

An Automatic Leaf Classification System Based on Shape Features

Larissa F. S. Britto¹, Luciano D. S. Pacifico¹, Teresa B. Ludermir²

¹Departamento de Computação (DC) – Universidade Federal Rural de Pernambuco (UFRPE) – Recife – PE – Brazil

²Centro de Informática (CIn) – Universidade Federal de Pernambuco (UFPE) – Recife - PE - Brazil

{larissa.feliciano, luciano.pacifico}@ufrpe.br, tbl@cin.ufpe.br

Abstract. *Plants are an indispensable part of our ecosystem, providing many irreplaceable resources for animals. But the identification of plant species is a hard task, and most people are not able to correctly recognize many plant species. In this work, we develop a plant recognition system, based on the extraction of shape features from plant leaf images. A new medicinal plant data set is proposed, and four classifiers are tested as the recognition module for the proposed automatic system. Experimental results showed that the best classifiers are able to obtain average accuracies over 98% on the proposed data set.*

Resumo. *As plantas são peças indispensáveis para nossos ecossistemas, provendo recursos insubstituíveis para os animais. Porém, a identificação de espécies de plantas é uma tarefa difícil, e a maioria das pessoas não está apta ao reconhecimento correto de muitas espécies de plantas. Neste trabalho, uma abordagem para o reconhecimento de espécies de plantas é desenvolvida, baseada na extração de características de forma das imagens de suas folhas. Uma base de dados de plantas medicinais é proposta, e quatro classificadores são testados como módulo de reconhecimento do sistema automático proposto. Os resultados experimentais mostram que os melhores classificadores são capazes de obter uma acurácia média acima de 98% na base de dados proposta.*

1. Introdução

As plantas desempenham papéis fundamentais em seus ambientes, sendo fundamentais para a sobrevivência dos animais no planeta Terra. Tais organismos são as maiores fontes de oxigênio e alimentos para os animais, promovendo ainda o balanceamento ecológico em seus ambientes. As plantas também servem como matéria-prima para diferentes setores da indústria, como o setor têxtil, produção de combustíveis, produção de cosméticos e produtos de higiene, etc.

Algumas espécies de plantas possuem também papel importante na medicina, servindo tanto como substratos para a indústria farmacêutica, ou sendo elas mesmas usadas como remédios e tratamentos médicos. Plantas usadas no tratamento ou prevenção de doenças e enfermidades são conhecidas como plantas medicinais [Sabu et al. 2017].

Embora as plantas estejam em todos os lugares, a classificação correta de espécies de plantas permanece sendo um grande desafio, tradicionalmente requerendo a consulta

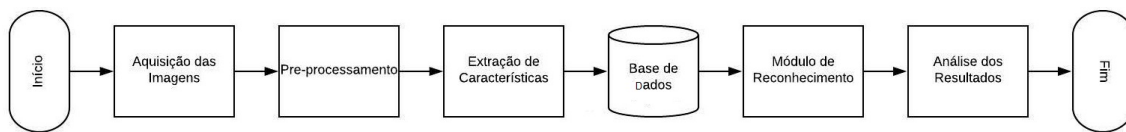


Figura 1. Etapas Básicas de um Sistema de Reconhecimento de Espécies de Plantas.

aos especialistas em botânica. Como os especialistas geralmente não estão disponíveis em todas as ocasiões, e além disso, um especialista em um tipo de família de espécie de plantas não é treinado para reconhecer outras espécies, o processo de identificação manual é bastante demorado e passível a erros humanos.

No intuito de evitar erros humanos, sistemas automáticos de reconhecimento de espécies de plantas têm sido propostos na literatura nos últimos anos [Agarwal et al. 2006, Kumar et al. 2012, Mallah et al. 2013, Jin et al. 2015, Jose et al. 2018, Adinugroho and Sari 2018], em decorrência do avanço significativo de áreas como a Aprendizagem de Máquina e Visão Computacional.

Neste trabalho, um sistema de reconhecimento de espécies de plantas medicinais é proposto. Uma nova base de dados é apresentada, formada pela extração de características relacionadas à forma das folhas das plantas. A base de dados proposta é composta por 1148 segmentos de imagens, obtidas de 15 espécies de plantas medicinais comumente encontradas no Brasil. Quatro algoritmos da Aprendizagem Supervisionada (*classificadores*) são testados: Árvore de Decisão, K-Vizinhos Mais Próximos, K-Vizinhos Mais Próximos Ponderado e Algoritmo de Floresta Aleatória. Os melhores classificadores serão selecionados para a composição do módulo de reconhecimento do sistema proposto.

