Indian Journal of Agricultural Sciences **91** (3): 388–92, March 2021/Article
https://doi.org/10.56093/ijas.v91i3.112496

Leaf area assessment using image processing and support vector regression in rice

TANUJ MISRA[1], SUDEEP MARWAHA[1]*, ALKA ARORA[1], MRINMOY RAY[1], SHAILENDRA KUMAR[1], SUDHIR KUMAR[1] and VISWANATHAN CHINNUSAMY[1]

ICAR-Indian Agricultural Statistics Research Institute, New Delhi 110 012, India

Received: 17 May 2020; Accepted: 12 November 2020

ABSTRACT

Crop growth, health, and correspondingly yield are much affected by abiotic environmental factors. Abiotic stress is considered as a threat to food security and has a disastrous consequence. Phenotyping parameters such as leaf area assessment is of utmost importance in determining the stresses due to water and environmental factors, micronutrients deficiencies, leaf diseases, pests, *etc.* In this study, a non-destructive approach through digital image analysis has been presented to assess the total leaf area of rice plants grown in pot culture. Images have been captured from four different angles with respect to the initial position of the camera. Support Vector Regression (SVR) and Tuned SVR have been employed by considering the pixel area of leaves obtained from different angles. Performance of Tuned SVR has been found better than the SVR on training and testing dataset based on RMSE values. A web-solution has been designed and developed to implement the presented approach using 3-tier architecture: Client-Side Interface Layer (CSIL), Database Layer (DL) and Server Side Application Layer (SSAL).

Keywords: Image analysis, Leaf area, Non-destructive phenotyping, Rice, Support Vector Regression (SVR), Tuned Support Vector Regression

Measurements of leaf parameters are important in the studies of plant biological characteristics (Sestak *et al.* 1971, Ali and Anjum 2004). Leaf area is important for crop light interception and therefore has a large influence on crop growth (Boote *et al.* 1988), transpiration (Enoch and Hurd 1979) and growth rate (Lieth *et al.* 1986). Many sophisticated electronic instruments provide accurate and fast leaf area measurement but they are expensive and destructive in nature (Deblonde *et al.* 1997). In the destructive methods, leaves are plucked and the area measurement is taken with the electronic devices. The most widely used digital device for measuring leaf area and dimensions is the LI-3100 leaf area meter (Tian and Wang 2009). Although the destructive method is a benchmark, it is laborious and time-consuming. Recently, image analysis techniques are widely used in plant phenotypic parameter estimation (Bi *et al.* 2010, Li *et al.* 2017, Sadeghi-Tehran *et al.* 2017, Hasan *et al.* 2018, Misra *et al.* 2019, 2020). A good account of research work is available in the field of leaf area assessment through image analysis (James and Newcombe 2000, Li *et al.* 2008, Tian and Wang 2009, Marcon and Mariano 2011, Patil and Bodhe 2011, Fanourakis *et al.* 2014). In almost all the mentioned studies, attempts were taken to estimate the area of single leaf. Calculating the area of a single leaf of plant is little bit easy as compared to calculation of area of all the leaves using color image analysis because of the complex growing nature of the plant. In this study, an experiment was conducted to estimate the total shoot area of rice plants of two varieties (IR 64 and N22) under pot cultivation using digital image analysis. Both varieties are of dense growing nature. Rice crop is opted for this work as it is a major food crop in India. Leaf area obtained from images as well as leaf area computed by Li-Cor meter (benchmark) was used to develop Support Vector Regression (SVR) models to predict the leaf area non-destructively. Results drawn from the models are compared and the best one has been used to design and develop the web based software.

MATERIALS AND METHODS

Rice plants of two genotypes (IR-64 and N22) were planted in each of 15 pots in the Phenomics laboratory of the Division of Plant Physiology of ICAR-Indian Agricultural Research Institute (IARI), New Delhi in the year of 2017. Plant images (Fig 1a) were captured from four sides (0°, 90°, 180°, 270°) with respect to the initial position of the plant by Nikon d7000 DSLR (16 megapixel) RGB camera maintaining 1 meter distance from the plant. Images were

[1]ICAR-Indian Agricultural Statistics Research Institute, New Delhi. *Corresponding author e-mail: Sudeep.Marwaha@icar.gov.in.

