

BRUNO GEIKE DE ANDRADE

**DISTINÇÃO ENTRE CARVÕES PROVENIENTES DE MADEIRA DE
EUCALYPTUS sp E DE ESPÉCIES NATIVAS POR ANÁLISE DE IMAGENS
DIGITAIS**

**Dissertação apresentada à
Universidade Federal de Viçosa,
como parte das exigências do
Programa de Pós-Graduação em
Ciência Florestal, para obtenção do
título de *Magister Scientiae*.**

VIÇOSA
MINAS GERAIS – BRASIL
2013

**Ficha catalográfica preparada pela Seção de Catalogação e
Classificação da Biblioteca Central da UFV**

T

A553d
2013

Andrade, Bruno Geike de, 1987-
Distinção entre carvões provenientes de madeira de
Eucalyptus sp e de espécies nativas por análise de imagens
digitais / Bruno Geike de Andrade. – Viçosa, MG, 2013.
xi, 66f. : il. ; 29cm.

Inclui anexo.

Orientador: Benedito Rocha Vital

Dissertação (mestrado) - Universidade Federal de Viçosa.

Referências bibliográficas: f. 60-65

1. Carvão vegetal. 2. Processamento de imagens - Técnicas
digitais. 3. Teoria bayesiana de decisão estatística.

I. Universidade Federal de Viçosa. Departamento de
Engenharia Florestal. Programa de Pós-Graduação em Ciência
Florestal. II. Título.

CDO adapt. CDD 634.98675

BRUNO GEIKE DE ANDRADE

**DISTINÇÃO ENTRE CARVÕES PROVENIENTES DE MADEIRA DE
EUCALYPTUS sp E DE ESPÉCIES NATIVAS POR ANÁLISE DE IMAGENS
DIGITAIS**

**Dissertação apresentada à
Universidade Federal de Viçosa,
como parte das exigências do
Programa de Pós-Graduação em
Ciência Florestal, para obtenção do
título de *Magister Scientiae*.**

APROVADA: 22 de fevereiro de 2013.

Carla Priscilla Távora Cabral

Ana Márcia Macedo Ladeira Carvalho

Angélica de Cássia Oliveira Carneiro
(Coorientadora)

Francisco de Assis de Carvalho Pinto
(Coorientador)

Benedito Rocha Vital
(Orientador)

*“Não são títulos que honram homens,
são homens que honram títulos”.*

Niccolò di Bernardo Machiavelli

AGRADECIMENTOS

À Deus pela vida.

À minha família, em especial aos meus pais, Fernando e Evônia, pelo apoio incondicional, que mesmo a distância me confortava a cada dificuldade e por serem, ambos, exemplos de determinação e perseverança. Aos meus irmãos Marília e Eduardo, aos meus tios Beto e Neide, e a todos os demais parentes, com os quais divido esta conquista.

À Universidade Federal de Viçosa, especialmente ao Departamento de Engenharia Florestal (DEF), por terem contribuído para a minha formação acadêmica.

Às escolas Zacheu Moreira da Fraga e Casemiro Costa, pela fundamental participação na gênese dessa trajetória e ao colégio Jesus Cristo Rei, por ter me preparado para a vida acadêmica.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES pelo apoio financeiro, através da concessão da bolsa de estudos, à FAPEMIG, à SECTES e à Embrapa Florestas.

Ao professor Benedito Rocha Vital, pela orientação, amizade e a atenção dedicada durante toda a pesquisa.

À professora Angélica de Cássia Oliveira Carneiro pela amizade, apoio e coorientação durante toda a pesquisa.

Ao professor Francisco de Assis de Carvalho Pinto, pela coorientação, atenção e valiosas contribuições ao longo da pesquisa.

Aos funcionários da secretária de pós-graduação em Ciência Florestal do DEF, Ritinha e Alexandre, pela atenção a cada dúvida em todas as etapas do processo de pós graduação.

À Vanessa, pela atenção, compreensão e paciência durante os dias e noites de construção da dissertação, seja na leitura e comentário ou pelo estímulo de continuidade dos estudos.

Aos companheiros do LAPEM e LPM e a todos que, de alguma forma, contribuíram para a realização deste trabalho.

BIOGRAFIA

Bruno Geike de Andrade, filho de Fernando Macedo de Andrade e Evônia Geike de Andrade, nasceu no dia 03 de maio de 1987, na cidade de Cachoeiro de Itapemirim – Espírito Santo.

Cresceu em um distrito de Cachoeiro de Itapemirim chamado Vargem Grande de Soturno, um pequeno vilarejo.

Estudou na escola estadual rural Zacheu Moreira da Fraga da primeira à quinta série, com exceção da segunda, quando, por motivo de greve, fora matriculado na escola municipal rural Casemiro Costa. Em seguida, da sexta série até o fim do ensino médio, estudou no colégio Jesus Cristo Rei, na cidade.

Desde a infância até a juventude, costumava ir passar os fins de semana, com os pais e os irmãos, no pequeno sítio da família. Lá, plantios de eucalipto e café lhe apresentavam as ciências agrárias, ao passo que os projetos do pai, entre os quais se destacam um carro elétrico e duas pequenas hidrelétricas, lhe despertavam o interesse pela engenharia.

Em 2005, ingressou no curso de Física da Universidade Federal de Viçosa. Foi reprovado em Cálculo e, embora tenha se tornado tutor de Física, desiludiu-se com o curso.

Fez novo vestibular, e, em 2007, ingressou no curso de Engenharia Florestal, também na UFV, concluindo a graduação em 2011.

Em Março de 2011, ingressou no programa de Pós Graduação, em nível de mestrado, em Ciência Florestal da Universidade Federal de Viçosa, submetendo-se à defesa da dissertação em Fevereiro de 2013.

CONTEÚDO

LISTA DE TABELAS	vii
LISTA DE FIGURAS	viii
RESUMO	x
ABSTRACT	xi
1. Introdução	1
2. Objetivos	3
3. Revisão de literatura.....	4
3.1. O carvão vegetal.....	4
3.1.1. A anatomia da madeira	5
3.1.2. A anatomia do carvão	6
3.1.3. A identificação do carvão vegetal	7
3.2. A análise de imagens digitais	12
3.2.1. Características baseadas na textura.....	13
3.2.2. A matriz de coocorrência de níveis de cinzas (MCNC)	15
3.3. Classificação de padrões	17
3.3.1. O classificador bayesiano para classes gaussianas	20
3.3.2. Redução da dimensionalidade dos dados	21
3.3.3. Avaliação de classificadores.....	23
4. Materiais e Métodos	26
4.1. Caracterização e preparo do material carbonizado	26
4.2. Aquisição de imagens digitais.....	27
4.3. Extração de atributos	27

4.4. Desenvolvimento dos classificadores.....	30
4.5. Avaliação dos classificadores.....	32
5. Resultados e discussão	35
5.1. Aquisição de imagens.....	35
5.2. Seleção de atributos.....	49
5.3. Avaliação dos classificadores.....	52
6. Conclusões	58
7. Referencias Bibliográficas	60

LISTA DE TABELAS

	Pg.
1 Principais abordagens utilizadas em reconhecimento de padrões.	18
2 Espécies e híbridos estudados.	26
3 Parâmetros das imagens-índice geradas no programa computacional <i>Matlab</i> ®.	28
4 Parâmetros das MCNCs geradas de acordo com Haralick et al. (1973).	28
5 Atributos utilizados no desenvolvimento dos classificadores.	30
6 Número de atributos selecionados para cada classificador.	50
7 Percentual das características básicas das MCNCs selecionadas para a construção dos modelos dos classificadores.	50
8 Percentual dos ângulos θ (graus) selecionados para a construção dos modelos dos classificadores.	51
9 Percentual dos valores de d selecionados para a construção dos modelos dos classificadores.	51
10 Exatidão global e acurácias por classe de carvão (<i>eucalipto e nativa</i>)	53
11 Estimativas dos coeficientes <i>Kappa</i> por classificador.	55
12 Características utilizadas no modelo do classificador 11.	57

LISTA DE FIGURAS

	Pg.
1 As quatro orientações propostas por Haralick et al. (1973).	16
2 Representação matricial de uma imagem com 4 níveis de cinzas.	16
3 MCNCs calculadas a partir da imagem contida na Figura 2.	16
4 Três classes de padrões distintas em um espaço bidimensional.	19
5 Modelo de reconhecimento de padrões.	20
6 Imagem da seção transversal do carvão de Angico-vermelho.	35
7 Imagem da seção transversal do carvão de Brauninha.	36
8 Imagem da seção transversal do carvão de Cedrinho.	36
9 Imagem da seção transversal do carvão de Mama de Porca.	37
10 Imagem da seção transversal do carvão de Pimenteira.	37
11 Imagem da seção transversal do carvão de Ruão.	38
12 Imagem da seção transversal do carvão de Casca Doce.	38
13 Imagem da seção transversal do carvão de Goiabeira.	39
14 Imagem da seção transversal do carvão de Pau Fumo.	39
15 Imagem da seção transversal do carvão de Quaresminha.	40
16 Imagem da seção transversal do carvão de Cedro.	40
17 Imagem da seção transversal do carvão de Casuarina.	41
18 Imagem da seção transversal do carvão de <i>E. camaldulensis</i> .	42
19 Imagem da seção transversal do carvão de <i>E. camaldulensis</i> x <i>E. grandis</i> .	42
20 Imagem da seção transversal do carvão de <i>E. camaldulensis</i> x <i>E. urophylla</i> .	43
21 Imagem da seção transversal do carvão de <i>E. grandis</i> x <i>E. urophylla</i> .	43
22 Imagem da seção transversal do carvão de <i>E. urophylla</i> (GG 150).	44
23 Imagem da seção transversal do carvão de <i>E. urophylla</i> (GG 100).	44
24 Imagens do carvão de angico vermelho e seus histogramas	46
25 Imagens do carvão de angico vermelho em diferentes resoluções de imagens (colunas) e profundidade de <i>pixels</i> (linhas).	47
26 Imagens do carvão de <i>Eucalyptus camaldulensis</i> em diferentes resoluções de imagens (colunas) e profundidade de <i>pixels</i> (linhas).	48
27 Porcentagem de atributos selecionados para diferentes resoluções e valores de <i>d</i> (<i>pixels</i>).	52
28 Participação percentual das espécies da classe <i>nativa</i> no erro dos	54

classificadores.

- 29 Participação percentual das espécies/híbridos da classe *eucalipto* no erro dos classificadores. 54

RESUMO

ANDRADE, Bruno Geike de, M. Sc., Universidade Federal de Viçosa, Fevereiro de 2013. **Discriminação entre carvões provenientes de madeira de *Eucalyptus* sp e de espécies nativas por análise de imagens digitais.** Orientador: Benedito Rocha Vital. Coorientadores: Angélica de Cássia Oliveira Carneiro e Francisco de Assis de Carvalho Pinto.

A distinção entre carvões produzidos com madeira de eucalipto daqueles produzidos com madeira de espécies nativas é importante para a comercialização, controle e fiscalização da produção de carvão vegetal no Brasil. Este trabalho propõe a aplicação de técnicas de análise de imagens como uma nova ferramenta para auxiliar na separação do carvão de eucalipto em relação àqueles provenientes de outras espécies. Para a realização do trabalho, 18 espécies lenhosas, 6 pertencentes ao gênero *Eucalyptus* e 12 espécies nativas da Zona da Mata Mineira, foram estudadas. Para cada espécie foram geradas 50 imagens da seção transversal do lenho carbonizado, totalizando 300 imagens para o gênero *Eucalyptus* e 600 para as espécies nativas. A partir destas imagens foram produzidas 24 imagens-índice, combinação de três diferentes resoluções (número de *pixels*) e oito diferentes profundidades de imagem (número de *bits* por *pixel*). Em seguida, 24 classificadores bayesianos para classes gaussianas foram desenvolvidos utilizando-se características de texturas extraídas das imagens-índice por meio da técnica da matriz de coocorrência de níveis de cinza. A avaliação da classificação foi realizada segundo o método da validação cruzada deixando um vetor de características de fora da fase de treinamento. Em seguida, por meio da matriz de erros, foram obtidos os índices de Exatidão Global e *Kappa* que foram utilizados para avaliar a significância dos classificadores por meio de um Teste Z. Concluiu-se que os classificadores apresentaram desempenho melhor que uma classificação ao acaso, ao nível de 5% de significância. Verificou-se que os 24 classificadores desenvolvidos acertaram, em média, 98,45% das imagens testadas. A utilização da análise de imagens digitais, portanto, permitiu distinguir satisfatoriamente o carvão produzido com madeira de eucalipto daqueles produzidos com madeiras de espécies nativas.

ABSTRACT

ANDRADE, Bruno Geike de, M. Sc., Universidade Federal de Viçosa, February, 2013. **Discrimination between charcoal from *Eucalyptus* sp and native species by digital image analysis.** Adviser: Benedito Rocha Vital. Co-Advisers: Angélica de Cássia Oliveira Carneiro and Francisco de Assis de Carvalho Pinto.

The distinction between charcoal produced with *Eucalyptus* wood from those produced with native wood is important for the commercialization, control and supervision of charcoal production in Brazil. This work proposes the application of image analysis techniques as a new tool to assist in the separation of *Eucalyptus*'s charcoal compared to those from other species. To conduct the study, we used 18 woody species, six belonging to the genus *Eucalyptus* and 12 native species in the Zona da Mata Mineira. For each species were generated 50 images of the cross section of charred wood, totaling 300 images for *Eucalyptus* and 600 for the native species. From these images were produced 24 images-index, combination of three different resolution (number of pixels) and eight different depths of image (number of bits per pixel). Then, 24 Bayesian classifiers for Gaussian classes were developed using texture features extracted from the images-index by the gray level co-occurrence matrix technique. The classification evaluation was performed according to the cross-validation method leaving one sample outside the training phase. Then, through the matrix of errors, were obtained Overall Accuracy and Kappa indexes that were used to evaluating the classifiers significance through a test Z. All classifiers performed better than a random classification, the 5% level of significance. The 24 classifiers developed hit, on average, 98.45% of the images tested. The use of digital image analysis, therefore, allowed to distinguish satisfactorily the charcoal produced from *Eucalyptus*'s wood from those produced with mixed native species.

1. Introdução

O Brasil é o maior produtor e consumidor de carvão vegetal do mundo, sendo a maior parte desta produção destinada ao mercado interno, principalmente para os setores de ferro-gusa e aço e, em menor escala, para o setor de ferro-ligas, o comércio e o consumo residencial (ABRAF, 2012).

Em 2011 foram produzidos 7,9 milhões de toneladas de carvão vegetal em nível nacional, um crescimento de 7,5% em relação ao ano anterior. Do total produzido, 74% foram consumidos pelo setor de ferro-gusa e aço (MINISTÉRIO DE MINAS E ENERGIA, 2012). Segundo estimativas do Sindicato da Indústria Mineral do Estado de Minas Gerais (2012), a participação do carvão vegetal na produção do ferro-gusa, em 2010, foi de 44% para o Estado de Minas Gerais e de 22% para todo o país.

A área de plantios de eucalipto no Brasil totalizou 4.873.952 ha em 2011, dos quais 28,5% situavam no Estado de Minas Gerais. Estima-se que 10% de toda a madeira de reflorestamento produzida em nível nacional, constituída predominantemente pelo gênero *Eucalyptus*, tenha sido utilizada para conversão em carvão vegetal (ABRAF, 2012).

Entretanto, apesar de ser o maior produtor de carvão vegetal do mundo, o Brasil ainda não é capaz de atender toda a sua demanda por meio da produção de carvão a partir de madeira de plantios florestais, tornando economicamente atrativa a exploração ilegal de florestas nativas para a produção clandestina.

Dados da AMS (2011) mostram que cerca de um terço de todo o carvão vegetal produzido no Brasil, em 2010, foram provenientes de madeira de florestas nativas, composta por madeira oriunda de planos de manejo e, majoritariamente, de extração ilegal. Contudo, conforme ABRAF (2011), nesse mesmo ano, a produção guseira a carvão vegetal apresentou um elevado índice de inatividade, fruto da crise econômica de 2008, de forma que foram utilizadas apenas um terço da capacidade instalada e somente 56% dos fornos. De forma geral, portanto, verifica-se que com recuperação do setor siderúrgico, e conseqüente aumento da demanda por carvão vegetal, a exploração de florestas nativas poderá se acentuar.

Para coibir a produção clandestina de carvão vegetal, existem vários dispositivos legais, dos quais se destaca, entre os mais recentes, o Documento de Origem Florestal (DOF). Instituído pelo Ministério do Meio Ambiente, através da Portaria nº253 de

2006, e normatizado pelo IBAMA, através da Instrução Normativa nº112 de 2006, o DOF trata-se de uma licença obrigatória para o controle do transporte e armazenamento de produtos e subprodutos florestais de origem nativa, incluído o carvão vegetal. Esse dispositivo ainda apresenta algumas vulnerabilidades, devendo o agente fiscalizador, portanto, verificar se o material transportado corresponde à quantidade e espécie descritas no documento, podendo haver autuação de toda a carga caso haja desacordo em relação à quantidade ou espécie.

No caso do carvão vegetal, a identificação da espécie é uma técnica complexa, baseada na anatomia da madeira que foi preservada no carvão e, geralmente, empregada apenas na paleontologia. Contudo, a diferenciação entre carvões oriundos de madeira de florestas equiâneas e aqueles oriundos de florestas nativas pode ser realizada mais facilmente e pode ocorrer em campo, durante a fiscalização.

Para isso é necessário uma inovação nas ferramentas de reconhecimento de espécies pelos agentes fiscalizadores, que deverão possuir rigor científico sem perder praticidade, uma vez que o reconhecimento é realizado em campo.

Nestas circunstâncias, a análise de imagens digitais surge como uma alternativa atrativa para auxiliar na distinção entre carvões provenientes de madeira de *Eucalyptus* sp e de espécies nativas, pois permite o desenvolvimento de sistemas automatizados capazes de tomar decisões com base em informações, ou características, visuais.

Dentre estas características, destaca-se a textura, tanto pelo potencial para a distinção de diferentes objetos e padrões, quanto pela relativa facilidade de representação por meio da técnica da Matriz de Coocorrência de Níveis de Cinzas. Além disso, visto que a textura do carvão resulta da organização e interação entre os diferentes elementos anatômicos, têm-se na análise da textura uma forma indireta de analisar a anatomia.

2. Objetivos

Este projeto teve como objetivo desenvolver procedimentos capazes de distinguir carvões produzidos com madeira de espécies nativas daqueles produzidos com madeira de eucalipto aplicando-se técnicas de análise de imagens da madeira carbonizada.

Especificamente buscou-se:

- a) Elaborar um sistema para extração de atributos de textura em imagens digitais com base nas matrizes de coocorrência de níveis de cinza (MCNC).
- b) Desenvolver classificadores, por meio da análise discriminante de atributos de texturas, a fim de separar o carvão em duas classes: *nativa* e *eucalipto*.
- c) Determinar a significância de cada classificador.
- d) Comparar, dois a dois, a significância entre os classificadores.

3. Revisão de literatura

3.1. O carvão vegetal

A determinação taxonômica do carvão vegetal é feito por meio da observação da anatomia da madeira preservada no carvão e, geralmente, têm o intuito de obter informações de ordem paleontológica e arqueológica (SCHEEL et al., 1996a; SCHEEL et al., 1996b). Contudo, a análise do lenho carbonizado tornou-se também uma ferramenta necessária para a identificação forense utilizada na fiscalização da produção, comércio e transporte do carvão vegetal, permitindo distinguir carvões produzidos ilegalmente com madeira de espécies nativas daqueles produzidos com madeira oriunda de plantios florestais, que são majoritariamente compostos por espécies do gênero *Eucalyptus*.

Apesar desta necessidade, ainda são recentes os estudos anatômicos abordando o propósito forense da análise do carvão vegetal. Trabalhos já foram feitos com o objetivo de caracterizar a anatomia da madeira carbonizada de espécies nativas, tais como aqueles realizados por Gonçalves (2010), Albuquerque (2012) e Muniz et al. (2012). Por outro lado, a caracterização anatômica de diversas espécies do gênero *Eucalyptus* já vem sendo realizada há mais tempo por diversos autores, tais como Tomazello Filho (1985), Oliveira & Hellmeister (1998), Silva (2002), Alzate (2004), Evangelista (2007), Evangelista (2010) e Brisola & Demarco (2011).

Contudo, a literatura apresenta apenas subsídios e carece de trabalhos englobando estes conhecimentos e se propondo a analisar, especificamente, as particularidades e potencialidades da discriminação entre o gênero *Eucalyptus* e espécies nativas a partir do lenho carbonizado. A ausência deste tipo de trabalho implica na falta de informações importantes e práticas, como por exemplo, as características que melhor separam estes dois grupos. Em função desta lacuna, o identificador forense fica condicionado a basear-se em informações dispersas pela literatura, em sua experiência própria, ou em coleções de referência, caso disponha de uma. Todavia, a fiscalização do carvão vegetal, principalmente quando realizada em campo, pode ser favorecida por metodologias de identificação rápidas, práticas e de baixo custo.

3.1.1. A anatomia da madeira

A anatomia da madeira é o estudo das funções, organizações e estruturas das células que constituem o lenho dos vegetais lenhosos. Por convenção, observações anatômicas adotam três planos básicos de observação: transversal, longitudinal tangencial e longitudinal radial (BURGUER; RICHTER, 1991).

A identificação anatômica de madeiras apresenta duas principais abordagens, a macroscópica e a microscópica. Para Wheeler & Baas (1998), as análises macroscópicas incluem as características anatômicas visíveis a olho “nu” ou com auxílio de uma lupa, tais como, arranjo e agrupamento de vasos, arranjo e abundância de parênquima axial e largura dos raios. Este tipo de análise geralmente é empregado para determinar rapidamente se uma dada amostra foi corretamente identificada ou a que família pertence.

“Na identificação microscópica são observadas as características dos tecidos e das células constituintes do lenho que não são distintas sem o uso de microscópio” (ZENID; CECCANTINI, 2007). A Associação Internacional de Anatomistas da Madeira (IAWA 1989) desenvolveu um conjunto de características microscópicas com o propósito de tornar-se uma lista concisa para objetivos de identificação de espécies madeireiras, reduzindo a ambiguidade das descrições anatômicas. Ao todo foram propostas 168 características anatômicas, qualitativas e quantitativas, tais como, anéis de crescimento, porosidade, arranjo e agrupamento de vasos e diâmetro médio de vasos, dentre outras. Em geral, algumas dessas características microscópicas, principalmente aquelas associadas aos vasos, podem ser vistas com auxílio de uma lupa.

