Image Classification using Support Vector Machine and Artificial Neural Network

Article *in* International Journal of Information Technology and Computer Science · May 2012

DOI: 10.5815/ijitcs.2012.05.05

5 authors, including:

Thai Hoang Le
VNU-HCM University of Science
80 PUBLICATIONS **798** CITATIONS
SEE PROFILE

Hai Tran
Ho Chi Minh City Pedagogical University
17 PUBLICATIONS **969** CITATIONS
SEE PROFILE

I.J. Information Technology and Computer Science, **2012, 5, 32-38**
Published Online May 2012 in MECS (http://www.mecs-press.org/)
DOI: 10.5815/ijitcs.2012.05.05

Image Classification using Support Vector Machine and Artificial Neural Network

Le Hoang Thai
Computer Science Department, University of Science, Ho Chi Minh City, Vietnam
Email: lhthai@fit.hcmus.edu.vn

Tran Son Hai
Informatics Technology Department, University of Pedagogy, Ho Chi Minh City, Vietnam, member of IACSIT
Email: haits@hcmup.edu.vn

Nguyen Thanh Thuy
University of Technology, Ha Noi City, Vietnam
Email: nguyenthanhthuy@vnu.edu.vn

Abstract—Image classification is one of classical problems of concern in image processing. There are various approaches for solving this problem. The aim of this paper is bring together two areas in which are Artificial Neural Network (ANN) and Support Vector Machine (SVM) applying for image classification. Firstly, we separate the image into many sub-images based on the features of images. Each sub-image is classified into the responsive class by an ANN. Finally, SVM has been compiled all the classify result of ANN. Our proposal classification model has brought together many ANN and one SVM. Let it denote ANN_SVM. ANN_SVM has been applied for Roman numerals recognition application and the precision rate is 86%. The experimental results show the feasibility of our proposal model.

Index Terms—image classification, support vector machine, artificial neural network

1. Introduction

Image classification is one of classical problems of concern in image processing. The goal of image classification is to predict the categories of the input image using its features. There are various approaches for solving this problem such as k nearest neighbor (K-NN), Adaptive boost (Adaboosted), Artificial Neural Network (NN), Support Vector Machine (SVM).

The k-NN classifier, a conventional non-parametric, calculates the distance between the feature vector of the input image (unknown class image) and the feature vector of training image dataset. Then, it assigns the input image to the class among its k-NN, where k is an integer [1].

Adaboosted is a fast classifier based on the set of weak classifiers. A weak classifier based on Haar-Like features could be defined [2] as:

$$h_j = \begin{cases} 1 \ if \quad p_j f(x) < p_j \theta_j \\ 0 \ otherwise \end{cases} \qquad (1)$$

Where x is a sub-window, and θ is a threshold. p_j indicating the direction of the inequality sign.

AdaBoost (Adaptive Boost) is an iterative learning algorithm to create a "strong" classifier using a training dataset and a "weak" learning algorithm. At every iterative step, the "weak" classifier with the minimum classification error is selected.

Artificial Neural Network (ANN), a brain-style computational model, has been used for many applications. Researchers have developed various ANN's structure in accordant with their problem. After the network is trained, it can be used for image classification.

SVM is one of the best known methods in pattern classification and image classification. It is designed to separate of a set of training images two different classes, (x_1, y_1), (x_2, y_2), ..., (x_n, y_n) where x_i in R^d, d-dimensional feature space, and y_i in {-1,+1}, the class label, with i=1..n [1]. SVM builds the optimal separating hyper planes based on a kernel function (K). All images, of which feature vector lies on one side of the hyper plane, are belong to class -1 and the others are belong to class +1.

Besides there are some integrated multi techniques model for classifying such as Multi Artificial Neural Network (MANN) applying for facial expression classification, and Multi Classifier Scheme applying for Adult image classification.

MANN model are shown in the following diagram:

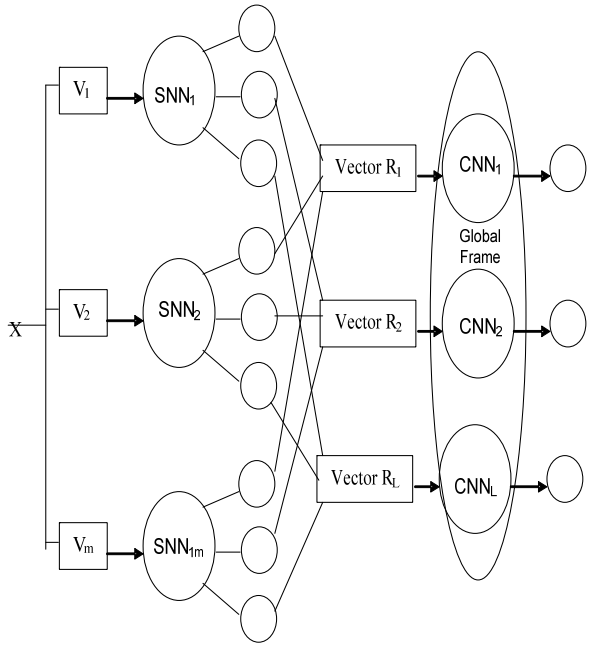

Fig. 1 Multi Artificial Neural Network model [3]

