Journal of
Information Technology & Software Engineering

Capillary Dynamolysis Image Discrimination Using Neural Networks

Mehmet S. Unluturk[1]*, Sevcan Unluturk[2], Fikret Pazir[3] and Firoozeh Abdollahi[4]

[1]Izmir University of Economics, Department of Software Engineering, Izmir, Turkey
[2]Izmir Institute of Technology,Food Engineering Department, Izmir, Turkey
[3]Ege University, Food Engineering Department, Izmir, Turkey
[4]1510 - 1360 York Mills Road- North York ON-Toronto-M3A 2A3-Canada

Abstract

Quality differences between organic and conventional fresh tomatoes (unprocessed) and frozen tomatoes (processed) are evaluated by using a capillary rising picture method (capillary dynamolysis). The best pictures showing the differences most sharply between organic and conventional samples were prepared with 0.25-0.75% silver nitrate, 0.25-0.75% iron sulphate and 30-100% sample concentration. But visual description and analysis of these images is a major challenge. Therefore, a novel methodology called Gram-Charlier Neural Network methodology (GCNN) has been studied to classify these images. Two separate GCNNs have been created for fresh and frozen cases. They are trained with the pictures of organic and conventional tomato samples from these two cases. The 2048 x 1536 pixel chromatogram images were acquired in a lab and cropped to 1400 x 900 pixel images depicting either a conventional tomato or an organic tomato for each case. A set of 20 images from each case was utilized to train each Gram-Charlier Neural Network. A new set of 4 images from each case was then prepared to test each GCNN performance. In addition, Hinton diagrams were utilized to display the optimality of the GCNN weights. Overall, the GCNN achieved an average recognition performance of 100%. This high level of recognition suggests that the GCNN is a promising method for the discrimination of capillary dynamolysis images and its performance does not depend on whether the tomato sample is fresh or frozen.

Keywords: Organic tomato; Conventional tomato; Capillary dynamolysis images; Hinton diagrams; Neural network; Bayes optimal decision rule.

Introduction

There is a rising awareness of health and the environment in worldwide. Consumers have increasing interest on the quality of foods. Especially organic farming receives a considerable attention in recent years. The quality of foods from conventional and organic production is compared in numerous studies. Most of these comparative studies are based on the physico-chemical tests examining desirable and undesirable ingredients, pesticide residues and environmental contaminants, in addition to sensory tests and feed experiments with animals. Besides, nutritional studies and investigations using holistic methods of analysis have been also included [1].

Holistic methods (picture developing methods) including biocrystallisation (copper chloride crystallization), capillary dynamolysis and the round-filter chromatography have been used to examine foods from organic cultivation [2]. These methods are developed on the basis of images of the 'inner structure' of an organism. The inner structure of a living organism is believed to be related to the growth (i.e. farming system) and development of the organism. Different imaging methods are developed to reveal this 'inner structure' [3]. The capillary dynamolysis is one of these methods. It began to be applied by [4] to plant sap in the twenties of the 20th century. The method involves capillary dynamolysis first with an aqueous extract of plant sap followed, with intermediate drying, by a metal salt, usually silver nitrate, which serves as an indicator to develop the structured pattern. Any kind of additive will change the chromatogram. The process is influenced by the qualitative and quantitative variations in the macromolecules of the biological extracts, thus allowing food quality assessment [5]. The pictures are evaluated qualitatively describing the different form elements and their interactions [6]. Figure 1 illustrates the chromatograms produced on the basis of biological substances, such as plant extracts, exhibit distinctive structures covering the chromatography paper [7].

The pictures obtained by capillary dynamolysis method are evaluated visually by trained people, using a 3-4 zone concept and characteristic criteria's for describing the differences (see Figure 2) [2,6,8]. Although capillary dynamolysis method is promising to distinguish between organic and conventional food, they need further development and validation for routine application in food quality assessment. Even if the methodological problems have been overcome there remains the cognitive task of how the images and patterns are to be read (pattern recognition). In literature several trained authors have shown to be capable of judging and interpreting capillary pictures by connecting them correctly to different farming systems [6,9]. Although visual examination by trained people can be used as a tool for evaluation of the images, it requires high expertise and practice. Additionally, there is a lack of validated methodology for the quality assessment. On the other hand, computerized image analysis can be an alternative approach increasing the objectiveness of the picture forming methods and allow the analysis of large numbers of chromatogram images [3,10]. It is based on the interpretation of the images by using the fundamental knowledge of texture analysis [5]. To the best of our knowledge, there is no available literature on the application of computerized image analysis over evaluation of capillary dynamolysis test results.

