Computers and Mathematics with Applications 63 (2012) 268–279

Contents lists available at SciVerse ScienceDirect

Computers and Mathematics with Applications

journal homepage: www.elsevier.com/locate/camwa

A real-time mathematical computer method for potato inspection using machine vision

Navid Razmjooy [a], B. Somayeh Mousavi [b,*], F. Soleymani [c]

[a] *Young Researchers Club, Majlesi Branch, Islamic Azad University, Isfahan, Iran*
[b] *Young Researchers Club, Zahedan Branch, Islamic Azad University, Zahedan, Iran*
[c] *Department of Mathematics, Zahedan Branch, Islamic Azad University, Zahedan, Iran*

ARTICLE INFO

Article history:
Received 14 September 2011
Received in revised form 11 November 2011
Accepted 11 November 2011

Keywords:
Potato
Otsu thresholding
Mathematical morphology
Defect detection
K nearest neighborhood
Support vector machines

ABSTRACT

Detection of external defects on potatoes is the most important technology in the realization of automatic potato sorting stations. This paper presents a hierarchical grading method applied to the potatoes. In this work a potato defect detection combining with size sorting system using the machine vision will be proposed. This work also will focus on the mathematics methods used in automation with a particular emphasis on the issues associated with designing, implementing and using classification algorithms to solve equations. In the first step, a simple size sorting based on mathematical binarization is described, and the second step is to segment the defects; to do this, color based classifiers are used. All the detection standards for this work are referenced from the United States Agriculture Department, and Canadian Food Industries. Results show that we have a high accuracy in both size sorting and classification. Experimental results show that support vector machines have very high accuracy and speed between classifiers for defect detection.

1. Introduction

Image processing and computer vision are typical important fields of information science and technology. Several areas of mathematics have contributed to essential progress of these fields. Nowadays, machine vision and image processing are improved as a key technology in quality control; however there are still challenges to be solved [1].

Formerly, color has not been vastly used in optical inspection because of its high cost, complicated equations and high processing power required. Albeit, as costs decrease and processing power ceases to be an issue, solution providers are looking at gathering color into machine vision optical inspection systems to obtain a high quality. In food industries, quality classification has been usually done by some experts who are drawn towards specific features of food that shows the product's quality. However, this method has some drawbacks such as non-uniform behavior; in other words, implementing a repetitious work results in making some mistakes.

World potato consumption is headed up, increasing at an annual rate of 4.5%. Therefore, computer vision technology which can detect some external characteristics of the potato has some advantages such as objectivity, low cost, and high accusation.

Generally, a computer vision algorithm for a defect detection system includes two principal stages; the first is that a proper segmentation algorithm should be applied on the input image to separate purpose objects from background and the second consists of a proper defect detection algorithm which is used on the target objects.

* Corresponding author.
E-mail addresses: navid.razmjooy@hotmail.com (N. Razmjooy), bbmoosavi@gmail.com (B.S. Mousavi), fazl_soly_bsb@yahoo.com (F. Soleymani).

0898-1221/$ – see front matter
doi:10.1016/j.camwa.2011.11.019

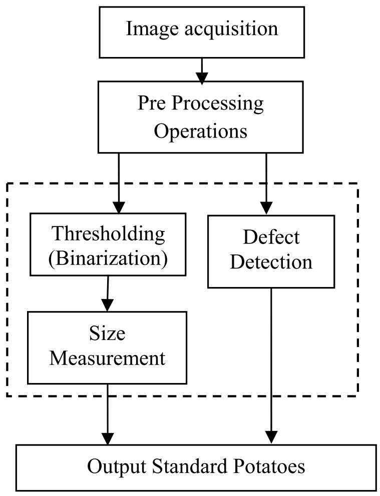

Fig. 1. Diagram of system overview.

Chalana et al. [2] used an HSV color space to segment the green zones which have about 50 images per second classification speed; an oval shape is then used to specify potato form and threshold it to segment green and unripe parts. This algorithm is done in an HSV color space. Recognition accuracy was 86.5%. In 2000 Noordam and Ottenfocused on high speed potato grading and quality inspection, metering the size, shape, and external defects [3].

In 2002 Noordam et al. used a new clustering method called CSI–FCM made as a suitable system by applying the algorithm on the RGB color space for image segmentation and detection; this system had a good performance for multivariate images [4].

In 2006 Franco Pedreschi et al. used an L^*a^*b color space for acquiring the images in a digital format in potato chips. Color values in $L^*a^*b^*$ units were recorded at different sampling times during frying at the four oil temperatures using the total color change parameter; in such condition chips fried at higher temperatures get darker in all 3 dimension of color spaces [5].

Li yang jun and Wang used HSI color space to orange segmentation; the basic idea was derived from computing mean difference and mean square difference into a current colorful image for all color parameters in the whole area of fruit. After that, it normalized them in order of their minimums and maximums [6].

Fang Junlong et al. used simple color segmentation as follows: noise filtering, RGB image dividing, and identifying them to classify the images with black spots [7].

Despite key technology towards the realization of an automatic potato grading and sorting station able to compute both sizes and defects together, it can be inferred from the foregoing works that there is no relevant segmentation system. In this work, we will employ both size and defect classification purposes. Fig. 1 shows the diagram of system overview.

2. Materials and methods

2.1. Image acquisition

A set of 50 bags of potatoes were randomly selected from the Ardabil (Iranian northern-west) champs from the 2011 harvest. These packs contained more varieties of healthy, unhealthy, standard and malformed potatoes.

