Modeling of Fried Potato Chips Color Classification using Image Analysis and Artificial Neural Network

T. Marique, A. Kharoubi, P. Bauffe, and C. Ducattillon

ABSTRACT: Quality of potatoes in chips industry is estimated from the intensity of darkening during frying. This is measured by a human jury, subject to numerous factors of variation. Gray level intensities were obtained for the apex, the center, and the basal parts of each chip using image analysis of frying assays. We then tested a feed-forward artificial neural network designed to associate these data with color categories. It behaved in good agreement with human estimations, obtaining correlation coefficients of 0.972 for training data and of 0.899 for validation data. A systematic study of the response of the model allowed understanding the criteria of evaluation used by the human operators.

Keywords: potato, chips, color, artificial neural network, image analysis

Introduction

In the potato chips industry, each batch of tubers must be tested for quality before processing, and visual aspect is, of course, of great importance. During frying, chips develop a dark coloration because of Maillard reaction (Cheftel and Cheftel 1980). The main protagonist molecules are known to be reducing sugars and proteins (Füller and Hughes 1984; Brown and others 1990; Richardson and others 1990; Gravoueille 1993). However, other components in potatoes, such as ascorbic acid and orthophenols, and external factors, such as respiration and storage temperature, may also play a role (Rodriguez-Saona and others 1997; Copp and others 2000; Blenkinsop and others 2002). To estimate this darkening, one performs a simple frying assay on 20 chips issued from the central part of 20 different potatoes. Each stick is then assigned a category by visual examination under standard white light. The jury builds its evaluation, with the help of a standard reference card, from both overall darkening of the stick and from contrast between the extremities (apex and base) and the center of the stick.

Of course, some problems are associated with this procedure. In particular, estimations may vary with the jury. Even for a given jury, sample variability can influence results because narrow distributions tend to be spread over the scale. We therefore tried to develop a model for reproducible estimation of chips color category.

Artificial Neural Networks (ANNs) attain very good performance when used to predict values for complex nonlinear systems (Wilkinson and Yuksel 1997; Mittal and Zhang 2000). They have been shown to perform better than classical regression models in numerous cases of all kinds, in particular for pattern recognition and classification (Briandet and others 1996; Amendolia and others 1998; Chtioui and others 1999; Fernandez-Caceres and others 2001; Polásková and others 2002; Tomida and others 2002). Moreover, they are endowed with broad capacities of generalization so that they can give useful information for cases that were not part of their training set (Schalkoff 1997; Wilkinson and Yuksel 1997; Marique and Wérenne 2001; Yang and others 2002). ANNs could help solve our specific problem; therefore we tested whether an ANN can provide successful prediction of darkening index for fried potatoes. We used image analysis to extract gray-level intensities from our image data bank gathered from routine frying assays. These were fed to the ANN, and the output values were compared with the corresponding color categories estimated by human evaluators.

MS 20030058 Submitted 1/31/03, Revised 3/5/03, Accepted 6/9/03. *Authors are with the Centre Agronomique de Recherches Appliquées du Hainaut (CARAH), 11 rue P. Pastur, 7800 Ath, Belgium. Direct inquiries to author Marique (E-mail: tmarique@ tvcablenet.be).*

Materials and Methods

Sample origin, preparation, and cooking

We used 12 different mealy potato cultivars from our assay fields: *Annabelle, Bintje, Cantate, Charmante, Cyclone, Daisy, Farmer, Innovator, Lady Olympia, Liseta, Markies, Victoria*. For each potato batch, 20 freshly harvested, calibrated tubers with a minimum diameter of 50 mm were selected at random. A potato kitchen cutter was used to divide the tuber into square sticks with edges 1 cm long; the central chips were isolated and gathered together. The epidermis at both ends was eliminated with a knife; the chips were rinsed in cold tap water and dried with blotting paper.

Frying was performed immediately for 3 min at 180 °C. We used a 17 kW deep fryer Frymaster H17BLSC (Enodis, Shreveport, La., U.S.A.) with 25 L peanut oil. Fried chips were then collected, examined under standard white light (daylight fluo tube 6500 °K, 20 W), and noted by comparison with the standard color card (Belgapom, Brussels, Belgium). All readings were performed by the same person, who assigned every chip to a color category, ranging from 0 (very pale) to 4 (very dark). The jury was instructed to proceed for each chip following a 2-step procedure. First, the chip was assigned to a given category on the basis of its apparent global gray level only. Next, if the chip possessed clearly contrasted dark ends (that is, more then 1 cm long), it was moved into the color category immediately superior to category it was assigned originally. These ratings, which are based on chips pictures from the reference card, are really much more complex due to expertise added by the jury. Training of the models will thus have to be performed using data from the jury rather than from the reference card.

