



Ministério da Educação
Universidade Tecnológica Federal do Paraná
Pró-Reitoria de Pesquisa e Pós-Graduação

Relatório Final de Atividades

Classificação de bioimagens utilizando descritores de forma **Vinculado ao projeto** **Métodos e técnicas para exploração e análise de bioimagens**

Anderson Brilhador
Voluntário

Tecnologia em análise e desenvolvimento de sistemas da informação

Prof. Dr. Fabrício Martins Lopes
Orientador

Prof. Dr. Pedro Henrique Bugatti
Co-orientador

Área do Conhecimento: Sistema de computação

CAMPUS CORNÉLIO PROCÓPIO, 2013

**ANDERSON BRILHADOR
FABRÍCIO MARTINS LOPES
PEDRO HENRIQUE BUGATTI**

RELATÓRIO FINAL PIBITI

Relatório Técnico do Programa de Iniciação
Tecnológica da Universidade Tecnológica Fe-
deral do Paraná.

CORNÉLIO PROCÓPIO, 2013

Sumário

Introdução	2
Materiais e Metodologia	2
Imagens digitais	2
Vizinhos de um pixel	3
Histograma	3
Filtros espaciais de suavização	4
Filtro da mediana	4
Segmentação de imagens	4
Segmentação de imagens	4
Representação e Descrição	7
Código da cadeia	7
Assinatura	9
Descritores de Fourier	10
Reconhecimento de Objetos	11
Classificação	12
Validação e Análise da precisão	12
Validação cruzada	12
Matriz de confusão	13
Precisão e <i>Recall</i>	13
<i>F-Measure</i>	13
Definição de características	14
Descrição da base de imagens	14
Implementação dos extratores	15
Resultados e discussões	16
Experimento	16
Conclusões	17

INTRODUÇÃO

Nos últimos anos, por conta do avanço tecnológico e popularização de dispositivos de mídia como microcomputadores, celulares e *tablets*, ocorreu uma verdadeira explosão na quantidade de dados disponíveis, prontos para serem analisados em virtude de gerarem informações importantes para um negócio.

Este novo cenário não se refere somente a volume de dados, mas também a velocidade, recuperação, complexidade e variedade de tipos de informação disponibilizadas, dentre elas arquivos multimídias (imagens e vídeos).

Devida a grande quantidade de informações, esses arquivos multimídias não estão disponibilizados de maneira estruturada, portanto antes de analisar esse conteúdo criou a necessidade de armazenar, indexar e recuperar essas imagens de forma organizada, assim grupos acadêmicos/científicos e empresas privadas se interessaram em desenvolvimento de tecnologias comportassem essas propriedades.

Parte desse conteúdo está relacionado a bioimagens, por exemplo a flora brasileira representa cerca de 19% da flora mundial com mais de 56.000 espécies (GIULIETTI et al., 2005), esse tipo de informação é de enorme valor ao homem, pois então ligadas a diversos nichos de mercado como alimentação, medicina, vestuário, habitação e atividade industrial.

O conhecimento exato de uma espécie de planta, distribuição geográfica e a utilização, é essencial, e raramente está disponível de forma prática, sendo necessário recorrer a grandes catálogos de espécies o que limita o conhecimento de diversos profissionais, pois a variabilidade entre as plantas possui proporções muito maiores que a biologia animal. Por isso, é necessário o uso de ferramentas auxiliares na identificação e recuperação dessas informações. Essa identificação e recuperação é realizada por meio da extração de características visuais como cor, forma e textura.

MATERIAIS E METODOLOGIA

Neste capítulo será apresentado os conceitos e técnicas de processamento de imagens digitais estudadas para o desenvolvimento do *framework* de extração de características proposto neste trabalho.

Imagens digitais

Uma imagem digital refere-se a uma função bidimensional, $f(x,y)$, em que x e y são as coordenadas espaciais de um plano, e o valor de f em qualquer ponto das coordenadas (x,y) representa a *intensidade* ou *nível de cinza* da imagem presente naquele ponto. Quando x,y e os valores de intensidade de f são quantidades finitas e discreta chamamos de *Imagem Digital*. Esses pontos possuem localização e valores específicos dentro da imagem e são chamados de *pixels* (contração do inglês de *picture element*) termo mais utilizado para representar elementos de uma imagem (GONZALEZ; WOODS, 2006). A Figura 1 mostra a representação de uma imagem discreta e sua subdivisão de pixels.

O nível de intensidade em um pixel varia entre 0 e 255 em imagem digital de 8 bits, essa imagem possui apenas tons de cinzas em suas tonalidades. Um dos modelos para representação de imagens coloridas é o modelo RGB (acrônimo do inglês *Red, Green e Blue*) aonde o número de bits para representar um pixel passar a ser chamado de *profundidade de pixel*. A profundidade de pixel em uma imagem RGB é de 24 bits, no qual um pixel corresponde a um trio de

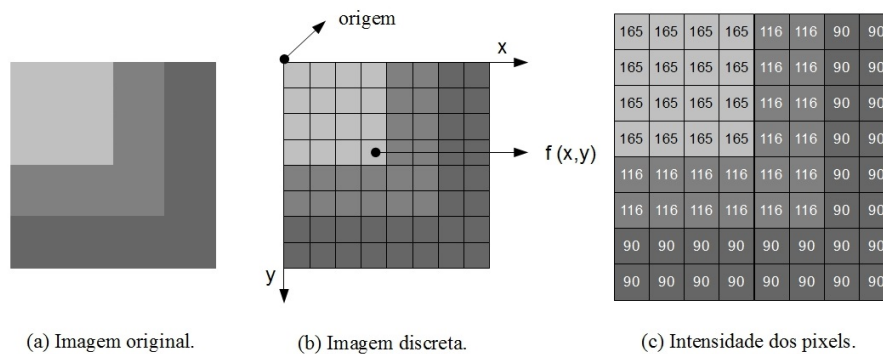


Figura 1: Descrição dos valores dos *pixels* de uma imagem digital.

valores $[(R, G, B)]$, aonde cada valor representa uma imagem de 8 bits em sua tonalidade. A junção dessas três imagens formam uma imagem RGB.

Para transformar uma imagem do modelo RGB para uma imagem de níveis de cinzas é necessário unir os canais para que os mesmo formem um só, para isso pode se utilizar de uma aproximação simples que é a soma dos valores de cada canal RGB pertencentes a um pixel dividido pelo quantidade de canais, o resultado é tonalidade de cinza daquele pixel, o processo deve ser realizado em toda imagem percorrendo cada pixel que deseja ser convertido.

Vizinhos de um pixel

Existem três tipos de vizinhanças entre os pixels: Conectividade 4 no qual um pixel p na coordenada (x,y) tem quatro vizinhos considerando que os pixels vizinhos podem estar na vertical, na horizontal e até mesmo na diagonal do pixel central. Caso p esteja na borda da imagem alguns destes vizinhos podem não pertencer a imagem; Conectividade 8 na qual a vizinhança considera os oito pixels em todas as direções do pixel central e a de Conectividade M na qual é modificado a conectividade de 8 produzindo conexões de formas diferentes.

Conhecer a conectividade entre os pixels, auxilia em diversas técnicas de processamento de imagem como determinação de rótulos em objetos e regiões pertencentes a imagem, através das características comuns extraídas dos pixels vizinhos. Após definir os rótulos é possível calcular distancias, enumerar a quantidade de elementos e até mesmo segmentar objetos da imagem.

