



Universidade Católica Dom Bosco
Centro de Ciências Exatas e Tecnológicas
Curso de Bacharelado em Engenharia de Computação

**Extração de Atributos e Análise
Discriminante Aplicadas na Detecção de
Defeitos em Couro Bovino**

Willian Paraguassu Amorim

Orientador: Prof. Dr. Hemerson Pistori

Relatório Final apresentado à Comissão de Projetos de Graduação do Curso de Engenharia de Computação da Universidade Católica Dom Bosco como parte dos requisitos para a obtenção do grau de Bacharel em Engenharia de Computação.

UCDB - Campo Grande - MS - Novembro/2006

Agradecimentos

Agradeço a Deus por iluminar meu caminho durante esses 5 anos, por ter me dado muita paz, amor, saúde, e de ter me protegido para que eu pudesse estar agora escrevendo essas palavras. Agradeço a Nossa Senhora Aparecida minha protetora de todos dias. Agradeço ao meu pai, por todo apoio e carinho que tem me dado, não somente durante esses 5 anos mas também pelo meus 22 anos. Agradeço a minha mãe, por seu carinho, seu afeto e pela sua preocupação diária pelo meu bem estar. Agradeço toda minha família por sempre terem me apoiado por mais distante que eu estava entre eles, mas sempre acreditando em mim. Agradeço também a Day (amiga nota 1.000.000) valeu day pela amizade e pelo bombons na faixa, Kely (Parceira), Tony (Grande Primo), Elvis (grande amigo), Vivi (Sempre Amiga), Grazi (Sempre me apoiando e que nunca vou esquecer de nossas conversas sobre nossas dificuldades), Thayla e Thayna (minhas sobrinhas que eu tanto amo), Rosi e Luce (Minha grandes irmãs), Kleber, Jessy e Laura (Trio parada dura, sem esses meu curso não seria o mesmo), tio Adalzizo (foi quem me fez prestar o vestibular e sempre acreditou no meu esforço), Lia (sempre disponível para minhas lamentações, meus desabafos diários e sempre me ajudando), pessoal do GPEC (grupo melhor que este não existe), Naty (“Amorzinho”, uma mulher maravilhosa). Agradeço principalmente aos professores Dr. Hemerson Pistori, pelo apoio em meu trabalho, pela confiança que foi me dado, e por todos os conselhos que me fizeram crescer pessoalmente e profissionalmente e a professor Amaury Jr. por ter me ensinado grandes lições de vida. Obrigado a todos que sempre estavam do meu lado e obrigado por uma frase que sempre me acompanhou, “A imaginação é mais importante do que o conhecimento”.

Resumo

Este trabalho apresenta a aplicação da análise discriminante de Fisher para seleção de atributos baseados em textura. Serão apresentados os módulos de geração de amostras e pré-processamento de imagens, que foram desenvolvidos para ajudar na realização dos testes de classificação. Serão apresentados também os experimentos e resultados da técnica de redução de atributos por Fisher aplicadas em características extraídas a partir de matriz de co-ocorrência, mapas de interação e atributos de cores (HSB), durante a classificação de defeitos em couros Wet-Blue. Esse trabalho faz parte do projeto DTCOURO, cujo objetivo principal é criar um sistema automático de extração de parâmetros, a partir de imagens digitais, capaz de tornar mais eficiente e preciso o processo de classificação de peles e couros bovino.

Abstract

This work presents the application of Fisher discriminant analysis to attribute selection based on textures. It presents the modules of sample generation and image processing, that were developed to help the classification stage. Fisher's method was used in experiments to reduce attributes extracted from co-occurrence matrix, interaction maps, and color attributes (HSB), performed in images of Wet-Blue leather during classification tests. This work is part of DTCOURO project, whose main goal is to create an automatic system to extract numerical parameters from digital images, in order to turn the process of bovine skins and leathers classification more efficient and accurate.

Conteúdo

1	Introdução	7
2	Fundamentação Teórica	10
2.1	Processamento Digital de Imagem	10
2.1.1	Imagem	12
2.1.2	Percepção visual humana	12
2.1.3	Cor	13
2.2	Reconhecimento de Padrões	15
2.3	Textura	16
2.4	Extração de Atributos	17
2.4.1	Matriz de Co-ocorrência	17
2.4.2	Mapas de Interação	17
2.5	Seleção de Atributos	19
2.6	Conceitos de Álgebra	20
2.6.1	Matriz	20
2.6.2	Operações com matrizes	21
2.6.3	Determinante	22
2.6.4	Matriz Inversa	23
2.6.5	Matriz de variância e covariância	23
2.6.6	Autovalores e Autovetores	24
2.7	Couro Bovino	24
3	Análise Discriminante	27
3.1	Redução de Atributos	28
4	Trabalhos Correlatos	31
4.1	Peles e Couros Bovino	31
4.2	Análise Discriminante	32
5	Desenvolvimento	34
5.1	IMAGEJ	34

5.2	WEKA	35
5.3	SIGUS	36
5.4	Ambiente de Aquisição e Marcação	36
5.5	Módulo de Experimentos	40
5.6	Módulo de Detecção de defeitos e Marcação Visual	41
6	Experimentos e Análise de Resultados	44
6.1	Experimentos	45
6.2	Resultados e Análise	49
7	Considerações Finais	52
A	Resultados	53
B	Defeitos Semelhantes	55
	Referências Bibliográficas	56

Capítulo 1

Introdução

Apesar de possuir o maior rebanho bovino do planeta, o Brasil ainda não consegue assumir um papel de destaque na indústria de couros bovino. A maior parte do couro produzido no país é de baixa qualidade, sendo que a maioria desses couros é comercializada nos estágios iniciais de processamento, com baixo valor agregado.

A baixa qualidade do couro bovino brasileiro tem limitado o desenvolvimento do setor coureiro, afetando também a capacidade do país agregar valor ao produto internamente e de usufruir das oportunidades oferecidas pela comercialização do couro acabado, de melhor cotação no mercado [13].

Pesquisas realizadas indicam que no Brasil, apenas 8,5% do couro produzido é de primeira qualidade e o restante é de baixa qualidade. Cerca de 60% dos defeitos ocorrem na fase produtiva, quando os bovinos são expostos aos ectoparasitas e ao manejo inadequado como: uso de ferrão, cerca de arame farpado, marcação a ferro quente em local impróprio, entre outros [11] [15] [8]. Os demais defeitos são ocasionados durante a fase de transporte dos animais ao frigorífico. A Figura 1.1 ilustra graficamente o perfil do couro bovino brasileiro, comparado ao dos EUA, a partir de suas classes de classificação.

O grande fator limitante para a melhoria da qualidade do couro bovino, é a inexistência de sistemas de remuneração diferencial pela qualidade da matéria prima produzida [18]. Atualmente, o couro é remunerado através do sistema “bica corrida”, isto é, em média, o pecuarista recebe pelo couro cerca de 7 a 8% do valor da arroba do boi gordo, independente de sua qualidade [12].

A partir das dificuldades encontradas para o problema da qualidade do couro bovino, foi proposto um sistema de classificação de couro verde e harmonizado. Esse sistema foi estabelecido pelo Ministério da Agricultura, Pecuária e Abastecimento por meio da Instrução Normativa Nº 12, em 18 de dezembro de 2002 [19]. Com essa norma foi estabelecido o critério de

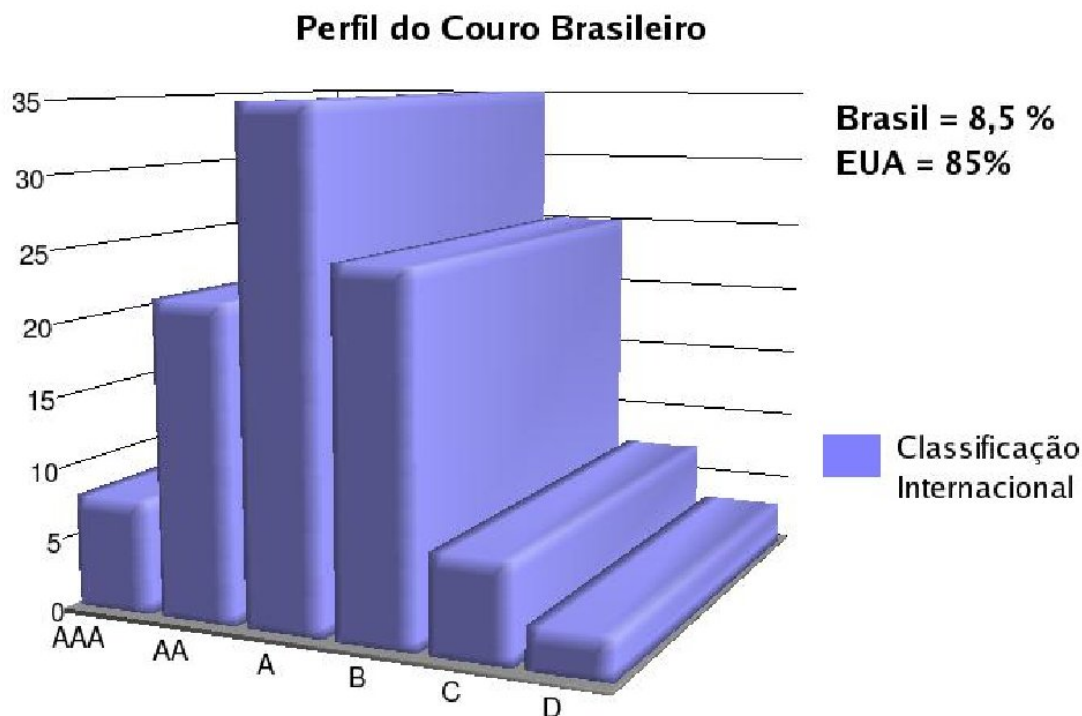


Figura 1.1: Perfil do couro bovino Brasileiro

classificação do couro bovino, através dos defeitos encontrados na pele do animal.

Para a viabilização de sistemas de remuneração, será preciso resolver alguns problemas operacionais importantes, como a detecção de defeitos e a classificação em diferentes estágios da cadeia produtiva. Para isso criou-se o projeto DTCOURO, um sistema em visão computacional que será capaz de realizar a detecção das diversas inconformidades apresentadas no couro, como: berne, placa de berne, carrapato, marca fogo, risco aberto, risco fechado e sarna. Através destas informações encontradas, realizar a classificação da peça do couro bovino. O objetivo principal do projeto DTCOURO é o desenvolvimento de um sistema automático de extração de parâmetros numéricos, a partir de imagens digitais, para que se torne mais eficiente e preciso o processo de classificação do couro bovino.

Este trabalho propõe a criação de um módulo computacional, utilizando análise discriminante, capaz de a partir de um conjunto de atributos, estabelecer combinações estatísticas entre eles, criando novos atributos e realizando dessa forma a seleção das melhores informações geradas. Uma vez que esse

processo seja realizado, será possível medir estatisticamente se a seleção de um conjunto de atributos, ou a geração de novos atributos, ou até mesmo a classificação utilizando análise discriminante de Fisher, aumentará o desempenho na classificação de peles e couros bovinos.

Para a realização desse trabalho, foram implementados 3 módulos auxiliares que ajudaram os testes de classificação, que são: o módulo de marcação, pré-processamento e geração de amostras, o módulo de extração e realização de experimentos e o módulo de marcação visual dos defeitos detectados automaticamente. A partir dos resultados dos módulos auxiliares utilizamos a análise discriminante de Fisher para a geração de novos atributos e seleção das melhores características extraídas. Através dessas informações geradas por Fisher, foi possível gerar os testes de desempenho utilizando alguns algoritmos de aprendizagem, mostrando resultados estatísticos do desempenho do classificador de peles e couros bovinos.

O texto desse trabalho está dividido em 6 partes. O capítulo 2, oferece uma breve abordagem dos principais conceitos sobre, processamento digital de imagens, reconhecimento de padrões, extração de atributos utilizando matriz de co-ocorrência e mapas de interação, seleção de atributos, conceitos de álgebra e o estudo do couro bovino. O capítulo 3, apresenta a definição e teoria da técnica de Fisher utilizada para redução de atributos. O capítulo 4, demonstra alguns trabalhos na área de análise do couro bovino e uso de análise discriminante de Fisher na seleção de atributos. O capítulo 5, demonstra todos os passos e ferramentas que foram utilizadas e integradas para a implementação dos módulos responsáveis pela realização dos testes. O capítulo 6, apresenta os experimentos e resultados obtidos a partir dos módulos desenvolvidos. Por fim, o capítulo 7 que é descrito as conclusões e sugestões para trabalhos futuros.

Capítulo 2

Fundamentação Teórica

2.1 Processamento Digital de Imagem

O processamento digital de imagens é uma área do conhecimento humano, que tem como objetivo a manipulação e análise de imagens por computador visando a extração de informações úteis para solução de determinados problemas. Os recursos disponibilizados pelo processamento de imagens são utilizados em várias atividades, entre as quais estão a Medicina, Meteorologia, Agricultura, Geologia, Cartografia, Robótica, entre outras.

Em outras palavras, o processamento digital de imagens entende-se como análise e manipulação de imagens por computador, com o objetivo de extrair as melhores informações ou transformá-las, de tal modo, que a informação seja mais facilmente discernível por um analista humano. As principais operações realizadas nas imagens para o seu processamento, podem ser classificada em 4 tipos básicos, que são [3]:

1. Modelagem e sua representação da imagem;
2. Realce de imagem;
3. Reconstituição de imagem;
4. Análise de imagem;

Modelagem e representação da imagem é a forma de representação de cada pixel¹ da imagem, e isso pode ser representado em termos de qualidade

¹Pixel é a abreviação de picture element(elemento da foto). É a menor parte de um imagem digitalizada, em que cada um destes pontos possui a informação que determina sua cor. Este conjunto de pontos é organizado na forma de matriz, formando assim uma imagem.

e quantidade. É importante que a representação da imagem seja capaz de armazenar todas as informações necessárias contidas em um objeto imageado. Realce de uma imagem tem como objetivo acentuar certos aspectos de uma imagem, a fim de melhorar sua análise e visualização. No processo de realce, nenhuma informação é acrescentada e sim enfatizada certas características. Restauração de imagem é o processo que visa minimizar ou remover certas degradações encontradas nas imagens. Essa degradação é normalmente decorrente da qualidade das câmeras que realizam a captura das imagens. As ferramentas que são utilizadas para a correção dessas degradações são denominadas filtros. Análise de imagens é o processo de medição quantitativa e qualitativa de um determinado aspecto da imagem, gerando sua descrição [3] [23]. A Figura 2.1 ilustra algumas dessas operações.