A escolha da abordagem a ser utilizada depende da habilidade e treinamento do identificador (ZENID; CECCANTINI, 2007). Uma dada característica microscópica geralmente não é melhor ou mais confiável que uma característica macroscópica. Contudo, quanto maior o número de características utilizadas para a identificação, maior será a chance de se obter uma combinação capaz de permitir um diagnóstico conclusivo, motivo pelo qual a análise microscópica se destaca por apresentar uma gama maior de características (WHEELER; BAAS, 1998).

3.1.2. A anatomia do carvão

“A carbonização é um processo de pirólise lenta, ou tratamento térmico, cujo principal objetivo é eliminar a maior parte do oxigênio pela ação do calor e, com isso, possibilitar a concentração do carbono na estrutura residual, que é o carvão vegetal” (VALE; GENTIL, 2008).

A estrutura anatômica do lenho altera-se em decorrência do processo de carbonização. Estas mudanças dependem principalmente das características da madeira, mas também são influenciadas pelo tempo e temperatura de carbonização, pelo teor de umidade inicial da madeira e pela dimensão das peças carbonizadas (PRIOR; ALVIN, 1986).

Prior & Alvin (1983) estudaram o efeito da carbonização em duas espécies cujas amostras, anidras, foram carbonizadas às temperaturas de 300, 400, 500, 600, 700 e 800°C por 60 minutos. Foi constatado que as células dos raios, vistas da seção tangencial, perderam seu formato ligeiramente alongado no sentido vertical e tornaram-se arredondadas quando em temperaturas de carbonização de 400°C. Os vasos, apesar de possuírem paredes relativamente finas, apresentaram pouca mudança estrutural com exceção dos colapsos ocorridos nas paredes entre vasos. Segundo os autores, as mudanças foram marcadamente diferentes entre as duas espécies, mas incapazes de invalidar o uso de características qualitativas para propósitos de identificação do carvão.

As mudanças anatômicas em *Quercus variabilis* em função de diferentes temperaturas finais de carbonização foram estudadas por Kwon et al. (2009). As amostras, secas ao ar, foram carbonizadas aumentando-se a temperatura em 5°C por minuto até atingir as temperaturas finais de 250, 300, 350, 400, 450 e 500°C. Os vasos apresentaram contração tangencial cerca de 7,5 vezes superior à contração radial para as temperaturas finais de 350-500°C. Até 300°C, estes elementos foram pouco modificados morfológicamente, mostrando satisfatória similaridade para com o estado *in natura*.

Gonçalves (2010) estudou as mudanças anatômicas do lenho carbonizado de dez espécies do cerrado paulista. As amostras, envoltas em papel alumínio, foram carbonizadas a 400°C por 40 minutos. Os resultados mostraram que a carbonização provocou apenas pequenas alterações em relação à estrutura original e que, dentre as características quantitativas, apenas a redução do diâmetro tangencial dos vasos foi

observada em todas as espécies. Segundo a autora, as variações morfométricas entre madeira e carvão não são um fator limitante para a identificação das espécies e permitem a aproximação entre a anatomia do carvão e a anatomia da madeira.

Carvões provenientes da madeira de 75 espécies da floresta Amazônica, carbonizadas a 400°C por 40 minutos, foi estudado por Albuquerque (2012), que obteve os seguintes resultados: I) mesmo quando escasso, o parênquima axial foi observado no carvão mantendo as mesmas características de interesse anatômico; II) o parênquima radial também foi observado no carvão e, em alguns casos, tornou-se até mesmo mais evidente em relação ao lenho; e III) os vasos apresentaram as mudanças mais nítidas, com distorções na seção transversal, tornando-se alongados e com as paredes irregulares e onduladas. Apesar da contração volumétrica observada, a carbonização também manteve satisfatoriamente a estrutura original da madeira, permitindo a análise anatômica do carvão.

Muñiz et al. (2012) estudaram a anatomia do carvão de 8 espécies florestais por meio de um estereomicroscópio de luz refletida utilizando-se aumentos de 10 e 25x, e verificaram que o aspecto anatômico geral de distribuição e tipos de células é bem evidente, sendo possível a identificação de material carbonizado com base na estrutura da madeira.

É importante ressaltar que, em trabalhos de pesquisa, as amostras utilizadas na carbonização são geralmente secas ao ar ou em estufa. No entanto, a madeira utilizada em fornos para produção de carvão possui, em geral, maior teor de umidade, algumas vezes, inclusive, acima do ponto de saturação das fibras. Deste modo, na prática, o carvão produzido comercialmente tende a ter uma estrutura anatômica mais danificada em função da eliminação da água da madeira, dificultando sua identificação. Dessa forma, conforme recomendado por Muñiz et al. (2012), devem ser estudadas as influências de diferentes taxas de aquecimento e umidade do material nas alterações estruturais do carvão.

3.1.3. A identificação do carvão vegetal

Dada a satisfatória preservação da anatomia da madeira após a carbonização, a identificação forense do carvão vegetal pode ser realizada através da análise de sua

anatomia. Tal como é feito para a madeira, a identificação pode ser alcançada por meio das abordagens macro ou microscópica.

Contudo, muito pouco tem sido proposto sobre a análise macroscópica, apesar de permitir agilidade e praticidade à identificação realizada em campo, onde geralmente ocorre a fiscalização. Por outro lado, muito tem sido proposto sobre a ótica microscópica que, com maior custo e limitada logística, é capaz de fazer a identificação do carvão até o nível de espécie com satisfatório grau de confiabilidade, embora isto nem sempre seja necessário para a finalidade a que a identificação forense do carvão vegetal se destina.

Para a análise anatômica, o preparo do carvão difere daquele convencionalmente utilizado para a madeira. Conforme Figueiral (1999), a friabilidade do carvão praticamente impossibilita qualquer tentativa de seccioná-lo em lâminas finas para análises microscópicas, a menos que sejam utilizados procedimentos para incorporação de resinas ou substâncias similares. Por este motivo, os estudos anatômicos do carvão são, tradicionalmente, feitos através de faces recém quebradas nos sentidos transversais, radiais e tangenciais, que são observadas em microscópios de dissecação, de luz refletida, ou de varredura. Na prática, segundo Hoadley (1933), superfícies planas e contínuas são mais facilmente obtidas trabalhando-se na seção transversal.

Para a Antracologia, ciência que visa obter dados paleoecológicos e paleoetnobotânicos a partir do carvão vegetal (GONÇALVES, 2010), a identificação da espécie carbonizada é realizada através da análise de fragmentos, quebrados manualmente segundo os planos da madeira, em microscópio óptico de luz refletida com campo claro e campo escuro. Para a determinação taxonômica dos carvões é feita uma comparação de sua estrutura anatômica com amostras de coleções de referência ou com descrições ou fotografias de obras da literatura (SCHEEL-YBERT, 2004).

Recentemente, alguns estudos têm sido realizados com o intuito de caracterizar a anatomia do lenho carbonizado de espécies nativas, visando produzir material referencial, fornecendo subsídios para a identificação forense do carvão. Destacam-se os trabalhos de Gonçalves (2010), Albuquerque (2012) e Muniz et al. (2012) que caracterizaram a madeira e o carvão de 10 espécies do cerrado paulista, 75 espécies da floresta Amazônica e 10 espécies nativas, respectivamente.

De forma geral, a identificação do carvão beneficia-se dos conhecimentos já adquiridos na área de identificação de madeiras. A respeito desta, uma importante

consideração feita por Wheeles & Baas (1998) trata do nível de dificuldade da identificação, que aumenta proporcionalmente ao número de diferentes e possíveis resultados para uma amostra desconhecida. Este nível apresenta, portanto, implicações logísticas, afetando a facilidade da identificação como também a escolha da abordagem a ser utilizada.

Assim, para a finalidade a que se destina a identificação forense do carvão vegetal, esta consideração pode contribuir para identificar possíveis estratégias. Considerando que grande parte do carvão produzido com madeira de florestas plantadas pertence ao gênero *Eucalyptus*, o nível de dificuldade da identificação forense pode ser substancialmente reduzido se apenas dois resultados forem admitidos ao avaliar uma amostra desconhecida, ou seja, pertencente ou não ao gênero *Eucalyptus*. Admitindo isto, a complexidade da identificação poderá então ser reduzida, em alguns casos, para o nível de família, facilitando a identificação. Dessa forma, a abordagem macroscópica pode ser utilizada como um procedimento inicial, identificando boa parte das amostras desconhecidas e resguardando a abordagem microscópica para os casos mais complexos.

Adicionalmente, a distinção entre o gênero *Eucalyptus* dos demais gêneros, como ponto de partida para a identificação forense do carvão, pode contribuir com outras potencialidades. Conforme Alfonso (1987), Oliveira & Hellmeister (1998) e Tomazello Filho (1985), o gênero *Eucalyptus* apresenta uma constituição anatômica muito homogênea entre as espécies, sob o nível morfológico, fator que dificulta a separação, com base apenas na composição e arranjo estrutural dos constituintes do lenho. Por outro lado, esta semelhança pode auxiliar na distinção entre este gênero para com os demais. Além disso, conforme Santos (2010), Castro (2011) e Pereira (2012), as principais espécies utilizadas para a produção de carvão vegetal são *Eucalyptus urophylla*, *E. grandis*, e os híbridos *E. urophylla* x *grandis*, *E. urophylla* x *camaldulensis* e *E. grandis* x *camaldulensis*.

Dentre as características anatômicas descritivas predominantes no gênero *Eucalyptus*, destacam-se a porosidade difusa, vasos em arranjo diagonal, vasos solitários, parênquima axial paratraqueal, parênquima radial apresentando raios unisseriados e a presença de tilose. A predominância dessas características já foi observada em diversos trabalhos.

A porosidade difusa foi verificada em *E. grandis* por Tomazello Filho (1985), Silva (2002), Alzate (2004) e Brisola & Demarco (2011); em *E. camaldulensis* por Evangelista (2007); em *E. urophylla* por Evangelista (2007) e Brisola & Demarco (2011); em *E. grandis* x *urophylla* por Alzate (2004) e Brisola & Demarco (2011); e em híbrido de *E. urophylla* e *E. grandis* x *camaldulensis* por Evangelista (2010).

O arranjo diagonal dos vasos foi observado em *E. grandis* por Tomazello Filho (1985), Silva (2002), Alzate (2004) e Brisola & Demarco (2011); em *E. camaldulensis* por Evangelista (2007); em *E. urophylla* por Evangelista (2007) e Brisola & Demarco (2011); em *E. grandis* x *urophylla* por Alzate (2004) e Brisola & Demarco (2011); e em híbrido de *E. urophylla* e *E. grandis* x *camaldulensis* por Evangelista (2010).

O agrupamento predominantemente solitário dos vasos foi verificado em *E. grandis* por Tomazello Filho (1985), Silva (2002), Oliveira & Hellmeister (1998), Alzate (2004) e Brisola & Demarco (2011); em *E. camaldulensis* por Evangelista (2007); em *E. urophylla* por Oliveira & Hellmeister (1998), Evangelista (2007) e Brisola & Demarco (2011); em *E. grandis* x *urophylla* por Alzate (2004) e Brisola & Demarco (2011); em híbrido de *E. urophylla* por Evangelista (2010); em *E. grandis* x *camaldulensis* por Evangelista (2010).

Os padrões observados para o parênquima axial de *E. grandis* foram paratraqueal escasso (ALZATE, 2004); vasicêntrico (BRISOLA; DEMARCO, 2011) e vasicêntrico e escasso, com alguns trechos confluentes nas cadeias diagonais e raros difusos (TOMAZELLO FILHO, 1985; SILVA, 2002). Para *E. urophylla* foi verificado parênquima paratraqueal vasicêntrico (BRISOLA; DEMARCO 2011) e predomínio de vasicêntrico com ocorrência de confluyente e unilateral (EVANGELISTA, 2007). Para *E. camaldulensis* foi observado parênquima paratraqueal vasicêntrico e, raramente, unilateral (EVANGELISTA, 2007). Para *Eucalyptus grandis* x *urophylla* foi verificado parênquima paratraqueal vasicêntrico (ALZATE, 2004) e vasicêntrico e aliforme (BRISOLA; DEMARCO, 2011) e para *E. grandis* x *camaldulensis* foi observado parênquima paratraqueal vasicêntrico, confluyente, aliforme e unilateral (EVANGELISTA, 2010).

O parênquima radial formado por raios predominantemente unisseriados foi verificado em *E. grandis* por Tomazello Filho (1985), Silva (2002, Alzate (2004) e Brisola & Demarco (2011); em *E. camaldulensis* por Evangelista (2007); em *E. urophylla* por Evangelista (2007) e Brisola & Demarco (2011); em *E. grandis* x

urophylla por Alzate (2004) e Brisola & Demarco (2011); e em *E. grandis* x *camaldulensis* por Evangelista (2010).

A presença de tilose foi observada em *E. grandis* por Oliveira & Hellmeister (1998) e Silva (2002); em *E. camaldulensis* por Evangelista (2007) e em *E. urophylla* por Oliveira & Hellmeister (1998) e Evangelista (2007).

Ressalta-se que todas as características anteriormente mencionadas são facilmente identificadas a partir da seção transversal. Adicionalmente, ressalta-se que conforme citado por Wheeler & Baas (1998), estas características podem ser observadas por meio da abordagem macroscópica.

Através das caracterizações anatômicas feitas por Gonçalves (2010), Albuquerque (2012) e Muniz et al. (2012), importantes considerações sobre a distinção anatômicas entre espécies de nativas e o gênero *Eucalyptus* podem ser feitas, embora comparações deste tipo devam utilizadas com cautela, em vista da subjetividade inerentes aos trabalhos anatômicos. Assim, considerando todas as 95 espécies nativas que ao todo foram estudadas por estes autores, nenhuma apresentou a combinação de características anatômicas comum ao gênero *Eucalyptus*, embora algumas espécies tenham apresentado estrutura anatômica similar, sendo distintas por uma ou duas diferenças. Destas, citam-se *Couepia grandiflora* (GONÇALVES, 2010), distinta devido à ausência de arranjo diagonal dos vasos e ao parênquima axial em faixas e *Ocotea porosa* (MUNIZ et al., 2012), distinta devido à ausência de arranjo diagonal dos vasos.

Por outro lado, a distinção entre o carvão produzido entre madeira de espécies nativas e madeira do gênero *Eucalyptus* também pode ser realizada com base na morfologia da peça carbonizada. Assim, no carvão de eucalipto, observam-se peças mais retas, sendo as bifurcações, decorrentes da inserção de galhos, pouco frequente. A casca das peças mais finas é sempre lisa. Geralmente, peças de carvão tortuosas, com muitas inserções de galhos, e casca rugosa nas peças de pequeno diâmetro são características de carvões produzidos com madeira nativa. No plano transversal, a observação da medula, também, pode oferecer indícios importantes. Os eucaliptos cultivados no Brasil possuem a medula pequena e por isso é mais difícil de ser observada após a carbonização. Por outro lado, algumas espécies nativas apresentam medula de maior dimensão podendo, inclusive, ser oca (VITAL; CARNEIRO, 2007).

Por fim, é importante que todo esse conhecimento abordado seja adaptado para condições de campo, pois, Conforme Hermanson & Wiedenhoeft (2011), quanto mais próximo da floresta, em tempo e espaço, a identificação for realizada, melhor protegida a floresta estará e mais rapidamente as medidas cabíveis serão tomadas. Entretanto, é possível que a necessidade de identificação rápida e precisa esteja superando a capacidade de formar pessoal capacitado para essa tarefa. Assim, o crescente aumento do poder computacional e da capacidade de aquisição de imagens, aliados à diminuição dos custos associados, fazem com que a automação do processo de identificação seja uma solução atrativa e plausível.

3.2. A análise de imagens digitais

Para Gonzalez et al. (2004), a visão computacional é um ramo da Inteligência Artificial que tem como principal objetivo permitir que os computadores imitem a visão humana e sejam capazes de aprender e tomar decisões com base em informações visuais. Segundo Pedrini & Schwartz (2008), em um sistema de visão computacional grande importância deve ser dada à análise de imagens digitais cujo principal objetivo é obter uma descrição que contenha as informações necessárias para distinguir entre diferentes objetos de interesse, de forma confiável e requerendo o mínimo de intervenção humana.

Uma imagem digital é uma matriz bidimensional. Se $L_x = \{1, 2, \dots, N_x\}$ e $L_y = \{1, 2, \dots, N_y\}$ são, respectivamente, seus domínios X e Y, então $N_x \times N_y$ é a resolução de *pixels* e a imagem digital I é uma função que atribui um valor de tom de cinza $G \in \{1, 2, \dots, N_g\}$ para cada *pixel* (HARALICK et al., 1973). O número de diferentes níveis de cinzas de uma imagem I é normalmente uma potência de 2. Assim, “se $G = 2^b$, G é o número de níveis de cinza da imagem e b é chamada de profundidade de *pixel*, que corresponde ao número de bits necessários para armazenar cada pixel da imagem digital” (PEDRINI; SCHWARTZ, 2008).

Para Pedrini & Schwartz (2008), o estudo das tarefas da visão computacional é comumente dividido em dois níveis de abstração: o processamento de imagens (baixo nível) e a análise de imagens (alto nível).

“O processamento de imagens digitais pode ser caracterizado como um conjunto de técnicas voltadas para melhorar o aspecto visual da imagem original, gerando produtos que possam ser submetidos a tratamentos posteriores” (REIS et al., 2001).

Por outro lado, a análise de imagens pode ser definida como uma técnica em que uma imagem é utilizada como entrada e têm-se, como resultado final, informações obtidas desta imagem (GONZALEZ et al., 2004) baseadas na forma, textura, níveis de cinzas ou nas cores dos objetos presentes na imagem (PEDRINI; SCHWARTZ, 2008).

Um sistema de análise de imagens é constituído de diversos processos, tais como: aquisição (ou digitalização), pré-processamento, segmentação, pós-processamento (ou representação), extração de atributos (ou descrição) e classificação (ou reconhecimento). A aquisição trata da captura da imagem. O pré-processamento refere-se a ações de baixo-nível, como processamento inicial de dados brutos para calibração radiométrica da imagem, a correção de distorções geométricas, a remoção de ruído e a suavização de determinadas propriedades da imagem. A segmentação consiste na identificação de áreas de interesse na imagem. O pós-processamento é a etapa em que são corrigidos os defeitos da segmentação, já que, muitas vezes, o resultado da segmentação não é adequado. A extração de atributos é a fase em que são obtidas características, tais como textura e cor, que possam ser utilizadas na discriminação entre classes de objetos. Tais propriedades são, em geral, descritas por atributos numéricos que formam um vetor de características. Por fim, na etapa de classificação é atribuído um identificador aos objetos da imagem, ou a determinadas imagens, baseado nas características providas pelos seus descritores (REIS et al., 2001; PEDRINI; SCHWARTZ, 2008).

3.2.1. Características baseadas na textura

Uma das tarefas mais complexas presente na análise de imagens é definir o conjunto de características que será utilizado para descrever de maneira efetiva a imagem de interesse ou de suas regiões. A textura encontra-se entre as características utilizadas pelo sistema visual humano, que a emprega com facilidade no reconhecimento e descrição de objetos. “As características de textura contêm informações sobre a distribuição espacial e a variação de luminosidade, além de

descrever o arranjo estrutural das superfícies e relações entre regiões vizinhas” (PEDRINI; SCHWARTZ, 2008).

Segundo Haralick et al. (1973), a textura é uma propriedade inata de praticamente todas as superfícies, tais como a grã da madeira, a trama de um tecido, o padrão de culturas em um campo, etc. Ela contém informações importantes sobre o arranjo estrutural de superfícies e sua relação com ambiente circundante.

A textura da imagem pode ser avaliada qualitativamente como tendo um ou mais das seguintes propriedades: finura, aspereza, suavidade, granulação, linearidade, aleatoriedade, regularidade, intensidade, entre outras características (HARALICK, 1979).

O maior propósito em se extrair atributos de textura é mapear as diferenças nas estruturas espaciais, representadas pela imagem, em diferenças de tons de cinza. A partir de então, métodos de segmentação podem analisar espacialmente estes atributos com o objetivo de determinar áreas homogêneas em uma imagem. A detecção de bordas e análise de regiões são consideradas as principais abordagens para extração e análise de textura, que aplicam-se também com objetivo de reconhecer imagens. (TAMURA et al., 1978; REED; HANS DU BUF, 1993).

Quando decomponível, uma dada textura apresenta duas dimensões básicas em que pode ser descrita. A primeira dimensão é a descrição das primitivas, elementos que repetem e arranjam-se espacialmente de acordo com uma regra de organização e geram o padrão da textura. A segunda dimensão é a descrição da dependência espacial ou interação entre estas primitivas. Assim, uma textura pode ser descrita pelo número e tipos de primitivas que a compõem, bem como pela organização espacial destas primitivas. Esta organização pode ser aleatória, pode apresentar uma dependência de emparelhamento entre primitivas vizinhas, ou pode ter uma dependência entre n primitivas. Esta dependência, por sua vez, pode ser estrutural, probabilística ou funcional (HARALICK, 1979).

Os principais métodos de extração de características de textura são divididos em três abordagens: espectrais, estruturais e estatísticas. A primeira é baseada no Processamento de Sinais. Os descritores são extraídos da representação obtida após a aplicação de transformações da imagem de entrada. A abordagem estrutural, ou geométrica, busca analisar a textura com base nas suas primitivas. Assim, após a identificação das primitivas, as características a serem extraídas tanto podem ser as

medidas das primitivas quanto as regras que descrevem a disposição e o relacionamento espacial entre elas (PEDRINI; SCHWARTZ, 2008).

“As abordagens estatísticas não tentam compreender explicitamente a estrutura hierárquica da textura. Elas representam a textura indiretamente por propriedades não determinísticas que definem as distribuições e relacionamentos entre os níveis de cinza de uma imagem.” Na abordagem estatística, os métodos baseados em características de segunda ordem, que levam em consideração a relação entre *pixels*, têm demonstrado alto poder discriminatório em grande diversidade de imagens (NASCIMENTO, 2003). Dentre estes, destaca-se o método baseado em Matrizes de Coocorrência de Níveis de Cinza (MCNC) proposto por Haralick et al. (1973) e discutido em detalhes no próximo tópico.

3.2.2. A matriz de coocorrência de níveis de cinzas (MCNC)

A MCNC é baseada na ocorrência repetida da configuração de alguns níveis de cinza na textura. “Dado um relacionamento espacial entre os *pixels* de uma imagem, os elementos de uma MCNC descrevem a frequência com que ocorrem as transições de nível de cinza entre pares de *pixels*”. Considerando-se que estes pares podem apresentar um relacionamento espacial com diferentes orientações e distâncias, várias MCNC podem ser obtidas para a análise de uma mesma textura (PEDRINI; SCHWARTZ, 2008).

Haralick et al. (1973) definiram dois parâmetros para caracterizar o relacionamento espacial entre *pixels*: a distância d , em *pixels*, e a orientação θ , em graus, sendo que para este último foram estabelecidos quatro ângulos: 0° , 45° , 90° e 135° . O número de linhas $i=(1,2,\dots,I)$ e colunas $j=(1,2,\dots,J)$ de uma MCNC serão iguais ao total de níveis de cinza presentes na imagem.

