

SELEÇÃO DE TOMATES PARA PROCESSAMENTO INDUSTRIAL POR MEIO DE REDES NEURAIS APLICADAS EM SISTEMA DE VISÃO COMPUTACIONAL

CLÁUDIO DENIS

*Faculdade de Tecnologia Termomecânica
Estrada dos Alvarenga, 4001, 09850-550 São Bernardo do Campo, SP, Brasil
cddenis@uol.com.br*

WÂNDERSON O. ASSIS

*Escola de Engenharia Mauá, Instituto Mauá de Tecnologia
Praça Mauá, 1, 09580-900, São Caetano do Sul, SP, Brasil
wanderson.assis@maua.br*

Abstract— This paper presents a method of tomatoes' selection based on images processing with emphasis on standard color treatment applying segmentation of colored images. The approach looks for a robust classification with respect to the variation of illumination and color brightness, which tries to be tolerant to errors in the sampling process. The classification system uses a webcam or digital camera, an algorithm developed in object oriented language and a Multilayer Perceptron neural network. The application is intended to automation of tomatoes' classification process where the tomatoes are used for industrial processing.

Keywords— Neural networks, computational vision, image processing.

Resumo— Este trabalho apresenta uma metodologia para a classificação de tomates baseada em processamento de imagens com ênfase para o tratamento do padrão cor e aplicando métodos de segmentação de imagens coloridas. A abordagem busca uma classificação robusta à variação de grandezas tais como iluminação ou brilho da cor, além de ser tolerante a erros no processo de amostragem. Utiliza-se na classificação um sistema constituído de webcam ou câmera digital, um algoritmo desenvolvido utilizando linguagem orientada a objetos para captação e processamento dos atributos de cores dos pixels da imagem e rede neural MLP-Multilayer Perceptron. A aplicação é destinada à automação do processo de classificação de tomates utilizados no processamento industrial.

Palavras Chave— Redes neurais, visão computacional, processamento de imagem.

1 Introdução

Mesmo depois da automação de muitos setores industriais, a inspeção de tomates é ainda realizada por humanos com melhor qualidade do que por meios artificiais. Grande parte das abordagens automáticas para esta tarefa concentram-se na análise de algum tipo de imagem digital do espectro visível ou não visível, tais como os raios X, para avaliar tais frutos.

As inspeções de tomates realizadas por humanos usualmente tomam como base padronizações realizadas por instituições especializadas. Para o caso específico do tomate no Brasil, uma das escassas referências é o Centro de Qualidade em Horticultura (CEA-GESP, 2006), que tornou-se padrão brasileiro e propõe a classificação de tomates baseada em padrões de cor e qualidade, o que também é proposto nas normas do departamento de agricultura dos Estados Unidos, conforme a norma 51.3310 revista em Janeiro de 1997 para classificação de tomates para processamento.

Alguns trabalhos têm sido propostos utilizando técnicas de processamento de imagens digitais para a análise da forma e tamanho de frutos e para a detecção de defeitos (Desmukh, 2005) (Usó et al, 2003).

Para o problema do reconhecimento de padrões no domínio da agricultura, algumas abordagens utilizam métodos estatísticos, redes neurais artificiais e análise de características (Kondo et al, 2000).

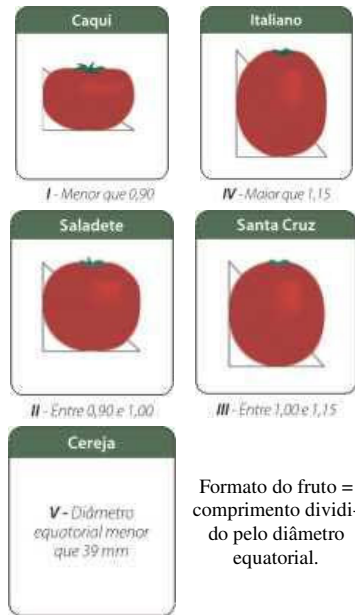
O presente trabalho propõe uma metodologia para a classificação de tomates para processamento industrial baseada em informações visuais, com ênfase para o tratamento do padrão de cor, e utilizando redes neurais.