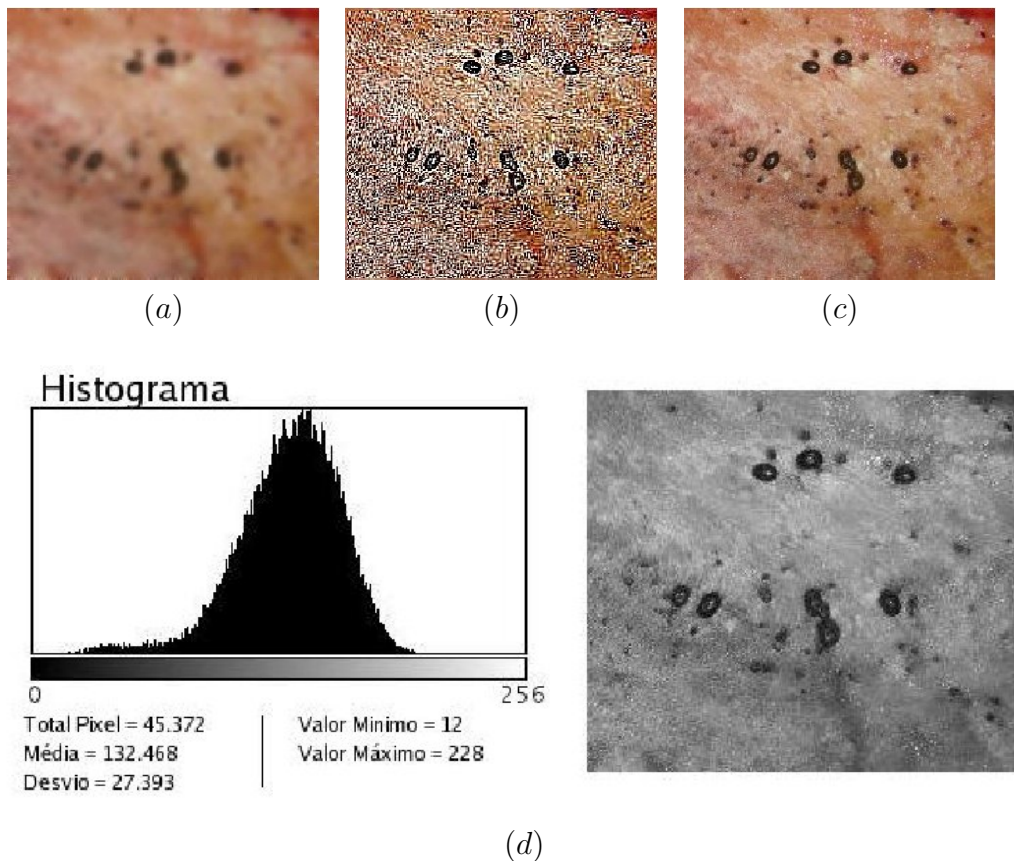


Figura 2.1: Tipos de operações para o processamento da imagem, (a) Modelagem e Representação, (b) Realce, (c) Reconstituição da imagem e (d) Análise da imagem

Para facilitar o estudo e o entendimento de processamento digital de imagens, será mostrada uma prévia abordagem de alguns conceitos que servirão como base para a compreensão dos tópicos aqui apresentados. A seguir são apresentados esses conceitos.

2.1.1 Imagem

Para a representação de uma imagem podemos utilizar uma função $f(x, y)$, bidimensional, em que os valores assumidos pela função quantificam a energia capturada por algum sistema de imageamento. Podemos dar como exemplo uma fotografia, em que a função $f(x, y)$ representa a intensidade luminosa referente ao ponto (x, y) do espaço considerado. Para que qualquer imagem possa ser manipulada pelo computador, é preciso que essa imagem seja discretizada ou digitalizada.

Uma imagem discreta é aquela em que uma função é definida em uma “grade” regular de pontos da forma $(m.d_x, n.d_y)$, em que d_x e d_y são os intervalos nas direções x e y , e (m, n) são valores inteiros nos intervalos $[0, m - 1]$ e $[0, n - 1]$. Em imagens de tons de cinza os valores que são assumidos pela função, são representados na forma de $k.d_z$, onde k está no intervalo $[0, k - 1]$ e d_z representa um dado valor, como mostra a Figura 2.2 [3] [23].

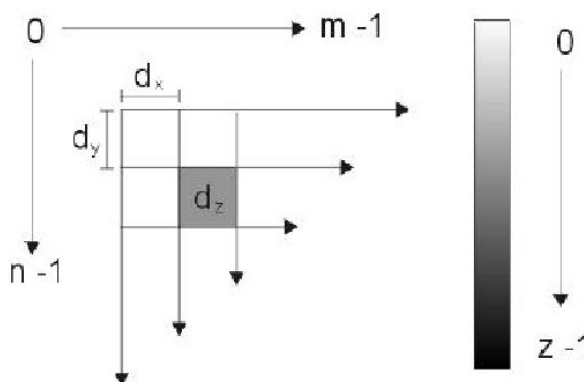


Figura 2.2: Representação discreta de uma imagem

2.1.2 Percepção visual humana

Sistemas para processamento digital de imagens devem considerar características da percepção visual humana. Muitas técnicas e algoritmos se baseiam em critérios elementares da estatística, como realce, restauração, entre

outros. Consideram também critérios de fidelidade de resultados, que dependem das características psicofísicas do sistema visual humano [3].

O sistema visual humano pode ser organizado em 3 níveis:

1. neuro-sensorial;
2. perceptivo;
3. cognitivo;

O nível neuro-sensorial envolve transformações dos traços elementares da estimulação visual em primitivas visuais. O nível perceptivo indica condições de aparecimento como, proximidade, similaridade, continuidade e conectividade. Os processos cognitivos se preocupam em montar graficamente a representação de um objeto.

2.1.3 Cor

Em tudo em que observamos a cor esta presente, tornando-se assim um elemento essencial em visão computacional². A cor transmite informações que dão realismo e naturalidade nas imagens, como condições de iluminação e a forma dos objetos visualizados. Isso significa que, para se ter uma representação fiel da cor, deverão ser empregues modelos de cores, corretos e precisos.

A cor pode ser formada por três tipos de processos, que são: processos aditivos, quando a ocorrência de dois ou mais raios luminosos de frequência diferentes; processo por subtração, quando a luz é transmitida de um filtro que absorve radiação luminosa de um determinado comprimento de onda, ou pelo processo por formação de pigmentação, que ocorre quando os pigmentos absorvem, refletem ou transmitem a radiação luminosa [10].

As cores podem ser divididas em três grupos distintos que são: cores primárias, cores secundárias e cores terciárias. Cores primárias e pigmentadas são formadas por meios das cores básicas, que são: vermelho, verde e azul (RGB). Para a criação das cores secundárias são utilizadas as combinações de duas cores primárias. A representação das cores terciárias são obtidos através da mistura de uma cor primária com uma ou mais cores secundárias.

²Visão Computacional é a área da ciência que se dedica a desenvolver teorias e métodos voltados à extração automática de informações úteis contidas em imagens

Modelos de cor

Diferentes maneiras podem ser utilizadas para a representação das cores de uma imagem no computador. São muitos os espaços de cores existentes, e entre eles podemos citar o RGB, CMY, YIQ, YUV, HSI, HSV, CIELAB, CI-ELUV, rgb , $c_1c_2c_3$, $l_1l_2l_3$, YCb Cr. Nesse trabalho apresentaremos somente duas dessas formas de representação de cores, que serão muito utilizadas em nossos experimentos, que são: RGB e HSB.

Modelo RGB

O modelo RGB é um modelo de cores com três cores primárias: vermelho, verde e azul. A sigla RGB deriva da junção das primeiras letras dos nomes das cores primárias em língua inglesa: Red, Green e Blue. Cada uma dessas três cores primárias tem um intervalo de valores de 0 até 255. Combinando os 256 possíveis valores de cada cor, o número total de cores é aproximadamente 16,7 milhões ($256 \times 256 \times 256$). A partir da combinação das intensidades dessas três cores é possível obter as demais cores. A representação gráfica do modelo RGB é formada por um cubo, como mostra a Figura 2.3.

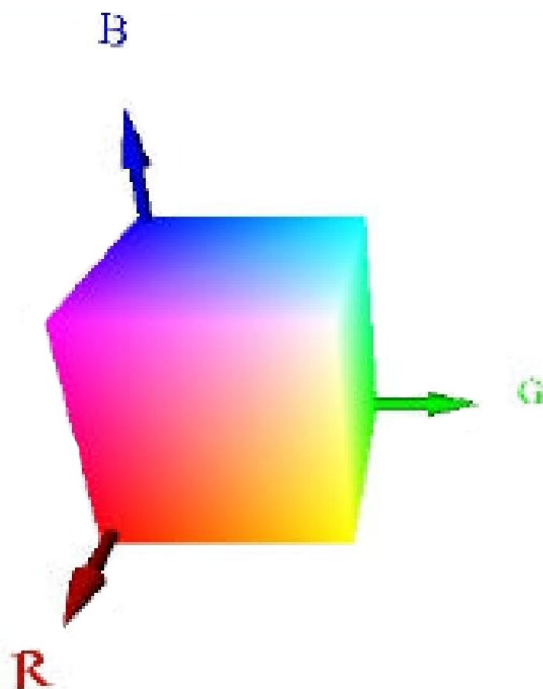


Figura 2.3: Representação gráfica do modelo RGB

Modelo HSB

O modelo de cores HSB se baseia na percepção humana das cores e não nos valores de RGB. Isso por que a mente humana não separa tão facilmente as cores em modelos de vermelho/verde/azul ou ciano/magenta/amarelo/preto. O olho humano vê cores como componentes de matiz, saturação e brilho. Em termos técnico matiz se baseia no comprimento da energia da onda de luz refletida de um objeto ou transmitida por ele. A Saturação é a quantidade de cinza de uma cor. Quanto mais alta a saturação, mais baixo é o conteúdo e mais intensa é a cor. O brilho é a medida de intensidade da luz em uma cor. Os valores médios de H(matiz), S(saturação) e B(brilho), serão capturadas das amostras que estarão no banco de imagens do projeto DTCOURO e utilizados como atributos de cores.

2.2 Reconhecimento de Padrões

Reconhecimento de Padrões é a ciência que estuda a forma de identificação ou classificação de um conjunto de informações. Esta área abrange desde a detecção de padrões até a escolha mais simples entre dois objetos [36]. O reconhecimento de padrões é de suma importância, devido às ocorrências na vida humana tomarem forma de padrões, como exemplo: formação da linguagem, modo de falar, desenho de figuras, entendimento das imagens, tudo se torna um padrão. O homem sempre avalia situações como padrões, a fim de descobrir relações existentes em seu meio.

A área de reconhecimento de padrões envolve inúmeros problemas de processamento de informações, indo desde reconhecimento de voz até detecção de erros em equipamentos ou inspeção visual para detecção de imperfeições em texturas. Na prática, algumas áreas podem ser citadas, sendo muito utilizadas em reconhecimento de padrões [37]:

1. Reconhecimento de Faces;
2. Identificação de impressões digitais;
3. Segmentação e pré-processamento de imagens;
4. Reconhecimento de Voz;
5. Detecção de defeitos;

O objetivo principal de reconhecimento de padrões é a distinção de diferentes tipos de padrões, tendo como capacidade a sua discriminação. A

maioria dos classificadores (estatísticos, sintáticos ou neurais), são baseadas em similaridade. Por exemplo, se quisermos classificar um padrão x com outro padrão y que pertence a classe c_1 , intuitivamente a tendência é classificar x como pertencente à classe c_1 [36].

2.3 Textura

Pela dificuldade do reconhecimento da textura em uma imagem, é difícil definir uma textura de forma precisa. A textura é um termo utilizado com conotação ligeiramente diferenciada em distintas áreas relacionadas com sistemas gráficos. A definição de uma textura está relacionada com a escala dos objetos visíveis numa imagem. Algumas definições de textura encontradas na literatura são:

1. variação local da luminância entre os pixels dentro de uma região pequena da imagem [9];
2. medidas das propriedades da superfície de um objeto [23];
3. conjunto de elementos mutuamente relacionados [34];

Textura é uma das características mais importantes para classificar e reconhecer padrões em uma imagem, e pode ser caracterizada por variações locais em valores de pixels que se repetem de maneira regular ou aleatória ao longo do objeto ou de uma imagem [30]. A Figura 2.4 ilustra alguns exemplos de textura.

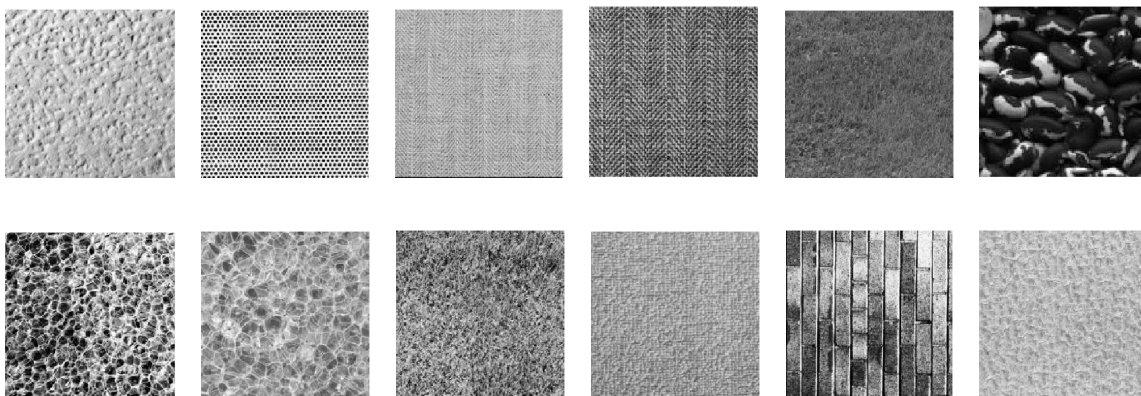


Figura 2.4: Exemplos de textura

2.4 Extração de Atributos

Extrair as características (atributos) mais importantes numa imagem evidencia as diferenças e similaridades entre os objetos. Algumas características são definidas por uma aparência visual na imagem, como: o brilho de uma determinada região, textura de uma região, amplitude do histograma, entre outros. O principal objetivo da extração de atributos é caracterizar os objetos para serem reconhecidos através de medidas sobre a imagem [24].

A seguir serão apresentadas duas técnicas de extração de atributos que foram implementadas e utilizadas para medir o desempenho do classificador do DTCOURO. Essas técnicas irão ajudar a análise discriminante de Fisher para realizar a geração de novos atributos e a seleção dos mais discriminantes.

2.4.1 Matriz de Co-ocorrência

A forma mais simples que pode ser definida em uma imagem digital em níveis de cinza é um pixel, que tem como propriedade principal seu nível de cinza. A matriz de co-ocorrência de níveis de cinza fornece relações espaciais entre os pixels de uma imagem, possibilitando a extração de atributos para a representação de suas características de textura, como exemplos: rugosidade, granulidade, aspereza, regularidade, direcionalidade, entre outras [20].

A matriz de co-ocorrência é calculada a partir dos índices das linhas e colunas que representam os diferentes valores de níveis de cinza. Dessa forma calcula-se a frequência que os mesmos ocorrem dois a dois em uma certa direção em uma certa distância [9]. As direções usualmente tratadas são as de 0° , 45° , 90° e 135° e as distâncias são escolhidas de acordo com a granularidade das imagens manipuladas [35]. Com esses parâmetros é calculada a matriz de co-ocorrência para cada direção e para cada distância.

Como exemplo, seja a imagem representada por uma matriz de pontos mostrada na Figura 2.5(a). Utilizando a direção 0° e a distância de 1 pixel, teremos a matriz de co-ocorrência mostrada na Figura 2.5(b). Como resultado a matriz sofre um processo de redução, sendo dividido cada elemento pelo número máximo de elementos em cada posição. A matriz $c[i, j]$ na linha i e coluna j , consiste no número de vezes em que o valor de i possui vizinhos mais próximos do valor j à esquerda ou direita da matriz $m[i, j]$ à uma distância de 1 pixel.