O trabalho está dividido como segue. Os principais componentes de um sistema automático de reconhecimento de plantas são descritos na Seção 2. Logo após, a metodologia utilizada para o desenvolvimento da base de dados proposta é definida (Seção 3). Por fim, os resultados experimentais são discutidos (Seção 4), assim como as conclusões e possíveis linhas para pesquisas futuras (Seção 5).

2. Sistemas de Reconhecimento de Espécies de Plantas

As etapas principais para um sistema automático de reconhecimento de plantas estão apresentadas na Fig. 1. São elas: aquisição das imagens, pré-processamento, extração de características e a etapa de reconhecimento (classificação) das espécies [Sahay and Chen 2016].

A aquisição das imagens das plantas é geralmente realizada sob ambientes controlados (como por exemplo, plano de fundo e fontes de iluminação controlados), para evitar problemas de aquisição, como objetos com variações de luminosidade, com oclusão parcial, dobras nas objetos, dentre outros [Kumar et al. 2012, Mallah et al. 2013, Sahay and Chen 2016]. Mas em algumas aplicações, as imagens são adquiridas no ambiente natural das plantas, como em [Cerutti et al. 2013, Sun et al. 2017].

O pré-processamento das imagens visa a padronização, tanto em termos de escala quanto em termos de orientação, das imagens antes da etapa de extração de características. Atividades comuns nessa etapa são à remoção de ruídos, correção do histograma, con-

Tabela 1. Distribuição Amostral por Espécie de Planta Medicinal.

Espécie	Número de Amostras
<i>Achillea millefolium</i>	12
<i>Aconitum napellus</i>	9
<i>Apium graveolens</i>	18
<i>Argemone mexicana</i>	25
<i>Bidens pilosa</i>	16
<i>Bixa orellana</i>	18
<i>Caesalpinia ferrea</i>	17
<i>Casearia sylvestris</i>	33
<i>Chondrodendron platyphyllum</i>	15
<i>Cinnamomum verum</i>	32
<i>Eugenia uniflora</i>	20
<i>Glycyrrhiza glabra</i>	24
<i>Illicium verum</i>	28
<i>Oxalis acetosella</i>	7
<i>Sambucus nigra</i>	13
Total	287

versão da imagem para outros sistemas de cores, conversão da imagem para uma imagem em níveis de cinza ou binária, aumento de contraste, etc [Sabu and Sreekumar 2017].

A extração de características geralmente considera características como cor, características das folhas, características das flores, características das sementes e características de outros órgãos das plantas [Cope et al. 2012, Rahmani et al. 2015, Sabu and Sreekumar 2017]. Muitas aplicações de reconhecimento de espécies de plantas focam na extração de características relacionadas às folhas, tais como forma, textura, venação, dimensões fractais, Características Robustas Aceleradas (*Speeded Up Robust Features*, ou SURF), Histogramas de Gradientes Orientados (*Histogram of Oriented Gradients*, ou HOG), dentre outras [Cope et al. 2012, Jin et al. 2015, Sabu and Sreekumar 2017, Sabu et al. 2017].

Como módulo de reconhecimento, vários algoritmos da literatura de Aprendizagem de Máquina têm sido aplicados ao reconhecimento de plantas: Árvores de Decisão [Rahmani et al. 2015], K-Vizinhos Mais Próximos [Mallah et al. 2013, Mallah and Orwell 2013, Britto and Pacifico 2018], Máquinas de Vetores de Suporte [Venkataraman and Mangayarkarasi 2016] e Redes Neurais Artificiais [Prasad and Singh 2017, Adinugroho and Sari 2018, Hu et al. 2018, Pacifico et al. 2018].

3. Metodologia

Neste trabalho, uma nova base de dados é apresentada. A base de dados proposta é composta de 287 amostras de imagens obtidas de 15 espécies de plantas medicinais facilmente encontradas no Brasil. As imagens foram obtidas através da internet e pela aquisição de imagens em campo. Tanto as imagens obtidas pela internet quanto as imagens obtidas em campo foram classificadas com a ajuda de especialistas. A distribuição amostral por classe é apresentada na Tabela 1.

Depois da aquisição, as etapas de pré-processamento (Seção 3.1), extração de características (Seção 3.2) e composição da base de dados (Seção 3.3) são executadas.