The objective of the present study is to develop a technique based on a Gram-Charlier Neural Network approach to judge and interpret the images (pictures) of conventionally and organically grown tomatoes obtained by capillary dynamolysis method. The efficiency and the

*Corresponding author: Mehmet S. Unluturk, Izmir University of Economics, Department of Software Engineering, Sakarya Cad. No.156, Balcova 35330, Izmir, Turkey, E-mail: suleyman.unluturk@ieu.edu.tr

performance of the developed technique is compared by using images of fresh (unprocessed) and frozen (processed) tomatoes.

Materials and Methods

Sample preparations

Conventional and organic tomato fruits were obtained from Ege University Agricultural Faculty Research, Applied and Production Farm in Menemen (Izmir, Turkey) during the period of 2005 and 2006. Tomato fruits were quickly transferred to the laboratory. The fresh fruits were cleaned, chopped into small pieces and passed through a kitchen type blender (Braun MR 404). The homogenate was first passed through a cloth to remove debris particles and then filtered. The fruit juice was diluted to 30-100% with bi-distilled water. Silver nitrate and iron sulphate solutions were also prepared at a 0.25-0.75% concentrations with bi-distilled water.

The combination ratios for the sample extract silver nitrate and iron sulphate influence specific patterns of the chromatograms. Therefore, different ratios of sample, silver nitrate and iron sulphate concentrations were studied. The concentration levels were between 30 to 100%, 0.25 to 0.75% and 0.25 to 0.75% for sample, silver nitrate and iron sulphate, respectively.

Sample extracts are prepared at different concentrations ranging from 30% to 100% with bi-distilled water. 0.6 ml of these aqueous extract is allowed to rise in chromatographic paper (Whatmann no. 2 CHR) up to 2cm from the base line and dried for 3 hours in a room at 20°C and 60 % RH. Next the chromatographic paper is placed in a solution of 0.7 ml of silver nitrate solution prepared at different concentrations (0.25-0.75 %). Silver nitrate is allowed to migrate vertically up to the level approximately 1 cm over the juice-line in a glass bowl for 30 min. After drying, 2 ml of iron sulphate solution (0.25-0.75 %) is added and allowed to rise nearly to the top of the paper (12 cm) in 0.5 to 2 hours (Figure 3). Then the paper is dried for 3 hours. The last two stages take place in a dark room [11]. A total of 23 paper chromatograms were produced from organic and conventional fresh tomato fruits.

Image acquisition and capture

Images were captured using a digital camera model DMC-FZ5 (5 Mp) (Panasonic, NJ, USA). The camera was positioned vertically over the sample at a certain distance. The angle between the camera lens, the lighting source and ambient illumination were fixed and kept the same for all the sample pictures. After then the images were transferred and stored in a PC as a JPEG format of 'high resolution' and 'superfine quality'.

Image processing

All the algorithms for pre-processing of full images and image segmentation were written in MATLAB 6.5 (The MathWorks, Inc., MA, USA). MATLAB code for pre-processing and image segmentation of a single full image is given below:

s = ['org_2005_' int2str(i) '.jpg']; I = imread(s); BW = rgb2gray(I); BW=BW(501:1900,301:1200); BW = double(BW); d=get_coeffs(BW)

Inside *get_coeffs(BW)* function, the gray scale image was converted to a row vector (*x=convert2vec(bw)*). This row vector was then used to calculate the Gram-Charlier coefficients of the corresponding gray scale image, as the purpose of calculating Gram-Charlier coefficients is to reduce the number of input nodes in our neural network.

Image segmentation

Each 2048 x 1536 pixel RGB images was turned into a gray scale image using the 'rgb2gray' command of MATLAB. The final gray scale image was then cropped to 1400 x 900 pixel size (*BW=BW(501:1900,301:1200)*; where *BW* is the gray scale image). These images were then used in the classification process, as the aim of these software-based procedures is to eventually replace the human visual decision-making process.