Histograma

De modo geral, o *Histograma* de uma imagem digital é o agrupamento dos pixels que possuem a mesma intensidade. A sua construção é dada por uma função discreta $h(r_k) = n_k$ onde r_k é o K -ésimo valor de intensidade e n_k é o numero de pixels da imagem com intensidade r_k (GONZALEZ; WOODS, 2006).

O histograma pode ser utilizado para diferentes finalidades como realce de imagens digitais, compressão, segmentação e classificação. Aplicar somente o histograma para caracterização de uma imagem não é o indicado, pois o mesmo não armazena informações espaciais e contém informações redundantes pelo grande número de valores gerados em sua descrição (LOPES, 2003).

Pelo nível de complexidade baixa em sua implementação, o histograma é uma técnica muito utilizada em aplicações que necessitem de economia de *hardware* e para processamento de

imagens em tempo real.

Filtros espaciais de suavização

Os filtros de suavização tem como resultado da aplicação em uma imagem digital um efeito de borramento e redução de ruídos. Esses filtros são aplicados com objetivo de remoção de pequenos detalhes da imagem antes da extração de objetos e conexão de pequenas discontinuidades em linhas e curvas (GONZALEZ; WOODS, 2006). Existem dois tipos de filtros espaciais de suavização que são os filtros filtragem linear e não linear.

Os filtros lineares são implementados por meio do calculo da média dos pixels contidos na vizinhança da máscara de filtragem. Esse filtros tem como efeito minimizar os ruídos sem alterar o nível médio de intensidades da imagem.

Enquanto os filtros não lineares são obtido por medidas estatísticas, como a mediana dos pixels presentes na vizinhança definida pela máscara. Seus resultados em relação a redução de ruídos são comumente superiores as filtros da média, mas sua aplicação altera o nível médio da imagem.

Filtro da mediana

O filtro da mediana consiste em substituir a intensidade do pixel central pela mediana dos pixels capturados pela máscara aplicada. Esse é o filtro de ordem estática mais útil no processamento de imagens, pois para certo tipos de ruídos como o conhecido *ruído de sal e pimenta* essa técnica proporciona excelentes resultados na redução de ruídos e preservação da agudeza de bordas e linhas (GONZALEZ; WOODS, 2006), fator importante para extração de objetos da imagem.

Para implementar a filtragem pela mediana em um pixel da imagem, deve primeiro ordena os valores dos pixels da vizinhança, de forma incremental ou decremental, incluindo o pixel central, então calcula-se a mediana dos pixels e atribui o valor ao pixel central, como por exemplo uma vizinhança de 3 x 3, a mediana é o 5º valor ordenado e de uma vizinhança de 5 x 5 é o 13º valor ordenado. Assim a principal função do filtro da mediana é obrigar que intensidades distintas se assemelhem com sua vizinhança.

Segmentação de imagens

Nesta seção serão apresentados os métodos de segmentação de imagens estudados e aplicados no conjunto de imagens, afim de separar o objeto do fundo para posterior análise.

Limiarização

É uma técnica de segmentação simples e computacionalmente rápida, por essas propriedades a limiarização é muito utilizada em aplicações que identificam e extraem objetos de uma imagem. A base da limiarização está relacionado ao histograma de intensidades explicado na seção .

O histograma da Figura 2 representa uma imagem, $f(x,y)$, composta por objetos claros sobre um fundo escuro, está composição gera um histograma com dois grupos de intensidades agrupados. A forma obvia de extrair os objetos do fundo e selecionar um limiar T que separe os

dois grupos. Então qualquer ponto (x, y) na imagem em que $f(x, y) > T$ é chamado de *ponto do objeto* e o caso contrario é chamado de *ponto de fundo*, isto é expressado pela seguinte forma:

$$g(x, y) = \begin{cases} 1 & \text{se } f(x, y) > T \\ 0 & \text{se } f(x, y) < T \end{cases} \quad (1)$$

Neste caso foi utilizado a convenção de usar 1 para objeto e 0 para fundo, mas 2 valores distintos quaisquer podem ser utilizados na função 1.

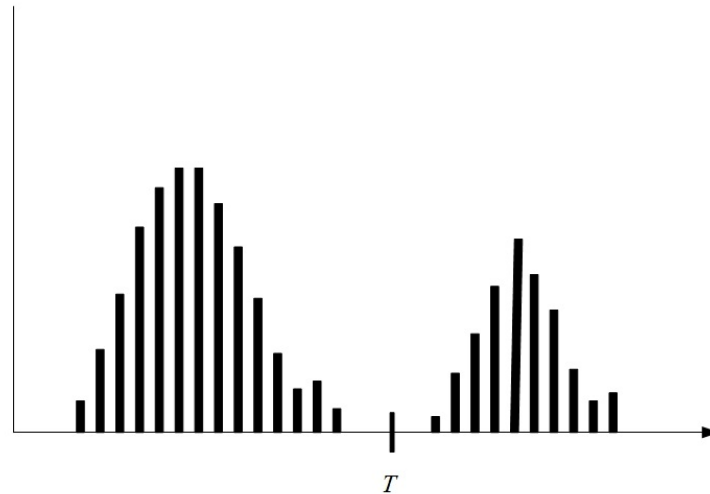


Figura 2: Histogramas de intensidades sendo dividido por um único limiar T .

Essa técnica possui variações e foi classificadas em dois grupos: *globais* e *locais* por Sahoo et al. (1988). Os limiares globais são o que dividem a imagem utilizando apenas um limiar. Os locais dividem a imagem em subimagem e determinar um limiar para cada subimagem.

Observando o método é fácil identificar que o sucesso da técnica está diretamente ligada a largura e a profundidade dos vales do histograma. Por sua vez, existem vários fatores que interferem nestas condições como índice de ruídos na imagem, o tamanho do objetos que deseja ser extraído e a uniformidade da luz e reflexão da imagem (GONZALEZ; WOODS, 2006).

Tendo em vista a dificuldade de identificação de um limiar e sendo impossível este reconhecimento ser realizado por uma pessoa em uma larga escala de imagens, foram proposto os métodos de limiarização global automáticos que tem como objetivo estabelecer um limiar para uma imagem de forma eficiente minimizando a taxa de erro entre a separação dos picos do histograma.

Método de Otsu

O método de Otsu (OTSU, 1979) é considerado um método ótimo para seleção automática de um limiar pois consegue maximizar a variância entre as classes, no caso separar objeto do fundo da imagem.