The computer vision system can be seen in Fig. 2 and is developed to detect defects of potatoes. Their sizes were composed of a CCD camera, a light source, and an Intel(R) core(TM) 2 duo CPU. A Sony cyber-shot DSC-W350 used to acquire images and a total of 500 pictures with 4320×3240 Jpeg resolution were taken from the 50 bags; because of CPU limitation, images were then converted to a 224×168 Jpeg resolution. For illumination two fluorescent lights were used.

After that a color space to commence the work is selected; this paper uses an RGB type of color space. In order to simplify the RGB color space, it has been changed to a popular version through computer vision purposes. Red, green and blue are three primary additive colors and are represented by a three-dimensional, Cartesian coordinate system; Fig. 3 shows this representation for some potato and background.

2.2. Contrast enhancement

The purpose of contrast enhancement in this work is to highlight the defect regions whilst leaving the unimportant background regions intact. This provides the defect detection point to better locate and represent each defect in the image. In our work global contrast enhancement by dynamic stretching (see the table), is applied to achieve this goal.

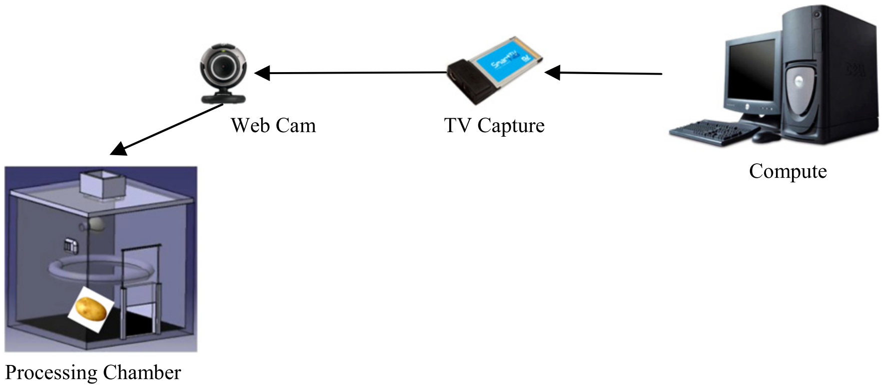

Fig. 2. The computer vision system developed to detect greened potato.

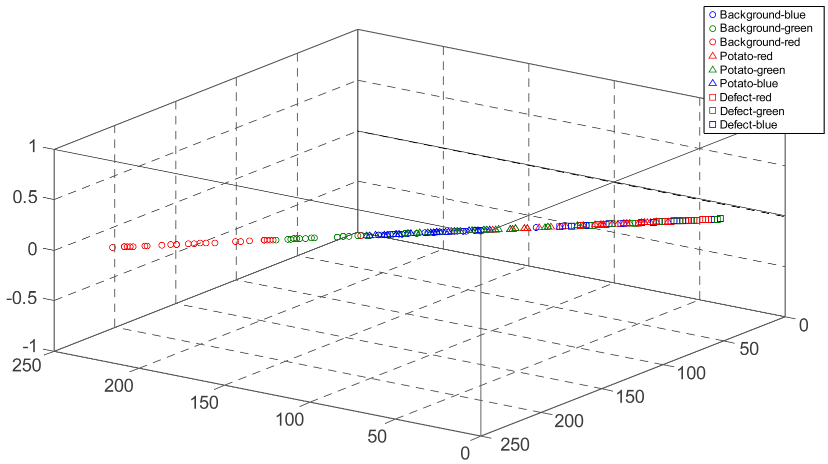

Fig. 3. Analysis of potato colors in the RGB color-space: 3D-plot of potato (blue), 3D-plot of non-potato (red); the potato colors approximately distributed in a linear fashion in the RGB color space. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

8-bits are applied throughout the segment and the resulting imagery is written back to disk. This allows bulk radiometric enhancement of image data.

The LUT transformation is given as follows:

$$\text{Output} = \frac{\text{Input} - \text{Min}}{\text{Max} - \text{Min}}. \tag{1}$$

Min and Max are the minimal and the maximal gray level values in the histogram of the input image respectively [8]. Fig. 4 shows a sample of contrast enhancement:

2.3. Area detection and thresholding

The mean value of the RGB (Intensity) color model is the most useful descriptor and is selected as the graying method of the potato image [8]; therefore the original image in RGB color space is then converted to the intensity dimension as below:

$$I = \frac{1}{3}(R + G + B). \tag{2}$$

After that, an Otsu threshold was applied to enhance the accuracy; using this threshold to discriminate the foreground and background shadow removed scanning images from the shadows. In Otsu's method we exhaustively search for the threshold that minimizes the intra-class variance, defined as a weighted sum of variances of the two classes:

$$\sigma_b^2(t) = \omega_1(t)\sigma_1^2(t) + \omega_2(t)\sigma_2^2(t). \tag{3}$$

Weights ω_i are the probabilities of the two classes separated by a threshold t and σ_i^2 variances of these classes. Otsu shows that minimizing the intra-class variance is the same as maximizing inter-class variance [9]:

$$\sigma_b^2(t) = \sigma^2 - \sigma_\omega^2(t) = \omega_1(t)\omega_2(t)[\mu_1(t) - \mu_2(t)]^2 \tag{4}$$

