Design and Implementation of an Automated Fruit Quality Classification System

Han Suk Choi*

Abstract

Most of fruit quality classification has been done by time consuming, inaccurate and intensive manual labor. This study proposed an automated fruit grading system based on appearances and internal flavors. In this study, image processing technique and a weight checker were used to measure the value of appearance features and the near infrared spectroscopy analysis method was used to estimate the value of internal flavors. Additionally, I suggested 8x8x5x5 ANN based fruit quality classifier model to grade fruits quality. The proposed automated fruit quality classification system is expected to be very beneficial for many farms where heavy manual labor is usually needed for fruit quality classification.

Keywords : automated fruit quality classification system | fruit appearance feature extraction | NIR spectroscopy analysis | fruit quality classifier

I. INTRODUCTION

In most fruit farms, the fruit quality classification has been made by human labor. The grading of fruit quality that persons take is very time consuming, inaccurate, and labor intensive work. Therefore, an automated fruit quality classification system is needed for grading fruits in real time at low cost. In this study, an automated fruit quality classification system utilizes the external appearances and internal characteristics of the fruit to improve grading performance[1]. Examples of appearance features are fruit weight, shape, size, volume, color, and outside bruise marks or scratches. Internal flavor factors are sweetness, hardness, crispness, bitterness, acidity, saltiness, moisture and nutrients[2].

This study proposes an automatic real-time fruits quality classification system applying by appearances features and internal flavor factors. In this study, weight, long axis length, short axis length, and volume were applied as external features and sugar content, hardness, acidity, and moisture were used as internal flavor factors. This study was performed using an image processing technique and a weight checker in order to calculate the value of external appearance features and the near-infrared(NIR) spectroscopy analysis method was used to estimate the value of internal flavors.

The proposed automated fruit quality classification system consists of hardware and software. The hardware components are a weight checker, a color CCD camera, a NIR spectroscopy, a data controller, and a LCD display on the automatic conveyer belt connected with a computing server. The software components are an external appearance measurer, defects detector, appearance classifier, internal flavor measurer, the final grading classifier, measured data analysis APIs.

The experiments of fruit quality grading were carried out with Korean pears. This study computed pear's weight, long axis length, short axis length, and volume as external features using color image processing methods. And the entropy image analysis technique was applied to find fruit bruises and scratch marks. The estimated values of internal flavors such

*Member, professor in department of Computer Engineering, Mokpo National University

Manuscript : 2018 11. 26
Revised : 2018. 12. 11.
Confirmation of Publication: 2018. 12. 21.
Corresponding Author :Han Suk Choi,
e-mail : chs@mokpo.ac.kr

as sweetness, hardness, acidity, and moisture were computed by the NIR spectral processing analysis techniques. And an artificial neural network model was suggested to sort fruits quality classification.

The proposed artificial neural network model consisted of 4 layers with 8 input nodes, 3hidden layers and 3 output nodes. In our experiments, we trained and tested the artificial neural network model with 1,800 numbers of Korean pears for quality classification. It has achieved the classification accuracy rate of 97.41% in our experiment. The automated fruit quality classification system proposed in this study is very beneficial for many farms which are doing large quantities of fruit quality classification by people. If the proposed system is commercialized, labor costs for fruit farms will decrease and productivity will increase significantly.

II. SYSTEM DESIGN AND METHODOLOGIES

2.1 System Design

Design architecture of the automated fruit quality classification system is shown in Fig 1.

