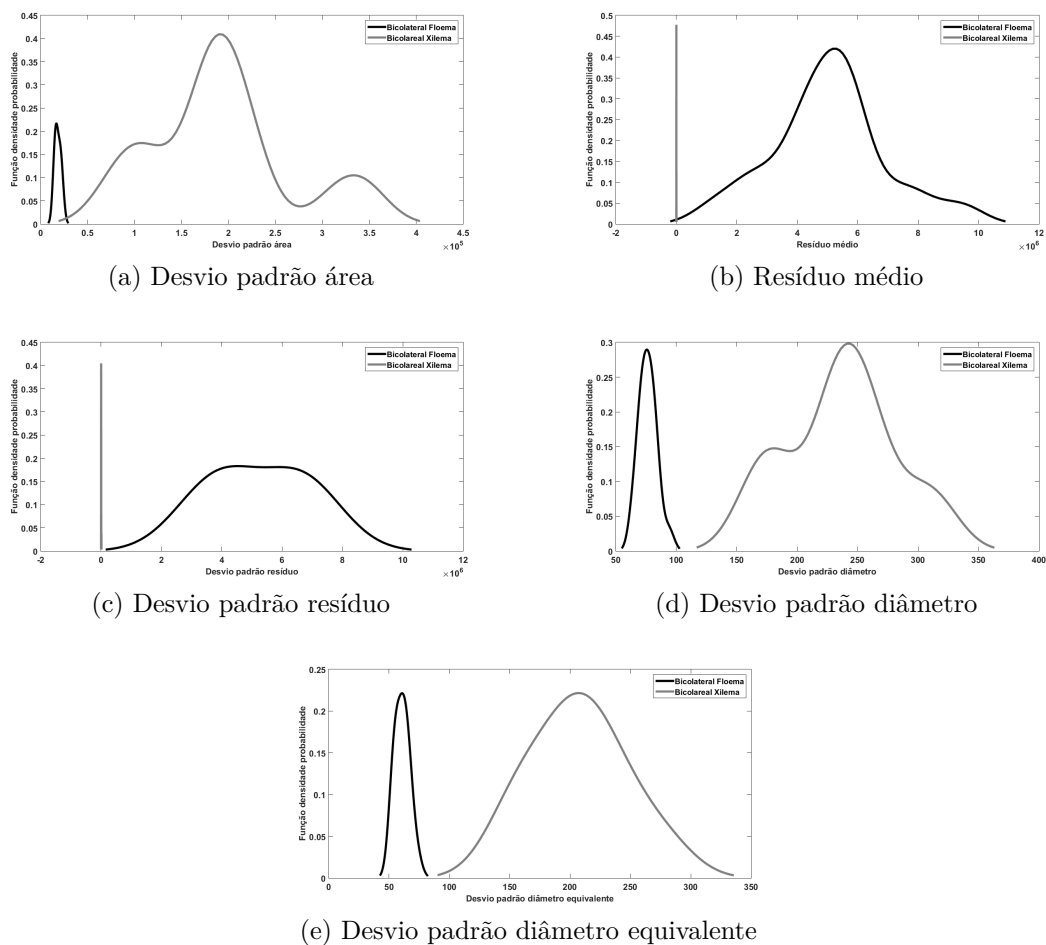


Figura 88 – Função densidade probabilidade para os atributos relacionados a forma capazes de segregar os tecidos dos feixes vasculares bicolaterais. (a) Desvio padrão da área; (b) Resíduo médio; (c) Desvio padrão do resíduo; (d) Desvio padrão diâmetro; (e) Desvio padrão do diâmetro equivalente.

Fonte: Autor

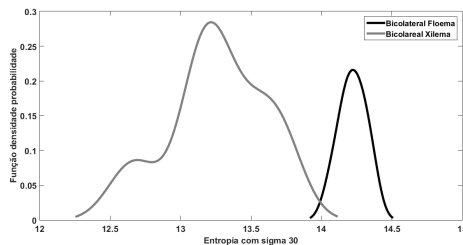
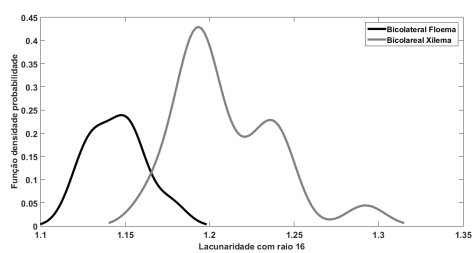
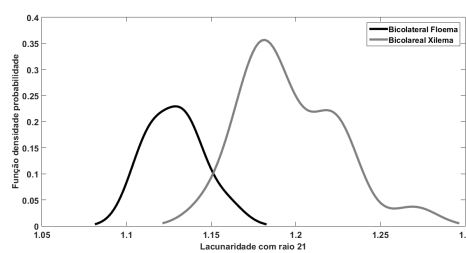


Figura 89 – Função densidade probabilidade para o atributo Uniformidade espacial com sigma 30 em relação aos tecidos vasculares dos feixes vasculares bicolaterais.

Fonte: Autor



(a) Lacunaridade com raio 16



(b) Lacunaridade com raio 21

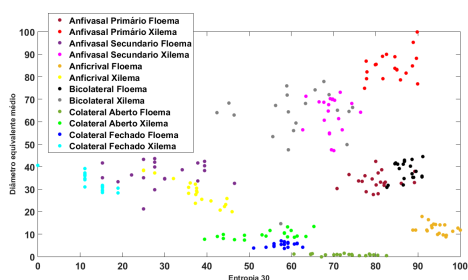
Figura 90 – Função densidade probabilidade para os atributos relacionados a regularidade estrutural capazes de segregar os tecidos vasculares dos feixes vasculares bicolaterais. (a) Lacunaridade com raio 16; (b) Lacunaridade com raio 21.

Fonte: Autor

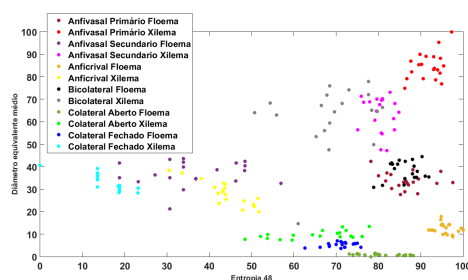
APÊNDICE B – Gráficos de dispersão selecionados

Essa seção é composta pelos gráficos de dispersão que foram selecionados por meio da abordagem matriz de espalhamento.

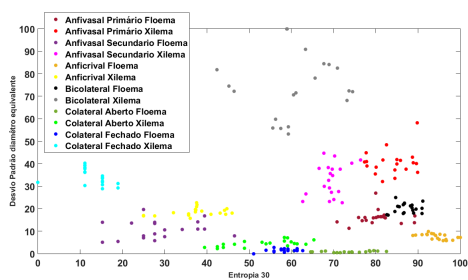
B.1 Todos os tecidos vasculares



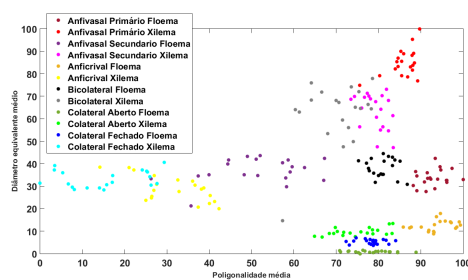
(a) Uniformidade espacial 30 x Diâmetro equivalente médio



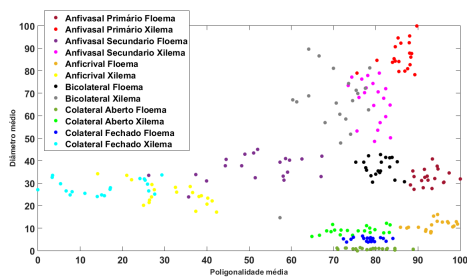
(b) Uniformidade espacial 48 x Diâmetro equivalente médio



(c) Uniformidade espacial 30 x Desv. P. diâmetro eq.



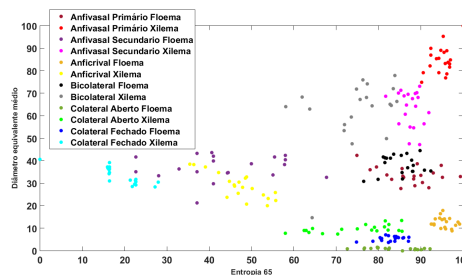
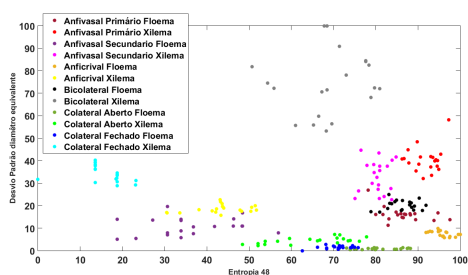
(d) Poligonalidade média x Diâmetro eq. médio



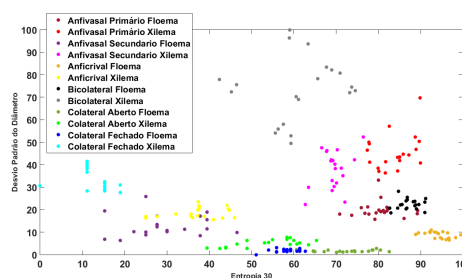
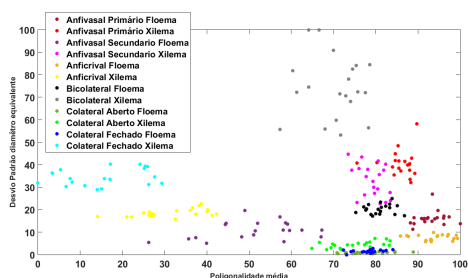
(e) Poligonalidade média x Diâmetro médio

Figura 91 – Gráficos de dispersão com os melhores atributos capazes de segregar os feixes vasculares vegetais estudados. (a) Uniformidade espacial 30 x Diâmetro equivalente médio; (b) Uniformidade espacial 48 x Diâmetro equivalente médio; (c) Uniformidade espacial 30 x Desvio Padrão diâmetro equivalente; (d) Poligonalidade média x Diâmetro equivalente médio; (e) Poligonalidade média x Diâmetro médio.

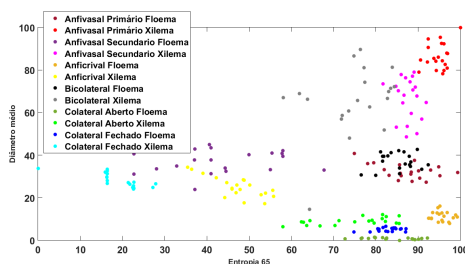
Fonte: Autor



(a) Uniformidade espacial 48 x Desv. P. diâmetro eq. (b) Uniformidade espacial 65 x Diâmetro eq. médio



(c) Uniformidade espacial 30 x Desv. P. diâmetro eq. (d) Uniformidade espacial 30 x Desv. P. diâmetro eq.

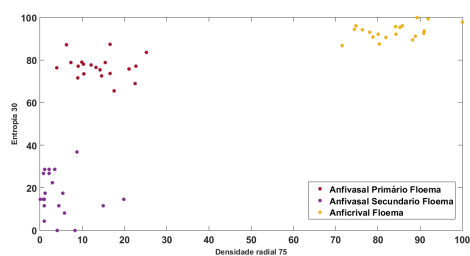


(e) Uniformidade espacial 65 x Diâmetro médio

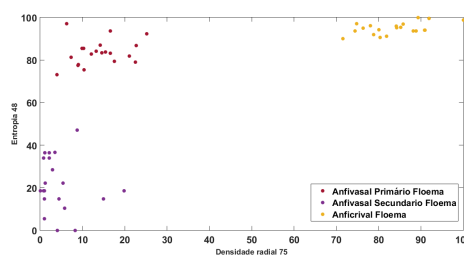
Figura 92 – Gráficos de dispersão com os melhores atributos capazes de segregar os feixes vasculares vegetais estudados (continuação). (a) Uniformidade espacial 48 x Desvio Padrão diâmetro equivalente; (b) Uniformidade espacial 65 x Diâmetro equivalente médio; (c) Uniformidade espacial 30 x Desv. P. diâmetro equivalente; (d) Uniformidade espacial 30 x Desvio Padrão diâmetro; (e) Uniformidade espacial 65 x Diâmetro médio.

