Computational Deep Intelligence Vision Sensing for Nutrient Content Estimation in Agricultural Automation

Susanto B. Sulistyo, Di Wu, Wai Lok Woo, *Senior Member*, S. S. Dlay, and Bin Gao, *Senior Member*

***Abstract*— This paper presents a novel computational intelligence vision sensing approach to estimate nutrient content in wheat leaves by analyzing color features of the leaves images captured on field with various lighting conditions. We propose the development of deep sparse extreme learning machines (DSELM) fusion and genetic algorithm (GA) to normalize plant images as well as to reduce color variability due to a variation of sunlight intensities. We also apply the DSELM in image segmentation to differentiate wheat leaves from a complex background. In this paper, four moments of color distribution of the leaves images (mean, variance, skewness, and kurtosis) are extracted and utilized as predictors in the nutrient estimation. We combine a number of DSELMs with committee machine and optimize them using the GA to estimate nitrogen content in wheat leaves. The results have shown the superiority of the proposed method in the term of quality and processing speed in all steps, i.e., color normalization, image segmentation, and nutrient prediction, as compared with other existing methods.**

***Note to Practitioners*—In order to support precision farming technology and agricultural automation, it is very essential to estimate the nutrient content in plants to prevent over fertilizing that will harm environment. Furthermore, the existing method to determine nutrient content in leaves is destructive as well as time consuming and requires special expertise to operate expensive devices. A number of methods have been developed to estimate nutrient content in leaves nondestructively based on the color features of the leaves. Most of these methods, however, are conducted in a controlled environment with artificial lighting. Such methods are not practical and need various equipment. We propose a low-cost, simple, and accurate technique to estimate nitrogen content in wheat leaves by analyzing RGB color of the leaves images. In this paper, we found that the developed DSELMs fusion has enabled better performance in normalizing images and is faster than other neural network types, i.e., backpropagation-based multilayer perceptron and original extreme learning machine. In image segmentation step, the established DSELM shows good performance to recognize and distinguish wheat leaves from other undesired background images. Furthermore, the developed weighted DSELMs have demonstrated enhanced ability in estimating nutrient content.**

Manuscript received December 27, 2016; revised July 1, 2017; accepted September 10, 2017. Date of publication December 21, 2017; date of current version July 2, 2018. This paper was recommended for publication by Associate Editor H. Qiao and Editor D. Tilbury upon evaluation of the reviewers' comments. *(Corresponding author: Wai Lok Woo.)*

The authors are with the School of Electrical and Electronics Engineering, Newcastle University, Newcastle upon Tyne NE1 7RU, U.K. (e-mail: s.b.sulistyo@ncl.ac.uk; d.wu3@ncl.ac.uk; w.l.woo@ncl.ac.uk; s.s.dlay@ncl.ac.uk; bin.gao@ncl.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TASE.2017.2770170

***Index Terms*— Agricultural automation, committee machines, computational intelligence, deep learning, deep neural networks, image processing.**

I. Introduction

NUTRIENT status analysis is one of the crucial stages in precision farming, which has become a topical issue in modern agriculture. Precision fertilizing requires farm machineries with capabilities to detect crop leaves from weeds as well as measure nutrient content in crops nondestructively and determine how much nutrient dose should be applied to plants [1]. Therefore, estimating nutrient content in plants is the key factor to determine the fertilizer dose. By defining the fertilizer dose more precisely, such practice will increase crop productivity as well as reduce production costs and prevent environment from excessive fertilization.

Nitrogen is one of the essential nutrients required by wheat crops in considerable amounts to ensure their growth since it has an important function in photosynthesis. In general, there are four common methods to assess nitrogen content in plants, as reported in [2], i.e., chemical and combustion tests, vegetation index (VI), chlorophyll meter (CM), and leaf color charts (LCC). Despite the high accuracy, the chemical and combustion tests are of destructive methods. The VI and CM methods are of nondestructive, yet they are expensive to obtain remote sensing images as well as the CM device. In spite of its simplicity, the LCC results are very subjective depending on the inspection of human visual.

Recently, due to the most recent developments in computer vision with rapidity and ease of image data acquisition, extensive studies on estimation of nitrogen status in crops have been conducted by using image-based analysis [3]. Nevertheless, most of these approaches are not practical and time consuming since they are typically performed in a controlled environment, such as in a closed image acquisition chamber with artificial lighting systems, as seen in Fig. 1 [4], [5].

It is a very challenging task to analyze nitrogen content in crops nondestructively according to the color features of plant leaves images which are captured in farm under sunlight. Characteristics of illuminant sources are very important in digital image processing systems, especially in image acquisition. Thus, the closed-chamber method cannot be applied to analyze nitrogen status of crops in open field as the sunlight intensity is continually changing. This issue, therefore, will cause inconsistent and unreliable image acquisition problems as well

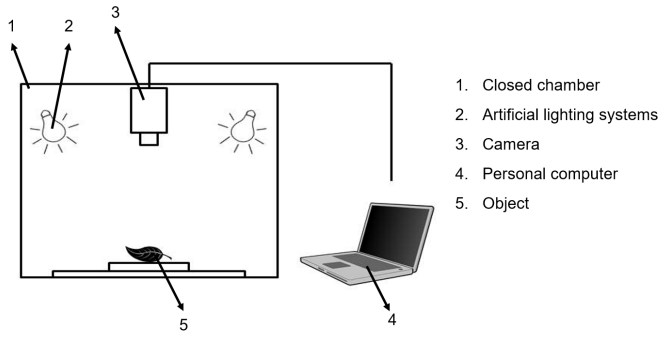


Fig. 1. Common image acquisition in a controlled environment with artificial lighting systems.

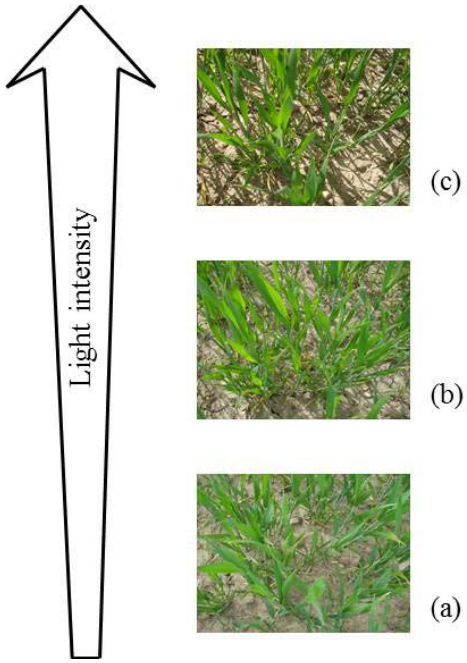


Fig. 2. Wheat crops images from the same field and the same fertilizer level captured on-field under various sunlight intensities. (a) Low light intensity. (b) Medium light intensity. (c) High light intensity.

as affect the results of nutrient estimation, which motivate the research focus of this paper. The problem faced in this research is how to automatically normalize images captured under various sunlight concentrations. The color normalization aims to reduce color variability among the images in order to resolve the image acquisition problem.

In this paper, we propose a low-cost and accurate nondestructive image-based technique to estimate nitrogen amount in wheat crops on field with various sunlight intensities using a conventional digital camera. As seen in Fig. 2, the intensity of the light source will dominantly affect the appearance of the wheat plants images even they are acquired from the same field with the same fertilizing level treatment. These images, therefore, should be normalized prior to subsequent image processing steps. The color variation of the wheat leaves after image normalization process is merely influenced by the difference of fertilizing concentrations. Generally, the lesser the nitrogen fertilizer applied to plants, the green color of the leaves will be much lighter. In this step, we introduce a novel method for color normalization by using fusion of deep sparse extreme learning machines (DSELMs) with genetic algorithm (GA) and 24-patch Macbeth color checker as the color reference. Extreme learning machine has been well-known as one of the recent successful approaches in machine learning with much faster training speed than the ordinary multilayer perceptron (MLP). The GA is one of the evolutionary algorithms which involves selection, crossover, and mutation operators to make its population more diverse and thus prevent the algorithm to be trapped in a local optimum. Theoretically, the diversity will increase the algorithm's speed to achieve global optimum since it will countenance the algorithm to explore the solution space faster. Macbeth color checker is usually used as a reference target for photographic and video production work as well as calibration process [6]. As it has various standard color patches, it is beneficial to use this chart as a color reference to normalize and correct plant images under various sunlight intensities. Studies on color normalization conducted by Min *et al.* [7] and Cheng *et al.* [8] utilized Macbeth color chart as color reference.

The next step is image segmentation and features extraction. In this step, we utilize DSELM to distinguish wheat leaves as the object of interest from undesired images, such as soil, weeds, dried leaves, stems and stones. After image segmentation, 12 statistical features from four types of moments of each RGB color channel, i.e., first moment (mean), second moment (variance), third moment (skewness), and fourth moment (kurtosis), are extracted from the segmented images as the nutrient estimation predictors. We propose the utilization of these statistical features as predictors since they indicate the color distributions of wheat leaves more significantly rather than a single color channel from a certain color model or combinations of some color channels.

In the nutrient estimation step, we combine several DSELMs with different hidden layer numbers by using committee machine and optimize the results with GA. The estimation results of the proposed approach are then compared with other common and existing nutrient estimation methods.

This paper proposes an innovative intelligent computational machine vision by using DSELMs fusion and GAs to estimate nitrogen status in wheat plants for agricultural automation. Furthermore, the novel contributions of our proposed method are summarized as follows.

1) A newly developed DSELM-based MLP is utilized to normalize the color constancy. Unlike other learning algorithms such as back-propagation-based MLP, DSELM is able to exploit more important information with much faster learning speed [9]–[11]. In the meanwhile, the utilized norm-1 (ℓ_1) regularization is able to generate more sparse representations than other MLP learning algorithm [9], [12], [13]. Thus, the problem of color variability due to various light intensities are able to be carried out by developing DSELMs fusion with 24-patch Macbeth color checker as the color reference, and its combination with GA for image normalization. Moreover, the developed GA is used to determine the correct output weights of each DSELM.
2) DSELM is also utilized in the image segmentation process to distinguish the desired target from noise

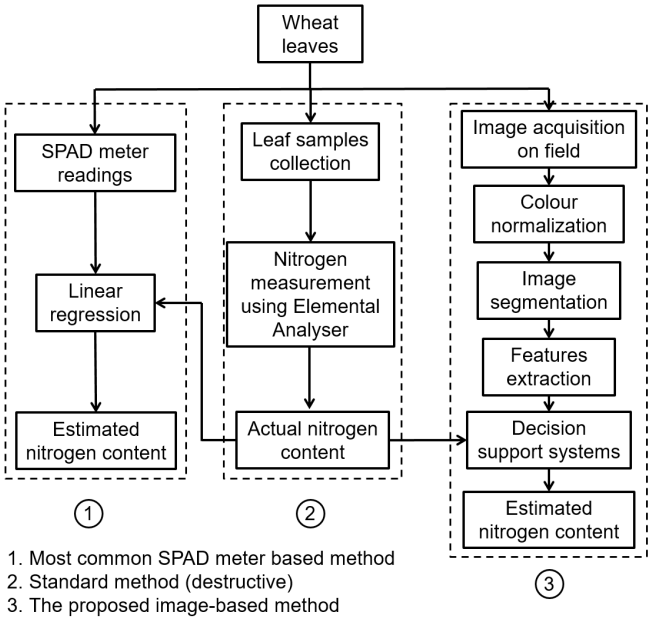


Fig. 3. Flowchart of the entire research to estimate nitrogen status.

images as DSELM is easy training and capable to extract more sparse and accurate features layer-by-layer with higher layers represent more abstract information than previous layers.