2 Classificação de Tomates

Classificar quer dizer separar o produto por variedade, tamanho, cor e qualidade de maneira que a separação resulte em conjunto de frutos com aparência uniforme. O tomate é classificado por: Grupos (formatos: I, II, III, IV e V); Cores (Vermelho, Rosado, Laranja e Amarelo; Subgrupos: I – Pintando; II – Colorido; III – Maduro) e Classes (0, 40, 50, 60, 70, 80, 90 e 100). A classificação por grupo é utilizada para caracterizar os grupos de cultivares conforme figura 1. A classificação por Classes é definida pelo diâmetro equatorial do fruto. Assim, o tomate será classificado em oito classes: 0 para diâmetro menor que 40mm; 40 para diâmetro maior ou igual a 40 até 50mm; 50 para diâmetro maior ou igual a 50 até

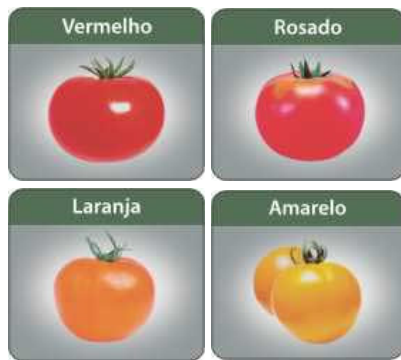
60mm; 60 para diâmetro maior ou igual a 60 até 70mm; 70 para diâmetro maior ou igual a 70 até 80mm; 90 para diâmetro maior ou igual a 90 até 100mm; 100 para diâmetro maior que 100mm (CE-AGESP, 2006).

Figura 1. Classificação de Tomates por Grupos.



Em relação à classificação das frutas pelo parâmetro cor, o padrão brasileiro de classificação de tomates segundo o (CEAGESP, 2006), propõe as seguintes categorias de cores: Vermelho, Rosado, Laranja e Amarelo; como ilustrado na figura 2.

Figura 2. Classificação de Tomates por Cores.



Existem dois tipos de tomates vermelhos sendo o primeiro deles usado para saladas e culinária e outro para processamento industrial visando obter produtos como molhos, polpa de tomate, ketchup, etc. O amadurecimento do tomate determina a mudança de coloração da sua casca e caracteriza três subgrupos: I – Pintando; II – Colorido; III – Maduro, como ilustrado na figura 3, sendo esta classificação tipicamente adotada para classificação do tomate para processamento industrial.

Dentro deste contexto, este trabalho apresenta uma metodologia para a classificação de tomates para processamento industrial baseada em informações visuais, com ênfase para o tratamento do padrão de cor.

Figura 3. Classificação de Tomates para Processamento Industrial.



3 Padronização de Cores

O órgão padronizador internacional na área de colorimetria é o CIE – Comissão Internacional de Iluminação (*Commission Internationale de l'Eclairage*). A maioria de suas padronizações foi estabelecida na década de 30 e permanece até hoje, tais como RGB e XYZ. Contudo, alguns outros padrões estabeleceram-se mundialmente no trabalho com cores, mesmo sem as padronizações do CIE, adotando até mesmo outras formas de representação que não a tricromática.

3.1 CIE - RGB

Uma vez que o sistema visual humano possui células capazes de detectar comprimentos de onda de três tamanhos distintos: vermelho, verde e azul, nas faixas alta, média e baixa do espectro, respectivamente, é imediata a utilização de comprimentos de onda em faixas similares para a composição de representação tripla. Embora com o passar dos anos os comprimentos de onda adotados tenham-se alterado consideravelmente atualmente definem-se os comprimentos de onda para o sistema RGB como os mostrados na tabela 1.

Tabela 1. Comprimento de onda por picos das curvas de sensibilidade das cores para o sistema CIE - RGB.