Fonte: Autor

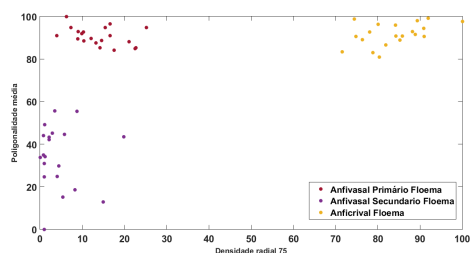
B.2 Feixes vasculares do tipo concêntrico: Floema



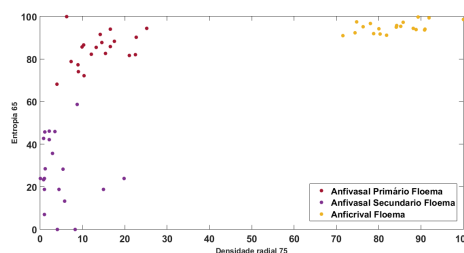
(a) Densidade radial 75 x Uniformidade espacial 30



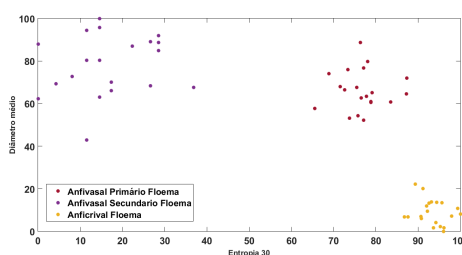
(b) Densidade radial 75 x Uniformidade espacial 48



(c) Densidade radial 75 x Poligonalidade média



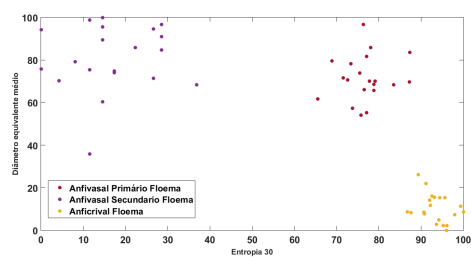
(d) Densidade radial 75 x Uniformidade espacial 65



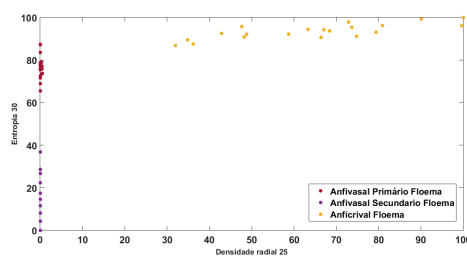
(e) Uniformidade espacial 30 x Diâmetro médio

Figura 93 – Gráficos de dispersão com os melhores atributos capazes de segregar os floemas dos feixes vasculares vegetais do tipo concêntrico. (a) Densidade radial 75 x Uniformidade espacial 30; (b) Densidade radial 75 x Uniformidade espacial 48; (c) Densidade radial 75 x Poligonalidade média; (d) Densidade radial 75 x Uniformidade espacial 65; (e) Uniformidade espacial 30 x Diâmetro médio.

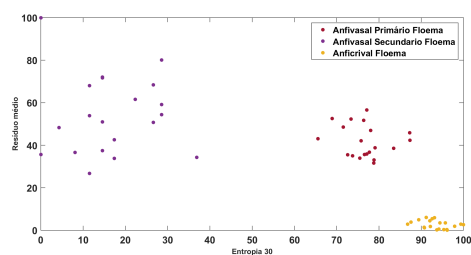
Fonte: Autor



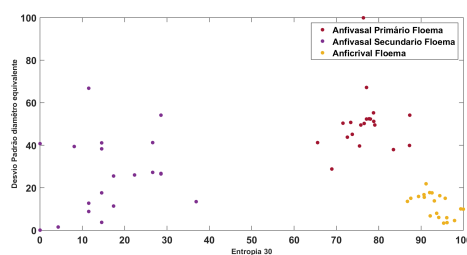
(a) Uniformidade espacial 30 x Diâmetro eq. médio



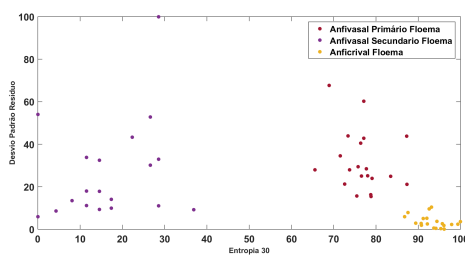
(b) Densidade radial 25 x Uniformidade espacial 30



(c) Uniformidade espacial 30 x Resíduo médio



(d) Uniformidade espacial 30 x Desv. P. diâmetro eq.

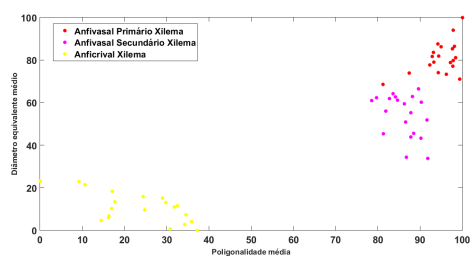


(e) Uniformidade espacial 30 x Desv. P. resíduo

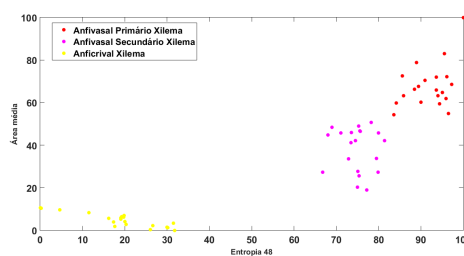
Figura 94 – Gráficos de dispersão com os melhores atributos capazes de segregar os floemas dos feixes vasculares vegetais do tipo concêntrico (continuação). (a) Uniformidade espacial 30 x Diâmetro equivalente médio; (b) Densidade radial 25 x Uniformidade espacial 30; (c) Uniformidade espacial 30 x Resíduo médio; (d) Uniformidade espacial 30 x Desvio Padrão diâmetro equivalente e (e) Uniformidade espacial 30 x Desvio Padrão resíduo.

Fonte: Autor

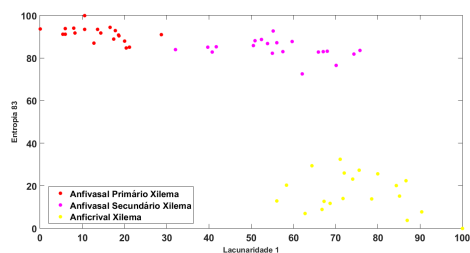
B.3 Feixes vasculares do tipo concêntrico: Xilema



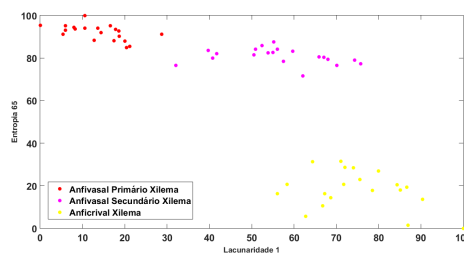
(a) Poligonalidade média x Diâmetro eq. médio



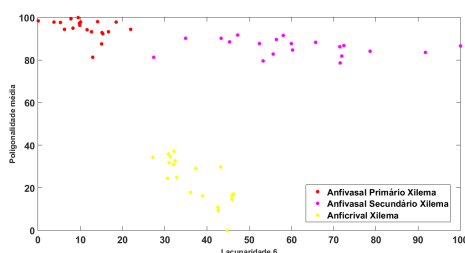
(b) Uniformidade espacial 48 x Área média



(c) Lacunaridade 1 x Uniformidade espacial 83



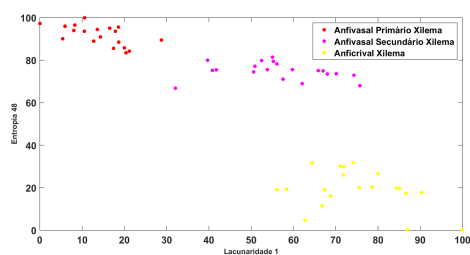
(d) Lacunaridade 1 x Uniformidade espacial 65



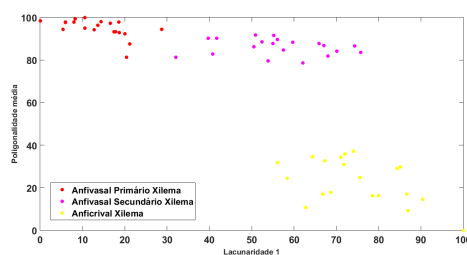
(e) Lacunaridade 6 x Poligonalidade média

Figura 95 – Gráficos de dispersão com os melhores atributos capazes de segregar os xilemas dos feixes vasculares vegetais do tipo concêntrico. (a) Poligonalidade média x Diâmetro equivalente médio; (b) Uniformidade espacial 48 x Área média; (c) Lacunaridade 1 x Uniformidade espacial 83; (d) Lacunaridade 1 x Uniformidade espacial 65; (e) Lacunaridade 6 x Poligonalidade média.

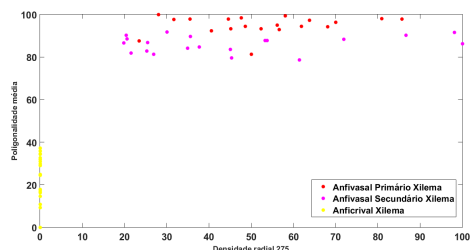
Fonte: Autor



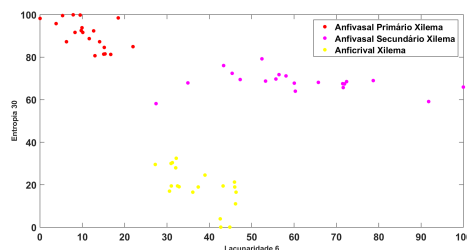
(a) Lacunaridade 1 x Uniformidade espacial 48



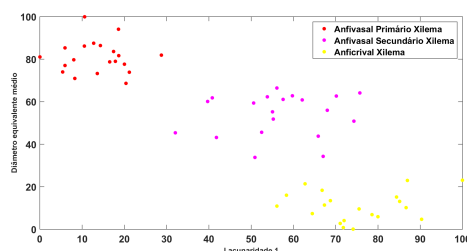
(b) Lacunaridade 1 x Poligonalidade média



(c) Densidade radial 275 x Poligonalidade média



(d) Lacunaridade 6 x Uniformidade espacial 30



(e) Lacunaridade 1 x Diâmetro eq. médio

Figura 96 – Gráficos de dispersão com os melhores atributos capazes de segregar os xilemas dos feixes vasculares vegetais do tipo concêntrico (continuação). (a) Lacunaridade 1 x Uniformidade espacial 48; (b) Lacunaridade 1 x Poligonalidade média; (c) Densidade radial 275 x Poligonalidade média; (d) Lacunaridade 6 x Uniformidade espacial 30; e (e) Lacunaridade 1 x Diâmetro equivalente médio.