Image Analysis: Theory

Bayes optimal decision rule

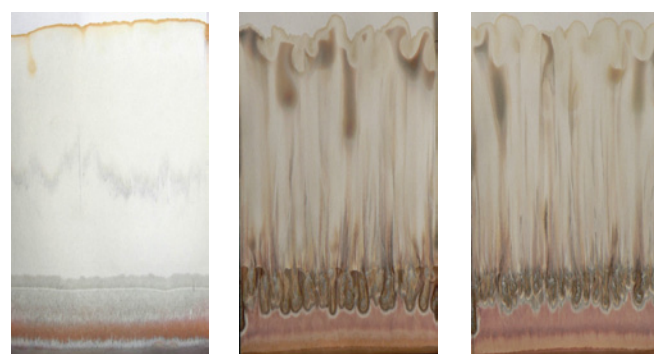Figures 4-7 show 2048 x 1536 pixel capillary rising pictures of

Figure 1: (a) Chromatogram obtained from basis of aqueous silver nitrate, iron sulphate(blank) (b) conventional growing fresh sweet red pepper chromatogram (c) organic growing fresh sweet red pepper chromatogram (Kuşçu, 2008).
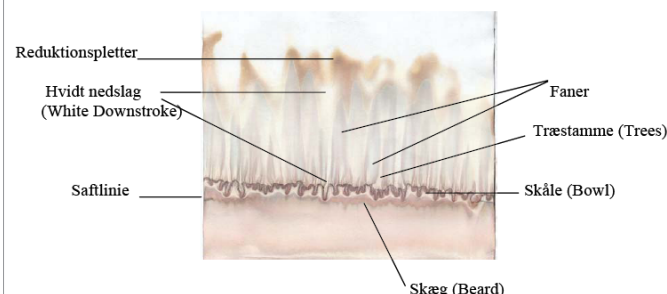

Figure 2: Definition of zone in capillary dynamolysis method [8].
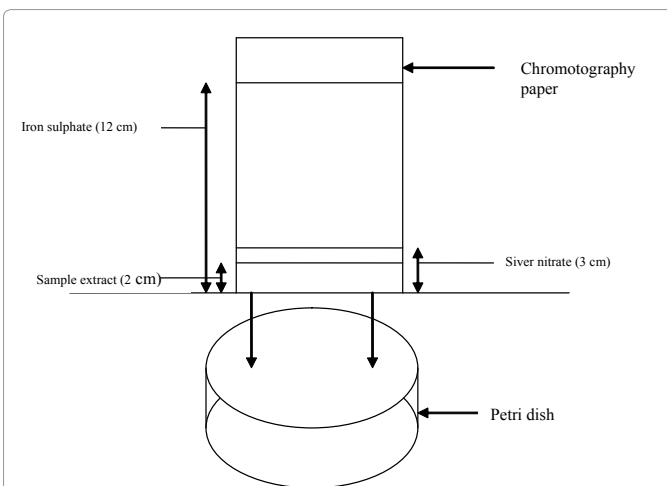

Figure 3: Application of capillary dynamolysis method

conventional and organic tomato for fresh and frozen cases respectively. Images were converted to grayscale images by eliminating the hue and saturation information while retaining the luminance using 'rgb2gray' command of MATLAB. These grayscale images were then cropped (see Figures 8-11). Let's represent one of these grayscale images as a row vector, obtained from the image matrix by combining its rows:

$$\vec{X} = \left[x(1), x(2), \cdots x(m) \right]$$

Where m is 1,260,000. Each element of the vector corresponds to a grayscale color value that is between 0 and 255. Then, there are two classes from which our input image might be drawn.

These classes are defined as:

H_o = Organic,

H_1 = Conventional.

The prior probability of an unknown image being drawn from class k is h_k (k=0, 1). The cost of making a wrong decision for class k is v_k. Note that the prior probability h_k and the cost probability v_k are taken as being equal, and hence can be ignored. The problem is to

Figure 6: Capillary rising picture of conventional frozen tomato (processed).

Figure 4: Capillary rising picture of conventional fresh tomato (unprocessed).

Figure 7: Capillary rising picture of organic frozen tomato (processed).

find a neural network model for determining the class from which an unknown image is taken. If we know the probability density functions $f_k(\vec{X})$ for all classes, the Bayes optimal decision rule [12] can be used to classify $\vec{X}$ into class k if

$$h_k \vartheta_k f_k(\vec{X}) > h_m \vartheta_m f_m(\vec{X}) \tag{1}$$

where $k \neq m$.

A major problem with Equation (1) is that the probability density functions of the classes are unknown. We can use Gram-Charlier series expansion to estimate these unknown probability density functions. In this case, the training set for the neural network consists of Gram-Charlier coefficients. Our objective in this study was to use these coefficients for classification. If the neural network is trained with known Gram-Charlier coefficients, then to determine the class of an unknown image, we need only to feed its Gram-Charlier coefficients. The next section describes the Gram-Charlier series expansion.