3) In the nitrogen estimation step, a committee machine is used to combine the output of DSELMs with various hidden layer numbers. In addition, the GA is applied to optimize the estimation results.

This paper is organized as follows. Experimental setup is explained in Section II. In Section III, the proposed DSELM and GA-based color normalization are presented. Section IV discusses the DSELM-based image segmentation and statistical color features extraction. Nitrogen content estimation using combined DSELM is proposed in Section V. Section VI analyzes and compares the results of the proposed work with other methods. Finally, we conclude our research in Section VII.

II. Experimental Setup and Data Acquisition

In general, this research can be divided into three major works as depicted in Fig. 3. Each work will be explained in more detail in Sections II-A–II-D.

A. Farm Experimental Materials and Design

An experiment to figure out the effect of fertilizer amount to crops growth has been conducted at Nafferton Experimental Farm, Newcastle University, from April to June. To produce variations in nitrogen levels, the treatments were set to three different nitrogen amounts, i.e., 0 (N1), 85 (N2), 170 (N3) kg/ha of NH_4NO_3 with each treatment replicated four times. Hence, there were 12 plots with each plot was (20×20) m^2 in dimension. The data collection was performed in three different periods, i.e., one week before fertilizing, two and four weeks after fertilizing. Therefore, in total 36 samples were used in this research.

B. Chlorophyll Meter Readings (Minolta SPAD-502)

The measurements of chlorophyll content using Minolta SPAD-502 were conducted on 30 leaf samples in each plot. The value showed on the SPAD meter screen signifies the chlorophyll content of the leaf, which is highly correlated with the nitrogen amount. To obtain the average value of SPAD of the plot, 30 leaves were randomly selected on one plot. Therefore, there were 12 SPAD values for 12 samples in one data collection time. In total, we have 36 SPAD values used for comparison.

C. Actual Nitrogen Measurement Using Elementar Analysis

The actual nitrogen amount was measured using an Elementar Vario Macro Cube. The procedure is as follows. First of all, around 50–60 wheat leaves per plot were taken as samples. These samples were then dried in a cabinet oven dryer at a temperature of 80 °C for about 48 h. The dried samples were ground afterward using an electric grinder at a speed of 14 000 r/min to pulverize the samples into powder. To measure the actual nitrogen amount, a sample of approximately 100 mg was weighed into a tin foil cup. The cup was then folded and squashed into a pellet to expel the air. This analysis involved the combustion method by burning the sample with a certain amount of oxygen in the cube. The nitrogen element was then analyzed automatically using special software and a percentage figure subsequently obtained.

D. On-Field Image Data Acquisition

Several sample images of wheat leaves were captured randomly at ten points on each plot. We obtained 360 images for the 12 set plots during three different periods. These image data collections were conducted in an open field with light intensity of the sun during this experiment ranged from around 7 to 80 Klux and the data collection time was from 10 am to 2 pm. This means that the color of the sunlight was relatively white, compared to its color in the morning or in the evening, when it is relatively reddish or yellowish. All images were grabbed from the top of the plants in a distance of 10–20 cm using a common digital camera (Sony DSC-W55). The images were captured at 1632×1224 pixels, but then down sampled to 448×336 pixels to assist with the effectiveness of the image processing.

III. Deep Sparse Extreme Learning Machine Fusion and Genetic Algorithm-Based Color Normalization

Since the changes of light intensity will affect the color of plant images, the images, therefore, should be normalized as if they are acquired under the same light intensity, in order to perform a consistent and reliable comparison among the images. In this research, we use a 24-patch Macbeth color checker with our proposed neural networks fusion-based color normalization approach. The Macbeth color checker is a square card that consists of 24 patches of color samples which represent natural objects, chromatic, primary, and grayscale colors, which are arranged in four rows (Fig. 4).

Fig. 4. Macbeth color checker.

Fig. 5. Photograph of digital light meter.

Neural networks have been extensively used for automation in many applications, such as iron and steel industry [14], financial service [15], plastic production [16], and energy generation prediction [17]. Neural networks are also used in digital signal processing [18], [19] and nonlinear systems control [20], [21]. However, the training procedure of neural networks with back-propagation algorithm is easy to fall into local minima and the training speed is generally slow, especially for MLP. To address this problem, we propose the DSELM algorithm to train the MLP. Compared with other machine learning algorithms, DSELM is much faster in training stage and has better generalization performance [9], [11].

Color normalization, also known as color constancy, is a computational technique to correct color deviations of images due to the differences of lighting conditions. Image colors of an object are considerably influenced by the direction and intensity of the light source, as well as illuminant color [23]. Moreover, an extensive research has been done to overcome the problem of color constancy [24]–[26]. In this paper, our color constancy concept differs from previous works given that the images are acquired under unconstrained daylight, as mentioned in Section II-D. This poses a more difficult challenge as plant images captured under various light intensities have to be normalized to a standard image that is captured under a standard light intensity.

Our proposed method of a combination of DSELMs fusion and GA for image normalization is described as follows. First of all, the images of the Macbeth color checker were captured under sunlight using a common digital still camera. The sunlight intensity, which was measured by using a digital light meter (Fig. 5), ranged between 7 to 82 Klux. The color checker images under light intensity of 50 ± 1 Klux were considered as the standard (target) images while the other images were used as the input of the developed DSELM. As for the target images, we captured five color checker images in a range of 49–51 Klux. In the meanwhile, 164 color checker images under light intensity of 7–48 Klux and 52–82 Klux were grabbed to be utilized as input images.

To obtain average color values of each patch, each color checker image was first cropped on each patch with a cropping window of 95×72 pixels. The average RGB color values of each patch of the cropped target and input images were then calculated. This step was done twice and thus we have $328(= 164 \times 2)$ RGB color training samples for each patch. Since there are 24 patches in the Macbeth color checker, therefore, there are 24 data sets in which each data set contains 328 RGB training samples. All the data sets are consequently combined to produce one large data set for DSELM fusion. This new data set, thus, consists of $7872(= 328 \times 24)$ RGB color samples.

To achieve a high generalization performance for color constancy, especially for large scales, appropriate representations are crucial. Single-layer feedforward neural network (SLFN) learning with back-propagation algorithm is an efficient method to learn compact features. However, the information learned from SLFN is not good enough to represent the input data, especially for large data set. Thus, multiple layer architecture of neural network is needed as the features extracted by multiple layer networks represent more abstract and accurate information than those from shallow ones. One of the most popular approaches is deep belief networks (DBNs) which can be done using stacked restricted Boltzmann machines. It has been shown to yield good performance in various areas; however, the training speed is generally very slow for the reason that all the parameters of the entire network need to be fine-tuned multiple times to achieve the criterions. Thus, the training of DBNs is too cumbersome and time consuming. To address this problem, we propose the DSELM. Compared with DBNs, DSELM has four notably attractive features. First, the hidden nodes can be randomly generated according to any continuous probability distribution without any prior knowledge, e.g., the uniform distribution [22]. Second, the only parameter that needs to be determined is the output weight which can be established by searching the path back from a random space. Third, once the feature of the previous hidden layer is learned, the parameters of current hidden layer will be fixed and need not to be fine-tuned [9], this is the major difference between DSELM and DBNs. Finally, more abstract and sparse hidden information is extracted using ℓ_1-regularization. Additionally, extreme learning machine is usually efficient, especially for small data set applications, as found in [47].

The DSELM consists of two phases: 1) unsupervised feature mapping and 2) supervised feature regression. In the former phase, an ELM-based sparse autoencoder is used to extract sparse features of the input layer by layer with higher layers represent more abstract and accurate features than that of previous layer. In the latter phase, an original ELM is stacked at the top of the learned deep structure to make the final decision. In the following, the overall architecture of DSELM will be introduced in detail and the description of the training procedure of Sparse ELM (SELM) is also presented as the

SELM is the building block which is used to construct the deep structure of DSELM.

In the initial ELM, given a set of N training data $(\boldsymbol{X}, \boldsymbol{T}) = \{\boldsymbol{x}_j, \boldsymbol{t}_j\}_{j=1}^{N}$, where $\boldsymbol{x}_j = [x_{j1}, x_{j2}, \ldots, x_{jP}] \in \mathbb{R}^P$ and $\boldsymbol{t}_j = [t_{j1}, t_{j2}, \ldots, t_{jQ}] \in \mathbb{R}^Q$ are the training data and the corresponding target, respectively. The parameters P and Q are the dimension of input and target vector, respectively. The output function $f(\boldsymbol{X})$ of ELM with K hidden nodes fully connect the input data to the outputs is represented by

$$f(\boldsymbol{x}_j) = \sum_{k=1}^{K} \psi_k\left(w_k^T \boldsymbol{x}_j\right) \cdot \beta_k, \quad j = 1, 2, \ldots N \tag{1}$$

where $\psi(\cdot)$ is the activation function which we used is the sigmoid function

$$\psi\left(w_k^T \boldsymbol{x}_j\right) = \frac{1}{1 + e^{-w_k^T \boldsymbol{x}_j}}. \tag{2}$$

$w_k \in \mathbb{R}^P$ is the randomly generated parameters connecting input layer and kth hidden node, $\beta_k \in \mathbb{R}^Q$ is the output weight vector connecting the kth hidden layer and the output layer.