Fonte: Autor

B.4 Feixes vasculares do tipo colateral: Floema

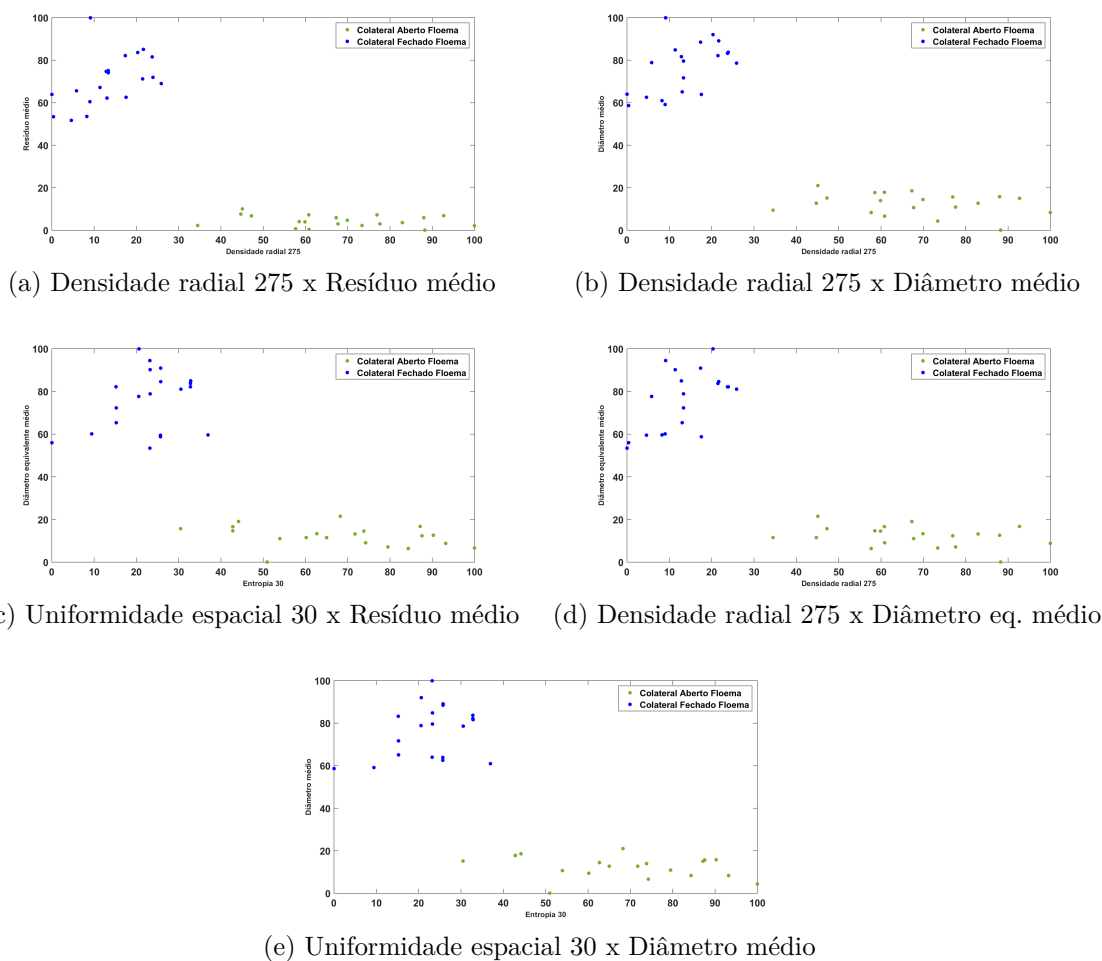
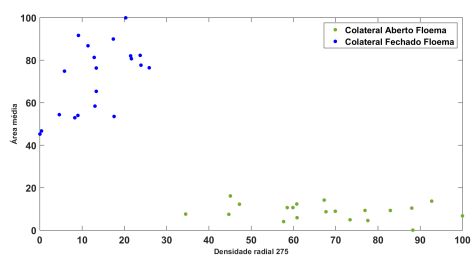
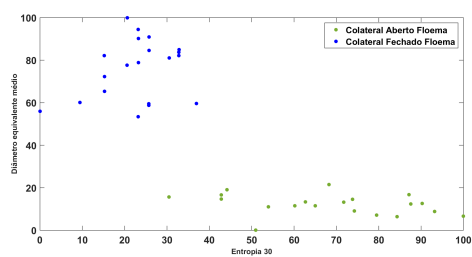


Figura 97 – Gráficos de dispersão com os melhores atributos capazes de segregar os floemas dos feixes vasculares vegetais do tipo colateral. (a) Densidade radial 275 x Resíduo médio; (b) Densidade radial 275 x Diâmetro médio; (c) Uniformidade espacial 30 x Resíduo médio; (d) Densidade radial 275 x Diâmetro equivalente médio; (e) Uniformidade espacial 30 x Diâmetro médio.

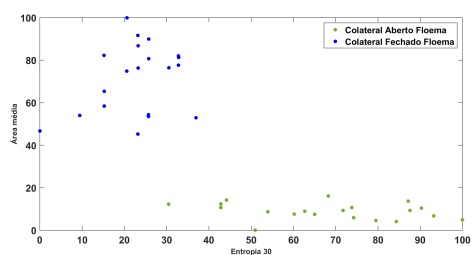
Fonte: Autor



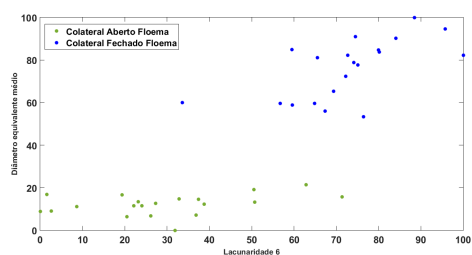
(a) Densidade radial 275 x Área média



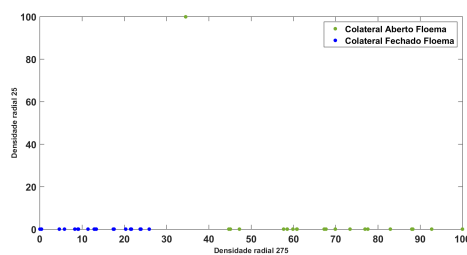
(b) Uniformidade espacial 30 x Diâmetro eq. médio



(c) Uniformidade espacial 30 x Área média



(d) Lacunaridade 6 x Diâmetro eq. médio



(e) Densidade radial 275 x Densidade radial 25

Figura 98 – Gráficos de dispersão com os melhores atributos capazes de segregar os floemas dos feixes vasculares vegetais do tipo colateral (continuação). (a) Densidade radial 275 x Área média; (b) Uniformidade espacial 30 x Diâmetro equivalente médio; (c) Uniformidade espacial 30 x Área média; (d) Lacunaridade 6 x Diâmetro equivalente médio e (e) Densidade radial 275 x Densidade radial 25.

Fonte: Autor

B.5 Feixes vasculares do tipo colateral: Xilema

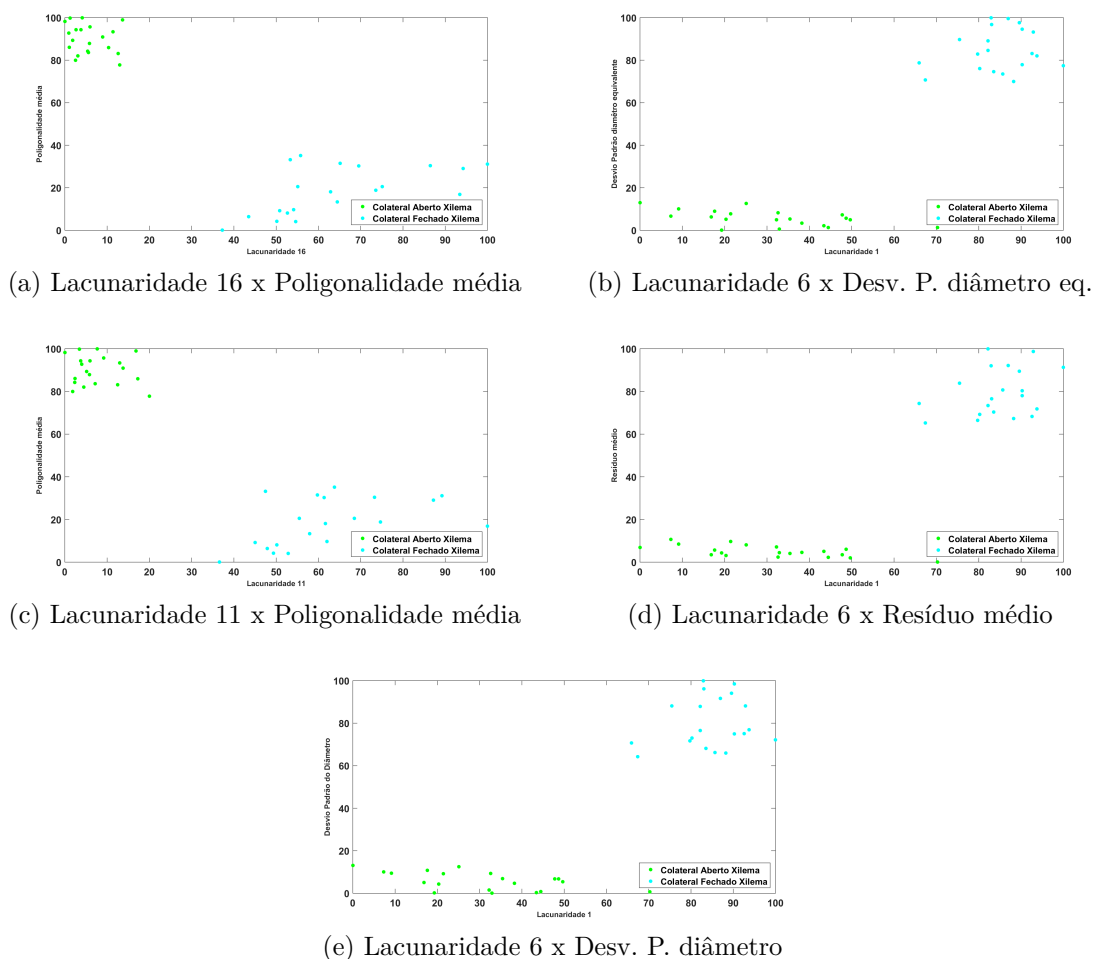
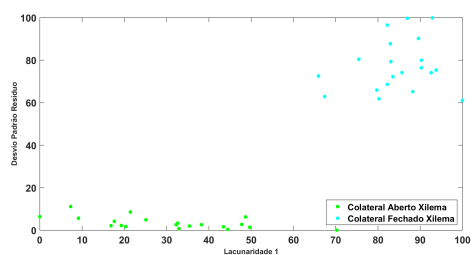
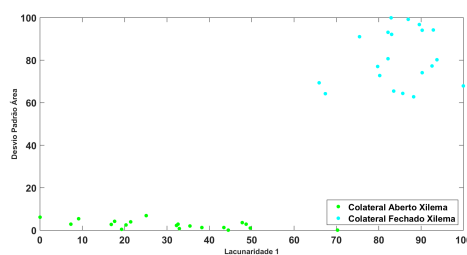


Figura 99 – Gráficos de dispersão com os melhores atributos capazes de segregar os xilemas dos feixes vasculares vegetais do tipo colateral. (a) Lacunaridade 16 x Poligonalidade média; (b) Lacunaridade 6 x Desv. P. diâmetro equivalente; (c) Lacunaridade 11 x Poligonalidade média; (d) Lacunaridade 6 x Resíduo médio; (e) Lacunaridade 6 x Desvio Padrão diâmetro.