taken at the difference of 90 degree to cover the entire leaves of the plants with minimal overlapping. Total 120 ([15+15]*4) images of 30 plants (15 plants of IR-64 and N22 each) were captured. The age of plants and time of imaging were kept constant while taking images. Images were taken in the morning between 10 am to 11 am at the vegetative stage of the plant. A uniform black background was maintained during image capturing to increase the accuracy of separation between background and the plant regions. After taking images, leaves were plucked for actual measurement of leaf area by LI-Cor 3100 Leaf Area Meter (Deblonde *et al.* 1997). Samples were placed between fixed guides on the lower transparent belt and allowed to pass through the LI-3100 meter. As the sample traveled under the fluorescent light source, the projected objects were reflected by a system of three mirrors to a linear array camera within the rear housing. An adjustable press roller flattens curled leaves and feeds them properly between the transparent belts. This makes it possible to accurately measure leaf area of small grasses, legumes, aquatic plants, and similar types of leaves. Leaf area taken by LI-C or 3100 was considered as a benchmark of leaf area measurement in the study.

Image analysis algorithm: The plant color images (RGB) were first converted into the grayscale image by simply averaging the Red (R), Green (G) and Blue (B) pixel values. Grayscale images were then converted into binary images using the Otsu's thresholding method (Xu *et al.* 2011) to separate plant area from the background. The thresholding mechanism is popularly known as image segmentation. The segmentation process was accomplished by selecting the pixels with values over the threshold value belonging to the plant region and other pixels to the background region. The resulting binary image consists of white and black pixels representing the plant and background regions, respectively. The number of pixels inside the plant region was counted and converted to square centimeter (cm^2) using the reference object (any object whose actual area is known previously) to get the projected leaf area. The steps involved in image analysis algorithm for projected leaf area estimation is represented in Fig. 1(b).

SVR model development: Projected leaf areas obtained from the four side view images of a specified plant and the corresponding ground truth leaf area recorded using Li-Cor meter were used to develop the Support Vector Regression (SVR) model. Nowadays, Support Vector Regression (SVR) has been widely used in divergent domain. A tremendous advantage of SVR is that it is data-driven and and can handle non-linear dynamics. A good account of SVR is given in Cortes and Vapnik (1995), Vapnik *et al.* (1997), David (2017) etc. The SVR performs the following nonlinear function mapping between the input and output:

$$f(y) = (w.\phi(x)) + c \quad \text{(i)}$$

where, w denotes the weights, c represents threshold value and ϕ (x) is known as kernel function. The objective is to find out the optimal values of w and c. In SVR, there are two main objectives, viz. flatness of weights and error which is also called empirical risk are to be minimized. However, the overall aim is to minimize the regularized risk which is sum of empirical risk and the half of the product of the flatness of weight and a constant term which is known as regularized constant. The regularized risk can be written as follows:

$$R_{reg}(f) = R_{emp}(f) + \frac{\tau}{2}\|w\|^2 \quad \text{(ii)}$$

where, $R_{reg}(f)$ is the regularized risk, $R_{emp}(f)$ denotes the empirical risk, τ is as constant which is called as regularized constant/capacity control term and $\|w\|$ is the flatness of weights. The regularization constant has a significant impact on a better fitting of the data and it can also be useful for the minimization of bad generalization effects.

The empirical risk can be represented as follows:

$$R_{emp}(f) = \frac{1}{N}\sum_{i=0}^{N-1} L\big(x(i), \alpha(i), f(x(i), w)\big) \cdots \quad \text{(iii)}$$

where, $\alpha(i)$ denotes the truth data of predicted value, $L(.)$ is known as loss function.