Para a construção de uma MCNC, são considerados dois *pixels* a cada vez, um chamado de *pixel* de referência e o outro de *pixel* vizinho. Considerando, na Figura 1, o *pixel* central como referência, os *pixels* vizinhos seriam determinados de acordo com a distância e a orientação estabelecida para o relacionamento espacial. Por exemplo, para $d=1$ e $\theta=45^\circ$, os *pixels* vizinhos seriam 3 e 6.

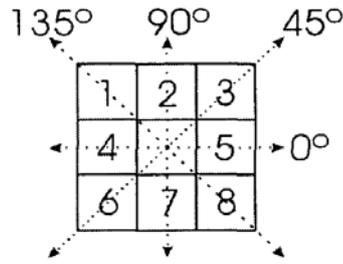


Figura 1 – As quatro orientações propostas por Haralick et al. (1973).

A MCNC, em sua forma geral, pode ser especificada como uma matriz de frequências relativas. Assim, um elemento $P(i,j,d,\theta)$ indica a frequência com que um *pixel* com nível de cinza i ocorre a uma distância de d *pixels* e orientação θ° de um *pixel* com nível de cinza j .

Um exemplo da MCNC, extraído do trabalho de Haralick et al. (1973), é mostrado a seguir. Levando em consideração a representação matricial de uma imagem (Figura 2) com resolução 4x4 *pixels* e 4 diferentes níveis de cinza [0,1,2,3], são mostradas quatro MCNCs obtidas a partir desta imagem na Figura 3.

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

Figura 2 – Representação matricial de uma imagem com 4 níveis de cinzas.

$$\begin{aligned}
 P(i,j;1,0^\circ) &= \begin{pmatrix} 4 & 2 & 1 & 0 \\ 2 & 4 & 0 & 0 \\ 1 & 0 & 6 & 1 \\ 0 & 0 & 1 & 2 \end{pmatrix} & P(i,j;1,45^\circ) &= \begin{pmatrix} 4 & 1 & 0 & 0 \\ 1 & 2 & 2 & 0 \\ 0 & 2 & 4 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \\
 P(i,j;1,90^\circ) &= \begin{pmatrix} 6 & 0 & 2 & 0 \\ 0 & 4 & 2 & 0 \\ 2 & 2 & 2 & 2 \\ 0 & 0 & 2 & 0 \end{pmatrix} & P(i,j;1,135^\circ) &= \begin{pmatrix} 2 & 1 & 3 & 0 \\ 1 & 2 & 1 & 0 \\ 3 & 1 & 0 & 2 \\ 0 & 0 & 2 & 0 \end{pmatrix}
 \end{aligned}$$

Figura 3 – MCNCs calculadas a partir da imagem contida na Figura 3.

Apesar de a MCNC ter sido definida como sendo composta pelo número de ocorrências entre dois pixels espacialmente relacionados, as características de textura são, na verdade, obtidas a partir da representação normalizada desta matriz. Assim, a

normalização da MCNC se faz dividindo cada um de seus elementos pelo total de ocorrências nela contida (PEDRINI; SCHWARTZ, 2008).

Para descrever os atributos das texturas, Haralick et al. (1973) propôs 28 medidas estatísticas para serem calculadas a partir da MCNC. Posteriormente, Baraldi e Parmiggiani (1995) afirmaram que apenas seis destas apresentam maior relevância: *segundo momento angular (SMA), entropia, contraste, variância, correlação e homogeneidade*.

Khalid et al. (2008) tentaram classificar espécies lenhosas por meio da análise de textura da madeira *in natura*. Utilizando uma câmera monocromática foram produzidas imagens com resoluções 768 (H) x 576 (V) *pixels* a partir da seção transversal de blocos de madeira. Ao todo, 20 espécies tropicais foram classificadas com uma exatidão global de 95%. O classificador foi construído em um modelo de redes neurais e utilizou-se de 20 características de textura obtidos de acordo com a técnica das MCNCs. Aproximadamente 90 imagens por espécie, totalizando 1753, foram utilizadas na fase de treinamento e 196 imagens foram utilizadas para avaliar o desempenho do classificador.

3.3. Classificação de padrões

A classificação de padrões é um campo de estudo desenvolvido amplamente na década de 1960. Era, e ainda é, um assunto interdisciplinar, cobrindo pesquisas nas áreas de engenharia, estatística, inteligência artificial, ciência computacional, psicologia, filosofia, entre outros. Dentre suas diversas aplicações, cita-se o reconhecimento de caracteres em imagens digitalizadas (WEBB, 2002).

O reconhecimento/classificação de uma padrão pode ser feito de duas formas: 1) supervisionada, na qual o padrão de entrada será identificado como pertencente a uma classe pré-estabelecida; ou 2) não supervisionada, na qual o padrão será alocado a uma classe desconhecida, tendo como base a similaridade entre os padrões analisados (WATANABE, 1985).

Para Jain et al. (2000), um sistema de reconhecimento de padrões envolve três aspectos essenciais: 1) a aquisição de informação; 2) a representação da informação e 3) a decisão. Espera-se que um sistema bem elaborado culmine no uso de padrões compactos e de uma estratégia de decisão simplificada. O aprendizado através de

exemplos (amostras de treinamento) é importante e buscado para a maioria dos sistemas de reconhecimentos de padrões. Dentre as diferentes abordagens que tratam do assunto, quatro podem ser destacadas: 1) casamento de padrões (*template matching*); 2) casamento estrutural; 3) redes neurais e 4) classificação estatística. Uma breve descrição entre estes é apresentada na Tabela 1.

Tabela 1 - Principais abordagens utilizadas em reconhecimento de padrões. FONTE: JAIN et al. (2000).

Abordagem	Representação	Função de reconhecimento	Critério de avaliação
Casamento de padrões	Amostras, pixels, curvas	Correlação, medidas de distância	Erro de classificação
Estrutural ou sintática	Primitivas	Regras, gramática	Erro de aceitação
Redes Neurais	Amostras, pixels, características	Função de rede	Erro médio quadrático
Estatística	Características	Funções discriminantes	Erro de classificação

O casamento de padrões é uma das abordagens mais antigas e simples. De maneira geral, dispõe-se de um modelo que represente o padrão a ser reconhecido. Para a classificação, é realizado o casamento do modelo com a imagem que apresenta o padrão a ser reconhecido, levando em conta variações como translação, rotação e mudanças na escala. “A medida de similaridade, normalmente relacionada com a ideia de correlação, deve ser otimizada com base no conjunto de treinamento” (NASCIMENTO, 2003).

A abordagem estrutural utilizada no reconhecimento de padrões parte da perspectiva de que um padrão é formado por sub-padrões, ou primitivas. Assim, de acordo com Jain et al. (2000), no reconhecimento estrutural de padrões, uma analogia formal é feita entre a estrutura de padrões e a sintaxe da linguagem. Os padrões são vistos como sentenças pertencentes a um idioma, primitivas são vistas como o alfabeto deste idioma e as sentenças são criadas de acordo com uma regra: a gramática. Deste modo, um grande grupo de padrões pode ser descrito por um pequeno grupo de primitivas e regras gramaticais.

“As redes neurais fundamentam-se em estudos sobre a estrutura do cérebro humano, constituindo um sistema de circuitos que tentam simular sua forma inteligente

de processar informação, inclusive seu comportamento, aprendendo, errando e fazendo novas descobertas” (NASCIMENTO, 2003). A principal característica das redes neurais é sua capacidade de aprender complexas relações não lineares de entrada e saída de informações e o potencial de adaptar-se a um determinado banco de dados (JAIN et al., 2000).

Na abordagem estatística, cada amostra é representada de acordo com suas características ou atributos. Estas características são expressas por meio de medidas compondo um espaço de características que apresenta dimensionalidade igual ao número de medidas utilizadas para descrição dos padrões (PEDRINI; SCHWARTZ, 2008).

Nos modelos de classificação de padrões baseados na abordagem estatística, um padrão é denotado como vetor p -dimensional $x = (x_1, \dots, x_p)$ de medidas, cujas componentes x_i são medidas de características, ou atributos, de uma amostra. Estas características são cuidadosamente escolhidas pelo investigador, que as considera relevantes para a classificação. Na discriminação entre diferentes amostras, assume-se que existam C classes, denotadas w_1, \dots, w_C , e associadas a cada padrão x está uma variável categórica z que denota a classe pertencente. Assim, se $z = i$, então o padrão pertence à classe w_i , $i \in \{1, \dots, C\}$ (WEBB, 2002).

O objetivo é fazer com que esse espaço seja gerado de forma que os vetores de padrões de diferentes classes sejam dispostos de maneira a simplificar o processo de classificação. Logo, é desejável que amostras de um mesmo padrão apresentem vetores que ocupem regiões compactas e disjuntas das regiões ocupadas por vetores de outros padrões (Figura 4) (NASCIMENTO, 2003).

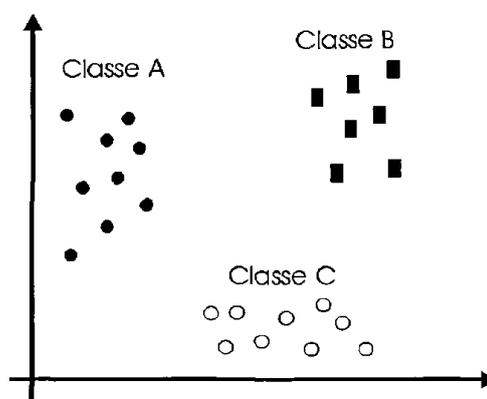


Figura 4 – Três classes de padrões distintas em um espaço bidimensional.
FONTE: NASCIMENTO (2003).

Conceitos de teoria da decisão estatística são utilizados para estabelecer limites de decisão entre as classes de padrões. O sistema de reconhecimento é operado em dois modos: treinamento (aprendizagem) e classificação (avaliação) (Figura 5). No modo de treinamento, é realizada a extração de características e, se necessário, sua seleção. No modo de classificação, o classificador, agora treinado, atribui um padrão a uma das classes de padrões com base nos descritores utilizados (JAIN et al., 2000).

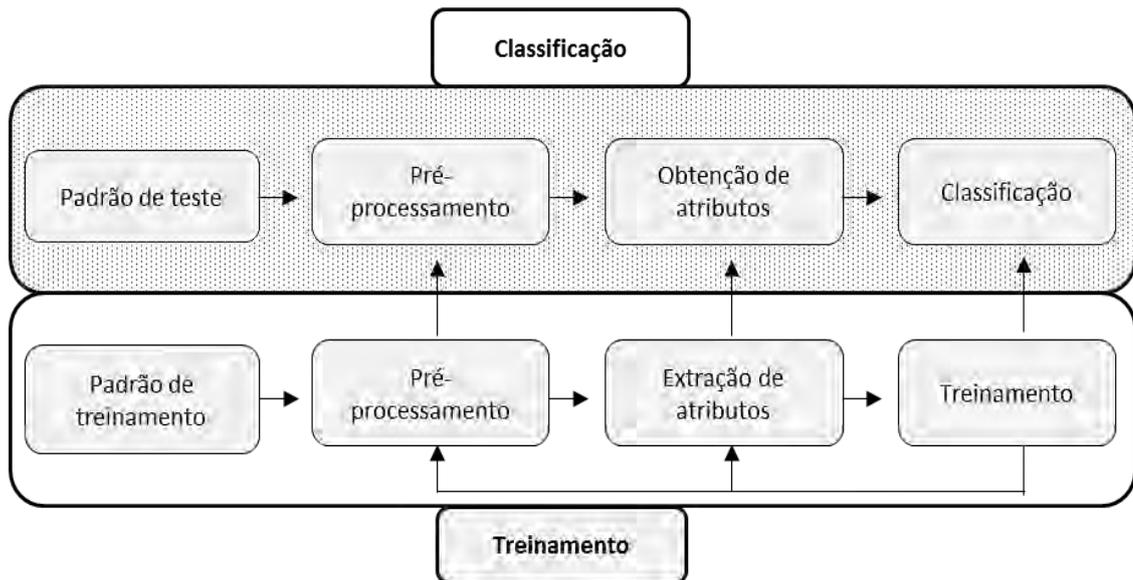


Figura 5 – Modelo de reconhecimento de padrões. FONTE: adaptado de JAIN et al. (2000).

3.3.1. O classificador bayesiano para classes gaussianas

Na classificação bayesiana um dado padrão será associado a uma dentre C classes w_1, w_2, \dots, w_C baseado em um vetor de p características $x = (x_1, x_2, \dots, x_p)$. Assumindo que as características apresentam uma função de densidade condicionada ao padrão da classe, um vetor padrão x , pertencente a classe w_i , é visto como uma observação aleatoriamente sorteada a partir da função de probabilidade classe-condicionada $p(x|w_i)$ (JAIN et al., 2000).

Nesse modelo, $p(x|w_i)$ denota a probabilidade de ocorrer x dado que o padrão pertence à classe w_i . Assim, a classificação de uma amostra segue a regra de decisão mostrada na Equação 1. Logo, x será atribuído à classe w_i caso $p(w_i|x)$, denominada

probabilidade *a posteriori*, seja maior que qualquer outra $p(w_j|x)$ (PEDRINI; SCHWARTZ, 2008).

$$p(w_i|x) > p(w_j|x) \quad j = 1, \dots, C; i \neq j \quad (1)$$

Segundo Webb (2002), essa regra de decisão particiona o espaço de medidas em C regiões $\Omega_1, \dots, \Omega_C$ de tal forma que se $x \in \Omega_i$, então x pertence a classe w_i . A probabilidade *a posteriori* $p(w_i|x)$ pode ser reescrita em termos das probabilidades *a priori*, $p(x)$ e $p(w_i)$, e da função de probabilidade classe-condicionada $p(x|w_i)$ usando o Teorema de Bayes:

$$p(w_i|x) = \frac{p(x|w_i)p(w_i)}{p(x)} \quad (2)$$

Assumindo que a função de probabilidade classe-condicional $p(x|w_i)$ é uma distribuição Gaussiana, Webb (2002) reescreve a probabilidade *a posteriori* $p(w_i|x)$ da seguinte forma:

$$F_i = \log(p(w_i)) - \frac{1}{2} m_i^t S_w^{-1} m_i + x^t S_w^{-1} m_i \quad (3)$$

Na Equação 3, m_i é o vetor de médias das características da classe w_i e S_w é a matriz de covariância comum das classes. Assim, a regra de decisão é atribuir x a classe w_i se $F_i > F_j$, para todo $j \neq i$. A utilização de uma matriz de covariância comum torna a função discriminante simplificada e linear (WEBB, 2002). Logo, um classificador bayesiano é composto por C funções de decisão que atuam em conjunto para estabelecer a regra decisão.

3.3.2. Redução da dimensionalidade dos dados

“Dimensionalidade é o número de características de uma representação de padrões, ou seja, a dimensão do espaço de características” (NASCIMENTO, 2003).

Dentre os principais motivos para reduzir a dimensionalidade dos dados em um sistema de classificação de padrões, Jain et al. (2000) destacam o custo de medição e a acurácia da classificação. Assim, com um conjunto limitado de atributos tanto a representação do padrão quanto os classificadores, que são construídos sobre tais representações, tornam-se simplificados, ou seja, mais rápidos e com menor uso de memória. Além disso, um pequeno número de características pode aliviar a ocorrência do fenômeno de Hughes, que trata de um fenômeno que ocorre quando um aumento no número de atributos pode degradar o desempenho de um classificador se o número de amostras de treinamento for pequena em relação à quantidade destes atributos. Contudo, é bem conhecido que a probabilidade de ocorrência dos erros de classificação não aumenta com o aumento do número de atributos desde que as probabilidades classe-condicional sejam completamente conhecidas ou, equivalentemente, se a quantidade de amostras de treinamento for arbitrariamente grande para representar estas probabilidades.

Para Jain & Chandrasekaran (1982), uma boa prática é adotar, para cada classe, dez amostras treinamento por atributo, a fim de evitar problemas acerca da dimensionalidade. Contudo, segundo Jain et al. (2000) e Webb (2002), apesar de algumas orientações serem sugeridas com relação ao tamanho das amostras e à dimensionalidade, é muito difícil de estabelecer uma relação exata entre a probabilidade de erro de classificação, o número de amostras e de atributos e os parâmetros das funções de probabilidade classe-condicional.

Deste modo, visando a redução da dimensionalidade, duas técnicas podem ser aplicadas, isoladas ou em conjunto, sendo elas a seleção e a extração de características. Para Jain et al. (2000), métodos que criam novos atributos baseados nas transformações ou combinações dos atributos do conjunto original fazem parte da extração de características, e seleção de atributos refere-se a métodos que selecionam, ou tentam selecionar, o melhor subconjunto de atributos presentes em um conjunto inicial.

Assim, para que seja possível selecionar melhor subconjunto de atributos, deve-se realizar uma análise entre todos os subconjuntos possíveis. “Entretanto, essa abordagem apresenta um custo computacional proibitivo, visto que o número de subconjuntos é exponencial em relação ao número de medidas obtidas na etapa de extração de características”. Deste modo, diversas técnicas propostas na literatura efetuam a redução da dimensionalidade, com satisfatória economia de gasto

computacional, por meios da seleção de um subconjunto sub-ótimo de atributos (PEDRINI; SCHWARTZ, 2008).

Dentre estas técnicas, podem ser citadas aquelas baseadas no modelo *Stepwise*, presentes em alguns *softwares* estatísticos, como o *Statistica*®. Conforme Hill & Lewicki (2007), os procedimentos básicos em um modelo *Stepwise* consistem em: 1) identificar um modelo inicial; 2) em várias etapas, e de forma iterativa, alterar o conjunto selecionado na etapa anterior através da entrada ou saída de características de acordo com um critério de seleção, como por exemplo, valores críticos da estatística *F*, e 3) encerrar o processo quando não for mais possível adicionar ou remover características em função do critério ou quando um número específico de etapas for atingido.

Dentre as técnicas baseadas no modelo *Stepwise* encontram-se: 1) *Forward entry*: o subconjunto começa vazio e a cada etapa uma estatística de entrada é calculada para cada característica, a fim de selecionar uma delas para fazer parte do subconjunto. O processo é encerrado após atingir um número de etapas especificado ou quando nenhuma característica restante fora do subconjunto apresentar um valor que supere o nível crítico da estatística para entrar. 2) *Backward removal*: o subconjunto começa com todas as características e a cada etapa uma estatística de saída é calculada para cada característica, a fim de selecionar uma delas para ser removida do subconjunto. O processo é encerrado após atingir um número de etapas especificado ou quando nenhuma característica restante dentro do subconjunto apresentar um valor inferior ao nível crítico da estatística para sair. 3) *Forward stepwise*: este método apresenta uma combinação entre os dois anteriores. O subconjunto começa vazio e na etapa 1 é executado o *forwad entry*. Nas etapas subsequentes executa-se o *forwad entry*, se possível, e o *backward removal*, se possível. Quando não for possível executar estes dois procedimentos, ou quando for atingido o limite de etapas, o processo encerra-se (HILL; LEWICKI, 2007).

3.3.3. Avaliação de classificadores

O erro de classificação, ou simplesmente, taxa de erro, é a principal medida de desempenho de um classificador e consiste na percentagem de amostras de validação classificadas erroneamente. A fim de que esta estimativa de erro seja confiável em

predizer o desempenho de classificação em aplicações futuras, não só o conjunto de treinamento e de validação dever ser suficientemente grandes, mas as amostras destes dois conjuntos devem ser independentes. Dessa forma, classificadores que competem por uma mesma aplicação podem ser comparados com base em suas probabilidades de erro. Outras medidas de desempenho podem incluir os custos para a mensuração dos atributos utilizados e os requerimentos computacionais para a aplicação da regra de decisão (JAIN et al., 2000).

Existem três principais abordagens sobre a forma como serão formados os conjuntos de treinamento e de validação, sendo eles a ressubstituição, a separação (Hold-Out), e a validação cruzada. Na ressubstituição, todas as amostras são utilizadas na fase de treinamento e também na avaliação do classificador (RAUDYS; JAIN, 1991). Isto gera estimativa de erro com viés otimista, principalmente quando a relação entre o tamanho de amostras e número de atributos é baixo (JAIN et al., 2000).

No método de separação, considerando um número total N de amostras, uma porção n é selecionada para fazer parte do conjunto de treinamento. As amostras restantes, $N - n$, é utilizada para compor o conjunto de validação (RAUDYS; JAIN, 1991). Diferentes particionamentos irão proporcionar diferentes estimativas de erro, que contêm viés pessimista. Além disso, este método obriga o investigador a confrontar um complexo questionamento: como as amostras disponíveis deveriam ser dispostas entre os conjuntos de treinamento e de validação? Se a amostras de treinamento for pequena, o classificador resultante poderá não ser muito robusto e ter uma baixa capacidade de generalização. Por outro lado, se o conjunto de validação for pequeno, também o será a confiabilidade da estimativa de erro (JAIN et al., 2000).

Na validação cruzada, $N-k$ amostras são utilizadas para treinar o classificador, que, em seguida, realiza a classificação das k amostras que não foram utilizadas. O processo se repete para todas as possíveis configurações do conjunto k e a taxa de erro média é calculada. Uma escolha popular é fazer $k=1$ resultando no conhecido método *deixe-um-fora* (*leave-one-out*), em que N classificações são realizadas (RAUDYS; JAIN, 1991). Uma grande vantagem deste método é produzir uma estimativa de erro livre de viés. Além disso, visa aproveitar um maior número de amostras nas fases de treinamento e validação sem, contudo, permitir que uma amostra seja avaliada por um classificador cuja etapa de treinamento tenha contemplado a amostra em questão.

Entretanto, este método gera estimativas com altas variâncias e exige um elevado custo computacional (JAIN et al., 2000).

Uma vez escolhida a forma como as amostras serão utilizadas, a etapa final é determinar a forma como os erros e acertos da classificação serão organizados. Uma forma útil de se fazer essa representação é através da matriz de erros, ou matriz de confusão, nesta, as exatidões são mostradas por classe, possibilitando determinar as acurácias do produtor e do usuário. A acurácia do produtor indica a porcentagem de amostras de uma dada classe que foram corretamente classificadas, já a acurácias do usuário indica, dentre as amostras classificadas como pertencentes a uma dada classe, a porcentagem daquelas que foram corretamente classificadas (CONGALTON, 1991).

“Uma matriz de erros é uma matriz quadrada C com um número de linhas i e colunas j equivalentes ao número de classes presentes no experimento. O elemento $c_{i,j}$ indica o número de amostras atribuídas à classe w_i dado que a classe correta seria w_j ” (PEDRINI; SCHWARTZ, 2008). A partir da matriz de erros é possível calcular uma série de estatísticas descritivas e analíticas, tal como o índice de Exatidão Global (CONGALTON, 1991) e o índice *Kappa* (HUDSON; RAMM, 1987).