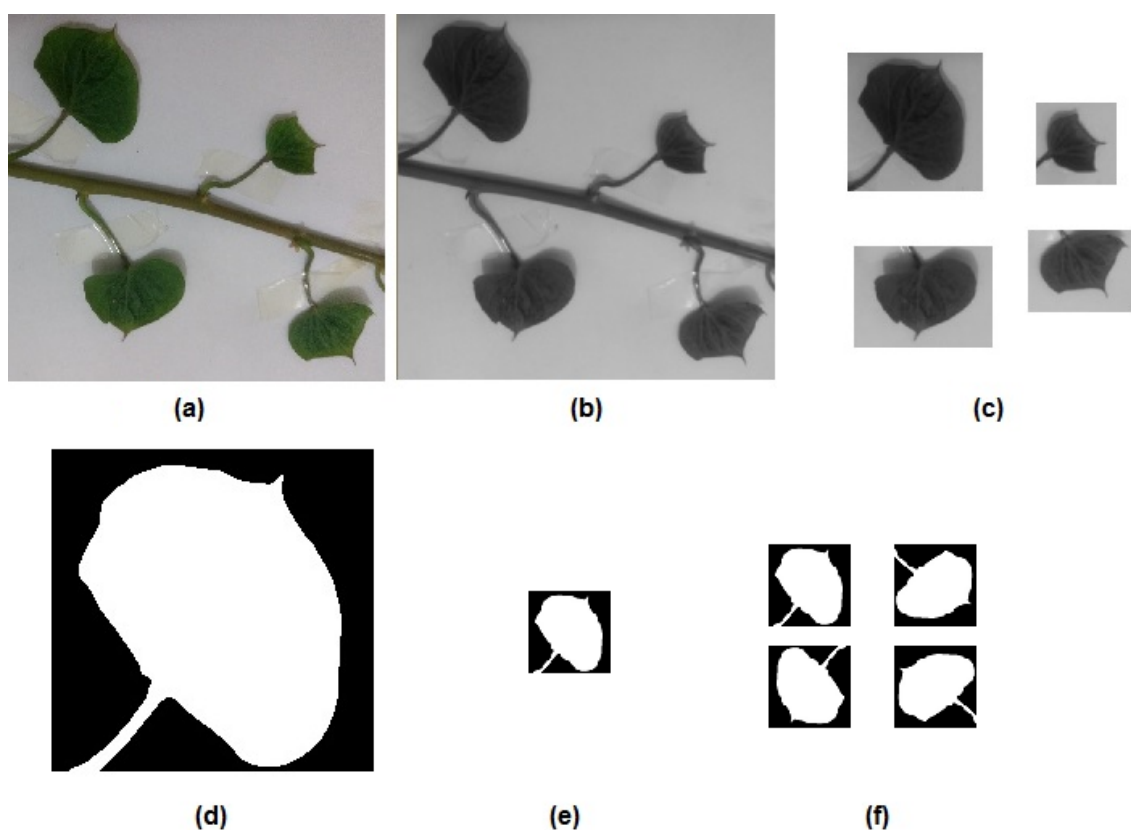


Figura 2. Etapas do Pré-Processamento: (a) Imagem Original, (b) Conversão para Níveis de Cinza, (c) Obtenção dos Segmentos de Imagens, (d) Conversão para Binário, (e) Ajuste de Escala, (f) Geração dos Segmentos Rotacionados.

3.1. Pré-Processamento das Imagens

Para cada imagem original, o pré-processamento foi executado de acordo com as etapas ilustradas na Fig 2.

As imagens originais (em RGB) são convertidas para imagens em níveis de cinza inicialmente. Dado que as imagens coletadas podem conter mais do que uma amostra das folhas, a segunda etapa do processo envolve a separação de cada amostra em um *segmento de imagem*, sendo essa separação realizada manualmente.

Após a obtenção dos segmentos de imagens, cada segmento é binarizado por uma operação de limiarização (*thresholding*). O operador de limiarização escolhido foi o algoritmo de Otsu [Otsu 1979], que é um dos algoritmos mais rápidos e efetivos para a binarização de imagens.

A etapa seguinte foi realizada pela padronização dos segmentos de imagens por um operador de redimensionamento. Uma vez que as imagens originais foram obtidas através de diferentes fontes, optou-se pela padronização das dimensões de cada segmento de imagem, como uma forma de tornar o treinamento dos modelos de reconhecimento mais balanceado, assim como evitar o aumento desnecessário no custo computacional do sistema.