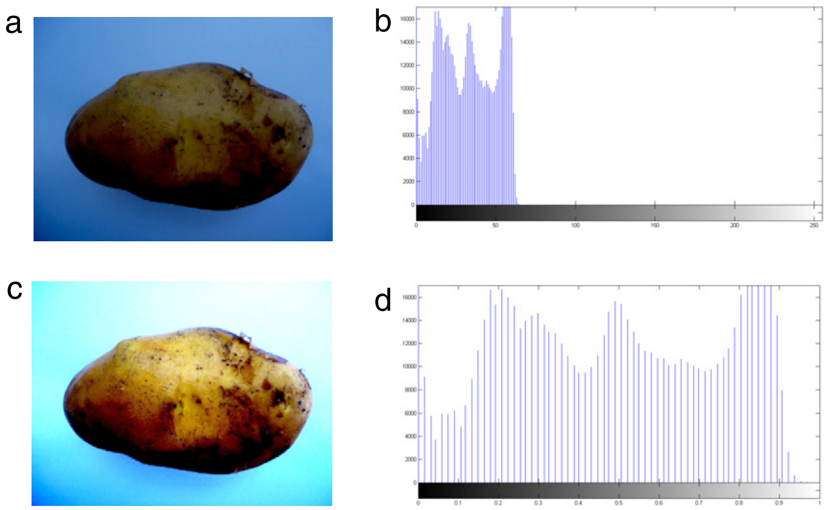

Fig. 4. (a) Input image. (b) Histogram of (a). (c) Image after contrast enhancement. (d) Histogram of (c).

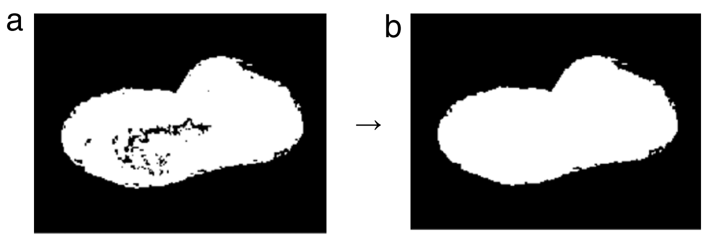

Fig. 5. Filling the thresholded image; (a) before filling operation, (b) after filling operation.

which is expressed in terms of class probabilities ω_i and class means μ_i which regularly can be updated iteratively. This idea yields an effective algorithm as below.

1. Compute histogram and probabilities of each intensity level.
   1. Set up initial $\omega_i(0)$ and $\mu_i(0)$ steps through all possible thresholds $t = 1, 2 \ldots$, *Maximum Intensity*.
   2. Update ω_i and μ_i.
   3. Compute $\sigma_b^2(t)$.
2. Desired threshold corresponds to the maximum $\sigma_b^2(t)$.

After all, mathematical morphology is applied as below [10].

2.4. Mathematical morphology

The proposed parts extracted by the color-space method were considered as potatoes. The potato image was then segmented by threshold to remove the background and normal potatoes. However due to the color discord in this work, the result of the threshold image was a discrete binary image. Three morphological operations including: closing, region filling and area open, are used in order to identify the suspect defects [6]. Region filling had to be adapted. The algorithms for region filling were based on set dilation, complementation, and intersections [8], as shown in:

$$X_k = (X_{k-1} \oplus B) \cap A^c, \quad k = 1, 2, 3, \ldots \tag{5}$$

where A is a set of boundaries and B is a structuring element. The algorithms terminate at iteration step k if $X_k = X_{k-1}$. Fig. 5. shows the filling operation of the threshold potato.

The opening of A by B is obtained by the erosion of A by B, followed by dilation of the resulting image by B, shown as follows:

$$A \circ B = (A \Theta B) \oplus B. \tag{6}$$

The aim of open area is to eliminate small area blemishes that can be ignored by the potato processing industries [11]. Opening of the closed image is shown in Fig. 6.

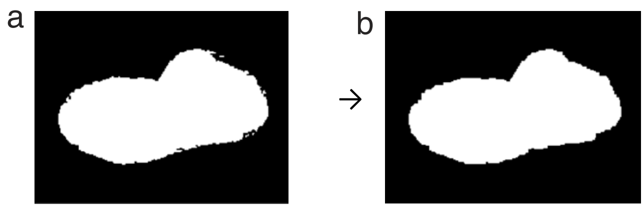

Fig. 6. Opening the closed image; (a) before opening operation, (b) after opening operation.

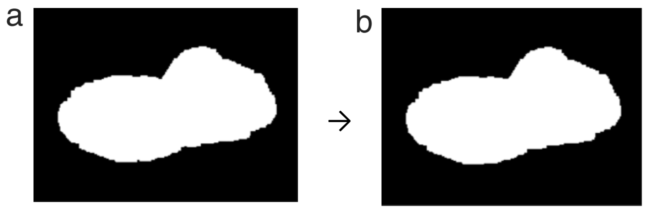

Fig. 7. Closing the filled image; (a) before closing operation, (b) after closing operation.

Closing typically makes counters smooth, fuses narrow breaks and long thin gulfs, eliminates small holes, and fills gaps in the contour. The closing of set A by structuring element B, denoted $A \cdot B$, is defined and mixing thin distances are as:

$$A \cdot B = (A \oplus B)\Theta B. \tag{7}$$

Closing might result in amalgamations of disconnected components, which generates new holes. Fig. 7. shows the result of closing applied to the thresholded image.

However, the morphological operations help to define the image contours and eliminate or reduce noise (undesired speckles outside the contour of the object). Afterwards, two principal properties must be measured as follows:

(1) Potato size.
  (1.1) Length/width diameter ratio.
  (1.2) Maximum size.
  (1.3) Minimum size.
(2) Potato defects.