Comprimento de onda (λ)	Cor
700 nm	Vermelho
546 nm	Verde
435,8 nm	Azul

Os modelos de especificação de cor baseados no espaço vetorial, como o modelo RGB, apesar de práticos do ponto de vista computacional, são inadequados do ponto de vista de comunicação com o usuário e com relação à especificação simples e intuitiva de uma determinada cor. Uma dificuldade de se especificar cor no sistema RGB ocorre quando se deseja determinar uma cor de tonalidade e luminância conhecidas e precisamos apenas dosar a saturação, diminuindo ou aumentando a quantidade de branco na

cor. Para realizar essa mudança devemos alterar três componentes de cor e essa alteração não tem qualquer relação intuitiva com os parâmetros percentuais de tonalidades, luminância e saturação (Gomes & Velho, 1998).

3.2 Sistema CIE XYZ

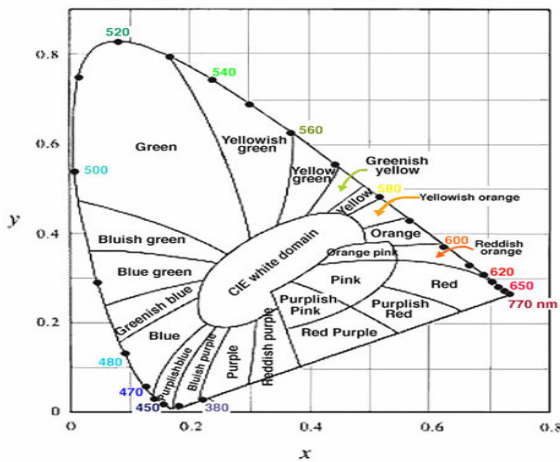
O modelo XYZ usa primárias imaginárias X, Y e Z para descrever o espaço de cores e com isso pode-se definir um gráfico de representação para que as grandezas fotométricas sejam encontradas mais facilmente, porém representando as cores que não estão definidas dentro do espectro visível. A representação desse sistema de cores é mostrado na figura 4, onde o formato de uma figura encontra-se inserido dentro de um triângulo representando as três primárias do sistema RGB e todas as cores visíveis são representadas com coordenadas positivas. Os espaços dentro da figura são locais de cores sem luminosidade e isto foi colocado ao longo de X para Z no sistema XYZ. Toda luminosidade é expressa em Y, sendo que o local de referência Y foi escolhido para enquadrar o domínio das cores reais. O ponto de energia branco (W) foi escolhido para ter coordenadas de cromaticidade iguais (0,33, 0,33). A soma das coordenadas de cromaticidade é igual a 1,0. Portanto, Z pode ser calculada pelo conhecimento das coordenadas X e Y.

A mudança de coordenadas entre os sistemas CIE-RGB e CIE-XYZ é dada por:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.049 & 0.31 & 0.20 \\ 0.17 & 0.81 & 0.01 \\ 0.00 & 0.01 & 0.99 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

Usando essa transformação de mudança de base podemos calcular grandezas colorimétricas do sistema CIE-XYZ a partir do sistema CIE-RGB.

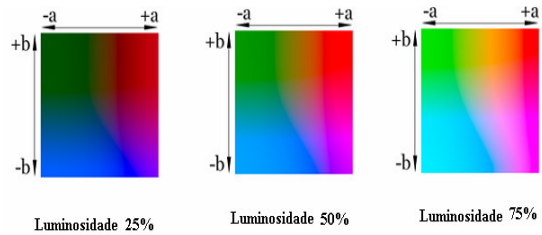
Figura 4. Diagrama de Cromaticidade CIE XYZ
Fonte: (Gomes & Velho, 1998).



3.3 Sistema CIE L*a*b*

CIE L*a*b*(ou CIE LAB) é o modelo de cor mais completo usado convencionalmente para descrever todas as cores visíveis pelo olho humano. O asterisco após o L, a e b são parte do nome completo L*, a* e b*. Os três parâmetros representam respectivamente L* luminosidade da cor (L* = 0 representa o preto e L* =100 representa o branco) conforme figura 5, a* está posicionado entre magenta e verde e b* está posicionado entre amarelo e azul.

Figura 5. Níveis de Luminosidade L* para CIE L*a*b*.