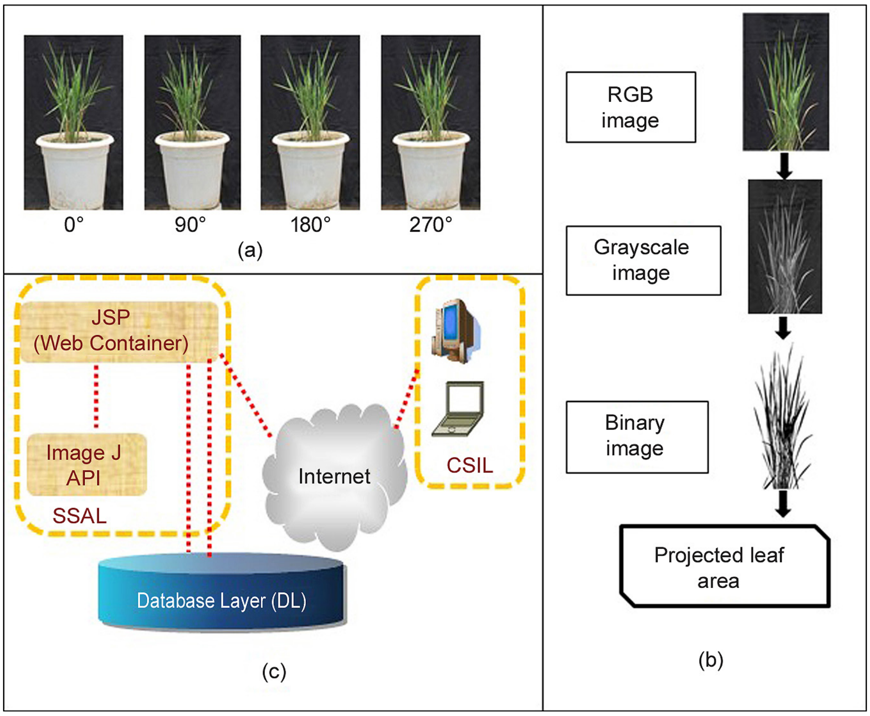

Fig 1 (a) Images taken from four direction of a single plant; (b) Image analysis algorithm for projected leaf area estimation; (c) 3-tier architecture of the web-based software consisting of Client Side Interface Layer (CSIL), Server Side Application Layer (SSAL) and Database Layer (DL).

There are various types of loss function in literature. But, two functions, viz. vapnik loss function and quadratic loss function are most popular and they are generally used. The quadratic programming problem has been made to minimize the regularised risk which is represented as follows:

$$\text{Minimize, } \frac{1}{2}w^2 + D\sum_{i=1}^{n} L\big(\alpha(i), f(x(i), w)\big) \tag{iv}$$

where,

$$L\big(\alpha(i), f(x(i), w)\big) = \big|\alpha(i) - f(x(i), w)\big| - \in \text{ if}$$

$$\big|\alpha(i) - f(x(i), w)\big| \geq \in \; = 0; \text{ otherwise.}$$

where, D is a constant which equals to the summation normalization factor and $\in$ represents the size of the tube. The computation of $\in$ and D is done empirically because the proper value of D and $\in$ are obtained by dual optimization problem using the lagrange multiplier which can be written as follows:

Maximize,

$$-\frac{1}{2}\sum_{i,j=1}^{N}\left(\beta_i - \beta_i^*\right)\left(\beta_j - \beta_j^*\right)x(i), x(j) - \in \sum_{i=1}^{N}\left(\beta_i - \beta_i^*\right) + \sum_{i=1}^{N}\alpha(i)\left(\beta_i - \beta_i^*\right) \tag{v}$$

$$\text{Subject to, } \sum_{i-1}^{N}\left(\beta_i - \beta_i^*\right) = 0\,;\; \beta_i, \beta_i^* \in [0, D] \tag{vi}$$

The function $f(x)$ is defined as;

$$f(x) = \sum_{i=1}^{N}\left(\beta_i - \beta_i^*\right)x, x(i) + C \tag{vii}$$

Karush–Kuhn–Tucker (KKT) conditions are used to get the solution of the weights. The significance of kernel function in non-linear support vector regression (NL-SVR) is very much imporatnt for mapping the data $y(i)$ into higher dimension feature space $\varnothing(x(i))$ in which the data becomes linear. Generally notation for kernel function is given as;

$$k\left(x, x'\right) = [\varnothing(x), \varnothing(x')] \tag{viii}$$

There are many methods in literature to solve the quadartic programming. However, the most used method is sequential minimization optimization (SMO) algorithm.

Measurement of the performance: Model performance was judged by computing Root Mean Square Error (RMSE). The model with less RMSE was preferred for prediction purposes.