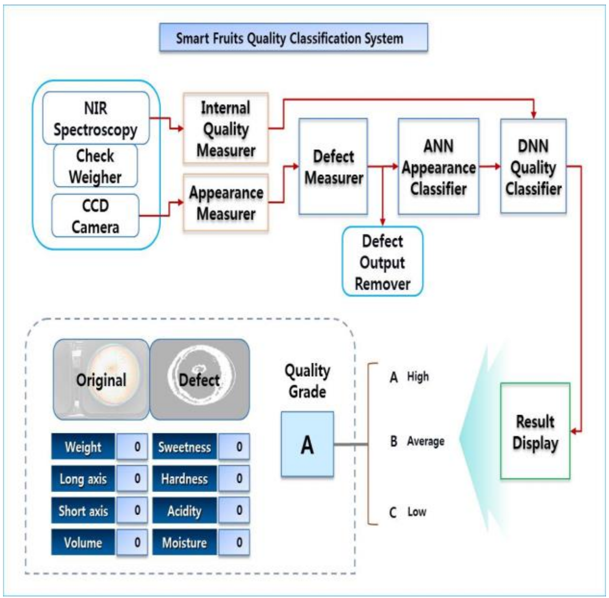


Fig 1. Design Architecture of the Automated Fruit Quality Classification System

Main components of the automated fruit quality classification system are a weight checker, a CCD camera, a NIR spectroscopy, an appearance measurer, an internal quality measurer, a defect measurer and defect output remover, an artificial neural network classifier, and a LCD display panel. The weight checker measures a fruit weight placed on the cup of the automatic conveyer belt in the proposed system. The CCD camera takes a fruit picture when the fruit is passed through the CCD camera box. The fruit color image is saved for computing values of appearance features such as long axis length, short axis length, and volume of the fruit. The saved fruit color image will be used for detection of the fruit bruise or scratches marks.

The NIR spectroscopy is used for estimating the values of internal flavor factors of the fruits such as sweetness, acidity, hardness, and moisture of the fruit. In our designed system, the spectrum range of the NIR wavelength is applied in 400~1,100nm. The LCD display panel shows all measured values of the fruit appearance features and the estimated values of internal flavors along with the original fruit image and preprocessed fruit image. And finally the grading values such as grade A, B, C are displayed in the LCD panel of the proposed system.

2.2 Methodologies

1) Fruit Appearance Feature Extraction

A fruit color image is captured by the CCD camera installed in the external appearance measurer box. The light intensity of the external appearance measurer box is very important to take high quality of the fruit image.

(1) Removing the back ground fruit image

The method of removing the background image from the original fruit image is to use a differential image processing technique. For the input image $I(x,y,t)$ and the initial background $B(x,y,t)$, the expression (1) is removing the back ground image[3].

$$F(x,y,t) = |I(x,y,t) - B(x,y,t)| \quad .... (1)$$

(2) Detection of the fruit image outline

The Sobel edge operator is used to detect the outline of a fruit image after removing the background fruit image. The advantage of using Sobel operator is highly resistant to noise in the edge detection. The Sobel operator is more sensitive to find the diagonal directions edges than the horizontal or vertical edges. The basic Sobel operator mask size is 3x3. The expression (2) is used to find the direction and the amplitude of the outline edges[3].

$$G = \sqrt{G_x^2 + G_y^2}$$
$$\theta = \tan^{-1}\left(\frac{G_y}{G_x}\right) \quad \cdots\cdots\cdots\cdots (2)$$

(3) Computation of a fruit volume from the length of long axis and short axis

In order to compute a fruit volume and the length of long axis and short axis, this study suggests to apply the first, second, third central moment calculation. The central moment computes the center of distribution and its surrounding moment. The central moment can be defined from the difference moment using the expression (3) and (4):

$$m_{p,q} = \sum_{i=1}^{n} F(x,y,t)\, x^p y^q \quad \cdots\cdots\cdots\cdots (3)$$
$$\mu_{p,q} = \sum_{i=0}^{n} F(x,y,t)(x-\bar{x})^p (y-\bar{y})^q \quad \cdots (4)$$

where $\bar{x} = m_{10}/m_{00}$, $\bar{y} = m_{01}/m_{00}$.