2.4.2 Mapas de Interação

Os mapas de interação são utilizados na análise de pares de pixel de uma imagem [16]. A utilização desse método leva em consideração as diferentes

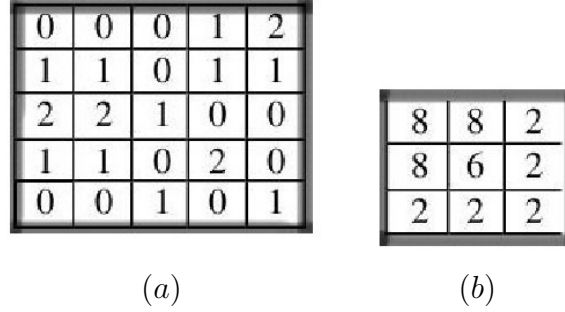


Figura 2.5: Exemplo utilizando matriz de co-ocorrência. (a) imagem $m[i, j]$ com níveis de cinza de 0,1 e 2. (b) matriz de co-ocorrência $c[i, j]$.

variações de textura encontradas no couro bovino. Como a matriz de co-ocorrência, o mapa de interação também trabalha com a intensidade em tons de cinza dos pixels de uma imagem. A Figura 2.6, ilustra como é aplicado essa técnica na análise de uma imagem.

No primeiro momento seleciona-se uma pequena amostra da imagem original como referência, para que a técnica crie uma nova imagem $img[m][n]$, com as intensidades originais de cada pixel. Em seguida, o método compara a variação de intensidade de um determinado pixel $img(i, j)$, com seus vizinhos, dado um deslocamento d e um ângulo a , criando uma nova matriz $map[d][a]$. Essa matriz chamada de mapa de interação, irá guardar apenas as informações da amostra capturada. A partir das informações do mapa polar, serão feitas comparações com as demais regiões da imagem original, ainda não analisada. Essas comparações permitirão indicar quais as similaridades e diferenças entre tais regiões sobre à amostra de referência [7] [6].

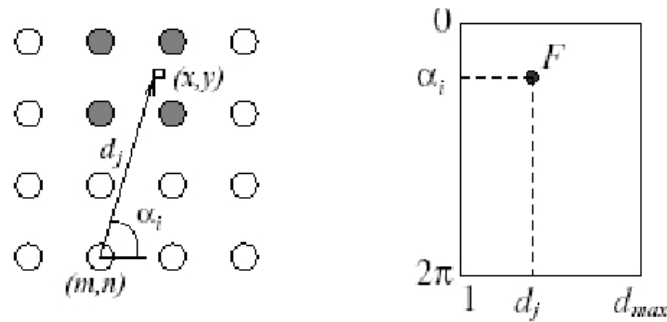


Figura 2.6: Aplicação do mapa de interação: seleção da amostra de imagem (à esquerda); criação do mapa polar (à direita).

A interação entre os pares de pixels é uma importante ferramenta da extração de características e análise de textura, no que diz respeito à detecção de defeitos. Os atributos que foram escolhidos para a extração sobre o mapa de interação e matriz de co-ocorrência foram os seguintes descritores: entropia, momento da diferença inversa, dissimilaridade, correlação, contraste, segundo momento angular e a diferença inversa.

Entropia em uma imagem é o nível de variação na intensidade entre os pixels da mesma. Imagens com pouca variação (ex.: floresta) apresentam um valor de entropia baixo e imagens com muita variação (ex.: couro) tem um valor maior de entropia. Momento da Diferença Inversa (MDI) é a medida da quantidade de homogeneidade local. Valores maiores de MDI apresentam maior similaridade entre os pixels do local. A dissimilaridade mede o grau de diferença nos níveis de cinza de pares de pixels. Quanto maior a diferença entre os pixels maior a dissimilaridade. Correlação é a medida das dependências lineares dos tons de cinza. Quanto maior a dependência maior a correlação. Contraste é a medida de variação local presente em uma imagem, geralmente utilizado para encontrar bordas na área de interesse. Segundo Momento angular é a medida de uniformidade de uma região da imagem. Essa uniformidade é calculada sobre a textura da imagem. O segundo momento angular tende a 1 quando a área é mais uniforme e tende a zero caso a área não seja uniforme [30].

2.5 Seleção de Atributos

A seleção de atributos é um problema de otimização que busca pelo menor subconjunto com a melhor eficiência no processo de classificação [24]. O método de seleção de atributos pode ser visto como um processo de busca onde um algoritmo usado deve encontrar os melhores atributos dentro de um conjunto. A dimensionabilidade do espaço de atributos pode resultar em dois tipos de problemas [33].

1. alto custo de processamento;
2. geração do fenômeno conhecido como maldição da dimensionabilidade.

A maldição da dimensionabilidade pode ser caracterizada como uma degradação nos resultados de classificação, com o aumento da dimensionabilidade dos dados [33]. Esse fenômeno ocorre quando o número de elementos de treinamento requeridos para um bom classificador, é uma função exponencial da dimensão do espaço de características.

O método de seleção de atributos pode ser dividido em 2 partes que são: métodos de busca e funções de avaliação, que é usada para medir a qualidade dos subconjuntos. Os métodos de busca podem ser considerados os melhores métodos, em que de modo exaustivo é gerado todas as combinações possíveis de atributos. O objetivo dos algoritmos de seleção de atributos é escolher o menor subconjunto que oferece o melhor desempenho na classificação em conformidade com custos computacionais razoáveis [5].

2.6 Conceitos de Álgebra

Métodos estatísticos multivariados são baseados em manipulação de matrizes. Uma ferramenta básica para o entendimento dessa área é a álgebra. A seguir serão apresentados alguns tópicos dessa área acompanhados de exemplos, que foram utilizados no entendimento da análise discriminante de Fisher.

2.6.1 Matriz

Matriz é considerada no modo algébrico como um arranjo bidimensional formado por elementos (i, j) , em que i é a representação da linha e j representação da coluna. Operações com matrizes podem ser utilizadas como uma forma compacta de realizar operações com vários números simultaneamente. Na ilustração 2.1 é mostrado um exemplo de matriz $M(i, j)$.

$$\mathbf{M} = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ m_{41} & m_{42} & m_{43} & m_{44} \end{pmatrix} \quad (2.1)$$

Uma matriz é representada na forma quadrada ou não. A forma quadrada é representada quando o número de linhas é igual ao número de colunas, e a forma não quadrada se baseia no número de linhas diferente do número de colunas. Em 2.2 é ilustrada uma matriz quadrada e 2.3 ilustra a matriz não quadrada.

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix} \quad (2.2)$$

$$\mathbf{Y} = \begin{pmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \end{pmatrix} \quad (2.3)$$

Uma matriz contendo somente uma linha é conhecida como vetor linha e quando contém somente uma coluna é chamada de vetor coluna. Em 2.4 e 2.5, são ilustrados esses dois tipos de situações.

$$\mathbf{X} = \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix} \quad (2.4)$$

$$\mathbf{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{pmatrix} \quad (2.5)$$

Em 2.6 é ilustrada a representação de uma matriz diagonal a partir da matriz quadrada, em que os elementos fora da diagonal principal são todos iguais a zero. Matriz identidade ou matriz unitária é uma matriz quadrada em que os elementos da diagonal principal são todos iguais a 1 e os demais igual a 0. A matriz 2.7 ilustra a matriz identidade.

$$\mathbf{X} = \begin{pmatrix} x_{11} & 0 & 0 \\ 0 & x_{22} & 0 \\ 0 & 0 & x_{33} \end{pmatrix} \quad (2.6)$$

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (2.7)$$

2.6.2 Operações com matrizes

Através da matriz é possível efetuar operações como: transposta, adição, subtração, multiplicação com escalar e multiplicação entre matrizes. A seguir serão explicadas cada forma de operação a partir de exemplos. A transposta é efetuada através da permuta das linhas e colunas da matriz e vice-versa. Em 2.8 é ilustrado um exemplo de transposta de uma matriz.

$$\text{Se } \mathbf{X} = \begin{pmatrix} 35 & 40 & 52 \\ 22 & 34 & 63 \\ 72 & 13 & 8 \end{pmatrix}, \text{ então } \mathbf{X}' = \begin{pmatrix} 35 & 22 & 72 \\ 40 & 34 & 13 \\ 52 & 13 & 8 \end{pmatrix} \quad (2.8)$$

Para a realização da soma e subtração de matrizes é necessário que o número de linhas e colunas sejam iguais nas duas matrizes que serão adicionadas ou subtraídas. Os exemplos 2.9 e 2.10 ilustram os cálculos de soma e subtração de matrizes.

$$\begin{pmatrix} 6 & 8 \\ 7 & 5 \end{pmatrix} + \begin{pmatrix} 2 & 3 \\ 4 & 1 \end{pmatrix} = \begin{pmatrix} 6+2 & 8+3 \\ 7+4 & 5+1 \end{pmatrix} = \begin{pmatrix} 8 & 11 \\ 13 & 6 \end{pmatrix} \quad (2.9)$$

$$\begin{pmatrix} 6 & 8 \\ 7 & 5 \end{pmatrix} - \begin{pmatrix} 2 & 3 \\ 4 & 1 \end{pmatrix} = \begin{pmatrix} 6-2 & 8-3 \\ 7-4 & 5-1 \end{pmatrix} = \begin{pmatrix} 4 & 5 \\ 3 & 4 \end{pmatrix} \quad (2.10)$$

Existem duas formas de multiplicação entre matrizes, que são: multiplicação escalar e multiplicação entre matrizes. Para as duas condições a multiplicação só é realizada se o número de linhas da matriz A for igual ao número de colunas da matriz B . Como resultado a matriz C , terá o mesmo número de linhas da matriz A e o mesmo número de colunas da matriz B . Para efetuar a multiplicação, segue a fórmula ilustrada em 2.11, em que c corresponde ao número de colunas de (A) ou número de linhas de (B) , ou seja, para se calcular C_{11} , é necessário multiplicar a primeira linha de (A) pela primeira coluna de (B) . Nos exemplos 2.12 e 2.13 são ilustrados a multiplicação de matrizes e multiplicação escalar [4].

$$C_{ij} = \sum_{z=1}^c A_{iz} * B_{zj} \quad (2.11)$$

$$\text{Matriz x Matriz} \quad \begin{pmatrix} 2 & 8 & 9 \\ 4 & 7 & 6 \end{pmatrix} * \begin{pmatrix} 6 & 5 \\ 1 & 3 \\ 6 & 7 \end{pmatrix} = \begin{pmatrix} 74 & 97 \\ 67 & 83 \end{pmatrix} \quad (2.12)$$

$$\text{Escalar x Matriz} \quad \begin{pmatrix} 3 \end{pmatrix} * \begin{pmatrix} 6 & 4 \\ 2 & 5 \end{pmatrix} = \begin{pmatrix} 18 & 12 \\ 6 & 15 \end{pmatrix} \quad (2.13)$$

2.6.3 Determinante

O determinante de uma matriz pode ser definida como os números associados as matrizes quadradas. O determinante é muito utilizada em resolução de sistemas lineares. Para se calcular a determinante, podem ser utilizados os seguintes métodos ilustrados em 2.15 para matrizes (2×2) e 2.17 para matrizes $(n \times n)$.

$$\begin{pmatrix} A \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad (2.14)$$

$$\det(A) = a_{11}a_{22} - a_{21}a_{12} \quad (2.15)$$

$$\begin{pmatrix} B \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} \quad (2.16)$$

$$\det(B) = b_{11}b_{22}b_{33} + b_{21}b_{32}b_{13} + b_{31}b_{12}b_{23} - b_{11}b_{32}b_{23} - b_{21}b_{12}b_{33} - b_{31}b_{22}b_{13} \quad (2.17)$$

Para se calcular a determinante de uma matriz, algumas propriedades devem ser seguidas, como: (1) Os determinantes de uma matriz e de sua transposta são iguais, isto é, $\det(A^T) = \det(A)$; (2) Se uma matriz B é obtida de uma matriz A trocando-se duas linhas (ou colunas) de A , então $\det(B) = -\det(A)$; (3) Se duas linhas (ou colunas) de A são iguais, então $\det(A) = 0$; (4) Se A tem uma linha (ou coluna) nula, então $\det(A) = 0$; (5) Se B é obtida de A multiplicando-se uma linha (ou coluna) de A por um número real c , então $\det(B) = c \cdot \det(A)$; (6) O determinante de um produto de matrizes é igual ao produto de seus determinantes, isto é, $\det(AB) = \det(A)\det(B)$.

2.6.4 Matriz Inversa

O inverso de uma matriz (A) é representado por $(A)^{-1}$ e para seu cálculo é necessário que satisfaça a condição de $(A) * (A)^{-1} = (I)$. Para alguns casos não existe a matriz inversa, onde a matriz é considerada invertível. Existem várias maneiras de se calcular a inversa de uma matriz, umas delas é estender a matriz (A) com a matriz identidade e aplicar em suas linhas e colunas as operações necessárias para que (A) se torne a matriz identidade. A ilustração 2.18, mostra um exemplo de cálculo da inversa de uma matriz [4].

$$\mathbf{A} = \left(\begin{array}{ccc} 2 & 1 & 1 \\ 1 & 1 & 1 \\ 2 & 3 & 2 \end{array} \right) = \left(\begin{array}{ccc|ccc} 2 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 2 & 3 & 2 & 1 & 1 & 1 \end{array} \right) = \left(\begin{array}{ccc} 1 & 1 & 0 \\ 0 & -2 & 1 \\ -1 & 4 & -1 \end{array} \right) \quad (2.18)$$

Para verificar se uma matriz quadrada é invertível, ou seja, se admite inversa, seu determinante deve ser diferente de *zero*, como é ilustrada em 2.19.

$$(A)^{-1} = \begin{cases} \exists, & \text{se } \det(A) \neq 0 \\ \nexists, & \text{se } \det(A) = 0 \end{cases} \quad (2.19)$$

2.6.5 Matriz de variância e covariância

A matriz de variância e covariância é utilizada para calcular a dispersão estatística entre variáveis. A variância de um vetor $x = [x_1, x_2, \dots, x_n]^t$, mede o afastamento dos seus elementos em torno da média (μ_x). A covariância é calculada de forma análoga à variância, mas envolvendo dois vetores de igual dimensão. Para o cálculo da matriz de variância e covariância (Σ) a seguinte

fórmula deve ser seguida, como: $\Sigma = E(X - \mu)(X - \mu)^T$, que é simétrica e definida positiva [32].

2.6.6 Autovalores e Autovetores

Um vetor coluna não nulo X de uma matriz quadrada A , é considerado um autovetor, se existir um escalar λ tal que, $AX = \lambda X \leftrightarrow AX = \lambda IX \leftrightarrow (A - \lambda I)X = 0$. Através da solução da equação característica da ilustração 2.20 de ordem $n \times n$, podemos encontrar os n autovalores de A , podendo assim ser reais ou complexos, inclusive nulos.

$$\det(A - \lambda I) = 0 \leftrightarrow \det \begin{pmatrix} a_{11} - \lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - \lambda & \dots & \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - \lambda \end{pmatrix} \quad (2.20)$$

2.7 Couro Bovino

O couro bovino é utilizado com material nobre em diversos segmentos industriais. Por ser um produto natural perecível, necessita de cuidados. Isso ocorre devido à sua constituição química, formada pelos seguintes componentes: água(60%), proteínas(33%), gorduras(2%), sais minerais (0,5%) e outras substâncias (0,5%) [14].