A separação é feita pela definição de um limiar que particione os pixels de uma imagem em duas classes C_0 sendo o fundo e C_1 o objeto, ou vice versa. Para isso é necessário maximizar uma das funções critérios exibidas na equação 2:

$$\lambda = \frac{\sigma_B^2}{\sigma_W^2} \quad \kappa = \frac{\sigma_T^2}{5\sigma_W^2} \quad \eta = \frac{\sigma_B^2}{\sigma_T^2} \quad (2)$$

mas primeiro deve calcular a probabilidade da distribuição dos níveis de *pixels* como definido na equação 3:

$$p_i = \frac{n_i}{N}, \quad (3)$$

onde n_i é o número de pixels no level i e N é o total de *pixels* da imagem. Com a probabilidade dos pixels, calcula-se a média total para todos os níveis de cinzas $[1, \dots, L]$ dada por

$$\mu_T = \mu(L) = \sum_{i=1}^L i p_i \quad (4)$$

e a variância total definida como

$$\sigma_T^2 = \sum_{i=1}^L (i - \mu_T)^2 p_i. \quad (5)$$

Em relação a classes (C_0 e C_1) são calculadas as probabilidades da ocorrências dos pixels (6) e (7) e a média dos níveis de cinza (8) e (9), utilizando as seguintes equações:

$$\omega_0 = Pr(C_0) = \sum_{i=1}^k p_i = \omega(k), \quad (6)$$

$$\omega_1 = Pr(C_1) = \sum_{i=k+1}^L p_i = 1 - \omega(k) \quad (7)$$

e

$$\mu_0 = \sum_{i=1}^k i Pr(i|C_0) = \frac{\sum_{i=1}^k i p_i}{\omega_0} = \frac{\mu(k)}{\omega(k)}, \quad (8)$$

$$\mu_1 = \sum_{i=1}^k i Pr(i|C_0) = \sum_{i=k+1}^L \frac{i p_i}{\omega_1} = \frac{\mu_T - \mu(k)}{1 - \omega(k)} \quad (9)$$

onde

$$\omega(k) = \sum_{i=1}^k p_i \quad (10)$$

e

$$\mu(k) = \sum_{i=1}^k i p_i, \quad (11)$$

onde a equação 11 é a somatória dos momentos do histograma de $[1, \dots, k]$.

Como foi descrito anteriormente, o limiar ótimo é o nível k que maximiza η . Caso essa máxima não seja única é possível aplicar uma forma alternativa (k^*) buscando de forma seqüencial em $[1, \dots, L]$ diversos valores máximos que k corresponde, utilizando a equação 12:

$$\sigma_B^2(k) = \frac{[\mu_T \omega(k) - \mu(k)]^2}{\omega(k)[1 - \omega(k)]}, \quad (12)$$

portanto o limiar k^* é definido como:

$$\sigma_B^2(k^*) = \max_{1 \leq k < L} \sigma_B^2(k). \quad (13)$$

Após o processo de definição do limiar é realizada a limiarização substituindo o valor de T .

Representação e Descrição

A representação e descrição de características de uma imagem são composta por três propriedades base textura, cor e forma. Essas representações estão diretamente ligadas a fase de segmentação de imagens como por exemplo extração da forma, tem a necessidade que a forma do objeto tenha sido segmentada para que o processo de caracterização seja efetuado.

Costa e Jr. (2009) propõem uma possível classificação das abordagens para representação de forma como poder ser visto na Figura 3 a forma é dividida em três abordagens: contorno, região e transformada.

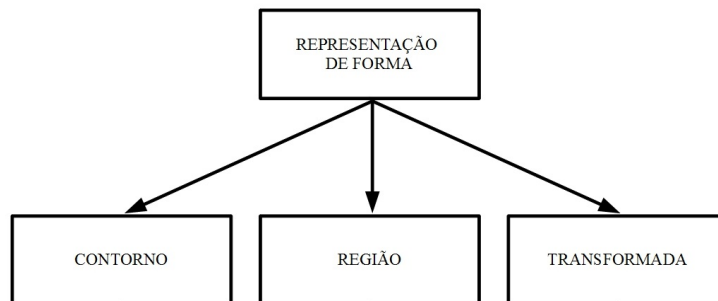


Figura 3: Classificação proposta por Costa e Jr. (2009) para técnicas de representação de formas.

Fonte: (COSTA; JR., 2009).

As abordagens baseadas em contorno tem como característica uma implementação simples e um custo computacional baixo, mas precisam ser invariantes a rotação, escala e translação. Algumas dessas técnicas de extração de contorno serão tratadas nas próximas seções.

Código da cadeia

Essa técnica foi proposta por Freeman (1961) e é considerada uma técnica simples para representação de contorno (COSTA; JR., 2009), seu objetivo é representar uma fronteira por meio de uma codificação numérica de valores os quais representam a conectividade entre cada segmento pertencente à fronteira. Essa representação é baseada em direções que tipicamente variam entre 4 ou 8 segmentos como na Figura 4.

Para aquisição dessa representação, as imagens digitais são adquiridas e processadas no formato de uma matriz contendo espaçamento igual nas direções de x e y , isto possibilita que o código da cadeia percorra toda a fronteira seguindo as coordenadas (x, y) em um sentido horário ou anti-horário a partir de um ponto inicial, o qual deve pertencer a fronteira do objeto.

Ao percorrer uma coordenada atribui-se o valor da direção que conecta um par de pixels ao código da cadeia, ou seja, o código da cadeia é a concatenação de todas as direções pertencentes a fronteira de um objeto ou região, como pode ser observado na Figura 5.

Esse método se torna inaplicável por dois motivos principais: a cadeia de código é longa e sua representação é super sensível ao ruído e a falhas de segmentação, atribuindo valores de

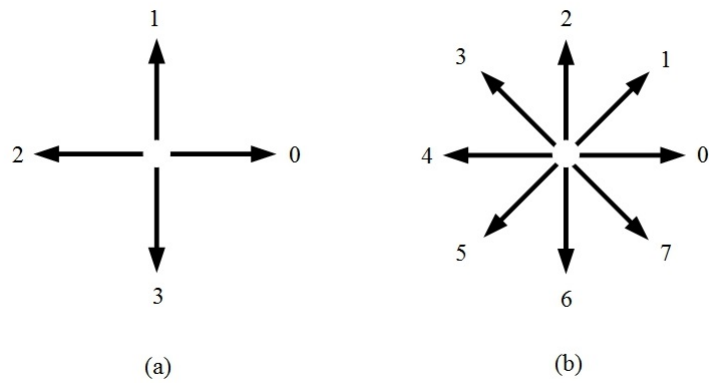


Figura 4: Exemplo de direções de segmentos, (a) cadeia direcional-de-4 e (b) cadeia direcional-de-8

Fonte: (GONZALEZ; WOODS, 2006).

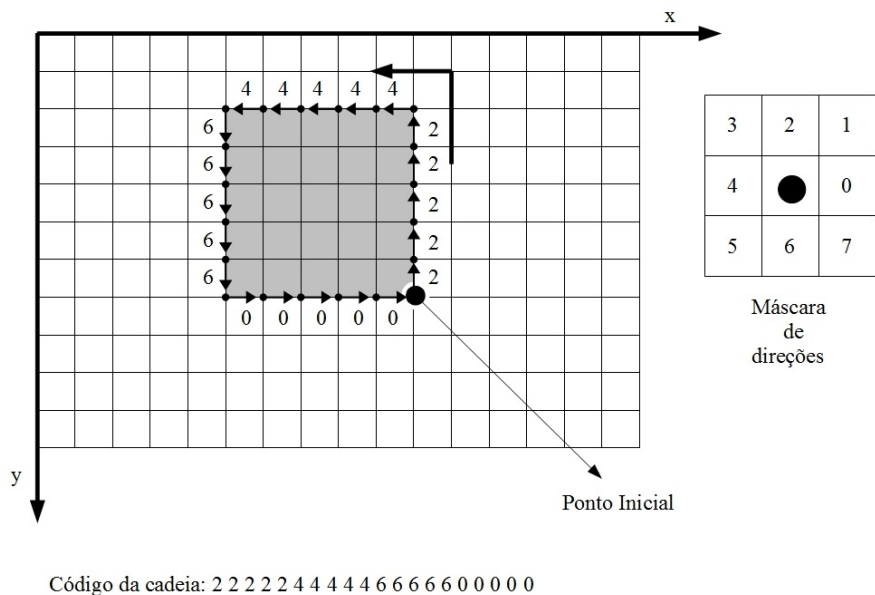


Figura 5: Representação do código da cadeia de um contorno simples.