Gram-charlier Series

The Gram-Charlier series expansion of the probability density

Figure 5: Capillary rising picture of organic fresh tomato (unprocessed).

function of a random variable with mean μ and variance σ^2 can be represented as

$$\rho(x) = \frac{1}{\sigma} \sum_{i=0}^{\infty} c_i \varphi^{(i)}(\frac{x-\mu}{\sigma}), \qquad (2)$$

where $\varphi(x)$ is a Gaussian probability density function and $\varphi^{(i)}(x)$ represents the i-th derivative of $\varphi(x)$. For normalized data where ($\mu = 0$, $\sigma^2 = 1$, $c_{0=1}$), the above equation can be simplified to

$$\rho(x) = (\varphi(x) + c_3 \varphi^{(3)}(x) + c_4 \varphi^{(4)}(x) + c_5 \varphi^{(5)}(x) + c_6 \varphi^{(6)}(x) + \cdots), \qquad (3)$$

where c_i coefficients are related to the central moments of $\varphi(x)$. In a sense, derivatives of the Gaussian function in Equation (3) provide us with the class type information for the tomato. Furthermore, $c_i \varphi^{(i)}$ are orthogonal functions that present unique information about the organic and conventional tomato class distribution. This leads us to conclude that an INN based on the decomposition of the probability density function by the Gram-Charlier series is well suited for organic/conventional tomato class discrimination. Let's define $\beta(x)$ as

$$\beta(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \qquad (4)$$
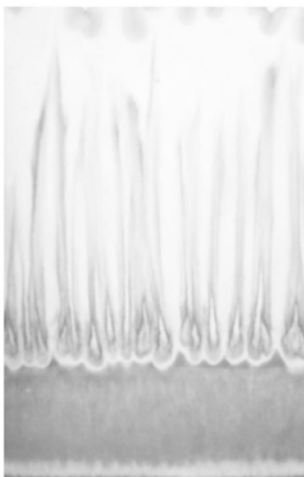

Figure 8: Cropped grayscale image of fresh conventional tomato (1400 pixel x 900 pixel).
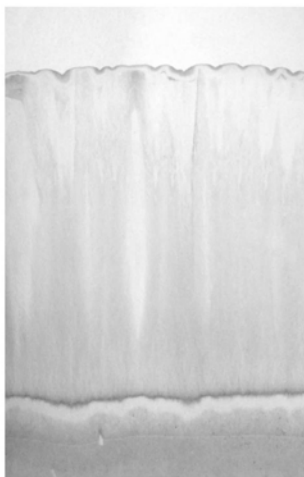

Figure 9: Cropped grayscale image of fresh organic tomato (1400 pixel x 900 pixel).
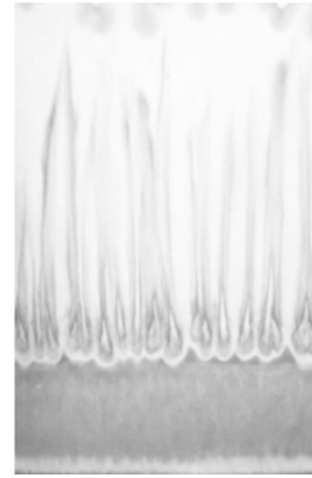

Figure 10: Cropped grayscale image of frozen conventional tomato (1400 pixel x 900 pixel).

Then, the probability density function becomes

$$\rho(x) = \beta(x)\left\{1 + \frac{1}{6}\Lambda_3 A_3 + \frac{1}{24}(\Lambda_4 - 3)A_4 + \cdots\right\} \qquad (5)$$

where

$$A_i(x) = x^i - \frac{i^{[2]}}{2.1!}x^{i-2} + \frac{i^{[4]}}{2^2.2!}x^{i-4} - \frac{i^{[6]}}{2^3.3!}x^{i-6} + \cdots \qquad (6)$$

$$i^{[k]} = \binom{i}{k} = \frac{i!}{(i-k)!}$$

$$m_k = \frac{1}{n}\sum_{i=1}^{n}\left[x(i)\right]^k \qquad (7)$$

$$\Lambda_3 = -\frac{m_3 - 3m_2 m_1 + 2m_1^3}{3!},$$

$$\Lambda_4 = \frac{m_4 - 4m_3 m_1 + 6m_2 m_1^2 - 3m_1^4}{4!},$$

$$\Lambda_5 = -\frac{m_5 - 5m_4 m_1 + 10m_3 m_1^3 - 10m_2 m_1^3 + 4m_1^5}{5!},$$

$$\Lambda_6 = \frac{m_6 - 6m_5 m_1 + 15m_4 m_1^2 - 20m_3 m_1^3 + 15m_2 m_1^4 - 5m_1^6}{6!}$$

Equation (7) is the so-called Gram-Charlier series [12] and the polynomial $Ai(x)$ is called the Tchebycheff-Hermite polynomial.