The expression (5) is to compute a fruit ellipse slope :

$$\theta = \frac{1}{2}\tan^{-1}\left(\frac{2\mu_{11}}{\mu_{20}-\mu_{02}}\right) \quad \cdots\cdots\cdots\cdots (5)$$

The expression (6) is to compute the length of long-axis and short-axis from a fruit image :

$$a = \left(\sum_x \sum_y (x-\bar{x})\cos\theta - (y-\bar{y})\sin\theta\right)^2$$
$$b = \left(\sum_x \sum_y (x-\bar{x})\sin\theta - (y-\bar{y})\cos\theta\right)^2$$
$$I_{max} = \sqrt[4]{\frac{4}{\pi}}\sqrt[8]{\frac{b^3}{a}}$$
$$I_{min} = \sqrt[4]{\frac{4}{\pi}}\sqrt[8]{\frac{a^3}{b}} \quad \cdots\cdots\cdots\cdots (6)$$

The value of a and b in the expression (6) refers to the radius of the long axis(vertical diameter) and the short axis(horizontal diameter) of the mapped fruit. I_{max} and I_{min} in the expression (6) can be calculated from the value of long axis length, a and short axis length, b. Finally, the expression (7) is to compute a fruit volume [3][4].:

$$V_{Fruit} = \frac{4}{3} I_{max}^2 I_{min} \pi \quad \cdots\cdots\cdots\cdots (7)$$

(4) Example of the elliptic curve of a fruit image

Fig 2 shows the elliptic curve of a Korean pear. The left figure in Fig 2 shows the vertical and horizontal diameter of an original Korean pear image. The fruit volume and lengths of long axis and short axis are computed by using the expression (2), (3), (4), (5), (6) and (7).

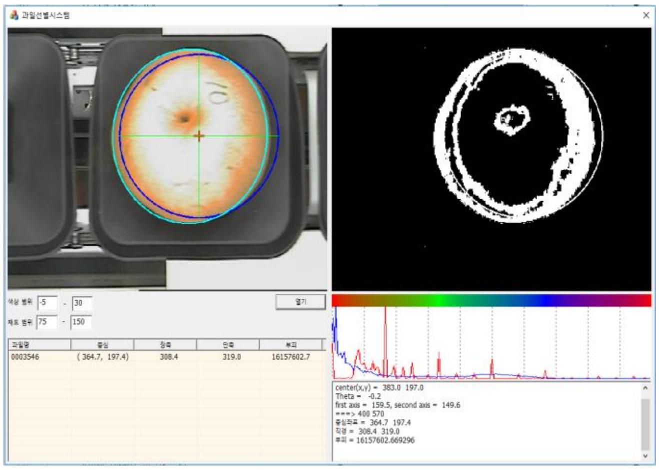


Fig 2. An example of the elliptic curve of a Korean pear image

(5) Volume estimation with multiple regression analysis

In this study, a fruit volume has been calculated by central moment computation method, but the calculated volume could be a little difference from an actual fruit shape. So we need to calibrate the computed value of a fruit volume applying a multiple regression analysis to increase the accuracy of the fruit volume values. This study has been applied two hundred pears to calibrate the length of long axis and short axis for fruit volume values. In our multiple regression analysis, the long axis and short axis are used as independent variables and volume is used as a dependent variable. The expression (8) is the multiple regression expression[3] :

$$Y = \beta_0 + \beta_1 I_{max} + \beta_2 I_{min} + \beta_3 I_{max} I_{min} \quad \cdots (8)$$

Where
$$\beta = [780.56, -8.88, -10.99, 0.17]$$.

2) Detection Method of a Fruit Bruise using Image Entropy Analysis

Bruises and scratches on the outside of the fruit must be found and removed during the fruit quality grading process. These bruise could be estimated by the edge extraction, different color distribution analysis, or fruit entropy analysis[5]. We used the entropy image analysis technique to find fruit bruises.