Fonte: Adaptado de (COSTA; JR., 2009).

direções não relacionadas à forma da fronteira. Uma abordagem utilizada para corrigir esses problemas é a reamostragem da fronteira por uma grade de espaçamento maior (GONZALEZ; WOODS, 2006).

O código da cadeia é dependente do ponto inicial, pois o mesmo interfere diretamente na representação do código da cadeia, para normalizar esses valores existem métodos simples como a substituição do ponto inicial arbitrário de uma sequência de valores para um ponto inicial pertencente a sequência que forme um inteiro de menor magnitude. Para normalizar a rotação do código da cadeia utiliza-se a diferença entre os pares de direções, considerando que a sequência de valores é circular, o primeiro elemento da sequência será calculado com o ultimo elemento, o resultado é um novo código da cadeia baseado na diferença entre as direções e normalizado em relação rotação. Mas essas normalizações só alcança um resultado satisfatório caso as fronteiras

já forem invariante a rotação e a mudança de escala (GONZALEZ; WOODS, 2006).

Outra forma de representação utilizando o código da cadeia é o histograma de direções que traz informações relacionadas a variação da fronteira do objeto. Também pode se utilizar o procedimento do código da cadeia para obter a sequência de coordenadas que formam a fronteira da região de um objeto, com essa lista de coordenadas é possível calcular por exemplo o centróide e área da forma segmentada.

Assinatura

É a representação de uma fronteira de forma unidimensional, a sua geração é baseada na fronteira bidimensional mais complexa para ser descrita. Essa representação pode ser gerada de diversas maneiras, uma das mais simples é obtida pelo gráfico de distâncias da fronteira em função do ângulo (GONZALEZ; WOODS, 2006), como pode ser observado na Figura 6.

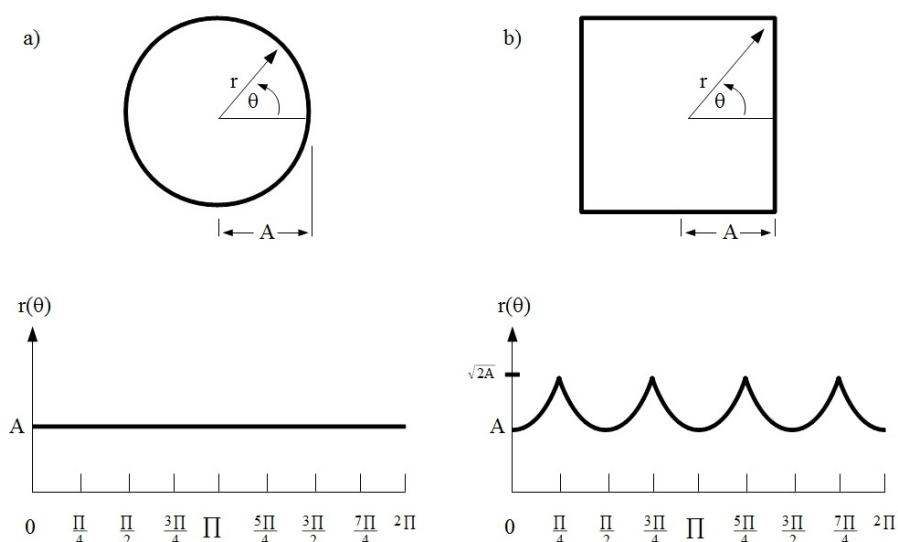


Figura 6: Assinaturas de distância em função do ângulo. Em (a), $r(\theta)$ é constante. Em (b), a assinatura consiste de repetições de forma padrão.

Fonte: (GONZALEZ; WOODS, 2006).

O exemplo de assinatura apresentado na Figura 6 é invariante à translação e dependente da rotação e a escala em relação a fronteira extraída da imagem digital. Para normalização em relação à rotação é possível aplicar uma técnica simples que tem como princípio selecionar o mesmo ponto para gerar a assinatura. Uma maneira simples é garantir que o ponto inicial seja o ponto mais distante do centróide, outra alternativa é extrair o código da cadeia e aplicar a normalização apresentada na seção do código da cadeia .

Considerando que as mudanças de escala sejam uniformes em relação ao eixo pertencente a fronteira e que as amostragens capturadas possuem o mesmo tamanho de intervalo entre elas, então o resultado da assinatura será proporcionalmente diferente em relação a escala das duas imagens. Para remover as dependências de tamanho em relação a escala, pode se adotar duas técnicas a primeira mais simples consiste em normalizar os valores representados na assinatura para o mesmo domínio entre 0 e 1, a segunda mais robusta consistem em dividir cada amostragem pela variância da assinatura (GONZALEZ; WOODS, 2006).

Após extrair a assinatura pode ser utilizar medidas de dispersão: como média, desvio padrão e variância, para quantificar a informação dispersa do formato do objeto, como por exemplo um

circulo possui baixa dispersão em relação a distância do centróide.

Existem outros métodos para obter a assinatura da fronteira como exemplo percorrer a fronteira e criar um gráfico em relação ao ângulo entre a linha tangente à fronteira e uma linha de referência ao longo da fronteira, uma variação dessa abordagem é conhecida como função de densidade de inclinação (*slope density function*) que é o gráfico dos ângulos das tangentes (GONZALEZ; WOODS, 2006).

Descritores de Fourier

Um descritor de forma amplamente utilizado é o descritor de Fourier, seus descritores são formados pelos coeficientes da transformada discreta de Fourier (DFT). O descritor de Fourier tem como característica expressar as informações globais da curvatura do objeto extraído da imagem.

A transformada discreta de Fourier basea-se no teorema de Fourier que têm como princípio que toda função contínua pode ser decomposta em uma função de frequências de senos e cossenos. Os dados de entrada de uma DFT são sequências de números discretos ou contínuos (GONZALEZ; WOODS, 2006).

A seqüência de números podem ser obtidas de um plano cartesiano através das coordenadas do contorno do objeto. Essas coordenadas podem ser representadas pela função unidimensional definida como:

$$s(k) = x(k) + jy(k) . \quad (14)$$

Sendo que $K = 0,1,2,\dots,N - 1$ é número de pontos do contorno, x é a coordenada real e y a coordenada imaginária de um plano cartesiano. Pode se utilizar a distância do centróide para forma o eixo real da função.

A transformada discreta de Fourier para o elemento $s(k)$ é definida como:

$$a(u) = \frac{1}{N} \sum_{k=0}^{N-1} s(k) \exp[-j2\pi uk]/N . \quad (15)$$

Os coeficientes $a(u)$ são chamados de descritores de Fourier do contorno, por meio destes coeficientes é possível aplicar a transformada inversa de Fourier para reconstruir $s(k)$:

$$s(k) = \sum_{u=0}^{N-1} a(u) \exp[j2\pi uk]/N . \quad (16)$$

Os descritores de Fourier de baixa frequência armazenam a essência da forma de um objeto, enquanto os descritores de altas frequência capturam os detalhes mais finos da forma. Então para reconstruir a forma de um objeto não são necessários todos os coeficientes de Fourier, podendo considerar apenas os coeficiente de baixa frequência para sua reconstrução.