Os segmentos de folhas binarizados resultantes são então rotacionados em três orientações diferentes (90°, 180° e 270°), de modo a aumentar a variabilidade de imagens na base proposta, resultando em um total de 1148 segmentos de imagens finais.

3.2. Extração de Características

Para o trabalho atual, dez características das formas das folhas foram extraídas de cada segmento de imagem binarizado:

1. Momentos de Hu [Hu 1962]: Os Momentos Invariantes de Hu são um dos mais consagrados operadores para a descrição da forma de objetos em imagens. Esses momentos se mantêm praticamente constantes diante de variações na escala dos objetos, assim como em translações e rotações desses objetos. Sete características da forma são obtidas pelo cálculo dos Momentos de Hu, sendo elas:

$$I_1 = \mu_{20} + \mu_{02} \quad (1)$$

$$I_2 = (\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2 \quad (2)$$

$$I_3 = (\mu_{30} - 3\mu_{12})^2 + (3\mu_{21} - \mu_{03})^2 \quad (3)$$

$$I_4 = (\mu_{30} + \mu_{12})^2 + (\mu_{21} + \mu_{03})^2 \quad (4)$$

$$I_5 = (\mu_{30} - 3\mu_{12})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] + (3\mu_{21} - \mu_{03})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \quad (5)$$

$$I_6 = (\mu_{20} - \mu_{02})[(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] - (\mu_{30} - 3\mu_{12})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \quad (6)$$

$$I_7 = (3\mu_{21} - \mu_{03})(\mu_{30} + \mu_{12})[(\mu_{30} + \mu_{12})^2 - 3(\mu_{21} + \mu_{03})^2] - (\mu_{30} - 3\mu_{12})(\mu_{21} + \mu_{03})[3(\mu_{30} + \mu_{12})^2 - (\mu_{21} + \mu_{03})^2] \quad (7)$$

onde μ_{ij} são momentos invariantes à escala [Hu 1962].

2. Perímetro da Folha ($leaf_{per}$): O perímetro da folha é calculado contando-se o número de pixels na margem da folha.
3. Área da Folha ($leaf_{area}$): A área da folha é calculada como o número de pixels tendo valor binário 1 no segmento de imagem binário.
4. Compacidade da Folha ($leaf_{comp}$): A compacidade da folha é calculada como:

$$leaf_{comp} = 4\pi \frac{leaf_{area}}{leaf_{per}^2} \quad (8)$$

3.3. Base de Dados Proposta

Após a conclusão das etapas de pré-processamento e extração de características, a base de dados de plantas medicinais é obtida. No total, 1148 padrões são obtidos (segmentos de imagens), e 10 características são extraídas para a composição da base, como indicado abaixo:

1. I_1
2. I_2
3. I_3
4. I_4
5. I_5
6. I_6
7. I_7
8. $lea.f_{per}$
9. $lea.f_{area}$
10. $lea.f_{comp}$
11. Espécie da Planta (rótulo da classe).

Após os resultados de uma abordagem de tentativa-e-erro, optou-se por fazer uso de segmentos de imagens de dimensões 50x50.

4. Avaliação Experimental

Nesta seção, os resultados experimentais serão apresentados. Quatro classificadores provenientes da literatura de Aprendizagem de Máquina são comparados: Árvore de Decisão (*Decision Tree*, ou DT), k -Vizinhos mais Próximos (KNN), k -Vizinhos mais Próximos Ponderado (WKNN) e o Classificador de Floresta Aleatória (RFC). Todas as variações do k -NN foram testadas com $k = 3, 4$ e 5 , usando a distância euclidiana como medida de dissimilaridade. Todos os algoritmos foram implementados na linguagem de programação Python, através do uso da biblioteca *scikit-learn* [Pedregosa et al. 2011, Buitinck et al. 2013], sendo a DT e o RFC testados com os parâmetros *default* dessa biblioteca. Todos os testes foram realizados em um computador com uma CPU i7-7700K, uma GPU NVIDIA GeForce GTX 1060 6GB e 32 GB RAM.

Os experimentos foram conduzidos através de um *framework* do tipo validação cruzada *10-fold*: a base de dados proposta é dividida aleatoriamente em dez partes (sem sobreposição entre essas partes), e em cada uma das dez etapas da repetição do experimento, uma dessas partes é usada como *conjunto de teste*, enquanto as outras nove partes são usadas como *conjunto de treinamento* dos modelos. Visando a obtenção de um conjunto maior de amostras, o processo de validação cruzada *10-fold* foi repetido 10 vezes, sendo em cada uma dessas vezes os dados redistribuídos aleatoriamente para a formação das *folds*. O objetivo dessa reamostragem é evitar resultados obtidos por sorte pelos classificadores, promovendo uma análise mais justa dos experimentos.