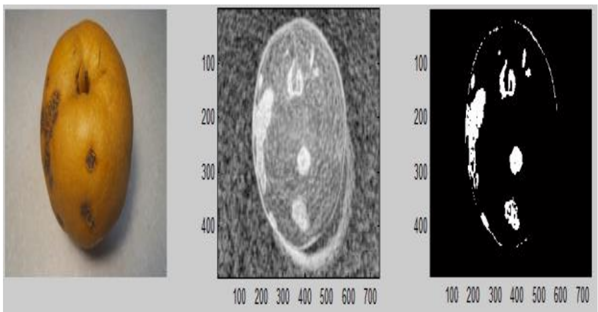
First of all, we compute the total color distribution or histogram of a fruit image. Then we compute the probability of bruising at each pixel from the fruit image. The expression (9) is a entropy computation method[5]:

$$\mathrm{E} = -\sum_{x} p(x) \log_2 p(x) \quad \cdots\cdots\cdots\cdots\cdots (9)$$

In general, the entropy of a portion of a bruise on the fruit is higher than other parts. We have to calculate the change ratio of local entropy to find bruises on a fruit. The local entropy is calculated as the sum of the entropy values of the surrounding entropies of a defined point. The expression (10) is the local entropy computation method :

$$\mathrm{E}_{\mathrm{local}} = -\sum_{x}\sum_{y} p(I(x,y)) \log_2 p(I(x,y)) \quad \cdots (10)$$

Figure 3 shows the process of finding the bruises on the Korean pear using the entropy image analysis technique. Figure 3a) is an original fruit color image with some bruises on the pear surface. Fig 3b) shows the pear's entropy image with bruises. The bruises parts of the pear shows more brightly compared to the other parts of the pear since the entropy of bruises part has higher local entropies. Fig 3c) shows a binary pear image Figure 3 shows a binary pear image obtained by AND operation with an entropy image and a mask image.



a) Pear color image with bruises b) Pear entropy image c) Pear binary image

Fig 3. Process of Finding Bruises on the Korean Pear using Entropy Image Analysis

3) NIR Spectroscopy Analysis Method for Measuring Fruit Internal Flavors

The main components of a NIR spectroscopy device for measuring fruit internal flavors such as sweetness, hardness, acidity, moisture, and so on consists of a light source analysis part with an optical fiber probe and a spectrometer for splitting and grating with a CCD array on the A/D board. The NIR spectroscopic analysis of the fruits applied in this study is as follows. One sample fruit is irradiated with a tungsten halogen lamp light, and the transmitted or reflected light is transmitted to the spectrometer using an optical fiber. The internal flavors of the sample fruit that have passed through the spectrometer slit are broken down by each wavelength through the grating[5].

The sample fruit was irradiated with a tungsten halogen lamp around the sample fruit and the transmitted or reflected light is transmitted to the real time spectrometer using an optical fiber. The light passing through the spectrometer slit is broken down by each wavelength through grating. As the wavelength for the internal flavors of the fruit is transferred to the CCD array, the wavelength energy is measured at each position of the internal components. The wavelength spectrum is produced in accordance with the various internal flavors of fruit. The wavelength spectrum for the internal flavors of the fruit will generally include noise. There are many factors that affect the making of noise in the wavelength spectrum during analysis of the internal flavors of the fruit[6].

In this study, we used various filtering techniques such as smoothing, multiple scatter correction, standard normal variation transform, and differentiation from the measured NIR spectrum to remove the noise in the wavelength spectrum. This study was applied to a predictive equation known as a calibration equation in order to estimate a specific internal flavor value in the Korean pear like sweetness. The statistical calibration equation such as multi−linear regression(MLR), principal component regression(PCR), and partial least square regression(PLS) methods[7] were used for the prediction expression development in this study. Fig 4 shows examples of NIR spectra and distribution of sweetness of 1,800 Korean pears.