CIE L*a*b* foi baseado diretamente no CIE XYZ espaço de cores como uma tentativa de linearizar a perceptividade das diferenças de cores. A relação não linear para L*, a* e b* são intencionais para imitar a resposta logarítmica do olho.

A conversão do sistema XYZ para CIE L*a*b* pode ser obtida por meio da transformação direta:

$$L^* = 116 * (Y / Y_n)^{1/3} \quad \text{para} \quad Y / Y_n > 0.008856 \quad (0 \leq L^* \leq 100) \quad (2)$$

$$L^* = 903.3 * Y / Y_n \quad \text{para} \quad Y / Y_n \leq 0.008856 \quad (3)$$

$$a^* = 500 [f(X/X_n) - f(Y/Y_n)] \quad (4)$$

$$b^* = 200 [f(Y/Y_n) - f(Z/Z_n)] \quad (5)$$

onde:

$$f(t) = t^{1/3} \quad \text{para} \quad t > 0.008856 \quad (6)$$

$$f(t) = 7.787t + 16/116 \quad \text{para} \quad t \leq 0.008856 \quad (7)$$

Aqui Xn, Yn e Zn são os valores de triestímulos de CIE XYZ da referência do iluminante (ponto branco). A divisão de f(t) em dois domínios foi feita para prevenir uma rampa com valor infinito em t = 0, sendo que f(t) foi assumido ser linear abaixo para t = t₀ para igualar o t^{1/3}, parte da função em t₀ para ambos os valores de rampa ou seja:

$$t_0^{1/3} = at_0 + b \quad (8)$$

$$1/(3t_0^{2/3}) = a \quad (9)$$

O valor de b foi escolhido para ser 16/116. As duas equações acima podem ser resolvidas para a e t₀.

$$a = 1/(3\delta^2) = 7.787037... \quad (11)$$

$$t_0 = \delta^3 = 0.008856... \quad (12)$$

onde $\delta = 6 / 29$ é um parâmetro constante. Observe que $16 / 116 = 2\delta / 3$

4 Segmentação de Imagens

A idéia da segmentação tem origem nos seres humanos que desenvolvem agrupamento visual baseado em características como proximidade, similaridade e continuidade. Na visão computacional, este processo é denominado segmentação, pelo qual entende-se a extração de parâmetros que permitam associar regiões da imagem com objetos na cena. Nesse sentido, busca-se dividir a imagem digital em regiões disjuntas, tais que uma *pixel* pertença apenas a uma das regiões.

Um dos algoritmos utilizados no processo de segmentação de imagens é o *crescimento de região* pela análise de vizinhança. Trata-se de um processo pelo qual *pixels* com características semelhantes em uma vizinhança são agregados em regiões maiores, sucessivamente, partindo-se de um conjunto de pontos denominado a "semente" da região (González & Woods, 2003). Propõe-se que, para a segmentação de imagens todo *pixel* de uma vizinhança que possuir uma diferença menor que um limite T do *pixel* semente, será enquadrado na região deste *pixel*. Observe-se que, embora faça uso da imposição de limiares, este método leva em consideração a informação da vizinhança do ponto, o que é desconsiderado pelo método anterior.

Tanto os algoritmos baseados em análise de descontinuidades quanto os que efetuam o agrupamento de regiões frequentemente utilizam-se da comparação de níveis de intensidade luminosa dos pixels da imagem para a tomada de decisões e da imposição de limiares. Contudo, tratando-se de imagens segmentadas por cores, nem sempre a diferença entre o valor luminoso de dois pontos é suficiente para proceder segmentações, visto que as noções cognitivas de cor não são homogeneamente distribuídas pelo espectro de cores. Tratando-se do paradigma da imposição de limiares, isto implica que as regiões representativas de cada cor pertencente ao espectro de cores não são necessariamente separáveis por sólidos de tamanhos iguais no espaço tricromático de cores e, mais que isso, não são necessariamente separáveis por sólidos regulares.

