Automatic Fruit Grading and Classification System Using Computer Vision: A Review

Seema
Department of Physics
National Institute of Technology
Kurukshetra-136119, India
E-mail:er.seema5@gmail.com

A. Kumar
Department of Physics
National Institute of Technology
Kurukshetra-136119, India
E-mail: ashavani@yahoo.com

G. S. Gill
Department of Instrumentation
Kurukshetra University
Kurukshetra-136119, India
E-mail: gsgill@kuk.ac.in

***Abstract*—Automation in agriculture comes into play to increase productivity, quality and economic growth of the country. Fruit grading is an important process for producers which affects the fruits quality evaluation and export market. Although the grading and sorting can be done by the human, but it is slow, labor intensive, error prone and tedious. Hence, there is a need of an intelligent fruit grading system. In recent years, researchers had developed numerous algorithms for fruit sorting using computer vision. Color, textural and morphological features are the most commonly used to identify the diseases, maturity and class of the fruits. Subsequently, these features are used to train soft computing technique network. In this paper, use of image processing in agriculture has been reviewed so as to provide an insight to the use of vision based systems highlighting their advantages and disadvantages.**

Keywords—Fruit grading; computer vision; morphological features; texture

I. INTRODUCTION

In recent years, use of image processing has been increasing day by day in different areas such as industrial image processing, medical imaging, real time imaging, texture classification, object recognition, etc. Image processing and computer vision in agriculture is another fast growing research field. It is an important analyzing tool for pre-harvest to post-harvest of crops. It has lots of applications in agriculture. In land identification, image processing is used for identification of land that will be suitable for agriculture [1]. In plant nitrogen identification, image processing can be used for estimation of plant nitrogen identification and chlorophyll identification [2]. In pest control, image processing is a good tool for identification of pest infected areas because it favors to build up the pest population [3] [4]. Using computer vision, image processing is used to automatic detection and classification of plant disease from color, texture and shape [5]. Food quality can be improved by quality inspection using computer vision [6].

II. COMPUTER VISION

Computer vision is used to gather the information from the images which are captured from the real time world. It is a field that includes methods for image acquisition, processing, analyzing and understanding the images in order to gather symbolic and numerical information. Basically its aim is to duplicate the effect of human vision by electronically perceiving, understanding and classification of images [7].Computer vision is widely used in post-harvest industries for quality inspection and grading of fruits and vegetables.

A. Computer Vision Setup

A computer vision system constitutes of five components: illumination, a camera, an image capture board also known as a frame grabber or digitizer, computer hardware and software [8]. Computer vision setup is shown in Fig 1.

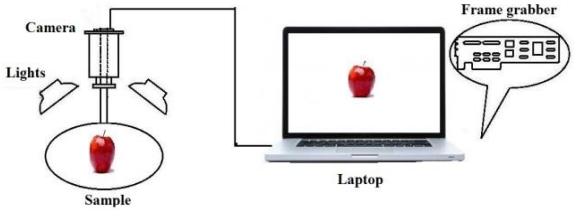

Fig1. The computer vision system's components

B. Color Features

Images are captured in RGB color models. It is most common color model in image processing and based on primary colors red (R), green (G), blue (B). Basically, for color features, every image is separated into red, green and blue planes, respectively, and through these planes, mean, median, standard deviation, etc. are calculated [9]. NTSC or YIQ color space consists of three components luminance (Y), which represents gray scale information, hue (I) and saturation (Q), which carry information of a signal. This color model is used in television in United State. YCbCr color space is used in digital video. Here, Y contains

978-1-4799-1734-1/15 $31.00 © 2015 IEEE
DOI 10.1109/ICACCE.2015.15

IEEE computer society

luminance information, Cb contains color information between blue component and reference value and Cr contains color information between blue component and reference value [10]. CIELAB color space is designed as perceptual uniformity by international standard. The objective of this color model is that, color differences a human perceives as equal correspond to equal Euclidean distances in CIELAB space. The coordinates of this model are denoted by L*, a*, b* where, L* is the measure of lightness; a* changes the red/green balance; and b* changes the green/blue balance [11]. HSV color space is used by people to select colors from a color wheel or palette. In this color space, H means hue that refers to tint; S means saturation that represents shade; V means value that refers to tone. HSI color space means hue, saturation and intensity. HSI is the best tool for developing image processing algorithm based on color that are natural and perceived by humans. HSV is as similar as HSI, but it basically focuses on color palette [10].