A anatomia do couro bovino pode ser dividida em três partes que são: epiderme³ (camada superior), a derme (camada intermediária, que resulta na pele comercializável) e a hipoderme⁴ (camada inferior, encostada na carne do animal). A derme é a camada industrializada que se transformará em couro, produto estável e resistente, sendo a camada mais importante do couro bovino.

A classificação brasileira das regiões do couro bovino são divididas em três partes: cabeça, grupão e flancos. Na classificação européia, a parte denominada cabeça é subdividida em paleta e colares e as demais partes são semelhantes. A parte mais nobre do couro bovino é o grupão (ou lombo ou dorso). Isso ocorre por que suas fibras têm a maior textura, maior uniformidade e resistência. A região da cabeça apresenta grande quantidade de rugas

³Camada formada por queratina, em que serão eliminados no processo de depilação

⁴Camada formada por tecido adiposo, sendo eliminada na operação de descarne

e pele de maior espessura. E por fim, os flancos que têm as fibras com textura inferior (vazias e abertas) às demais partes e mais finas, proporcionando mais facilidade de rupturas. A Figura 2.7 mostra essa divisão [12].

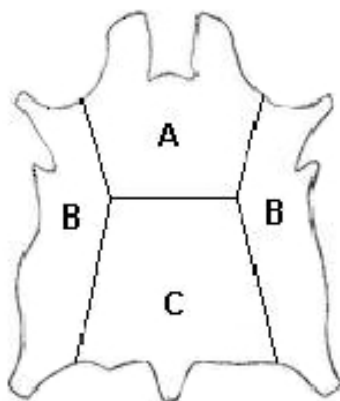


Figura 2.7: Partes do Couro Bovino (Classificação Brasileira). (A) Cabeça, (B) Flanco, (C) Grupão

Dois tipos de estágios do couro bovino serão descritos nesse trabalho: couro cru e wet-blue. Couro cru é o couro em sua fase inicial, que ainda não passou por nenhuma fase de curtimento. Wet-Blue é o couro já obtido na fase de curtimento, mais especificamente na operação chamada rebaixa. A Figura 2.8 ilustra esses estágios.

Um dos grandes problemas em couros bovinos são seus inúmeros tipos de defeitos. Entre eles se destacam: berne (defeitos similares a furos, causados pela larva da mosca conhecida como “berne”), Carrapato (marcas feitas pelo “carrapato”, e aparecem nos couros com maior flor⁵ lixada), cortes de esfolia (são cortes que aparecem no couro, causados por faca, quando os animais são abatidos), marca de fogo (defeitos causados por marcas de identificação do animal, causando grandes prejuízos nos couros), riscos (defeitos normalmente causados por chicote ou arame farpado) e veias (problemas arteriais do animal, em que problemas de estrutura ou rompimento se alargam e ficam perto da flor, aparecendo após o curtimento) [14]. A Figura 2.9 mostra alguns tipos de defeitos em Couro Cru e Wet-Blue.

⁵Parte externa do couro bovino, em que antes do uso é submetida a tratamentos especiais.

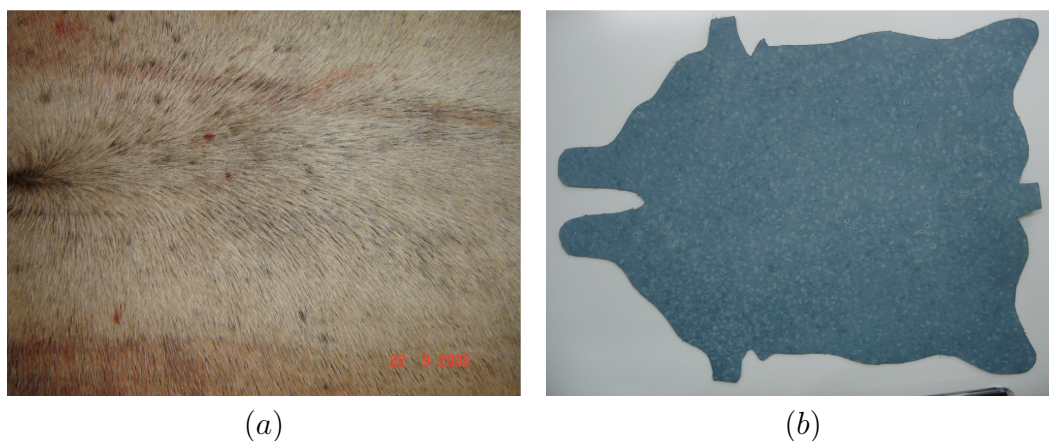


Figura 2.8: Imagem de couros bovino. (a) Imagem couro cru, (b) Imagem couro no estado Wet-Blue.

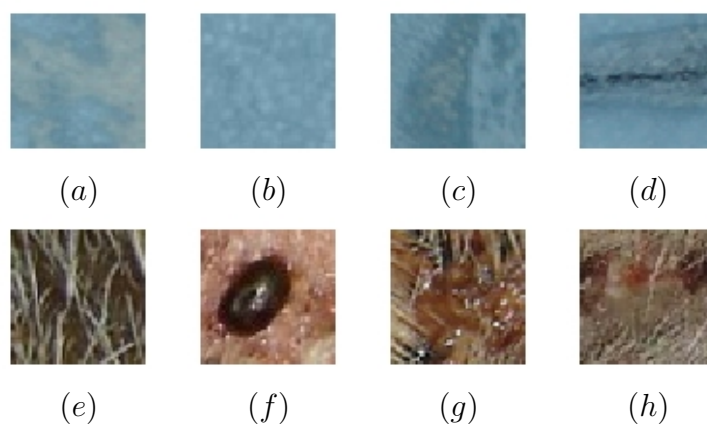


Figura 2.9: Imagem de defeitos de couros bovino nos processos (Couro Cru e Wet-Blue). (a) defeito sarna (Couro Cru), (b) defeito carrapato (Couro Cru), (c) defeito marca fogo (Couro Cru), (d) defeito risco (Couro Cru), (e) defeito sarna (wet-Blue), (f) defeito carrapato (wet-Blue), (g) defeito marca fogo (wet-Blue), (h) defeito risco (wet-Blue).

Capítulo 3

Análise Discriminante

A análise discriminante, também conhecida como discriminante linear de Fisher, é uma técnica que está se tornando muito comum em aplicações de visão computacional. A análise discriminante trata-se de um método que utiliza informações das categorias associadas a cada padrão para extrair linearmente as características mais discriminantes [3].

Através da análise discriminante podemos também realizar a discriminação entre classes, através de processos supervisionados (quando se conhece o padrão) ou através de processos não supervisionados, em que é utilizado quando não se tem um padrão conhecido. A análise discriminante é um método supervisionado na concepção estatística [18]. Ela deve ser empregada quando as seguintes condições forem atendidas [12].

1. As classes sob investigação são mutuamente exclusivas;
2. Cada classe é obtida de uma população normal multivariada;
3. As matrizes de covariância relativas a cada classe são iguais;
4. Devem existir no mínimo duas classes: $c \geq 2$, onde c é número de classes;
5. Devem existir pelo menos dois indivíduos por classe: $N_i > 2$, onde N_i é o número de indivíduos da classe i ;
6. Duas medidas não podem ser perfeitamente correlacionadas;
7. O número máximo de variáveis é igual ao número de observações menos dois: $0 < n < (N - 2)$;

Nos trabalho de [17] e [22], são apresentados de forma detalhada a utilização da análise discriminante de Fisher. Essa técnica consiste na computação de uma combinação linear de m variáveis quantitativas que mais eficientemente separam grupos de amostras em um espaço m -dimensional de modo que a razão da variância entre-grupos e inter-grupos é maximizada [28]. A separação intra-classe é medida através da seguinte fórmula estabelecida por Fisher.

1. Espalhamento intra-classes;

$$S_w = \sum_{j=1}^c \sum_{i=1}^{T_j} (x_i^j - u_j).(x_i^j - u_j)^t, \quad (3.1)$$

em que x_i^j é a i -ésimo exemplo da classe j , u_j é a média da classe j , T_j é o número de amostras da classe j e c é a quantidade de classes;

2. Espalhamento inter-classes;

$$S_b = \sum_{j=1}^c (u_j - u).(u_j - u)^t, \quad (3.2)$$

em que u é a média de todas as classes.

A partir do cálculo de espalhamento intra-classe e inter-classe de um conjunto de amostras, será possível seguir o critério de fisher que é maximizar a medida inter-classes e minimizar a medida intra-classes. Uma forma de fazer isso é maximizar a taxa $S_f = \det(S_b)/\det(S_w)$. Essa fórmula mostra que se S_w é uma matriz não singular, então essa taxa é maximizada quando os vetores colunas da matriz S_f são autovetores de $S_w^{-1}.S_b$.

3.1 Redução de Atributos

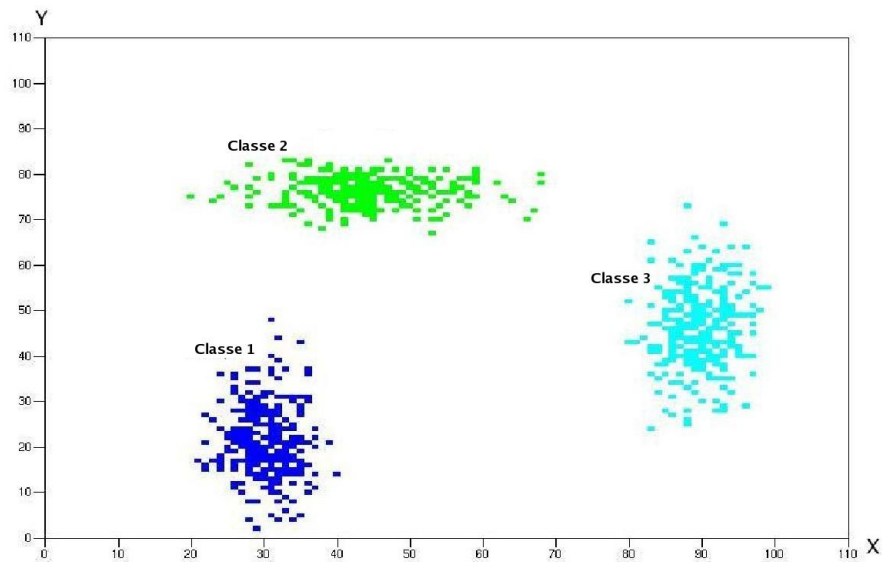
Em análise discriminante de Fisher, a redução de atributos é realizada a partir de um conjunto de amostras para n classes, tendo p variáveis, com o objetivo de reduzir para m atributos. Para a redução de atributos segue-se a seguinte linha de cálculo;

1. Calcular o espalhamento S_w e S_b para n classes;
2. Maximizar a medida inter-classes e minimizar a medida intra-classes S_f a partir de $S_w^{-1}.S_b$.

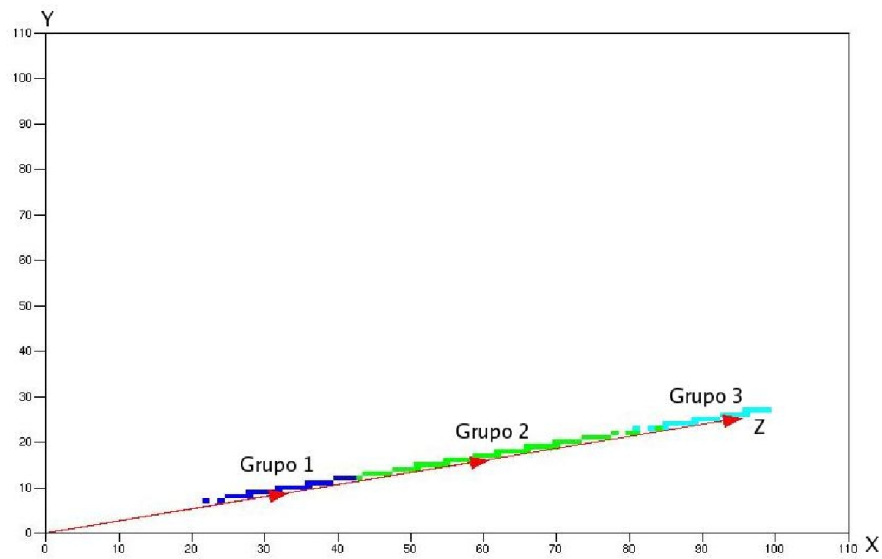
A partir de S_f é possível a redução de atributos utilizando os auto-vetores e auto-valores, em que os m atributos selecionados serão os m maiores auto-valores com maior grau de seus auto-vetores.

A Figura 3.1(a) ilustra 3 conjuntos de indivíduos representando cada classe de classificação e utilizando dois atributos (x, y) para sua representação. A Figura 3.1(b) mostra um exemplo usando Fisher para a redução de 2 atributos (x, y) para um atributo (z) . O objetivo da seleção de atributos utilizando análise discriminante de Fisher é realizar a redução de atributos, com o objetivo de manter a qualidade de sua discriminação.

A Figura 3.1(a) foi gerada a partir da contribuição do prof^o Dr. Hemerson Pistori, na criação do script em scilab da técnica de redução de atributos por Fisher. Para a visualização da projeção dos atributos reduzidos esse trabalho contribuiu com ajustes no script para uma melhor visualização, como é ilustrado na Figura 3.1(b).



(a)



(b)

Figura 3.1: Exemplo utilizando análise discriminante de Fisher para redução de variáveis (a) Conjunto de amostras utilizando 2 atributos (x, y) e (b) Redução de atributos para um único atributos (z)

Capítulo 4

Trabalhos Correlatos

Nesse capítulo serão apresentados alguns trabalhos, que têm como objetivo o estudo do couro bovino e técnicas que foram implementadas utilizando análise discriminante para solucionar problemas em visão computacional. Para uma melhor análise essa seção será dividida em duas partes que são: trabalhos sobre peles e couros bovino e trabalhos utilizando a análise discriminante de Fisher.

4.1 Peles e Couros Bovino

Pesquisas foram realizadas para a identificação das principais dificuldades que o setor de classificação do couro bovino sofre, e alguns trabalhos foram selecionados. Gomes [13], mostra uma pesquisa e avaliação do sistema de classificação de couros bovino proposto pelo MAPA, ao longo de dois anos em frigoríficos e curtumes de Campo Grande e Nova Andradina/MS. Através de avaliações de viabilidade das técnicas dos critérios de qualidade do couro estabelecido pelo Sistema de Classificação de Couro Bovinos, esse trabalho propôs ajustes no sistema para suprir as principais necessidades encontradas. Nesse trabalho foram propostas algumas mudanças no sistema de classificação de peles e couros bovino, como: mudança da categoria C (inferior) para a categoria (tipo D) de desclassificados, para aqueles que não se enquadram na categoria A ou B (classificatórios).

Outro trabalho muito importante é de Pereira [25]. O objetivo desse trabalho foi verificar a ocorrência de defeitos oriundos na propriedade rural e sua influência na classificação do couro bovino tanto no frigorífico quanto no curtume, e analisar a adequação da metodologia de classificação proposta pelo MAPA. Com esse trabalho teve-se como conclusão que o couro bovino em geral, possui uma baixa qualidade, e que os principais fatores que comprometem essa baixa qualidade é a marcação inadequada na parte nobre da

pele bovina. Sendo causado intencionalmente pelo homem.