Esses descritores devem ser insensíveis a translação, rotação e diferentes escalas, os descritores por si só não são insensíveis as essas mudanças, mas é possível aplicar transformações simples nos descritores para obter tal resultado.

Para que seja possível realizar a comparação entre o contorno de objetos semelhantes os descritores de Fourier devem ser invariantes a translação, rotação e escala. Essas invariâncias são fáceis de serem obtidas por meio de transformações simples aplicadas aos coeficientes de Fourier. Essas transformações são apenas aproximações da normalizações e apesar de Bartolini, Ciaccia e Patella (2005) demonstrar que usando o ângulo da fase e amplitude obtêm resultados

melhores o seu custo computacional ainda é alto, por isso foi adotado por utilizar apenas a amplitude.

Os coeficientes de Fourier são a somatório dos números complexos pertencentes ao vetor S_k que contem as coordenadas do contorno. A transformada discreta de Fourier é representada na forma:

$$\sum_{k=0}^{N-1} (s_k + s) \exp[-j2\pi uk]/N = \sum_{k=0}^{N-1} (s_k) \exp[-j2\pi uk]/N + (s) \sum_{k=0}^{N-1} \exp[-j2\pi uk]/N . \quad (17)$$

Esse somatório representa que o primeiro elemento do vetor não é transladado enquanto o segundo é:

$$\sum_{k=0}^{N-1} \exp[-j2\pi uk]/N = \begin{cases} 1 - \exp[-j2\pi uk] \\ 1 - \exp[-j2\pi uk]/N \end{cases} = 0, N-1, k \neq 0, k = 0 . \quad (18)$$

Isto indica que a translação afeta somente o primeiro coeficiente da transformada de Fourier, então a invariância à translação pode ser obtida eliminando o elemento S_0 . Se entrada da função for uma sequência de distâncias do centróide não é necessário aplicar essa normalização pois a assinatura baseada na distância do centróide já é invariante a translação.

Considerando θ é o ângulo do objeto e que este pode ser descrito pela multiplicação de cada elemento do vetor S_k . A transformada de Fourier é definida como:

$$\sum_{k=0}^{N-1} (S_k \exp^{\theta u}) \exp[-j2\pi uk]/N = \exp^{\theta u} \sum_{k=0}^{N-1} (s) \exp[-j2\pi uk]/N . \quad (19)$$

Como $|\exp^{-\theta u}| = 1$ pode considerar apenas o valor absoluto do elemento para obter a invariância à rotação.

A escala pode ser descrita por uma constante c multiplicadora do vetor s_n , então a transformada de Fourier pode ser definida como:

$$\sum_{k=0}^{N-1} (cs_k) \exp[-j2\pi uk]/N = c \sum_{k=0}^{N-1} (s_k) \exp[-j2\pi uk]/N . \quad (20)$$

Com esse principio pode se obter a invariância definindo a constante c como um valor absoluto de um elemento pertencente ao vetor S . A invariância a translação é obtida eliminando o elemento S_0 do vetor S , então esse elemento não pode ser utilizado para normalização da escala. Um padrão utilizado para realizar essa normalização é dividir todos os elementos do vetor S por S_1 , S_1 é considerado a frequência fundamental (RAFIEI; MENDELZON, 2002) e possui a maior amplitude quando o contorno é capturado no sentido anti-horário (WALLACE; WINTZ, 1980). Caso se utilize a distância do centróide para representar o contorno, o elemento a ser usado é o S_0 pois o mesmo não sofre influência de translação.

Reconhecimento de Objetos

O **reconhecimento** é um processo que tem como objetivo atribuir um rótulo a um objeto (também conhecido como padrões) com base nos descritores (GONZALEZ; WOODS, 2006). Esses descritores devem caracterizar um objeto de maneira que se torne distinto de outros padrões existentes dentro do contexto do trabalho.

Para que o reconhecimento destes padrões seja realizado por uma máquina, é necessário uma técnica de atribuição de padrões para às suas respectivas classes sendo realizada de forma

automática ou semi-automática. Essa atribuição é composta por um vetor de características (*feature vector* ou *pattern vector*) representado da seguinte maneira:

$$x = [x_1, x_2, \dots, x_n], \quad (21)$$

onde x_i representa o i -ésimo descritor e n é o número total de descritores relacionados ao padrão.

O vetor de características têm influencia importante no desempenho final do reconhecimento de objetos, pois a escolha dos descritores que compõem o vetor devem conter características com um nível elevado de separabilidade entre as *classes* (família de padrões que compartilham propriedades comuns).

Portanto para avaliar o grau de separabilidade das classes por meio dos descritores aplica-se métodos de classificação aos vetores de características, após essa classificação é realizada uma análise dos resultados para avaliar o grau de separabilidade e desempenho dos classificadores. Essas técnicas serão abordadas nas próximas seções e subseções.

Classificação

A classificação é um processo de análise do vetor de características que permite extrair modelos para obter uma classe. Esse processo pode ser dividida em duas fases: a primeira fase é conhecida como treinamento na qual o objetivo é determinar um grupo de classes por meio de atributos de entradas (HAN; KAMBER, 2006). Dentro desta fase é possível fragmentá-la em 2 estilos: **aprendizagem supervisionada** - na qual as instâncias a serem classificadas contém classes de exemplos, assim a classificação é gerada de forma que o resultado chegue o mais próximo possível das classes de exemplo e a **aprendizagem não supervisionada** as instâncias não possuem classes de exemplos, sendo necessário o uso de raciocínio probabilístico para o conhecimento das classes (DUDA; HART; STORK, 2001).

A segunda fase consiste em utilizar o modelo criado na primeira fase para rotulação de novas instâncias. Para essa classificação é possível utilizar diversos métodos. Os métodos utilizados neste trabalho são *Instance-based learning algorithms* (IBK) (AHA; KIBLER, 1991), *Naive Bayes* (NB) (JOHN; LANGLEY, 1995), J48 classificador C4.5 (QUINLAN, 1993), *Random Forest* (STATISTICS; BREIMAN, 2001), *MultiLayer Perceptron* (RUSSELL et al., 1996) e *Support Vector Machines* (SVM) (ABE, 2010).

Validação e Análise da precisão

Afim de validar e analisar os resultados dos classificadores executados em conjunto com as características geradas pelo trabalho, foi necessário utilizar métodos para validação das classificações e posteriormente métodos para análise dos resultados gerados. Os métodos utilizados serão apresentados nas próximas subseções.

Validação cruzada

O método de validação cruzada (*Cross-validation*) permite utilizar todos os elementos disponíveis, não tendo necessidade de separar amostras para fase de treinamento ou classificação final. Segundo Santos et al. (2005) a validação cruzada consiste em dividir as amostras (P) em K subconjuntos mutuamente exclusivos (P_1, P_2, \dots, P_k) e de dimensões aproximadamente

iguais (*K-folds*). Os modelos são executados K vezes pelo o algoritmo de classificação e cada execução do modelo é reservada uma parte P_i para ser realizada a classificação. A acurácia ou erro final é dada pela média dos resultados das execuções. O valor de K pode variar entre 2 e N , mas o valor 10 é o mais utilizado (*Ten-Fold Cross-Validation*).

Após a classificação é necessário analisar e validar os resultados, por meio de técnicas como matriz de confusão, e análise de precisão e *recall*.

Matriz de confusão

A matriz de confusão é uma matriz de erros gerados por um classificador, onde os dados variam entre verdadeiro positivo, falso positivo, verdadeiro negativo e falso negativo, como pode ser observado na Tabela 1.