The N equations in (1) can be written compactly as

$$F(\boldsymbol{X}) = \boldsymbol{\Psi}\boldsymbol{\beta} \tag{3}$$

where $\boldsymbol{\beta} = \{\beta_k\} \in \mathbb{R}^{K \times Q}$ is the output weight matrix, $\boldsymbol{\Psi}$ is the $N \times K$ hidden feature mapping matrix with respect to input $\boldsymbol{X}$. The elements of $\boldsymbol{\Psi}$ can be described as follows:

$$\begin{aligned} \boldsymbol{\Psi} &= \left[\psi_1\left(w_1^T \boldsymbol{X}\right), \psi_2\left(w_2^T \boldsymbol{X}\right), \ldots, \psi_K\left(w_K^T \boldsymbol{X}\right)\right] \\ &= \begin{bmatrix} \psi_1\left(w_1^T x_1\right) & \cdots & \psi_K\left(w_K^T x_1\right) \\ \vdots & \ddots & \vdots \\ \psi_1\left(w_1^T x_N\right) & \cdots & \psi_K\left(w_K^T x_N\right) \end{bmatrix}_{N \times K}. \end{aligned} \tag{4}$$

An ELM learns the parameters in two sequential stages: 1) random feature mapping and 2) linear parameter solving. In the first stage, with randomly initialized parameters, the input data are projected into an ELM feature space using the activation function $\psi(\cdot)$. Huang *et al.* [27] have proved that ELM is able to approximate any continuous function with randomly initialized parameters. Therefore, the only parameter that needs to be calculated is output weight $\boldsymbol{\beta}$. In the second stage, an ELM aims to reach the smallest training error and the smallest norm of output weights using the following equation to optimize the output weight $\boldsymbol{\beta}$:

$$\begin{aligned} &\underset{\boldsymbol{\beta} \in \mathbb{R}^{K \times Q}}{\operatorname{argmin}} \ \frac{1}{2}\|\boldsymbol{\beta}\|_{1_2}^2 + \frac{C}{2}\sum_{j=1}^{N}\|\boldsymbol{e}_j\|^2 \\ &\text{s.t. } \boldsymbol{\Psi}(x_j)\boldsymbol{\beta} = \boldsymbol{t}_j - \boldsymbol{e}_j, \quad j = 1, 2, \ldots, N \end{aligned} \tag{5}$$

where the first term is ℓ_2 optimization to avoid over fitting and obtain compact hidden information, C is a tradeoff coefficient which are chosen experimentally, $\boldsymbol{e}_j = (\boldsymbol{t}_j - f(\boldsymbol{x}_j)) \in \mathbb{R}^Q$ is the error vector with respect to jth input data. Equation (5) can be rewritten as an unconstrained optimization problem

$$\underset{\boldsymbol{\beta} \in \mathbb{R}^{K \times Q}}{\operatorname{argmin}} \ \frac{1}{2}\|\boldsymbol{\beta}\|_{1_2}^2 + \frac{C}{2}\|\boldsymbol{\Psi}\boldsymbol{\beta} - \boldsymbol{T}\|^2 \tag{6}$$

where $\boldsymbol{T} = [\boldsymbol{t}_1, \boldsymbol{t}_2, \ldots, \boldsymbol{t}_N] \in \mathbb{R}^{N \times Q}$ is the target. As ELM autoencoder is designed to encoded outputs to approximate the original inputs by minimizing the reconstruction errors, therefore, the target $\boldsymbol{T}$ is set to original inputs $\boldsymbol{X}$. In the meanwhile, due to the use of ℓ_2 regularization, the extracted features of ELM autoencoder may have redundancy and not sparse enough to represent the input data, therefore, a more sparse solution is needed. Compared with ℓ_2 norm, ℓ_1 regularization is able to induce more sparsity in the optimal solution of (6). Also, ℓ_1 regularization is less sensitive to output. Therefore, ℓ_1 penalty is applied. Equation (6) can be rewritten as

$$\underset{\boldsymbol{\beta} \in \mathbb{R}^{K \times Q}}{\operatorname{argmin}} \ \|\boldsymbol{\beta}\|_{1_1} + \|\boldsymbol{\Psi}\boldsymbol{\beta} - \boldsymbol{X}\|^2 \tag{7}$$

where $\|\boldsymbol{\beta}\|_{\ell_1}$ stands for the sum of the absolute values of the components of $\boldsymbol{\beta}$. Equation (7) can also be adapted to semisupervised and unsupervised learning [48] and solved by a gradient projection algorithm. Of which the most popular methods is the class of *Iterative Shrinkage-Thresholding Algorithm (ISTA)*, where each iteration involves matrix-vector multiplication involving $\boldsymbol{\Psi}$ and $\boldsymbol{\Psi}^T$ followed by a shrinkage/soft-threshold step [12], [13]. The general step of *ISTA* is

$$\boldsymbol{\beta}_{i+1} = \Upsilon_t(\boldsymbol{\beta}_i - 2t\boldsymbol{\Psi}^T(\boldsymbol{\Psi}\boldsymbol{\beta}_i - \boldsymbol{X})) \tag{8}$$

where t is an appropriate step size and $\Upsilon_\alpha : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is the shrinkage operator defined by

$$\Upsilon_\alpha(\boldsymbol{\beta}_i) = (|\boldsymbol{\beta}_i| - \alpha)_+ \operatorname{sgn}(\boldsymbol{\beta}_i). \tag{9}$$

One of the simplest methods for solving an unconstrained minimization problem $\min\{p(x) : x \in \mathbb{R}^n, p : \mathbb{R}^n \longrightarrow \mathbb{R}^n\}$ is the gradient algorithm which generates a sequence $\{x_i\}$ via

$$x_0 \in \mathbb{R}^n, \quad x_i = x_{i-1} - t_i \nabla p(x_{i-1}) \tag{10}$$

where t_i is an appropriate step size. Equation (10) can be viewed as a proximal regularization of the linearized function p at x_{i-1}, thus, (10) can be rewritten as

$$\begin{aligned} x_i = \underset{x}{\operatorname{argmin}} \Big\{ & p(x_{i-1}) + \langle x - x_{i-1}, \nabla p(x_{1-1})\rangle \\ & + \frac{1}{2t_i}\|x - x_{i-1}\|^2 \Big\}. \end{aligned} \tag{11}$$

Rewriting (7) as

$$\min\{p(\boldsymbol{\beta}) + q(\boldsymbol{\beta}), \boldsymbol{\beta} \in \mathbb{R}^{K \times Q}\} \tag{12}$$

where $p(\boldsymbol{\beta}) = \|\boldsymbol{X} - \boldsymbol{\Psi}\boldsymbol{\beta}\|^2$ and $q(\boldsymbol{\beta}) = \|\boldsymbol{\beta}\|_{\ell_1}$. Therefore, $\boldsymbol{\beta}$ can be calculated as

$$\begin{aligned} \boldsymbol{\beta}_i = \underset{\boldsymbol{\beta}}{\operatorname{argmin}} \Big\{ & p\left(\boldsymbol{\beta}_{i-1}\right) + \langle \boldsymbol{\beta} - \boldsymbol{\beta}_{i-1}, \nabla p(\boldsymbol{\beta}_{1-1})\rangle \\ & + \frac{1}{2t_i}\left\|\boldsymbol{\beta} - \boldsymbol{\beta}_{i-1}\right\|^2 + \|\boldsymbol{\beta}\|_{1_1} \Big\}. \end{aligned} \tag{13}$$

The constant terms can be ignored, thus, (13) can be written as

$$\boldsymbol{\beta}_i = \underset{\boldsymbol{\beta}}{\operatorname{argmin}} \left\{ \frac{1}{2t_i}\|\boldsymbol{\beta} - (\boldsymbol{\beta}_{i-1} - t_i \nabla p(\boldsymbol{\beta}_{i-1})\|^2 + \|\boldsymbol{\beta}\|_{1_1} \right\}. \tag{14}$$

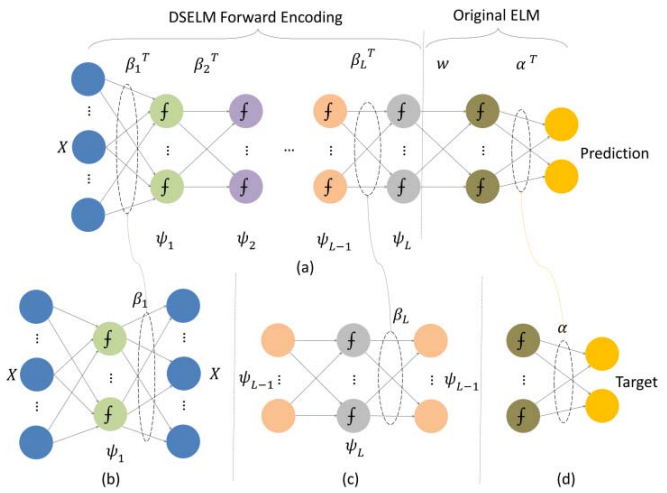


Fig. 6. (a) Layer wise training of DSELM, which is consists of two phases: deep forward learning followed by the original ELM classification. (b) Implementation of first hidden layer SELM autoencoder. (c) Training procedure of Lth hidden layer SELM autoencoder. (d) Analytically calculate the output weights of original ELM with labeled target using randomly initialized parameters.

Let $\nabla p = 2\boldsymbol{\Psi}^T(\boldsymbol{\Psi}\boldsymbol{\beta} - \boldsymbol{X})$ denotes the gradient of p and $L := L(p) = 2(\boldsymbol{\Psi}^T\boldsymbol{\Psi})$ is the Lipschitz constant of ∇p. Define operator $\varphi_L : \mathbb{R}^n \longrightarrow \mathbb{R}^n$, $\varphi_L(\boldsymbol{\beta}) = \Upsilon_t(\boldsymbol{\beta}_i - 2t\boldsymbol{\Psi}^T(\boldsymbol{\Psi}\boldsymbol{\beta}_i - \boldsymbol{X}))$, $t = 1/L(p)$. The computation of output weight $\boldsymbol{\beta}$ using ISTA algorithm with constant step size can be represented as follows:

$$\boldsymbol{\beta}_i = \varphi_L(\boldsymbol{\beta}_{i-1}). \tag{15}$$

However, ISTA algorithm shares a sublinear global rate of convergence $O(1/z)$, where z is the iteration times and it appears to be a slow method. To improve the complexity result, a *Fast ISTA (FISTA)* is used [12]. FISTA keeps the simplicity of ISTA but shares complexity of $O(1/z^2)$ for minimizing smooth convex problems. It computes $\boldsymbol{\beta}_i$ using based on the following:

$$\boldsymbol{\beta}_i = \varphi_L(y_i) \tag{16}$$

where y_i is a new point which is a specific linear combination of previous two points $\{\boldsymbol{\beta}_{i-1}, \boldsymbol{\beta}_{i-2}\}$.

As the SELM is the building block of DSELM, the learned output $\boldsymbol{\beta}$ with respect to the input data is the first-layer weight of DSELM. After the output of the first hidden layer sparse representations are obtained, a new SELM autoencoder is stacked at the top to learn the second layer parameters with the same procedure. In this manner, all parameters of the DSELM can be computed sequentially and all parameters can be fixed without iteratively fine-tuning. At last, an original ELM classifier is stacked at the top of the deep network to make the final decision.

The steps taken in the DSELM, as seen in Fig. 6, can be described as follows.