Web-solution development: Based on the proposed approach, web based software has been developed. The software consists of 3-tier architecture: Client Side Interface Layer (CSIL) is implemented by Hyper Text Markup Language (HTML), Cascading Style Sheet (CSS) and JavaScript (JS); Database Layer (DL) has been developed for storing the user information, images and experimental information carried out by users under various experiments, implemented using Microsoft Structured Query Language (MSSQL) Server 2008; JSP (Java Server Pages) technology has been used to implement Server Side Application Layer (SSAL). ImageJ library is integrated with SSAL for the computation of leaf area. ImageJ is an open source library developed in Java which provides features for image analysis through its API. For developing the web-solution, NetBeans 7.0.1 IDE (Integrated Development Environment) has been used. Fig. 1(c) shows the architecture of the web based software for leaf area estimation.

RESULTS AND DISCUSSION

The experimental dataset consists of images of 30 plants (15 for IR-64 and N22 each) taken from 4 side views. Therefore, the image dataset comprises 120 images ([15+15]*4). Randomly selected 80% of the dataset, i.e. 96 images of 24 plants were used to develop the SVR model and the remaining (24 images of 6 plants) were used to test the accuracy and efficiency of the developed model.

There is a multi-collinearity problem in the data as explanatory variables are highly correlated. Conventional regression model cannot deal with multicollinearity. Hence, in this study instead of conventional regression, Support Vector Regression (SVR) model has been applied as SVR can deal with multicollinearity by using regularization. In addition, by tuning the hyperparameters (cost and epsilon) of SVR, viz. Tuned SVR model has also been applied. For SVM fitting, several kernel functions namely Linear, Polynomial, Sigmoid and Radial Basis Function (RBF) have been applied for the SVM models. However, RBF was found superior as compared to other kernel functions for the considered regression problems. For Tuned SVR model also RBF was found superior as compared to other kernel functions. The estimated parameters of SVR model and tuned SVR model are given in Table 1.

Table 1 Parameter estimation of SVR and Tuned SVR

Parameter estimation of SVR		Parameter estimation of Tuned SVR	
Sampling method	10-fold cross validation	Sampling method	10-fold cross validation
Epsilon	0.1	Epsilon	0
Cost	1	Cost	11.31371
Gamma	0.25	Gamma	0.25
Number of Support Vectors	20	Number of Support Vectors	24
SVR-Type	eps-regression	SVR-Type	eps-regression
SVR-Kernel	Radial Basis Function	SVR-Kernel	Radial Basis Function