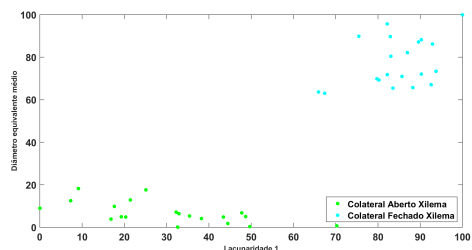
Fonte: Autor



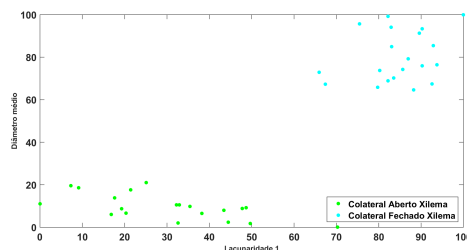
(a) Lacunaridade 6 x Desv. P. resíduo



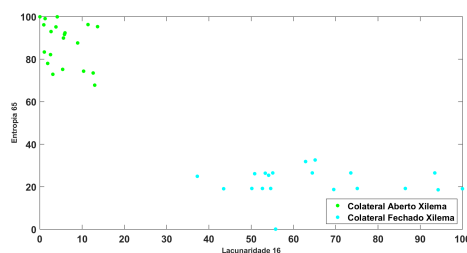
(b) Lacunaridade 6 x Desv. P. área



(c) Lacunaridade 6 x Diâmetro eq. médio



(d) Lacunaridade 6 x Diâmetro médio

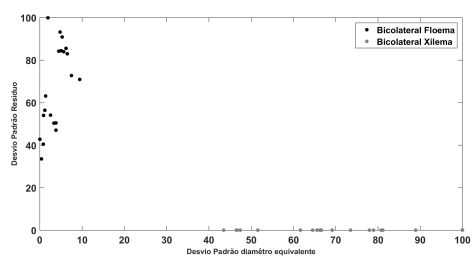


(e) Lacunaridade 16 x Uniformidade espacial 65

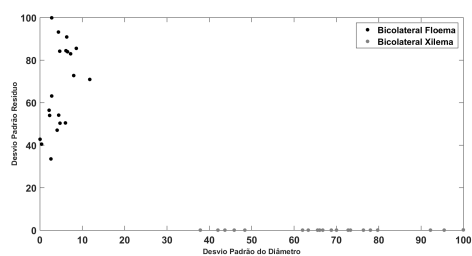
Figura 100 – Gráficos de dispersão com os melhores atributos capazes de segregar os xilemas dos feixes vasculares vegetais do tipo colateral (continuação). (a) Lacunaridade 6 x Desvio Padrão resíduo; (b) Lacunaridade 6 x Desvio Padrão área; (c) Lacunaridade 6 x Diâmetro equivalente médio; (d) Lacunaridade 6 x Diâmetro médio e (e) Lacunaridade 16 x Uniformidade espacial 65.

Fonte: Autor

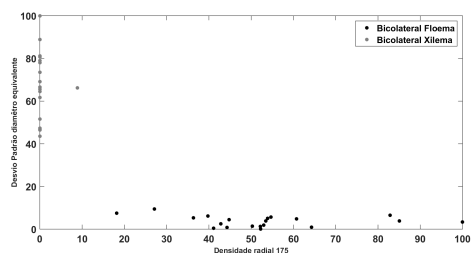
B.6 Feixes vasculares do tipo bicolateral



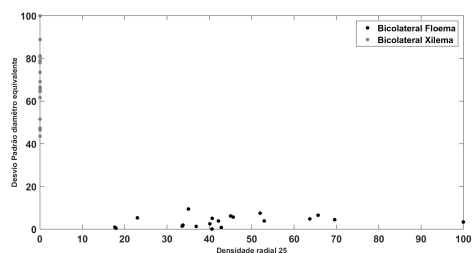
(a) Desv. P. diâmetro eq. x Desv. P. resíduo



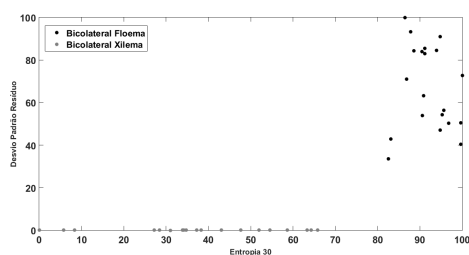
(b) Desv. P. diâmetro x Desv. P. resíduo



(c) Densidade radial 175 x Desv. P. diâmetro eq.



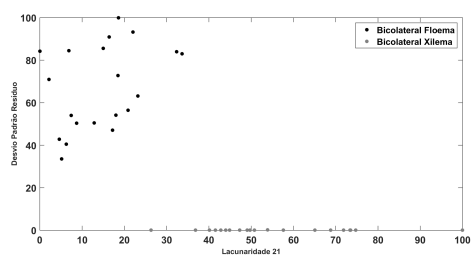
(d) Densidade radial 25 x Desv. P. diâmetro eq.



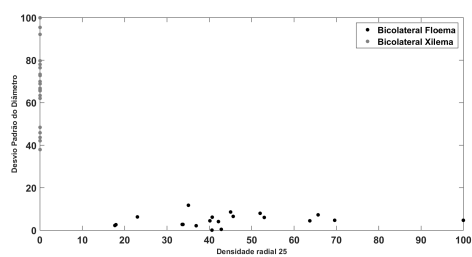
(e) Uniformidade espacial 30 x Desv. P. resíduo

Figura 101 – Gráficos de dispersão com os melhores atributos capazes de segregar os feixes vasculares vegetais do tipo bicolateral. (a) Desvio Padrão diâmetro equivalente x Desvio Padrão resíduo; (b) Desvio Padrão diâmetro x Desvio Padrão resíduo; (c) Densidade radial 175 x Desvio Padrão diâmetro equivalente; (d) Densidade radial 25 x Desvio Padrão diâmetro equivalente; (e) Uniformidade espacial 30 x Desvio Padrão resíduo.

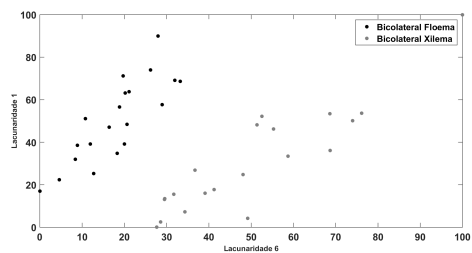
Fonte: Autor



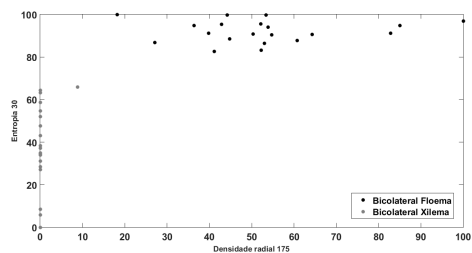
(a) Lacunaridade 21 x Desv. P. resíduo



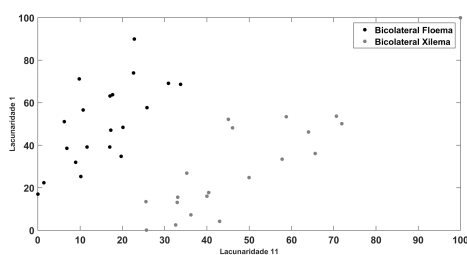
(b) Densidade radial 25 x Desv. P. diâmetro



(c) Lacunaridade 6 x Lacunaridade 1



(d) Densidade radial 175 x Uniformidade espacial 30



(e) Lacunaridade 11 x Lacunaridade 1

Figura 102 – Gráficos de dispersão com os melhores atributos capazes de segregar os feixes vasculares vegetais do tipo bicolateral (continuação). (a) Lacunaridade 21 x Desvio Padrão resíduo; (b) Densidade radial 25 x Desvio Padrão diâmetro; (c) Lacunaridade 6 x Lacunaridade 1; (d) Densidade radial 175 x Uniformidade espacial 30 e (e) Lacunaridade 11 x Lacunaridade 1.

Fonte: Autor

APÊNDICE C – *k*-NN: Matrizes de confusão

C.1 Forma

C.1.1 Espaço original

O espaço original dessa categoria de atributos é composto por 10 dimensões. A matriz de confusão referente ao espaço original da categoria Forma é mostrada na Figura 103. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anfici. Flo.	Anfici. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	2	0	1	0	0	1	1	0	0	0	0	0
	Anf. Prim. Xi.	0	4	0	1	0	0	0	0	0	0	0	0
	Anf. Sec. Flo.	1	0	4	0	0	0	0	0	0	0	0	0
	Anf. Sec. Xi.	0	4	0	1	0	0	0	0	0	0	0	0
	Anfici. Flo.	0	0	0	0	4	0	0	0	0	1	0	0
	Anfici. Xi.	2	0	0	0	0	3	0	0	0	0	0	0
	Bico. Flo.	3	0	0	0	0	0	2	0	0	0	0	0
	Bico. Xi.	0	0	0	0	0	0	0	5	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	0	0	0	0	2	0	3	0
	Col. Ab. Xi.	0	0	0	0	1	0	0	0	0	4	0	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	0	0	5	0
	Col. Fech. Xi.	1	0	0	0	0	0	0	0	0	0	0	4

Figura 103 – Matriz de confusão da abordagem *k*-NN para a categoria Forma no espaço original.

Fonte: Autor

A acurácia dessa matriz de confusão é 66,67%.

C.1.2 Dimensionalidade reduzida: 2D

A matriz de confusão referente ao primeiro e ao segundo componentes principais da categoria Forma é mostrada na Figura 104. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anfici. Flo.	Anfici. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	2	0	1	0	0	2	0	0	0	0	0	0
	Anf. Prim. Xi.	0	1	0	4	0	0	0	0	0	0	0	0
	Anf. Sec. Flo.	1	0	2	0	0	2	0	0	0	0	0	0
	Anf. Sec. Xi.	0	0	0	5	0	0	0	0	0	0	0	0
	Anfici. Flo.	0	0	0	0	0	0	0	0	0	5	0	0
	Anfici. Xi.	0	0	0	0	0	3	0	0	0	0	0	2
	Bico. Flo.	0	0	3	0	0	0	2	0	0	0	0	0
	Bico. Xi.	0	1	0	0	0	0	1	3	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	0	0	0	0	0	2	3	0
	Col. Ab. Xi.	0	0	0	0	1	2	0	0	0	2	0	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	2	0	3	0
	Col. Fech. Xi.	0	0	0	0	0	1	0	0	0	0	0	4

Figura 104 – Matriz de confusão da abordagem *k*-NN para a categoria Forma após aplicação da ACP com dois componentes principais.

Fonte: Autor

A acurácia dessa matriz de confusão é de 45%.

C.1.3 Dimensionalidade reduzida: 3D

A matriz de confusão referente ao primeiro, segundo e terceiro componentes principais da categoria Forma é mostrada na Figura 105. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anfici. Flo.	Anfici. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	1	0	2	0	0	2	0	0	0	0	0	0
	Anf. Prim. Xi.	0	1	0	4	0	0	0	0	0	0	0	0
	Anf. Sec. Flo.	2	0	1	0	0	2	0	0	0	0	0	0
	Anf. Sec. Xi.	0	0	0	5	0	0	0	0	0	0	0	0
	Anfici. Flo.	0	0	0	0	5	0	0	0	0	0	0	0
	Anfici. Xi.	2	0	0	0	0	3	0	0	0	0	0	0
	Bico. Flo.	0	0	3	0	0	0	2	0	0	0	0	0
	Bico. Xi.	0	1	0	0	0	0	1	3	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	0	0	0	0	0	5	0	0
	Col. Ab. Xi.	0	0	0	0	1	0	0	0	0	4	0	0
	Col. Fech. Flo.	0	0	0	0	3	0	0	0	0	2	0	0
	Col. Fech. Xi.	1	0	0	0	0	0	0	0	0	0	0	4

Figura 105 – Matriz de confusão da abordagem *k*-NN para a categoria Forma após aplicação da ACP com três componentes principais.