1) *Normalize Inputs (X_i):* RGB input colors should be normalized by dividing their values with the maximum value 255. Thus, $X_1 = (\mathrm{R}/255)$, $X_2 = (\mathrm{G}/255)$, $X_3 = (\mathrm{B}/255)$; $X_i \in [0, 1]$, $i = 1, 2, 3$.
2) *Randomly Initialize the Input Weight $\boldsymbol{w}$:* Set weight connecting input layers and hidden layer small random values (between -1 and 1).
3) *Calculate the Hidden Output $\boldsymbol{H}$:* With randomly initialized weights, the kth hidden node output h_k can be calculated as:
$$h_k = \psi_k\left(w_k^T \cdot \boldsymbol{X}\right)$$
where $\boldsymbol{X}$ is input data. The activation function that we use is sigmoid function.
4) Calculate the Lipschitz constant L of the gradient of smooth convex function ∇p
$$L := L(p) = 2(\boldsymbol{\Psi}^T\boldsymbol{\Psi}). \tag{17}$$
5) Calculate the output weights $\boldsymbol{\beta}_i$ using SELM.
   a) Take the initial value $y_1 = \boldsymbol{\beta}_0 \in \mathbb{R}^n$, $t_1 = 1$.
   b) for $i \geq 1$, compute
$$\begin{aligned}\boldsymbol{\beta}_i &= \varphi_L(y_i)\\ &= \Upsilon_t(y_i - t\nabla p(y_i))\\ &= \underset{\boldsymbol{\beta}}{\operatorname{argmin}}\left\{\frac{L}{2}\left\|\boldsymbol{\beta} - \left(\boldsymbol{\beta}_{i-1} - \frac{1}{L}\nabla p(\boldsymbol{\beta}_{i-1})\right)\right\|^2 + \|\boldsymbol{\beta}\|_{1_1}\right\}\\ t_{i+1} &= \frac{1+\sqrt{1+4t_i^2}}{2}\\ y_{i+1} &= \boldsymbol{\beta}_i + \left(\frac{t_i - 1}{t_{i-1}}\right)(\boldsymbol{\beta}_i - \boldsymbol{\beta}_{i-1}).\end{aligned}$$
6) Recompute the hidden output $\boldsymbol{H}$ with the learned output weight $\boldsymbol{\beta}$ instead of the randomly initialized weight $\boldsymbol{w}$
$$\boldsymbol{H} = \sigma\ (\boldsymbol{\beta}^T \cdot \boldsymbol{X}).$$
7) Stack a new SELM at the top with the input is the previous hidden layer output $\boldsymbol{H}$.
8) Repeat the above processes (2–7) until the desired deep structure is achieved.
9) Stack an original ELM which is trained with labeled data at the top of the learned deep structure to make final decisions.

The next step is combining all the extreme learning machines into one neural network system. The proposed DSELMs fusion, as shown in Fig. 7, is developed to generate new RGB outputs. The proposed approach differs from the method in [28] by altering the conventional backpropagation neural network with DSELM. The final output RGB values from the networks fusion is obtained as follows:

$$\begin{aligned}\mathbf{Z} &= \alpha \cdot \boldsymbol{O}\\ &= [\propto_1, \propto_2, \propto_3, \ldots, \propto_{24}] \cdot [\boldsymbol{O}_1, \boldsymbol{O}_2, \boldsymbol{O}_3, \ldots, \boldsymbol{O}_{24}]^T\end{aligned} \tag{18}$$

where α is the weight matrix of each network output, $\mathbf{O}$ is the output matrix of each network, and $\mathbf{Z}$ is the final output matrix of the neural networks fusion.

From (18), it can be seen that matrix α consists of 24 diagonal α matrices with dimension of 3×3. Similar to matrix α, matrix $\mathbf{O}$ also has 24 matrices with a dimension of

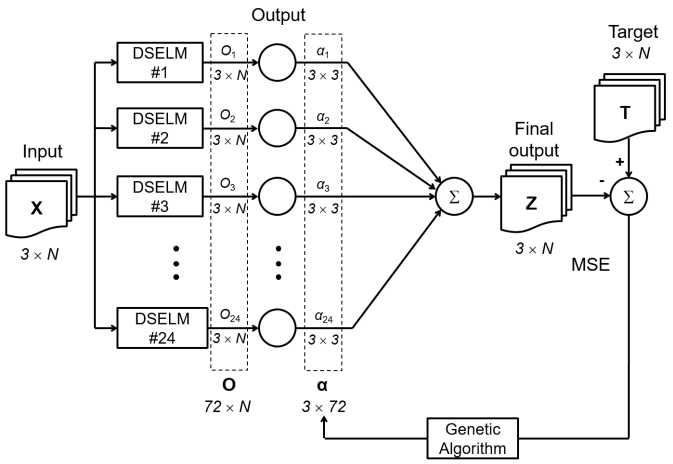


Fig. 7. Proposed DSELMs fusion using Macbeth color checker and GA-based optimization for image color normalization.

$3 \times N$ for each matrix $\boldsymbol{O}_i$, and N is the number of training samples.

In this paper, the GA is utilized to find the optimum value for each of the 24 α matrices (see Fig. 7). The GA is an algorithm based on the Darwin principle of evolution as well as natural selection and biological systems. It has been widely used for optimization in applications involving patient flow distribution [29], thin film transistor-liquid crystal display module assembly scheduling [30], parallel robotic mechanisms for ankle rehabilitation treatments [31], and semiconductor wafer fabrications scheduling [32]. Basically, GA encompasses a population with a certain number of individuals. Each individual in a population has the possibility of being the solution to the optimization problem. Hence, by applying crossing over and mutation among individuals, a new generation is produced. This process is repeated several times until a new individual provides the most appropriate solution for the problem.

Based on our experiments, the proposed DSELMs fusion can be optimized by means of GA with the following steps.

1) *Define Fitness Function:* The fitness function of the color normalization is to minimize the mean square error (mse) between the target RGB colors of the Macbeth color checker (**T**) and the final output RGB colors of the DSELM fusion (**Z**) (see Fig. 7)

$$\underset{\alpha}{\operatorname{argmin}} \frac{1}{N} \sum_{n} (T_n - Z_n)^2$$

where N is the number of RGB color samples in the data set.

2) *Determine the Initial Population:* In this step, we set the initial population size ($S_{\text{pop}}^{(1)}$) of the developed GA to 500 individuals. These individuals performed as the first generation

$$S_{\text{pop}}^{(1)} = 500.$$

3) *Encoding and Decoding:* Encoding is used to express each individual in the population by binary strings of 0s and 1s. The $\alpha = [\propto_1, \propto_2, \propto_3, \ldots, \propto_{24}]$ matrix has a dimension of 3×72 and each element of the matrix $\propto_k$ is expressed by a 8-bit string of binary numbers (0s and 1s). Thus, each individual of matrix α has 1728 ($= 8 \times 3 \times 72$) bits length.
Decoding is used to change the encoding binary strings into decimal number. The decoded numbers are then computed into the defined fitness function.

4) *Set Boundary Conditions:* We set boundary conditions of each element in the matrix α to be always positive in order to prevent RGB colors of the final output (**Z**) have negative values. In particular, the boundary of each element of matrix $\propto_k$, i.e., $a_{k,ij}$ with $i, j = 1, 2, 3$, is set as follows:

$$\text{if } i = j \text{ then } a_{k,ij} \in [0, 1]$$
$$\text{else } a_{k,ij} = 0.$$

Thus, each matrix $\propto_k$ has a structure as follows:

$$\propto_k = \begin{bmatrix} a_{k,11} & 0 & 0 \\ 0 & a_{k,22} & 0 \\ 0 & 0 & a_{k,33} \end{bmatrix}.$$

5) *Reproduce:* Reproduce next generations ($S^{(2)}, S^{(3)}, S^{(4)}, \ldots$) by running the selection, crossover and mutation operators.
According to [33], selection operator attempts to give "a pressure" to the population as the same as natural selection in the biological life. Individuals with better performance, or fitter, will be kept to the next generations. Otherwise, they will be wiped out. In crossover, two individuals exchange some bits of the same section one another to create two offspring, while mutation turns over bits in a chromosome (a 0 to a 1 and *vice versa*). Furthermore, the existence of mutation depends on the mutation rate or probability of mutation (ρ) set in the algorithm and a random number given by the computer (ω). In this step, the ρ value was set to 0.005. According to [34], probability of mutation normally ranges from 0.001 to 0.01, as found in [35] and [36] in which used ρ values of 0.003 and 0.01, respectively. In addition, Coley [33] suggested that small mutation rate will be less disastrous than high one in most common problems.
The mutation operator, thus, can be defined as follows:

$$\text{mutation} = \begin{cases} 1\,(\text{mutates}), & \text{if } \omega \le \rho \\ 0\,(\text{not mutate}), & \text{if } \omega > \rho. \end{cases}$$

6) *Repeat:* Repeat the selection, crossover, and mutation processes until the best individual achieved, which has mse less than 0.0001.

Once the optimum value of matrix α is achieved, the next step is applying the developed neural networks fusion and matrix α to adjust the RGB color of the wheat plants. In this paper, a wheat plant image has a dimension of 448×336 pixels. Through this developed color adjusting system, each pixel of a plant image acquired under various light intensities is transformed to the equivalent pixel of the image under the standard light intensity, i.e., 50 Klux.

IV. DSELM-Based Image Segmentation and Statistical Color Feature Extraction

In a machine vision system, image segmentation plays an important role since it aims to classify each pixel in an image into a certain part. In this research every pixel will be divided into two segments, i.e., object (wheat leaves) and background (nonleaf images). Commonly, in a controlled environment (see Fig. 1), an object is laid down on a white paper background with certain illuminations. In this method, the object of the captured image can easily be distinguished from its background by applying the conventional Otsu algorithm with a simple threshold value. In this research, however, the problem is more complicated. The images of the wheat leaves are captured directly on field and contain leaves as the object and some unwanted parts, such as soil, stones, weeds, and dried and semidried leaves in the background. Many of the unwanted parts (especially the weeds and semidried leaves) have a similar color to the wheat plant.

A DSELM was used for image segmentation to remove the nonleaf images and keep the wheat leaves as the region of interest. Basically, the developed DSELM for this step is similar to that for the color normalization process as explained in Section III. The network has three hidden layers with three nodes of input layer, which indicate red, green, and blue values (RGB) for each pixel of plant images. However, the output layer has only one unit, which signifies whether each pixel is a part of a leaf or not. The output value of the network is equal to 1 if the corresponding pixel is a part of a leaf, otherwise the value is 0.