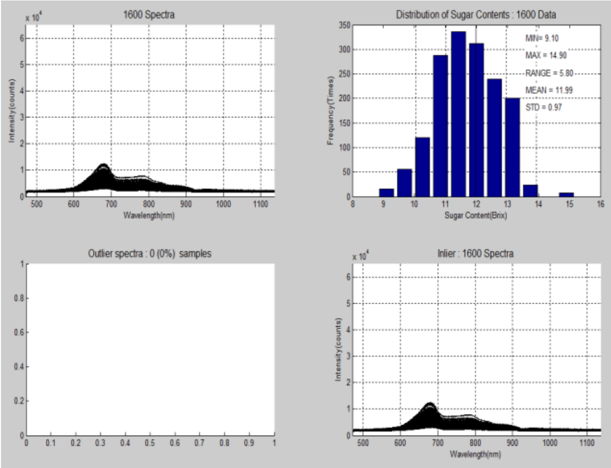


Fig 4. NIR spectra and distribution of the sweetness of 1,800 Korean pears

4) Fruit Quality Classifier based on Artificial Neural Networks(ANN)

This study proposed an artificial neural network model as the fruit quality classifier. The ANN is a weighted connection network system with multiple artificial neurons on the input, output, and hidden layers. There are various kinds of artificial neural network structures according to their topologies and learning algorithms[8]. This study proposed an ANN feed−forward model for an automated fruit quality classification.

Fig 5 shows a fruit quality classifier based on the ANN. This ANN based fruit quality classifier has input layer with 8 nodes, output layer with 3 nodes, and 2 hidden layers. The first hidden layer had 8 nodes and the second layer had 5 nodes. Eight input values consists of the fruit appearance features and internal flavors such as weight, long axis, short axis, fruit volume, sweetness, hardness, acidity, and moisture. Three output nodes are mapped into a grade ‘A, ‘B’, and ‘C’ as the best quality, the medium quality, and the lowest quality.

This ANN based fruit quality classifier model is a fully connected feed−forward supervised learning model. It has been trained and tested with 1,800 Korean pears. The dataset of eight input values was first standardized by removing the mean and scaling to unit variance. Each input value has been centered and scaled for the relevant statistics standardization on samples of the training set. Standardization of dataset is a common requirement for the ANN learning estimators[9].

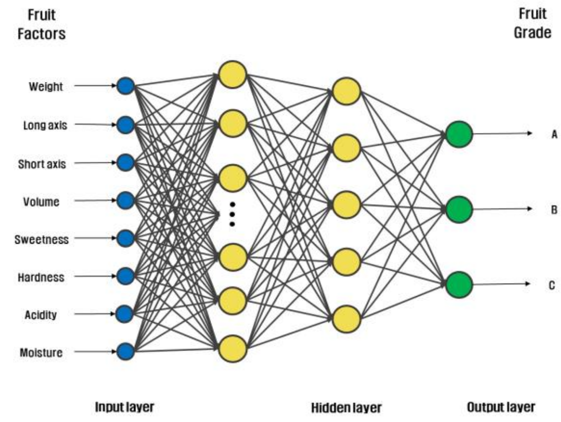


Fig 5. ANN based Fruit Quality Classifier

In this study, we choose the ReLU(Rectified Linear Unit) activation function which gives an output x if x is positive and 0 otherwise. And the classical stochastic gradient descent is the following expression (11).

$$w \leftarrow w - \eta\left(\alpha \frac{\partial R(w)}{\partial w} + \frac{\partial Loss}{\partial w}\right) \quad \cdots(11),$$

where η is a control parameter for training speed and $Loss$ is a loss function. The Adam optimization algorithm was used to update network weights iterative based in training data. Adams optimizer is a popular algorithm because it can achieve the significant result in small time compared to the classical stochastic gradient descent procedure[10] [11].

III. IMPLEMENTATION RESULT

3.1 System Implementation Result

Fig 6 shows a manufactured prototype of Korean pear quality classification system. It was built according to the schematic design architecture of the automated fruit quality classification system proposed in this study[12].