Fonte: Autor

A acurácia dessa matriz de confusão é de 48,33%.

C.2 Densidade

C.2.1 Espaço original

O espaço original dessa categoria de atributos é composto por 16 dimensões. A matriz de confusão referente ao espaço original da categoria Densidade é mostrada na Figura 106. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anfici. Flo.	Anfici. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	2	0	0	0	0	0	0	0	3	0	0	0
	Anf. Prim. Xi.	0	0	0	4	0	0	0	0	0	1	0	0
	Anf. Sec. Flo.	0	0	4	0	0	0	0	0	0	1	0	0
	Anf. Sec. Xi.	0	0	0	2	0	0	0	1	0	2	0	0
	Anfici. Flo.	0	0	0	0	5	0	0	0	0	0	0	0
	Anfici. Xi.	0	0	0	0	0	4	0	0	0	0	0	1
	Bico. Flo.	0	0	0	0	0	0	4	1	0	0	0	0
	Bico. Xi.	0	0	0	0	0	0	0	5	0	0	0	0
	Col. Ab. Flo.	1	0	0	0	1	0	0	0	3	0	0	0
	Col. Ab. Xi.	1	0	0	1	0	0	0	0	0	3	0	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	0	0	5	0
	Col. Fech. Xi.	0	0	0	0	0	0	0	0	0	0	0	5

Figura 106 – Matriz de confusão da abordagem *k*-NN para a categoria Densidade no espaço original.

Fonte: Autor

A acurácia dessa matriz de confusão é 70%.

C.2.2 Dimensionalidade reduzida: 2D

A matriz de confusão referente ao primeiro e ao segundo componentes principais da categoria Densidade é mostrada na Figura 107. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anfici. Flo.	Anfici. Xi.	Bico. Flo	Bico. Xi	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	2	0	0	0	0	0	0	0	3	0	0	0
	Anf. Prim. Xi.	0	0	0	1	0	0	1	3	0	0	0	0
	Anf. Sec. Flo.	0	0	4	0	0	0	0	0	0	0	0	1
	Anf. Sec. Xi.	0	0	0	1	0	0	2	2	0	0	0	0
	Anfici. Flo.	0	0	0	0	5	0	0	0	0	0	0	0
	Anfici. Xi.	0	0	0	0	0	5	0	0	0	0	0	0
	Bico. Flo	0	0	0	2	0	0	0	1	0	2	0	0
	Bico. Xi	0	0	0	0	0	0	0	2	0	3	0	0
	Col. Ab. Flo.	0	0	0	0	1	0	1	0	2	1	0	0
	Col. Ab. Xi.	1	0	0	0	0	0	1	1	0	2	0	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	0	0	5	0
	Col. Fech. Xi.	0	0	0	0	0	0	0	0	0	0	0	5

Figura 107 – Matriz de confusão da abordagem *k*-NN para a categoria Densidade após aplicação da ACP com dois componentes principais.

Fonte: Autor

A acurácia dessa matriz de confusão é de 55%.

C.2.3 Dimensionalidade reduzida: 3D

A matriz de confusão referente ao primeiro, segundo e terceiro componentes principais da categoria Densidade é mostrada na Figura 105. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anfici. Flo.	Anfici. Xi.	Bico. Flo	Bico. Xi	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	2	0	0	0	0	0	0	0	3	0	0	0
	Anf. Prim. Xi.	0	0	0	3	0	0	1	0	0	1	0	0
	Anf. Sec. Flo.	0	0	4	0	0	0	0	0	0	1	0	0
	Anf. Sec. Xi.	0	0	0	2	0	0	0	1	0	2	0	0
	Anfici. Flo.	0	0	0	0	5	0	0	0	0	0	0	0
	Anfici. Xi.	0	0	0	0	0	5	0	0	0	0	0	0
	Bico. Flo	0	1	0	0	0	0	1	3	0	0	0	0
	Bico. Xi	0	0	0	0	0	0	1	4	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	1	0	1	0	3	0	0	0
	Col. Ab. Xi.	1	0	0	1	0	0	0	0	0	3	0	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	0	0	5	0
	Col. Fech. Xi.	0	0	0	0	0	0	0	0	0	0	0	5

Figura 108 – Matriz de confusão da abordagem *k*-NN para a categoria Densidade após aplicação da ACP com três componentes principais.

Fonte: Autor

A acurácia dessa matriz de confusão é de 65%.

C.3 Regularidade estrutural

C.3.1 Espaço original

O espaço original dessa categoria de atributos é composto por 7 dimensões. A matriz de confusão referente ao espaço original da categoria Regularidade estrutural é mostrada na Figura 109. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anfici. Flo.	Anfici. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	1	0	0	0	0	0	4	0	0	0	0	0
	Anf. Prim. Xi.	0	0	0	0	0	0	0	5	0	0	0	0
	Anf. Sec. Flo.	0	0	3	0	0	1	0	1	0	0	0	0
	Anf. Sec. Xi.	0	0	1	4	0	0	0	0	0	0	0	0
	Anfici. Flo.	0	0	0	0	4	0	0	0	1	0	0	0
	Anfici. Xi.	0	0	1	0	0	4	0	0	0	0	0	0
	Bico. Flo.	0	0	0	0	0	0	5	0	0	0	0	0
	Bico. Xi.	0	0	0	0	0	0	5	0	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	0	0	0	0	3	1	1	0
	Col. Ab. Xi.	0	0	0	0	0	0	0	0	0	5	0	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	0	4	1	0
	Col. Fech. Xi.	0	0	1	0	0	1	0	0	0	0	0	3

Figura 109 – Matriz de confusão da abordagem *k*-NN para a categoria Regularidade estrutural no espaço original.

Fonte: Autor

A acurácia dessa matriz de confusão é 63,33%.

C.3.2 Dimensionalidade reduzida: 2D

A matriz de confusão referente ao primeiro e ao segundo componentes principais da categoria Regularidade estrutural é mostrada na Figura 110. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anfici. Flo.	Anfici. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	1	0	0	0	0	0	4	0	0	0	0	0
	Anf. Prim. Xi.	0	2	0	0	0	0	3	0	0	0	0	0
	Anf. Sec. Flo.	0	0	3	1	0	1	0	0	0	0	0	0
	Anf. Sec. Xi.	0	0	1	4	0	0	0	0	0	0	0	0
	Anfici. Flo.	0	0	0	0	2	0	0	0	1	1	1	0
	Anfici. Xi.	0	0	2	0	0	3	0	0	0	0	0	0
	Bico. Flo.	3	0	0	0	0	0	2	0	0	0	0	0
	Bico. Xi.	0	0	0	1	0	0	0	4	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	0	0	0	0	4	1	0	0
	Col. Ab. Xi.	0	0	0	0	0	0	0	0	0	4	1	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	0	1	4	0
	Col. Fech. Xi.	0	0	1	0	0	1	0	0	0	0	0	3

Figura 110 – Matriz de confusão da abordagem *k*-NN para a categoria Regularidade estrutural após aplicação da ACP com dois componentes principais.

Fonte: Autor

A acurácia dessa matriz de confusão é de 60%.

C.3.3 Dimensionalidade reduzida: 3D

A matriz de confusão referente ao primeiro, segundo e terceiro componentes principais da categoria Regularidade estrutural é mostrada na Figura 111. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anfici. Flo.	Anfici. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	0	0	0	0	0	0	5	0	0	0	0	0
	Anf. Prim. Xi.	0	2	0	0	0	0	0	3	0	0	0	0
	Anf. Sec. Flo.	0	0	3	1	0	0	0	1	0	0	0	0
	Anf. Sec. Xi.	0	0	0	5	0	0	0	0	0	0	0	0
	Anfici. Flo.	1	0	0	0	3	0	0	0	0	0	1	0
	Anfici. Xi.	0	0	2	0	0	3	0	0	0	0	0	0
	Bico. Flo.	0	0	0	0	0	0	5	0	0	0	0	0
	Bico. Xi.	0	0	1	1	0	0	0	3	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	1	0	0	0	2	0	2	0
	Col. Ab. Xi.	0	0	0	0	0	0	0	0	0	4	1	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	1	1	3	0
	Col. Fech. Xi.	0	0	1	0	0	0	0	1	0	0	0	3

Figura 111 – Matriz de confusão da abordagem *k*-NN para a categoria Regularidade estrutural após aplicação da ACP com três componentes principais.

Fonte: Autor

A acurácia dessa matriz de confusão é de 60%.

C.4 Multiescala

C.4.1 Espaço original

O espaço original dessa categoria de atributos é composto por 21 dimensões. A matriz de confusão referente ao espaço original da categoria Multiescala é mostrada na Figura 112. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anfici. Flo.	Anfici. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	3	0	0	0	0	0	0	0	2	0	0	0
	Anf. Prim. Xi.	0	0	0	3	0	0	0	1	0	1	0	0
	Anf. Sec. Flo.	0	0	5	0	0	0	0	0	0	0	0	0
	Anf. Sec. Xi.	0	0	0	5	0	0	0	0	0	0	0	0
	Anfici. Flo.	0	0	0	0	5	0	0	0	0	0	0	0
	Anfici. Xi.	0	0	0	0	0	5	0	0	0	0	0	0
	Bico. Flo.	0	0	0	0	0	0	4	1	0	0	0	0
	Bico. Xi.	0	0	0	0	0	0	0	5	0	0	0	0
	Col. Ab. Flo.	1	0	0	0	1	0	0	0	3	0	0	0
	Col. Ab. Xi.	1	0	0	0	0	0	0	0	0	4	0	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	0	1	4	0
	Col. Fech. Xi.	0	0	1	0	0	0	0	0	0	0	0	4

Figura 112 – Matriz de confusão da abordagem *k*-NN para a categoria Multiescala no espaço original.

Fonte: Autor

A acurácia dessa matriz de confusão é 78,33%.

C.4.2 Dimensionalidade reduzida: 2D

A matriz de confusão referente ao primeiro e o segundo componentes principais da categoria Multiescala é mostrada na Figura 113. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anfici. Flo.	Anfici. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	2	0	0	0	0	0	0	0	3	0	0	0
	Anf. Prim. Xi.	0	1	0	2	0	0	1	1	0	0	0	0
	Anf. Sec. Flo.	0	0	3	0	0	1	0	0	0	1	0	0
	Anf. Sec. Xi.	0	0	0	4	0	0	0	1	0	0	0	0
	Anfici. Flo.	0	0	0	0	5	0	0	0	0	0	0	0
	Anfici. Xi.	0	0	1	0	0	4	0	0	0	0	0	0
	Bico. Flo.	0	0	0	2	0	0	2	1	0	0	0	0
	Bico. Xi.	0	0	0	3	0	0	0	2	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	1	0	0	0	3	1	0	0
	Col. Ab. Xi.	2	0	0	0	0	0	1	0	0	2	0	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	0	0	5	0
	Col. Fech. Xi.	0	0	1	0	0	0	0	0	0	0	0	4

Figura 113 – Matriz de confusão da abordagem *k*-NN para a categoria Multiescala após aplicação da ACP com dois componentes principais.

Fonte: Autor

A acurácia dessa matriz de confusão é de 61,67%.