2.5. Size sorting

To measure the diameter of the potato, we have to compute some denoted cases; to obtain the length/width ratio in a pattern, because of potatoes rotatory movement on the roll, we must first find the length and width of the potato correctly. For this, a four way projection is applied to the binary image.

Formerly, to obtain the standard diameter, a laser beam system was used [11]; instead of that, we use a programming system to find maximum and minimum diameters to decrease the cost.

This algorithm uses a differentiation technique [12] and can notice horizontal, vertical and oblique edges.

In this method, we characterized the rate of black and white pixels of the binary image in four directions as follows. Fig. 8 shows a sample of the work; after analyzing the results, we could notice that the system has some numbers for vertical (2 numbers) and some for horizontal (20 numbers). These lines are maximum and minimum lines (i.e. the oblique diagonals are among these pixels and are not necessary to process). All of maximum lines are equal and also for minimum ones. In this condition a proper accuracy for proper ratio finding is defined.

After that, it is time to find the size of image. For this, we must have a definitive distance for image acquisition. We used United States Department of Agriculture and Canadian Food Industries standards for distant measurement, leading to a standard ratio between size and number of pixels in the image, as follows [13]:

1. Minimum diameter, unless otherwise specified, shall not be less than 38.1 mm in diameter.
2. Maximum diameter, unless otherwise specified, shall not be less than 82.6 mm in diameter.
3. Ratio of maximum and minimum diameter shall not be less than $\alpha = 2.167979$.

To compute the dimensions, we must have a definite distance to compare the numbers of object defined pels (pixels). Fig. 9. shows a misshapen potato. It is definitely that because the ratio ($(208.03/47.02) > \alpha$) of the potato is not standard.

2.6. Defect detection

To recognize the defects of potatoes in a fast and real-time method, supervised classifiers are used; the aim of classification is to categorize all pixels in a digital image into one of several classes; by using supervised classification, we recognize examples of the information classes (i.e. defected pixels) of interest in the image; there are several techniques for

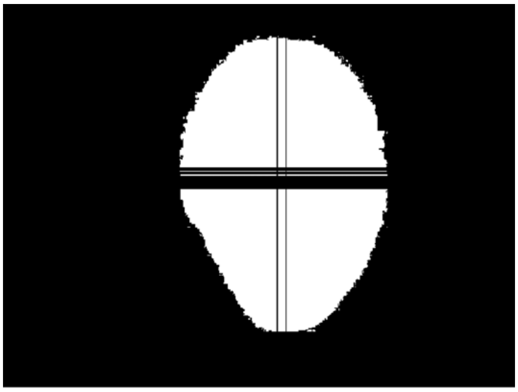

Fig. 8. Size projected image.

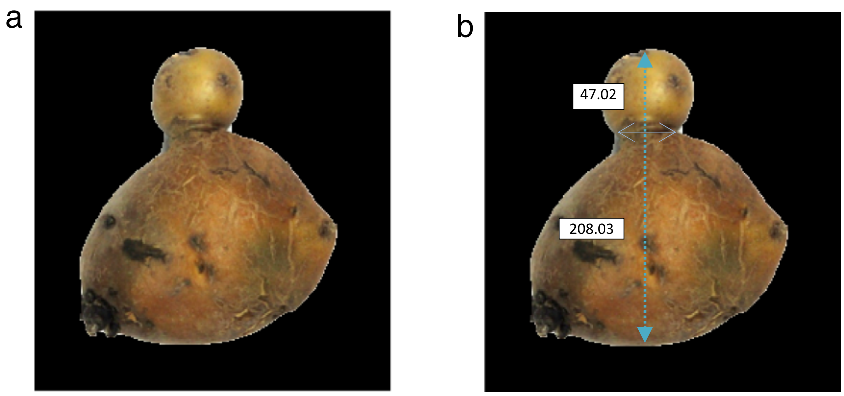

Fig. 9. (a) Segmented image. (b) Detection of a misshapen ratio.

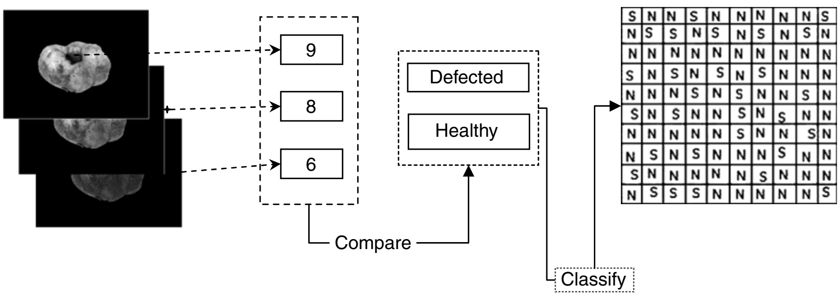

Fig. 10. Steps in supervised classification.

this purpose [14]. Potato defect color is one of the considered cases which can become a classification problem. The purpose of potato defect color pixel classification is to determine whether a color pixel is a defected mode color or not. Good potato color pixel classification should provide coverage of healthy and defected types.

Various classification methods have been developed from Knn classifier to artificial neural networks (Ann) classifiers.

This paper emphasizes the analysis and usage of a popular type of neural networks, Multi Layer Perceptron (MLP). The benefits of the neural network are the generalization authority [15] about the untrained samples due to the massively parallel interconnections and ease of implementation simply by training with training samples for any complicated rule or mapping problem.

However, the performance of the neural network is authentic upon the assumption that we have enough training samples [16,17]. Support vector machines are also new classifiers which are well known nowadays [18]. We also used support vector machines for the comparisons.