Dentre outras técnicas importantes para a avaliação de classificadores, cita-se o teste estatístico pareado para testar a significância de um classificador frente a uma classificação ao acaso. O valor Z de uma classificação pode ser computado a partir da estimativa do índice *Kappa* dividida pela raiz quadrada de sua variância, conforme citado por Vieira (2000). Assim, para um nível de confiança α , se o Z calculado for maior que o valor crítico $Z_{\alpha/2}$, a classificação é significativamente melhor que a classificação ao acaso. Conforme Congalton & Mead (1983), também é possível estabelecer a significância entre dois classificadores por meio do teste Z .

4. Materiais e Métodos

4.1. Caracterização e preparo do material carbonizado

Este trabalho foi desenvolvido no Laboratório de Painéis e Energia da Madeira (LAPEM) da Universidade Federal de Viçosa. O material estudado foi obtido na carboteca do LAPEM, constituída por amostras circulares de lenho carbonizado de espécies e híbridos do gênero *Eucalyptus* e de nativas típicas da Zona da Mata Mineira - Mata Atlântica.

Nesta carboteca, cada espécie ou híbrido é representado por uma amostra proveniente de uma única árvore, sem informação de idade ou de posição no tronco. Os carvões oriundos da carboteca foram carbonizadas em um forno elétrico tipo mufla, seguindo uma marcha com temperatura inicial de 150°C, incremento de 50°C por hora e temperatura final de 450°C, totalizando 7 horas de carbonização.

As espécies e híbridos utilizados neste estudo são apresentados na Tabela 2.

Tabela 2 - Espécies e híbridos estudados.

Identificação	Nome vulgar	Nome científico
1	Angico vermelho	<i>Anadenanthera peregrina</i> Speg.
2	Brauninha	<i>Dyctiolum vandellianum</i> A. Juss.
3	Camaldulensis	<i>Eucalyptus camaldulensis</i> Dehnh.
4	Cedrinho	<i>Trattinnickia ferrugínea</i> Kuhlmann
5	Grandis x Camaldulensis	<i>Eucalyptus grandis</i> x <i>Eucalyptus camaldulensis</i>
6	Mama de porca	<i>Zanthoxylum rhoifolium</i> Lam.
7	Pimenteira	<i>Xylopia sericea</i> A. St.-Hil.
8	Ruão	<i>Vismia martiana</i> H. G. Reich
9	Urocam	<i>Eucalyptus urophylla</i> x <i>Eucalyptus camaldulensis</i>
10	Urograndis	<i>Eucalyptus. urophylla</i> x <i>Eucalyptus grandis</i>
11	Urophylla (GG 157)	<i>Eucalyptus urophylla</i> S. T. Blake
12	Urophylla (GG100)	<i>Eucalyptus urophylla</i> S. T. Blake
13	Casca doce	<i>Alchornea glandulosa</i> Poepp. & Endl.
14	Goiabeira	<i>Psidium guajava</i> L.
15	Pau fumo	<i>Piptocarpha macropoda</i> Baker
16	Quaresminha	<i>Miconia candolleana</i> Triana
17	Cedro	<i>Cedrella fissilis</i> Vellozo
18	Casuarina	<i>Casuarina equisetifolia</i> L.

Para obtenção das imagens digitais dos carvões vegetais, primeiramente foi realizado um protocolo de preparo das superfícies das amostras a serem utilizadas para a

captura de imagem. As peças de carvão foram seccionadas, no plano transversal, com auxílio de uma lâmina de serra manual metálica e posteriormente lixadas, em ambas as faces, usando, em sequência, lixas d'água de grãos 600 e 1500 de forma a mitigar as marcas e demais padrões não condizentes com a anatomia do carvão, tais como aqueles causados pela serra metálica.

Todos os procedimentos descritos acima foram realizados sob água corrente, de modo a facilitar as operações e atenuar o entupimento dos poros por partículas menores. Depois, as amostras de carvão foram submetidas a ar pressurizado, para remoção das partículas que por ventura tenham permanecido, e, por fim, armazenadas em potes plásticos com identificação.

4.2. Aquisição de imagens digitais

A aquisição de imagens foi realizada com uma câmera *Pixelink*, modelo PL-A662 acoplada a uma lupa estereoscópica ZEISS Stemi 2000-C e conectada a um computador por meio de uma placa LSI 1394 OHLC. Por meio da lupa as imagens das amostras de carvão foram ampliadas em 25 vezes.

As imagens foram produzidas pelo *software Pixelink capture* e salvas com extensão JPEG, dimensões 1280 (H) x 1024 (V) *pixels*, resolução espacial de 3,8 *Micrometros/pixel* e profundidade de *pixel* de 8 bits.

Foram utilizadas duas fontes de luz incandescentes de 20 Watts, distantes aproximadamente 4 cm do carvão e cobertas com papel “manteiga” para evitar a incidência direta. As projeções de luz incidiram em sentidos opostos, fazendo ângulo de 45° com a superfície.

Toda peça de carvão foi fotografada na seção transversal, na região de cerne e alburno, de modo que seus raios apresentassem verticalidade no centro da imagem e origem na porção inferior desta. Evitaram-se as áreas do carvão que apresentavam descaracterização da estrutura anatômica provenientes da carbonização.

4.3. Extração de atributos

Para cada espécie e híbrido mostrado na Tabela 2, foram geradas 50 imagens da seção transversal das amostras de carvão, que foram utilizadas na forma como foram

produzidas, ou seja, sem qualquer tipo de pré-processamento. Dessa forma, foram produzidas, ao todo, 300 fotos dos carvões produzidos com madeira de espécies e híbridos do gênero *Eucalyptus* e 600 fotos dos carvões produzidos com madeira de espécies nativas.

A extração de atributos foi realizada através de um algoritmo implementado no programa computacional Matlab®, que, a partir de cada imagem inserida, gerou-se 24 imagens-índice com três diferentes resoluções de imagem (número de *pixels*) e oito níveis de profundidades de *pixel* (*bits*), conforme detalhado na Tabela 3.

Tabela 3 - Parâmetros das imagens-índice geradas no programa computacional Matlab®.

Resoluções (número de <i>pixels</i>)	1280x1024; 320x256; 160x128
Profundidades de <i>pixel</i> (<i>bits</i>)	1, 2, 3, 4, 5, 6, 7 e 8

Para a alteração da resolução utilizou-se o comando *imresize*, que utiliza, por padrão, o método de interpolação bicúbico. Para a alteração da profundidade da imagem, os valores dos *pixels* foram divididos por um fator (Equação 4) e o resultado foi arredondado para baixo.

$$Fator = \frac{2^8}{2^c} \quad (4)$$

Em que,

c = número de bits da imagem modificada.

Em seguida, para cada imagem-índice, foram criadas, de acordo com Haralick et al. (1973), 20 matrizes de coocorrência de níveis de cinza (MCNCs) com cinco e quatro valores para d (*pixels*) e θ (*graus*), respectivamente (Tabela 4).

Tabela 4 - Parâmetros das MCNCs geradas, de acordo com Haralick et al. (1973).

d (<i>pixels</i>)	1, 2, 3, 4 e 5
θ (<i>graus</i>)	0, 45, 90 e 135

Depois, para cada MCNC criada, foram calculadas suas seis características básicas de textura, segundo as Equações 5 a 10, conforme proposto por Haralick et al.

(1973). Assim, de cada imagem-índice foram obtidas, ao todo, 120 características de textura (produto de 20 MCNCs e 6 características básicas de textura), que formaram um vetor de atributos. Para cada imagem de carvão obteve-se, portanto, 24 vetores.

$$C_{sma} = \sum_{i=1}^L \sum_{j=1}^L [p(i,j)]^2 \quad (5)$$

$$C_{var} = \sum_{i=1}^L \sum_{j=1}^L (i - \mu)^2 p(i,j) \quad (6)$$

$$C_{ent} = \sum_{i=1}^L \sum_{j=1}^L p(i,j) \ln p(i,j) \quad (7)$$

$$C_{con} = \sum_{i=1}^L \sum_{j=1}^L (i - j)^2 p(i,j) \quad (8)$$

$$C_{corr} = \frac{\sum_{i=1}^L \sum_{j=1}^L (i - j)^2 p(i,j)}{C_{var}} \quad (9)$$

$$C_{hom} = \sum_{i=1}^L \sum_{j=1}^L \frac{p(i,j)}{1 + (i - j)^2} \quad (10)$$

Em que,

C_{sma} = Segundo momento angular;

C_{var} = Variância;

C_{ent} = Entropia;

C_{con} = Contraste;

C_{corr} = Correlação;

C_{hom} = Homogeneidade;

$p(i,j)$ = Elemento da matriz de coocorrência na i -ésima linha e j -ésima coluna;

L = quantidade de níveis de cinza da imagem, e

μ = média determinada pela Equação 11.

$$\mu = \sum_{i=1}^L \sum_{j=1}^L i p(i,j) \quad (11)$$

4.4. Desenvolvimento dos classificadores

Em seguida, foi desenvolvido um classificador para cada combinação de Resolução de imagem e Profundidade de *pixel*, ou seja, 24 combinações, conforme apresentado na Tabela 5.

Tabela 5 – Atributos utilizados no desenvolvimento dos classificadores.

Classificadores	Parâmetros das imagens-índices		Parâmetros das MCNCs		Características básicas das MCNCs
	Resolução (pixels)	Profundidade de imagem (bits)	d (<i>pixels</i>)	θ (graus)	
1	1280x1024	1	1, 2, 3, 4 e 5	0, 45, 90 e 135	Segundo momento angular, variância, entropia, contraste, correlação, homogeneidade
2		2			
3		3			
4		4			
5		5			
6		6			
7		7			
8		8			
9	320x256	1	1, 2, 3, 4 e 5	0, 45, 90 e 135	
10		2			
11		3			
12		4			
13		5			
14		6			
15		7			
16		8			
17	160x128	1	1, 2, 3, 4 e 5	0, 45, 90 e 135	
18		2			
19		3			
20		4			
21		5			
22		6			
23		7			
24		8			

Dentre os 120 atributos gerados para cada classificador, foram selecionados aqueles que viriam compor o modelo da função discriminante, escolhendo aqueles que

apresentavam a contribuição mais significativa para a discriminação das classes de carvões estudadas.

A seleção de atributos foi realizada no programa computacional Statistica 7 por meio do procedimento *Forward Stepwise* utilizando como critério de seleção o valor F obtido por meio da λ de Wilk, conforme Rao (1951). O valor limítrofe de F para a entrada de um atributo foi arbitrariamente escolhido como $F_{5\%}(1, 900) = 3,85$. O valor crítico de F para a saída de um atributo foi arbitrariamente escolhido como o valor imediatamente inferior a F para entrada do atributo. Estabeleceu-se um número máximo de etapas igual a 120.

Depois, os atributos selecionados foram utilizados na construção dos classificadores segundo o modelo Bayesiano, considerando uma matriz de covariância comum entre as classes *eucalipto* e *nativa* (S_w). Assim, foram criadas duas funções de decisão (Equação 12) para cada classificador, de acordo com Webb (2002).

$$F_i = \log(p(w_i)) - \frac{1}{2} m_i^t S_w^{-1} m_i + x^t S_w^{-1} m_i \quad (12)$$

Em que,

F_i = função de decisão da classe i ;

$p(w_i)$ = probabilidade *a priori* de ocorrência da classe i ;

m_i^t = transposta do vetor média dos atributos da classe i ;

S_w^{-1} = matriz de covariância comum inversa;

x^t = transposta do vetor de atributos a ser classificado;

A matriz de covariância comum entre as classes foi calculada conforme a Equação 13 (WEBB, 2002).

$$S_w^{-1} = \sum_{i=1}^2 \frac{n_i}{n} S_i \quad (13)$$

Em que,

n_i = número de vetores da classe i ;

n = número total de amostras;

S_i = matriz de covariância da classe i , na qual cada elemento S_{xy} denota a covariância entre o atributo x e o atributo y .

4.5. Avaliação dos classificadores

Através de um algoritmo implementado no programa computacional Matlab® cada classificador foi avaliado empregando-se a validação cruzada, de forma que, dentre os 900 vetores de atributos disponíveis, 899 foram utilizadas para desenvolver as funções discriminantes que, em seguida, foram utilizadas para classificar o vetor não incluído na etapa anterior. Esse procedimento foi repetido 900 vezes, sendo aplicado a um diferente vetor em cada repetição. Assim, assumindo que os dados apresentavam distribuição normal, um vetor “ x ” foi associado à classe cuja função de decisão resultasse no maior valor.

Com os resultados obtidos pela validação cruzada, foram construídas matrizes de erros para cada classificador, conforme descrito por Congalton (1991). Com base nessas matrizes, foram calculados os índices de Exatidão Global e as Acurácias do Produtor e do Usuário para cada classe, conforme Equações 14, 15 e 16, respectivamente (CONGALTON, 1991).

$$EG = \frac{\sum_{i=1}^c x_{ii}}{n} \quad (14)$$

$$AP_w = \frac{x_w}{x_{\oplus w}} \quad (15)$$

$$AU_w = \frac{x_w}{x_{w\oplus}} \quad (16)$$

Em que,

EG = exatidão global da classificação;

AP_w = estimativa da acurácia do produtor para a classe W;

AU_w = estimativa da acurácia do usuário para a classe W;

x_{ii} = valor da célula na linha i e coluna i (diagonal) da matriz de erros;

n = número total de vetores utilizados na validação;

c = número total de classes.

x_w = número de observações da classe W classificadas como W;

$x_{\oplus w}$ = total de observações da classe W inseridas no classificador.

$x_{w\oplus}$ = total de observações classificadas como W.

Através das matrizes de erro, foram ainda calculadas as estimativas dos coeficientes *Kappa* (HUDSON; RAMM, 1987) e suas variâncias (BISHOP et al., 1975 citados por HUDSON; RAMM, 1987), segundo as Equações 17 e 18, respectivamente.

$$K_j = \frac{n \sum_{i=1}^c x_{ii} - \sum_{i=1}^c x_{i\oplus} x_{\oplus i}}{n^2 - \sum_{i=1}^c x_{i\oplus} x_{\oplus i}} \quad (17)$$

Em que,

- K_j = estimativa do coeficiente kappa do classificador j ;
- x_{ii} = valor na linha i e coluna i (diagonal) da matriz de erros;
- $x_{i\oplus}$ = total da linha i ;
- $x_{\oplus i}$ = total da coluna i ;
- n = número total de vetores; e
- c = número total de classes.

$$var(K_j) = \frac{1}{n} \left[\frac{\theta_1(1-\theta_1)}{(1-\theta_2)^2} + \frac{2(1-\theta_1)(2\theta_1\theta_2 - \theta_3)}{(1-\theta_2)^3} + \frac{(1-\theta_1)^2(\theta_4 - 4\theta_2^2)}{(1-\theta_2)^4} \right] \quad (18)$$

Em que,

$var(K_j)$ = variância do coeficiente kappa do j -ésimo classificador;

$$\theta_1 = \frac{1}{n} \sum_{i=1}^c x_{ii}$$

$$\theta_2 = \frac{1}{n^2} \sum_{i=1}^c x_{i\oplus} x_{\oplus i}$$

$$\theta_3 = \frac{1}{n^2} \sum_{i=1}^c x_{ii}(x_{i\oplus} + x_{\oplus i})$$

$$\theta_4 = \frac{1}{n^3} \sum_{i=1}^c \sum_{j=1}^c x_{ij}(x_{j\oplus} + x_{\oplus j})^2$$

Em seguida, através do Teste Z bilateral (Equação 19) ao nível de confiança de 5%, foi verificado se os classificadores apresentaram capacidade de distinguir os carvões das classes *eucalipto* e *nativa* com desempenho melhor que uma distinção ao acaso ($K = 0$). Assim, se $Z_{calculado} < Z_{tabelado(2,5\%)}$, então $K \neq 0$ ao nível de 5% de probabilidade.

$$Z = \frac{K}{\text{var}(K)} \quad (19)$$

Posteriormente, ao nível de confiança de 5%, os coeficientes *Kappa* dos classificadores foram, dois a dois, comparados (Equação 20). Então, se $Z_{\text{calculado}} < Z_{(2,5\%)}$, $K_1 = K_2$ ao nível de 5% de probabilidade.

$$Z = \frac{K_1 - K_2}{\sqrt{\text{var}(K_1) + \text{var}(K_2)}} \quad (20)$$

5. Resultados e discussão

5.1. Aquisição de imagens

O uso de lixas d'água permitiu um adequado acabamento da superfície do carvão para o nível de detalhamento anatômico necessário. Por meio da metodologia empregada, a seção transversal de uma amostra de carvão gastou em média 10 minutos para ser preparada.

Mesmo com o uso de ar comprimido sobre a superfície do carvão, verificou-se que alguns vasos ficaram obstruídos por partículas oriundas do processo de polimento, dificultando a avaliação segura da presença de tilose com o nível de detalhamento utilizado.

Nas Figuras 6 e 17 são apresentadas as imagens da seção transversal do carvão de espécies da classe *nativa*. Nas Figuras 18 e 23 são apresentadas imagens da seção transversal do carvão de espécies da classe *eucalipto*.

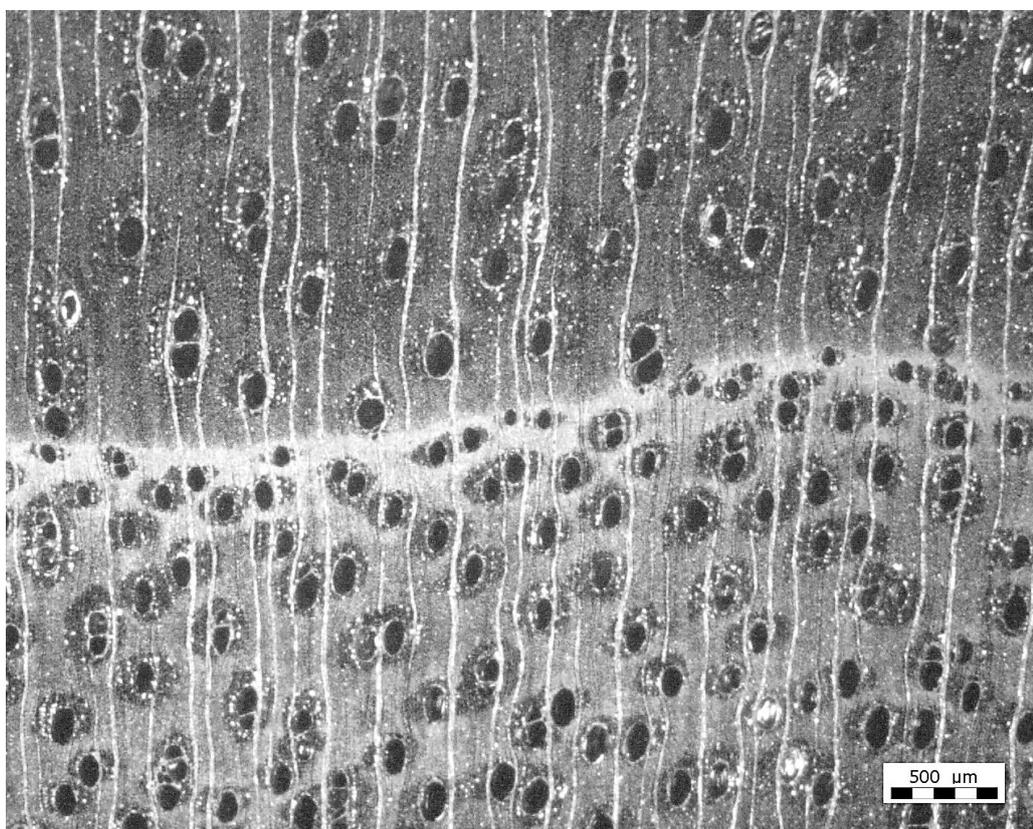


Figura 6 – Imagem da seção transversal do carvão de Angico-vermelho.

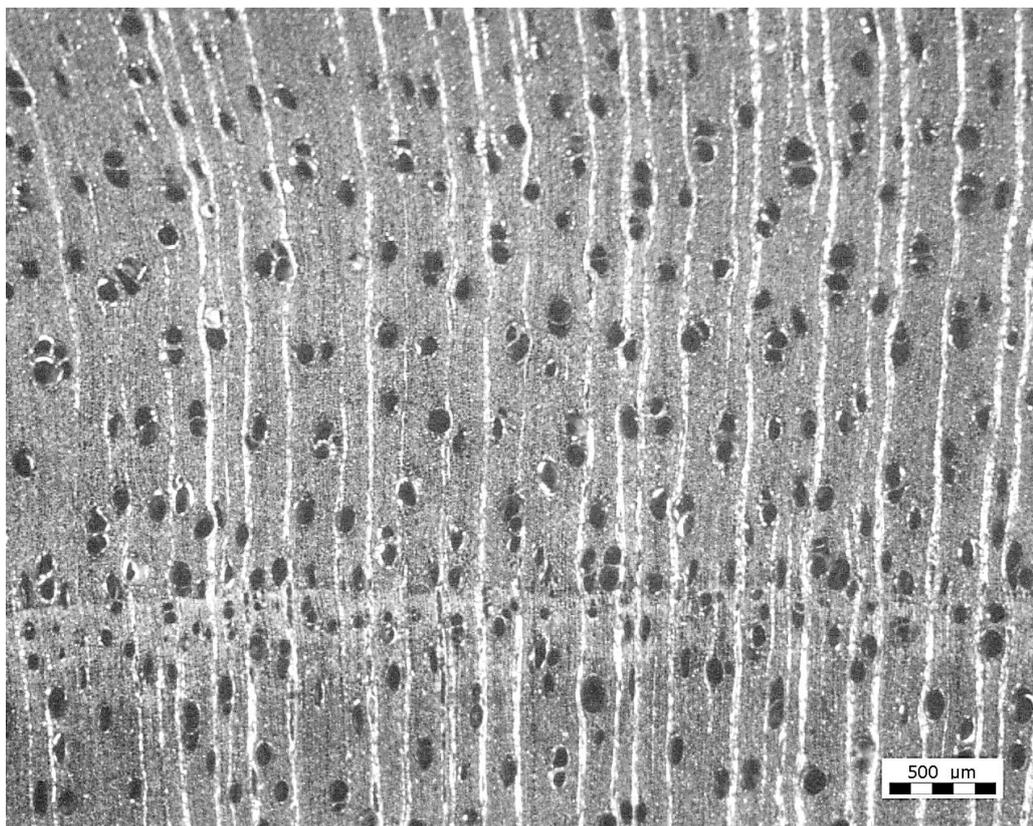


Figura 7 - Imagem da seção transversal do carvão de Brauningha.