C.4.3 Dimensionalidade reduzida: 3D

A matriz de confusão referente ao primeiro, segundo e terceiro componentes principais da categoria Multiescala é mostrada na Figura 114. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anfici. Flo.	Anfici. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	3	0	0	0	1	0	0	0	1	0	0	0
	Anf. Prim. Xi.	0	1	0	0	0	0	3	1	0	0	0	0
	Anf. Sec. Flo.	0	0	5	0	0	0	0	0	0	0	0	0
	Anf. Sec. Xi.	0	0	0	3	0	0	0	2	0	0	0	0
	Anfici. Flo.	0	0	0	0	5	0	0	0	0	0	0	0
	Anfici. Xi.	0	0	0	0	0	5	0	0	0	0	0	0
	Bico. Flo.	0	1	0	0	0	0	4	0	0	0	0	0
	Bico. Xi.	0	0	0	1	0	0	0	4	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	1	0	0	0	4	0	0	0
	Col. Ab. Xi.	1	0	0	0	0	0	0	0	1	3	0	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	1	0	4	0
	Col. Fech. Xi.	0	0	1	0	0	0	0	0	0	0	0	4

Figura 114 – Matriz de confusão da abordagem *k*-NN para a categoria Multiescala após aplicação da ACP com três componentes principais.

Fonte: Autor

A acurácia dessa matriz de confusão é de 75%.

C.5 Todos

C.5.1 Espaço original

O espaço original dessa categoria de atributos é composto por 33 dimensões. A matriz de confusão referente ao espaço original da categoria *Todos* os atributos é mostrada na Figura 115. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anfici. Flo.	Anfici. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	5	0	0	0	0	0	0	0	0	0	0	0
	Anf. Prim. Xi.	0	3	0	2	0	0	0	0	0	0	0	0
	Anf. Sec. Flo.	0	0	5	0	0	0	0	0	0	0	0	0
	Anf. Sec. Xi.	0	0	0	5	0	0	0	0	0	0	0	0
	Anfici. Flo.	0	0	0	0	5	0	0	0	0	0	0	0
	Anfici. Xi.	0	0	0	0	0	5	0	0	0	0	0	0
	Bico. Flo.	0	0	0	0	0	0	5	0	0	0	0	0
	Bico. Xi.	0	0	0	0	0	0	0	5	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	1	0	0	0	4	0	0	0
	Col. Ab. Xi.	1	0	0	0	0	0	0	0	0	4	0	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	0	1	4	0
	Col. Fech. Xi.	0	0	1	0	0	0	0	0	0	0	0	4

Figura 115 – Matriz de confusão da abordagem *k*-NN para a categoria *Todos* os atributos no espaço original.

Fonte: Autor

A acurácia dessa matriz de confusão é 90%.

C.5.2 Dimensionalidade reduzida: 2D

A matriz de confusão referente ao primeiro e o segundo componentes principais da categoria *Todos* os atributos é mostrada na Figura 116. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anficri. Flo.	Anficri. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	4	0	0	0	1	0	0	0	0	0	0	0
	Anf. Prim. Xi.	0	1	0	4	0	0	0	0	0	0	0	0
	Anf. Sec. Flo.	0	0	2	0	0	3	0	0	0	0	0	0
	Anf. Sec. Xi.	0	0	0	5	0	0	0	0	0	0	0	0
	Anficri. Flo.	0	0	0	0	5	0	0	0	0	0	0	0
	Anficri. Xi.	0	0	1	0	0	4	0	0	0	0	0	0
	Bico. Flo.	0	0	0	0	0	0	5	0	0	0	0	0
	Bico. Xi.	0	0	0	0	0	0	0	5	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	2	0	0	0	2	1	0	0
	Col. Ab. Xi.	0	0	0	0	1	2	0	0	0	2	0	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	0	1	4	0
	Col. Fech. Xi.	0	0	2	0	0	0	0	0	0	0	0	3

Figura 116 – Matriz de confusão da abordagem *k*-NN para a categoria *Todos* os atributos após aplicação da ACP com dois componentes principais.

Fonte: Autor

A acurácia dessa matriz de confusão é de 70%.

C.5.3 Dimensionalidade reduzida: 3D

A matriz de confusão referente ao primeiro, segundo e terceiro componentes principais da categoria *Todos* os atributos é mostrada na Figura 117. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anficri. Flo.	Anficri. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	3	0	0	0	2	0	0	0	0	0	0	0
	Anf. Prim. Xi.	0	2	0	0	0	0	0	3	0	0	0	0
	Anf. Sec. Flo.	0	0	5	0	0	0	0	0	0	0	0	0
	Anf. Sec. Xi.	0	0	0	5	0	0	0	0	0	0	0	0
	Anficri. Flo.	0	0	0	0	5	0	0	0	0	0	0	0
	Anficri. Xi.	0	0	0	0	0	5	0	0	0	0	0	0
	Bico. Flo.	1	0	0	0	0	0	4	0	0	0	0	0
	Bico. Xi.	0	0	0	0	0	0	0	5	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	2	0	0	0	2	1	0	0
	Col. Ab. Xi.	0	0	1	0	0	0	0	0	1	3	0	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	0	0	5	0
	Col. Fech. Xi.	0	0	2	0	0	0	0	0	0	0	0	3

Figura 117 – Matriz de confusão da abordagem *k*-NN para a categoria *Todos* os atributos após aplicação da ACP com três componentes principais.

Fonte: Autor

A acurácia dessa matriz de confusão é de 78,33%.

C.6 Selecionados

C.6.1 Espaço original

O espaço original dessa categoria de atributos é composto por 14 dimensões. A matriz de confusão referente ao espaço original da categoria *Selecionados* é mostrada na Figura 118. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anfici. Flo.	Anfici. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	5	0	0	0	0	0	0	0	0	0	0	0
	Anf. Prim. Xi.	0	4	0	1	0	0	0	0	0	0	0	0
	Anf. Sec. Flo.	0	0	5	0	0	0	0	0	0	0	0	0
	Anf. Sec. Xi.	0	0	0	5	0	0	0	0	0	0	0	0
	Anfici. Flo.	0	0	0	0	5	0	0	0	0	0	0	0
	Anfici. Xi.	0	0	0	0	0	5	0	0	0	0	0	0
	Bico. Flo.	3	0	0	0	0	0	2	0	0	0	0	0
	Bico. Xi.	0	0	0	0	0	0	0	5	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	0	0	0	0	5	0	0	0
	Col. Ab. Xi.	0	0	0	0	0	0	0	0	0	5	0	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	0	0	5	0
	Col. Fech. Xi.	0	0	1	0	0	0	0	0	0	0	0	4

Figura 118 – Matriz de confusão da abordagem *k*-NN para a categoria *Selecionados* no espaço original.

Fonte: Autor

A acurácia dessa matriz de confusão é 91,67%.

C.6.2 Dimensionalidade reduzida: 2D

A matriz de confusão referente ao primeiro e o segundo componentes principais da categoria *Selecionados* é mostrada na Figura 119. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anfici. Flo.	Anfici. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	4	0	0	0	0	0	1	0	0	0	0	0
	Anf. Prim. Xi.	0	1	0	4	0	0	0	0	0	0	0	0
	Anf. Sec. Flo.	0	0	3	0	0	2	0	0	0	0	0	0
	Anf. Sec. Xi.	0	0	0	5	0	0	0	0	0	0	0	0
	Anfici. Flo.	0	0	0	0	5	0	0	0	0	0	0	0
	Anfici. Xi.	0	0	4	0	0	1	0	0	0	0	0	0
	Bico. Flo.	2	0	0	0	0	0	3	0	0	0	0	0
	Bico. Xi.	0	1	0	0	0	0	0	4	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	1	0	0	0	1	3	0	0
	Col. Ab. Xi.	0	0	0	0	0	0	0	0	0	5	0	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	0	0	5	0
	Col. Fech. Xi.	0	0	1	0	0	1	0	0	0	0	0	3

Figura 119 – Matriz de confusão da abordagem *k*-NN para a categoria *Selecionados* após aplicação da ACP com dois componentes principais.

Fonte: Autor

A acurácia dessa matriz de confusão é de 66,67%.

C.6.3 Dimensionalidade reduzida: 3D

A matriz de confusão referente ao primeiro, segundo e terceiro componentes principais da categoria *Selecionados* é mostrada na Figura 120. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anfici. Flo.	Anfici. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	5	0	0	0	0	0	0	0	0	0	0	0
	Anf. Prim. Xi.	0	1	0	4	0	0	0	0	0	0	0	0
	Anf. Sec. Flo.	0	0	5	0	0	0	0	0	0	0	0	0
	Anf. Sec. Xi.	0	0	0	5	0	0	0	0	0	0	0	0
	Anfici. Flo.	0	0	0	0	5	0	0	0	0	0	0	0
	Anfici. Xi.	0	0	0	0	0	5	0	0	0	0	0	0
	Bico. Flo.	3	0	0	0	0	0	2	0	0	0	0	0
	Bico. Xi.	0	0	0	1	0	0	1	3	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	1	0	0	0	0	1	3	0
	Col. Ab. Xi.	0	0	0	0	0	0	0	0	0	5	0	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	0	0	5	0
	Col. Fech. Xi.	0	0	1	0	0	1	0	0	0	0	0	3

Figura 120 – Matriz de confusão da abordagem *k*-NN para a categoria *Selecionados* após aplicação da ACP com três componentes principais.

Fonte: Autor

A acurácia dessa matriz de confusão é de 73,33%.

APÊNDICE D – PMC: Matrizes de confusão

D.1 Forma

D.1.1 Espaço original

O espaço original dessa categoria de atributos é composto por 10 dimensões. A matriz de confusão referente ao espaço original da categoria Forma é mostrada na Figura 121. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anfieri. Flo.	Anfieri. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	3	0	1	0	0	1	0	0	0	0	0	0
	Anf. Prim. Xi.	0	1	0	4	0	0	0	0	0	0	0	0
	Anf. Sec. Flo.	0	0	5	0	0	0	0	0	0	0	0	0
	Anf. Sec. Xi.	0	0	0	5	0	0	0	0	0	0	0	0
	Anfieri. Flo.	0	0	0	0	5	0	0	0	0	0	0	0
	Anfieri. Xi.	2	0	0	0	0	3	0	0	0	0	0	0
	Bico. Flo.	3	0	2	0	0	0	0	4	0	0	0	0
	Bico. Xi.	0	1	0	0	0	0	0	0	4	0	0	0
	Col. Ab. Flo.	0	0	0	0	0	0	0	0	5	0	0	0
	Col. Ab. Xi.	0	0	0	0	0	0	0	0	0	5	0	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	0	0	5	0
	Col. Fech. Xi.	0	0	0	0	0	1	0	0	0	0	0	4

Figura 121 – Matriz de confusão da abordagem PMC para a categoria Forma no espaço original.

Fonte: Autor

A acurácia dessa matriz de confusão é 75%.

D.1.2 Dimensionalidade reduzida: 2D

A matriz de confusão referente ao primeiro e ao segundo componentes principais da categoria Forma é mostrada na Figura 122. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anficri. Flo.	Anficri. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	1	0	1	0	0	2	1	0	0	0	0	0
	Anf. Prim. Xi.	0	1	0	4	0	0	0	0	0	0	0	0
	Anf. Sec. Flo.	0	0	2	0	0	2	1	0	0	0	0	0
	Anf. Sec. Xi.	0	0	0	5	0	0	0	0	0	0	0	0
	Anficri. Flo.	0	0	0	0	2	1	0	0	0	2	0	0
	Anficri. Xi.	0	0	0	0	0	1	0	2	0	0	0	2
	Bico. Flo.	0	0	1	0	0	0	4	0	0	0	0	0
	Bico. Xi.	0	5	0	0	0	0	0	0	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	0	0	0	1	0	4	0	0
	Col. Ab. Xi.	0	0	0	0	0	0	0	3	0	2	0	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	0	5	0	0
	Col. Fech. Xi.	0	0	0	0	0	0	0	2	0	0	0	3

Figura 122 – Matriz de confusão da abordagem PMC para a categoria Forma após aplicação da ACP com dois componentes principais.