Afterwards, the image is then classified by trying the performance for each pixel and making a decision about which of the signatures it resembles most; Fig. 10 shows the steps of classification.

2.6.1. Support vector machines

Instead of employing simple classifiers for detecting defected pixels in an image it is possible to use a support vector machine (SVM); In essence, an SVM is a mathematical entity, an algorithm for maximizing a specific mathematical function with respect to a given collection of data. The SVM will in theory give a better result than using only thresholds for each channel as the network also takes into account the relationship between the channels [18,19]. Support vector machines are very popular for discrimination roles because they can accurately combine a lot of features to find an optimal separating hyper plane, since they are used in some automation purposes like biscuits [20]. SVMs minimize the classification error rate based on two contemporary constraints. They both seek a hyper plane with a major margin (i.e. the distance from the nearest example for segregating the hyper plane) and minimize the number of false classified training samples, using slack variables. If a sample is utterly classifiable in feature space, then the second constraint is not necessary. Albeit this is not the case in our issue, so SVMs minimize both the error on the training set and maximize the margin, increasing their generalization ability [21–23]. SVMs deliver a unique solution, since the optimality problem is convex. This is an advantage compared to neural networks, which have multiple solutions associated with local minima and for this reason may not be robust over different samples.

SVMs use the results of statistical learning and optimization theories in order to maximize their generalization authority for samples. These features showed us that SVM could improve performance on our potato classification. The main purpose is to find a decision surface that "best" classifies the data points into two classes. The decision function in SVM is shown below:

$$y = \operatorname{sgn}\left(\sum_{i=1}^{N} y_i \alpha_i K(x, x_i) + b\right) \tag{8}$$

where x is the d-dimensional vector of a test example, $y \in \{-1, 1\}$ is a class label, x_i is the vector for the ith training example, N is the number of training examples, $K(x, x_i)$ is a kernel function, $\alpha = \{\alpha_1 \dots \alpha_N\}$ and b are the parameters of the model. α_i can lead from [24,25]. In recent years, kernel methods have received major attention, particularly due to the increased popularity of the support vector machines. Kernel functions can be used in many applications as they provide a simple bridge from linearity to non-linearity for algorithms which can be expressed in terms of dot products. In this article, we used 3 types of these kernels for our work that are listed below:

$$\text{Linear-kernel } k(x, y) = x^T y + c \tag{9}$$

$$\text{Polynomial-kernel } k(x, y) = (\alpha x^T y + c)^d \tag{10}$$

$$\text{Quadratic-kernel } k(x, y) = 1 - \frac{\|x - y\|^2}{\|x - y\|^2 + c} \tag{11}$$

where c is a constant, and d is a polynomial degree.

We also used *sequential minimal optimization* (SMO) to have a better result for the work; SMO is a new SVM learning algorithm which is conceptually simple, easy to implement, often faster and has better scaling properties than a standard SVM algorithm. SVMs are starting to enjoy increasing adoption in the machine learning communities, but two of their major weaknesses have meant limited use by engineers. First, the training of SVM is slow, especially for large problems. Second, SVM training algorithms are complicated, tenuous and sometimes difficult to implement. The *quadratic programming* (QP) problem to train an SVM is shown below: we need to find suitable Lagrange multipliers α_i to get the following function to reach its maximum value.

$$\mathrm{Max}W(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} y_i y_j k(x_i, x_j) \alpha_i \alpha_j, \quad \forall i, 0 \le \alpha_i \le C. \tag{12}$$

Subject to

$$\sum_{i=1}^{n} y_i \alpha_i = 0. \tag{13}$$

Unlike standard SVM learning algorithms, which use numerical quadratic programming as an inner loop, SMO uses an analytic QP step. Because SMO spends most of its time evaluating the decision function, rather than performing QP. In such condition, a point is an optimal point if and only if the KKT (Karush–Kuhn–Tucker) conditions are fulfilled. The KKT conditions can be shown as follows:

$$\alpha_i = 0 \Rightarrow y_i f(x_i) \ge 1 \tag{14}$$

$$0 \le \alpha_i \le C \Rightarrow y_i f(x_i) = 1 \tag{15}$$

$$\alpha_i = C \Rightarrow y_i f(x_i) \le 1. \tag{16}$$

The KKT conditions can be evaluated as one example for one Lagrange multiplier once, which is useful in the structure of the SMO algorithm. At every step, SMO selects two Lagrange multipliers to jointly optimize (using KKT condition above), finds the optimal values for these multipliers and updates the SVM to reflect the new optimal deals. The advantage of SMO lies in the fact that solving for two Lagrange multipliers can be done analytically. An entire inner iteration due to numerical

QP optimization is avoided. In addition SMO does not require extra matrix memory (to store previous α_1 α_2 and current α_1 α_2 we only need 2×2 matrices). Therefore, very large SVM training issues can fit inside of the memory of a personal computer. The two Lagrange multipliers must fulfill all of the constraints of the full issue. The inequality constraints cause the Lagrange multipliers to lie in the box. Therefore, one step of SMO must find an optimum of the cost function on a diagonal line segment.

2.6.2. *K-nearest-neighborhood (Knn)*

The *K-nearest-neighbor* (Knn) algorithm is a popular supervised classification algorithm and is defined as a nonparametric supervised pattern classification method. Given N prototype patterns (represented by their feature vectors of dimension D called training sets) and their good classification into several classes, the Knn rule assigns every unclassified pattern (in this work corresponding to a pixel) to the class that is most heavily represented among the k closest prototypes in the feature space [26,27].