Marçal [21] descreve a falta de preocupação do pecuarista brasileiro no cuidado do couro bovino e destaca algumas medidas preventivas para ajudar no manejo sanitário. Uma delas é evitar que os animais sejam infectados por carrapatos e bernes, ou até mesmo, problemas como fotossensibilidade, mencionando que o aspecto importante para isso, seria a orientação de peões e capatazes. Na realização de marcações de identificação do boi, a proposta dada por este trabalho seria a marcação sobre a face, assim evitando sempre as partes altas, o que desvalorizam completamente a matéria prima. Como conclusão esse trabalho, mostra que os danos causados ao couro bovino é por consequência dos criados e podem ser minimizados seguindo algumas orientações como: exame periódico das cercas da fazenda, sendo trocadas por arames lisos; retirada de objetos que possam machucar o gado no pasto; não utilizar ferrão pontiagudo para conduzir o gado; inspecionar a grade do caminhão, verificando se a carroceria está em boas condições de uso; cuidar imediatamente dos ferimentos acidentais; conservar corretamente o couro e principalmente, participação efetiva das escolas de medicina veterinária orientando futuros colegas sobre a importância dessa matéria prima, não só gerando divisas e sim proporcionando novos empregos.

4.2 Análise Discriminante

Diversos trabalhos foram encontrados em diversas áreas de aplicação utilizando análise discriminante. No trabalho [5], mostra os estudos realizados de métodos de extração de características e de classificação estatística para a aplicação em reconhecimento de faces. Mais especificamente essa tese de mestrado focaliza seu trabalho em técnicas para redução de dimensionabilidade utilizando principalmente seleção de características e entre elas técnicas para extração de atributos como a análise discriminante. Com esse trabalho tem-se como resultado de que se utilizando métodos de análise estatística é possível obter um grande aumento no poder de reconhecimento de padrões e uma maior velocidade em seu reconhecimento.

Amado [1], utiliza também análise discriminante para resolver problemas de seleção de atributos, afim de melhorar o desempenho em termos de classificação. São demonstradas também técnicas que foram utilizadas para análise dos resultados utilizando métodos discriminante como: os métodos de seleção teste F, teste Bootstrap e métodos robustos. Como resultado esse trabalho mostra que a metodologia bootstrap fornece uma ferramenta importante para análise de múltiplos problemas que surgem no dia a dia, uma vez que permite “estimar” a distribuição de estimadores sem qualquer

conhecimento sobre a distribuição da população ou populações.

Em Santos [34] é apresentado um estudo sobre o modelo de composição de especialista locais (CEL) e um estudo sobre análise discriminante para classificação de dados. Nesse trabalho, experimentos com base em dados reais foram utilizados para medição de desempenho e comparação entre o modelo CEL e análise discriminante. Para a realização dos testes utilizando análise discriminante foram demonstrados e utilizados três técnicas de discriminação: Análise Discriminante de Fisher, Regressão Logística e Extended. Os resultados desse trabalho mostram que o modelo de classificação CEL e análise discriminante, tiveram um grande êxito na classificação dos seus dados, em que por uma pequena diferença na taxa de classificação correta a análise discriminante, obteve o maior sucesso, chegando a 91,6% de acerto.

Capítulo 5

Desenvolvimento

Para a realização dos experimentos de detecção e classificação do couro bovino utilizando a análise discriminante de Fisher, foram implementados 3 módulos. O primeiro módulo implementado tem o objetivo de criar um ambiente de fácil interação para a realização de marcação e rotulação no banco de dados, das principais áreas do couro bovino. Com essas imagens pré-cadastradas e pré-processadas o módulo é capaz de realizar a geração de amostras referentes as marcações sobre as imagens do couro bovino.

O segundo módulo implementado tem como objetivo criar um ambiente de experimentos, dando opções de extração e seleção das melhores informações, criação de arquivos de interpretação para experimentos no *WEKA*, realização de testes de classificação, entre outros. Esse módulo é responsável por criar a base de conhecimentos para a aprendizagem, passando essas informações para o próximo módulo.

O terceiro módulo é o de marcação visual, esse módulo utiliza as bases de conhecimento para aprendizagem gerada, e através dela realiza a varredura na imagem para a detecção das diversas incoformidades. Para a implementação desses três módulos foram utilizadas e integradas três ferramentas livres, que são: *IMAGEJ*, *WEKA* e *SIGUS*. A seguir serão descritas a implementação de cada módulo e ferramentas utilizadas. A Figura 5.1 mostra a tela principal do DTCOURO.

5.1 IMAGEJ

Para a utilização de recursos de processamento digital de imagens e visão computacional, foi proposta a utilização do pacote *ImageJ*, que é uma versão multiplataforma, do software *NIH Image*, para *Macintosh*. Entre os recursos oferecidos pelo pacote, destacamos a disponibilidade de programas-fonte



Figura 5.1: Tela inicial do DTCOURO

abertos de diversos algoritmos como: manipulação dos mais variados formatos de arquivo de imagens, detecção de bordas, melhoria de imagens, cálculos diversos (áreas, médias, centróides) e operações morfológicas. Esse software disponibiliza também um ambiente gráfico que simplifica a utilização de tais recursos, além de permitir a extensão através de plugins escritos em Java. Um outro fator importante para a sua utilização é a existência de uma grande comunidade de programadores trabalhando em seu desenvolvimento, com novos plugins sendo disponibilizados freqüentemente [27].

5.2 WEKA

O segundo pacote proposto para a utilização no desenvolvimento desse módulo foi o *WEKA* (Waikato Environment for Knowledge Analysis)[18], também escrito em Java e com programas-fonte abertos. O *WEKA* é um ambiente bastante utilizado em pesquisas na área de aprendizagem de máquina, pois oferece diversos componentes que facilita a implementação de classificadores. Além disto, esse ambiente permite que novos algoritmos sejam comparados a outros algoritmos já consolidados na área de aprendizagem, como o C4.5, o Backpropagation, o KNN e o naiveBayes, entre outros [27]. Esse pacote será utilizado para ajudar nos testes que serão executados, para verificar se com uma grande base de amostras e com um conjunto de atributos, se aumentará o desempenho do classificador de couros bovino do projeto DTCOURO. Com essa ferramenta é possível também obter facilmente dados como: tempo de aprendizagem, tempo de classificação, porcentagem de acerto, porcentagem de erro, falso positivo, falso negativo, matriz de confusão, entre outros.

5.3 SIGUS

O terceiro pacote utilizado para facilitar o desenvolvimento do projeto foi o *SIGUS*. A plataforma *SIGUS* é um ambiente computacional de apoio ao desenvolvimento de sistema para inclusão digital de pessoas com necessidades especiais, com objetivo de aumentar a quantidade de programas destinados a essas pessoas [26]. Através deste pacote serão utilizados os métodos de extração de atributos implementados, como: matriz de co-ocorrência, mapas de interação e atributos de cores (HSB).

5.4 Ambiente de Aquisição e Marcação

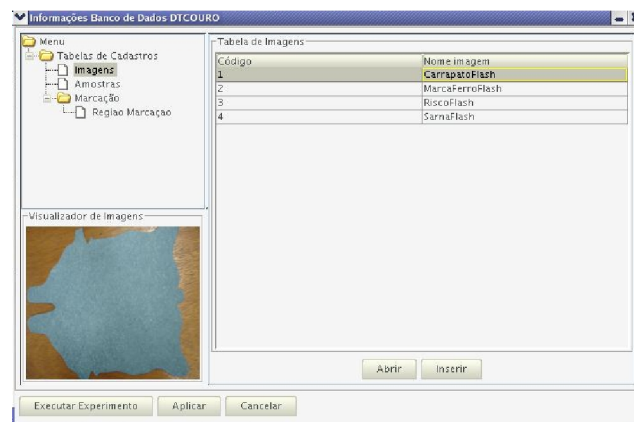
O primeiro módulo implementado auxilia nas seguintes fases de processamento: (1) fase de aquisição de imagens, (2) fase de pré-processamento, (3) fase de marcação e rotulação das principais áreas que concentram os defeitos e a (4) fase de geração de imagens para testes.

Na fase de aquisição de imagens, foi criado de forma ilustrativa, uma janela de seleção e inserção, em que o usuário poderá pré-visualizar e cadastrar as imagens que servirão como base de trabalho. A janela também da opções de visualização das amostras geradas e as regiões de marcação. A Figura 5.2 ilustra com mais detalhes a fase de aquisição de imagens, com suas opções auxiliares.

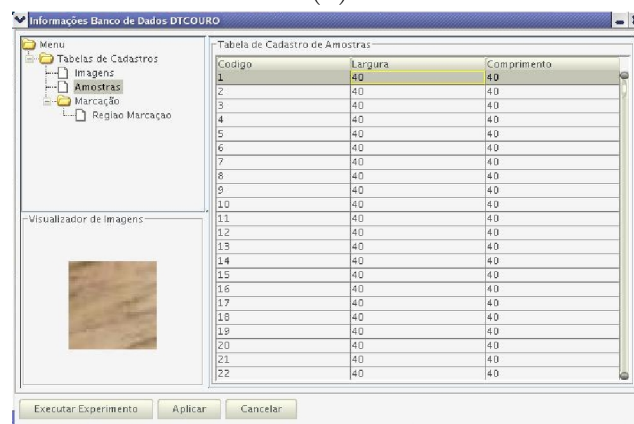
Na fase de pré-processamento o módulo disponibiliza ao usuário a opção de realizar processamentos nas imagens como: filtros (convolução, filtros de Fourier), limiarização, detector de bordas, conversão de formato de cores(RGB, HSB), entre outros. Na fase de marcação e rotulação o usuário realiza marcações sobre regiões de interesse. A cada marcação, o módulo cria um registro no banco de imagens, contendo informações capturadas do usuário como: nome do defeito marcado, tipo de defeito selecionado e região da marcação. A Figura 5.3, mostra a fase de pré-processamento.

A marcação é feita utilizando o conceito de máscara, em que somente é criada uma camada de visualização sobre a imagem, assim não afetando a imagem original. Em qualquer marcação realizada sobre a imagem, o módulo captura todas as informações referente a marcação, como: tipo de marcação e a região marcada. O módulo também oferece a opção de calibragem da marcação, podendo dessa forma capturar a região em centímetros ou milímetros. A Figura 5.4, mostra o processo de marcação.

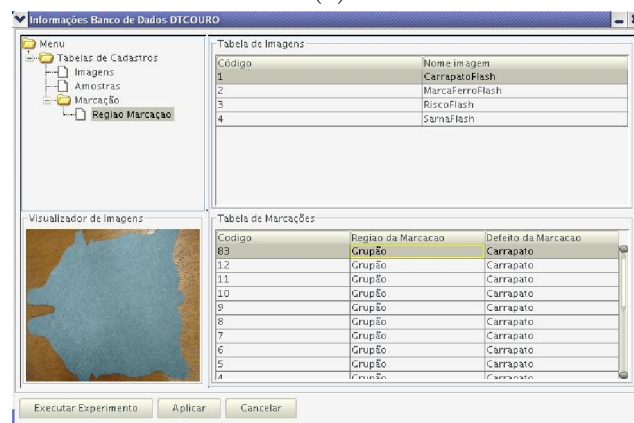
Todas essas informações são armazenadas automaticamente a cada marcação, fazendo com que a imagem original não seja modificada, apenas que seja criada uma máscara sobre a mesma. Dessa forma o módulo



(a)

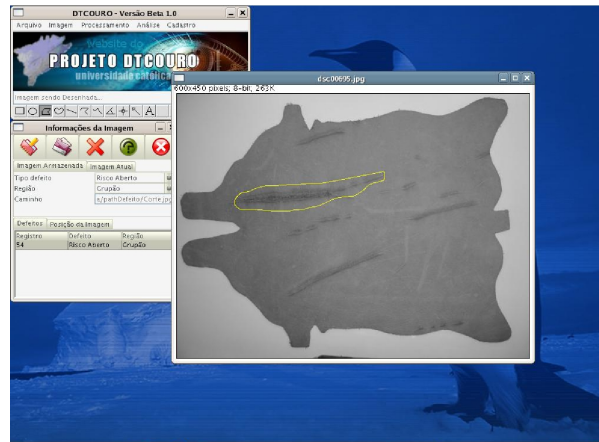


(b)

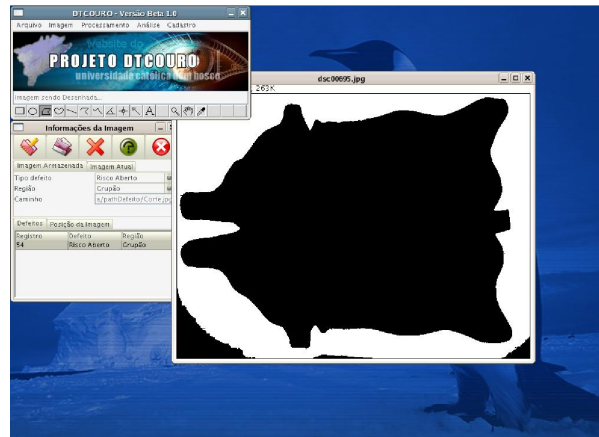


(c)

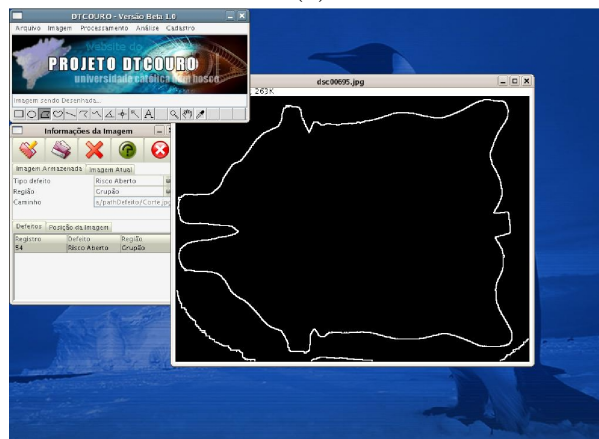
Figura 5.2: Imagens da fase de aquisição e registro de imagens: (a) Seleção e inserção de imagens, (b) Seleção e visualização de amostras, (c) Seleção e visualização de regiões de marcação



(a)



(b)



(c)

Figura 5.3: Imagens do módulo de pré-processamento e geração de imagens para o treinamento e testes de classificação de peles e couros bovinos: (a) Imagem sendo transformada para 8-bit (b), Imagem sendo binarizada, (c) Imagem sendo passada pelo processo de convolução para detecção de bordas

realiza a marcação visual a medida que a imagem é solicitada, caso isso ocorra, o módulo buscará no banco de dados todas as informações referentes as marcações. Caso exista alguma marcação referente a imagem, o módulo ira realiza-la com base nas informações armazenadas.