C. Morphological Features

The most commonly used features for classification of fruits is morphological features, i.e. shape and size. Size features are physical dimensional measurements that tells about the appearance of an object. Area, perimeter, major and minor axis lengths, and aspect ratio are commonly used as morphological features. Morphological features are widely used in automatic sorting purpose in industries [9]. The area is a scalar quantity that is the actual number of pixels in the region. Perimeter is a scalar quantity and it is the distance around the boundary of the region. Major axis length is scalar quantity that is specifying the length of major axis (in pixels) of the major axis of the ellipse that has the same normalized second central moment as the region. Minor axis length is also scalar that is specifying the length of minor axis (in pixels) of the minor axis of the ellipse that has the same normalized second central moment as the region. Shape features are measured by roundness ($4\pi \times \text{Area}/\text{Perimeter}^2$), aspect ratio (Major Axis/Minor Axis) and compactness ($\text{Perimeter}^2/\text{Area}$) [12].

D. Texture Features

Texture represents the surface appearance and distribution of elements. It is an important feature in machine vision and predicts surface in form of contrast, roughness, orientation, entropy, etc. A wide variety of techniques have been proposed for describing image texture. In model based approach, a set of parameters is derived from variation of pixel elements that is used to define an image model such as Gaussian Markov Random Field (GMRF), fractional Brownian motion (fBm). In GMRF, conditional probability of a certain pixel depends upon the value of the neighboring pixels, while fBm exploits the self-similarity of texture at varying scales. Statistical based methods are derived by analyzing the spatial distribution of pixel gray level values such as Gray level co-occurrence matrix (GLCM), run-length and auto covariance function [13].

E. Application of Computer Vision in Fruit Grading Sysytem

Sorting of agriculture products automatically is more efficient as compared to the current manual system which is very slow, tedious, labor intensive and error prone. However, there is a need of an automatic sorting system which can identify the agricultural products based on their characteristics. Computer vision has been widely used in fruit grading system. Grading is based upon shape, size, color, intensity and texture of the fruits.

Savakar had graded five different types of fruit images (Apple, Chickoo, Orange, Mango and Sweet Lemon). Total 5000 sample images had been captured, i.e. 1000 images of each type of fruit. The algorithm had been developed by extracting 18 colors and 27 texture features. The color features were calculated by separating RGB (Red, Green and Blue) components. The RGB image was then converted into HSI model and its components were separated. Mean, variance and range were calculated for each RGB and HSI component individually. Texture features were calculated using Gray Level Co-occurrence Matrices (GLCM). The study revealed that classification of chickoo, apple, sweet lemon, orange and mango were 94%, 93%, 93%, 92%, 92% respectively [14]. Deepa proposed a method to evaluate the extracted features used for grading and classification of defected and non-defected fruits. The image database included 200 mosambi fruits. Shape features, intensity features and texture features were calculated. This database was then classified based on PNN and the result showed that shape, intensity and texture features gave 100%, 92%, 96%identification rate respectively [15]. Mustafa et al. presented a novel approach for fruit grading system. In this paper five fruits (apples, bananas, carrots, mangoes and oranges) were analyzed. Shape and color features were extracted from fruit sample images. Morphological features were used to distinguish between almost similar shapes and sizes such as apple and orange or bananas and carrot. Color features were used to remove misclassification between apple and orange or banana and carrot and increase the accuracy to 79-90% [9].

Khojastehnazhand et al. proposed an algorithm for sorting and classification of lemon fruits based upon the color and size in Visual Basic 6. Volume of sample image had been calculated and RGB images were converted into HSI images. HSI values were estimated and these data were stored in a database. During the sorting stage, calculated volume and color are compared with the saved information in the database. The overall system gave 94.04% accuracy [16]. Jackman et al. proposed a system for food quality assessment based on computer vision. In this paper, to extract surface texture feature of food, three approaches had been suggested which were pixel co-occurrence, Run length and Difference Histogram methods. Fourier Transform was also proposed as a method to extract texture feature in term of convolution of sinusoidal waves but it had a limitation that it is used only where a small number of frequencies can reproduce the surface image. Another classical approach Wavelet Transform was also proposed for texture

analysis. An artificial intelligence was provided to computer to analysis the texture which was efficient and robust [17]. Liming et al. presented a system for automatic grading of strawberry. In this RGB image was changed into L*a*b* color model. The size was calculated using the major axis length and color features were extracted from the dominant color model on a* channel. K-means clustering method was used for classification purpose and it gave 90% accuracy for shape classification and 88.8% accuracy for color grading [18].