In the above Fig. 1, Multi Artificial Neural Network (MANN) [4], applying for pattern or image classification with parameters (m, L), has m Sub-Neural Network (SNN) and a global frame (GF) consisting L Component Neural Network (CNN). In particular, m is the number of feature vectors of image and L is the number of classes. This model uses many Neural Networks so that the training phrase is complex and long. Besides, it is not suitable in case the number of classes L is high. MANN is the 2-layers classifier model using Neural Network.

Besides multi classifier scheme has just been proposed for Adult image classification with low level feature in 2011[5]. This model contains two-layers classifier. Layer 1 uses Support Vector Machine (SVM) classifier and AdaBoost classifier. Layer 2 is the majority base classifier integrating the classified results of layer 1. Multi Classifier Scheme model is shown in the following diagram:

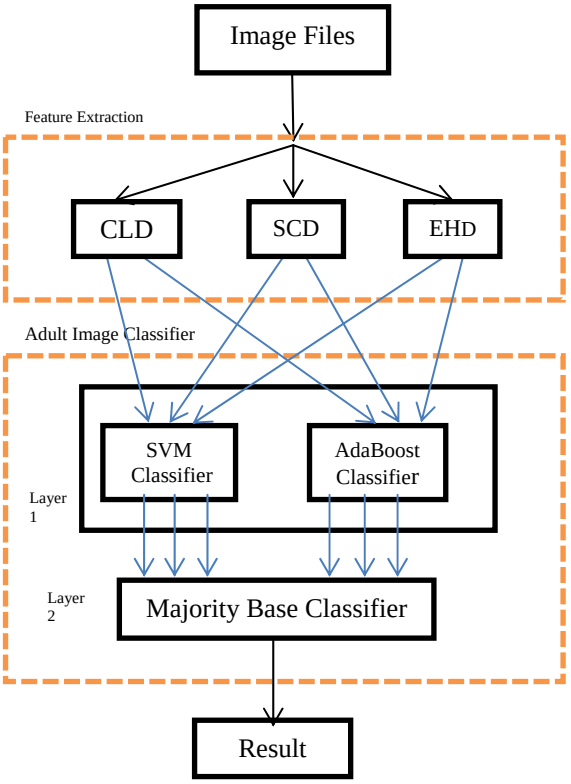

Fig. 2 Multi Classifier Scheme model [5]

In the above Fig. 2, the Multi Classifier Scheme model is two layers classifier. The output of SVM classifier and AdaBoost classifier has been combined by Majority Base Classifier. This experiment has showed that we need to choose the appropriate classifiers for the feature extraction to increase the precision of image classification. On the other hand, the precision of classification system depends on the feature extraction and the classifier.

The remainder of this paper is organized as follows: Section 2 devoted to study of image classification process and its problems. Section 3 provides a detailed exposition of our proposal model ANN_SVM which has been compiled many Artificial Neural Networks and the Support Vector Machine. Section 4 contains a discussion of the experiments and evaluation of Roman numeral recognition application using our proposal model ANN_SVM. Conclusion and future work are given in the final section.

2. Background and Related Work

2.1 The stages of image classification

The main steps in the image classification process are shown in the following diagram:

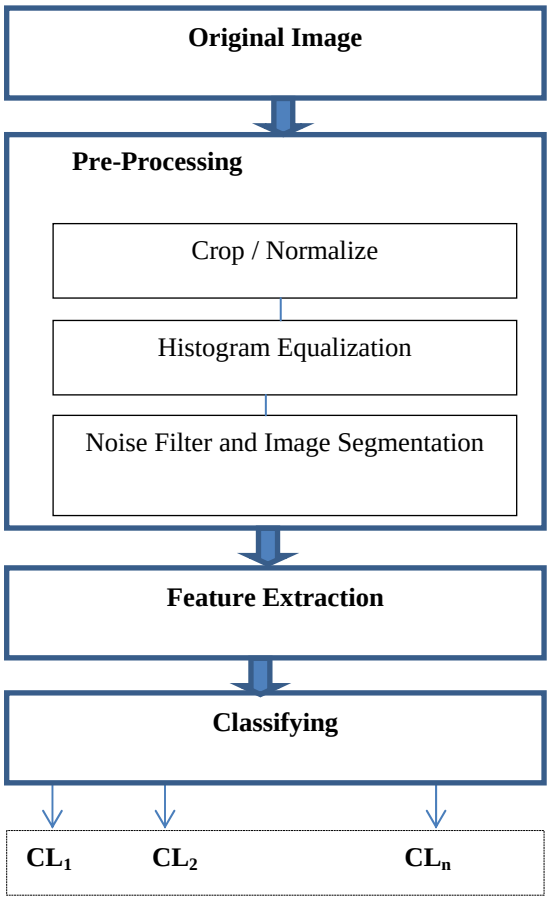

Fig. 3 Image Classification Process

In the Fig. 3 above, CL_1, CL_2, .., CL_n refers to the classes or categories that images are classified into. Step 1, pre-processing, is required before applying any image analysis methods. The images are normalized, performing histogram equalization, applying the noise filter and segmenting. In the step 2, feature extraction, using the suitable transform to decompose an image for example, wavelet, PCA, ICA... The features of images are the input of our classification system. Finally, images are classified into the responsive classes by the suitable techniques (K-NN, NN, SVM ...).