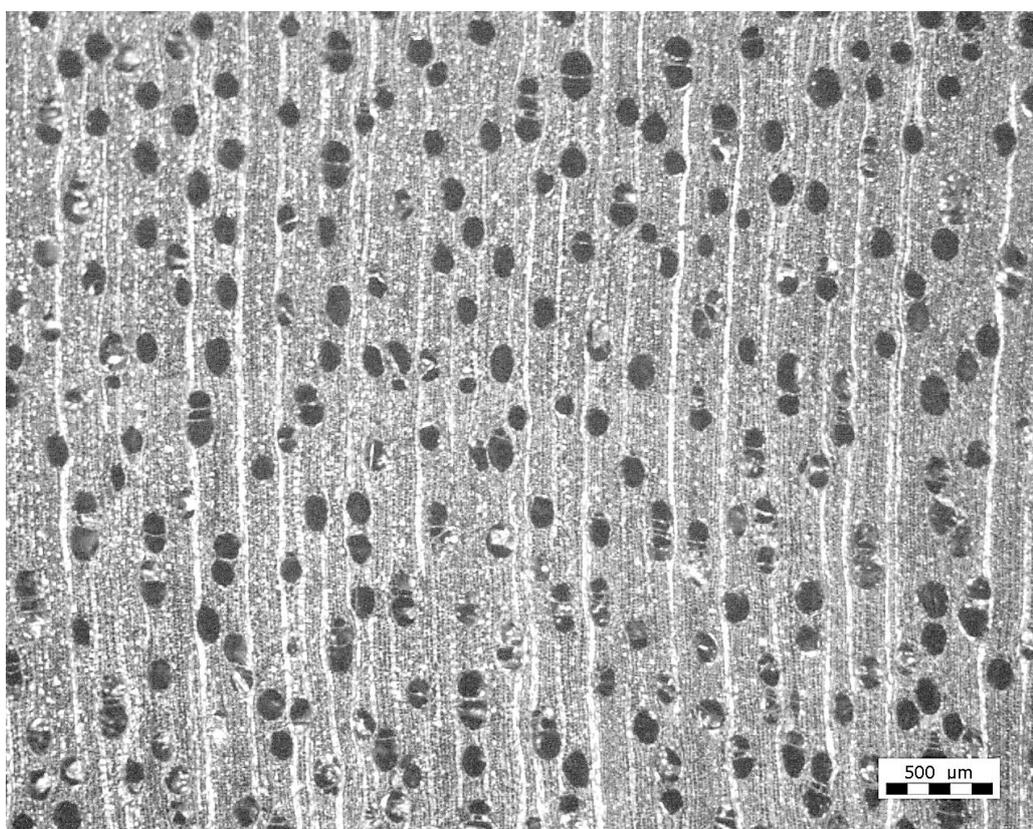


Figura 8 - Imagem da seção transversal do carvão de Cedrinho.

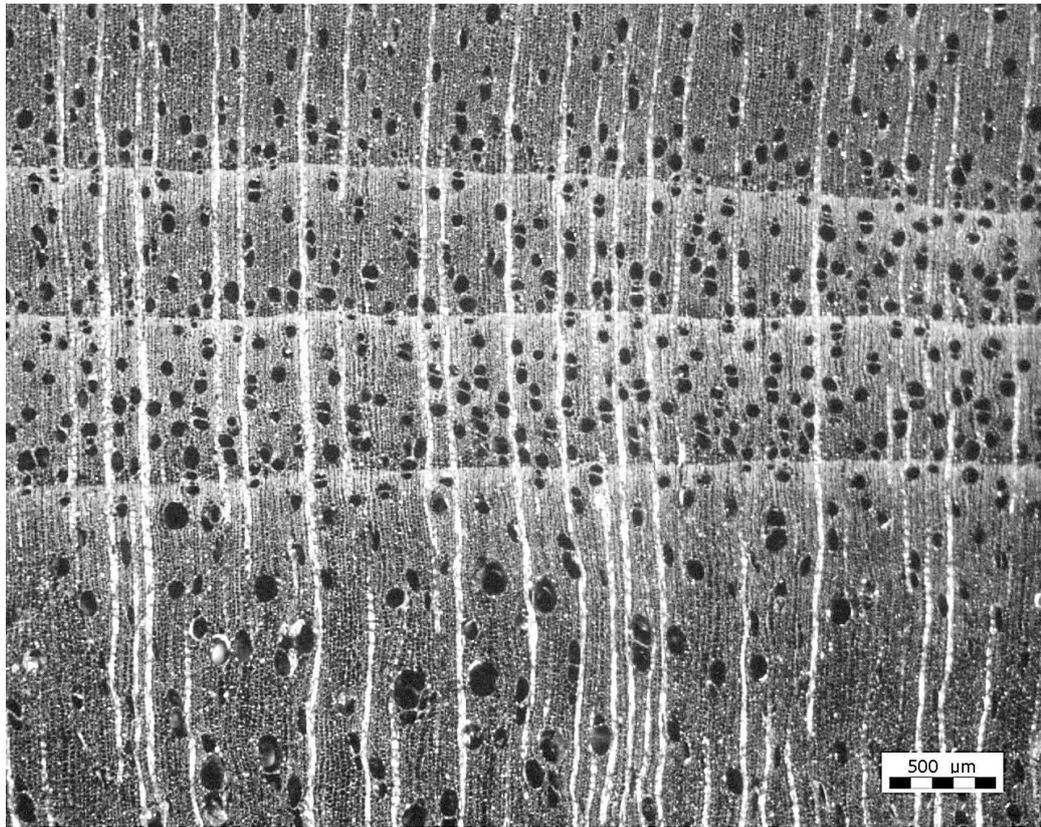


Figura 9 - Imagem da seção transversal do carvão de Mama de Porca.

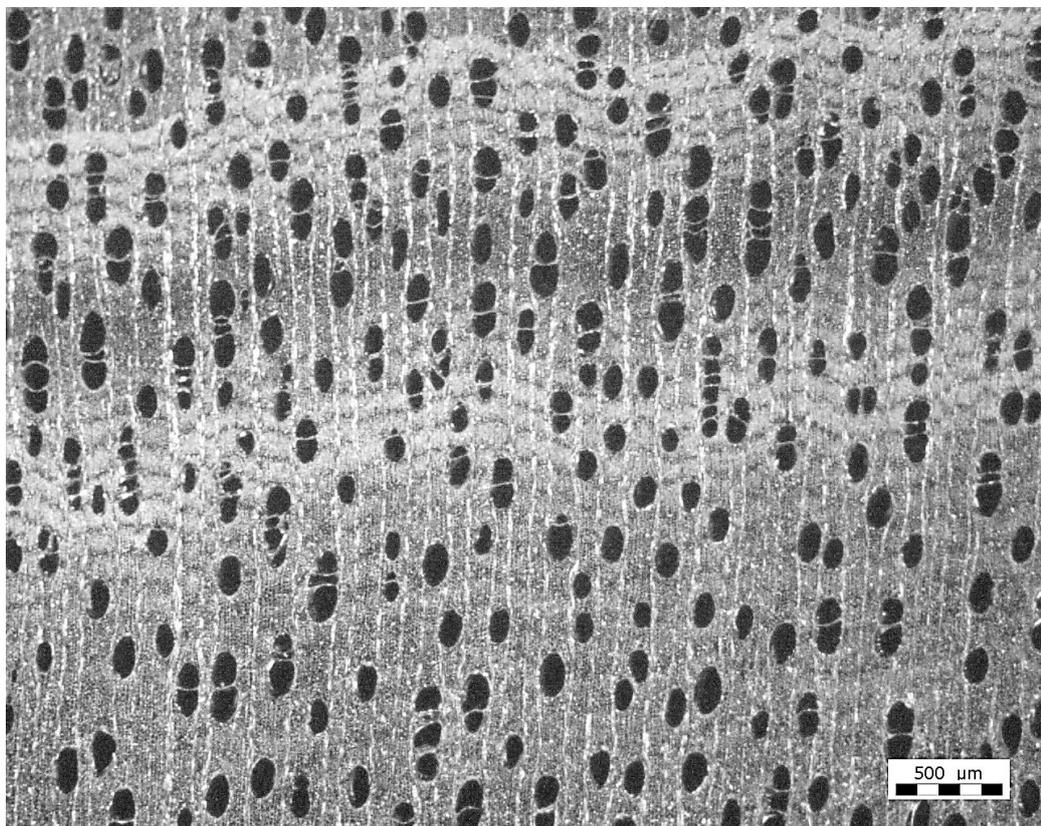


Figura 10 - Imagem da seção transversal do carvão de Pimenteira.

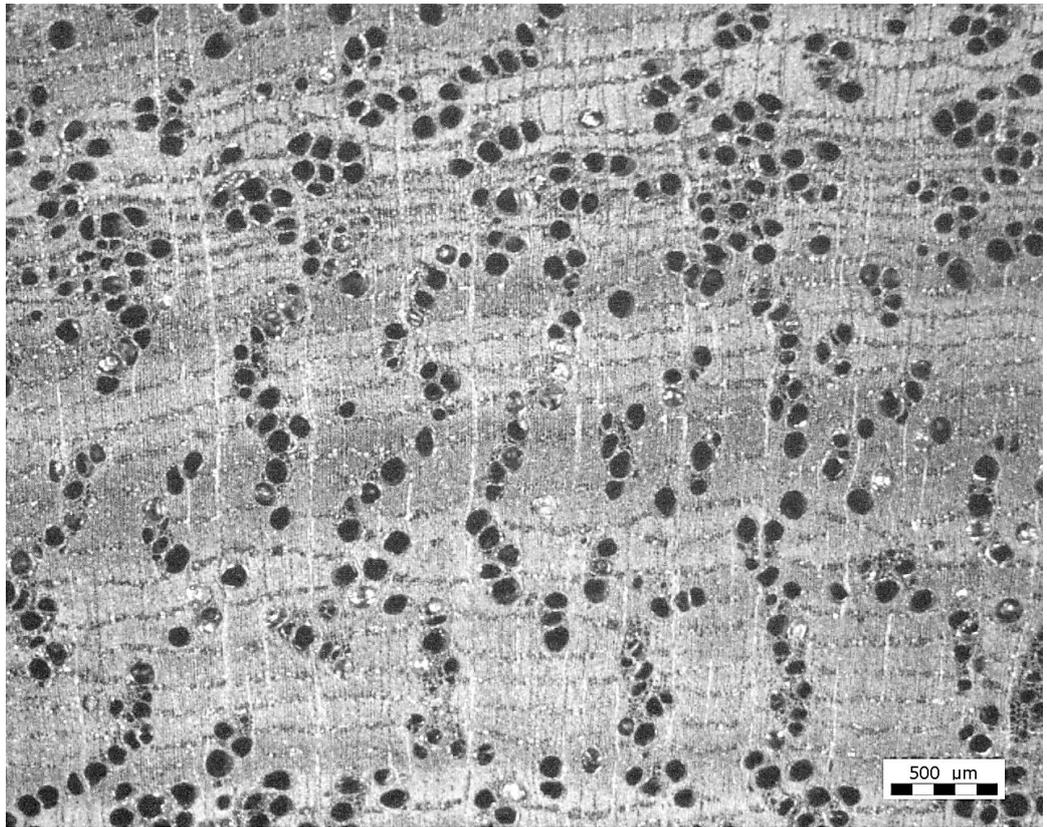


Figura 11 - Imagem da seção transversal do carvão de Ruão.

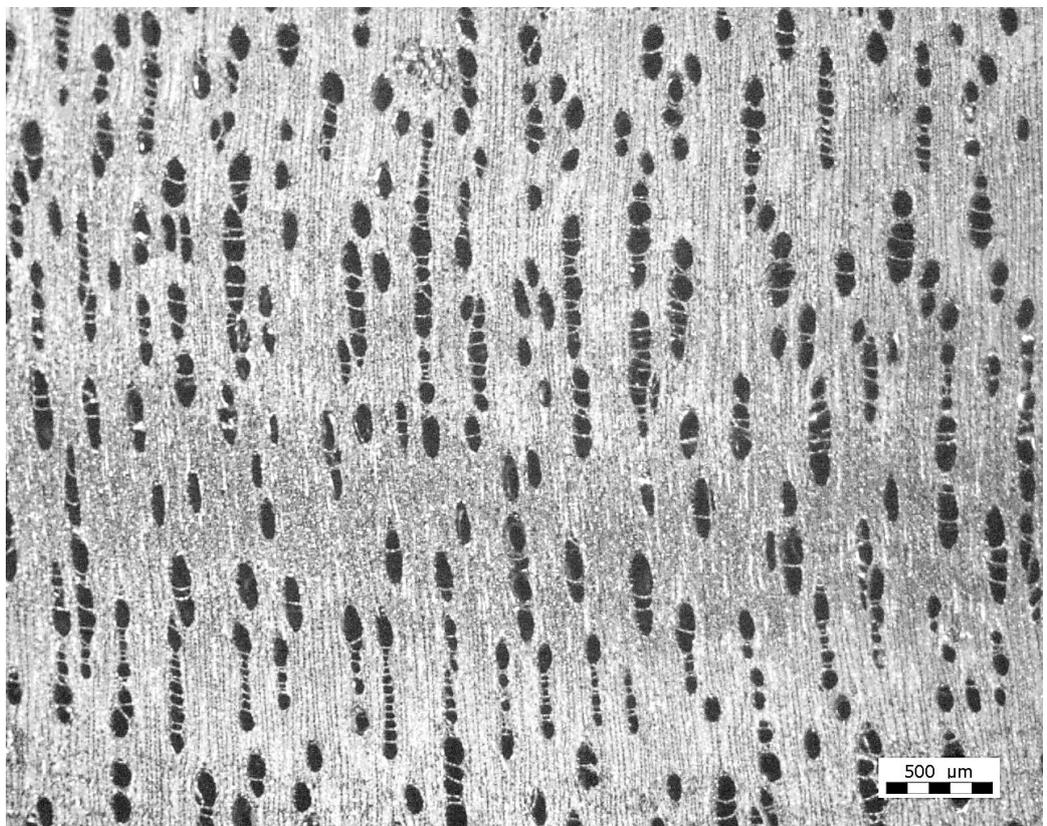


Figura 12 - Imagem da seção transversal do carvão de Casca Doce

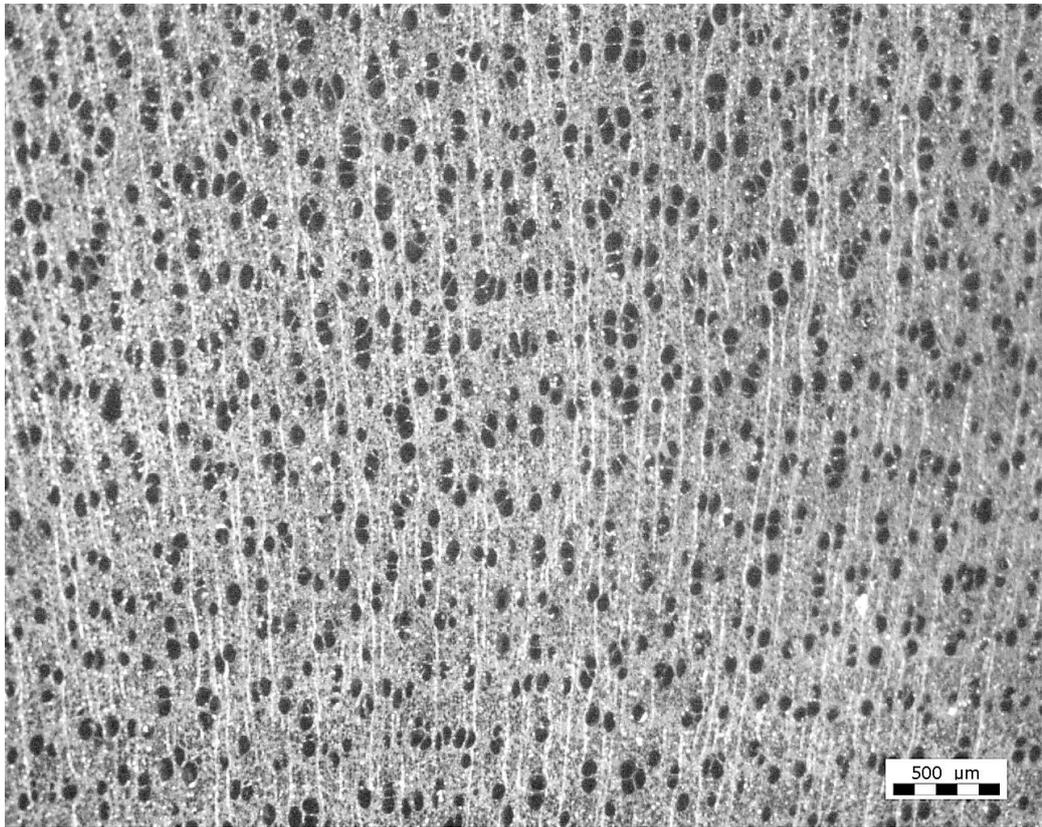


Figura 13 - Imagem da seção transversal do carvão de Goiabeira.

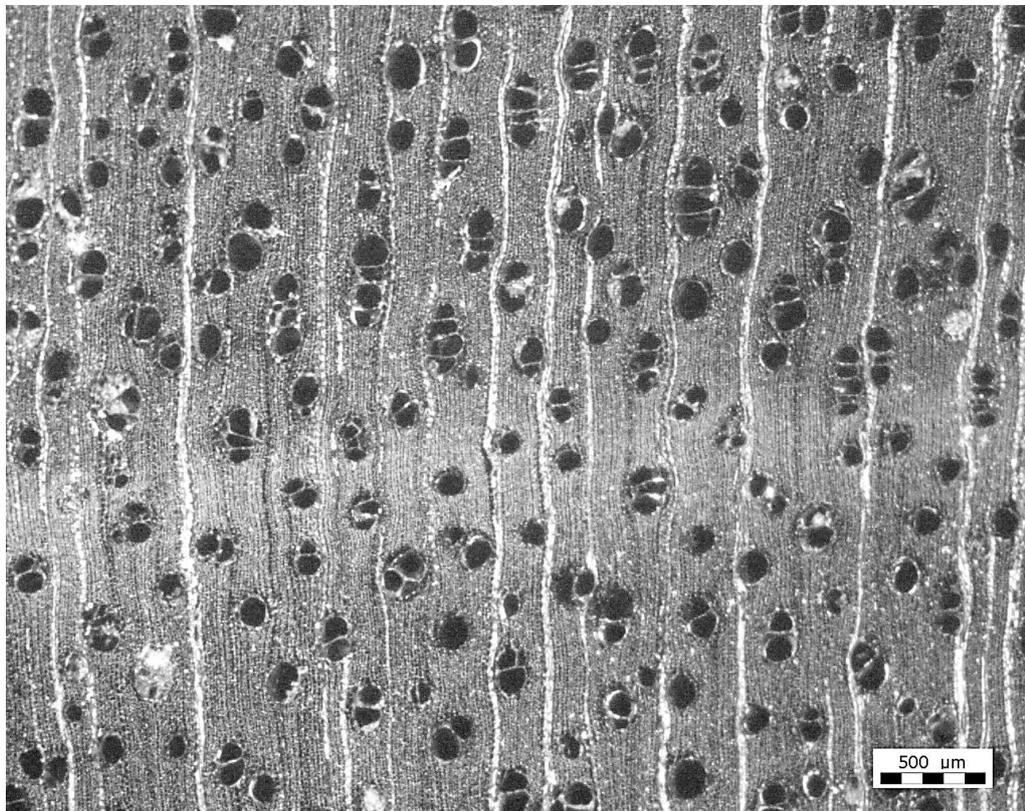


Figura 14 - Imagem da seção transversal do carvão de Pau Fumo.

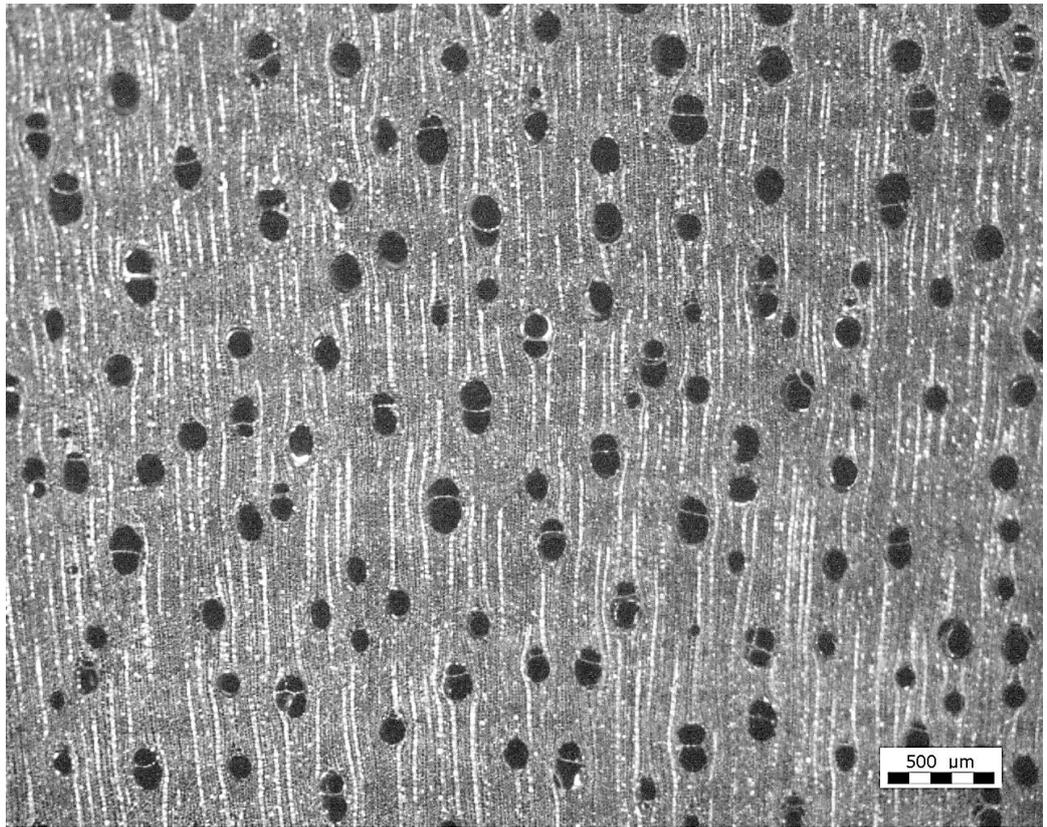


Figura 15 - Imagem da seção transversal do carvão de Quaresminha.

