

Reconhecimento Automático de Espécies de Pimentas Utilizando Redes Neurais Artificiais

Paulo Renato Baliza Silva¹

Escola de Ciências Exatas e da Computação, PUC Goiás, Goiânia, GO

Arlindo Rodrigues Galvão Filho²

Escola de Ciências Exatas e da Computação, PUC Goiás, Goiânia, GO

Heber Valdo Nogueira³

Divisão de Inovação e Resultados, Assembleia Legislativa do Estado de Goiás

Clarimar José Coelho⁴

Escola de Ciências Exatas e da Computação, PUC Goiás, Goiânia, GO

Resumo. Este trabalho apresenta uma metodologia para classificação de espécies pimentas utilizando redes neurais artificiais perceptron multicamadas. São analisadas 360 imagens obtidas em ambiente controlado utilizando uma câmera multiespectral no visível para 24 espécies de pimentas divididas em 15 amostras por espécie. As imagens passam por uma etapa de pré-processamento onde os ruídos são eliminados e a região da pimenta é detectada. As características utilizadas para a classificação das pimentas são o tamanho da pimenta e momentos de cores obtidos das imagens pré-processadas em diferentes espaços de cores. Os resultados obtidos mostram que em média a classificação das imagens tem uma taxa de acerto de 97,36%. **Palavras-chave.** Redes Neurais Artificiais, Classificação de Imagens, Reconhecimento de Padrões

1 Introdução

O conhecimento das características morfológicas de pimentas é de grande relevância para a caracterização de suas espécies. Existem diversas variações nessas características que são determinantes para suas classificações visuais. Algumas espécies podem apresentar características similares, como coloração, textura, tamanho e outras, o que pode resultar em classificações erradas. Além da análise clássica visual [10, 12], outras metodologias já foram aplicadas, tais como a classificação utilizando um coeficiente de variação [4], determinação por espectrofotometria de capsaicinoides [17], ou ainda baseada em análise de imagens multi/hiperespectrais [1, 6].

Nos estudos feitos para o desenvolvimento deste trabalho, não foi encontrado nenhum modelo menos custoso, ou ainda baseado em uma estratégia mais simples do que a classificação utilizando imagens digitais no espectro do visível. Tal estratégia utilizaria imagens

¹paulorenatoobs@gmail.com

²arlindopucgo@gmail.com

³heber.nogueira@al.go.leg.br

⁴clarimarc@gmail.com

obtidas por câmeras digitais convencionais em conjunto com modelos de classificação que destacaria as características obtidas pela análise clássica visual automaticamente.

Estratégias semelhantes já foram aplicadas na agricultura, como [8], que classifica o tipo de plantação utilizando imagens de sensoriamento remoto e *Deep Learning*, [9] que criou um sistema de classificação de frutas baseado em entropia fracional de Fourier e algoritmos genéticos e [7], que classifica sementes de pimentas utilizando uma rede neural.

Neste cenário, o objetivo deste trabalho é apresentar um banco de imagens contendo múltiplas espécies de pimentas, juntamente com uma estratégia para classificação de tais espécies. Esta estratégia aplica um pré-processamento às imagens e extrai as características que serão utilizadas na classificação através de uma rede neural artificial perceptron multicamadas (PMC).

Este artigo está organizado da seguinte forma: Na seção 2 é apresentado banco de dados proposto. Nas seções 3 e 4 são descritos o pré-processamento e extração de características, respectivamente. A configuração da rede neural utilizada é descrita na seção 5. Os resultados são discutidos na seção 6. A seção 7 traz as conclusões do trabalho.

2 Banco de Dados

A avaliação do método proposto requer o uso de uma base de dados catalogada e com boa diversidade de imagens para cada pimenta. Não foi encontrado nenhum banco de imagens que atendesse a tais requisitos. Neste caso, a alternativa foi construir o próprio banco de dados. Todas as amostras foram digitalizadas individualmente, em ambiente controlado. Tais cuidados possibilitaram a geração de um banco de dados onde características como tamanho e forma das pimentas estão em escala.

As imagens das pimentas foram obtidas utilizando uma câmera multiespectral JAI AD-080GE com a mesma configuração de controle de foco e iluminação para todas amostras. A base de dados é composta por 360 imagens no espectro visível, sendo 15 amostras de cada uma das 24 espécies de pimentas. A Figura 1 apresenta três espécies de pimentas: Pimenta de Cheiro Vermelha (a), Pimenta Dedo de Moça Amarela (b) e (c) Pimenta Pitanga.