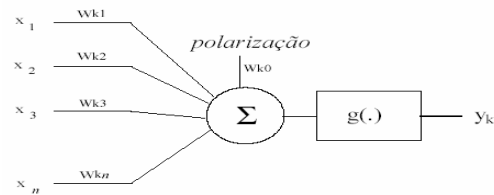
O método da imposição de limiares pode ser considerado um método de classificação, já que subdivide o espaço de entrada em regiões discretas que podem ser interpretadas como classes. A fim de obter melhores classificações, ferramentas mais elaboradas podem ser utilizadas, em especial, ferramentas do âmbito da inteligência artificial, onde o problema de classificação é de grande relevância. Uma abordagem

natural e com grande habilidade de generalização é a das *redes neurais artificiais* (RNAs).

5 Redes Neurais Artificiais (RNAs)

O neurônio pode ser entendido como um dispositivo que tem muitas entradas e apenas uma saída. Nesse sentido, Rummelhart, Hinton e Williams (Rummelhart et al, 1986) propuseram um modelo não-linear para o neurônio biológico, que é amplamente aceito pela comunidade científica:

Figura 6. Modelo Não-Linear de um Neurônio.



onde:

- X é o vetor de entrada do neurônio k ;
- x_i é a excitação de entrada na sinapse i ;
- y_k é a resposta (ou saída) do neurônio k ;
- w_{ki} é o peso sináptico da entrada i do neurônio k ;
- w_{k0} é o peso sináptico da entrada de polarização do neurônio k ;
- $g(.)$ é a função de ativação do neurônio.

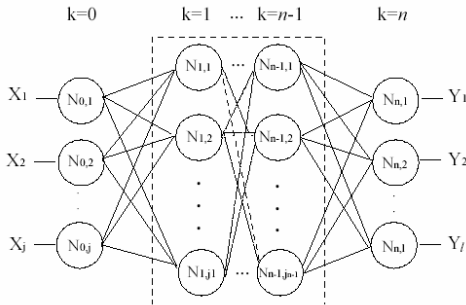
Esse modelo é basicamente um *perceptron* com uma função de ativação semilinear (Haykin, 2004). O *perceptron* modela um neurônio processando uma soma ponderada de suas entradas e submetendo o resultado a uma camada de processamento de limiares ($g(.)$) (figura 6). No caso dos *perceptrons*, essa camada consiste basicamente de uma função degrau, mapeando as saídas nos níveis binários 0 ou 1, naturalmente uma função altamente não-linear.

O *perceptron* foi desenvolvido por Rosenblatt em 1962 (Haykin, 2004). Rummelhart, Hinton e Williams (RUMMELHART et al, 1986), anos mais tarde, desenvolveram um princípio de aprendizado para redes com funções de ativação semilineares que foi denominado *error backpropagation*, ou a retropropagação do erro, que propiciou a arquitetura e o treinamento do *Multilayer perceptron* (MLP) ou *perceptron* multicamada.

O MLP é uma rede com a topologia apresentada na figura 7. Os MLPs podem ser entendidos como uma coleção de neurônios dispostos em k camadas, usualmente classificadas como camada de entrada, primeira camada da rede, camada de saída, última camada da rede e camadas ocultas, todas as demais. As j dimensões do estímulo são apresentadas aos j neurônios da camada de entrada. As saídas de cada

neurônio são propagadas como entradas para todos os neurônios da camada seguinte, sucessivamente, até a resposta final da rede, de dimensão l igual ao número de neurônios da camada de saída. Por N_{kj} entende-se o neurônio disposto na j -ésima posição da camada k e por w_{kja} entende-se o peso sináptico associado à entrada a do neurônio na posição j da camada k .

Figura 7. A Rede Neural Multicamada.