Gram-charlier neural network architecture

This novel neural network has three processing stages (see Figure 12). The first stage preprocesses capillary rising organic and conventional fresh and frozen tomato-images and the second stage converts these grayscale images into row vectors and estimates Gram-Charlier coefficients [13] corresponding to these row vectors (see Figure 13). In the third processing block, these parameters are applied to a back propagation neural network to classify the images (see Figure 14). This neural network is a fully connected feed forward neural network. When the Gram-Charlier coefficients are presented as an input to the neural network, each output of the hidden neurons is a weighted sum of the input nodes and the bias node passes through a
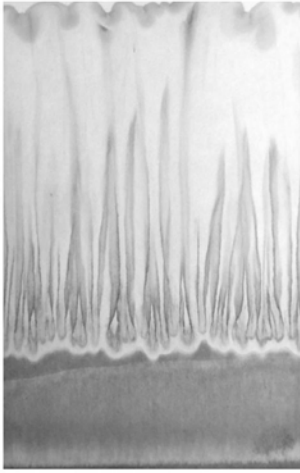
Figure 11: Cropped grayscale image of frozen organic tomato (1400 pixel x 900 pixel).
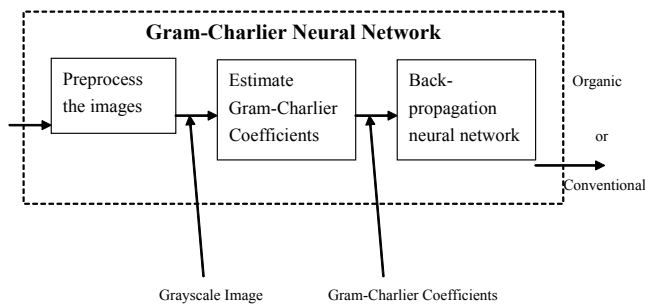

Figure 12: Gram-Charlier Neural Network (GCNN) Architecture for fresh or frozen case.

Figure 13: Cropped grayscale 1400 x 900 pixel image converted to a 1 x 1,260,000 pixel row vector (X) (where m is 1,260,000). This row vector is normalized before its Gram-Charlier coefficients are calculated ($x(i)$, $i = 1...m$ is normalized (mean value is removed and standard deviation is made 1).

Figure 14: Gram-Charlier neural network (If Output ≥ 0, then input type of capillary rising fresh or frozen sample image belongs to organic class)

hyperbolic tangent function. The hidden neuron's output of this neural network, y_j, $j = 1 \cdots L$, is given as

$$y_j = \varphi(\sum_{i=1}^{M} x(i) w_{ji}^{h} + \theta^{j}) \qquad (8)$$

where L is 10, M is 4, and the activation function for the hidden and output layer, $\varphi()$, is a tangent hyperbolic function defined as