In this image segmentation step, we develop a data set from 7200 samples of RGB color and binary values (0 or 1) as the input and target values, respectively. The data set was acquired from 24 normalized images. On each image, 150 pixels in the leaf region and 150 pixels in other parts of the region were selected manually. The RGB color values of the selected pixels were then used as inputs of the developed DSELM.

After image segmentation, some small noises will be presented in the color segmented image. These noises should be removed prior to features extraction. In the majority of images, weeds are also present which need to be eliminated from the segmented images, as they can influence the color information of the wheat leaves. To overcome this obstacle, we use the largest part of the leaves which has the highest number of object pixels. This algorithm can be seen in Fig. 8. Some examples of the proposed color normalization and image segmentation results can be seen in Fig. 9. As revealed, the proposed color constancy method for image normalization and the DSELMs-based image segmentation can be used in an automated manner to normalize the images of the plants and to remove the unwanted parts from the image, as indicated by the black circles.

In the features extraction step, several statistical color features pertaining to the final color segmented images are calculated. These features are used for nutrient estimation in the next step. Four moments of statistical data distribution are used in this research, i.e., mean (first moment), variance (second moment), skewness (third moment) and kurtosis (fourth moment). Thus, there are 12 statistical features for

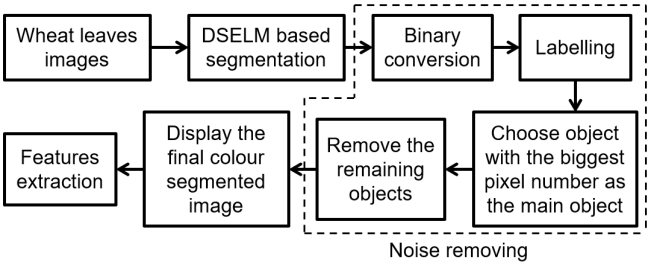


Fig. 8. Image segmentation algorithm.

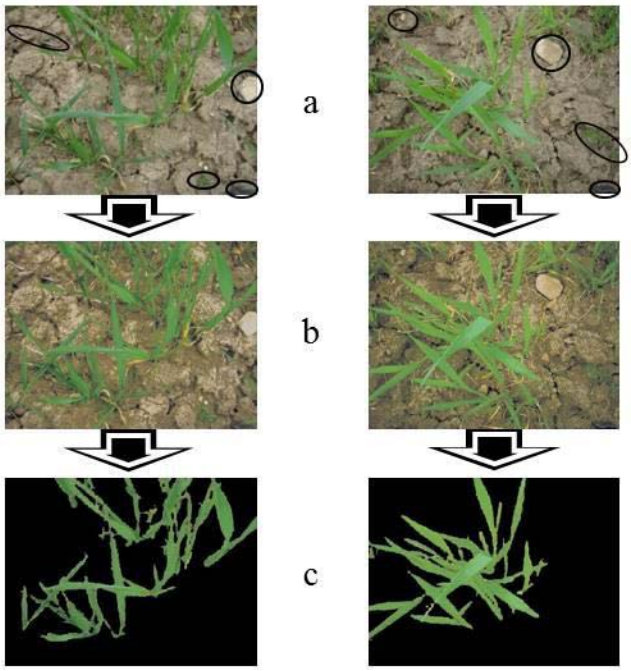


Fig. 9. Two examples of the proposed color normalization and image segmentation results. (a) Original images with some spots of unwanted parts. (b) Normalized images. (c) Segmented images.

three color channels (red, green and blue) which represent the color distributions related to the segmented images. According to [37], the first order (mean), the second order (variance), and the third order (skewness) of color moments have been demonstrated efficiently and effectively to represent color distribution of images.

In addition, in most natural scene images such as in the proposed research, it is found that the pixel intensity follows a non-Gaussian distribution. Non-Gaussian distributions are typically characterized with high order moments. The first- and second-order moments (mean and variance) measures the mean and average power pixel intensity of the image, respectively. The third-order moment (i.e., skewness) enables the neural network to distinguish mirror images of otherwise identical images while the fourth-order moment (i.e., kurtosis) captures the features derived from the pixel intensities whose value has large dynamic range. In addition, the fourth-order moment can identify outliers in the images which contain information about the edges of the wheat leaves.

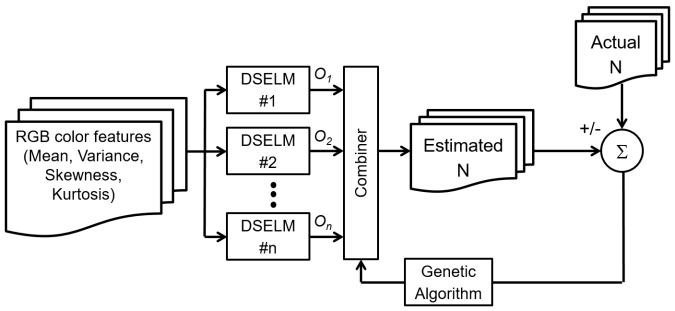


Fig. 10. Combination of several DSELM with committee machine and GA-based optimization for nitrogen content estimation.

The statistical color features can be achieved by using the following formulas:

$$\text{mean} = \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{19}$$

$$\text{variance} = \sigma^2 = \frac{1}{n}\sum_{i=1}^{n} (y_i - \bar{y})^2 \tag{20}$$

$$\text{skewness} = \text{skew} = \frac{\frac{1}{n}\sum_{i=1}^{n} (y_i - \bar{y})^3}{\sigma^3} \tag{21}$$

$$\text{kurtosis} = \text{kurt} = \frac{\frac{1}{n}\sum_{i=1}^{n} (y_i - \bar{y})^4}{\sigma^4} \tag{22}$$

where y refers to each color channel (red, green, and blue), n is the number of object pixels, and σ is the standard deviation.

V. Nitrogen Content Estimation Using Weighted DSELMs

In this section, we will explain further the final step of our proposed method, i.e., nitrogen estimation. Several DSELMs are utilized to perform nutrient estimation by combining them into a committee machine, as seen in Fig. 10. We also employ GA to optimize the estimation results. In this paper, we limit our experiments up to five different DSELMs with five hidden layer numbers. Each network is repeated 100 times to eliminate the effect of the random bias numbers and initial weights. Subsequently, a number of neural networks which provided the minimum mse are chosen to achieve the final estimation.

Committee machine can yield significant improvements in the estimation since it can minimize the effect of a random component due to data noise in the generalization performance of a single network [38]. Basically, the concept of a committee machine is to combine outputs of several expert systems with the same input data, with the aim of producing a new output. In this paper, we use ensemble averaging as the combination method. Committee machine with ensemble averaging produces the same if not smaller amount of error as compared to single neural network. This can be shown as follows. Suppose that there are P neural network systems to approximate a target vector T. Each neural network has output vector O_i and error e_i

$$O_i = T + e_i. \tag{23}$$

Thus, the sum of the squared error for the ith neural network y_i is

$$E_i = \xi[(O_i - T)^2] = \xi\left[e_i^2\right] \tag{24}$$

where $\xi[\cdot]$ denotes the statistical expectation. The average error of each neural network (E_{ave}) is then

$$E_{\text{ave}} = \frac{1}{P}\sum_{i=1}^{P} E_i = \frac{1}{P}\sum_{i=1}^{P} \xi\left[e_i^2\right]. \tag{25}$$

In other words, by means of committee machine, the output value Y can be achieved by simply averaging the output vector O_i

$$Y = \frac{1}{P}\sum_{i=1}^{P} O_i. \tag{26}$$

Thus, the squared error of the committee machine (E_{COM}) is

$$\begin{aligned} E_{\text{COM}} &= \xi[(Y - T)^2] \\ &= \xi\left[\left(\frac{1}{P}\sum_{i=1}^{P} O_i - T\right)^2\right] = \xi\left[\left(\frac{1}{P}\sum_{i=1}^{P} e_i\right)^2\right]. \end{aligned} \tag{27}$$

But since

$$\xi\left[\left(\frac{1}{P}\sum_{i=1}^{P} e_i\right)^2\right] \le \frac{1}{P}\sum_{i=1}^{P} \xi\left[e_i^2\right].$$

Hence, we have

$$E_{\text{COM}} \le E_{\text{ave}}. \tag{28}$$

Thus, the calculated error of the committee machine is always smaller than if not equal to that of the single neural network.

In this paper, we perform ensemble averaging using the DSELMs combiner to achieve more enhanced generalization and system performance. The estimated nitrogen amount (N_e) of wheat leaves is calculated by using committee machine with simple averaging method as the combiner of P DSELMs as

$$N_{e_{\text{ave}}} = \frac{1}{P}\sum_{p=1}^{P} O_p. \tag{29}$$

The simple averaging method as expressed in (28) indicates that each DSELM has the same weight to produce a new output. Concurrently, we also investigate the possibility that each DSELM has a different weight. We apply a weighted averaging method according to

$$N_{e_{\text{weigh}}} = \sum_{p=1}^{P} (w_p \times O_p) \quad \text{s.t.} \sum_{p=1}^{P} w_p = 1 \tag{30}$$

where w_p and O_p are the weight and the output of pth single network, respectively.

A GA is utilized to discover the optimum value of the weights in the developed committee learning. Basically, the procedures of the GA developed in this step are the same as those in the color normalization. The fitness function used in this optimization is to minimize mse between the actual (N_a)

and the estimated nitrogen content (N_e) of n samples (see Fig. 10), which can be expressed as follows:

$$\underset{\boldsymbol{w}}{\operatorname{argmin}} \ \frac{1}{n}\sum_{i} \ (N_{a_i} - N_{e_i})^2.$$

In addition, the initial population size $S_{\text{pop}}^{(1)}$ and the mutation rate are set to 500 individuals and 0.006, respectively. Each individual is expressed by 16 ($= 2 \times 8$) b length of binary numbers (0s and 1s).

The level of the estimation accuracy is measured by calculating the error value of the actual and estimated nitrogen content. In this research, we use the mean absolute percentage error (MAPE) for the performance assessment. The less the error is, the superior the prediction is. For a comparison, several types of error are also measured, i.e., mean absolute error (MAE), *mse*, root-mean-square error (RMSE), and sum of squared error (SSE). The error types used in this research can be expressed as follows:

$$\text{MAPE} = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{N_{a_i} - N_{e_i}}{N_{a_i}}\right| \tag{31}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|N_{a_i} - N_{e_i}| \tag{32}$$

$$\text{mse} = \frac{1}{n}\sum_{i=1}^{n}(N_{a_i} - N_{e_i})^2 \tag{33}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(N_{a_i} - N_{e_i})^2} \tag{34}$$

$$\text{SSE} = \sum_{i=1}^{n}(N_{a_i} - N_{e_i})^2 \tag{35}$$

where n is the number of samples and N_a and N_e are the actual and the estimated [using either simple (28) or weighted average (29)] nitrogen content, respectively.