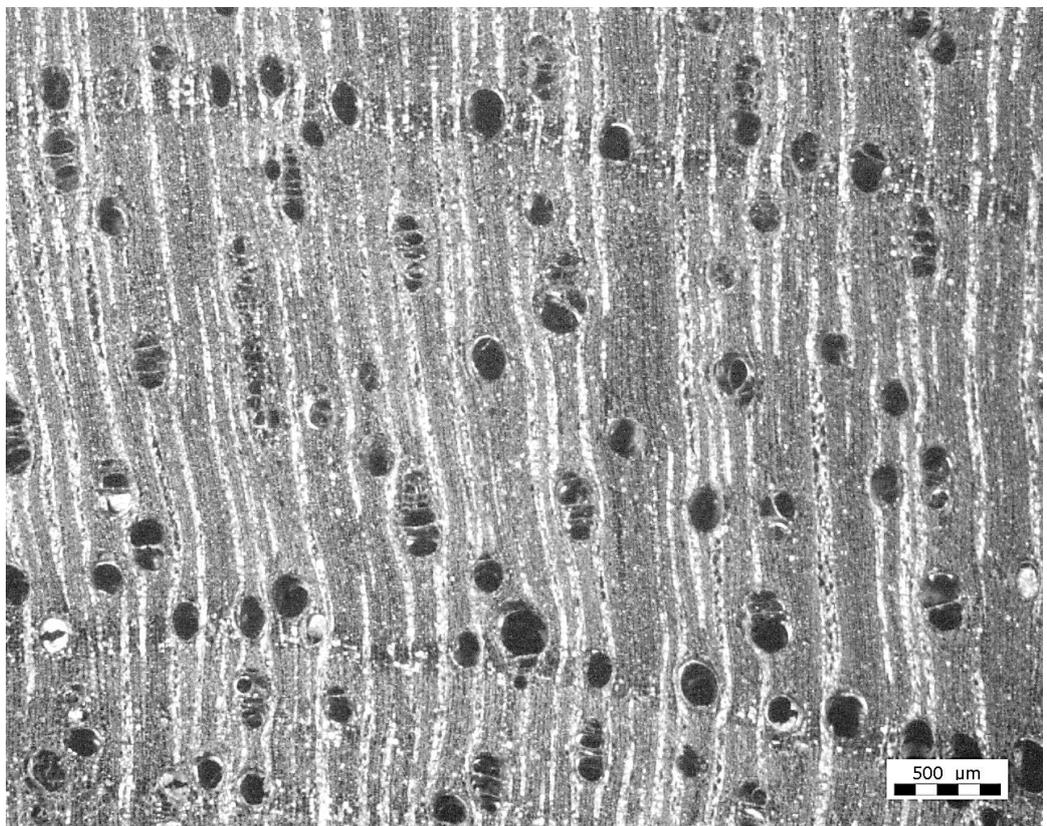


Figura 16 - Imagem da seção transversal do carvão de Cedro.

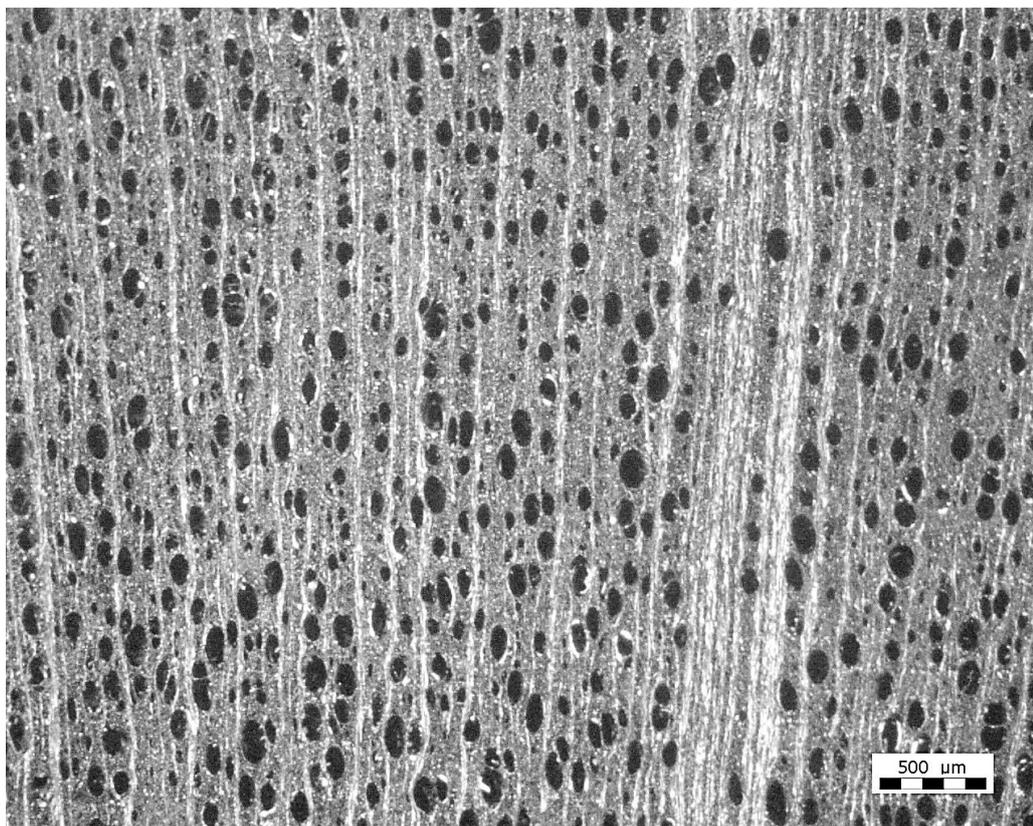


Figura 17 - Imagem da seção transversal do carvão de Casuarina.

Nos carvões da classe *nativa* predominou a porosidade difusa, mas ocorrendo também as porosidade semi-difusa e em anel. O Ruão foi a única espécie desta classe com poros exclusivamente solitários, embora próximos entre si. Segundo IAWA (1989), este tipo de proximidade, comum aos poros solitários do gênero *Eucalyptus*, não constitui agrupamento múltiplo. Nas demais espécies nativas, visualizou-se, além de poros solitários, a ocorrência de geminados, destacando-se a Quaresminha e a Pimenteira, e múltiplos, como em Casca Doce e Cedro. Quanto ao arranjo dos poros, nas espécies nativas foi observado o tipo radial, evidente em poros agrupados. O arranjo em diagonal e dendrítico também foram verificados para o Ruão.

Em geral, na classe *nativa* foram observados raios reflectivos e largos em sete das 12 espécies estudadas, destacando-se o Cedro, a Brauninha e o Pau Fumo. Nas demais espécies nativas verificaram-se raios pouco reflectivos e finos, com exceção da Casca Doce, cujos raios foram não reflectivos e, portanto, de difícil observação. Em geral, o processo de carbonização dificultou a observação do parênquima axial, mas foi possível visualizar os tipos confluyente e vasicêntrico em Angico Vermelho, em faixas regulares em Cedro e reticulado em Ruão e Pimenteira.

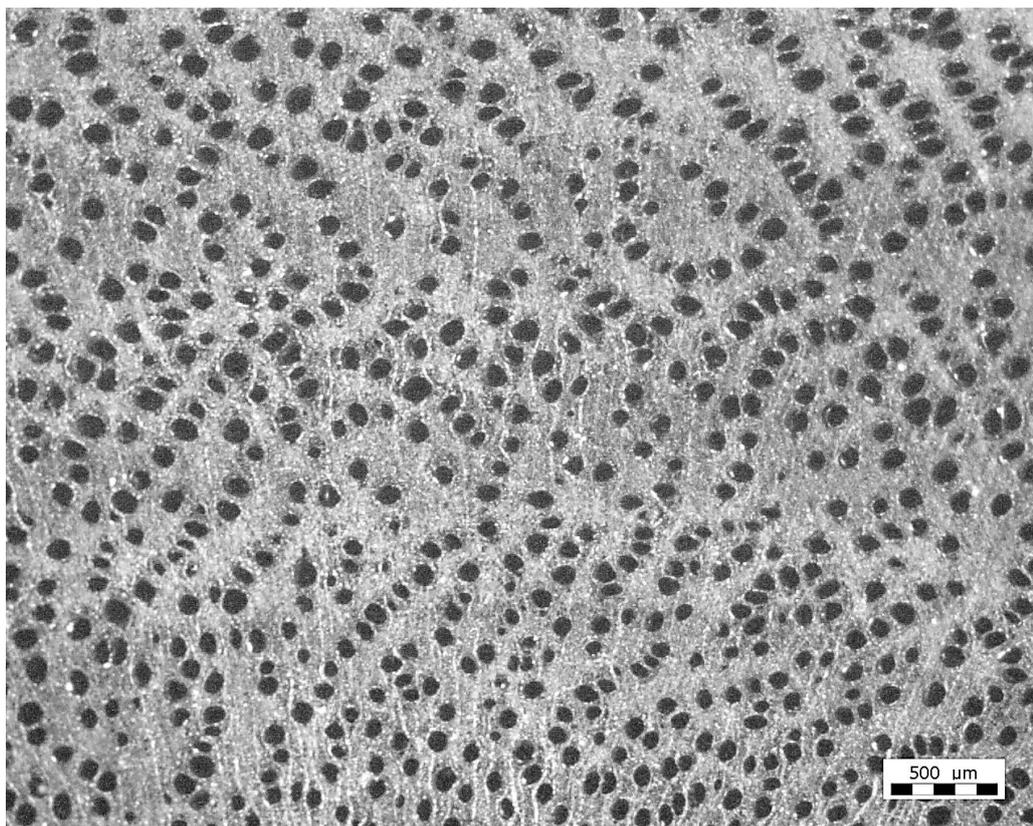


Figura 18 - Imagem da seção transversal do carvão de *E. camaldulensis*.

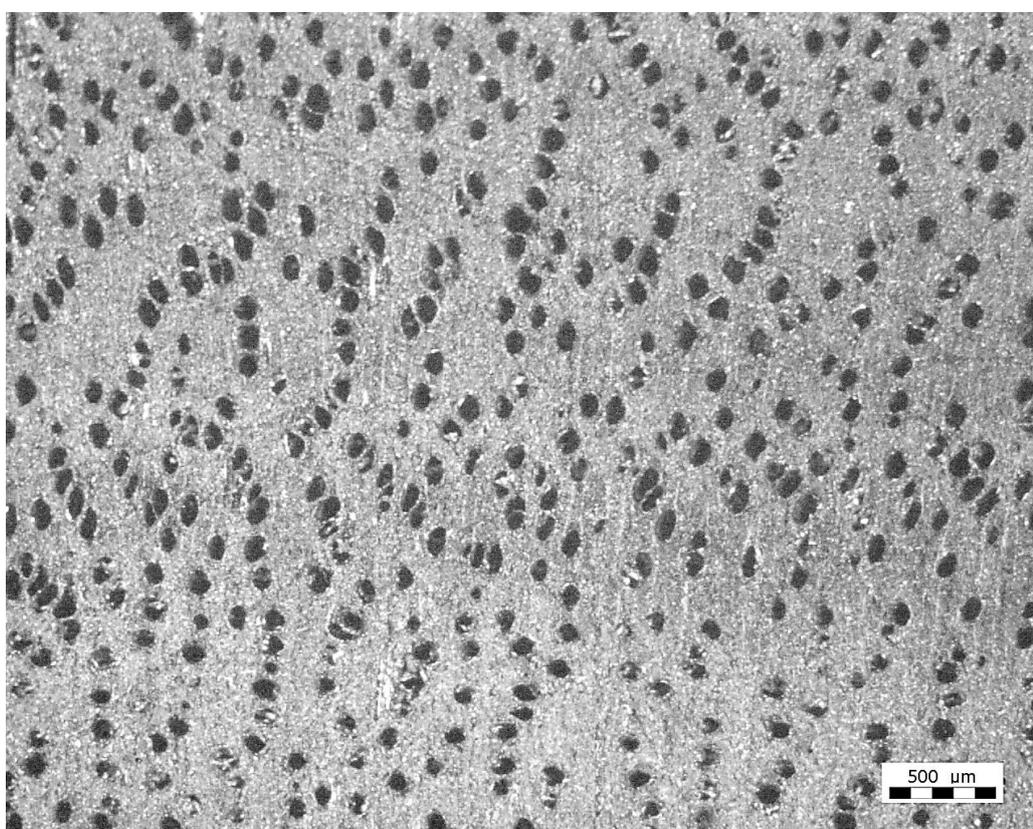


Figura 19 - Imagem da seção transversal do carvão de *E. camaldulensis* x *E. grandis*.

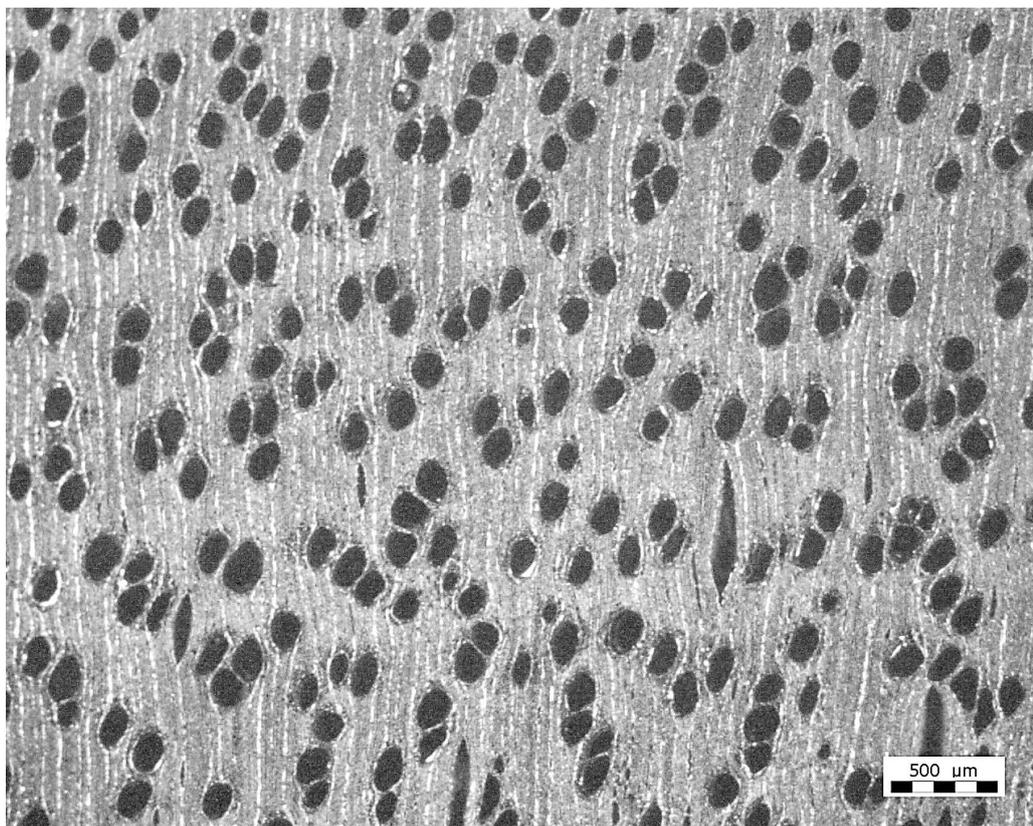


Figura 20 - Imagem da seção transversal do carvão de *E. camaldulensis* x *E. urophylla*.

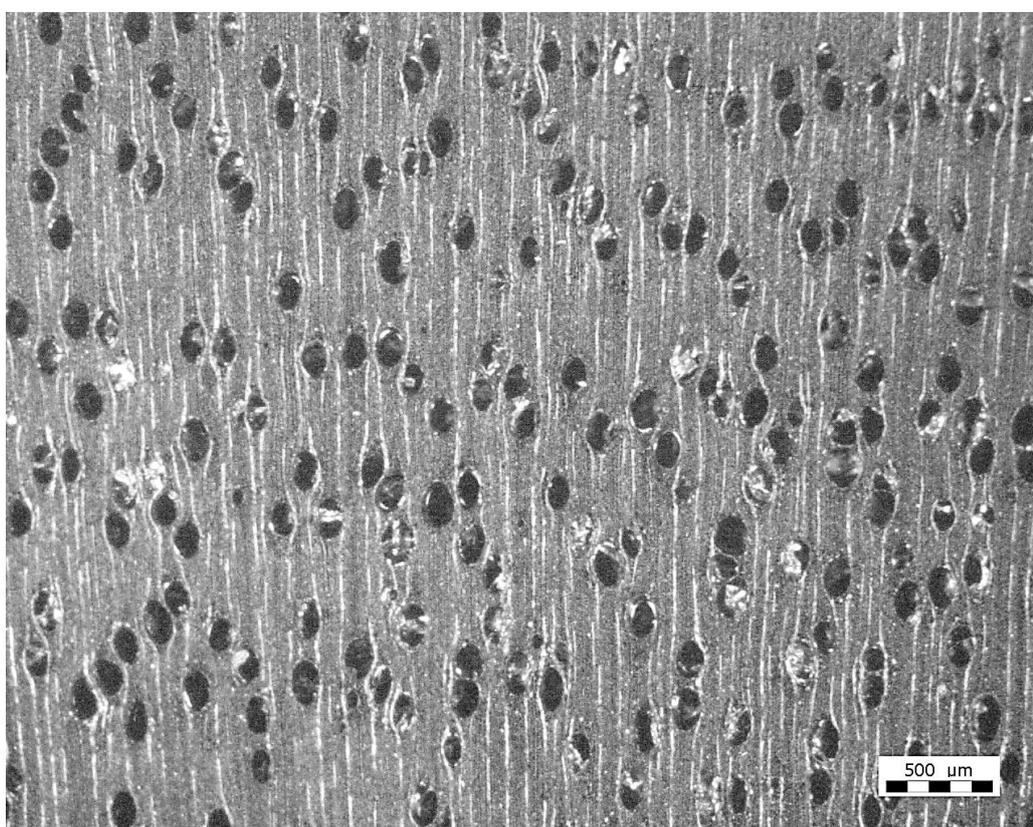


Figura 21 - Imagem da seção transversal do carvão de *E. grandis* x *E. urophylla*.

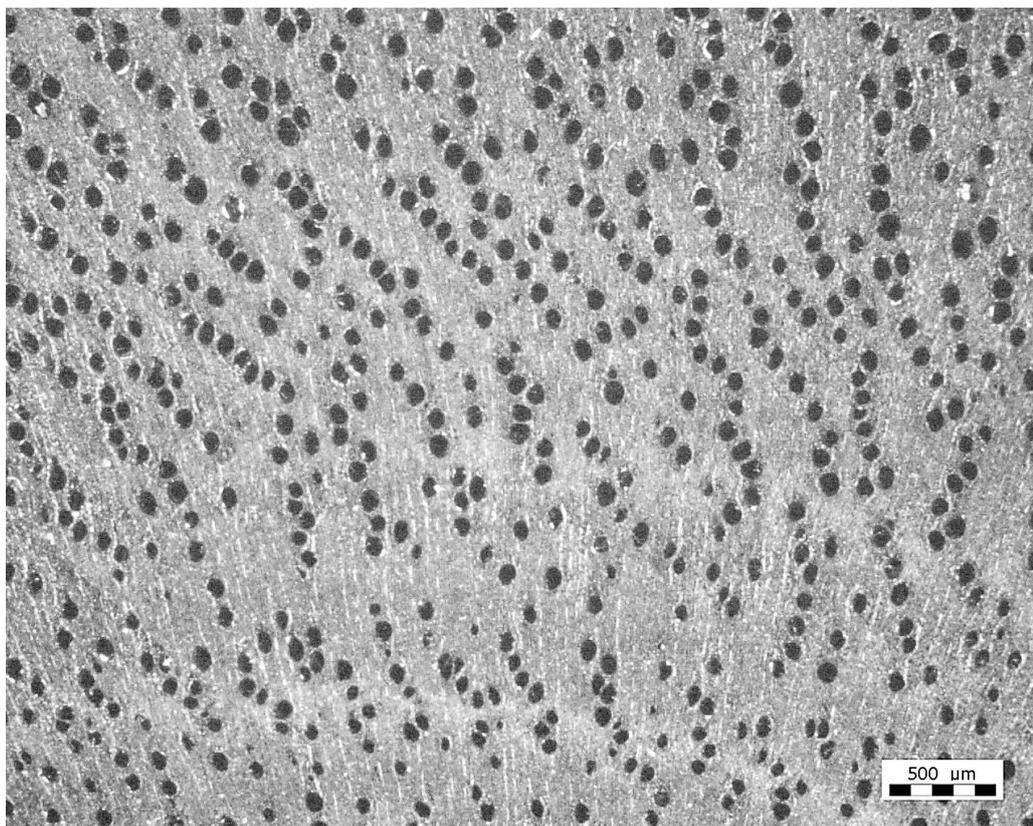


Figura 22 - Imagem da seção transversal do carvão de *E. urophylla* (GG 150).

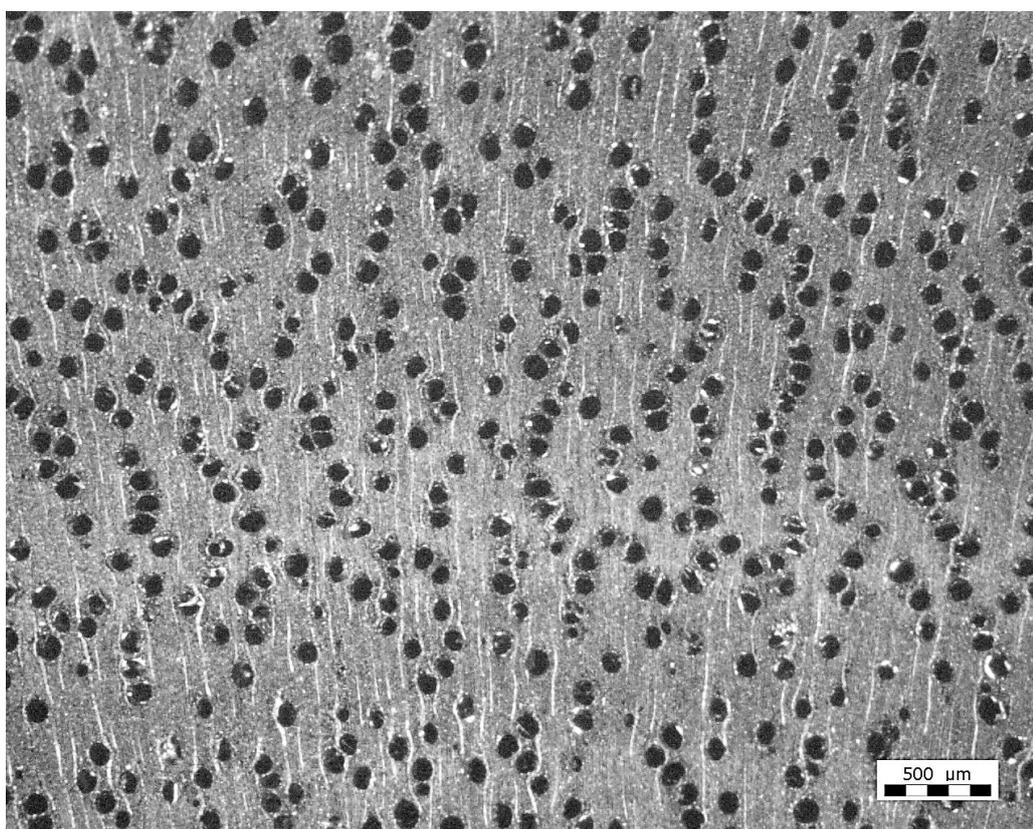


Figura 23 - Imagem da seção transversal do carvão de *E. urophylla* (GG 100).