Alavi proposed a system for grading of Mozafati dates. Total 100 fruits were taken and three quality parameters were calculated such as quantity of juice, size and freshness. These fruits were graded using both fuzzy interference system and human experts for comparison and graded using fuzzy showed 86% conformity result as compared to human experts [19]. Suresha et al presented an automatic grading of apples with the help of support vector machines (SVM). In this, apple images were captured into RGB color model and threshold based segmentation was used to extract the region of interest from the background. The RGB color model was then converted into HSV model and average red and green color components were determined for classification. This classifier gave 100% accuracy in grading [20]. Leemans et al. proposed apples grading method and two types of apple were used-Golden Delicious and Jonagold. Features were extracted from shape, color, texture and stem position. This method for apple external quality grading gave 72% accuracy for Golden Delicious and 78% Jonagold apples. The grading of healthy fruits was better and an error rate decreases to 5 and 10%, respectively [21]. Nozari et al. presented an algorithm for grading of Mozafati dates. Total 100 fruits were taken and classified based on length, width and thickness. These fruits were graded using both ANFIS and human experts for comparison and ANFIS showed 93.5% conformity result as compared to human experts [22]. Razak et al. presented automatic grading of mango using fuzzy analysis. In this size, color and skin features were extracted. Size of mango was determined by calculating area of sample image. Then RGB component was extracted from image and mean of three color components was detected. For shape analysis, edge detection algorithm had been implemented. Fuzzy inference rules were applied for mango grading into different classes and it gave 80% overall accuracy [23].

Kavdir et al. proposed a method of apple grading in which some quality features were extracted such as color, size and defects of apples. These features were gathered and evaluated using the fuzzy system and this gave 89% accuracy in classification [24]. Zhang proposed that a new method to differentiate apple stem-end/calyx from true defect according to their depth information on apple surface. The projector was designed to generate NIR structured light based on position encoding. The NIR structured light was used to sense the depth information of apple surface, then the stem-end/calyx region was identified. It gave an overall 95.24% detection rate [25].

Brief summary of the literature review described above has been tabulated in Table I.

TABLE I. BRIEF SUMMARY OF DIFFERENT FRUITS AND ITS PARAMETERS

Types Of Fruits	Parameters	Efficiency	References
Apples	Color feature	100% Accuracy	Suresha et al. [20]
	Color, defect, shape and size	89% Accuracy	Kavidar et al. [24]
	Stem-end/Calyx	95.24% Accuracy	Zhang et al.[25]
	Color, shape, texture and stem position	73% Accuracy	Leeman et al. [21]
Fruits	Shape, intensity, GLCM texture features	GLCM: 96% Accuracy Shape: 100% Accuracy Intensity: 92% Accuracy	Deepa [15]
	Shape and color	79-90% Accuracy	Mustafa et al. [9]
	Color and texture	Chickoo: 94% Accuracy Apple and Sweet lemon: 93% Accuracy Mango and Orange:92% Accuracy	Savakar [14]
	Texture feature	Guava and Lemon: 96% Accuracy	Khoje et al. [26]
Date	length, width and thickness	93.5% Accuracy	Nozari et al. [22]
	Quantity of juice, size and freshness	86% Accuracy	Alavi [19]

	Length, width, thickness	90% Accuracy	Pourjafar et al. [27]
Strawberry	Size and color features	Color: 88.8% Accuracy Size: 90% Accuracy	Liming et al [18]
Lemon	Color and size features	94.04% Accuracy	Khojastehnazhand et al. [16]
Mango	Size, color and skin	80%. Accuracy	Razak et al. [23]
	Fractal dimension and color	Factral Dimension: 85.19% Accuracy Color: 88.89% Accuracy	Zheng et al. [28]

III. Classifier Models

This section includes the most popular classifier models that include fuzzy logic, artificial neural network, support vector machine and adaptive network-based fuzzy interference system.