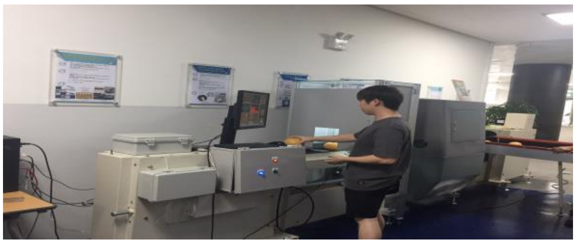


Fig 6. A Manufactured Prototype of Korean Pear Quality Classification System

3.2 Training and Testing Result

In this study, we were trained for inputs as well as outputs to adjust weights for the Korean pear quality grading. These weights along with different input values were then fed to the proposed ANN based fruit quality classifier for testing. The input values were weight, long axis, short axis, volume, sweetness, acidity, hardness, and moisture. Outputs were three grade classes like grade A (Best Quality), grade B(Medium Quality), and grade C(Low Quality).

The total number of pears for training and testing was 1,800. Then the 1,260(70% of all experimental pears) were used for training purpose while 540(30% of experimental pears) pears for testing. In these dataset, we adjusted the train weight values in some important attributes such as pear's physical weight, volume, sweetness because there were no standard grading values of Korean pears in other references. So we set the measured attribute dada in real for our experiment, For example, the average weight was 903.47 gram with 29.53 gram standard deviation and the average sweetness was 14.22 Brix with 0.19 standard deviation for grade A. The average weight was 814.90 gram with 28.25 gram standard deviation and the average sweetness was 14.13 Brix with 0.18 standard deviation for grade B. And the average weight was 733.01 gram with 18.13 gram standard deviation and the average sweetness was 14.03 Brix with 0.17 standard deviation for grade C, and so on.

Table 1 shows the testing result of the ANN based Korean pears quality classifier. As shown in this table 1, 3 pears out of 101 in class A were misclassified to class B, so the error classification rate of class A instead of class B was 2.97%. The only 2 pears out of 163 in class C were mismatched to class B, so the error classification rate of class C instead of class B was 1.23%. And 9 pears out of 276 in class B were misclassified to class A(4) and class C(5), so the error classification rate of class B instead of class A and class C was 3.26%.

It had achieved total accuracy rate of 97.41% in our experiment. This testing was performed with normal pears because the bruised and cracked pears were get rid of in preprocessing process of the proposed smart fruit quality grading system.

Table 1. Testing Result of Korean Pears Quality Classifier

Pear grade (numbers)	A	B	C	Mismatched Pears
A(101)	98	3	0	3
Error rate	−	2.97	0.0	
B(276)	4	267	5	9
Error rate	1.45		1.81	−
C(163)	0	2	161	2
Error rate (	0.0	1.23	161	
Total accuracy rate	−			97.41 (14/540)

3.3 Discussion

This study shows a good possibility of development method for a practical automated fruit quality classification system using appearance features and internal flavor factors in order to decrease human labor cost in fruit industry. However, the proposed system should be tested in a variety of real fruit farms and adapted by a number of real testing process according to various kinds of fruits. The system should be improved by testing on a variety of fruit growth environments and farmers effort.

IV. CONCLUTION

In most fruit farms, the fruit quality classification has been made by human labor. The grading of fruit quality that persons take is time consuming, inaccurate, and labor intensive work. This study proposed an automated fruit quality classification system applying by appearances features and internal flavor factors.

In this study, weight, long axis length, short axis length, and volume were applied as external appearance features and sweetness, hardness, acidity, and moisture were used as internal flavor factors. This study was performed using an image processing technique and a weight checker in order to calculate the value of external appearance features and the near-infrared(NIR) spectroscopy analysis method was used to estimate the value of internal flavors. And 8x8x5x5 ANN based fruit quality classifier model was suggested to classify fruits quality grading.