Em geral, as características anatômicas observadas nos carvões da classe *eucalipto* foram porosidade difusa, arranjo de poros em diagonal, poros exclusivamente solitários e raios finos. Nos carvões das espécies desta classe, verificou-se parênquima axial indistinto.

No híbrido *E. camaldulensis* x *E. urophylla* a proximidade dos poros sugere a ocorrência de agrupamento geminado e múltiplo, embora sejam predominantes os poros solitários. A ocorrência de múltiplos geminados já foi observada em *E. camaldulensis* por Evangelista (2007), de forma que o agrupamento de poros observado em *E. camaldulensis* x *E. urophylla* possivelmente divergiu do padrão do gênero em função da hibridação.

Observa-se, nas imagens obtidas, que os elementos anatômicos que, visualmente, contribuem para a discriminação entre as duas classes de carvão são os raios e os poros. Assim, nos carvões da classe *eucalipto* verificou-se raios finos e pouco reflectivos e predominância de poros solitários em arranjo diagonal, porém, para os carvões da classe *nativa*, foram observados raios largos e poros com agrupamento múltiplo e arranjo radial.

A possibilidade de distinção entre as duas classes com base nos poros e nos raios do carvão apresenta-se como uma vantagem para a análise de imagens, pois estes elementos anatômicos possuem reflectâncias completamente distintas, segundo a metodologia utilizada. Conforme se observa na Figura 24, os poros e os raios constituíram os limites inferior e superior dos histogramas das imagens, respectivamente.

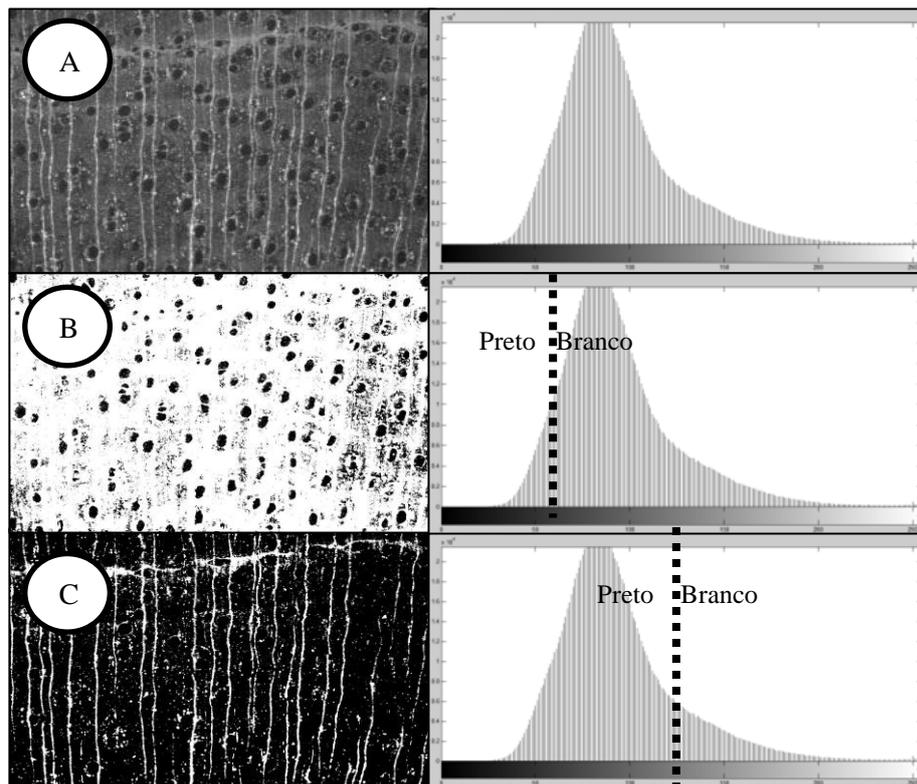


Figura 24 – Imagens do carvão de angico vermelho e seus histogramas: A - imagem original; B - imagem limiarizada usando um limite de 0,25; C - imagem limiarizada usando um limite de 0,5.

Nas Figuras 25 e 26 são apresentadas as imagens dos carvões de angico vermelho e *Eucalyptus camaldulensis*, respectivamente, em diferentes resoluções de imagem e profundidades de *pixels*.

Figura 25: Imagens do carvão de angico vermelho em diferentes resoluções de imagens (colunas) e profundidade de *pixels* (linhas).

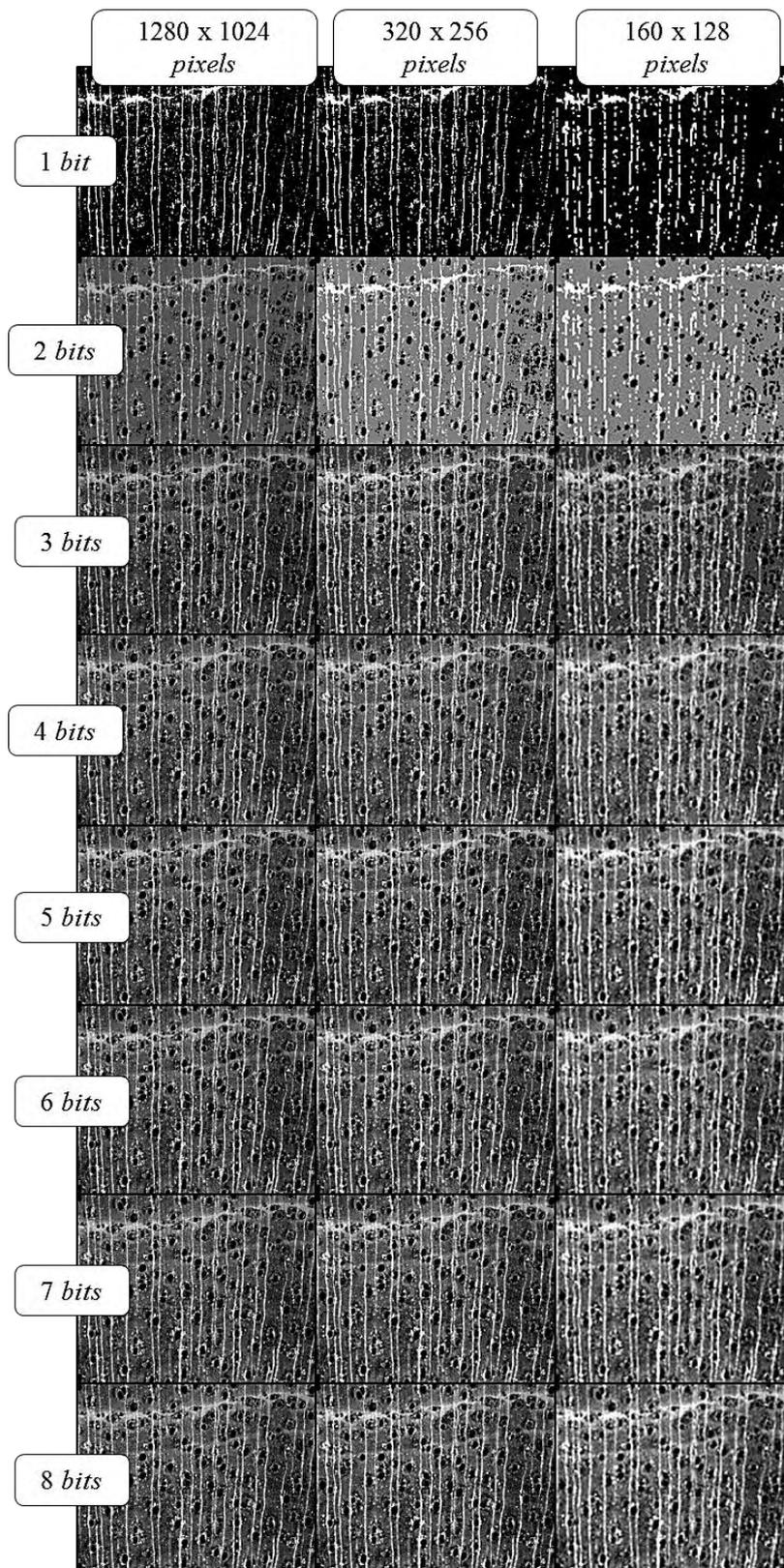
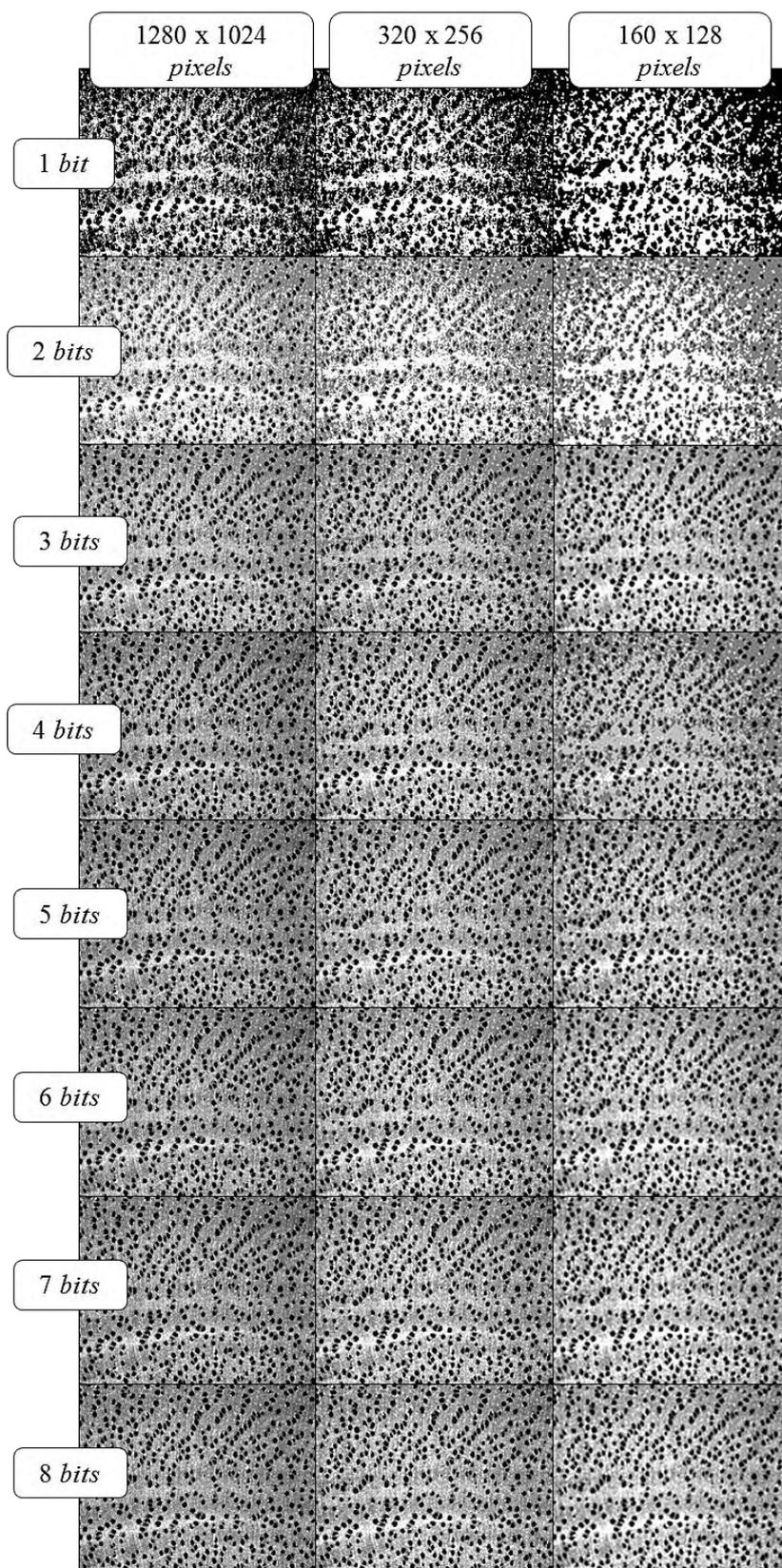


Figura 26: Imagens do carvão de *Eucalyptus camaldulensis* em diferentes resoluções de imagens (colunas) e profundidade de *pixels* (linhas).



Verifica-se, nas figuras 25 e 26, que a redução da resolução da imagem não impediu a observação de poros e raios. As diferenças entre imagens com 5 a 8bits foram pouco perceptíveis. Em imagens com 1 bits, ou seja, com 2 tons de cinza, os raios, quando altamente reflexivos, foram destacados, tal como em angico vermelho. Nas imagens com 2 e 3bits nota-se que os raios e poros foram formados por *pixels* de tons claros e escuros, respectivamente, sendo as fibras e o parênquima axial formados pelos *pixels* de tonalidades intermediárias.

5.2. Seleção de atributos

O número de atributos selecionados pelo procedimento *Forward Stepwise*, para a construção das funções de decisão de cada classificador, pode ser visualizada na Tabela 6. Dentre as 120 características de textura analisadas, foram selecionadas, em média, 29. Os classificadores que utilizaram a menor e a maior quantidade de características em seus modelos foram os de número 9 e 15 contendo, respectivamente, 10 e 42 atributos cada um.

Tabela 6 – Número de atributos selecionados para cada classificador.

Classificador	Resolução (<i>pixels</i>)	Profundidade de imagem (bits)	Nº de atributos
1	1280 x 1024	1	20
2		2	33
3		3	19
4		4	26
5		5	25
6		6	35
7		7	25
8		8	37
9	320 x 256	1	10
10		2	39
11		3	37
12		4	41
13		5	28
14		6	36
15		7	42
16		8	35
17	160 x 128	1	14
18		2	28
19		3	27
20		4	34
21		5	27
22		6	28
23		7	22
24		8	24

Os percentuais das características básicas das MCNCs selecionadas para a construção dos modelos dos classificadores estão apresentados na Tabela 7.

Tabela 7 – Percentual das características básicas das MCNCs selecionadas para a construção dos modelos dos classificadores.

Classificadores	1 a 8	9 a 16	17 a 24	Média
Resolução (<i>pixels</i>)	1280 x 1024	320 x 256	160 x 128	
Segundo momento angular	21,4%	19,8%	14,2%	18,6%
Variância	0,9%	4,1%	5,4%	3,5%
Entropia	14,5%	19,0%	15,2%	16,5%
Contraste	14,5%	18,7%	21,6%	18,2%
Correlação	24,1%	23,9%	26,0%	24,6%
Homogeneidade	24,5%	14,6%	17,6%	18,6%

Observa-se, na Tabela 7, que a variância apresentou a menor participação nos modelos desenvolvidos. Por outro lado, a correlação foi, em geral, a característica mais frequentemente selecionada.

Na Tabela 8 são apresentados os percentuais dos ângulos θ (graus) selecionados para a construção dos modelos dos classificadores.

Tabela 8 – Percentual dos ângulos θ (graus) selecionados para a construção dos modelos dos classificadores.

Classificadores	1 a 8	9 a 16	17 a 24	Média
Resolução (pixels)	1280 x 1024	320 x 256	160 x 128	
θ (graus)				
0°	35,9%	33,2%	32,4%	33,8%
45°	19,5%	20,1%	25,5%	21,5%
90°	22,7%	26,1%	24,0%	24,4%
135°	21,8%	20,5%	18,1%	20,2%

Observa-se que a ocorrência de atributos baseados em um ângulo de 0° foi numericamente superior aos demais em todas as resoluções de imagem. Isto possivelmente está relacionado com a importância da largura e frequência dos raios e dos poros para a discriminação das duas classes de carvões.

Na Tabela 9 são apresentados os percentuais dos valores da distância d (pixels) selecionados para a construção dos modelos dos classificadores.

Tabela 9 – Percentual dos valores da distância d (pixels) selecionados para a construção dos modelos dos classificadores.

Classificadores	1 a 8	9 a 16	17 a 24	Média
Resolução (pixels)	1280 x 1024	320 x 256	160 x 128	
d (pixels)				
1	20,5%	22,8%	23,0%	22,1%
2	14,5%	20,9%	26,5%	20,5%
3	15,5%	13,1%	20,6%	16,0%
4	20,5%	17,9%	13,7%	17,5%
5	29,1%	25,4%	16,2%	23,8%

Em geral, nota-se que a participação de atributos com valores de d acima de 3 foi maior quando foram utilizadas imagens com maiores resoluções (*pixels*), e vice-versa, conforme pode ser observado na Figura 25. Uma possível explicação reside no fato de que, quanto maior a resolução da imagem, menor a probabilidade de que dois pixels vizinhos representem elementos anatômicos distintos. Logo, maiores resoluções exigiriam um valor de d maior para que a análise da coocorrência abrangesse pixels de células distintas, melhorando, portanto, o potencial discriminatório do classificador uma vez que a interação entre diferentes elementos anatômicos apresenta grande relevância para a identificação anatômica.

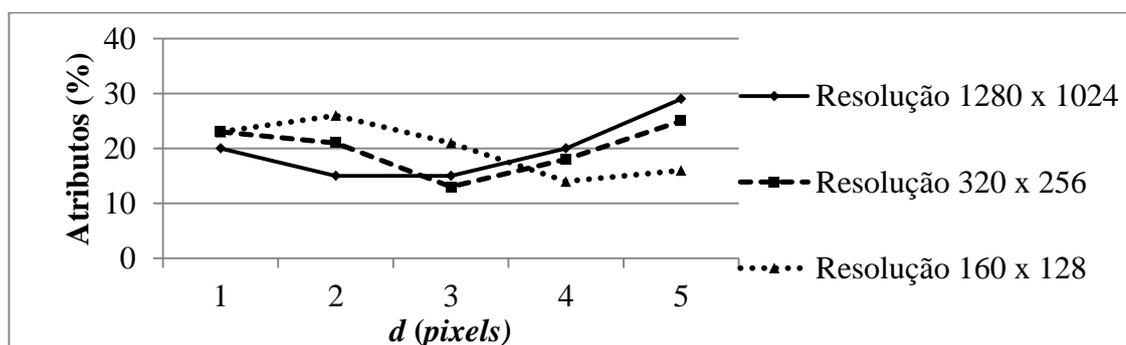


Figura 27 – Porcentagem de atributos selecionados para diferentes resoluções e valores de distância d (*pixels*).

5.3. Avaliação dos classificadores

Na Tabela 10 são apresentados, por classificador, a Exatidão Global e as Acurácias do Produtor e do Usuário para a classe *nativa* e *eucalipto*.

Tabela 10 – Exatidão global e acurácias por classe de carvão (*eucalipto* e *nativa*).

Classificador	Exatidão Global (%)	Acurácia (%)			
		Do produtor		Do usuário	
		Classe <i>nativa</i>	Classe <i>eucalipto</i>	Classe <i>nativa</i>	Classe <i>eucalipto</i>
1	98,7	99,0	98,0	99,0	98,0
2	99,2	99,3	99,0	99,5	98,7
3	98,8	99,2	98,0	99,0	98,3
4	99,6	99,3	100,0	100,0	98,7
5	99,9	99,8	100,0	100,0	99,7
6	100,0	100,0	100,0	100,0	100,0
7	99,9	99,8	100,0	100,0	99,7
8	100,0	100,0	100,0	100,0	100,0
9	87,3	83,0	96,0	97,7	73,9
10	98,6	98,2	99,3	99,7	96,4
11	99,7	99,7	99,7	99,8	99,3
12	99,9	100,0	99,7	99,8	100,0
13	99,7	99,8	99,3	99,7	99,7
14	99,9	100,0	99,7	99,8	100,0
15	99,8	99,8	99,7	99,8	99,7
16	99,9	100,0	99,7	99,8	100,0
17	91,2	91,3	91,0	95,3	84,0
18	97,6	97,0	98,7	99,3	94,3
19	98,8	99,3	97,7	98,8	98,7
20	98,9	99,2	98,3	99,2	98,3
21	99,1	99,2	99,0	99,5	98,3
22	98,9	99,3	98,0	99,0	98,7
23	98,9	99,2	98,3	99,2	98,3
24	98,8	99,3	97,7	98,8	98,7
Média	98,5	98,4	98,6	99,3	97,1

Conforme é observado na Tabela 10, em média, os classificadores apresentaram uma Exatidão Global de 98,45%, mostrando que, em geral, apenas 1,55% das tentativas de distinguir o carvão de *nativa* do carvão de *eucalipto* resultaram em erro. Esse desempenho é compatível com os 95% obtido por Khalid et al. (2008) ao avaliar um classificador desenvolvido para discriminar 20 espécies lenhosas através de 20 características de textura extraídas de imagens do lenho *in natura*.

Com relação às acurácias do produtor, em média, foram classificadas de forma correta 98,37 e 98,61% das imagens das classes *nativa* e *eucalipto*, respectivamente. Logo, a chance dos classificadores errarem foi, em geral, aproximadamente a mesma para as duas classes. Por outro lado, com relação às acurácias do usuário, 99,28% das

imagens classificadas como *nativa* estavam corretas, contra 97,13% das imagens classificadas como *eucalipto*.

A participação percentual das espécies no erro médio, considerando todos os classificadores, é apresentada na Figura 28, para a classe *nativa*, e na Figura 29, para a classe *eucalipto*.

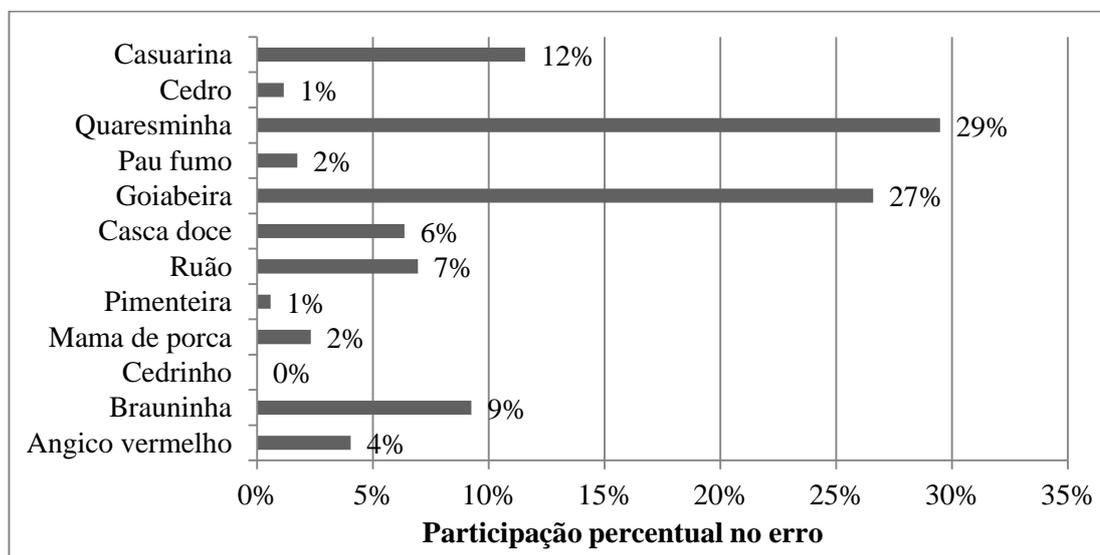


Figura 28 – Participação percentual das espécies da classe *nativa* no erro dos classificadores.