Figura 5.4: Imagens do módulo de pré-processamento e geração de imagens para o treinamento e testes de classificação de peles e couros bovinos: (a) Marcação sendo realizada (b) Cadastro no banco da marcação realizada

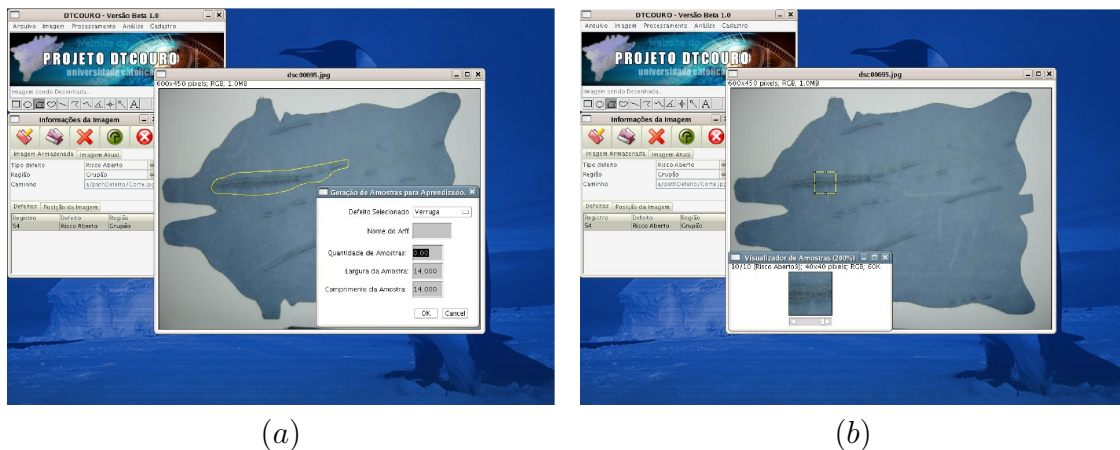


Figura 5.5: Imagens do módulo de pré-processamento e geração de imagens para o treinamento e testes de classificação de peles e couros bovinos: (a) Tela de opções para a geração de amostras (b) Visualização das amostras geradas

Na fase de geração de imagens de testes, o módulo utiliza a base de imagens marcadas e armazenadas no banco de dados. Para a geração de

amostras é utilizado o modo de varredura. Esse modo de geração tem como objetivo principal varrer a imagem ativa, e através das informações capturas do usuário como: marcação selecionada, quantidade de amostras a serem geradas, largura e comprimento em pixel de cada amostra, percorrer a imagem e verificando os quatro pontos de cada amostra se estão localizados dentro da marcação selecionada. Caso esteja localizado dentro da região, o modo varredura cria uma cópia dessa amostra e armazena no banco de dados, caso não esteja localizado dentro da região, o sistema ira ignorar, e continuará sua varredura até o fim da imagem. Dessa forma o módulo cria um banco de amostras específicas para cada tipo de defeito selecionado. A Figura 5.5, mostra a fase de geração de amostras.

5.5 Módulo de Experimentos

O ambiente de experimentos foi criado com o objetivo de facilitar o uso desde a manipulação de amostras, extração e seleção de atributos e testes estatísticos na classificação utilizando opções de classificadores. Esse ambiente se divide em três seções, que são: seção atributos, seção experimentos e seção resultados. A seção de atributos disponibiliza ao usuário a opção de escolha das amostras dos tipos de defeitos cadastrados. É também disponibilizada, uma lista de todos os tipos de métodos de extração de atributos que estão localizados dentro do pacote *SIGUS*. Dessa forma, a partir das amostras selecionadas e dos métodos escolhidos de extração de atributos, ja é possível a geração do arquivo de interpretação do *WEKA*. A Figura 5.6 ilustra a seção de atributos.

A seção de experimento, foi criada a partir da integração dos componentes do *WEKA* “SetupModePanel” e “RunPanel”. SetupModelPanel tem a função da criação de novos experimentos, dando opções de escolha, como: tipos de algoritmos de aprendizagem, seleção dos arquivos “.arff” que serão analisados, tipos de experimentos, entre outros. A Figura 5.7 ilustra a seção de experimentos.

O componente RunPanel tem somente a função de capturar o experimento ativo criado, e a partir das propriedades configuradas realizar o experimento. A seção de resultados também integrada do *WEKA*, tem a função de capturar a saída do “RunPanel” executado, e partir dele, visualizar seus resultados. A Figura 5.8 ilustra a seção de resultados. A partir do ambiente de experimentos será criado a base de conhecimento que será utilizada na fase de detecção automática e marcação visual.

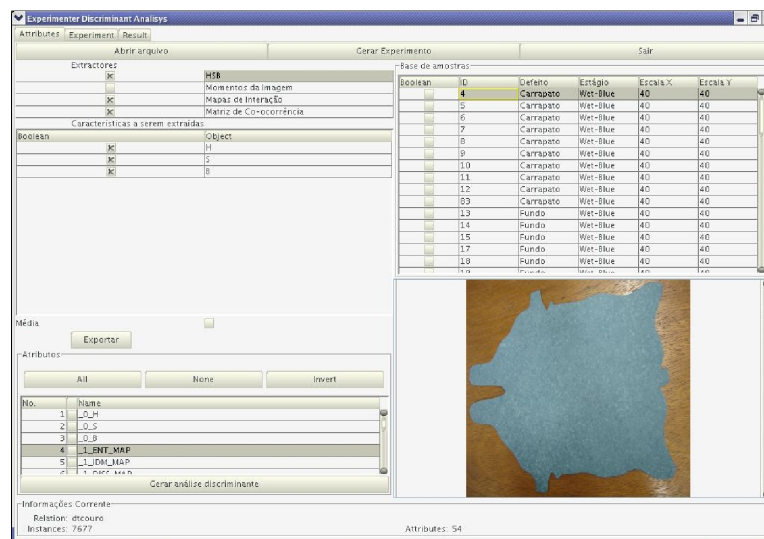


Figura 5.6: Seção de Atributos

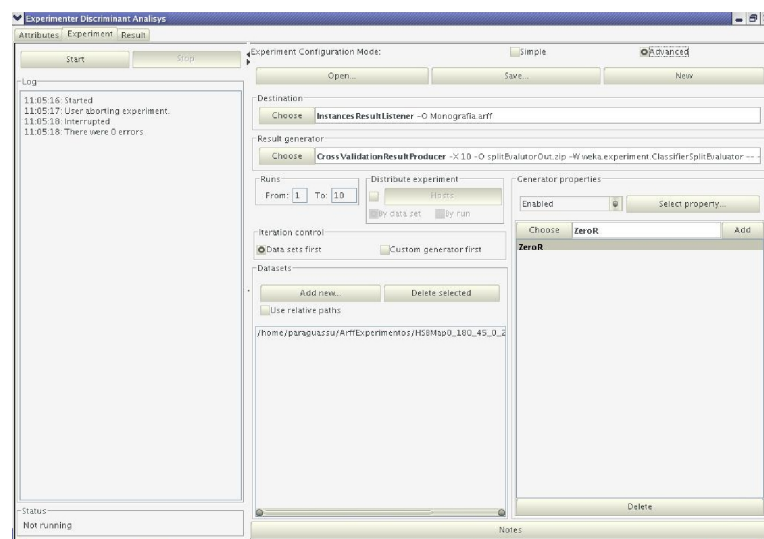


Figura 5.7: Seção de Experimentos

5.6 Módulo de Detecção de defeitos e Marcação Visual

O módulo de detecção de defeitos foi desenvolvido com a finalidade de realizar a marcação visual dos defeitos encontrados em imagens, a partir da base de conhecimentos gerada. A partir da base de conhecimento e através das propriedades informadas pelo usuário, o módulo é capaz de varrer a imagem

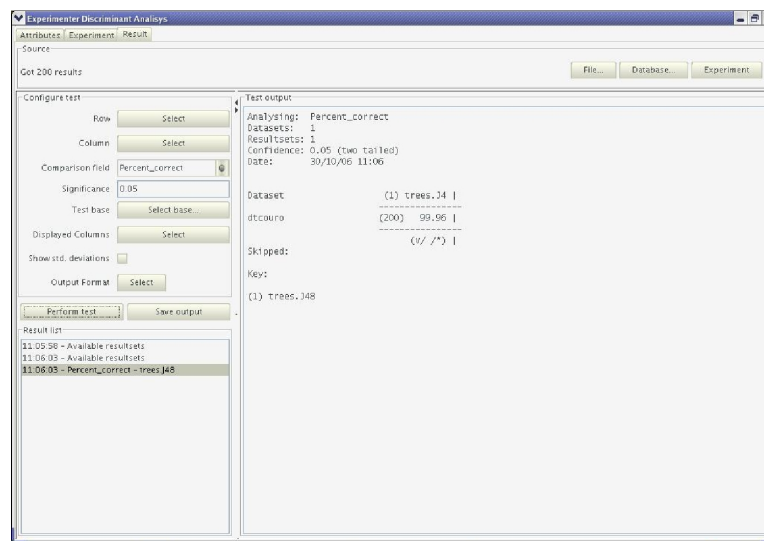


Figura 5.8: Seção de Resultados

ativa, extraindo os atributos selecionados e realizar a redução dos mesmos através da técnica de Fisher. Foi desenvolvida também a opção de escolha do algoritmo de aprendizagem que será utilizado na classificação, como: árvore de decisão, naive Bayes, SVM, redes neurais, entre outros. A Figura 5.9 ilustra o módulo de detecção e marcação visual.

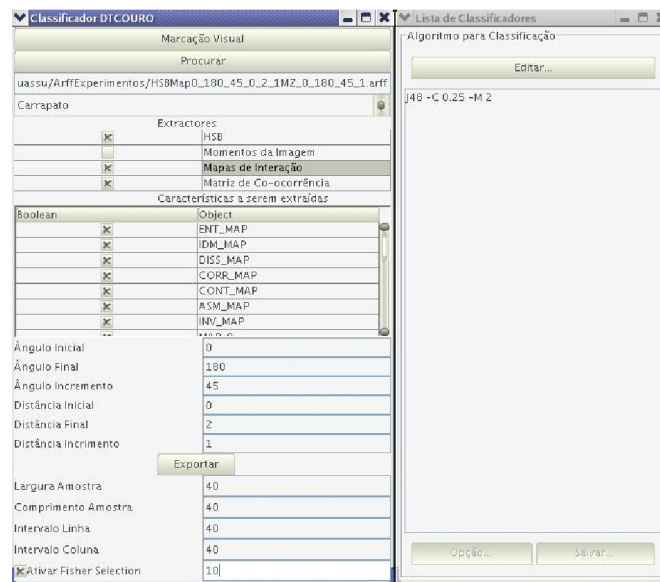


Figura 5.9: Módulo de Detecção de defeitos e marcação visual

Para que se tenha uma melhor visualização dos defeitos marcados, foi criada a opção de cores para classificação, em que a partir da base de aprendizagem, o módulo lista as classes que serão utilizadas e através dela é dada uma sugestão de cores para a marcação visual de cada defeito. Desse modo será fácil a percepção visual dos acertos e erros de classificação, ajudando a melhorar a base de conhecimento.

Capítulo 6

Experimentos e Análise de Resultados

Para a realização dos experimentos dos módulos implementados, foi utilizado um computador com processador Pentium 4 - 2.3 GHz com 512 MB de memória RAM. Junto com a pesquisadora Mariana de Aragão Pereira da EMBRAPA e acadêmicos de pesquisa do grupo GPEC (Grupo de Pesquisa em Engenharia de Computação), foi criado um banco de imagens, com diferentes tipos de defeitos de couros bovinos no estágio Wet-Blue.

Todas essas imagens foram cadastradas no módulo de geração de amostras, e em seguida foram realizadas as marcações referentes aos seus defeitos. Com essas informações armazenadas no banco de dados, foi possível realizar então a fase de geração de amostras. Nessa fase o módulo utilizou todos os dados das marcações, para realizar uma varredura sobre os tipos de defeitos pré-cadastrados, criando dessa forma um maior conjunto de amostras específicas para cada tipo de defeito.

Inicialmente para a realização desses experimentos, dois testes iniciais foram feitos utilizando os módulos e os métodos de extração de atributos implementados, que foram Amorim [2] e Ramirez [29]. O objetivo do trabalho de [2], foi demonstrar o uso da ferramenta de marcação e geração de amostras, na criação de um banco de imagens e geração de amostras específicas para cada tipo de defeito. Com esse trabalho foram realizados testes com imagens do couro bovino no processo Wet-Blue, com dois tipos de iluminação, que foram: com *flash* e sem *flash*. A partir desse trabalho foi verificado que o módulo de geração de amostras auxiliou com grande êxito o classificador do DTCOURO, chegando como resultado de classificação de até 99.6% de acerto.

Em Ramirez [29], foram demonstrados os resultados obtidos na comparação das técnicas de extração de atributos utilizando matriz de co-

ocorrência e mapas de interação. O objetivo desse trabalho foi reunir duas técnicas muito utilizadas em análise de textura para medir o desempenho na classificação de diversos tipos de defeitos em peles e couros bovinos. As imagens que foram utilizadas nesse trabalho foram as mesmas de [2], tentando dessa forma demonstrar somente a comparação dos resultados de classificação. Como resultado esse trabalho mostrou que a técnica de matriz de co-ocorrência com (96% de acerto) mostrou superioridade sobre a técnica de mapas de interação com (81.125% de acerto).

6.1 Experimentos

Para a realização dos experimentos, utilizamos as técnicas implementadas dos trabalhos anteriores para a realização dos experimentos usando a análise discriminante de Fisher. Como esse trabalho tem o objetivo de avaliar o desempenho da redução de atributos, iremos utilizar somente imagens de couros bovinos no estágio Wet-Blue, localizadas no banco de imagens do projeto DTCOURO. Para a realização desses experimentos foram utilizadas 4 imagens do couro bovino no estágio Wet-Blue, sendo uma para cada defeito, que são: marca ferro, risco, sarna e carrapato. A partir do módulo de marcação, foram realizados para cada imagem, três tipos de regras de marcação, que foram: defeito específico da imagem, região sem defeito, e a região que se encontra o couro bovino. O objetivo para esse tipo de regra de marcação foi a necessidade de se obter uma maior variação de regiões sem defeito, fundo, e regiões com defeitos, próximos um do outro.

Durante a fase de marcação houve uma grande preocupação nos defeitos que são visualmente similares, para que quando fossem realizados os testes, não tivesse tanto efeito em sua classificação. Dessa forma foi realizada a marcação especificamente na parte do defeito que mais se destacava, possibilitando assim uma melhora na base dados do classificador. A Figura 6.1 ilustra as imagens utilizadas, com suas marcações realizadas manualmente.

Em seguida, para cada tipo de marcação cadastrada no banco de dados, foram geradas todas as amostras possíveis referentes a cada marcação na escala de 40X40 pixel e com um intervalo de uma amostra a outra de 5 pixels. A quantidade de amostras geradas, foram: 1.368 amostras para carrapato, 411 amostras para risco, 863 amostras para sarna, 778 amostras para marca ferro, 3.218 amostras para fundo e 1.042 amostras para sem defeito.