$$\varphi(y) = \frac{e^{y} - e^{-y}}{e^{y} + e^{-y}} \qquad (9)$$

The term $x(i)$ is the input vector that consists of the Gram-Charlier coefficients, and w_{ji}^{h} is the weight from the jth hidden neuron to the ith input neuron. Then, the output of the neural network is the weighted sum of all the hidden neurons and the bias node. This actual output of the neural network is then compared to the desired output for every set of input values. Furthermore, the neural network error is defined as the difference in these outputs. Then, the weights of the neural network are updated using the gradient of this output error [14,15]. The training matrix was prepared using the Gram-Charlier coefficients. Each column in the training matrix represented one set of these Gram-Charlier coefficients with a length of 4. Note that the amount of columns in the training set could be any number. The bigger the number, the better the discrimination achieved by the GCNN for the target capillary rising organic and conventional fresh and frozen tomato-images.
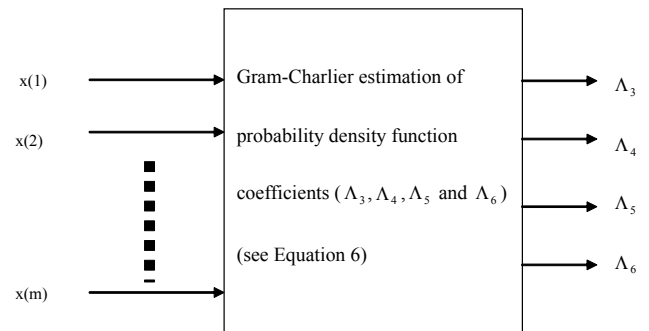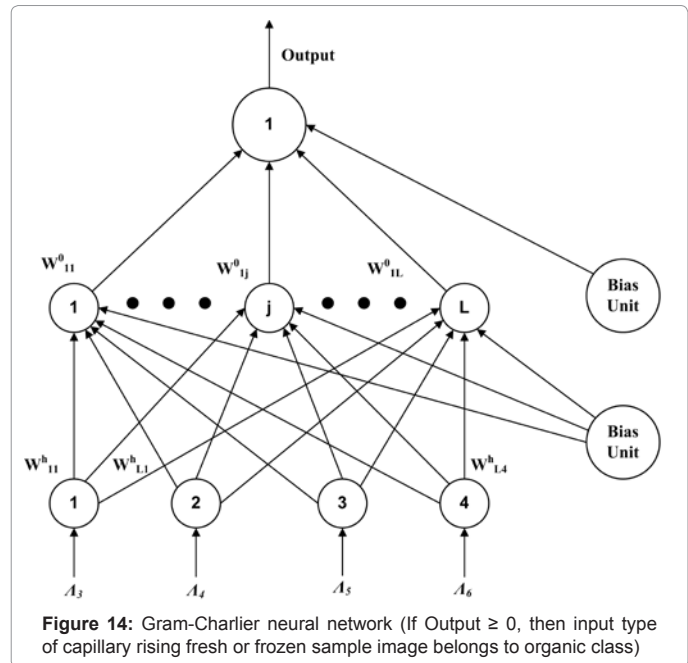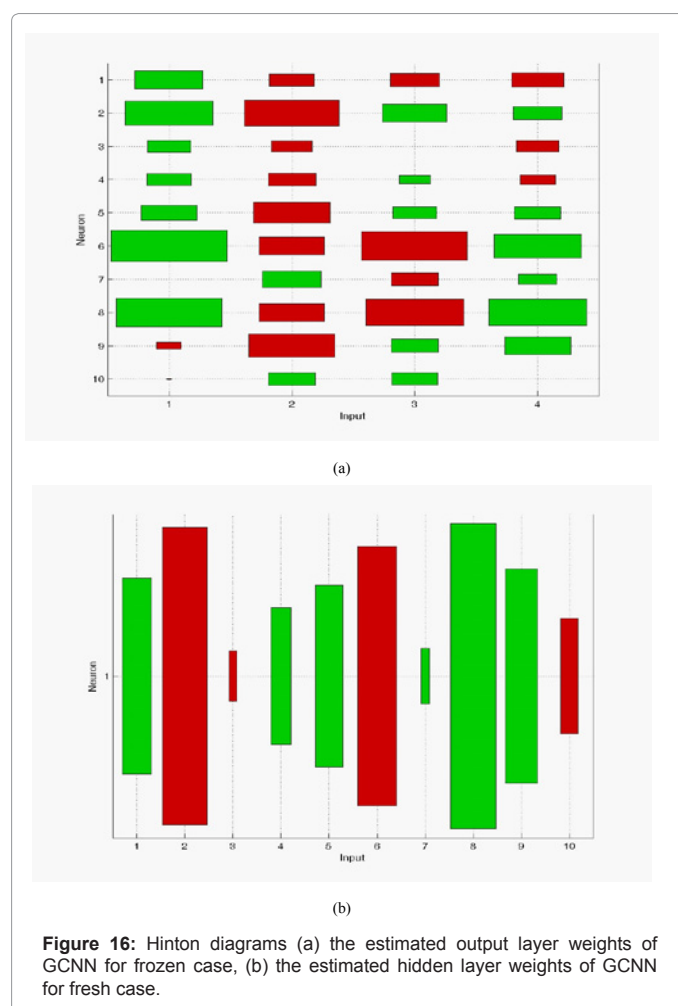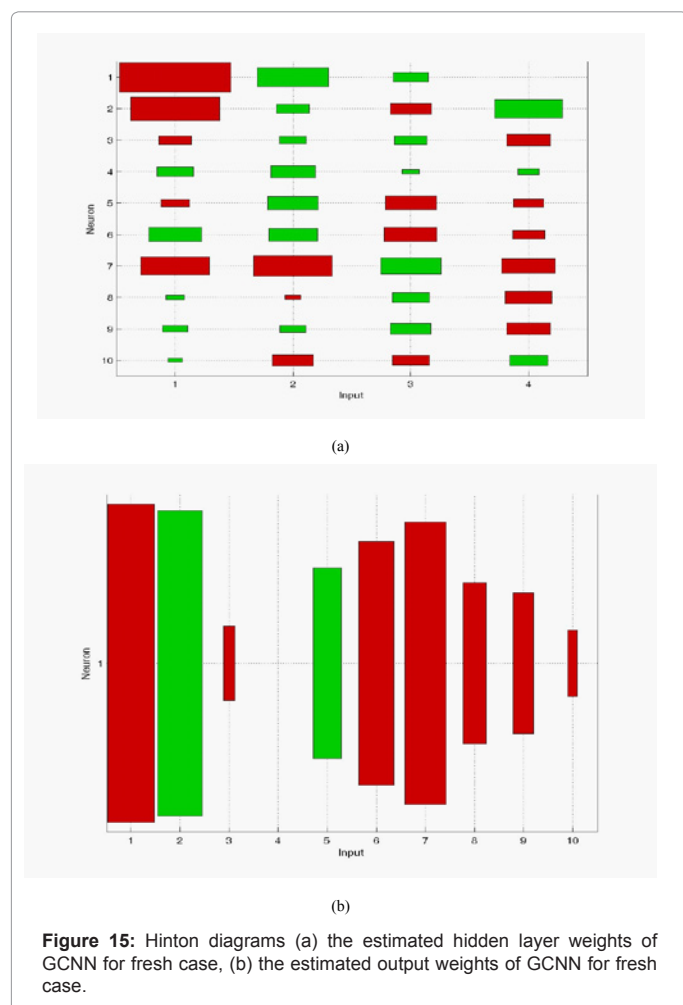
Results and Discussion