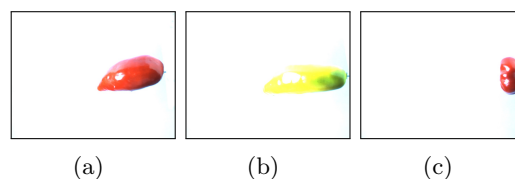


Figura 1: Pimentas sem Pré-Processamento. (a) Pimenta de Cheiro Vermelha; (b) Pimenta Dedo de Moça Amarela; (c) Pimenta Pitanga.

A Figura 1 mostra a presença de plano de fundo branco e ruídos gerados pela luminosidade. Nesse cenário, faz-se necessária a utilização de técnicas de pré-processamento para eliminar ruídos e detectar a pimenta nas imagens, possibilitando a eliminação do plano de fundo das imagens e preparando-as para a extração de características.

3 Pré-Processamento das Imagens

O pré-processamento de imagens é essencial para o sucesso da classificação das diferentes espécies de pimenta, pois nesta etapa as imagens são preparadas para a extração de características, eliminando ruídos e identificando a região de interesse das imagens [13].

A metodologia proposta neste trabalho é iniciada pela detecção das estruturas que compõem a imagem com base em sua conectividade. Para isso, primeiramente a imagem é submetida a um operador de suavização com filtro gaussiano com o objetivo de remover ruídos presentes no fundo da imagem. Sendo a imagem original $I(x, y)$ e $H(u, v)$ o filtro gaussiano, a imagem suavizada é obtida pela convolução $G(x, y) = I(x, y) \cdot H(u, v) = \sum_{u,v \in H} [I(x-u, y-v) \cdot H(u, v)]$, onde $H(u, v) = (2\pi\sigma^2)^{-1/2} \exp -\frac{(u+v)^2}{2\sigma^2}$ e σ^2 é a variância do filtro. O processo de suavização resulta em uma imagem com menor influência de ruídos, prevalecendo as estruturas macroscópicas que da imagem [14]. Os melhores resultados foram obtidos com um filtro de dimensão 18×18 e um exemplo da suavização é apresentado na Figura 2(a).

Em seguida a segmentação do fundo é obtida pela binarização da imagem de modo que $B(x, y) = 1$ se $G(x, y) \leq T$ e $B(x, y) = 0$ se $G(x, y) > T$, onde B é a imagem binária e T é o limiar de binarização que separa regiões claras (objeto de interesse) de escuras (fundo). No estudo de caso, o limiar de binarização foi definido empiricamente em $T = 230$, um exemplo do resultado dessa etapa é apresentado na Figura 2(b). É possível observar também a presença de imperfeições na máscara binária devido a regiões da pimenta que refletiram a luz durante o processo de digitalização. Assumindo que o objeto de interesse são regiões conectadas, a imagem binarizada é submetida a um processo morfológico matemático de preenchimento de buracos para remover *pixels* provenientes do reflexos da luz. O algoritmo de preenchimento de buracos é o complemento da operação de reconstrução morfológica por dilatação, resultante de operação morfológica recursiva de dilatações geodésicas da imagem B com o elemento estruturante S , usando o marcador F como ponto de partida. Seja F uma imagem binária de mesmo tamanho de B , denominada imagem marcador, construída de modo que $F(x, y) = 1 - B(x, y)$, se (x, y) for um pixel sobre a borda da imagem B , e $F(x, y) = 0$, caso contrário. A dilatação geodésica da imagem B de tamanho 1 é denotada por $D^{(1)}(F) = (F \oplus S) \cap B$, onde o símbolo \oplus é o operador de dilatação, e $D_B^{(n)}(F) = D_B^{(1)}(D_B^{(n-1)}(F))$. A Figura 2(c) apresenta a máscara binária resultante da aplicação do filtro de preenchimento de buracos.

Com a máscara binária resultante das etapas anteriores é possível encontrar a região da pimenta através das coordenadas do primeiro e último pixels brancos na horizontal (xme e xma) e vertical (yme e yma). Com a localização destes quatro pixels é possível determinar as coordenadas que delimitam a região da pimenta.