Tabela 1: Exemplo de matriz de confusão.

Classes Reais	Classes Preditas	
Classe 1	Verdadeiro positivo (VP)	Falso negativo (FN)
Classe 2	Falso positivo (FP)	Verdadeiro negativo (VN)

A partir desta matriz é possível extrair métricas de qualidade para as classificações como acurácia e taxa de erros:

$$Acurácia : Acc = \frac{vp + vn}{vp + fp + vn + fn} , Taxa de erro : E = \frac{fp + fn}{vp + fp + vn + fn} . \quad (22)$$

Essas duas métricas são muito utilizadas para análise de desempenho, mas quando as instâncias entre as classes não estão balanceadas, essas medidas podem inferir dados enganosos. Para evitar problemas desse porte a base de imagens foi normalizada como descrita na seção descrição da base de imagens.

Precisão e Recall

A precisão e *Recall* (BAEZA-YATES; RIBEIRO-NETO, 1999) são métricas frequentemente utilizadas em trabalhos relacionados ao reconhecimento de padrões. Ao utilizar essas métricas o conjunto de rótulos possíveis é dividido em dois subgrupos, um dos quais é considerado relevante para os objetivos da métrica. Então o *Recall* é calculado como a fração de instâncias corretas entre todas as instâncias que pertencem ao subconjunto relevante e a Precisão é a fração de instâncias correta entre aquelas que o algoritmo julgou pertencer ao subconjunto relevante. A precisão é vista como medida de exatidão ou fidelidade e *Recall* como medida de completude. Assim Precisão e *Recall* foram definidas pelas seguintes equações:

$$Precisão = \frac{vp}{vp + fp} , \quad (23)$$

$$Recall = \frac{vp}{vp + fn} . \quad (24)$$

F-Measure

A métrica *F-Measure* (RIJSBERGEN, 1979) tem como objetivo sintetizar as informações das medidas de Precisão e *Recall*, apresentadas na seção anterior, obtendo dessa maneira a média harmônica ponderada entre elas. Isto é utilizado quando é necessária a análise do desempenho de vários classificadores através de um único valor facilitando e melhorando a comparação entre eles.

A equação da *F-Measure* é dada por:

$$F\text{-Measure} = \frac{(1 + \beta^2) * \text{Precisão} * \text{recall}}{\beta^2 * \text{Precisão} + \text{recall}}, \quad (25)$$

sendo β o coeficiente que ajusta a importância relativa de Precisão *versus* *Recall*, sendo normalmente $\beta = 1$.

Definição de características

O trabalho teve como ênfase a extração de características de forma, nas quais podem ser divididas em 3 tipos: contorno, região e transformadas. Foram utilizados somente descritores de contorno para a representação das bioimagens, os quais são Código da cadeia, Assinatura em função da distância do ângulo e Descritores de Fourier aplicado ao sinal complexo da fronteira.

Descrição da base de imagens

O Conjunto de dados utilizados foi o mesmo usado pelo concurso de identificação de plantas ImageCLEF 2012, que faz parte da CLEF 2012 (GOËAU et al., 2012). Esse conjunto de dados é baseado no *Pl@ntLeaves dataset 2* e todos os dados foram disponibilizados com licença *creative commons*. O conjunto de dados conta com 126 espécies diferentes de árvores francesas localizadas na área do Mediterrâneo, o número total de imagens é de 11.527 que são subdivididas em três grupos de imagens: *scan* (57%), *scan-like* (24%) e *free natural photos* (19%).



Figura 7: Exemplo de imagens de scan.

Fonte: Base de imagens ImageCLEF 2012 (GOËAU et al., 2012).

Para o trabalho realizado foram selecionadas somente as classes de treinamento de imagens de *scan* (Figura 7), que possui em seu total 4870 imagens, divididas de maneira não uniforme entre as 126 espécies. Com objetivo de normalizar essa distribuição aplicou-se a seguinte regra: (1) calcula-se a média aritmética x da quantidade de imagens por espécie e_n , onde n é o número total de espécies. (2) com base nessa média verifica-se a quantidade de imagens q_i

da espécie e_i , onde $i = 0, 1, 2, \dots, N - 1$, é igual ou superior a x . As espécies com q_i inferior a x são desconsideradas, com isso o número de amostragem por espécie é balanceado. Mas o número de espécies e imagens é reduzido, passando a ter 54 espécies diferentes e 3345 imagens de amostragem.

Implementação dos extratores

Após a definição das características, foi inicializada a implementação do *framework* para a classificação das bioimagens. Para isso foi utilizada a linguagem JAVA que proporciona ao software uma propriedade importante, a portabilidade entre os sistemas operacionais, por meio da máquina virtual. Como ferramenta de desenvolvimento foi utilizado o NetBeans IDE versão 7.3.1 e todos os componentes são de uso livre.

Seguindo a implementação notou-se que era necessário filtrar as imagens de entradas, pois as mesmas continham ruídos que poderiam interferir na etapa de segmentação. No momento inicial se utilizou o filtro da média para realizar a redução dos ruídos, mas o resultado do filtro não foi satisfatório, pois dava um efeito de “embaçamento” na imagem eliminando os ruídos e a agudeza das bordas, ponto importante para os algoritmos de segmentação e descrição de forma. Tendo como necessidade manter a nitidez nas bordas, aplicou-se o filtro da mediana e o resultado foi satisfatório utilizando uma máscara de 5 x 5 sobre a imagem.

Com a remoção das interferências externas sobre as imagens, passou para etapa de extração do objeto também conhecida como segmentação, a primeira tentativa de extração do objeto foi aplicando as técnicas de detecção de bordas mas os resultados não foram adequados, considerando que apresentaram pequenas falhas na detecção do contorno tornando impossível a aplicação de outras técnicas, como por exemplo o código da cadeia, que tem como particularidade percorrer as fronteiras do objeto. Buscando outra abordagem para extração de objeto aplicou-se duas técnicas de limiarização Máxima entropia e método de Otsu. A primeira não conseguiu limiarizar as imagens com perfeição enquanto a segunda alcançou resultados excelentes sendo aplicado sobre as imagens de scan do banco de imagens.

Desta forma, passou para etapa mais importante do trabalho a implementação dos descritores de forma. O primeiro descritor implementado foi o código da cadeia e foi utilizado o histograma de direções para representar essa característica, pois o uso direto do código de direções atinge proporções elevadas em sua representação. O segundo descritor foi a Assinatura em função da distância do ângulo, na qual foi utilizada medidas de dispersão sobre o sinal da assinatura por dois motivos: (1º) evitar o aumento de dimensionalidade do vetor de características, pois o sinal da assinatura também atinge proporções elevadas em sua representação e (2º) a assinatura deve ser invariante ao ponto inicial, mas a grande dissimetria existente entre folhas da mesma espécie impossibilitou aplicação dessa invariação de forma perfeita. O Terceiro descritor foi a transformada de Fourier aplicado ao sinal complexo da fronteira do objeto.

Após essas implementações, era necessário exportar as características extraídas por esses descritores para um formato padrão para ser utilizada na suíte WEKA (*Waikato Environment for Knowledge Analysis*) (HALL et al., 2009), para isso desenvolveu uma interface permitindo que o usuário selecione o banco de imagens, o local que deseja salvar o arquivo e os descritores que irão compor as características inseridas no arquivo, o final de todo o processo pode ser visualizado na Figura 8.