A. Fuzzy Logic

Zadeh proposed the theory of fuzzy sets. This theory introduced making of the membership function operations over the range of real numbers [0, 1]. To calculate logic membership functions, new operations were proposed and showed to be a reasonable tool to generalize classic logic. Fuzzy logic is non-parametric classification procedure which can deduce nonlinear classification between input and output, and maintain flexibility in decision making even on complex problems. Fuzzy systems provide the means of translating the expert knowledge of humans about the process in terms of fuzzy (IF–THEN) rules. A fuzzy rule is the basic unit to gain knowledge in fuzzy systems. Like a conventional rule in artificial intelligence, fuzzy system has two components: an 'if' part and a 'then' part which are also known as antecedent and consequent, respectively [19]. Fuzzy is easy to train and requires less number of samples. However, it gives a lower accuracy rate as compared to any other classifier [30].

Kavdir et al. proposed a method of apple grading in color, size and defects of apples are extracted. These features were gathered and evaluated using fuzzy system and this gave 89% accuracy in classification [24]. Date fruits were graded using Fuzzy by extracting some features like quality of juice, size and freshness and it gave 86% accuracy [19]. Rokunuzzman et al. presented an algorithm to classify tomatoes and it gave 84% accuracy rate [29].

B. Artifical Neural Network(ANN)

ANN is a massively parallel distributed information processing system that is made up of artificial neurons has certain performance characteristics resembling biological neurons of the human brain [31]. A neural network is characterized by its architecture that represents in which pattern it is connected between nodes and which method is used to determine the connection weights and the activation function. Feed forward hierarchical architecture is the most commonly used ANN.

A typical three-layered feed-forward neural network is made of a multiple elements called nodes, and connection pathways that link them. When an input signal is received at a neuron, then neuron processes it, and sends an output signal to other interconnected neurons [31].

ANN is suitable to analyze complex problems. However, it has some demerits such as it needs large training set and it is a time consuming process. ANN gave high accuracy for training set but inadequate interpolation for testing data [32].

Mustafa et al. proposed a method to determine the size and ripeness of banana. Shape and color features were extracted. Then ANN was used for classification and it gave accuracy of 79-90% [9]. Rokunuzzman et al. presented an algorithm to classify tomatoes ANN also and it gave 87.5% accuracy rate which was more than the accuracy rate given by Fuzzy logic [29]. Alipasandi et al. proposed a method in which peach was classified and it gave 99.3% accuracy rate [33].

C. Adaptive Neural Fuzzy Interference System

ANFIS was first introduced by Jang which is capable of approximating any real and continuous function on a compact set to any degree of accuracy. ANFIS is a fusion of artificial neural network and fuzzy interference system. In ANFIS, a fuzzy inference system (FIS) has been constructed, then membership function parameters are tuned using either a back propagation algorithm alone or in combination with a least squares method. This adjustment allows to FIS to train from the extracted data. Normally, an ANFIS is made up of the membership layer, the fuzzification layer, the normalization layer, the defuzzification layer and the fifth layer, the output layer [34]. ANFIS is represented faster than ANN. However, when the trained parameters are applied to checking data, total error of ANFIS is larger than ANN

Nozari et al. presented an algorithm for grading of Mozafati dates and classified these based on weight, length, width and thickness. These fruits were graded using both ANFIS and human experts for comparison and ANFIS showed 93.5% accuracy as compared to human experts [22].

D. Support Vector Machine

In machine learning, support vector machines are supervised learning systems based on the statistical learning theory for data

classification and pattern recognition in classification and regression analysis problems. This method is used for classification of both linear and nonlinear data. In support vector machine model, samples are presented as points in space, and mapped such that the samples of the distinct classes are separated by a clear boundary which is as wide as possible. In this method, an optimal boundary of two data sets in a vector space is known as a hyper plane, which is obtained independent of the probabilistic distribution of training vectors in the set. Originally, SVM was introduced for the problems related to two classes only, but we can extend it for multi-class problem using near one plot against one class or one against all strategies. If the space between two data sets is not linearly separable i.e. there is no separating hyper plane to distinguish the classes, a kernel function may be used. The kernel function evaluates the relationships within the data and makes complex divisions in the space [20]. SVM has the highest accuracy rate as compared to other techniques [35].

Suresha et al. proposed an Apple classifier using SVM and it gave 100% accuracy when only color features are compared [20]. Zheng et al. presented an algorithm of mango grading. Fractal dimension and L*a*b* color model are used for grading purposes. Using SVM, fractal dimension and color gave 85.19% and 88.89% accuracy respectively [28].