Como métricas de avaliação dos experimentos, quatro índices comumente aplicados a problemas de classificação foram adotados: a Acurácia, a Precisão (*Precision*), a Revocação (*Recall*) e a *F-Measure*. A avaliação inclui uma análise empírica dos resultados obtidos para o conjunto de teste dos experimentos, assim como uma análise do tempo médio de execução de cada um dos algoritmos.

A avaliação também inclui um sistema de *ranks* obtido através do uso do teste de Friedman [Friedman 1937, Friedman 1940] em relação à avaliação global da média no conjunto de teste para cada métrica. O teste de Friedman é um teste de hipóteses não-paramétrico que calcula valores de *ranks* para os algoritmos para cada base de dados separadamente. Se a hipótese nula de que os *ranks* não são significativamente diferentes for rejeitada, o teste de Nemenyi [Nemenyi 1962] é adotado com um teste *post hoc* para o teste de Friedman. De acordo com o teste de Nemenyi, a performance de dois algoritmos é considerada significativamente diferente se a diferença entre seus valores médios de *rank* for ao menos maior que uma *diferença crítica* dada por:

$$CD = q_\alpha \sqrt{\frac{n_{alg}(n_{alg} + 1)}{6n_{bases}}} \quad (9)$$

onde n_{bases} representa o número de bases de dados, n_{alg} representa o número de algoritmos comparados e q_α são valores críticos baseados em estatísticas nos limites do modelo t de *Student* divididas por $\sqrt{2}$ [Demšar 2006]. Como as métricas de avaliação são de *maximização* (quanto mais alto o valor obtido, melhor), os melhores algoritmos irão obter os valores mais altos para o *rank* médio.

Os resultados experimentais são apresentados na Tabela 2.

Tabela 2. Resultados Experimentais. Média: resultado média para a métrica; Std: desvio padrão para a métrica; Tempo: tempo médio de execução (em segundos). A Acurácia, Precisão, Revocação e *F-Measure* levam em consideração os resultados obtidos para o conjunto de testes.

Algoritmo	Acurácia		Precisão		Revocação		<i>F-Measure</i>		Tempo
	Média	Std	Média	Std	Média	Std	Média	Std	
DT	0.9482	0.0188	0.9546	0.0172	0.9482	0.0188	0.9480	0.0189	0.0081
KNN ₃	0.9555	0.0221	0.9634	0.0176	0.9555	0.0221	0.9554	0.0223	0.0027
KNN ₄	0.8693	0.0359	0.8918	0.0306	0.8693	0.0359	0.8677	0.0364	0.0030
KNN ₅	0.7912	0.0518	0.8221	0.0484	0.7912	0.0518	0.7900	0.0513	0.0030
WKNN ₃	0.9773	0.0132	0.9810	0.0105	0.9773	0.0132	0.9774	0.0131	0.0027
WKNN ₄	0.9738	0.0138	0.9782	0.0106	0.9738	0.0138	0.9739	0.0137	0.0031
WKNN ₅	0.9707	0.0147	0.9751	0.0123	0.9707	0.0147	0.9706	0.0148	0.0032
RFC	0.9834	0.0153	0.9862	0.0126	0.9834	0.0153	0.9835	0.0153	0.0285

Da Tabela 2, em uma análise empírica, podemos observar que o RFC obteve os melhores resultados em respeito a todos os índices de classificação selecionados, em comparação aos demais classificadores. O WKNN₃ e o WKNN₄ obtiveram, respectivamente, o segundo e o terceiro melhores desempenhos médios. Pelos experimentos, podemos observar que tanto no KNN quanto no WKNN o desempenho médio nos índices considerados decaiu com o aumento do valor do parâmetro k . Considerando o fato de que a base de dados proposta possui quinze classes, sendo essas classes não balanceadas entre si, esse decaimento na performance é completamente esperado, corroborando resultados obtidos na literatura [Rahmani et al. 2015, Pacifico et al. 2018]. Ainda de acordo com a Tabela 2, pode-se observar que o KNN tradicional é mais sensível às variações no valor do número de vizinhos sendo considerados do que sua versão ponderada (o WKNN).