Antes de recortar a imagem, é realizada a multiplicação *pixel a pixel* da imagem original pela máscara, resultando em uma imagem com a pimenta e fundo preto (Figura 2(d)), sem os ruídos presentes no fundo branco das imagens brutas. Por fim, a imagem é recortada de acordo com as coordenadas que delimitam a pimenta na imagem (Figura 2(e) e o fundo preto gerado na multiplicação pela máscara binária é substituído por um fundo branco, resultando na imagem final pré-processada (Figura 2(f)). Ao final do pré-processamento, as imagens estão prontas para a extração de características.

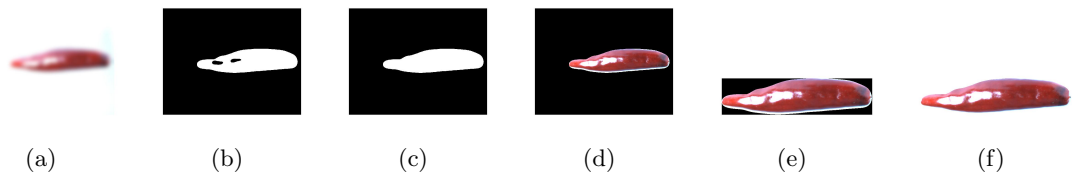


Figura 2: (a)Aplicação do filtro Gaussiano; (b)Binarização; (c)Preenchimento Morfológico de Buracos; (d)Multiplicação da Imagem pela Máscara; (e)Corte Final; (f)Substituição do Fundo.

4 Extração de Características

Como o presente trabalho tem o objetivo de identificar espécies de pimentas a partir de imagens, o vetor de características (Φ) será composto apenas por dados relacionados aos atributos físicos das pimentas.

A construção do vetor de características é realizada utilizando as imagens resultantes da etapa de pré-processamento. A partir de tais imagens, os dois primeiros atributos do vetor de características estão relacionados às dimensões das pimentas. Várias pimentas podem ser diferenciadas pelos seus formatos, dessa maneira é possível utilizar o número de linhas x e o número de colunas y de cada imagem representada na forma matricial como uma aproximação da altura e largura da pimenta.

As demais características do vetor são os momentos de cores das imagens. Imagens coloridas são representadas pelo computador como matrizes de três dimensões $F(x, y, z)$, onde as dimensões x e y representam as linhas e colunas da matriz, respectivamente, e z varia dependendo do espaço de cores da imagem. Os momentos de cores, por sua vez, são informações estatísticas relacionadas ao valor dos *pixels* das imagens [14]. Neste trabalho são utilizados três momentos de cores, a média $\mu = \frac{1}{xy} \sum_{i=1}^x \sum_{j=1}^y f_{ij}$, o desvio padrão $\sigma = \frac{1}{xy} \sum_{i=1}^x \sum_{j=1}^y (f_{ij} - \mu)^{1/2}$ e a simetria $\gamma = \frac{1}{xy} \sum_{i=1}^x \sum_{j=1}^y (f_{ij} - \mu)^{1/3}$, onde $f_{i,j}$ é o valor do *pixel* na posição i, j .

As imagens originais foram obtidas no padrão RGB. Contudo, para obter maior variedade de características, as imagens foram convertidas para outros dois padrões de cores via software: HSV e CIE Lab. A extração dos momentos de cores resultam um total de 27 características, portanto, o vetor de características (Φ) pode ser descrito da seguinte forma:

$$\Phi = \left(\begin{array}{c} x, y, \mu_r, \sigma_r, \gamma_r, \mu_g, \sigma_g, \gamma_g, \mu_b, \sigma_b, \gamma_b, \mu_h, \sigma_h, \gamma_h, \\ \mu_s, \sigma_s, \gamma_s, \mu_v, \sigma_v, \gamma_v, \mu_L, \sigma_L, \gamma_L, \mu_a, \sigma_a, \gamma_a, \mu_b, \sigma_b, \gamma_b \end{array} \right) \quad (1)$$

cada vetor de características é a representação de uma imagem. Os vetores serão normalizados e utilizados para o treinamento, validação e testes da rede neural.

5 Classificação com Redes Neurais PMC

Redes neurais artificiais são modelos computacionais inspirados no sistema nervoso humano e possuem a capacidade de aquisição e manutenção do conhecimento. O processo de aprendizagem se dá pela apresentação de conjuntos de exemplos e resultados esperados.