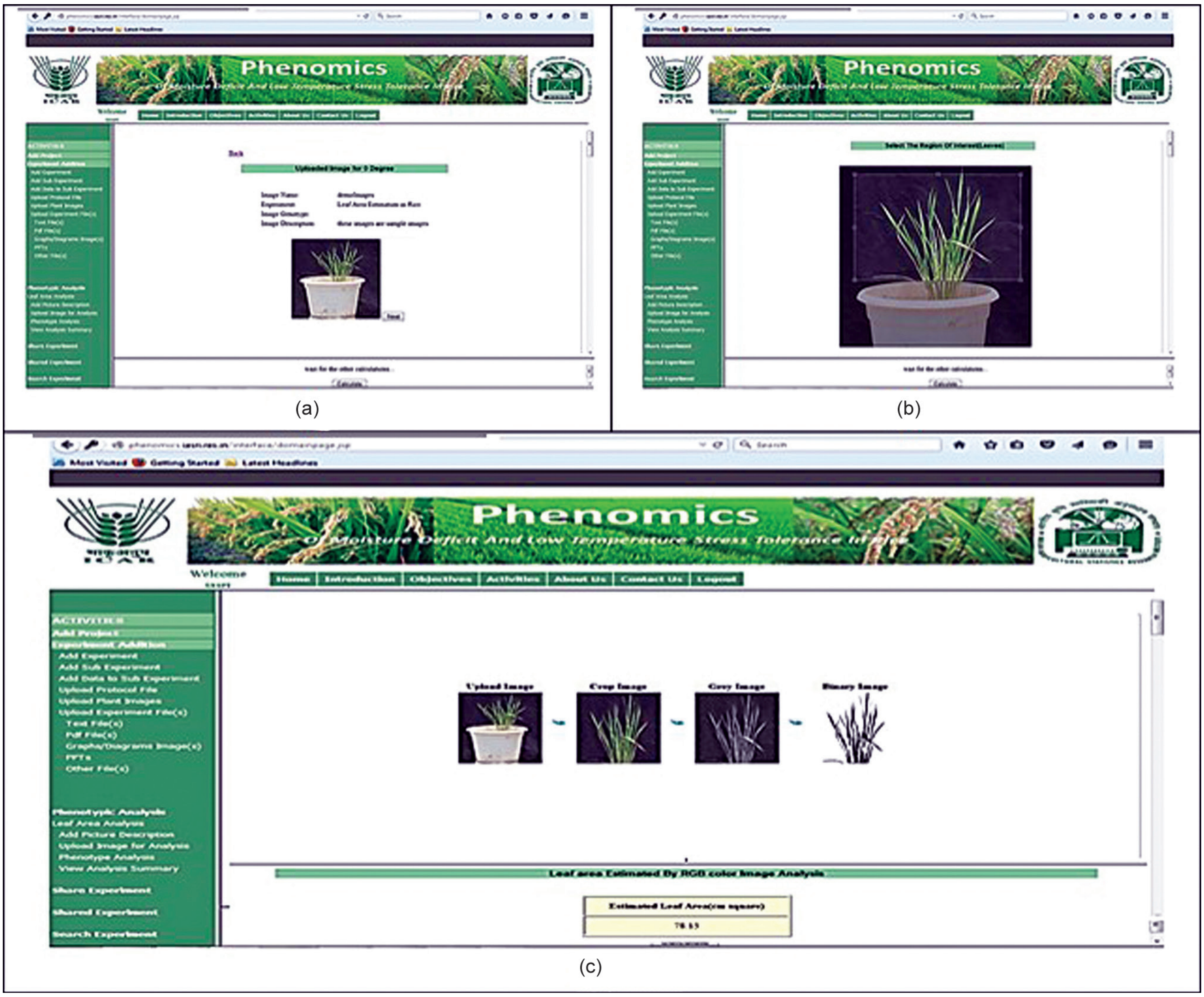

Fig 2 (a) upload image; (b) crop image; (c) leaf area calculation using tuned SVR model.

Further, the hyperparameters of SVR model has been tuned by employing grid search algorithm.

Table 2 shows that RMSE values for training and testing are comparatively lower in case of Tuned SVR than SVR. It reflects that the performance of Tuned SVR is much better than the SVR. So, the projected leaf area through the image analysis algorithm and the ground truth Leaf area meter values are suitably explained by the Tuned SVR.

Steps of Leaf Area Estimation through web-based software:

Step 1: Upload/Choose Plant image [Fig 2 (a)]

Step 2: Crop the region of interest [Fig 2(b)]: After uploading the image user has to select and crop the area covered by the leaves in the image.

Table 2 Comparison between SVR and Tuned SVR

Model	RMSE	
	Training	Testing
SVR	26.43	42.25
Tuned SVR	17.56	33.31

Step 3: Estimation of Leaf area [Fig 2(c)]: System receives the cropped image and the proposed image analysis algorithm will be applied to extract the projected leaf area which will be used as input to the tuned SVR model to estimate the total leaf area of the specified plant.

Most of the methods and electronic equipment are available for leaf area estimation is based on destructive method which is very expensive and slow. In this study, we have presented a methodology for estimating the leaf area of rice plant in pot culture non-destructively through visual image analysis. The developed tuned SVR model is outperformed over SVR model with RMSE 17.56 and 33.31 for the training and test dataset, respectively. This approach may have wider applicability in the assessment of leaf area in other cereal crops like wheat, barley, maize *etc.* as they have similar characteristics. A web-solution has been developed by implementing the proposed methodology which is user friendly and will be very helpful for the plant physiologist.

ACKNOWLEDGEMENTS

The first author acknowledges the fellowship received from IASRI, New Delhi, India to undertake the research work. This work was supported by National Agriculture Science Fund (NASF), ICAR, Grant No.NASF/Phen-6005/2016-17.

REFERENCES

Ali H and Anjum M A. 2004. Aerial growth and dry matter production of potato (*Solanum tuberosum* L.) cv. Desiree in relation to phosphorus application. *International Journal of Agriculture and Biology* **6**: 458–61.