Image analysis

A digital camera (Kodak DC3200) was used to take an image of every assay, in jpeg format. Each image included a reference scale for gray levels.

Images were analyzed using the Image-Pro Plus 4.5 software from MediaCybernetics (Carlsbad, Calif., U.S.A.). They were converted from color to 8-bit gray images, and contrast was optimized so that the reference scale stretched from 0 (black) to 255 (white). There were 20 chips on every image, which were then scanned and mean gray levels on about half their width were measured, all along their length.

Raw data processing and neural network modeling or multiple linear regression (MLR)

The values issued from image analysis were transferred, together with the corresponding color categories assigned by the jury, to Microsoft Excel software (Redmond, Wash., U.S.A.). Three mean values from 26 pixels each were computed per chip, corresponding to the apex, center, and base of the sticks, respectively. This data was then divided at random between 2 sets; the larger set (2/3 of the data) was used for training. The remaining set was used for validation of either a MLR or the ANN.

MLR was performed using the MATLAB 6.1 software (The MathWorks, Inc., Natick, Mass., U.S.A.). We built a 2-layer feed-forward ANN (Figure 1) using the MATLAB 6.1 software.

Three inputs correspond to the mean gray values from the 3 regions of a given chip. The output layer consists of a single linear neuron with bias that gives the estimated value of the color category (from 0 to 4). The hidden layer consists of neurons, with sigmoid transfer functions, and bias. The number of hidden neurons used for ANN construction affects the network performance; too few lead to weak prediction performance and too many result in overtraining and bad generalization capacities. In this case, best correlations for validation data were obtained using 4 hidden neurons.

The ANN was trained using different algorithms. Gradient-descent showed slow learning and poor correlations. We finally retained Lenvenberg-Marquardt algorithm ("trainlm" function with standard built-in parameter values), because it gave fast convergence, as is usually the case for small networks (Schalkoff 1997).

Results and Discussion

Image analysis

Figure 2 shows the way the jury assigns a particular color value to any chip. The different color categories are distributed throughout the gray values scale, succeeding one another with partial overlapping. This comes from the fact that when a particular chip stands exactly between 2 categories, the jury will select a category randomly, and so either undervalue or overvalue it, hence the overlapping.

Chips are assigned to category 0 if they appear both very pale (gray levels over 150) and rather paler at the extremities than in the center (population laying at the right side of the diagonal, Figure 2). A chip will be assigned to category 1 either if it appears clearer in the center but possesses contrasted dark ends or if it appears clearly darker in the center, with pale ends. Category 1 is thus clearly split into 2 subpopulations (ellipsoids in Figure 2), flanking both sides of category 0. Categories 2, 3, and 4 then progressively regroup darker chips, generally clearly pale in the center with more or less contrasted dark ends. Thus, it is only for the color categories 0 and 1 that the jury will overvalue a stick possessing dark contrasted extremities. For color categories 2, 3, and 4, estimation is based mostly on the chip global (center) appearance, dark contrasted ends being considered as "normal."

Compared with this complexity, the reference card appears to be a very crude instrument. It only presents a few samples of chips (2 per color category) that are supposed to provide the jury with a broad guideline for general darkening level appreciation. In particular, all the chips pictures are of a homogenous color, whereas "real" experimental samples show all kinds of visual aspects, for example, brown apex, brown base and apex, and so on. As the reference card does not account for such variability, we obvious-

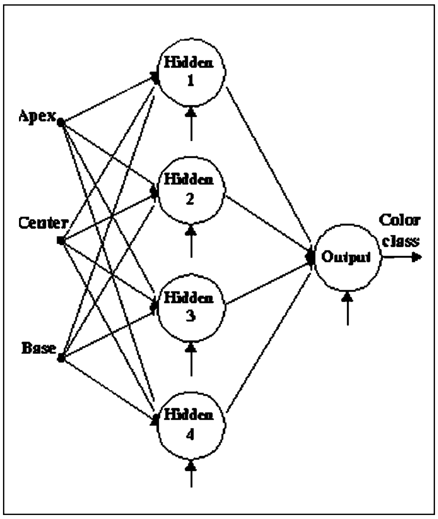

Figure 1—Structure of a 2-layer feed-forward ANN with 3 inputs, 4 hidden neurons with bias, and 1 output neuron with bias.