A Tabela 3 apresenta o valor dos *ranks* médios obtidos pelos testes de Friedman/Nemenyi. Tais resultados mostram que o RFC obteve os melhores valores médios

Tabela 3. Avaliação Geral: *ranks* médios para o teste de Friedman/Nemenyi com CD = 3.3202. Rank_A: *rank* relativo à Acurácia; Rank_P: *rank* relativo à Precisão; Rank_R: *rank* relativo à Revocação; Rank_F: *rank* relativo à *F-Measure*.

Algoritmo	Rank _A	Rank _P	Rank _R	Rank _F
DT	17.74	17.21	17.74	17.88
KNN ₃	19.91	20.15	19.91	19.88
KNN ₄	7.82	7.86	7.82	7.82
KNN ₅	3.44	3.44	3.44	3.46
WKNN ₃	29.40	29.50	29.40	29.44
WKNN ₄	27.19	27.57	27.19	27.08
WKNN ₅	25.75	25.26	25.75	25.69
RFC	32.75	33.01	32.75	32.75

para o *rank*, de acordo com uma avaliação global, para todas as métricas, apresentando diferença estatística significativa em relação aos demais classificadores, exceto para a *F-Measure*, no qual o RFC foi estatisticamente equivalente ao WKNN₃ (porém, com um *rank* um pouco melhor nessa métrica). Os testes de Friedman/Nemenyi mostram ainda que, em relação a todas as métricas de avaliação, o WKNN₃ e WKNN₄ obtiveram o segundo melhor desempenho em conjunto, tendo em vista que ambos os modelos não apresentaram diferenças estatísticas significativas entre seus resultados, embora o WKNN₃ tenha atingido *ranks* um pouco melhores do que o WKNN₄. Os resultados para os testes de Friedman/Nemenyi são ilustrados na Fig. 3, do pior algoritmo (à esquerda), ao melhor (à direita).

Os melhores tempos médios de execução foram encontrados pelas variações do KNN e do WKNN, tendo em vista que o treinamento dessas técnicas ocorre pelo simples armazenamento das amostras conhecidas (rotuladas).

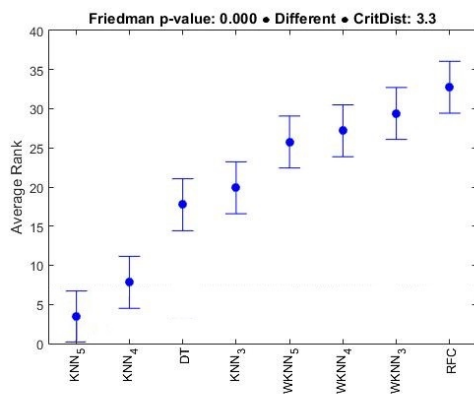
Em uma avaliação geral, levando em consideração tanto os valores das métricas de classificação quanto o tempo médio de execução, podemos concluir que o classificador de Floresta Aleatória seria uma escolha mais adequada como módulo de reconhecimento do sistema proposto. O RFC foi responsável pelas melhores performances entre todos os algoritmos, assim como também obteve um tempo de execução médio aceitável para o desenvolvimento de uma aplicação de tempo real.

5. Conclusões

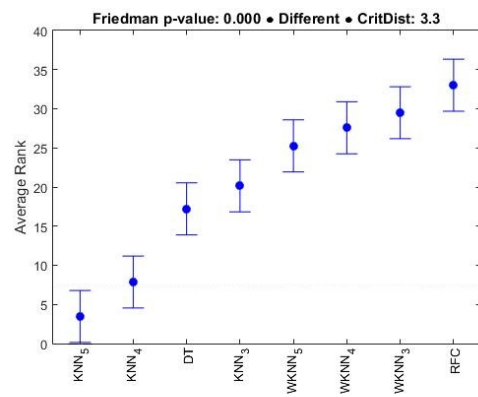
Neste trabalho, uma abordagem para o desenvolvimento de um sistema automático de reconhecimento de imagens de plantas medicinais é apresentada. Uma base de dados de plantas medicinais é proposta, elaborada através da extração de características de forma de segmentos de imagens das plantas.

Como módulo de reconhecimento, quatro classificadores bastante conhecidos da literatura de Aprendizagem de Máquina foram testados: Árvore de Decisão, K-Vizinhos Mais Próximos, K-Vizinhos Mais Próximos Ponderado e o classificador de Floresta Aleatória.