O *framework* implementado neste trabalho, denominado de *ImageLeaf*, pode ser encontrado no *Google Code*, no seguinte link: <https://code.google.com/p/imageleaf/>. Com o *ImageLeaf* é possível além de executar o ciclo principal no qual é a extração dos descritores, ele permite abrir diversas imagens e visualizar o resultado de cada técnica implementada de forma independente

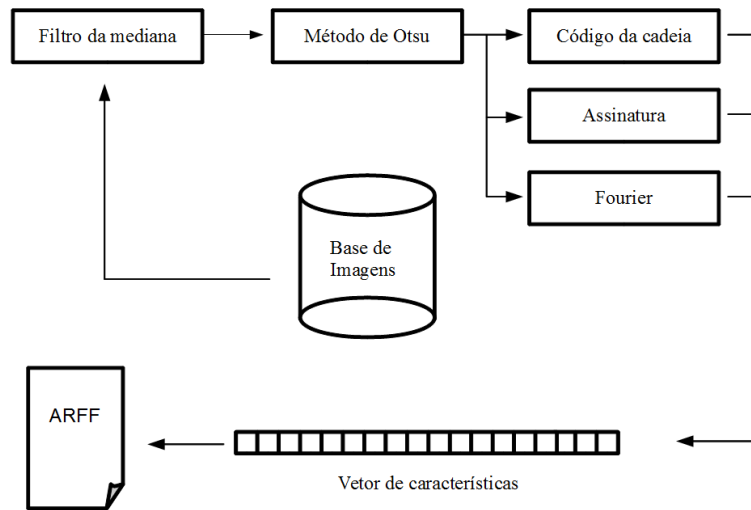


Figura 8: Representação do fluxo principal de extração de características do *framework*

e ainda comparar duas imagens distintas utilizando a distância euclidiana e os descritores de forma.

RESULTADOS E DISCUSSÕES

Nesse capítulo é apresentado o resultado do comparativo entre os classificadores utilizando as características forma de maneira separada e em conjunto. Para validação dos resultados foi aplicada a técnica de validação cruzada com o valor de K igual a 10 (*Ten-Fold Cross-Validation*).

Experimentos

Neste primeiro caso foram utilizados três vetores de características, denominados de Vetor de característica de Assinatura (VCA), Vetor de características do Código da cadeia (VCC) e Vetor de características de Fourier (VCF), possuindo os seguintes atributos:

VCA: [1 - 5] medidas de dispersão em função da assinatura. Um total de 6 atributos contando com a classe que rótula as instâncias.

VCC: [1 - 8] histograma do código da cadeia. Um total de 9 atributos contando com a classe que rótula as instâncias.

VCF: [1 - 126] descritores de Fourier. Um total de 127 atributos contando com a classe que rótula as instâncias.

Com os vetores estabelecidos, foi aplicado a classificação utilizando os seguintes métodos: IBK, *Naive Bayes* (NB), J48, *Random Forest* (RF), *MultiLayer Perceptron* (MLP) e *Support Vector Machines* (SVM).

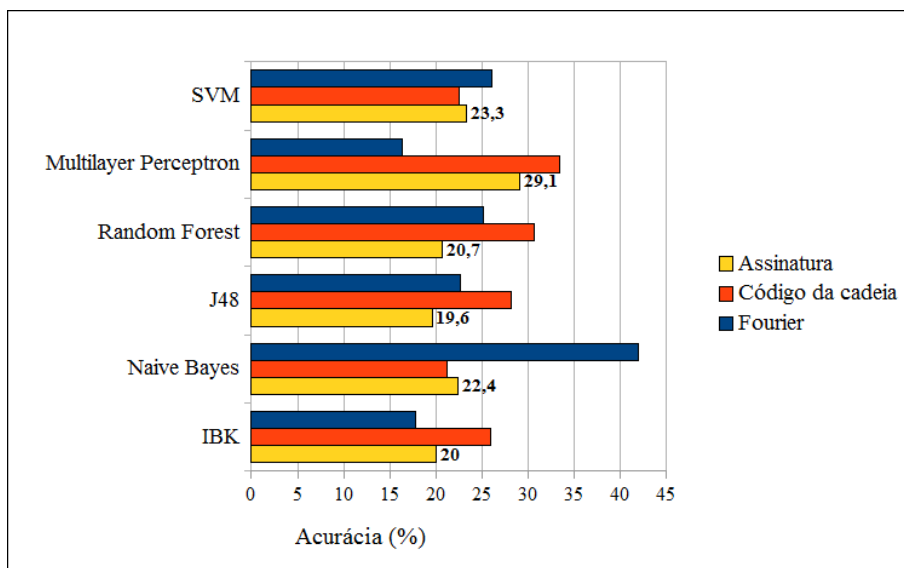


Figura 9: Gráfico de acurácias por classificador e tipo de característica.

A Figura 9 apresenta a acurácia obtida por cada classificador, ao analisar os dados da Figura pode se notar de forma clara que o resultados das classificações não foram satisfatórios, obtendo o melhor resultado com o classificador *Random Forest* com 42% de acurácia.

Esse experimento demonstrou que uso dos descritores de forma apresentados neste trabalho aplicados de maneira separa não foram capazes de discretizar as folhas em uma grau elevado de separabilidade entre as classes.

No segundo experimento foram agrupados os descritores de forma em um único vetor de características e aplicado a ele duas técnicas de aprimoramento *Adaptive Boosting* (Adaboost) que têm o objetivo de aprimorar o desempenho de qualquer algoritmo de aprendizado (FREUND; SCHAPIRE et al., 1996) e *Correlation-based feature selection* (CFS) proposto em (HALL, 1998) é uma técnica de seleção de atributos na qual parte da hipótese que o melhor subconjunto de atributos para uma classificação é aquele que contém atributos com alta correlação com as classes a serem classificadas, mas que possuem baixa correlação entre os atributos.

O vetores utilizados no segundo experimento foram o Vetor de características agrupadas (VCA) e o Vetor de características selecionadas (VCS) e eles possuem as seguintes configurações:

VCA: [1 - 5] medidas de dispersão em função da assinatura, [6 - 13] histograma do código da cadeia e [14 - 139] descritores de Fourier do sinal complexo. Um total de 140 atributos contando com a classe que rótula as instâncias.

VCS: [1 - 3] medidas de dispersão em função da assinatura, [4 - 7] histograma do código da cadeia e [8 - 17] descritores de Fourier do sinal complexo. Um total de 18 atributos contando com a classe que rótula as instâncias.

Analisando a Figura 10 observa-se que em todas as classificações os resultados foram superiores a 40% de acurácia no qual foi melhor resultado da classificação anterior, indicando que a junção de características e aplicação das técnicas de aprimoramento apresentaram ganhos consideráveis. Sendo o melhor resultado novamente com o classificador *Random Forest* com 59,7% de acurácia.