The K-nearest-neighbor algorithm in image classification measures the distance between a query pixel and a set of pixels in the data set.

The distance between two pixels can be computed using some distance function $d(x, y)$, where x, y are pixels composed of N features (*in this work* N is 3), such that $x = \{x_1, \dots, x_N\}$, $y = \{y_1, \dots, y_N\}$.

Distance function is used in this work is Euclidean distance and is as below:

$$d_E(x, y) = \sum_{i=1}^{N} \sqrt{x_i^2 - y_i^2}. \tag{17}$$

Because the distance between two pixels is dependant of the intervals, it is recommended that resulting distances be scaled such that the arithmetic mean across the dataset is 0 and the standard deviation 1. This can be accomplished by replacing the scalars x, y with x', y' using the following function:

$$x' = \frac{x - \bar{x}}{\sigma(x)} \tag{18}$$

where x is the unscaled value, $\bar{x}$ is the arithmetic mean of feature x across the data set is its standard deviation, and is the resulting scaled value. The arithmetic mean is defined as:

$$\bar{x} = \frac{1}{N} \sum_{i=1}^{N} x_i. \tag{19}$$

The standard deviation then computes as follows:

$$\delta(x) = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (x_i - \bar{x})^2}. \tag{20}$$

Data set can be represented as a matrix $D = N \times P$, containing P pixels $s^1, \dots, s^p$, where each pixel s^i contains N features $s^i = \left\{s_1^i, \dots, s_N^i\right\}$. A vector o with length P of output values $o = \left\{o^1, \dots, o^p\right\}$ square this matrix, tabulating the output value o^i for each pixel s^i.

It should be noted that the vector o can also be seen as a column matrix; if multiple output values are desired, the width of the matrix may be extended.

Knn in image classification can be run in these steps:

1. Store the output values of the M nearest neighbors to query pixel q in vector $r = \left\{r^1, \dots, r^M\right\}$ by repeating the following loop M times:
   a. Go to the next pixel s^i in the data set, where the current iteration within the domain is $\{1, \dots, P\}$.
   b. If q is not set or: $q < d\,(q, s^i) : q \leftarrow d(q, s^i), t \leftarrow o^i$.
   c. Loop until we reach the end of the data set (i.e. $i = P$).
   d. Store q into vector and into vector r.
2. Calculate the arithmetic mean output across as follows:

$$\bar{r} = \frac{1}{M} \sum_{i=1}^{M} r_i. \tag{21}$$

3. Return $\bar{r}$ as the output value for the query scenario q.

2.6.3. *Multi Layer Perceptrons (MLPs)*

The back propagation algorithm is a generalization of the least mean square algorithm that improves network weights to minimize the mean squared error between the desired and actual outputs of the network [28,29]. Back propagation employs supervised learning in which the network is trained using data for which inputs as well as desired outputs are known. Once

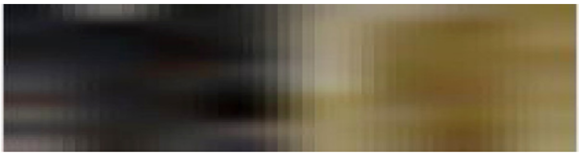

Fig. 11. Test set: (a) defected sets, (b) healthy sets.

trained, the network weights are vapid and can be used to calculate the output values for new input samples. Model Standard *Multi Layer Perceptron* (MLP) architecture consists of more than 2 layers; An MLP can have any number of layers, units per layer, network inputs, and network outputs.

MLP in image classification can be run in some steps as described below.

Start with randomly chosen weights.

While *MSE* is unsatisfactory and computational bounds are not exceeded, do for each input pattern.

(a) Compute hidden node inputs.
(b) Compute hidden node outputs.
(c) Compute inputs to the outputs nodes.
(d) Compute the network outputs.
(e) Compute the error between network outputs, and desired output.
(f) Modify the weights between hidden and output nodes.
(g) Modify the weights between input and hidden nodes.
End

where R is the input vector, IW is the weight vector, and b is the bias vector [29].

2.6.4. Apply classification

A two zone classification (as healthy or defected individually) is used in this paper; we discuss *pixel-based* potato detection classifiers, which classify each pixel, independently from its neighbors. Typically databases are in a 448 × 336 and the others convert into these two sizes for easier comparisons. The structure of the classifiers is a vector of $3 \times n$ pixel coefficient vectors from each face or non face image where n is the number of neurons in the hidden layer: since the transfer function used is a sigmoid function (for MLP); it will produce an output between 0 and 255 (uint8 mode). Thus, the output of the neural network needs to be modified so that it is either 0 or 1. In this work, a single threshold value is used in determining the potato and non-potato classes; if the MLP network output is higher than the threshold value, then the output is 1. Otherwise, the output is 0. The threshold value used in this work is 128. After all, morphological processing consists of filling holes, and opening followed by closing is applied to the resulting images [28]. The Present work has implemented an artificial neural network and consists of three methods as follows: Standard three layer MLP with a back propagation algorithm, with features of image RGB color pixel values are taken as input values, and defected or healthy potatoes as the output. Here we use log-sigmoid as the activation function. A batch steepest mode gradient descent, because of less of a requirement to update [30] (that inputs and biases are applied to network before they updated) is selected, and other parameters are chosen as: Learning rate = 0.01, Epochs = 1000, Target error goal: 10^{-4}.