Fonte: Autor

A acurácia dessa matriz de confusão é de 35%.

D.1.3 Dimensionalidade reduzida: 3D

A matriz de confusão referente ao primeiro, segundo e terceiro componentes principais da categoria Forma é mostrada na Figura 123. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anficri. Flo.	Anficri. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	1	0	3	0	1	0	0	0	0	0	0	0
	Anf. Prim. Xi.	0	5	0	0	0	0	0	0	0	0	0	0
	Anf. Sec. Flo.	0	0	2	0	1	1	1	0	0	0	0	0
	Anf. Sec. Xi.	0	5	0	0	0	0	0	0	0	0	0	0
	Anficri. Flo.	0	0	0	0	5	0	0	0	0	0	0	0
	Anficri. Xi.	3	0	0	0	0	2	0	0	0	0	0	0
	Bico. Flo.	0	0	5	0	0	0	0	0	0	0	0	0
	Bico. Xi.	0	4	0	0	0	0	0	1	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	0	0	0	0	0	5	0	0
	Col. Ab. Xi.	0	0	0	0	1	0	0	0	0	4	0	0
	Col. Fech. Flo.	0	0	0	0	4	0	0	0	0	1	0	0
	Col. Fech. Xi.	1	0	0	0	0	0	0	0	0	0	0	4

Figura 123 – Matriz de confusão da abordagem PMC para a categoria Forma após aplicação da ACP com três componentes principais.

Fonte: Autor

A acurácia dessa matriz de confusão é de 40%.

D.2 Densidade

D.2.1 Espaço original

O espaço original dessa categoria de atributos é composto por 16 dimensões. A matriz de confusão referente ao espaço original da categoria Densidade é mostrada na Figura 124. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anferi. Flo.	Anferi. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	4	0	0	0	0	0	0	0	1	0	0	0
	Anf. Prim. Xi.	0	0	0	3	0	0	0	0	0	2	0	0
	Anf. Sec. Flo.	0	0	5	0	0	0	0	0	0	0	0	0
	Anf. Sec. Xi.	0	0	0	1	0	0	0	2	0	2	0	0
	Anferi. Flo.	0	0	0	0	5	0	0	0	0	0	0	0
	Anferi. Xi.	0	0	0	0	0	5	0	0	0	0	0	0
	Bico. Flo.	0	0	0	0	0	1	4	0	0	0	0	0
	Bico. Xi.	0	0	0	0	0	0	0	4	0	0	0	1
	Col. Ab. Flo.	1	0	0	0	0	0	0	0	4	0	0	0
	Col. Ab. Xi.	0	0	0	1	0	0	0	0	0	4	0	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	0	0	5	0
	Col. Fech. Xi.	0	0	0	0	0	0	0	0	0	0	0	5

Figura 124 – Matriz de confusão da abordagem PMC para a categoria Densidade no espaço original.

Fonte: Autor

A acurácia dessa matriz de confusão é 76,67%.

D.2.2 Dimensionalidade reduzida: 2D

A matriz de confusão referente ao primeiro e ao segundo componentes principais da categoria Densidade é mostrada na Figura 125. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anfieri. Flo.	Anfieri. Xi.	Bico. Flo	Bico. Xi	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	0	0	0	0	0	0	0	0	5	0	0	0
	Anf. Prim. Xi.	0	1	0	2	0	0	0	2	0	0	0	0
	Anf. Sec. Flo.	0	0	5	0	0	0	0	0	0	0	0	0
	Anf. Sec. Xi.	0	0	0	1	0	0	0	3	0	1	0	0
	Anfieri. Flo.	0	0	0	0	5	0	0	0	0	0	0	0
	Anfieri. Xi.	0	0	0	0	0	5	0	0	0	0	0	0
	Bico. Flo	0	1	0	0	0	0	0	4	0	0	0	0
	Bico. Xi	0	0	0	0	0	0	0	4	0	1	0	0
	Col. Ab. Flo.	0	0	0	0	1	0	2	0	2	0	0	0
	Col. Ab. Xi.	1	0	0	0	0	0	1	1	0	2	0	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	0	0	5	0
	Col. Fech. Xi.	0	0	0	0	0	0	0	0	0	0	0	5

Figura 125 – Matriz de confusão da abordagem PMC para a categoria Densidade após aplicação da ACP com dois componentes principais.

Fonte: Autor

A acurácia dessa matriz de confusão é de 66,67%.

D.2.3 Dimensionalidade reduzida: 3D

A matriz de confusão referente ao primeiro, segundo e terceiro componentes principais da categoria Densidade é mostrada na Figura 123. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anfieri. Flo.	Anfieri. Xi.	Bico. Flo	Bico. Xi	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	5	0	0	0	0	0	0	0	0	0	0	0
	Anf. Prim. Xi.	0	1	0	3	0	0	1	0	0	0	0	0
	Anf. Sec. Flo.	0	0	5	0	0	0	0	0	0	0	0	0
	Anf. Sec. Xi.	0	0	0	2	0	0	0	0	0	3	0	0
	Anfieri. Flo.	0	0	0	0	5	0	0	0	0	0	0	0
	Anfieri. Xi.	0	0	0	0	0	5	0	0	0	0	0	0
	Bico. Flo	0	0	0	0	0	0	1	4	0	0	0	0
	Bico. Xi	0	0	0	0	0	0	0	5	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	1	0	1	0	3	0	0	0
	Col. Ab. Xi.	1	0	0	1	0	0	0	0	0	3	0	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	0	0	5	0
	Col. Fech. Xi.	0	0	1	0	0	0	0	0	0	0	0	4

Figura 126 – Matriz de confusão da abordagem PMC para a categoria Densidade após aplicação da ACP com três componentes principais.

Fonte: Autor

A acurácia dessa matriz de confusão é de 76,33%.

D.3 Regularidade estrutural

D.3.1 Espaço original

O espaço original dessa categoria de atributos é composto por 7 dimensões. A matriz de confusão referente ao espaço original da categoria Regularidade estrutural é mostrada na Figura 127. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anfici. Flo.	Anfici. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	1	0	0	0	0	0	4	0	0	0	0	0
	Anf. Prim. Xi.	0	0	0	0	0	0	0	5	0	0	0	0
	Anf. Sec. Flo.	0	0	5	0	0	0	0	0	0	0	0	0
	Anf. Sec. Xi.	0	0	2	3	0	0	0	0	0	0	0	0
	Anfici. Flo.	0	0	0	0	3	0	1	0	1	0	0	0
	Anfici. Xi.	0	0	1	0	0	4	0	0	0	0	0	0
	Bico. Flo.	0	0	0	0	0	0	5	0	0	0	0	0
	Bico. Xi.	0	0	1	0	0	0	0	4	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	0	0	0	0	4	1	0	0
	Col. Ab. Xi.	0	0	0	0	0	0	0	0	0	5	0	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	1	4	0	0
	Col. Fech. Xi.	0	0	1	0	0	1	0	0	0	0	0	3

Figura 127 – Matriz de confusão da abordagem PMC para a categoria Regularidade estrutural no espaço original.

Fonte: Autor

A acurácia dessa matriz de confusão é 61,73%.

D.3.2 Dimensionalidade reduzida: 2D

A matriz de confusão referente ao primeiro e ao segundo componentes principais da categoria Regularidade estrutural é mostrada na Figura 128. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anficri. Flo.	Anficri. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	0	0	0	0	0	0	5	0	0	0	0	0
	Anf. Prim. Xi.	0	2	0	1	0	0	0	2	0	0	0	0
	Anf. Sec. Flo.	0	0	3	0	0	1	0	1	0	0	0	0
	Anf. Sec. Xi.	0	0	0	5	0	0	0	0	0	0	0	0
	Anficri. Flo.	0	0	0	0	1	0	0	0	2	1	1	0
	Anficri. Xi.	0	0	2	0	0	3	0	0	0	0	0	0
	Bico. Flo.	1	0	0	0	0	0	4	0	0	0	0	0
	Bico. Xi.	0	0	0	1	0	0	0	4	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	0	0	0	0	4	0	1	0
	Col. Ab. Xi.	0	0	0	0	0	0	0	0	0	4	1	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	0	1	4	0
	Col. Fech. Xi.	0	0	0	0	0	1	0	0	0	0	0	4

Figura 128 – Matriz de confusão da abordagem PMC para a categoria Regularidade estrutural após aplicação da ACP com dois componentes principais.

Fonte: Autor

A acurácia dessa matriz de confusão é de 63,33%.

D.3.3 Dimensionalidade reduzida: 3D

A matriz de confusão referente ao primeiro, segundo e terceiro componentes principais da categoria Regularidade estrutural é mostrada na Figura 129. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anficri. Flo.	Anficri. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	0	0	0	0	1	0	4	0	0	0	0	0
	Anf. Prim. Xi.	0	1	0	1	0	0	0	3	0	0	0	0
	Anf. Sec. Flo.	0	0	4	1	0	0	0	0	0	0	0	0
	Anf. Sec. Xi.	0	0	0	5	0	0	0	0	0	0	0	0
	Anficri. Flo.	0	0	0	0	5	0	0	0	0	0	0	0
	Anficri. Xi.	0	0	0	0	0	5	0	0	0	0	0	0
	Bico. Flo.	0	0	0	0	0	0	5	0	0	0	0	0
	Bico. Xi.	0	0	0	1	0	0	0	4	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	1	0	0	0	2	1	1	0
	Col. Ab. Xi.	0	0	0	0	0	0	0	0	0	4	1	0
	Col. Fech. Flo.	0	0	0	0	1	0	0	0	1	1	2	0
	Col. Fech. Xi.	0	0	0	0	0	0	0	1	0	0	0	4

Figura 129 – Matriz de confusão da abordagem PMC para a categoria Regularidade estrutural após aplicação da ACP com três componentes principais.

Fonte: Autor

A acurácia dessa matriz de confusão é de 68,33%.

D.4 Multiescala

D.4.1 Espaço original

O espaço original dessa categoria de atributos é composto por 21 dimensões. A matriz de confusão referente ao espaço original da categoria Multiescala é mostrada na Figura 130. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anfici. Flo.	Anfici. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	4	0	0	0	0	0	0	0	1	0	0	0
	Anf. Prim. Xi.	0	1	0	3	0	0	0	1	0	0	0	0
	Anf. Sec. Flo.	0	0	5	0	0	0	0	0	0	0	0	0
	Anf. Sec. Xi.	0	0	0	5	0	0	0	0	0	0	0	0
	Anfici. Flo.	0	0	0	0	5	0	0	0	0	0	0	0
	Anfici. Xi.	0	0	0	0	0	5	0	0	0	0	0	0
	Bico. Flo.	0	0	0	0	0	0	5	0	0	0	0	0
	Bico. Xi.	0	0	0	0	0	0	0	5	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	0	0	0	1	4	0	0	0
	Col. Ab. Xi.	1	0	0	0	0	0	0	0	0	4	0	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	0	1	4	0
	Col. Fech. Xi.	0	0	1	0	0	0	0	0	0	0	0	4

Figura 130 – Matriz de confusão da abordagem PMC para a categoria Multiescala no espaço original.