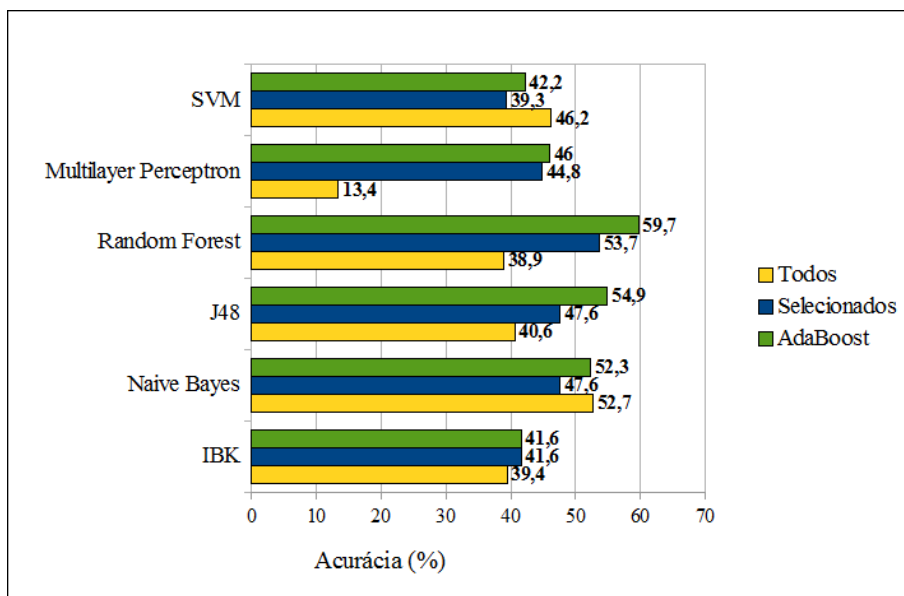


Figura 10: Gráfico de acurácias por classificador e tipo de característica.

CONCLUSÕES

O objetivo do trabalho foi desenvolver um *framework* para extração de características de bioimagens, mais especificamente folhas de diversas espécies, para posterior classificação.

As classificações utilizando os descritores de forma separada não apresentaram resultados satisfatórios em todos os métodos de classificação aplicados, mostrando que o uso desse arranjo de características não é suficiente para determinar o melhor nível de separabilidade entre as classes de espécies.

Enquanto os resultados das classificações utilizando as características em conjunto se apresentaram superiores aos descritores separados mas podem ser melhorados em grande proporção considerando que o melhor resultado alcançou a marca de 59,7% de acurácia com o classificador *Random Forest*.

Esse resultados motivaram a realização de outro trabalho que teve como objetivo a junção das características de forma e textura e aplicação de técnicas de aprimoramento os quais apresentaram resultados consideráveis para a classificação de bioimagens, levando a escrita de um artigo para apresentação de tais resultados. O artigo (BRILHADOR et al., 2013) foi submetido e aceito para apresentação na 18º Congresso Ibero-americano de Reconhecimento de Padrões CIARP 2013.

Referências Bibliográficas

- ABE, S. *Support vector machines for pattern classification*. [S.l.]: Springer, 2010.
- AHA, D.; KIBLER, D. Instance-based learning algorithms. *Machine Learning*, v. 6, p. 37–66, 1991.
- BAEZA-YATES, R. A.; RIBEIRO-NETO, B. *Modern Information Retrieval*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1999. ISBN 020139829X.
- BARTOLINI, I.; CIACCIA, P.; PATELLA, M. Warp: Accurate retrieval of shapes using phase of fourier descriptors and time warping distance. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, IEEE, v. 27, n. 1, p. 142–147, 2005.
- BRILHADOR, A. et al. Combining texture and shape descriptors for bioimages classification: A case of study in imageclef dataset. In: *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications, Proceedings*. [S.l.]: Springer, 2013. (Lecture Notes in Computer Science), p. 1 – 8. 18th Iberoamerican Congress on Pattern Recognition, CIARP 2013.
- COSTA, L. da F.; JR., R. M. C. *Shape Classification and Analysis: Theory and Practice*. 2nd. ed. Boca Raton, FL, USA: CRC Press, Inc., 2009. ISBN 0849379296, 9780849379291.
- DUDA, R.; HART, P.; STORK, D. *Pattern classification*. 2nd. ed. [S.l.]: Wiley, 2001. (Pattern Classification and Scene Analysis: Pattern Classification). ISBN 9780471056690.
- FREEMAN, H. Techniques for the Digital Computer Analysis of Chain-Encoded Arbitrary Plane Curves. In: *Proc. Nat. Electronics. Conf.* [S.l.: s.n.], 1961. v. 17, p. 412–432.
- FREUND, Y.; SCHAPIRE, R. E. et al. Experiments with a new boosting algorithm. In: MORGAN KAUFMANN PUBLISHERS, INC. *MACHINE LEARNING-INTERNATIONAL WORKSHOP THEN CONFERENCE*. [S.l.], 1996. p. 148–156.
- GIULIETTI, A. M. et al. Biodiversidade e conservação das plantas no brasil. *Megadiversidade*, v. 1, n. 1, 2005.
- The ImageCLEF 2012 Plant Identification Task*. Disponível em: <http://www.imageclef.org/system/files/TheImageCLEF2012PlantTaskOverview.pdf>.
- GONZALEZ, R. C.; WOODS, R. E. *Digital Image Processing (3rd Edition)*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2006. ISBN 013168728X.
- HALL, M. et al. The weka data mining software: an update. *ACM SIGKDD Explorations Newsletter*, ACM, v. 11, n. 1, p. 10–18, 2009.

- HALL, M. A. *Correlation-based feature selection for machine learning*. [S.l.], 1998.
- HAN, J.; KAMBER, M. *Data mining: concepts and techniques*. [S.l.]: Morgan Kaufmann, 2006.
- JOHN, G. H.; LANGLEY, P. Estimating continuous distributions in bayesian classifiers. In: MORGAN KAUFMANN PUBLISHERS INC. *Proceedings of the eleventh conference on uncertainty in artificial intelligence*. [S.l.], 1995. p. 338–345.
- LOPES, F. M. Um modelo perceptivo de limiarização de imagens digitais. *Universidade Federal do Paraná*, 2003.
- OTSU, N. A Threshold Selection Method from Gray-level Histograms. *IEEE Transactions on Systems, Man and Cybernetics*, v. 9, n. 1, p. 62–66, 1979. Disponível em: <<http://dx.doi.org/10.1109/TSMC.1979.4310076>>.
- QUINLAN, R. *C4.5: Programs for Machine Learning*. San Mateo, CA: Morgan Kaufmann Publishers, 1993.
- RAFIEI, D.; MENDELZON, A. O. Efficient retrieval of similar shapes. *The VLDB Journal*, v. 11, p. 17–27, 2002.
- RIJSBERGEN, C. J. V. *Information Retrieval*. 2nd. ed. Newton, MA, USA: Butterworth-Heinemann, 1979. ISBN 0408709294.
- RUSSELL, S. J. et al. *Artificial intelligence: a modern approach*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1996. ISBN 0-13-103805-2.
- SAHOO, P. K. et al. A survey of thresholding techniques. *Comput. Vision Graph. Image Process.*, Academic Press Professional, Inc., San Diego, CA, USA, v. 41, n. 2, p. 233–260, fev. 1988. ISSN 0734-189X. Disponível em: <[http://dx.doi.org/10.1016/0734-189X\(88\)90022-9](http://dx.doi.org/10.1016/0734-189X(88)90022-9)>.
- SANTOS et al. *Data mining*. Lisboa: FCA, 2005. (Sistemas de informação).
- STATISTICS, L. B.; BREIMAN, L. Random forests. In: *Machine Learning*. [S.l.: s.n.], 2001. p. 5–32.
- WALLACE, T.; WINTZ, P. An efficient three-dimensional aircraft recognition algorithm using normalized fourier descriptors. v. 13, n. 2, p. 99–126, June 1980.

Nome Orientador

Nome Aluno