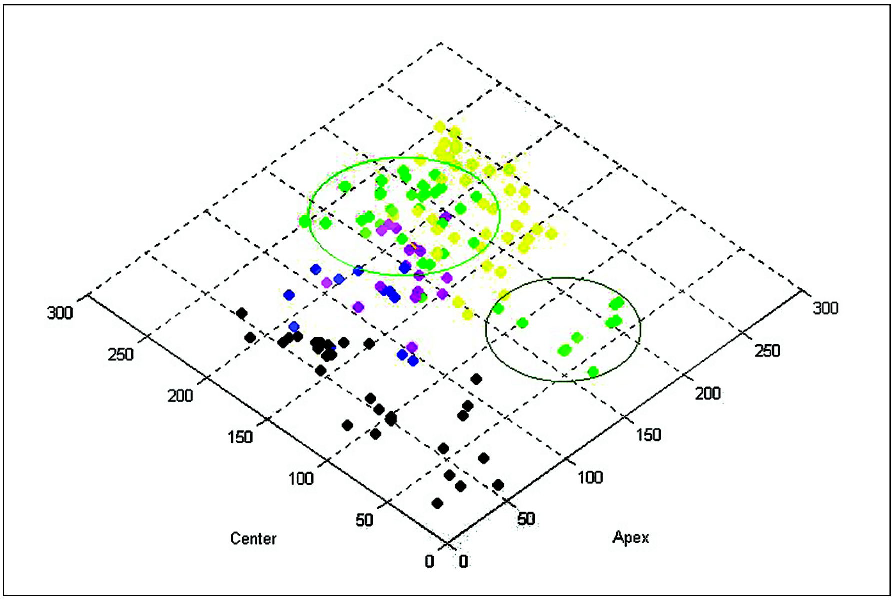

Figure 2—Mean gray values for the center and the apex of the chips. The color code indicates the class of color. Yellow, 0; green, 1; magenta, 2; blue, 3; black, 4. The ellipsoids contain the 2 subpopulations of class 1: in dark green, globally darker sticks with pale ends; in light green, clearer sticks with contrasted dark ends.

ly cannot use it to train our models. This is why we used the ratings from the human jury instead. We thus lost the advantage of obtaining objective data values, but gained the irreplaceable expertise that only a jury can bring.

Training and validation of MLR or ANN

The data obtained from image analysis were used to train and validate a MLR. We obtained good performance for training data, with correlation coefficients of 0.939. For validation data, we obtained a less satisfactory correlation coefficient of 0.878.

The same data were used to train and validate an ANN. Best correlations were obtained with 4 neurons in the hidden layer. The ANN shows better performance than MLR, with a correlation coefficient of 0.972 for training data and a less satisfactory correlation coefficient of 0.899 for validation data, possibly as a consequence of inaccuracies in the target values decided by the human jury.

We used the trained ANN to generate a complete set of predictions, for the different possible combinations of gray levels of the center and of the apex of the chips. This is illustrated in Figure 3, where computing was performed using equal gray values for both the base and the apex of the chips.

We observe that the network displays complex and continuous behavior for the low color categories 0, 1, and 3 but operates a discrete classification between categories 3 and 4 only. This could be a consequence of both the greatest number of data points for high color categories and of the more complex behavior of the jury for low color categories.

A more complete simulation is shown in Figure 4, illustrating the discrete color categories (the values predicted from the ANN have been rounded to the nearest integer) obtained for all the possible combinations of gray levels of the 3 regions of the chips.

We observed again a more complex behavior for low values color categories. Intermediate categories 2 and 3 also can be seen to extend between 2 broad, different regions, 1 globally paler with dark ends and another globally darker with pale ends. Color classification varies most with apex and center gray values.

In this study, we worked only on gray level data values obtained from original color images. We still need to investigate whether we could obtain better performances by using color data, at the expense of a greater network complexity. Indeed, as different potato varieties possess more or less yellow flesh, the jury could take this factor into account.