In our experiments, we trained and tested the proposed ANN based fruit quality classifier model with 1,800 numbers of Korean pears for quality classification. It has achieved the classification accuracy rate of 97.41% in our experiment. The reasons of the error classification were caused by the inaccurately measured dataset in this study. So it is very important to get the real accurate dataset for the commercialization.

The automated fruit quality classification system proposed in this study is very beneficial for many farms which are doing large quantities of fruit quality classification by people. If the proposed system is commercialized, labor costs for fruit farms will decrease and productivity will increase significantly.

REFERENCES

[1] Han S. Choi, Hong S. Choi, and H. D. Mun, " A Smart Fruits Quality Classification Hardware Design using the Near-Infraed Spectroscopy and Image Processing Technologies," *International Conference on Convergence Content(ICCC 2016)*, pp. 9-10, 2016.

[2] S. Naik and B. Patel, "Machine Vision based Fruit Clssification and Grading - A Review," *International Journal of Computer Applications(0975-8887)*, vol. 170, No.9, pp. 22- 34, July, 2017.

[3] Jong H. Shin, Digital Image processing, *Hanbit Academy*, 2015

[4] Sang Cheol Park, Kang Han Oh, In Seop Na, Soo Hyung Kim, Hyung Jeong Yang, Guee Sang Lee, "Noise Removal for Level Set based Flower Segmentation," *Smart Media Journal*, vol. 1, no. 2, pp.34-39, June, 2012.

[5] Mert R. Sabuncu, Entropy-based Image Regestration, *Ph.D. Dissertation, Prinston University*, Nov., 2006.

[6] J. Yoon, M. So, C. Han, H. Y. Kim, "Comparison of Prediction Models of the Sugar Content for a Nondestructuve Sorting System Based on Near-infrared," *Proceedings of KIIT Summer Conference*, pp. 414-417, 2014.

[7] Jacob Cohen, P. Cohen, S.G. West, L. S. Aiken, *"Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences ,"* 3_{rd} Edition, lawrence Erlbaum Associates, Publishers, 2003

[8] I. Sa, Z. Ge, F. Dayoub, B. Upcroft, T. Perez, and C. McCool, " DeepFruit: A Fruit Detrection System using Deep Neural Networks," *sensors 2016, 16, 1222*, doi:10.3390/s116081222.

[9] Hae-Min Moon, Jin-Won Park, Sung Bum Pan, "Performance Analysis of Face Recognition by Face Image resolutions using CNN without Backpropergation and LDA," *Smart Media Journal*, vol. 5, no. 1, pp.24-29, March, 2016.

[10] Kaifeng Ly, Shunhua, Jian Li, "Learning Gradient Descent:Better Generalization and Longer Horizons," *Proceedings of the 34^{th} International Conference on machine Learning*, Sydney Australia, PMIR 70 2017.

[11] Sun Park, Jongwon Kim, "Red Tide Algea Image Classification using Deep Learning based Open Source," *Smart Media Journal*, vol. 7, no. 2, pp34-39, June, 2018.

[12] Han Suk Choi, Je Bong Cho, Sang Gyun Kim, Hong Seok Choi, "A Real-Time Smart Fruit Quality Grading System Classifying by External Appearance and Internal Flavor Factors," *IEEE 2018 International Conference on Industrial Technology(ICIT)*, pp. 321-326. Feb., 2018.

Authors

Han Suk Choi

He received his B.S. degree in Mathematics Education from Chonnam National University, Korea in 1980, M.S. degree in Computer Science from Western Illinois University, USA in 1986, and Ph.D degree in Computer Science from Chunbuk National University in 1997. Since 1989, he has been a professor in department of Computer Engineering, Mokpo National University, Korea. His research interests are Big Data Analysis, Deep Learning, and System Integration of Computer Applications.