The results obtained from testing the GCNN indicate that capillary rising picture method used for discrimination of organic and conventional fresh and frozen tomato images can be fully distinguished with 100% accuracy. Out of the 23 images, 19 were used for training and 4 were used for testing. The number of hidden neurons for the back propagation neural network was chosen to be 10, based on experience. In testing, if the output of the GCNN ≥ 0, then the input Gram-Charlier coefficients belonged to class H_o = Organic. If the output of the GCNN < 0, then the input Gram-Charlier coefficients belonged to class H_1 = Conventional (see Figure 14). The success of the GCNN may be obscured if there is no understanding of the optimality of the neural network. In this study, Hinton diagrams were used to explain the optimal characteristics of the neural network. Hinton diagrams are named for Geoffrey Hinton, who used this method to display neural network weights. In the Hinton diagrams, the size of each square is proportional to the strength of each weight. The green color indicates a positive magnitude weight, and the red color indicates a negative magnitude weight. Figures 15a-16a depicts the Hinton diagram of

the two GCNNs. The first row in each figure shows the weights from the input nodes to the first hidden neuron, the second row depicts the weights from the input nodes to the second hidden neuron, and so on. Note that the number of rows in Hinton diagrams is equal to the number of hidden neurons. The first column shows the weights between each hidden neuron and the first input neuron. Figures 15b-16b depicts the weights from the hidden neurons to the output nodes. Since there is one output node, the number of rows in this Hinton diagram is 1. Again, the first column depicts the weights between the output nodes and the first hidden neuron. Inspection of these Hinton diagrams reveals that all of the hidden nodes contributed to the decision process and that the optimal design of the neural network was realized. Furthermore, there is no situation suggesting that a set of weights yields very large values, very low values, or predictable values.

In summary, this novel neural network has the following properties: 1) the response is real-time once it is trained, 2) all hidden neurons have the same structure, which makes its hardware implementation easy using VLSI design techniques [16,17].

Although similar neural network designs are used in medical ultrasonic imaging for tissue characterization and diagnosis (ref…), based on experimental observations obtained throughout this study, the GCNN can also be efficiently used to discriminate the capillary rising picture method of conventionally and organically grown tomatoes. Furthermore, the Gram-Charlier neural network renders practical advantages such as real-time processing, adaptability, and training capability.

(a)

(b)

Figure 15: Hinton diagrams (a) the estimated hidden layer weights of GCNN for fresh case, (b) the estimated output weights of GCNN for fresh case.

(a)

(b)

Figure 16: Hinton diagrams (a) the estimated output layer weights of GCNN for frozen case, (b) the estimated hidden layer weights of GCNN for fresh case.

Conclusion

In this study we developed a neural network methodology that is designed to classify the capillary rising picture method of conventional and organic tomatoes for fresh and frozen cases. The images taken from a lab-bear information related to the class type of tomato. However, this information is not readily quantifiable and lacks uniquely recognizable features. Therefore, a neural network becomes appealing for classifying these images, because it is trainable. The optimal values for the neural network weights were estimated using the back propagation algorithm. Experimental measurements of the tomato were utilized to train and test the image neural network. This network showed a remarkable 100% classification performance. This success of GCNN does not depend on whether tomato is fresh or frozen. Parallel classification performance was also achieved when training the neural network.