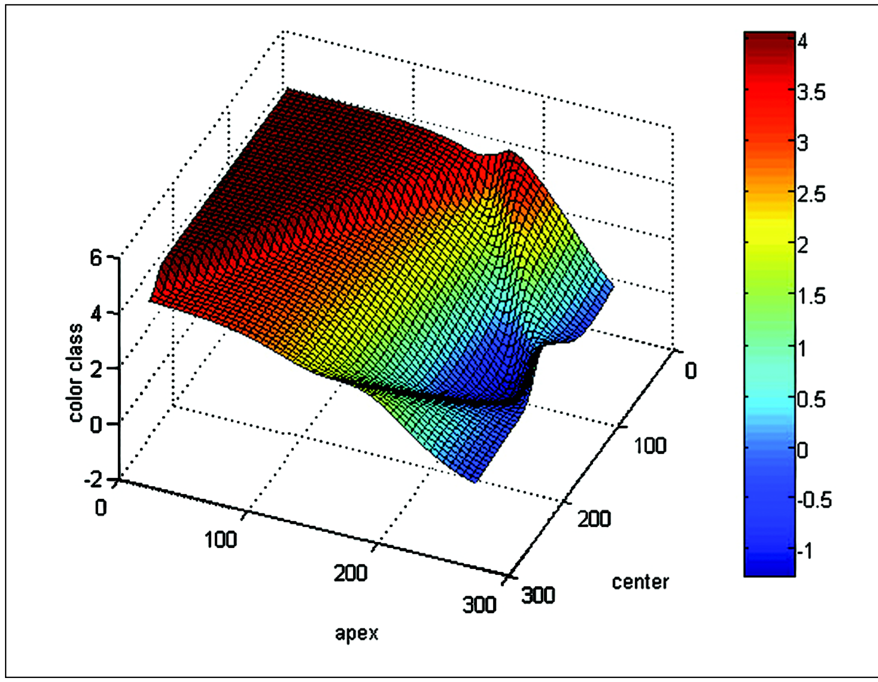

Figure 3—Response of the ANN: color class categories as a function of all possible combinations of gray levels of the center and the apex of the chips.

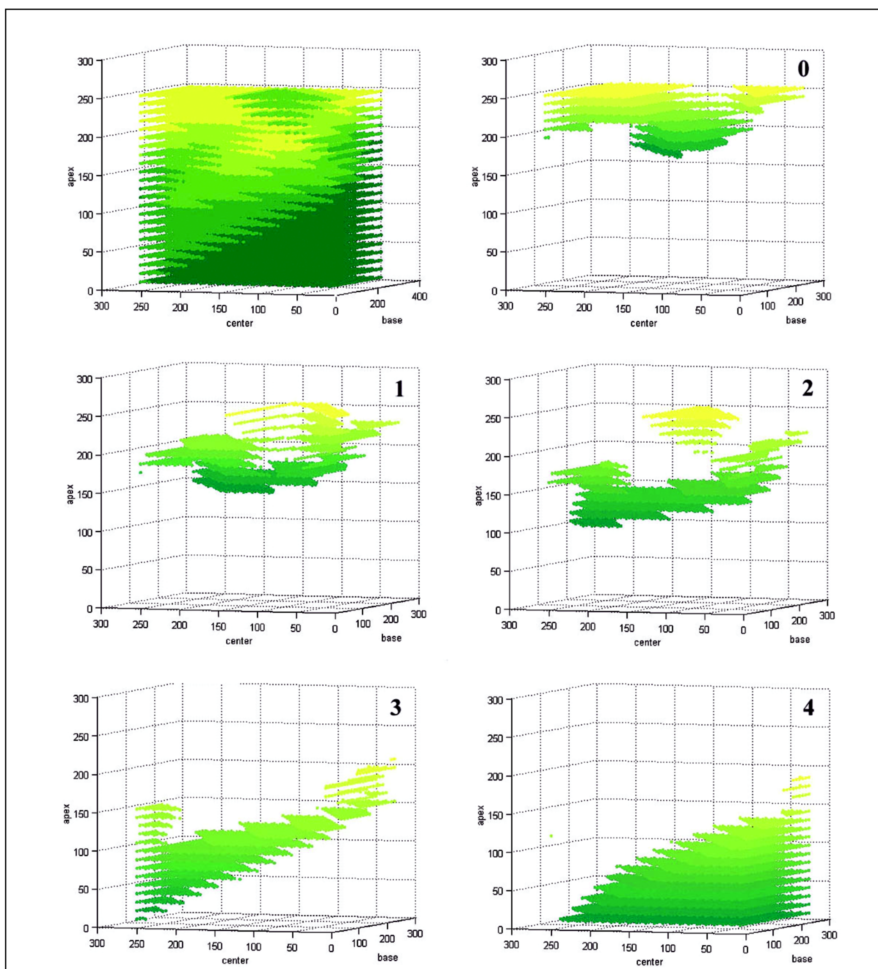

Figure 4—Discrete color categories obtained for all the possible combinations of gray levels for the 3 regions of the chips. Upper left: general presentation. Others: partial views of each color category, from 0 (upper right) to 4 (bottom right).