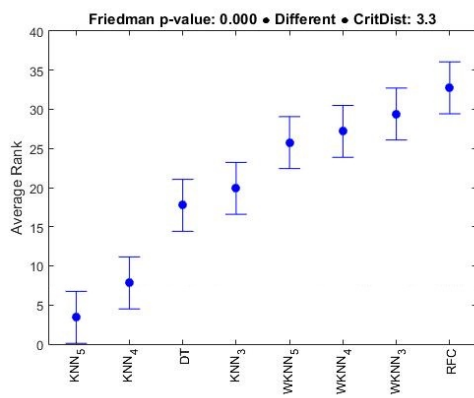
Como forma de avaliação dos classificadores selecionados, uma análise empírica complementada por testes de hipóteses do tipo Testes de Friedman/Nemenyi foi aplicada, levando-se em consideração quatro métricas tradicionais de classificação (Acurácia, Pre-



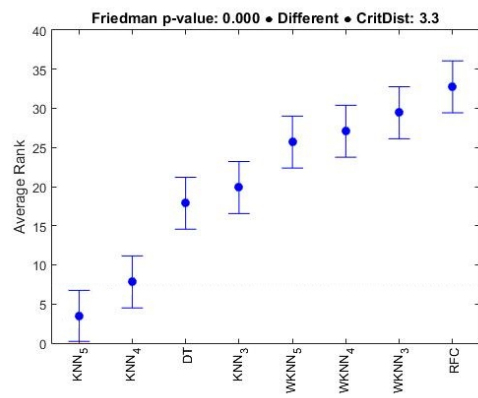
(a)



(b)



(c)



(d)

Figura 3. Resultados para os Testes de Friedman/Nemenyi: (a) Acurácia, (b) Precisão, (c) Revocação, (d) *F-Measure*.

ção, Revocação e *F-Measure*). O tempo de execução médio também foi usado como métrica da avaliação empírica.

Os resultados experimentais apontaram que o RFC e o WKNN (com $k = 3$) obtiveram os melhores desempenhos em relação à cada uma das métricas de classificação selecionadas, de acordo com os testes de Friedman/Nemenyi executados. Com uma avaliação levando em consideração tanto os resultados das métricas de classificação quanto o tempo de execução, concluímos que o RFC pode ser considerado uma boa escolha para a composição do módulo de reconhecimento de um sistema de reconhecimento de espécie de plantas em tempo real, dado que esse algoritmo apresentou bons valores de acurácia (acurácia média de 98,34%), assim como um tempo de execução aceitável.

Como trabalhos futuros, pretendemos expandir a base de dados proposta, tanto pelo acréscimo de novas espécies de plantas, quanto pelo aumento do número de imagens por espécie. Imagens de outros órgãos das plantas (como flores, sementes e frutos) também serão adquiridas e agregadas ao sistema, em uma tentativa de aumentar a separabilidade entre as espécies de plantas reconhecidas pelo sistema, assim como também tornar os sistemas a serem desenvolvidos mais precisos e completos. Também pretendemos expandir a análise pela avaliação de outras características que possam ser extraídas das imagens das folhas (tais como venação, características de margem, entre outras). Por fim, uma ferramenta de propósito geral será desenvolvida, que além de oferecer um reconhecimento automático e preciso das espécies de plantas medicinais, também apresentará informações sobre o uso correto dessas plantas, como forma de difusão dos conhecimentos a cerca da importância dessas espécies na prevenção e no tratamento de doenças e problemas de saúde.

Agradecimentos

Os autores gostariam de agradecer ao CNPq e a CAPES pelo suporte financeiro.

Referências

- Adinugroho, S. and Sari, Y. A. (2018). Leaves classification using neural network based on ensemble features. In *2018 5th International Conference on Electrical and Electronic Engineering (ICEEE)*, pages 350–354. IEEE.
- Agarwal, G., Belhumeur, P., Feiner, S., Jacobs, D., Kress, W. J., Ramamoorthi, R., Bourg, N. A., Dixit, N., Ling, H., Mahajan, D., et al. (2006). First steps toward an electronic field guide for plants. *Taxon*, 55(3):597–610.
- Britto, L. F. and Pacifico, L. D. (2018). Plant classification using weighted k-nn variants. In *Anais do XV Encontro Nacional de Inteligência Artificial e Computacional*, pages 58–69. SBC.
- Buitinck, L., Louppe, G., Blondel, M., Pedregosa, F., Mueller, A., Grisel, O., Niculae, V., Prettenhofer, P., Gramfort, A., Grobler, J., Layton, R., VanderPlas, J., Joly, A., Holt, B., and Varoquaux, G. (2013). API design for machine learning software: experiences from the scikit-learn project. In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning*, pages 108–122.
- Cerutti, G., Tougne, L., Mille, J., Vacavant, A., and Coquin, D. (2013). Understanding leaves in natural images—a model-based approach for tree species identification. *Computer Vision and Image Understanding*, 117(10):1482–1501.