Bi K, Jiang P, Li L, Shi B and Wang C. 2010. Non-destructive measurement of wheat spike characteristics based on morphological image processing. *Transactions of the Chinese Society of Agricultural* **12**: 212–16.

Boote K J, Jones J W and Hoogenboom G. 1988. Research and management application of the 'PNUTGRO' crop growth model. *Proceedings of the American Peanut Research Educational Society* **20**: 57.

Cortes C and Vapnik V. 1995.Support-vector network. *Machine Learning* **20**: 1–25.

David M. 2017. E1071: Misc Functions of the Department of Statistics. *Probability Theory Group R package version* **1**: 6–8.

Deblonde G, Penner M and Royer A. 1994. Measuring leaf area index with the LI-COR LAI-2000 in pine stands. *Ecology* **75**(5): 1507–11.

Enoch H Z and Hurd R G. 1979. The effect of elevated CO_2 concentrations in the atmosphere on plant transpiration and water use efficiency. A study with potted carnation plants. *International Journal of Biometeorology* **23**(4): 343–51.

Fanourakis D, Briese C, Max J F, Kleinen S, Putz A, Fiorani F, Ulbrich A and Schurr U. 2014. Rapid determination of leaf area and plant height by using light curtain arrays in four species with contrasting shoot architecture. *Plant Methods* **10**(1): 1–11.

Hasan M M, Chopin J P, Laga H and Miklavcic S J. 2018. Detection and analysis of wheat spikes using Convolutional Neural Networks. *Plant Methods* **14**(1): 1–13.

James R R and Newcombe G. 2000. Defoliation patterns and genetics of insect resistance in cottonwoods.*Canadian Journal of Forest Research* **30**(1): 85–90.

Li Q, Cai J, Berger B, Okamoto M and Miklavcic S J. 2017. Detecting spikes of wheat plants using neural networks with laws texture energy. *Plant Method* **13**:1–13.

Li Z, Ji C and Liu J. 2008. Leaf area calculating based on digital image.*Computer and Computing Technologies in Agriculture* **259**: 1427–33.

Lieth J H, Reynolds J F and Rogers H H. 1986. Estimation of leaf area of soybeans grown under elevated carbon dioxide levels. *Field Crops Research* **13**: 193–203.

Marcon M and Mariano K. 2011. Estimation of total leaf area in perennial plants using image analysis. *R. Bras. Eng. Ambiental* **15**: 96–101.

Misra T, Arora A, Marwaha S, Chinnusamy V, Rao A R, Jain R, Sahoo R N, Ray M, Raju D, Jha R R and Goel S. 2020. SpikeSegNet-a deep learning approach utilizing encoder-decoder network with hourglass for spike segmentation and counting in wheat plant from visual imaging. *Plant Methods* **16**(1): 1–20.

Misra T, Arora A, Marwaha S, Ray M, Raju D, Kumar S, Goel S, Sahoo R N and Chinnusamy V. 2019. Artificial neural network for estimating leaf fresh weight of rice plant through visual-nir imaging.*Indian Journal of Agricultural Sciences* **89**: 1698–1702

Patil S B and Bodhe S K. 2011.Betel Leaf Area Measurement Using Image Processing.*International Journal on Computer Science and Engineering (IJCSE)* **3**(7): 2656-2660

Sadeghi-Tehran P, Sabermanesh K, Virlet N and Hawkesford M J. 2017. Automated method to determine two critical growth stages of wheat: heading and flowering. *Frontiers in Plant Science* **8**: 252.

Tian Y W and Wang X J. 2009.Analysis of leaf parameters measurement of cucumber based on image processing. (In) *2009 WRI World Congress on Software Engineering, IEEE* **3**: 34–37.

Tian Y and Wang X. 2009. Development of Hand-held Leaf Area Meter Based on S3C2440.*Proc.The 9th International Conference on Electronic Measurement and Instruments*, Beijing, China.

Vapnik V, Golowich S E and Smola A J. 1997. Support vector method for function approximation, regression estimation and signal processing. *Advances in Neural Information Processing Systems* 281–87.

Xu X, Xu S, Jin L and Song, E. 2011. Characteristic analysis of Otsu threshold and its applications. *Pattern Recognition Letters* **32**(7): 956–61.