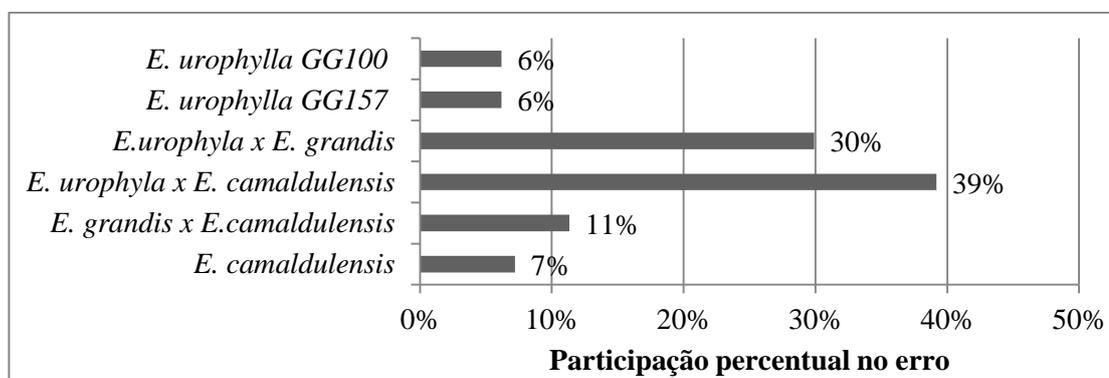


Figura 29 – Participação percentual das espécies/híbridos da classe *eucalipto* no erro dos classificadores.

Observa-se, na Figura 28, que o Cedrinho foi a única espécie da classe *nativa* cujas imagens do carvão foram classificadas corretamente em todos os casos. Por outro lado, as imagens dos carvões de Quaresminha e Goiabeira somaram mais da metade das imagens da classe *nativa* erroneamente classificadas como *eucalipto*, totalizando 56%

de participação no erro. A presença de raios finos e pouco reflectivos nessas duas espécies, característica comum da classe *eucalipto*, provavelmente contribuiu com este resultado.

Conforme pode ser visualizado na Figura 29, dentre as imagens da classe *eucalipto* classificadas incorretamente, 69% correspondem aos híbridos de *E. urophylla* x *E. camaldulensis* e *E. urophylla* x *E. grandis*. Estes híbridos apresentaram os raios mais reflectivos da classe *eucalipto*, fato que pode ter contribuído para dificultar a distinção com a classe *nativa*.

Na Tabela 11 são apresentados os classificadores e seus respectivos índices *kappas*.

Tabela 11 – Estimativas dos coeficientes *Kappa* por classificador.

Classificador	Nº de atributos	Resolução	Profundidade de imagem (bits)	Kappa	
1	20	1280 x 1024 <i>pixels</i>	1	0,970*	f g
2	33		2	0,983*	c d e f
3	19		3	0,972*	e f g
4	26		4	0,990*	b c d e
5	25		5	0,998*	a b
6	35		6	1*	a
7	25		7	0,998*	a b
8	37		8	1*	a
9	10	320 x 256 <i>pixels</i>	1	0,735*	
10	39		2	0,968*	f g
11	37		3	0,993*	a b c d
12	41		4	0,997*	a b
13	28		5	0,992*	a b c d
14	36		6	0,997*	a b
15	42		7	0,995*	a b c
16	35		8	0,997*	a b
17	14	160 x 128 <i>pixels</i>	1	0,807*	
18	28		2	0,946*	g
19	27		3	0,972*	e f g
20	34		4	0,975*	d e f
21	27		5	0,980*	c d e f
22	28		6	0,975*	d e f
23	22		7	0,975*	d e f
24	24		8	0,972*	e f g

*: estatisticamente diferente de zero, pelo Teste Z, ao nível de 5% de probabilidade.

Obs: índices seguidos pela mesma letra, na coluna, não diferem estatisticamente entre si, pelo Teste Z, ao nível de 5% de probabilidade.

Observa-se, na Tabela 11, que, ao nível de 5% de probabilidade, todos os classificadores apresentaram desempenho significativamente superior ao de uma classificação ao acaso.

Verifica-se, de forma geral, que os piores resultados couberam aos classificadores que utilizaram o menor número de características, tais como o número 9, com dez características, e o número 17, com 14. Contudo, os melhores resultados não foram necessariamente acompanhados por um alto número de atributos selecionados, como mostram os classificadores 5 e 7, ambos com 25 atributos e estatisticamente superiores aos classificadores 10 e 20, com 39 e 34 atributos, respectivamente.

Ao todo, dois classificadores apresentaram uma classificação perfeita ($Kappa = 1$) e outros oito classificadores apresentaram índices $Kappas$ numericamente inferiores, mas estatisticamente iguais a 1.

Considerando os classificadores que usaram imagens com resolução de 1280 x 1024 *pixels*, destacou-se o classificador 5, por apresentar um desempenho superior, juntamente com outros três classificadores (6 a 8), e por utilizar imagens com a menor profundidade de imagem, 5 *bits*, contribuindo para um menor custo computacional.

Considerando os classificadores que usaram imagens com resolução de 320 x 256 *pixels*, destacou-se o classificador 11, por apresentar um desempenho superior, juntamente com outros cinco classificadores (12 a 16), e por utilizar imagens com a menor profundidade de imagem, 3 *bits*.

Considerando os classificadores que usaram imagens com resolução de 160 x 128 *pixels*, destacou-se o classificador 21, por apresentar um resultado estatisticamente superior aos demais utilizando imagens com 5 *bits*.

Entre estes destaques, o classificador 21 apresentou resultado significativamente inferior aos demais. Entre os classificadores 5 e 11, o primeiro utilizou um menor número de atributos, contudo, o classificador 11 apresentou uma maior economia de gasto computacional para gerar as MCNCs, devido ao menor número de *bits* e de resolução de imagem, tendo, portanto, melhor desempenho.

As características que foram selecionadas para compor o modelo do classificador 11 são apresentadas na Tabela 12.

Tabela 12: Características utilizadas no modelo do classificador 11.

Característica básica da MCNC	Parâmetros da MCNC	
	Distância d (<i>pixels</i>)	Ângulo θ (graus)
Contraste	1	0
Homogeneidade	1	0
Contraste	1	45
Variância	1	90
Contraste	1	90
Segundo Momento Angular	1	135
Contraste	1	135
Homogeneidade	1	135
Segundo Momento Angular	2	0
Entropia	2	0
Homogeneidade	2	0
Segundo Momento Angular	2	45
Segundo Momento Angular	2	90
Correlação	2	90
Homogeneidade	2	90
Entropia	2	135
Contraste	2	135
Correlação	2	135
Entropia	3	45
Segundo Momento Angular	4	0
Segundo Momento Angular	4	0
Contraste	4	0
Segundo Momento Angular	4	45
Contraste	4	45
Entropia	4	135
Correlação	4	135
Homogeneidade	4	135
Entropia	5	0
Correlação	5	0
Segundo Momento Angular	5	45
Entropia	5	45
Contraste	5	45
Correlação	5	45
Variância	5	90
Contraste	5	90
Correlação	5	90
Correlação	5	135

6. Conclusões

A análise de imagens digitais se mostrou uma potencial e promissora ferramenta para auxílio na identificação da origem do carvão vegetal

Os 24 classificadores desenvolvidos acertaram, em média, 98,45% das 900 imagens testadas, permitindo distinguir satisfatoriamente carvões produzido com madeira de 6 espécies do gênero *Eucalyptus* daquele produzido com a madeira de 12 espécies nativas da Zona da Mata Mineira.

O melhor desempenho foi obtido pelo classificador 11, que analisou imagens com 320 x 256 *pixels* e 3 *bits* e apresentou um índice *Kappa* de 0,993 utilizando-se de 39 características de textura.

7. Recomendações

Trabalhos futuros deverão contemplar uma gama maior de espécies, abordar materiais carbonizados em diferentes temperaturas e analisar as variabilidades da madeira oriundas de diferentes idades e posições no tronco.

Futuramente, comprovada sua eficiência nestes novos estudos, a análise de textura poderá ser adaptada para condições de campo, visando semi-automatizar a classificação do carvão em postos de fiscalização.

8. Referencias Bibliográficas

ABRAF - ASSOCIAÇÃO BRASILEIRA DE PRODUTORES DE FLORESTAS PLANTADAS. **Anuário estatístico da ABRAF: ano base 2011**. Brasília, 2012. 145p. Disponível em: <<http://www.abraflor.org.br/estatisticas/ABRAF12/ABRAF12-BR.pdf>> Acesso em: 2 de outubro de 2012.

ABRAF - ASSOCIAÇÃO BRASILEIRA DE PRODUTORES DE FLORESTAS PLANTADAS. **Anuário estatístico da ABRAF: ano base 2010**. Brasília, 2011. 130p.

ALBUQUERQUE, A. R. **Anatomia comparada do lenho e do carvão aplicada na identificação de 75 espécies da floresta Amazônica, no estado do Pará, Brasil**. Piracicaba: ESALQ, 2012. 250p. Dissertação (Mestrado em Ciências) Escola Superior de Agricultura “Luiz de Queiroz”.

ALFONSO, V. A. **Caracterização anatômica do lenho e da casca das principais espécies de *Eucalyptus L'Hérit* cultivados no Brasil**. São Paulo, USP: 1987. 188f. Tese (Doutorado em Biociências) Universidade de São Paulo.

ALZATE, S. B. A. **Caracterização da madeira de árvores de clones de *Eucalyptus grandis*, *E. saligna* e *E. grandis x urophylla***. Piracicaba: ESALQ, 2004. 150p. Tese (Doutorado em Recursos Florestais) Escola Superior de Agricultura “Luiz de Queiroz”.

AMS - ASSOCIAÇÃO MINEIRA DE SILVICULTURA. **Anuário Estatístico 2010**. Belo Horizonte, 2011.

ANGELES, G. New techniques for the anatomical study of charcoalfied wood. **IAWA Journal**, v.22, p.245–254. 2001.

BARALDI, A.; PARMIGGIANI, F. An investigation of the textural characteristics associated with gray level coocurrence matrix statistical parameters. **IEEE Transactions on Geoscience and Remote Sensing**, n. 33 (2), p.293-304. 1995.

BRISOLA, S. H.; DEMARCO, D. Análise anatômica do caule de *Eucalyptus grandis*, *E. urophylla* e *E. grandis x urophylla*: desenvolvimento da madeira e sua importância para a indústria. *Revista Scientia Forestalis*, Piracicaba, v. 39, n. 91, p.317-330, set. 2011.

BURGUER, L. M.; RICHTER, H. G. **Anatomia da Madeira**. São Paulo: Nobel, 1991. 154p.

CASTRO, A. F. M. N. **Efeito da idade e de materiais genéticos de *Eucalyptus* sp. Na madeira e carvão vegetal.** Viçosa: UFV, 2011. 98p. Dissertação (Mestrado em Ciência Florestal) Universidade Federal de Viçosa.

CONGALTON, R. G. A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. **Remote Sensing of Environment.** n.37. p.35-46. 1991.

CONGALTON, R.G.; MEAD, R.A. A quantitative method to test for consistency and correctness in photointerpretation. **Photogrammetric engineering & remote sensing,** Bethesda, v.49 n.1, p.69-74, 1983.

EVANGELISTA, W. V. Caracterização da madeira de clones de *Eucalyptus camaldulensis* Dehnh. *E. urophylla* S.T. Blake, oriunda de consórcio agrossilvipastoril. Viçosa: UFV, 2007. 141p. Dissertação (Mestrado em Ciência Florestal) Universidade Federal de Viçosa.

EVANGELISTA, W. V. Penetração e retenção de arseniato de cobre cromatado em madeira de eucalipto. Viçosa: UFV, 2011. 126p. Tese (Doutorado em Ciência Florestal) Universidade Federal de Viçosa.

FIGUEIRAL, I. Lignified and charcoalified fossil wood. Em: JONES, T.P. & ROWE, N.P. (eds), **Fossil plants and spores: modern techniques:** p.92–96. The Geological Society, London. 1999.

FUKUNAGA, K.; HAYES R. R. Effects of Sample Size in Classifier Design. **IEEE Transaction on Pattern Analysis and Machine Intelligence,** v.11, n.8, p.873-885. 1989.

GONÇALVES, T. A. P. **Anatomia do lenho e do carvão de espécies arbóreas do cerrado no estado de São Paulo, Brasil.** Botucatu, UNESP: 2010. 113p. Dissertação (Mestrado em Ciência Florestal) Universidade Estadual Paulista “Júlio de Mesquita Filho”.

GONZALEZ, R. C., WOODS, R.E., EDDINS, S.L. **Digital image processing using matlab.** Upper Saddle River, Nova Jersey: Pearson Education, 2004. 344p.

HARALICK, R. M. Statistical and Structural Approaches to Texture. **Proceedings of the IEEE,** v.67, n.5, p.786- 804. 1979.

HARALICK, R. M.; SHANMUGAM, K.; DINSTEN, I. Textural Features For image classification. **IEEE Transactions on Systems, Man, and Cybernetics.** v.smc-3, n.6, p.610-621. 1973.

HERMANSON, J. C.; WIEDENHOEFT, A. C. A brief review of machine vision in the context of automated wood identification systems. **IAWA Journal**, v.32, n.2, p.233-250. 2011.

HILL, T.; LEWICKI, P. **Statistics: methods and applications**. Tulsa, OK: StatSoft, 2007. 719p.

HOADLEY, R. B. **Identifying wood: accurate results with simple tools**. Newtown, Connecticut, USA: The Taunton Press, 1933. 215p.

HUDSON, W.D.; RAMM, C.W. Correct formulation of the kappa coefficient of agreement. **Photogrammetric engineering & remote sensing**, Bethesda, v.53 n.4, p.421-422, 1987.

IAWA Committee. List of microscopic features for hardwood identification. **IAWA Bulletin**, Leuven, v. 10, n. 2, p. 219-332. 1989.

JAIN, A. K.; CHANDRASEKARAN, B. Dimensionality and sample size considerations in pattern recognition practice. In: KRISHNAIAH, P. R.; KANAL, L. N. **Handbook of Statistics 2: Classification, Pattern Recognition and Reduction of Dimensionality**. Amsterdam: North-Holland. v.2, p.835-855. 1982.

JAIN, A. K.; DUIN, R. P. W.; MAO, J. Statistical Pattern Recognition: A Review. **IEEE Transactions on Pattern Analysis and Machine Intelligence**. v. 22, n.1, p.4-37 2000.

KHALID, M.; YUSOF, R.; LIEW, E.; NADARAJ, M. Design of an intelligent Wood species recognitions system. **International Journal of Simulation System, Science and Technology**, v. 9, n. 3, p.9-19. 2008.

KWON, S-M.; KIM, N-H.; CHA, D.-S. An investigation on the transition characteristics of the wood cell walls during carbonization. **Wood Science and Technology**, Heidelberg, v. 43, p. 487-498, 2009.

MME/EPE – MINISTÉRIO DE MINAS E ENERGIA / EMPRESA DE PESQUISA ENERGÉTICA. **Balanco Energético Nacional – Ano base 2011**. Rio de Janeiro, 2012.

MUÑIZ, G. I. B. de; NISGOSKI, S.; HARDOSIN, F. Z. S.; FRANÇA, R. F. Anatomia do carvão de espécies florestais. **Cerne**, Lavras, v.18, n.3, p. 471-477. 2012.

NASCIMENTO, J. P. R. do. **Análise e classificação de imagens baseadas em características de textura utilizando matrizes de coocorrência**. Curitiba, UFPR: 2003. 116p. Dissertação (Mestrado em Informática) Universidade Federal do Paraná.

OLIVEIRA, J. T. da S.; HELLMEISTER J. C. Caracterização da madeira de Eucalipto para a Construção Civil. São Paulo, **Boletim Técnico da Escola Politécnica da USP**. 52p. 1998.

PEDRINI, H.; SCHWARTZ, W. R. **Análise de imagens digitais: princípios, algoritmos e aplicações**. São Paulo: Thomson Learning, 2008. 508p.

PEREIRA, B. L. C. **Qualidade da madeira de *Eucalyptus* para a produção de carvão vegetal**. Viçosa: UFV, 2012. 102p. Dissertação (Mestrado em Ciência Florestal) Universidade Federal de Viçosa.

PRIOR, J.; ALVIN, K. L. Structural changes on charring woods of *Dichrostachys* and *Salix* from southern Africa. **IAWA Bulletin**, Leuven, v. 4, n. 4, p. 197-206, 1983.

PRIOR, J.; GASSON, P. Anatomical changes on six African hardwoods. **IAWA Journal**, Leuven, v. 14, n. 1, p. 77-86, 1993.

RAO, C. R. An asymptotic expansion of the distribution of Wilks' criterion. **Bulletin of the International Statistical Institute**, 33, 177-181. 1951.

RAUDYS, S. J.; JAIN, A. K.. Small sample size effects in statistical pattern recognition: recommendations for practitioners. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v.13, n.3, p.252-264. 1991.

REED, T. R.; HANS DU BUF, J. M. A review of recent texture segmentation and feature extraction. **CVGIP: Image Understanding**. v. 57, n. 3, p. 359-372. 1993.

REIS, C. F.; ALBUQUERQUE, M. P.; CASTRO, S. B. Introdução ao reconhecimento de padrões utilizando redes neurais. Rio de Janeiro: **Centro Brasileiro de Pesquisas Físicas** – CBPF. 17p. 2001. Disponível em: <<http://funk.on.br/esantos/doutorado/INTELIG%CANCIA%20ARTIFICIAL/T%C9CNICAS/REDES%20NEURAS/RN%20UFSC/8.pdf>> Acesso em: 7 de novembro de 2012.

SANTOS, R. C. dos. **Parâmetros de qualidade da madeira e do carvão vegetal de clones de eucalipto**. Viçosa: UFV, 2010. 173p. Tese (Doutorado em Ciência Florestal) Universidade Federal de Viçosa.

SCHEEL, R.; GASPAR, M.D.; YBERT, J.-P.; Antracologia, uma nova fonte de informações para a Arqueologia Brasileira. **Revista do Museu de Arqueologia e Etimologia**, São Paulo, v.6: p.3-9, 1996a.

SCHEEL, R.; GASPAR, M. D. ; YBERT, J.-P. A anatomia dos carvões pré-históricos. Arqueologia encontra respostas em restos de fogueiras e incêndios florestais. **Ciência Hoje**, Rio de Janeiro, v. 21, n. 122, p. 66-69, 1996b.

SCHEEL-YBERT, R. Teoria e Métodos em Antracologia. 2 - Técnicas de Campo e de Laboratório. **Arquivos do Museu Nacional**, Rio de Janeiro, v.62, n.4, p.343-356. 2004.

SILVA, J. de C. Caracterização da madeira de *Eucalyptus grandis* Hill ex. Maiden, de diferentes idades, visando a sua utilização na indústria moveleira. Curitiba: UFP, 2002. 181p. Tese (Doutorado em Ciência Florestal) Universidade Federal do Paraná.

SINDICATO DA INDÚSTRIA MINERAL DO ESTADO DE MINAS GERAIS - SINDIEXTRA. **Ferro Gusa**. Belo Horizonte, 2012. Disponível em: <http://www.sindiextra.org.br/arquivos/2012_02_14_00_50_10_Ferro_Gusa.pdf> Acesso em: 15 de outubro de 2012.

TAMURA, H.; MORI, S.; YAMAWAKI, T. Textural Features Corresponding to Visual Perception. **IEEE Transactions on Systems, Man, and Cybernetics**, v.smc-8, n.6. p. 460-473. 1978.

TOMAZELLO FILHO, M. Estrutura anatômica da madeira de oito espécies de eucalipto cultivadas no Brasil. **IPEF**, Piracicaba, n.29, p.25-36. 1985.

VALE, A. T.; GENTIL, L. V. Produção e uso energético de biomassa e resíduos agroflorestais. Em: OLIVEIRA, J. T. S; FIEDLER, N. C.; NOGUEIRA, M. (Ed.). **Tecnologias aplicadas ao setor madeireiro III**. Jerônimo Monteiro - ES: 2008. p. 196-246.

VITAL, B. R.; CARNEIRO, A. C. O. Como facilitar a identificação da origem do carvão: floresta nativa ou exótica. **Jornal SIF**, Viçosa, n. 85, pg.6, 2007.

WATANABE, S. **Pattern recognition: humam and mechanical**. New York: Wiley, 1985. 570p.

WEBB, A. R. **Statistical Pattern Recognition** (Second Edition). Sussex do Oeste, Inglaterra: John Wiley & Sons, 2002. 504p.

WHEELER, E. A. ; BAAS, P. Wood Identification -A Review. **IAWA Journal**. v.19, n.3, p.241-264. 1998.

ZENID, G. J.; CECCANTINI, G. C. T. **Identificação Macroscópica de Madeiras.** Laboratório de Madeira e Produtos Derivados, Centro de Tecnologia de Recursos Florestais, Instituto de Pesquisas Tecnológicas do Estado de São Paulo – IPT, São Paulo. 2007.

9. Glossário

Classificador bayesiano: classificador cuja regra de decisão é avaliada de acordo com o teorema de Bayes.

Índice Kappa: índice que avalia o desempenho de um classificador. É uma estimativa do acerto real menos o acerto por acaso.

Matriz de coocorrência de níveis de cinza: é uma matriz numérica que denota a probabilidade de ocorrência de um *pixel* valor i com um *pixel* de valor j sendo que o relacionamento espacial entre ambos é caracterizado por uma distância d e um ângulo θ .

Pixel: menor elemento constituinte de uma imagem digital.

Primitiva: é o menor elemento que constitui uma textura. A frequência e organização de diferentes primitivas da origem ao padrão de uma dada textura.

Profundidade de *pixel*: número de *bits* necessários para armazenar cada *pixel* da imagem digital. Em uma imagem com n *bits*, os *pixels* podem apresentar 2^n diferentes valores.

Resolução de imagem: quantidade de *pixels* contida em uma imagem digital.

Textura: é uma característica visual de uma superfície que contém informações importantes sobre sua estrutura.