Fonte: Autor

A acurácia dessa matriz de confusão é 85%.

D.4.2 Dimensionalidade reduzida: 2D

A matriz de confusão referente ao primeiro e o segundo componentes principais da categoria Multiescala é mostrada na Figura 131. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anfici. Flo.	Anfici. Xi.	Bico. Flo	Bico. Xi	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	2	0	0	0	0	0	0	0	3	0	0	0
	Anf. Prim. Xi.	0	2	0	2	0	0	0	1	0	0	0	0
	Anf. Sec. Flo.	0	0	3	0	0	1	1	0	0	0	0	0
	Anf. Sec. Xi.	0	0	0	5	0	0	0	0	0	0	0	0
	Anfici. Flo.	0	0	0	0	5	0	0	0	0	0	0	0
	Anfici. Xi.	0	0	0	0	0	4	0	0	0	0	1	0
	Bico. Flo	0	2	0	1	0	0	2	0	0	0	0	0
	Bico. Xi	0	0	0	3	0	0	0	2	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	1	0	0	0	3	1	0	0
	Col. Ab. Xi.	1	0	0	0	0	0	1	0	1	2	0	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	0	0	5	0
	Col. Fech. Xi.	0	0	2	0	0	0	0	0	0	0	0	3

Figura 131 – Matriz de confusão da abordagem PMC para a categoria Multiescala após aplicação da ACP com dois componentes principais.

Fonte: Autor

A acurácia dessa matriz de confusão é de 63,33%.

D.4.3 Dimensionalidade reduzida: 3D

A matriz de confusão referente ao primeiro, segundo e terceiro componentes principais da categoria Multiescala é mostrada na Figura 132. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anfici. Flo.	Anfici. Xi.	Bico. Flo	Bico. Xi	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	3	0	0	0	2	0	0	0	0	0	0	0
	Anf. Prim. Xi.	0	1	0	0	0	0	1	3	0	0	0	0
	Anf. Sec. Flo.	0	0	4	0	0	0	0	0	0	1	0	0
	Anf. Sec. Xi.	0	0	0	5	0	0	0	0	0	0	0	0
	Anfici. Flo.	0	0	0	0	5	0	0	0	0	0	0	0
	Anfici. Xi.	0	0	0	0	0	5	0	0	0	0	0	0
	Bico. Flo	0	0	0	0	0	0	5	0	0	0	0	0
	Bico. Xi	0	0	0	0	0	0	0	5	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	1	0	0	0	4	0	0	0
	Col. Ab. Xi.	0	0	0	0	0	0	0	0	1	4	0	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	0	0	5	0
	Col. Fech. Xi.	0	0	2	0	0	0	0	0	0	0	0	3

Figura 132 – Matriz de confusão da abordagem PMC para a categoria Multiescala após aplicação da ACP com três componentes principais.

Fonte: Autor

A acurácia dessa matriz de confusão é de 81,67%.

D.5 Todos

D.5.1 Espaço original

O espaço original dessa categoria de atributos é composto por 33 dimensões. A matriz de confusão referente ao espaço original da categoria *Todos* os atributos é mostrada na Figura 133. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anfici. Flo.	Anfici. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	5	0	0	0	0	0	0	0	0	0	0	0
	Anf. Prim. Xi.	0	4	0	1	0	0	0	0	0	0	0	0
	Anf. Sec. Flo.	0	0	5	0	0	0	0	0	0	0	0	0
	Anf. Sec. Xi.	0	0	0	5	0	0	0	0	0	0	0	0
	Anfici. Flo.	0	0	0	0	5	0	0	0	0	0	0	0
	Anfici. Xi.	0	0	0	0	0	5	0	0	0	0	0	0
	Bico. Flo.	0	0	0	0	0	0	5	0	0	0	0	0
	Bico. Xi.	0	0	0	0	0	0	0	5	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	1	0	0	0	4	0	0	0
	Col. Ab. Xi.	1	0	0	0	0	0	0	0	0	4	0	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	0	1	4	0
	Col. Fech. Xi.	0	0	1	0	0	0	0	0	0	0	0	4

Figura 133 – Matriz de confusão da abordagem PMC para a categoria *Todos* os atributos no espaço original.

Fonte: Autor

A acurácia dessa matriz de confusão é 91,67%.

D.5.2 Dimensionalidade reduzida: 2D

A matriz de confusão referente ao primeiro e o segundo componentes principais da categoria *Todos* os atributos é mostrada na Figura 134. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anfici. Flo.	Anfici. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	5	0	0	0	0	0	0	0	0	0	0	0
	Anf. Prim. Xi.	0	2	0	3	0	0	0	0	0	0	0	0
	Anf. Sec. Flo.	0	0	1	0	0	4	0	0	0	0	0	0
	Anf. Sec. Xi.	0	0	0	5	0	0	0	0	0	0	0	0
	Anfici. Flo.	0	0	0	0	5	0	0	0	0	0	0	0
	Anfici. Xi.	0	0	1	0	0	4	0	0	0	0	0	0
	Bico. Flo.	0	0	0	0	0	0	5	0	0	0	0	0
	Bico. Xi.	0	0	0	1	0	0	0	4	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	1	0	0	0	3	1	0	0
	Col. Ab. Xi.	0	0	0	0	0	0	0	0	1	4	0	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	0	0	5	0
	Col. Fech. Xi.	0	0	0	0	0	0	0	0	0	0	0	5

Figura 134 – Matriz de confusão da abordagem PMC para a categoria *Todos* os atributos após aplicação da ACP com dois componentes principais.

Fonte: Autor

A acurácia dessa matriz de confusão é de 80%.

D.5.3 Dimensionalidade reduzida: 3D

A matriz de confusão referente ao primeiro, segundo e terceiro componentes principais da categoria *Todos* os atributos é mostrada na Figura 135. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anfici. Flo.	Anfici. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	5	0	0	0	0	0	0	0	0	0	0	0
	Anf. Prim. Xi.	0	4	0	0	0	0	0	1	0	0	0	0
	Anf. Sec. Flo.	0	0	4	0	0	0	0	0	0	1	0	0
	Anf. Sec. Xi.	0	0	0	5	0	0	0	0	0	0	0	0
	Anfici. Flo.	0	0	0	0	5	0	0	0	0	0	0	0
	Anfici. Xi.	0	0	0	0	0	5	0	0	0	0	0	0
	Bico. Flo.	1	0	0	0	0	0	4	0	0	0	0	0
	Bico. Xi.	0	0	0	0	0	0	0	5	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	2	0	0	0	2	1	0	0
	Col. Ab. Xi.	1	0	0	0	0	0	0	0	0	4	0	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	0	0	5	0
	Col. Fech. Xi.	0	0	1	0	0	0	0	0	0	0	0	4

Figura 135 – Matriz de confusão da abordagem PMC para a categoria *Todos* os atributos após aplicação da ACP com três componentes principais.

Fonte: Autor

A acurácia dessa matriz de confusão é de 86,67%.

D.6 Selecionados

D.6.1 Espaço original

O espaço original dessa categoria de atributos é composto por 14 dimensões. A matriz de confusão referente ao espaço original da categoria *Selecionados* é mostrada na Figura 136. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anfici. Flo.	Anfici. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	5	0	0	0	0	0	0	0	0	0	0	0
	Anf. Prim. Xi.	0	4	0	1	0	0	0	0	0	0	0	0
	Anf. Sec. Flo.	0	0	5	0	0	0	0	0	0	0	0	0
	Anf. Sec. Xi.	0	0	0	5	0	0	0	0	0	0	0	0
	Anfici. Flo.	0	0	0	0	5	0	0	0	0	0	0	0
	Anfici. Xi.	0	0	0	0	0	5	0	0	0	0	0	0
	Bico. Flo.	1	0	0	0	0	0	4	0	0	0	0	0
	Bico. Xi.	0	1	0	0	0	0	0	4	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	0	0	0	0	5	0	0	0
	Col. Ab. Xi.	0	0	0	0	0	0	0	0	1	4	0	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	0	0	5	0
	Col. Fech. Xi.	0	0	1	0	0	0	0	0	0	0	0	4

Figura 136 – Matriz de confusão da abordagem PMC para a categoria *Selecionados* no espaço original.

Fonte: Autor

A acurácia dessa matriz de confusão é 91,67%.

D.6.2 Dimensionalidade reduzida: 2D

A matriz de confusão referente ao primeiro e o segundo componentes principais da categoria *Selecionados* é mostrada na Figura 137. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anfici. Flo.	Anfici. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	5	0	0	0	0	0	0	0	0	0	0	0
	Anf. Prim. Xi.	0	2	0	3	0	0	0	0	0	0	0	0
	Anf. Sec. Flo.	0	0	4	0	0	1	0	0	0	0	0	0
	Anf. Sec. Xi.	0	0	0	5	0	0	0	0	0	0	0	0
	Anfici. Flo.	0	0	0	0	5	0	0	0	0	0	0	0
	Anfici. Xi.	0	0	1	0	0	4	0	0	0	0	0	0
	Bico. Flo.	3	0	0	0	0	0	2	0	0	0	0	0
	Bico. Xi.	0	0	0	0	0	0	0	5	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	1	0	0	0	4	0	0	0
	Col. Ab. Xi.	0	0	0	0	0	0	0	0	1	4	0	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	0	0	5	0
	Col. Fech. Xi.	0	0	0	0	0	2	0	0	0	0	0	3

Figura 137 – Matriz de confusão da abordagem PMC para a categoria *Selecionados* após aplicação da ACP com dois componentes principais.

Fonte: Autor

A acurácia dessa matriz de confusão é de 80%.

D.6.3 Dimensionalidade reduzida: 3D

A matriz de confusão referente ao primeiro, segundo e terceiro componentes principais da categoria *Selecionados* é mostrada na Figura 138. Os acertos (em verde) estão localizados na diagonal principal dessa matriz. Em vermelho, estão destacados os erros classificatórios.

		Classificação atual											
		Anf. Prim. Flo.	Anf. Prim. Xi.	Anf. Sec. Flo.	Anf. Sec. Xi.	Anfici. Flo.	Anfici. Xi.	Bico. Flo.	Bico. Xi.	Col. Ab. Flo.	Col. Ab. Xi.	Col. Fech. Flo.	Col. Fech. Xi.
Classificação prevista	Anf. Prim. Flo.	4	0	0	0	0	0	1	0	0	0	0	0
	Anf. Prim. Xi.	0	5	0	0	0	0	0	0	0	0	0	0
	Anf. Sec. Flo.	0	0	5	0	0	0	0	0	0	0	0	0
	Anf. Sec. Xi.	0	0	0	5	0	0	0	0	0	0	0	0
	Anfici. Flo.	0	0	0	0	5	0	0	0	0	0	0	0
	Anfici. Xi.	0	0	0	0	0	5	0	0	0	0	0	0
	Bico. Flo.	0	0	0	0	0	0	2	0	0	3	0	0
	Bico. Xi.	0	5	0	0	0	0	0	0	0	0	0	0
	Col. Ab. Flo.	0	0	0	0	1	0	0	0	0	0	4	0
	Col. Ab. Xi.	0	0	0	0	0	0	0	0	0	4	1	0
	Col. Fech. Flo.	0	0	0	0	0	0	0	0	0	0	5	0
	Col. Fech. Xi.	0	0	0	0	0	2	0	0	0	0	0	3

Figura 138 – Matriz de confusão da abordagem PMC para a categoria *Selecionados* após aplicação da ACP com três componentes principais.

Fonte: Autor

A acurácia dessa matriz de confusão é de 71,67%.