SVM has been applied to different applications such as texture classification, as in [17–21]. In this paper, we used 3 different kernels and Sequential Minimal Optimization for check the best classifier for potato classification purposes. In this work, C (optimal parameter) is equal to 100, and end-accuracy is 10^{-4}. The evolutional situation was quite the same as each model. We also used a 4 nearest neighborhood for classification; Fig. 11 shows a part of the dataset which is used to compare the methods.

To evaluate the performance of the proposed algorithm towards gradient descent, three performance metrics are defined. The first metric is the correct detection rate (*CDR*) and is given in Eq. (22). The false acceptance rate (*FAR*) is the percentage of identification moments in which false acceptance occurs. The false rejection rate (*FRR*) is the percentage of identification moments in which false rejection occurs. The FAR and FRR are expressed in Eqs. (23) and (24), respectively:

$$CDR = \frac{No.of.Pixels.Correctly.Classified}{Total.Pixels.in.the.Test.Dataset} \tag{22}$$

$$FAR = \frac{No.of.non-Potato.Pixels.Classified.as.Potato.PixelsClassified}{Total.Pixels.in.the.Test.Dataset} \tag{23}$$

$$FRR = \frac{No.of.Potato.Pixels.Classified.as.non-potato.PixelsClassified}{Total.Pixels.in.the.Test.Dataset}. \tag{24}$$

Tables 1 and 2 present the performance of compared classifiers in accuracy and speed.

It is obvious from the above that a support vector machine with sequential minimal optimization has better performance in both accuracy and time; Fig. 11 shows the final results of the paper containing both size sorting and defect detection by SMO. Finally, 96.86% accuracy for sizing and about 95% for the best classifier is reached (see Figs. 12–15).

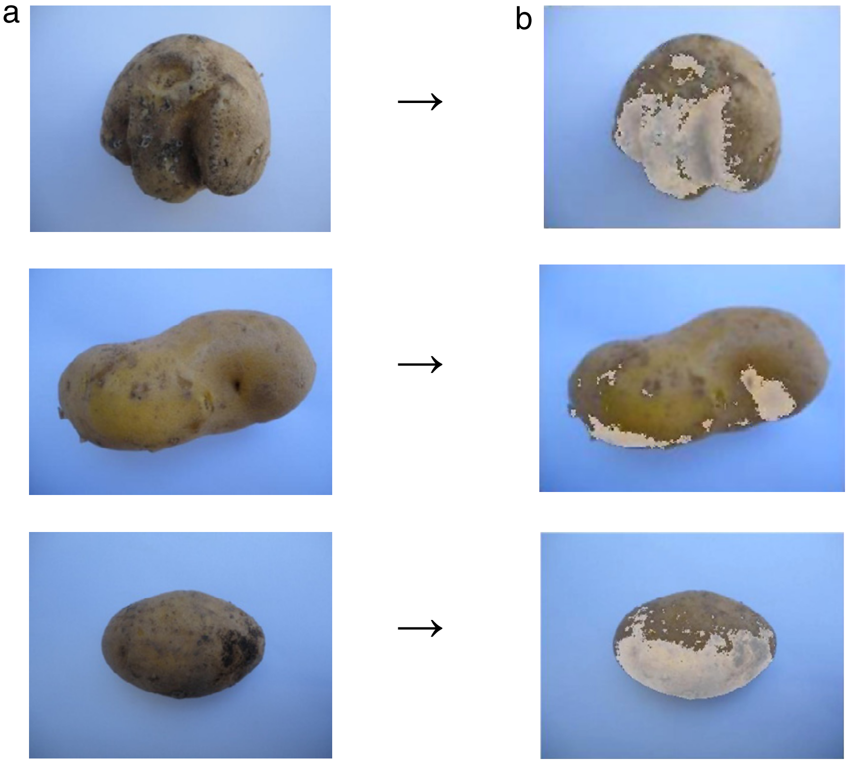

Fig. 12. Sample images from the defect detection database. (a) Original image, (b) defect segmented result.

Fig. 13. Sample images for size sorting. (a) Original image, (b) binarized image, (c) size detecting.

Fig. 14. Sample images for size sorting. (a) Original image, (b) binarized image, (c) size detecting.

Table 1
Classification comparison of performance in the presented methods.

	Method					
	Knn	MLP	SVM-linear kernel	SVM-poly nominal kernel	SVM-quadratic kernel	***SVM–SMO based***
CDR (%)	91.25	95	95	95	95	***95***
FAR (%)	6	2	3	3	3	***4***
FRR (%)	2.75	3	2	2	2	***1***

Fig. 15. Sample images for size sorting. (a) Original image, (b) binarized image, (c) size detecting.

Table 2
Classification comparison of speed in the presented methods.

Method	Operation time (s)
Knn	0.8750
MLP	0.1131
SVM-linear kernel	0.0571
SVM-poly nominal kernel	0.1076
SVM-quadratic kernel	0.0970
SVM–SMO based	***0.0211***

3. Conclusions

A system for identifying surface defects on potatoes was designed, based on analyzing images acquired while potatoes were rotating in front of the camera. When multiple images were combined and adjustments made for rotation, dark areas caused by defects would appear with almost the same shape and at the same place in three or more frames. The proposed algorithm was able to detect defects. To increase the standard mode for the system, in parallel with defect detection, a simple size sorting is also applied. The best algorithm method in classification has an accuracy of about 95% and the size grading section has an accuracy of about 96.86% for the samples in these experiments.