A partir de cada amostra gerada, o módulo de experimentos do DTCOURO realizou a extração dos atributos. Os atributos que foram capturados pelo módulo foram, matriz de co-ocorrência com variações de (0° a 360° com intervalo de 45° e distância de 1 pixel) extraíndo 56 atributos, mapas

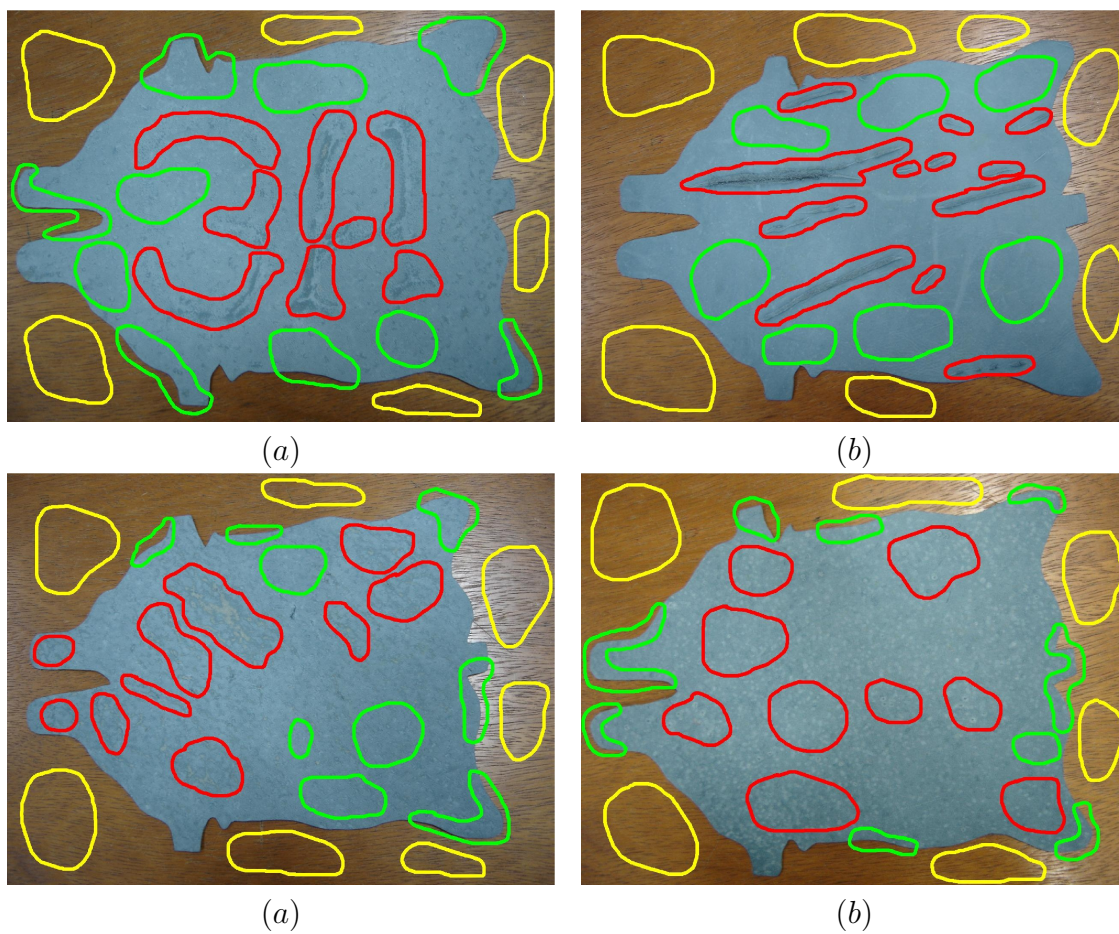


Figura 6.1: Imagens marcadas manualmente pelo módulo de pré-processamento e geração de amostras, com as seguintes cores, (vermelho = defeito, amarelo = fundo e verde = sem defeito). (a) defeito marca ferro, (b) defeito risco, (c) defeito sarna e (d) defeito carrapato.

de interação com variações de (0° a 360° com intervalo de 45° e distância inicial de 0 a 2 com intervalo de 1 pixel) extraíndo 7 atributos e as médias de H(matiz), S(saturação) e B(brilho) com 3 atributos. Para cada matriz gerada foram capturados também atributos, como: entropia (*ENT*), momento da diferença inversa (*IDM*), dissimilaridade (*DISS*), correlação (*CORR*), contraste (*CONT*), segundo momento angular (*ASM*) e a diferença inversa (*INV*), totalizando 63 atributos de textura e 3 atributos de cor. A Tabela 6.1 mostra os atributos extraídos

Tabela 6.1: Atributos extraídos (MAT = Matriz de Co-ocorrência e MAP = Mapas de Interação)

H(matiz)	IDM-MAT-90	ASM-MAT-180	DISS-MAT-315
S(saturação)	DISS-MAT-90	INV-MAT-180	CORR-MAT-315
B(brilho)	CORR-MAT-90	ENT-MAT-225	CONT-MAT-315
ENT-MAP	CONT-MAT-90	IDM-MAT-225	ASM-MAT-315
IDM-MAP	ASM-MAT-90	DISS-MAT-225	INV-MAT-315
DISS-MAP	INV-MAT-90	CORR-MAT-225	ENT-MAT-360
CORR-MAP	ENT-MAT-135	CONT-MAT-225	IDM-MAT-360
CONT-MAP	IDM-MAT-135	ASM-MAT-225	DISS-MAT-360
ASM-MAP	DISS-MAT-135	INV-MAT-225	CORR-MAT-360
INV-MAP	CORR-MAT-135	ENT-MAT-270	CONT-MAT-360
ENT-MAT-45	CONT-MAT-135	IDM-MAT-270	ASM-MAT-360
IDM-MAT-45	ASM-MAT-135	DISS-MAT-270	INV-MAT-360
DISS-MAT-45	INV-MAT-135	CORR-MAT-270	
CORR-MAT-45	ENT-MAT-180	CONT-MAT-270	
CONT-MAT-45	IDM-MAT-180	ASM-MAT-270	
ASM-MAT-45	DISS-MAT-180	INV-MAT-270	
INV-MAT-45	CORR-MAT-180	ENT-MAT-315	
ENT-MAT-90	CONT-MAT-180	IDM-MAT-315	

A partir dos 66 atributos extraídos das amostras, utilizando-se o módulo de experimentos do DTCOURO e junto com a técnica implementada de Fisher para redução de atributos, foram criadas 33 novas bases de aprendizagem no formato de interpretação do *WEKA*, sendo cada uma contendo

$\sum_{i=1}^{66} i + 2$ atributos. Utilizamos a sessão de experimentos do *WEKA*, para a medição do desempenho do classificador. Para isso foram escolhidos 2 algoritmos de aprendizagem e classificação, que podem ser visualizados na Tabela 6.2.

O algoritmo J48 ou C4.5 constrói uma árvore de decisão a partir de um conjunto de treinamento previamente fornecido. Este conjunto é composto por várias instâncias do problema, onde para cada uma, são fornecidos valores para atributos e para uma classe. Os atributos podem conter valores discretos, contínuos e também o valor “nulo”, que representa valores não especificados. As classes são representadas por valores discretos, e não admitem valores nulos [31].

O classificador IBk é uma técnica de aprendizagem baseada em instâncias, que utiliza o algoritmo k-Nearest Neighbour (kNN). Seu funcionamento se baseia no aprendizado supervisionado não paramétrico, em que um atributo é associado a uma classe, dependendo assim do número de K atributos vizinhos que pertençam a essa classe, seguindo algum critério de distância [4]. Em nossos experimentos utilizamos o valor de $K = 5$.

Para a realização dos experimentos, o *WEKA* escolhe como padrão, o modo de validação cruzada em 10 dobras e 10 validações. Esse modo de validação pega o total de amostras, mantém 90% para a aprendizagem e os 10% restante para a realização dos testes. Isso ocorre 10 vezes para cada validação, sendo uma para cada 10% seguintes. Com esses 100 resultados é calculada sua média indicando sua porcentagem de acerto. Nesse trabalho foi utilizado o modo de validação cruzada com 5 dobras e 3 validações.

Tabela 6.2: Algoritmos de Aprendizagem

Sigla	Descrição	Configuração
J48	árvore de decisão, apenas para classe nominal	padrão <i>WEKA</i>
IBk	classificação baseada em instâncias	Exponencial = 5

Para a visualização dos resultados, foram selecionados como critérios para comparação de desempenho algumas características, que são ilustrados na Tabela 6.3. Dessa forma será possível analisar se o desempenho do classificador irá aumentar, a medida que os atributos serão reduzidos utilizando análise discriminante de Fisher.

Tabela 6.3: Características para a medição de desempenho

Porcentagem de acerto de classificação(%)
Tempo de treinamento dos algoritmos de aprendizagem(ms)
Tempo de classificação das amostras(ms)
Matriz de confusão

6.2 Resultados e Análise

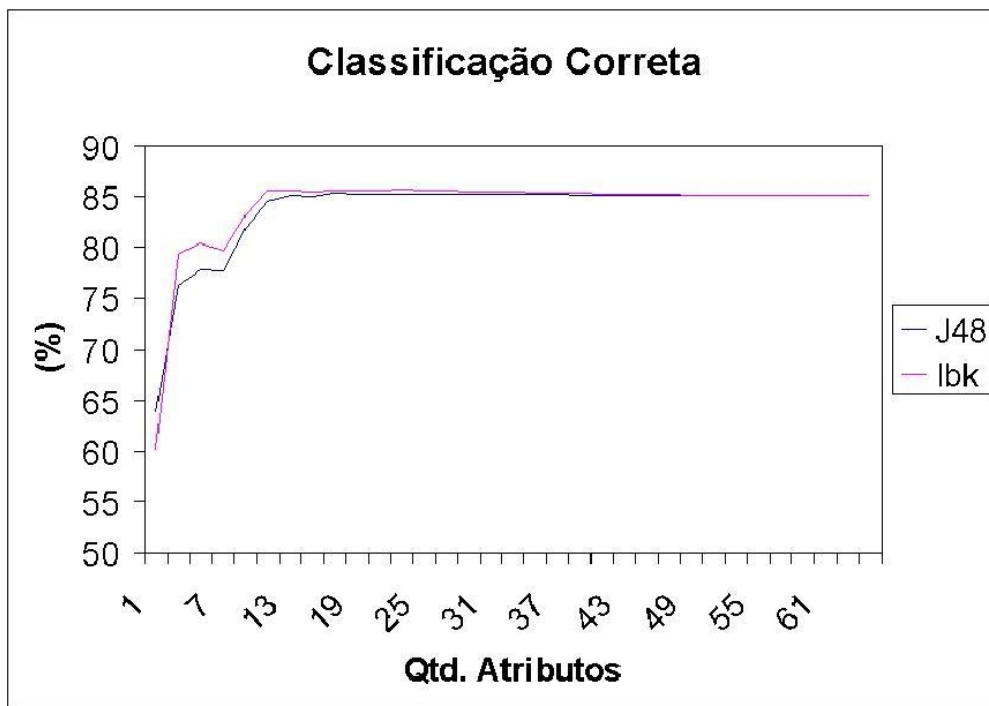
Para avaliar os resultados, mostraremos graficamente o comportamento dos classificadores a medida que o número de atributos é reduzido. Nos experimentos foram utilizadas 7.680 amostras, para a classificação de 6 classes (carrapato, sarna, risco, marca ferro, fundo e sem defeito), utilizando os algoritmos de aprendizagem (J48 e IBk). O tempo total para a realização dos 33 experimentos foram de (J48 = 280 minutos) e (IBk = 360 minutos). Os Gráficos 6.2(a), 6.2(b), 6.3 e as tabelas do Anexo A ilustram os resultados com base na porcentagem de acerto, tempo de treinamento e tempo de classificação para cada um dos classificadores.

Com base nos resultados, podemos observar que a taxa de acerto do classificador não teve um grande êxito, isso se deve a baixa quantidade de imagens utilizadas e pequena variação nos defeitos marcados. Analisando o comportamento do gráfico 6.2(a), podemos verificar que mesmo com a redução de atributos, mas com pequenas variações da taxa de acerto, a análise discriminante de Fisher apresentou um bom desempenho, mantendo o nível de classificação quase constante. Pelo gráfico também pode ser percebido que a porcentagem de acerto melhora a medida que a quantidade de atributos é aumentada, isto até uma certa quantidade, em que há uma estabilidade. Com a utilização de dois classificadores conseguimos também visualizar que independente do algoritmo de classificação o comportamento da técnica de Fisher é a mesma.

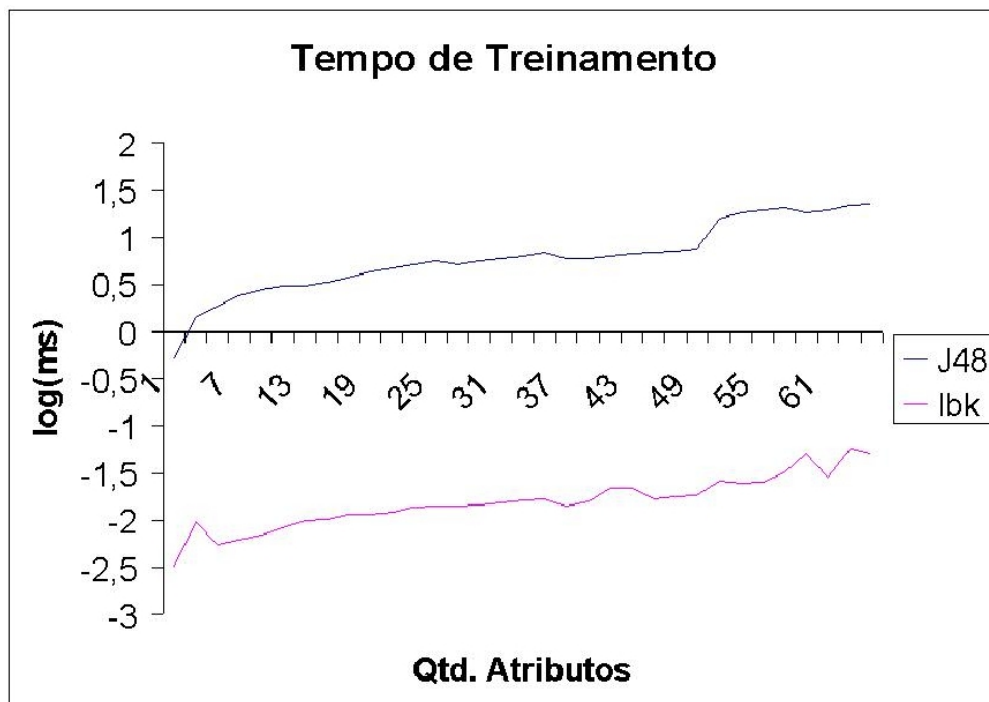
Os gráficos 6.2(a) e 6.3 foram ilustrados com o objetivo de mostrar que a redução de atributos é uma etapa muito importante para classificação de imagens. A medida que a quantidade de atributos é acrescentada o tempo de treinamento e classificação aumentam, mostrando assim que a utilização de um técnica para redução de atributos, como análise discriminante de Fisher, pode dar eficiência, velocidade e confiança para um classificador automático.

Nas Tabelas 6.4 e 6.5 são ilustrados as matrizes de confusão do J48 e IBk, a partir do conjunto com 13 atributos, que foi o ponto em que houve a estabilidade da porcentagem acerto. Analisando as matrizes de confusão, pode-se perceber que houve grande discriminação entre as amostras com defeito e até mesmo regiões sem defeito. O fato de algumas amostras terem sido classificadas incorretamente, mostra que alguns tipos de defeitos apresentam valores semelhantes nos atributos selecionados para a discriminação.

Espera-se uma elevação no grau de confusão com o aumento da quantidade de diferentes defeitos em diferentes situações. Também se pode constatar que alguns defeitos, como é ilustrado no Anexo B, apresentam semelhanças, por apresentarem uma textura de cicatriz muito semelhante.