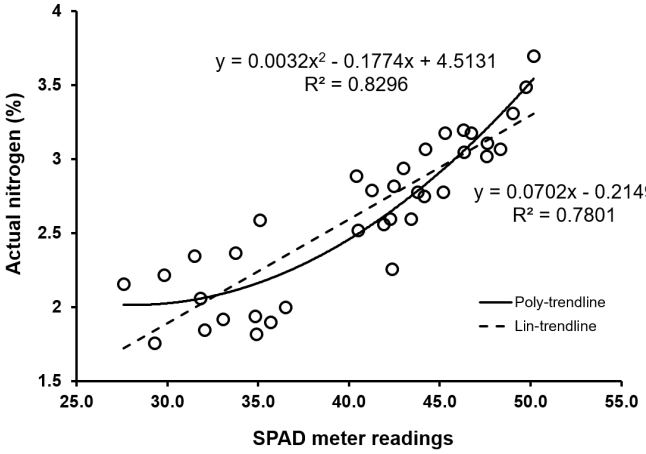


Fig. 11. Relationship between the SPAD meter value and actual nitrogen content.

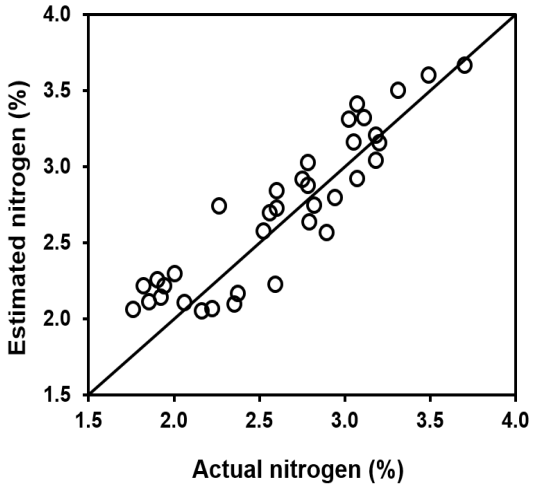


Fig. 12. Fitting plot of the actual and nonlinearly regressed estimated nitrogen content of SPAD meter-based analysis.

VI. Results and Discussion

A. Chlorophyll Meter-Based Nitrogen Content Estimation

The CM Minolta SPAD-502 was widely used to determine the chlorophyll content in leaves by measuring the absorbance of the leaf in two wavelength regions, i.e., red and infrared. After the signal processing steps, the absorbance was displayed in a units range from −9.9 to 199. Moreover, the chlorophyll amount was highly correlated with the nitrogen content. In many research, SPAD meter value is linearly correlated with nitrogen content; however, our paper found that the coefficient of determination (R^2) between them using nonlinear regression is slightly bigger than that using linear regression.

Based on the experiments conducted with 36 samples of wheat leaves, the R^2 value of SPAD and nitrogen content is 0.78 using linear regression and 0.83 using nonlinear regression, as seen in Fig. 11. It means that the relationship between the SPAD and nitrogen amount is fairly strong. The estimated nitrogen amount can then be calculated by using the regression equation. The error estimation of this method is 8.30%, as seen in Fig. 12. Similar research with relatively close relationships between the SPAD and nitrogen have been reported in rice ($R^2 = 0.89$) [39] and oilseed rape ($R^2 = 0.74$) [40]. The correlation between the SPAD meter readings and nitrogen percentage in leaves was strongly affected by leaf thickness. The variation in leaf thickness can influence the accuracy of SPAD meter readings, as this device works based on the leaf's capacity to absorb red and infrared lights.

B. Image Analysis-Based Nitrogen Status Estimation

1) DSELMs Fusion-Based Color Normalization: Based on our experiments, the proposed DSELMs fusion-based color constancy can be used to normalize wheat plant images captured under various light intensities. After image normalization, we can assume that all images are captured under the same light intensity and compared with each other.

In this paper, we also compare our results with other existing methods, i.e., MLPs fusion [28], gray world assumption [41], white-patch algorithm [42], linear regression model [43], and

TABLE I
COMPARISON OF COLOR NORMALIZATION RESULTS

Methods	ΔE_{RGB}	Time (s)*	Ref.
Gray world assumption	22.96	0.1156	[41]
White patch algorithm	14.25	0.1251	[42]
Linear regression model	10.38	0.1468	[43]
Single MLP	6.52	1455	[44]
MLPs fusion	4.29	21861	[28]
The proposed method (DSELMs fusion)	3.86	312	

* All algorithms were performed by using a desktop PC with 3.2 GHz Intel Core i5 processor and 8 GB of RAM.

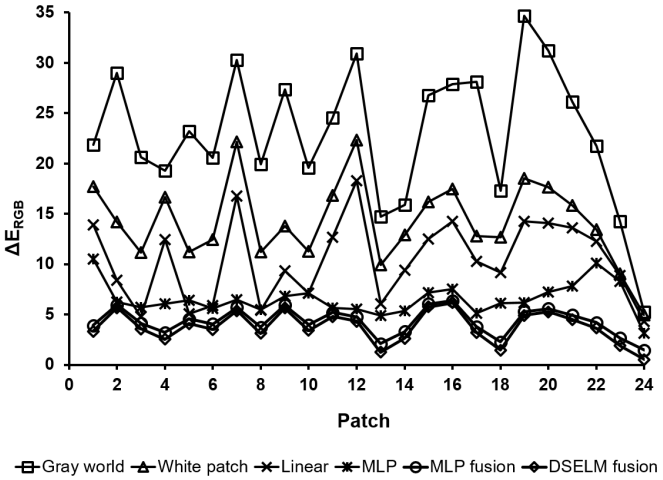


Fig. 13. ΔE_{RGB} of each patch using all methods.

single MLP [44]. The parameters used in this comparison are accuracy level and processing speed. The accuracy level is measured by calculating average Euclidean distance of the target and the output RGB color of all patches by using the following formula:

$$\Delta E_{\mathrm{RGB}} = \sqrt{(R_t - R_e)^2 + (G_t - G_e)^2 + (B_t - B_e)^2} \quad (36)$$

while the processing speed is measured by calculating the time (in seconds) required to process color normalization in each method. As seen in Table I, our proposed method using DSELMs fusion is better than the aforementioned methods in both accuracy and speed. The comparison of ΔE_{RGB} of each patch by using all methods is also presented in Fig. 13.

Basically, in the gray world as well as white patch and linear model algorithms, color normalization is accomplished by scaling color value of each pixel with a certain constant. It does not require training process as that in neural networks. This is why the processing times of color normalization using these methods are very quick. On the other hand, the output colors are significantly different compared to the target. These algorithms, therefore, are not suitable to normalize plants images.

All the neural network types used in this research has three input and output nodes, which represent red, green and blue color, respectively. In the MLP, we use one hidden layer with 92 nodes, according to the following empirical formula:

$$n_h = \left(\frac{n_i + n_o}{2}\right) + \sqrt{n_p} \quad (37)$$

where n_i, n_h, and n_o are the number of input, hidden, and output layer nodes, respectively, and n_p is the number of input patterns in the training set (number of training samples).

Furthermore, this neural network architecture is also used in both the MPLs fusion and DSELMs fusion methods. Since we fuse 24 MLPs, the process time, therefore, is considerably longer than that of single MLP. The results, however, is quite better, as indicated by the ΔE_{RGB} value. The challenge to produce better results and fast training on the color normalization process has been overcome by applying DSELMs fusion (see Table I). The color normalization error is less than that of MLPs fusion with processing speed 70 times faster.

In this proposed method, we utilize the GA to optimize 24 α values as the output weights of each DSELM. These α matrices are then applied to adjust wheat plant images by using the developed DSELMs fusion method. An example of the α matrices used to obtain a new image is as follows:

$$\begin{aligned} Z = &\begin{bmatrix} 0.039 & 0 & 0 \\ 0 & 0.069 & 0 \\ 0 & 0 & 0.007 \end{bmatrix} \cdot O_1 \\ &+ \begin{bmatrix} 0.053 & 0 & 0 \\ 0 & 0.025 & 0 \\ 0 & 0 & 0.005 \end{bmatrix} \cdot O_2 \\ &+ \begin{bmatrix} 0.047 & 0 & 0 \\ 0 & 0.014 & 0 \\ 0 & 0 & 0.065 \end{bmatrix} \cdot O_3 \\ &+ \ldots + \begin{bmatrix} 0.064 & 0 & 0 \\ 0 & 0.015 & 0 \\ 0 & 0 & 0.008 \end{bmatrix} \cdot O_{24}. \end{aligned}$$

By applying the developed DSELMs fusion-based color adjusting system and the optimized αmatrices, wheat crop images with a variation of light intensities will be transformed to that with standard light intensity (50 Klux). In this research, a wheat plant image has a dimension of 448 × 336 pixels. The parameter used for measuring the effectiveness of color normalization is color variability of the images. This parameter can be measured by calculating standard deviation of the original images and the corrected images. The smaller the standard deviation of image colors, the smaller the color variability of the images. As an example, from 30 plants images from the same field and the same fertilizer level, the standard deviations of the original leaves images are 23.61, 14.54, and 28.52 for red, green, and blue, respectively, while that of the corrected images are 5.28, 4.15, and 6.32, respectively. This example shows that the proposed method can be used to lower the color variability of wheat plants images captured on field with various light intensity.

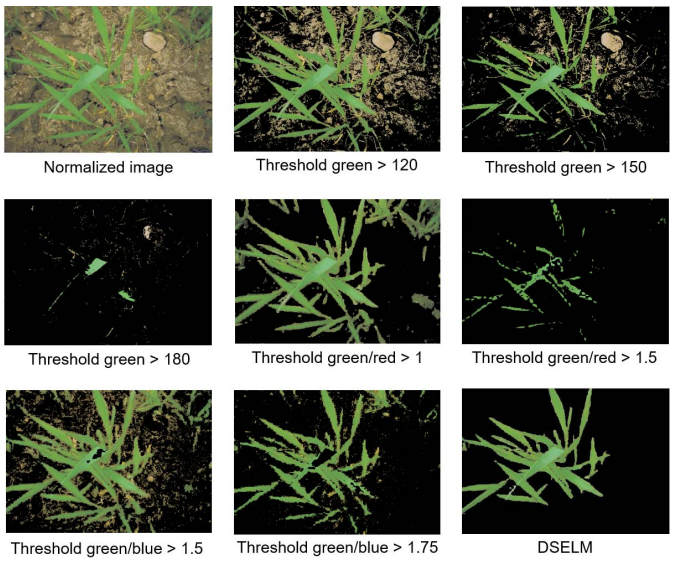


Fig. 14. Comparison of some threshold-based segmentation and the developed DSELM-based image segmentation.