- Cope, J. S., Corney, D., Clark, J. Y., Remagnino, P., and Wilkin, P. (2012). Plant species identification using digital morphometrics: A review. *Expert Systems with Applications*, 39(8):7562–7573.
- Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *The Journal of Machine Learning Research*, 7:1–30.
- Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, 32(200):675–701.
- Friedman, M. (1940). A comparison of alternative tests of significance for the problem of m rankings. *The Annals of Mathematical Statistics*, 11(1):86–92.
- Hu, J., Chen, Z., Yang, M., Zhang, R., and Cui, Y. (2018). A multiscale fusion convolutional neural network for plant leaf recognition. *IEEE Signal Processing Letters*, 25(6):853–857.
- Hu, M.-K. (1962). Visual pattern recognition by moment invariants. *IRE transactions on information theory*, 8(2):179–187.
- Jin, T., Hou, X., Li, P., and Zhou, F. (2015). A novel method of automatic plant species identification using sparse representation of leaf tooth features. *PloS one*, 10(10):e0139482.
- Jose, C.-R., Erick, M.-M., and Herve, G. (2018). Hidden biases in automated image-based plant identification. In *2018 IEEE International Work Conference on Bioinspired Intelligence (IWOBI)*, pages 1–9. IEEE.
- Kumar, N., Belhumeur, P. N., Biswas, A., Jacobs, D. W., Kress, W. J., Lopez, I. C., and Soares, J. V. (2012). Leafsnap: A computer vision system for automatic plant species identification. In *Computer Vision—ECCV 2012*, pages 502–516. Springer.
- Mallah, C., Cope, J., and Orwell, J. (2013). Plant leaf classification using probabilistic integration of shape, texture and margin features. *Signal Processing, Pattern Recognition and Applications*, 5(1).
- Mallah, C. D. and Orwell, J. (2013). Probabilistic classification from a k-nearest-neighbour classifier. *Computational Research*, 1(1):1–9.
- Nemenyi, P. (1962). Distribution-free multiple comparisons. In *Biometrics*, volume 18, page 263. INTERNATIONAL BIOMETRIC SOC 1441 I ST, NW, SUITE 700, WASHINGTON, DC 20005-2210.
- Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, 9(1):62–66.
- Pacifico, L. D., Macario, V., and Oliveira, J. F. (2018). Plant classification using artificial neural networks. In *2018 International Joint Conference on Neural Networks (IJCNN)*, pages 1–6. IEEE.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830.

- Prasad, S. and Singh, P. P. (2017). Medicinal plant leaf information extraction using deep features. In *Region 10 Conference, TENCON 2017-2017 IEEE*, pages 2722–2726. IEEE.
- Rahmani, M. E., Amine, A., and Hamou, M. R. (2015). Plant leaves classification. *ALL-DATA 2015*, 82.
- Sabu, A. and Sreekumar, K. (2017). Literature review of image features and classifiers used in leaf based plant recognition through image analysis approach. In *Inventive Communication and Computational Technologies (ICICCT), 2017 International Conference on*, pages 145–149. IEEE.
- Sabu, A., Sreekumar, K., and Nair, R. R. (2017). Recognition of ayurvedic medicinal plants from leaves: A computer vision approach. In *Image Information Processing (ICIIP), 2017 Fourth International Conference on*, pages 1–5. IEEE.
- Sahay, A. and Chen, M. (2016). Leaf analysis for plant recognition. In *Software Engineering and Service Science (ICSESS), 2016 7th IEEE International Conference on*, pages 914–917. IEEE.
- Sun, Y., Liu, Y., Wang, G., and Zhang, H. (2017). Deep learning for plant identification in natural environment. *Computational intelligence and neuroscience*, 2017.
- Venkataraman, D. and Mangayarkarasi, N. (2016). Computer vision based feature extraction of leaves for identification of medicinal values of plants. In *Computational Intelligence and Computing Research (ICCIC), 2016 IEEE International Conference on*, pages 1–5. IEEE.