6 Sistema de Classificação de Tomates com RNA

O sistema de seleção de tomates para processamento industrial é constituído de uma câmera do tipo webcam que é responsável pela captura da imagem e transmite esta imagem digitalizada para o computador. Um programa foi desenvolvido em Visual Basic especialmente para capturar esta imagem e criar uma interface para o usuário como ilustrado na figura 8; para elaborar este programa utilizou-se da linguagem orientada a objetos. Quando solicitado por meio de um botão, o programa analisa a imagem capturada e efetua a varredura nos *pixels*. Um método de segmentação da imagem baseado no agrupamento de regiões é utilizado para separar a imagem do fundo utilizando os dados de cores RGB. Em seguida os dados são convertidos para o sistema de cores CIE $L^*a^*b^*$. No passo seguinte, estes dados são tratados, e enviados para um classificador formado por rede neural artificial. A rede neural utilizada (figura 9) é a MLP-*Multilayer Perceptron* desenvolvida em software Matlab® que tem três entradas: L, a, b, uma camada oculta de neurônios e três saídas: P(pintando), C(colorido), V(Vermelho) que apresentam a classificação do tomate.

7 Resultados Obtidos

A rede foi treinada utilizando imagens de 10 exemplares de tomates com classificação Pintando, 10 exemplares de tomates com classificação Colorido e 10 exemplares de tomates com classificação Maduro, todos estes com a variação de luminosidade de 100lux até 1.000lux, totalizando 300 imagens. Para cada imagem foram obtidas as médias dos valores correspondentes em R, G e B, e $L^*a^*b^*$, obtendo excelente convergência na rede neural com o erro

Figura 8. Interface do Programa Desenvolvido em Visual Basic.

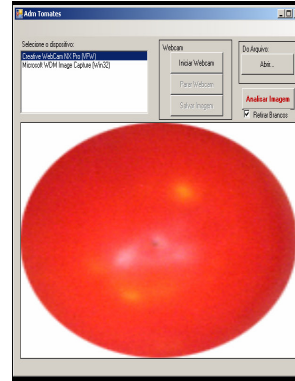
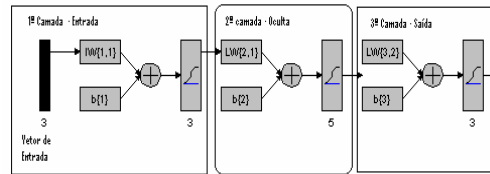


Figura 9. A Rede Neural do Sistema de Seleção de Tomates.



quadrático após treinamento atingindo o valor de 1.52726×10^{-12} após 10000 épocas, portanto tendendo para zero.

No treinamento utilizando a função guide *ntool* do Matlab® foram utilizados os seguintes parâmetros de configuração: épocas (passos: 10000; μ :0.001; μ_{dec} : 0.1; μ_{inc} :10; μ_{max} : 1000000000; show: 25; time: inf. O resultado do treinamento é um vetor com pesos neurais(pesos sinápticos) e bias dados pela tabela 2.

Tabela 2. Pesos Sinápticos e Bias da Rede Neural após Treinamento.

Pesos Sinápticos e Bias	
Entrada IW{1,1}=Weight to layer1 from input1	IW{1,1}: [18.3306 20.6468 -25.8618; 10.1673 0.98737 43.5514; 25.2936 30.7558 -20.2817]
Intermediária IW{2,1}Weight to layer2 from output layer1	IW{2,1}: [-1.8863 8.186 4.4592; 5.1632 -4.7303 -7.9301; 4.006 -3.8519 23.1306; -9.2816 -0.3223 9.1606; - 10.9538 -4.8961 -10.1071]
Saída IW{3,2}Weight to layer3 from output layer 2	IW{3,2}: [53.8761 -5.7055 28.4308 - 1.0163 0.70641; -19.0448 0.26531 22.3815 -10.9573 11.4022; -2.4678 3.3907 -2.5071 15.6495 -11.3612]
Bias entrada b{1}	b{1}: [-4.7668; -33.3561; -1.4579]
Bias intermed. b{2}	b{2}: [8.1358; -4.6668; -4.2325; - 12.9001; 8.2349]
Bias Saída b{3}	b{3}: [54.3266; -20.3186; -7.5748]

Alguns resultados abaixo ilustram a eficiência da rede neural e do sistema de seleção desenvolvido.