Por meio de tais informações são realizadas mudanças nos valores dos pesos sinápticos que são responsáveis pelas alterações no comportamento da rede. Dessa forma, o processo de ajuste dos pesos sinápticos tem a finalidade de sintonizar a rede para que seja capaz de classificar novas informações [3].

Neste trabalho foi utilizada uma rede do PMC, que tem sido utilizada com relativo sucesso em problemas de regressão, aproximação de funções e reconhecimento de padrões [5, 15]. Esta classe de redes neurais possui uma arquitetura de alimentação para frente, tendo pelo menos uma camada oculta e uma camada de saída em sua composição [5].

Alguns trabalhos mostram que uma única camada oculta já é suficiente para obtenção de bons resultados [2, 5, 13], e por esse motivo esta configuração foi utilizada. A função de ativação utilizada na camada oculta foi a sigmóide, sendo escolhida por ser totalmente diferenciável e por resultar em uma normalização adequada (entre 0 e 1). Já a função de ativação utilizada na camada de saída foi a *softmax*. Tal função transforma o vetor de saída em uma distribuição de probabilidade, onde a soma dos valores dos neurônios da camada de saída é igual a 1. Cada posição do vetor representa uma espécie de pimenta e o valor nesta posição representa a probabilidade da pimenta pertencer à espécie em questão. A posição com o valor mais alto é definida como a espécie da pimenta que se deseja classificar. O sinal de saída da camada oculta é utilizado como entrada para a camada de saída, composta por 24 neurônios, onde cada neurônio representa uma das espécies de pimenta presentes na base de dados.

Antes de executar o treinamento da PMC é preciso determinar os conjuntos de treinamento, validação e teste. Uma vez que o banco de dados é formado por 360 imagens, sendo 15 imagens de cada uma das 24 espécies de pimenta, optou-se por utilizar 9 imagens de cada espécie para treinamento, 3 para validação e outras 3 para teste. Na etapa de treinamento da rede PMC foi utilizado aprendizado supervisionado e o algoritmo *backpropagation* com gradiente conjugado escalonado, que, de acordo com [11], minimiza o erro mais rapidamente que outras opções como *backpropagation* com *momentum* ou gradiente conjugado.

A definição de critérios de parada para a etapa de treinamento da PMC é de suma importância para garantir que a rede possua boa eficiência na classificação e boa capacidade de generalização, evitando o *overfitting*. Neste sentido, foram definidos como 6 o número máximo de erros de validação e como 1^{-6} e menor erro de gradiente.

6 Resultados

Os resultados apresentados na Tabela 1 foram obtidos com a rede PMC contendo 10, 20, 29 e 50 neurônios na camada oculta e 24 neurônios na camada de saída. Para cada uma das configurações, a PMC foi executada 30 vezes, com os conjuntos de treinamento, validação e teste sendo separados de forma aleatória em todas as execuções.

É possível observar que a configuração com melhor precisão e acurácia é aquela que possui 29 neurônios na camada oculta. As PMC's com menores quantidades de neurônios na camada escondida apresentaram taxas de acerto médio menores e maior variação na classificação, assim como a configuração com 50 neurônios. Tais resultados já eram

esperados, uma vez que um maior número de neurônios na camada oculta não implica em melhores precisão e acurácia do classificador [16]. Por outro lado, [3] aponta que uma estratégia é utilizar a quantidade de características na camada de entrada como parâmetro para determinar a quantidade de neurônios na camada escondida. De fato, a configuração com 29 neurônios apresentou a maior taxa de acerto médio na classificação e o menor índice de variação nas 30 realizações, o que indica que essa pode ser uma estratégia válida.

Configuração	Média(%)	Desvio Padrão(%)	Mínimo(%)	Máximo(%)
10 Neurônios	94,21	4,08	83,33	100
20 Neurônios	95,55	2,2	90,28	100
29 Neurônios	97,36	1,6	94,44	100
50 Neurônios	96,89	1,95	93,05	100

Tabela 1: Taxa de acerto na classificação para 10, 20, 29 e 50 neurônios na camada oculta

7 Conclusões

Este trabalho propôs uma metodologia para a identificação automática de espécies de pimentas. Para validar o modelo proposto foi construída uma base de dados de imagens no espectro visível contendo 24 diferentes espécies de pimentas. A partir dos testes realizados foi possível obter, por mais de uma vez, 100% de taxa de acerto na classificação das espécies de pimenta. Na média, a taxa de acerto foi de 97,36% para uma PMC com 29 neurônios na camada oculta. Os resultados obtidos indicam que a metodologia é válida e suficiente para classificar diversas espécies de pimentas.