References

[1] D. Nieto, E.V. Fernández, A. Formella, F. Martin, S.T. Guijarro, H.G. Jorge, A genetic algorithm approach for feature selection in potatoes classification by computer vision, in: 2009 Industrial Electronics, IECON'09, 35th Annual Conference of IEEE, Porto, pp. 1995–1960.
[2] V. Chalana, Y. Kim, L. Zhou, PC-based machine vision for real-time computer-aided potato inspection, International Journal of Imaging Systems and Technology 9 (1998) 423–433.
[3] J.C. Noordam, G.W. Otten, A color vision system for high speed sorting of potatoes, Agrotechnological Research Institute No: 00-AE-002, Netherlands, 2000.
[4] J.C. Noordam, W.H.A.M. van den Broek, L.M.C. Buydens, Multivariate image segmentation with cluster size insensitive fuzzy C-means, Chemometrics and Intelligent Laboratory Systems 64 (2002) 65–78.
[5] F. Pedreschi, J. Leo'n, D. Mery, P. Moyano, Development of a computer vision system to measure the color of potato chips, in: 2006 Food Research International Conference, vol. 39, pp. 1092–1098.
[6] L. yang jun, W. Jiawen, Graph and Image Processing Based on MATLAB 7.0, National Defense Industry Press, 2006, pp. 226–276.
[7] F. Junlong, W. Shuwen, Z. Changli, Automatic identification and classification of tomatoes with bruise using computer vision, Transactions of the CSAE 21 (8) (2005) 98–101. Harbin, China.
[8] Jing Jin, Jinwei Li, Guiping Liao, Xiaojuan Yu, Leo Christopher C. Viray, Methodology for potatoes defects detection with computer vision, in: International Symposium on Information Processing, Huangshan, 2009, pp. 346–351. ISBN: 978-952-5726-02-2.
[9] Andrew Greensted, 2010. http://www.labbookpages.co.uk/software/imgProc/otsuThreshold.html.
[10] Payman Moallem, Bibi Somayeh Mousavi, S. Amirhassan Monadjemi, A novel fuzzy base system for pose independent faces detection, Applied Soft Computing 1 (2) (2011) 1801–1810. Amsterdam.
[11] J.C. Noordam, G.W. Otten, A.J.M. Timmermans, B.H. van Zwol, High-speed potato grading and quality inspection based on a color vision system, in: Proc. SPIE 3966 Conference, USA, 2000. doi:10.1117/12.380075.
[12] Milan Sonka, Vaclav Hlavac, Roger Boyle, Image Processing, Analysis and Machine Vision, 2nd ed., PWS, Pacific Grove, ISBN: 053495393X, 1999.
[13] Mahler Sarah, US potato board, Guide to US seed potato export varieties, International Marketing Manager, Seed and Fresh Programs 7555 East Hampden Ave., Suite 412 Denver, Colorado 80231, USA.
[14] D. Zhang, Md. Monirul Islam, G. Lu, A review on automatic image annotation techniques, Pattern Recognition (2011) doi:10.1016/j.patcog.2011.05.013.
[15] W.F. Schmidt, et al. Initialization, back propagation, and generalization of feed forward classifier, in: IEEE Int. Conf. Neural Networks, 1993, pp. 598–604.
[16] R.P. Lippmann, B. Gold, Neural-net classifiers useful for speech recognition, in: IEEE Int. Conf. Neural Networks, San Diego, CA, vol. 4, 1987, pp. 417–425.
[17] Joongho Chang, Gunhee Han, Jos'e M. Valverde, Norman C. Griswold, Cork quality classification system using a unified image processing and fuzzy-neural network methodology, IEEE Transactions on Neural Networks 8 (4) (1997).
[18] S. Theodoridis, K. Koutroumbas, Pattern Recognition, third ed., Academic Press, CA, 2006.
[19] R.G. Brereton, G.R. Lloyd, Support vector machines for classification and regression, Analyst 135 (2) (2010) 230–267.
[20] S. Nashat, A. Abdullah, S. Aramvith, M.Z. Abdullah, Support vector machine approach to real-time inspection of biscuits on moving convey or belt, Computers and Electronics in Agriculture (2010) doi:10.1016/j.compag.2010.10.010.
[21] K.P. Sudheer, S.K. Jain, Radial basis unction neural network for modeling rating curves, Journal of Hydrologic Engineering—ASCE 8 (3) (2003) 161–164.
[22] C. Cortes, V. Vapnik, Support vector networks, Machine Learning 20 (1995) 273–297.
[23] Yih-Lon Lin, Jer-Guang Hsieh, Hsu-Kun Wu, Jyh-Horng Jeng, Three-parameter sequential minimal optimization for support vector machines, Neuro computing (2011) doi:10.1016/j.neucom.2011.06.011.

[24] Rong Yan, Yan Liu, Rong Jin, Alex Hauptmann, On predicting rare classes with SVM ensembles in scene classification, in: IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP'03, April 6–10, 2003.
[25] K.I. Kim, K. Jung, S.H. Park, H.J. Kim, Support vector machines for texture classification, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (11) (2003) 1542–1550.
[26] P.-N. Tan, M. Steinbach, V. Kumar, Introduction to Data Mining, Pearson Addison-Wesley, 2006.
[27] Shakhnarrovish, Darrell Indyk, Nearest Neighbor Methods in Learning and Vision, MIT Press, 2005.
[28] Rong Yan, Yan Liu, Rong Jin, Alex Hauptmann, On predicting rare classes with SVM ensembles in scene classification, in: IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP'03, April 6–10, 2003.
[29] Matlab Tutorial, Mathworks, 2007.
[30] Paul J. Werbos, Beyond regression "New tools for prediction and analysis in the behavioral sciences", Ph.D. Thesis, Harvard University, 1974.