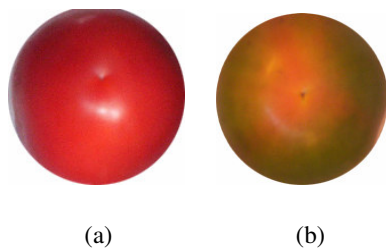
Exemplo1: Tomate Maduro a 100 lux (figura 10(a)).

Vetor de solução: [0.01183;0.13343;0.94463]. O valor máximo obtido 0.94463 representa a identificação do tomate com classificação Maduro.

Exemplo 2: Tomate Pintando a 1000 lux (figura 10(b)).

Vetor de solução: [0.95342;0.083159;0.05455]. O valor máximo obtido 0.95342 representa a identificação do tomate com classificação Pintando.

Figura 10. Exemplos ilustrativos: (a) Tomate Maduro a 100 lux (b) Tomate Pintando a 1000 lux.



Os resultados dos testes resultaram em um índice de acerto que aproxima-se de 100%, se assemelhando aos resultados que seriam obtidos com a seleção não automática, efetuada por humanos.

3 Conclusão

Um método de seleção de tomates baseado no processamento de imagens, aplicando técnicas de segmentação de imagens coloridas e utilizando redes neurais artificiais foi proposto neste trabalho. A solução utiliza um sistema formado por uma câmera do tipo *webcam* que transmite a imagem para o computador e um programa desenvolvido especialmente para efetuar o processamento da imagem em linguagem orientada a objetos. A rede neural utilizada é a MLP-Multilayer Perceptron que tem três entradas: L^* , a^* , b^* , uma camada oculta de neurônios e três saídas: P(pintando), C(colorido), M(Maduro) que apresentam o tipo do tomate que foi classificado. Aplicando este sistema de classificação em um ambiente que apresenta variações de luminosidade observa-se que a rede neural apresenta um bom rendimento e excelente índice de acerto, atendendo a proposta do trabalho, além de utilizar componentes de baixo custo no desenvolvimento da proposta.

Referências Bibliográficas

- Barnard, E. and Casanet, D. (1991) Invariance and neural nets. *IEEE Transactions on neural networks* Vol.2, n.5, September.
- Bo L., Huosheng H., Libor S. (2003) An adaptive color segmentation algorithm for Sony legged robots. *Proceedings of the 21st IASTED International Conference Applied Informatics*. February 10-13, Innsbruck, Austria.
- CEAGESP – Centro de Qualidade em Horticultura. (2006) Classificação do tomate. *ABH*, São Paulo.
- Desmukh K. S; Shinde G. N. (2005) An adaptive color image segmentation. *ELCVIA - Electronic Letters on Computer Vision and Image Analysis*, Computer Vision Center da Universidade Autònoma de Barcelona, Barcelona, Spain.
- Gomes, J.; Velho, L. (1998) Computação Gráfica: Imagem. *Instituto de Matemática pura e aplicada - IMPA*. Rio de Janeiro.
- Gonzales, R. C. and Woods, R. E. (2003) Processamento de Imagens Digitais. *Editora Edgard Blucher Ltda*, SP.
- Haykin, S. (2004) Redes neurais: princípios e prática. 2^o ed. *Editora Bookman*. Porto Alegre.
- Kondo, N.; Ahmad, U.; Monta, M. and Murase, H. (2000). Machine vision based quality evaluation of Iokan orange fruit using neural networks. *Elsevier. Computers and electronics in agriculture*. Vol. 29.
- Littman, E.; Ritter, H. (1997) Adaptive color segmentation - A comparison of neural and statistical methods. *IEEE Transactions on Neural Networks*. Vol. 8, N.1, January.
- Rummelhart, D., Hinton, W. (1986) Learning internal representations by error propagation. *Parallel distributed processing*. MIT Press, Cambridge, 1986.
- Usó, M. A.; Pla, F.; Sevilha, P. G. (2003) Multispectral image segmentation by energy minimization for fruit quality estimation. Disponível em: <<http://www.vision.uji.es>> Acesso em: 29 janeiro 2006.