IV. Conclusion

In this paper, different image processing based classification techniques for fruit grading and sorting system is reviewed. Machine vision based fruit grading systems are capable of replacing labor work for inspection of fruit grading. Out of morphological, color and texture feature, morphological gave highest accuracy rate. In color model, HIS (Hue, Saturation, Intensity) color model is commonly used for grading because it is related to human perception. In machine learning techniques, SVM (Support Vector Machine) gave highest accuracy, but ANFIS (Adaptive Neuro Fuzzy Interference System) showed the best result out of these techniques. Further, Fuzzy gave lowest accuracy rate result, but it is easy to implement.

References

[1] B. Erdenee, T. Ryutaro, G. Tana, "Particular Agricultural Land Cover Classification Case Study Of Tsagaannuur, Mongolia," in IEEE International Geoscience & Remote Sensing Symposium (IGRASS), pp. 3194-3197, 2010.

[2] V. K. Tewari, A. K. Arudra, S. P. Kumar, V. Pandey, N. S. Chandel, "Estimation of plant nitrogen content using digital image processing," in International Commission of Agricultural and Biosystems Engineering, vol. 15(2), pp. 78-86, July 2013.

[3] M. Krishna, Jabert G., "Pest Control in Agriculture Plantation using Image Processing," in IOSR Journal of Electronic and Communication Engineering(IOSR-JECE), vol.6(4), pp. 68-74, June 2013.

[4] G. Bhandane, S. Sharma, V.B. Nerkar, "Early Pest Identification in Agriculture Crop using Image Processing ," in International Journal of Electrical, Electronics and Computer Engineering, vol.2(2), pp. 77-82, 2008.

[5] J. K. Patil, R. Kumar, "Advances in Image Processing for Detection of Plant Diseases," in Journal of Advanced Bioinformatics Applications and Research ISSN, vol2(2), pp.135-141, June 2011.

[6] T. Bronsnan, D. W. Sun, "Improving quality inspection of food products using computer vision- a review," in Journal of food engineering, vol. 61, pp. 3-16, 2004.

[7] Da-Wen Sun, "Computer Vision- an Objective, Rapid and Non-Contact Quality Evaluation Tool for the Food Industry," In Journal of Food Engineering, vol.61, pp.1-2, 2014

[8] H. H. Wang, D. W. Sun, "Evaluation of the Functional Properties of Cheddar Cheese using a Computer Vision Method", in Journal of Food Engineering, vol. 49(1), pp.47–51, 2002.

[9] N. B. A. Mustafa, K. Arumugam, S. K. Ahmed,Z. A. M. Sharrif, "Classification of Fruits using Probabilistic Neural Networks-Improvement using Color Features", IEEE International Conference on TENCON, 2011.

[10] R. C. Gonzalez and R. E. Woods, Digital Image Processing. Prentice Hall, 2nd ed., 2002.

[11] M. W. Schwarz, W. B. Cowan, J. C. Beatty, "An Experimental Comparison of RGB, YIQ, LAB, HSV, and Opponent Color Models", in ACM Transactions on Graphics, vol. 6(2), April 1987

[12] A. Hazra, K. Deb, S. Kundu, P. Hazra, "Shape Oriented Feature Selection for Tomato Plant Identification", in International Journal of Computer Applications Technology and Research, vol 2(4), pp. 449 - 454, 2013.

[13] O.S. Al-Kadi, "Combined Statistical and Model based Texture Features for Improved Image Classification", in Advances in Medical, Signal and Information Processing, MEDSIP, 4th International Conference, pp 1-4, 2008

[14] D. Savakar, "Identification and Classification of Bulk Fruits Images using Artificial Neural Networks," in International Journal of Engineering and Innovative Technology (IJEIT), vol. 1(3), March 2012.

[15] P. Deepa, "A Comparative Analysis of Feature Extraction Methods for Fruit Grading Classifications," in International journal of emerging technologies in computational and applied sciences (IJETCAS) vol.13(138), 2013.

[16] M. Khojastehnazhand, M. Omid, A. Tabatabaeefar, "Development of Lemon Sorting System based on Color and Size," in African Journal of Plant Science vol. 4(4), pp. 122-127, April 2010.