Overall, the results of present study are encouraging and suggesting that neural networks are potentially useful for detecting conventional/ organic tomatoes and can be used in food quality control.

Acknowledgement

References

1. Woese K, Lange D, Boess C, Bögl KW (1997) A comparison of organically and

conventionally grown foods: Results of a review of the relevant literature. J Food and Agri 74: 281-293.

2. Zalecka A, Skjerbaek K, Doesburg P, Pyskow B, Huber M, et al. (2006) The capillary dynamolysis method as a characterized tool for crop quality determination. Paper presented at Joint Organic Congress, Odense, Denmark, May 30-31.

3. Doesburg P (2007) Holistic (picture forming) methods for measuring quality. Seminar (Organised by Lucy P.L. van de Vijver, PhD); measuring food quality: concepts, methods and challenges, February, Louis Bolk Institute, Netherlands.

4. Kolisko L (1929) Das Silber und der Mond - Experimentelle Studien aus dem biologishen Institut am Goetheanum - Schriftenreihe der Natura III. Orient-Occident Verlag, Stuttgart.

5. Meelursarn A (2006) Statistical Evaluation of Texture Analysis from the Biocrystallization Method: Effect of Image Parameters to Differentiate Samples from Different Farming Systems, Ph.D. thesis, University of Kassel, Department of Organic Food Quality and Food Culture, Witzenhausen 220.

6. Skjerbaek K, Pyskow B, Zalecka A, Kahl J, Huber M, et al. (2004) Development and Characterization of the Capillary Dynamolysis Method for Food Quality Analysis. International Research Project, Triangle Report Nr. 2, Retrieved May 14, 2009.

7. Kuşçu A (2008) Organik ve Konvansiyonel Kırmızıbiber ve Ürünlerinin Ayırt Edilebilme Yöntemleri ve Kalite Özelliklerinin İncelenmesi, Ph.D. Thesis, Ege Üniversitesi Fen Bilimleri Enstitüsü, Izmir, 420.

8. Andersen JO, Skjerbaek K, Paulsen M, Pyskow B (2004) Kvaliteten Af Udvalgte Økologisk Dyrkede Vårhvedesorter, Belyst Ved Billeddannende Metoder, Rapport nr. 2, Biodynamisk Forskningsforening 38.

9. Mäder P, Pfiffner L, Niggli U, Balzer U, Balzer F, et al. (1993) Effect of three farming systems (bio-dynamic, bio-organic, conventional) on yield and quality of beetroot (beta vulgaris l. Var. Esculenta l.) in a seven year crop rotation, ISHS Acta Horticulturae 339: 11-31.

10. Andersen JO, Henriksen CB, LaursenJ, Nielsen AA (1999) Computerised Image Analysis of Biocrystallograms Originating from Agricultural Products, ComP and Electra in Agri 22: 51-69.

11. Bloksma J, Northolt M, Huber M (2001) Parameters for Apple Quality Part-1, Louis Bolk Institute, 115, Retrieved May 14, 2009.

12. Masters T (1993) Practical Neural Networks, Academic Press, Inc.

13. Kendall MG, Stuart A (1963) The Advanced Theory of Statistics, Charles Griffin & Company Limited.

14. Cichocki A, Unbehauen R (1992) Neural Networks for Optimization and Signal Processing, John Wiley and Sons.

15. Freeman JA, Skapura DM (1991) Neural Networks Algorithms, Applications and Programming Techniques, Addison-Wesley Publishing Company.

16. Hollis PW, Paulos JJ (1990) Artificial Neural Networks Using MOS Analog Multipliers. IEEE J Solid State Circuits 25: 849-855.

17. Hollis PW, Paulos JJ (1994) A Neural Network Learning Algorithm Tailored for VLSI Implementation. IEEE Transactions On Neural Netw 5: 784-791.

18. Botros N, Tee TH (1991) Ultrasound Tissue Characterization Using Artificial Neural Network, Medical Ultrasound. 13: 165-166.

19. Lu HHS, Chen CM, Yang IH (1998) Cross-reference weighted least square estimates for positron emission tomography, IEEE Trans Med Imaging 17: 1-8.

20. Shrimali V, Anand RS, Kumar V (2009) Current Trends in Segmentation of Medical Ultrasound B-mode Images: A Review, IETE Technical Review 26: 8-17.