TABLE II
COMPARISON OF IMAGE SEGMENTATION RESULTS

Methods	Accuracy (%)	Time (s)*
MLP	98.8	6.349
Original ELM	99.7	1.154
The proposed method (DSELM)	99.8	0.998

* All algorithms were performed by means of a desktop PC with 3.2 GHz Intel Core i5 processor and 8 GB of RAM.

2) DSELM-Based Image Segmentation and Feature Extraction: After image correction using the developed DSELMs fusion, the next step is image segmentation. In Fig. 14, we show the comparison of image segmentation. The DSELM-based image segmentation, as described in Section IV, has successfully distinguished wheat leaves from other parts, such as weeds, soil, stones, and dried leaves. The proposed segmentation method is superior to the conventional Otsu algorithm (threshold-based segmentation) and standard neural network learning algorithms such as MLP and ELM. The database pertaining to the color of leaves and nonleaves provides sufficient data to train the plant images. Therefore, the DSELM can precisely classify whether a pixel belongs to the leaves or nonleaves region.

Furthermore, the developed DSELM method can accomplish image segmentation process much faster than the standard MLP and ELM, as shown in Table II. Table II shows the obtained accuracy and time taken for these methods which indicated that the developed DSELM can precisely classify leaves or nonleaves region within less than one second. The image segmentation results using three different neural network types are all good as indicated by the accuracy level. However, the developed DSELM can perform the image segmentation 6 times faster than the MLP and slightly faster than the original ELM. Once all images have been segmented, 12 statistical color features as described previously in Section IV are then extracted to be utilized as predictors in the developed nitrogen estimation.

TABLE III
COMPARISON OF NITROGEN AMOUNT ESTIMATION ERRORS

Methods	MAPE	MAE	MSE	RMSE	SSE
SPAD meter	8.30%	0.1985	0.0520	0.2281	1.1610
1 DSELM	3.42%	0.0477	0.0288	0.0556	0.1599
Simple averaged 2 DSELMs	3.28%	0.0445	0.0277	0.0508	0.1205
Simple averaged 3 DSELMs	3.26%	0.0440	0.0275	0.0500	0.1148
Simple averaged 4 DSELMs	3.23%	0.0433	0.0273	0.0492	0.1091
Simple averaged 5 DSELMs	3.28%	0.0444	0.0276	0.0503	0.1169
Weighted averaged 4 DSELMs	2.78%	0.0379	0.0230	0.0431	0.0915

TABLE IV
COMPARISON OF ESTIMATION ERRORS USING COLOR FEATURES AND THE PROPOSED METHOD

Features	R	G	B	G/R	G/B
MAPE (%)	11.04	14.43	16.69	11.80	12.78
Features	G-R	2G-R-B	I_{kaw}	I_{PCA}	Proposed method
MAPE (%)	12.73	17.41	10.36	9.69	2.78

3) DSELM and Committee Machine-Based Image Nutrient Estimation: In this step, several DSELMs with various numbers of hidden layers are combined and then trained to determine which gives the best results. We establish that the combination of four DSELMs results in the minimum generalization error of networks performance compared to other possible combinations. The first combination used in this step is a simple average. The estimated nitrogen amount is then calculated, as follows:

$$Y_{\text{ave}} = \frac{1}{4}\sum_{i=1}^{4} O_i. \tag{38}$$

The best result relates to the first type of committee machine, i.e., simple average, is subsequently compared to that of the second type combiner, i.e., weighted average, which is optimized by the GA. By using this method, the weights of the output of the network system are 0.127, 0.385, 0.193, and 0.295, respectively, for the first until the fourth DSELM. The estimated nitrogen content can, therefore, be expressed as follows:

$$Y_{\text{GA}} = (0.127 \times O_1) + (0.385 \times O_2) + (0.193 \times O_3) + (0.295 \times O_4). \tag{39}$$

Table III shows the results on nitrogen estimation using the DSELM and the SPAD meter under various criteria such as MAPE, MAE, MSE, RMSE, and SSE. The SPAD

meter is the current method used to determine the nitrogen content of plant leave. Our results suggest that the proposed DSELM which extracts statistical features from the images has led to improved performance over the SPAD meter in all criterions.

Moreover, from Table III, we can perceive that by using a simple average combiner, the combination of four DSELMs provides the best results compared to other network combinations. The MAPE of this combination is around 3%. However, the weighted average combiner with GA optimization offers enhanced results. As seen in Table III, the MAPE of the GA-based committee machine with four DSELMs is smaller than the simple average method, i.e., 2.78%. It means that the deviation of the estimated nitrogen using this method is approximately 2.78% of the actual nitrogen content. For instance, if the actual nitrogen content is 3%, then the estimated nitrogen is between 2.917% and 3.083%. Thus, the error noted is relatively small.

We also investigated the relationship between nitrogen content and each color channel as well as a number of combinations of them. A research has been established that there are significant correlations between chlorophyll content in the maize leaf and the averages of the R and G components, as well as 2G-R-B of the linear transformation [4]. We also estimated the nitrogen content using the greenness index developed by Pagola *et al.* [45] (I_{kaw}). I_{kaw} is defined as follows:

$$I_{\text{kaw}} = \frac{R - B}{R + B}. \tag{40}$$

Based on the I_{kaw} formula, Pagola *et al.* [46] modified the greenness index to estimate nitrogen content in barley leaves and use the principal component analysis (PCA) to produce a new greenness index (I_{PCA}), as follows:

$$I_{\text{PCA}} = 0.7582|R - B| - 0.1168|R - G| + 0.6414|G - B|. \tag{41}$$

According to our investigation, single color features and their combinations, including I_{PCA} and I_{kaw}, are not suitable for nitrogen estimation. The estimation errors of those analyses are too high, compared to our proposed method, as seen in Table IV. The RGB values in those analyses are only obtained from the mean value of the observed leaves color. This value is not sufficient to represent the color distribution of leaves color. In the proposed method, we utilize not only mean value, but also variance, skewness, and kurtosis of the observed leaves color. The use of these statistical features is more effective to describe the color distribution of leaves color. As seen in Table IV, our proposed method is superior to all the discussed methods, as it provides an estimation error of 2.78%.

VII. Conclusion

A novel computational intelligence vision sensing has been proposed to acquire plant images and estimate nutrient content in wheat leaves based on color features of plant images captured on field with significant variation of sunlight intensities. The developed algorithm focuses on the development of DSELM and GA to overcome the problems on wide color variability due to different lighting conditions, image segmentation to distinguish crop leaves from complex background, and optimization of the nutrient estimation. The proposed method has been successfully demonstrated to normalize images as well as lower the color deviation and perform image segmentation much faster than other neural network methods. Furthermore, the combination of DSELMs with committee machine and GA has shown very promising results in estimating nitrogen content in wheat leaves compared with existing methods.

References

[1] Y. He, F. Liu, and D. Wu, "Nutrition management and automation," in *Agricultural Automation: Fundamentals and Practices*. Boca Raton, FL, USA: CRC Press, 2013, pp. 231–262.

[2] P. Auearunyawat, T. Kasetkaem, A. Wongmaneeroj, A. Nishihara, and R. Keinprasit, "An automatic nitrogen estimation method in sugarcane leaves using image processing techniques," in *Proc. Int. Conf. Agricult. Environ. Biol. Sci. (ICAEBS)*, 2012, pp. 39–42.

[3] X. Yao, W. Du, S. Feng, and J. Zou, "Image-based plant nutrient status analysis: An overview," in *Proc. IEEE Int. Conf. Intell. Comput. Intell. Syst. (ICIS)*, Oct. 2010, pp. 2492–2495.

[4] Y. Xu, X. Wang, H. Sun, H. Wang, and Y. Zhan, "Study of monitoring maize leaf nutrition based on image processing and spectral analysis," in *Proc. IEEE World Autom. Congr. (WAC)*, Sep. 2010, pp. 465–468.

[5] X. Yao, W. Luo, and Z. Yuan, "An adaptive and quantitative rubber nutrient status analyzing system by digital foliar images," in *Proc. IEEE Int. Congr. Image Signal Process. (CISP)*, Oct. 2010, pp. 2492–2495.

[6] D. Pascale. (2006). *The BabelColor Company*. Accessed: 2016. [Online]. Available: http://www.babelcolor.com

[7] H. Min, C. H. Jin, and Y. Sang-Hee, "A neural network approach to color constancy," in *Proc. 11th Int. Conf. Control, Autom. Syst. (ICCAS)*, Oct. 2011, pp. 1678–1681.

[8] D. Cheng, B. Price, S. Cohen, and M. S. Brown, "Beyond white: Ground truth colors for color constancy correction," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 298–306.

[9] J. Tang, C. Deng, and G.-B. Huang, "Extreme learning machine for multilayer perceptron," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 4, pp. 809–821, Apr. 2016.

[10] L. L. C. Kasun, H. Zhou, G.-B. Huang, and C. M. Vong, "Representational learning with extreme learning machine for big data," *IEEE Intell. Syst.*, vol. 28, no. 6, pp. 31–34, Nov. 2013.

[11] E. Cambria *et al.*, "Extreme learning machines [Trends & Controversies]," *IEEE Intell. Syst.*, vol. 28, no. 6, pp. 30–59, Nov./Dec. 2014.

[12] A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imag. Sci.*, vol. 2, no. 1, pp. 183–202, 2009.

[13] A. Beck and M. Teboulle, "A fast iterative Shrinkage–Thresholding algorithm with application to wavelet-based image deblurring," in *Proc. IEEE Int. Conf. Acoust., Speech, Sign. Process.*, Apr. 2009, pp. 693–696.

[14] H. X. Tian and Z. Z. Mao, "An ensemble ELM based on modified AdaBoost.RT algorithm for predicting the temperature of molten steel in ladle furnace," *IEEE Trans. Autom. Sci. Eng.*, vol. 7, no. 1, pp. 73–80, Jan. 2010.

[15] D. Niu, Y. Sun, and F. Wang, "Optimization of advertising budget allocation over time based on LS-SVMR and DE," *IEEE Trans. Autom. Sci. Eng.*, vol. 11, no. 4, pp. 1076–1082, Oct. 2014.

[16] Y. Liu, F.-L. Wang, Y.-Q. Chang, and C. Li, "A SNCCDBAGG-based NN ensemble approach for quality prediction in injection molding process," *IEEE Trans. Autom. Sci. Eng.*, vol. 8, no. 2, pp. 424–427, Apr. 2011.

[17] Z. Y. A. Ang, W. L. Woo, and E. Mesbahi, "Artificial neural network based prediction of energy generation from thermoelectric generator with environmental parameters," *J. Clean Energy Technol.*, vol. 5, no. 6, pp. 458–463, 2017.

[18] C. Wei, W. L. Woo, and S. S. Dlay, "Nonlinear underdetermined blind signal separation using Bayesian neural network approach," *Digit. Signal Process.*, vol. 17, no. 1, pp. 50–68, 2007.