[17] P. Jackman, Da-Wen Sun, "Recent Advances in Image Processing using Image Texture Features for Food Quality assessment," trends in Food Science & Technology vol. 29(1), pp. 35-43, 2013.

[18] Xu Liming and Zhao Yanchao, "Automated strawberry grading system based on image processing," in Computers and Electronics in Agriculture, vol. 71, no. Supplement 1, pp. S32-S39, April 2010.

[19] N. Alavi, "Date Grading using rule based Fuzzy Interference System," in Journal of Agriculture Technology, vol.8 (4), pp. 1243-1254, 2012.

[20] M. Suresha, N.A. Shilpa, B. Soumaya, "Apples Grading based on SVM classifier," in International Journal of Computer Applications on National Conference on Advanced Computing and Communication-NCACC, April 2012.

[21] V. Leemans, H. Mageinb, M.F. Destain, "On-line Fruit Grading according to their External Quality using Machine Vision", Journal of Automation and Emerging Technologies, Belgium. Biosystems Engineering, pp. 397–404, 2002

[22] V. Nozari, M. Mazlomzadeh, , "Date Grading based on some Physical properties," in Journal of Agriculture Technology, vol. 9(7), pp. 1703-1713, 2013.

[23] T. R. B. Razak, M.B. Othman, M. N. A. Bakar, K. A. Ahmed, A. R. Mansor, "Mango Grading using Fuzzy Image Analysis," in International Conference on Agricultural, Environment and Biological Sciences (ICAEBS),2012.

[24] I. Kavdir, D.E. Guyer, "Apple Grading using Fuzzy Logic," in Journal of Agric Turk, vol. 27, pp. 375-382, 2003

[25] C. Zhang , L. Chen, W. Huang , Z. Guo, Q. Wang, "Apple Stem-end/Calyx Identification Using a Speckle-array Encoding Pattern," in ISCP proceeding, 2012.

[26] S. A. Khoje, S. K. Bodhe , A. Adsul, "Automated Skin Defect Identification System for Fruit Grading Based on Discrete Curvelet

Transform" in International Journal of Engineering and Technology (IJET), vol. 5(4), 2013.
[27] M. Pourjafar, M. R. Benam, S. M. Mazloumzadeh, "Using Adaptive Neuro-Fuzzy Interference System for Classify Date Fruits," in Journal of Agriculture Technology, vol. 9(5), pp.1309-1318, 2013.
[28] H. Zheng, H. Lu, "A least-squares support vector machine (LS-SVM) based on fractal analysis and CIELab parameters for the detection of browning degree on mango (Mangifera indica L.)," in Computers and Electronics in Agriculture, vol. 83, pp. 47-51, January 2012.
[29] M. Rokunuzzaman, H. P.W. Jayasuriya, "Development of a Low Cost Machine Vision System for Sorting of Tomatoes", in Agric Eng Int: CIGR Journal, vol.15 (1), March 2013.
[30] H. Atmaca, B. Cetisli, H. S. Yavuz, "The Comparison of Fuzzy Inference Systems and Neural Network Approaches with ANFIS Method for Fuel Consumption Data" in Second International Conference on Electrical and Electronics Engineering ELECO, Bursa-Türkiye, Nov. 2001.
[31] S. Haykin, Neural Networks: A Comprehensive Foundation, 2nd Edition, Prentice-Hall, 1999.
[32] Y. Huang, "Advances in Artificial Neural Networks-Methodological Development and Application", in algorithms ISSSN 1999-4893, vol.2, pp. 973-1007, 2009.
[33] A. Alipasandi, H. Ghaffari and S. Z. Alibeyglu, "Classification of Three Varieties of Peach Fruit Using Artificial Neural Network Assisted with Image Processing Techniques", in International Journal of Agronomy and Plant Production, vol. 4(9), pp. 2179-2186, 2013.
[34] J. S. R. Jang.,"ANFIS: Adaptive-Network-based Fuzzy Inference Systems," IEEE Transactions on Systems, Man, and Cybernetics, vol. 23(3), pp. 665-685, May 1993..
[35] H. Zhibin, W. Xiaohu, L. Hu, D. Jun , "A Comparative Study Of Artificial Neural Network, Adaptive Neuro Fuzzy inference System and Support Vector Machine for Forecasting River Flow in the Semiarid Mountain Region", in Journal of Hydrology, vol.509, pp. 379–386,2014.