Conclusions

THE ANN DEVELOPED HERE SHOWED GOOD performance, learning from a relatively small number of data values. The ANN model behaved better than multiple linear regression analysis. Predicted categories appear to reproduce the pattern of the experimental data issued from the jury, revealing nonlinear mapping, existence of subregions and partial overlapping of categories. Moreover, the generalization capacities of the network allowed to simulate plausible predictions for the whole set of parameter combinations. The present work is to be considered as a 1st step toward a practical ANN model that will be used for objective, precise, and accurate online prediction of chips quality.

References

Amendolia SR, Doppiu A, Ganadu MLG. 1998. Classification and quantitation of 1H NMR spectra of alditols binary mixtures using artificial neural networks. Analyt Chem 70:1249–54.

Blenkinsop RW, Copp LJ, Yada RY, Marangoni AG. 2002. Changes in compositional parameters of tubers of potato (*Solanum tuberosum*) during low-temperature storage and their relationship to chip processing quality. J Agric Food Chem 50:4545–53.

Briandet R, Kemsley EK, Wilson RH. 1996. Approaches to adulteration detection in instant coffees using infrared spectroscopy and chemometrics. J Sci Food Agric 71:359–66.

Brown J, MacKay GR, Bain H, Griffith DW, Allison MJ. 1990. The processing potential of tubers of the cultivated potato, *Solanum tuberosum* L., after storage at low temperatures. 2. Sugar concentration. Potato Res 33:219–27.

Cheftel JC, Cheftel H. 1980. Introduction à la biochimie et à la technologie des aliments. Vol. 1. Paris: Technique et documentation. 381 p.

Chtioui Y, Panigrahi S, Francl L. 1999. A generalized regression neural network and its application for leaf wetness prediction to forecast plant disease. Chemom Intell Lab Syst 48:47–58.

Copp LJ, Blenkinsop RW, Yada RY, Marangoni AG. 2000. The relationship between respiration and chip color during long-term storage of potato tubers. Am J Potato Res 77:279–87.

Fernandez-Caceres PL, Martin MJ, Pablos F, Gonzalez AG. 2001. Differentiation of tea (*Camellia sinensis*) varieties and their geographical origin according to their metal content. J Agric Food Chem 49:4775–9.

Füller TJ, Hughes JC. 1984. Factors influencing the relationships between reducing sugars and fry color of potato tubers of c.v. record. J Food Technol 19:455–67.

Gravoueille JM. 1993. Les sucres de la pomme de terre. Pomme Terre fr. 477:133–40.

Marique T, Wérenne J. 2001. A general artificial neural network for the modelization of culture kinetics of different CHO strains. Cytotechnology 36:55–60.

Mittal GS, Zhang J. 2000. Use of artificial neural network to predict temperature, moisture, and fat in slab-shaped foods with edible coatings during deep-fat frying. J Food Sci 65:978–83.

Polasková P, Bocaz G, Li H, Havel J. 2002. Evaluation of calibration data in capillary electrophoresis using artificial neural networks to increase precision of analysis. J Chromatogr A 979:59–67.

Richardson DL, Davis HV, Ross HA. 1990. Potato tuber sugar content during development and storage (10 °C): possible predictors of storage potential and the role of sucrose in storage hexose accumulation. Potato Res 33:241–5.

Rodriguez-Saona LE, Wrolstad RE, Pereira C. 1997. Modeling the contribution of sugars, ascorbic acid, chlorogenic acid and amino acids to non-enzymatic browning of potato chips. J Food Sci 62:1001–10.

Schalkoff RJ. 1997. Artificial neural networks. New York: McGraw-Hill. 448 p.

Tomida S, Hanai T, Koma N, Suzuki Y, Kobayashi T, Honda H. 2002. Artificial neural network predictive model for allergic disease using single nucleotide polymorphisms data. J Biosci Bioeng 93:470–8.

Wilkinson C, Yuksel D. 1997. Using artificial neural networks to develop prediction models for sensory-instrumental relationships: an overview. Food Qual Prefer 8:439–45.

Yang C-C, Prasher SO, Landry J, Ramaswamy HS. 2002. Development of neural networks for weed recognition in corn fields. Trans ASAE 45:859–64.