[19] W. L. Woo and S. S. Dlay, "Neural network approach to blind signal separation of mono-nonlinearly mixed sources," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 52, no. 6, pp. 1236–1247, Jun. 2005.

[20] D. Liu, D. Wang, D. Zhao, Q. Wei, and N. Jin, "Neural-network-based optimal control for a class of unknown discrete-time nonlinear systems using globalized dual heuristic programming," *IEEE Trans. Autom. Sci. Eng.*, vol. 9, no. 3, pp. 628–634, Jul. 2012.

[21] H. Zhang, C. Qin, and Y. Luo, "Neural-network-based constrained optimal control scheme for discrete-time switched nonlinear system using dual heuristic programming," *IEEE Trans. Autom. Sci. Eng.*, vol. 11, no. 3, pp. 839–849, Jul. 2014.

[22] T. T. Teo, T. Logenthiran, and W. L. Woo, "Forecasting of photovoltaic power using extreme learning machine," in *Proc. IEEE Innov. Smart Grid Technol.-Asia (ISGT Asia)*, Bangkok, Thailand, Nov. 2015, pp. 1–5.

[23] S. B. Sulistyo, W. L. Woo, and S. S. Dlay, "Regularized neural networks fusion and genetic algorithm based on-field nitrogen status estimation of wheat plants," *IEEE Trans. Ind. Informat.*, vol. 13, no. 1, pp. 103–114, Feb. 2016.

[24] A. Gijsenij, T. Gevers, and J. van de Weijer, "Computational color constancy: Survey and experiments," *IEEE Trans. Image Process.*, vol. 20, no. 9, pp. 2475–2489, Sep. 2011.

[25] S. Bianco, G. Ciocca, C. Cusano, and R. Schettini, "Improving color constancy using indoor–outdoor image classification," *IEEE Trans. Image Process.*, vol. 17, no. 12, pp. 2381–2392, Dec. 2008.

[26] K. Barnard, V. Cardei, and B. Funt, "A comparison of computational color constancy algorithms—Part I: Methodology and experiments with synthesized data," *IEEE Trans. Image Process.*, vol. 11, no. 9, pp. 972–983, Sep. 2002.

[27] G.-B. Huang, L. Chen, and C.-K. Siew, "Universal approximation using incremental constructive feedforward networks with random hidden nodes," *IEEE Trans. Neural Netw.*, vol. 17, no. 4, pp. 879–892, Jul. 2006.

[28] S. B. Sulistyo, W. L. Woo, and S. S. Dlay, "Computational intelligent color normalization for wheat plant images to support precision farming," in *Proc. IEEE Int. Conf. Adv. Comput. Intell.*, Feb. 2016, pp. 130–135.

[29] J. Song, Y. Qiu, and Z. Liu, "Integrating optimal simulation budget allocation and genetic algorithm to find the approximate Pareto patient flow distribution," *IEEE Trans. Autom. Sci. Eng.*, vol. 13, no. 1, pp. 149–159, Jan. 2016.

[30] C.-W. Chou, C.-F. Chien, and M. Gen, "A multiobjective hybrid genetic algorithm for TFT-LCD module assembly scheduling," *IEEE Trans. Autom. Sci. Eng.*, vol. 11, no. 3, pp. 692–705, 2014.

[31] P. K. Jamwal, S. Hussain, and S. Q. Xie, "Three-stage design analysis and multicriteria optimization of a parallel ankle rehabilitation robot using genetic algorithm," *IEEE Trans. Autom. Sci. Eng.*, vol. 12, no. 4, pp. 1433–1446, Oct. 2015.

[32] F. Qiao, Y. M. Ma, L. Li, and H. X. Yu, "A Petri net and extended genetic algorithm combined scheduling method for wafer fabrication," *IEEE Trans. Autom. Sci. Eng.*, vol. 10, no. 1, pp. 197–204, Jan. 2013.

[33] D. A. Coley, *An Introduction to Genetic Algorithms for Scientists and Engineers*. Singapore: World Scientific, 1999.

[34] N. Mohamad, M. K. A. Ariffin, A. Ali, F. Mustapha, and I. M. Salleh, "Development of genetic algorithm toolbox using MATLAB in cutting tool path optimization," *Sci. Res. Essays*, vol. 8, no. 38, pp. 1848–1857, 2013.

[35] A. Dacal-Nieto, E. Vázquez-Fernández, A. Formella, F. Martin, S. Torres-Guijarro, and H. González-Jorge, "A genetic algorithm approach for feature selection in potatoes classification by computer vision," in *Proc. IEEE 35th Annu. Conf. Ind. Electron.*, Nov. 2009, pp. 1955–1960.

[36] F. J. Lin, P. K. Huang, and W. D. Chou, "Recurrent-fuzzy-neural-network-controlled linear induction motor servo drive using genetic algorithms," *IEEE Trans. Ind. Electron.*, vol. 54, no. 3, pp. 1449–1461, Jun. 2007.

[37] Q. K. Man, C. H. Zheng, X. F. Wang, and F. Y. Lin, "Recognition of plant leaves using support vector machine," in *Advanced Intelligent Computing Theories and Applications. With Aspects of Contemporary Intelligent Computing Techniques* (Communications in Computer and Information Science), vol. 15, D. S. Huang, D. C. Wunsch, D.S. Levine, and K. H. Jo, Eds. Berlin, Germany: Springer, 2008, pp. 192–199.

[38] S. B. Sulistyo, W. L. Woo, and S. S. Dlay, "Ensemble neural networks and image analysis for on-site estimation of nitrogen content in plants," in *Proc. SAI Intell. Syst. Conf. (Intellisys)*, London, U.K., 2016, pp. 103–118.

[39] H. Yang, J. Yang, Y. Lu, and J. He, "SPAD values and nitrogen nutrition index for the evaluation of Rice nitrogen status," *Plant Prod. Sci.*, vol. 17, no. 1, pp. 81–92, 2014.

[40] H.-Y. Song, Z.-L. Guo, Y. He, H. Fang, and Z.-Y. Zhu, "Non-destructive estimation oilseed rape nitrogen status using chlorophyll meter," in *Proc. IEEE Int. Conf. Mach. Learn. Cybern.*, Aug. 2006, pp. 4252–4256.

[41] M. T. T. Nguyen, C. L. D. A. Mai, and N. M. Kwok, "Estimating image illuminant color based on gray world assumption," in *Proc. IEEE Int. Congr. Image Signal Process.*, Oct. 2011, pp. 989–993.

[42] J. Y. Cepeda-Negrete and R. E. Sanchez-Yanez, "Color constancy algorithms in practice," in *Proc. Robot. Summer Meet. (ROSSUM)*, 2011, pp. 78–79.

[43] Y. A. Sari, R. V. H. Ginardi, and N. Suciati, "Color correction using improved linear regression algorithm," in *Proc. IEEE Conf. Inform. Commun. Technol. Syst. (ICIS)*, Sep. 2015, pp. 73–78.

[44] C. Shengxian, D. Bangkui, S. Jiawei, L. Fan, Y. Shanrang, and X. Zhiming, "A colour constancy algorithm based on neural network and application," in *Proc. IEEE World Congr. Intell. Control Autom.*, Jul. 2010, pp. 3100–3103.

[45] S. Kawashima and M. Nakatani, "An algorithm for estimating chlorophyll content in leaves using a video camera," *Ann. Botany*, vol. 81, no. 1, pp. 49–54, 1998.

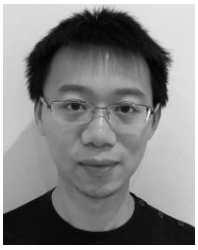
[46] M. Pagola *et al.*, "New method to assess barley nitrogen nutrition status based on image colour analysis: Comparison with SPAD-502," *Comput. Electron. Agricult.*, vol. 65, no. 2, pp. 213–218, 2009.

[47] Y. Chen, H. Yu, C. Miao, B. Chen, X. Yang, and C. Leung, "Using motor patterns for stroke detection," *Sci. Adv. Comput. Psychophysiol.*, vol. 350, no. 6256, pp. 12–14, 2015.

[48] G. Huang, S. Song, J. N. D. Gupta, and C. Wu, "Semi-supervised and unsupervised extreme learning machines," *IEEE Trans. Cybern.*, vol. 44, no. 12, pp. 2405–2417, Dec. 2014.

Susanto B. Sulistyo received the B.Sc. (Hons.) degree in agricultural technology and the M.Sc. degree in agricultural engineering sciences from Bogor Agricultural University, Bogor, Indonesia, in 2003 and 2008, respectively. He is currently pursuing the Ph.D. degree with Newcastle University, Newcastle upon Tyne, U.K.

His research interests include image processing and artificial intelligence applications in agricultural systems.



Di Wu received the B.S. degree from the Minzu University of China, Beijing, China, and the M.Sc. degree in communications and signal processing from Newcastle University, Newcastle upon Tyne, U.K., in 2010 and 2012, respectively, where he is currently pursuing the Ph.D. degree.

His research interests include deep learning, audio and image processing, source separation, and object recognition.

Wai Lok Woo (M'09–SM'11) received the B.Eng. degree in electrical and electronics engineering and the Ph.D. degree from Newcastle University, Newcastle upon Tyne, U.K.

He is currently a Reader in Intelligent Signal Processing and the Director of Operations for the international branch of the university in Singapore. His major research is in mathematical theory and algorithms for machine learning and multidimensional signal processing.

Dr. Woo is a member of the Institution Engineering Technology. He was a recipient of the IEE Prize and the British Scholarship. Currently, he serves on the editorial board of the several international signal processing journals.

S. S. Dlay received the B.Sc. (Hons.) degree in electrical and electronic engineering and the Ph.D. degree in VLSI design from Newcastle University, Newcastle upon Tyne, U.K.

In 1984, he was a Lecturer with the Department of Electronic Systems Engineering, University of Essex, Colchester, U.K. In 1986, he joined Newcastle University as a Lecturer, and then in 2006, he became a Personal Chair in Signal Processing Analysis.

Prof. Dlay is a College Member of the Engineering and Physical Science Research Council (EPSRC). He was a recipient of a Scholarship from the Engineering and Physical Science Research Council (EPSRC) and the Charles Hertzmann Award.

Bin Gao (M'12–SM'14) received the B.S. degree in communications and signal processing from Southwest Jiao Tong University, Chengdu, China, in 2005, and the M.Sc. degree (Hons.) in communications and signal processing and the Ph.D. degree from Newcastle University, Newcastle upon Tyne, U.K., in 2007 and 2011, respectively.

He is currently a Professor with the School of Automation Engineering, University of Electronic Science and Technology of China, Chengdu. His research interests include sensor signal processing, machine learning, nondestructive testing and evaluation.