Como trabalhos futuros espera-se aperfeiçoar as técnicas de extração de características e de segmentação de imagens para viabilizar aplicações de classificação com imagens obtidas fora de ambiente controlado, assim como testar metodologias diferentes para a classificação das pimentas. Adicionalmente, espera-se enriquecer a base de dados com mais pimentas e espécies.

Referências

- [1] M. Ataş, Y. Yardimci and A. Temizel. *A new approach to aflatoxin detection in chili pepper by machine vision*. Computers and electronics in agriculture, 87:129-141, 2012.
- [2] C. M. Bishop. *Neural Networks for Pattern Recognition*. Oxford university press, 1995, ISSN: 0198538642.
- [3] A. de P. Braga, A. C. P. L. F. Carvalho and T. B. Ludermir. *Redes Neurais Artificiais—Teoria e Prática*. Editora LTC, 2000, ISSN: 8573077182.
- [4] A. R. da Silva, P. R. Cecon, E. R. do Rêgo and M. Nascimento. *Avaliação do coeficiente de variação experimental para caracteres de frutos de pimenteiras*. Ceres, 58, 2015.

- [5] H. B. Demuth, M. H. Beale, O. De Jess and M. T. Hagan. *Neural Network Design*. 2014, ISSN: 0971732108.
- [6] H. Kalkan, P. Beriat, Y. Yardimci and T. C. Pearson. *Detection of contaminated hazelnuts and ground red chili pepper flakes by multispectral imaging*. Computers and Electronics in Agriculture, 77:28-34, 2011.
- [7] F. Kurtulmus, I. Alibas and I. Kavdir. *Classification of pepper seeds using machine vision based on neural network*. International Journal of Agricultural and Biological Engineering, 9(1):51, 2016
- [8] N. Kussui, M. Lavreniuk and S. Skakun. *Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data*. IEEE Geoscience and Remote Sensing Letters. 14(5):778-782, 2017.
- [9] Z. Lu, S. Lu, S. Wang, Y. Li, Y. Zhang and H. Lu. *A fruit sensing and Classification System by Fractional Fourier Entropy and Improved Hybrid Genetic Algorithm*. In 5th International Conference on Industrial Application Engineering (IIAE). Kitakyushu, Institute of Industrial Applications Engineers, Japan. 2017.
- [10] C. H. C. Matos, A. Pallini, M. Vezon, D. D. M. Rezende and R. C. P. Freitas. *Caracterização morfológica e classificação da superfície foliar de pimentas quanto à presença de tricomas e domácias*. Horticultura Brasileira, 29:181-186, 2011.
- [11] M. F. Møller. *A scaled conjugate gradient algorithm for fast supervised learning*. Neural networks, 6:525-533, Elsevier, 1993.
- [12] R. S. Neitzke, S. Z. Fischer, R. L. Barbieri and R. O. Treptow. *Pimentas ornamentais: aceitação e preferências do público consumidor*. Horticultura Brasileira, 34:102-109, 2016.
- [13] H. V. Nogueira, A. R. Galvão Filho, S. Cabanelas, G. T. Laureano and C. J. Coelho. *Verificação da Autenticidade de Assinaturas Manuscritas Utilizando Redes Neurais Artificiais*. Congresso Brasileiro de Automática, 2012.
- [14] T. D. Ping. *A review on image feature extraction and representation techniques*. International Journal of Multimedia and Ubiquitous Engineering, 8:385-396. 2013.
- [15] H. Ramchoun, M. A. J. Idrissi, Y. Ghanou and M. Ettanouil. *Ramchoun, Hassan, et al. "Multilayer Perceptron: Architecture Optimization and Training*. IJIMAI, 4(1), 26-30, 2016.
- [16] Silva, IN da and Spatti, Danilo Hernane and Flauzino, Rogério Andrade *Redes neurais artificiais para engenharia e ciências aplicadas* Editora: Artliber, 2010.
- [17] P. T. Souza, A. V. Rossi. *Determinação espectrofotométrica indireta de capsaicinoides em pimentas Capsicum a partir da reação com o complexo de Co (II) com 4-(2-piridilazo) resorcinol*. Química Nova, 2014.