(a)



(b)

Figura 6.2: Comportamento do classificador, (a) classificação correta(%) e (b) tempo de treinamento(ms)

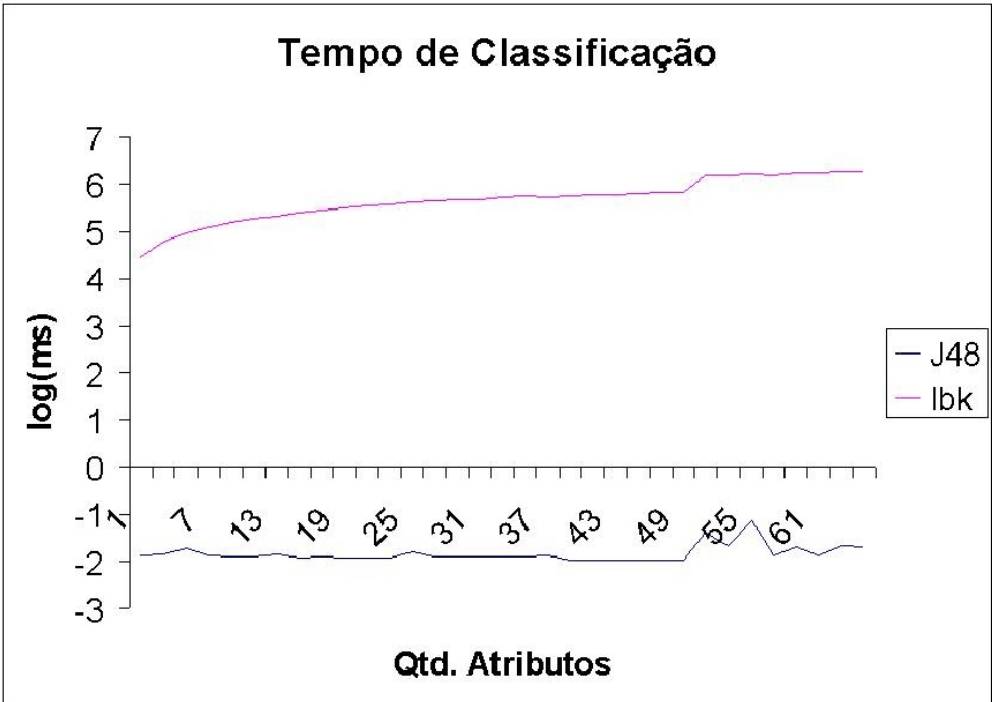


Figura 6.3: Tempo de classificação (ms)

Tabela 6.4: Matriz de confusão J48 com (13 atributos)

a	b	c	d	e	f	
1129	5	4	5	160	65	a-Carrapato
7	3070	48	56	0	37	b-Fundo
5	15	630	43	85	0	c-Marca Ferro
15	25	78	274	14	5	d-Risco Aberto
128	0	77	10	638	10	e-Sarna
87	143	0	3	28	781	f-Sem Defeito

Tabela 6.5: Matriz de confusão IBk com (13 atributos)

a	b	c	d	e	f	
1195	2	4	3	136	28	a-Carrapato
5	3061	57	50	0	45	b-Fundo
10	12	648	29	79	0	c-Marca Ferro
15	31	85	260	16	4	d-Risco Aberto
152	0	73	10	619	9	e-Sarna
94	140	0	2	30	776	f-Sem Defeito

Capítulo 7

Considerações Finais

A detecção de defeitos e a classificação do couro bovino, com base em inspeção visual, é uma importante etapa da cadeia produtiva do boi. A automatização desse processo, através de sistemas computacionais, pode torná-lo mais preciso, menos subjetivo, menos sujeito a falhas e mais uniforme.

Com base nos experimentos realizados e resultados gerados, a análise discriminante de Fisher demonstrou um bom desempenho, mostrando que a partir de um conjunto de atributos, podem ser selecionados um sub-conjunto com a mesma eficiência de classificação. Foi também verificado que os módulos implementados facilitaram os experimentos para testes de classificação, reforçando as expectativas de que será possível a implantação de um classificador automático de couros bovinos. A baixa quantidade de imagens e a pequena variação de semelhança dos defeitos, favoreceram o baixo nível de acerto, porém para um experimento ideal, seria necessário um conjunto maior de imagens, em diferentes ambientes.

Durante os experimentos, alguns problemas foram encontrados, como: baixa quantidade de imagens de tipos de defeitos e problemas com um número muito alto de atributos utilizando a técnica de Fisher, provocando assim, problemas de matrizes singulares.

A partir desses resultados alcançados, foi possível abrir novas perspectivas para trabalhos futuros, como: resolver o problema de matriz singular da redução de atributos por Fisher, para uma quantidade de atributos muito alta; Utilizar técnicas de segmentação para a remoção do fundo das imagens do couro bovino para realização de novos experimentos; utilizar a técnica de Fisher em outras aplicações, como: reconhecimento de textura, gestos, faces, entre outros.

Anexo A

Resultados

Tabela A.1: Resultado de acerto na classificação

Qtd. Atributos	J48(%)	IBk(%)	Qtd. Atributos	J48(%)	IBk(%)
1	63.88	60.27	35	85.28	85.39
3	76.33	79.30	37	85.28	85.42
5	77.83	80.47	39	85.26	85.43
7	77.69	79.71	41	85.26	85.36
9	82.05	83.31	43	85.26	85.34
11	84.65	85.65	45	85.26	85.31
13	85.07	85.63	47	85.26	85.29
15	85.06	85.56	49	85.24	85.20
17	85.43	85.59	51	85.19	85.22
19	85.33	85.62	53	85.19	85.16
21	85.31	85.57	55	85.17	85.14
23	85.32	85.67	57	85.17	85.10
25	85.30	85.61	59	85.18	85.10
27	85.36	85.58	61	85.18	85.11
29	85.34	85.52	63	85.18	85.11
31	85.31	85.55	65	85.19	85.10
33	85.31	85.51			

Tabela A.2: Resultado do tempo de treinamento dos classificadores

Qtd. Atributos	J48(ms)	IBk(ms)	Qtd. Atributos	J48(ms)	IBk(ms)
1	0.5187	0.0033	35	6.8039	0.0169
3	1.4436	0.0096	37	6.0201	0.0143
5	1.8451	0.0055	39	6.1107	0.0158
7	2.3969	0.0062	41	6.3190	0.0222
9	2.8341	0.0069	43	6.6712	0.0215
11	3.0159	0.0083	45	6.9349	0.0168
13	2.9939	0.0098	47	7.2745	0.0178
15	3.3519	0.0104	49	7.5467	0.0183
17	3.8121	0.0117	51	15.9060	0.0261
19	4.2727	0.0115	53	18.1355	0.0246
21	4.7173	0.0123	55	19.4431	0.0249
23	5.1411	0.0135	57	20.3589	0.0324
25	5.6087	0.0143	59	18.3857	0.0516
27	5.2697	0.0141	61	19.4025	0.0287
29	5.6844	0.0147	63	21.8179	0.0566
31	5.9941	0.0155	65	22.5771	0.0521
33	6.3655	0.0167			

Tabela A.3: Resultado do tempo de classificação

Qtd. Atributos	J48(ms)	IBk(ms)	Qtd. Atributos	J48(ms)	IBk(ms)
1	0.0131	2.8529	35	0.0125	55.3917
3	0.0137	6.0435	37	0.0129	52.8410
5	0.0188	9.0279	39	0.0108	57.4315
7	0.0135	12.1379	41	0.0108	59.3123
9	0.0128	15.4582	43	0.0108	59.9587
11	0.0127	18.7630	45	0.0108	63.8077
13	0.0139	21.1979	47	0.0109	67.8377
15	0.0119	24.3172	49	0.0110	69.8418
17	0.0124	27.5314	51	0.0413	150.1045
19	0.0120	30.7841	53	0.0217	154.4247
21	0.0121	35.2910	55	0.0759	161.0137
23	0.0121	38.0315	57	0.0134	152.8639
25	0.0159	42.5215	59	0.0209	176.8072
27	0.0124	44.5075	61	0.0133	177.6694
29	0.0123	46.8898	63	0.0213	183.3813
31	0.0127	49.4690	65	0.0196	184.2655
33	0.0127	52.4004			

Anexo B

Defeitos Semelhantes

Tabela B.1: Defeitos semelhantes analisados

Defeito	Defeitos semelhantes
Carrapato	Sarna
Fundo	Risco e Marca Ferro
Marca Ferro	Sarna e Risco
Risco	Marca Ferro
Sarna	Carrapato e Marca Ferro
Sem Defeito	Fundo e Carrapato

Referências Bibliográficas

- [1] C. Amado and A. M. Pires. Uso de técnicas de reamostragem para selecção de variáveis em análise discriminante. *V Congresso Anual da Sociedade Portuguesa de Estatísticas*, 1997.
- [2] W. P. Amorim, R. H. R. Viana, R. C. B. Rodrigues, and H. Pistori. Desenvolvimento de um software de processamento e geração de imagens para classificação de couro bovino. in: Workshop de iniciação científica. *SIBGRAPI, 2006, Manaus. Workshop de Iniciação Científica - SIBGRAPI*, 2006.
- [3] A. J Antunes. Processamento digital de imagen - uma abordagem indisciplinar aplicada a correção de prognósticos meteorológicos. *Universidade Federal de Pelotas - Instituto de Física e Matemática*, 1999.
- [4] R. M. Moraes L. S. Machado B. B. Gnecco, M. C. Cabral. Um sistema de visualização imersivo e interativo de apoio ao ensino de classificação de imagens. *Brazilian Workshop on Virtual Reality*, 2001.
- [5] T. E. Campo. Técnicas de seleção de características com aplicações em reconhecimento de faces. Master's thesis, Universidade de São Paulo - USP, Instituto de Matemática e Estatística - Ciência da Computação, 2001.
- [6] D. Chetverikov. Structural filtering with texture feature based interaction maps: Fast algorithms and applications. In *ICPR96*, pages II: 795–799, 1996.
- [7] D. Chetverikov. Texture analysis using feature-based pairwise interaction maps. *Pattern Recognition*, 32(3):487–502, 1999.
- [8] A. Gomes e V. S. Lário E. Cardoso. Análise da cadeia produtiva de peles e couros no brasil. *Embrapa - Comunicado Técnico*, pages (68):1–4, 2001.

- [9] J. C. Felipe and A. J. M. Traina. Utilizando características de textura para identificação de tecidos em imagens médicas. *Instituto de Ciências Matemáticas e de Computação - Universidade de São Paulo*, 2002.
- [10] D. A. Forsyth. Computer vision a modern aproach. *Prentice Hall*, 2003.
- [11] A. Gomes. Como melhorar a qualidade do couro. *Gado de corte informa*, page 10, 1987.
- [12] A. Gomes. Aspectos da cadeia produtiva do couro no brasil e em mato grosso do sul. *Embrapa Gado de Corte*, pages 61–72, 2002.
- [13] A. Gomes. Avaliação técnica e operacional do sistema de classificação do couro bovino. *Embrapa Gado de Corte, Relatório Final do Projeto (CNPQ)*, 2005.
- [14] Renato Sproeser Ido Michels, Francisco Bayardo. *Couro Bovino*. Editora UFMS, 2003.
- [15] IEL. Instituto euvaldo lodi, cna confederação nacional da agricultura, e sebrae/nacional serviço brasileiro de apoio ás micro e pequenas empresas. estudo sobre a eficiência econômica e competitividade da cadeia da pecuária de corte no brasil. page 398, 2000.
- [16] R. Jobanputra and D.A. Clausi. Texture analysis using gaussian weighted grey level co-occurrence probabilities. In *CRV04*, pages 51–57, 2004.
- [17] R. A. JOHNSON and D. W. WICHERN. Applied multivariate statistical analysis. *New Jersey: Prentice Hall*, page 642, 1982.
- [18] A Lavoura. 60page 105, 2002.
- [19] MAPA. Ministério da agricultura, pecuária e abastecimento. instrução normativa nº 12, de 18 de dezembro de 2002. estabelece critérios de classificação para qualificação do couro bovino visando a sua valorização comercial e das outras províncias. diário oficial da república federativa do brasil. 2002.
- [20] E. Martins, P. Azevedo-Marques, L. Oliveira, R. Pereira Jr, and C. M. Trad. Caracterização de lesões intersticiais de pulmão em radiograma de tórax utilizando análise local de textura. *Radio Bras*, 2005.
- [21] W. S. Marçal. Considerações técnicas sobre o couro bovino. *Universidade Estadual de Londrina*, 2005.

-
- [22] D. F. MORRISON. Multivariate statistical methods. *New York: McGraw - Hill*, page 495, 1976.
- [23] E.O Nunes and A. Conci. Segmentação por textura e localização do contorno de regiões em imagens multibandas. *UFF*, 2000.
- [24] J. A. Olivera, L. V. Dutra, and C. D. Renn. Aplicações de métodos de extração e seleção de atributos para classificação de regiões. *Instituto Nacional de Pesquisas Espaciais - INPE*, 2005.
- [25] M. A. Pereira and M. A. C. Jacinto. Qualidade e classificação do couro bovino em mato grosso do sul. *42ª Reunião Anual da Sociedade Brasileira de Zootecnia*, 2005.
- [26] H. Pistori. Plataforma de apoio e desenvolvimento de sistema guiadas por sinais visuais. <http://www.gpec.ucdb.br/sigus>, 2004.
- [27] H. Pistori and M. C. Pereira. Integração dos ambientes livres weka e imagej na construção de interfaces guiadas por sinais visuais. In *Anais do V Workshop de Software Livre - WSL*, Porto Alegre, RS, 2-5 de Junho 2004.
- [28] J. C. B. Queiroz, L. Costa, and H. P. Moura. Caracterização de indivíduos por métodos multivariados. *Departamento de Estatística - Universidade Federal do Pará-UFPA*, 2005.
- [29] L. Ramirez, W. P. Amorim, and P. S. Martins. Comparação de matrizes de co-ocorrência e mapas de interação na detecção de defeitos em couro bovino. *SIBGRAPI, 2006, Manaus. Workshop de Iniciação Científica - SIBGRAPI*, 2006.
- [30] A. R. Rocha and N. J. Leite. Classificação de texturas a partir de vetores de atributos e função de distribuição de probabilidades. *UNICAMP, Universidade Estadual de Campinas - Instituto de Computação*, 2002.
- [31] Salvatore Ruggieri. Efficient C4.5. (TR-00-01), 2, 2000.
- [32] R. J. B. Sampaio. Estudo sobre o cálculo de autovalores e autovetores de matrizes simétricas dada por produtos de matrizes, com o uso de matrizes cíclicas, e sua aplicação no problema de correlação canônica. *Pontifícia Universidade Católica do Paraná, Departamento de Engenharia de Produção*, 2005.

- [33] J. C. Santos, J. R. F. Oliveira, and L. V. Dutra. Uso de algoritmos genéticos na seleção de atributos para classificação de regiões. *Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos - SP - Brasil*, 2004.
- [34] O. J. S. Santos and A. Z. Milioni. Composição de especialistas locais para classificação de dados. *Instituto Tecnológico de Aeronáutica (ITA) - Divisão de Engenharia Mecânica-Aeronáutica - São José dos Campos - Brasil*, 2005.
- [35] W. R. Schwartz and H. Pedrini. Métodos para classificação de imagens baseada em matrizes de co-ocorrência utilizando características de textura. *Anais do III Colóquio Brasileiro de Ciências Geodésicas*, 2002.
- [36] J. A. Souza. *Reconhecimento de Padrões usando Indexação Recursiva*. PhD thesis, Universidade Federal de Santa Catarina, Departamento de Engenharia de Produção e Sistemas, 1999.
- [37] Sergios Theodoridis and Konstantinos Koutroumbas. *Pattern recognition*. Academic press, 1999.