2.2 Image Feature Extraction

The extraction of image features is the fundamental step for image classification. There are various types of features for image classification's aim as follow: color and shape features, statistical features of pixels, and transform coefficient features [6, 7, 8]. In addition, some researchers have used algebraic feature for image recognition and image classification.

The output of image's feature extraction is often a vector or multi vectors. In this research, an image is extracted to k feature vectors based on k representing sub-space.

2.3 Image classification

There are various approaches for image classification. Most of classifiers, such as maximum likelihood, minimum distance, neural network, decision tree, and support vector machine, are making a definitive decision about the land cover class and require a training sample. On the contrary, clustering based algorithm, e.g. K-mean, K-NN or ISODATA, are unsupervised classifier, and fuzzy-set classifier are soft classification providing more information and potentially a more accurate result. Besides, the knowledge based classification, using knowledge and rules from expert, or generating rules from observed data, is becoming attractive. We refer to D. Lu and Q. Weng [1] for complete treatment of image classification approaches.

In recent years, combine of multiple classifiers have received much attention. Some researchers combine NN classifier [9], SVM classifier [10] or AdaBoost classifier for image classification. The aim of this paper is bring together two areas in which are Artificial Neural Network (ANN) and Support Vector Machine (SVM) applying for image classification.

3. A novel combination model (ANN_SVM) apply for image classification

After the images were preprocessed and extracted features, they would present in the large representation space. Thus, they would be projected into the Sub-space in order to analysis easily and reduce dimensions of image's feature.

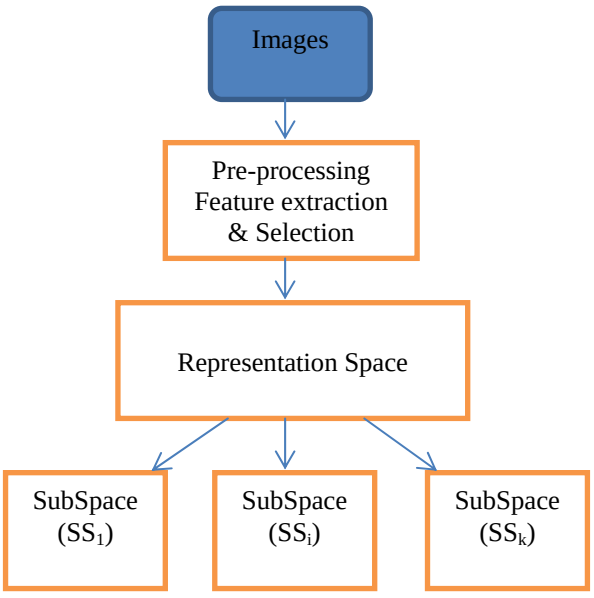

Fig. 4 k Sub-Spaces creation

3.1 Using ANN to classify on each sub-image

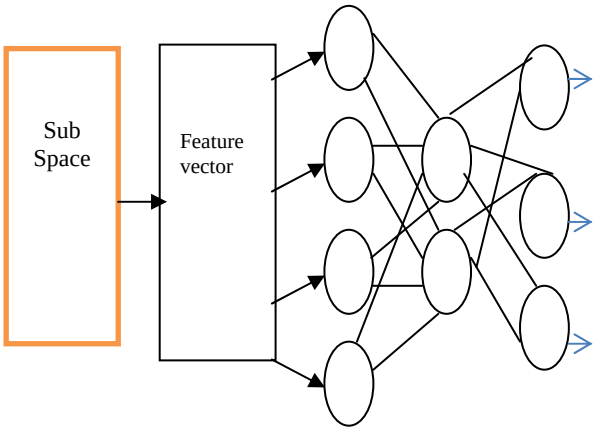

Fig. 5 ANN for classifying

In the above Fig. 5, for each sub-space, an image would be extracted the feature vector. This feature vector is the input of ANN for image classification based on a sub-space. Every ANN has 3 layers: input, hidden and output. The number nodes of input layer are equal to the dimension of feature vector, called in. The number nodes of output are equal to n, the number of classes.

We have k sub-spaces so that there are k classification results of sub-space, called CL_SS_1, CL_SS_2, ..., CL_SS_k . Thus the problem is how to integrate all of those results. The simple integrating way is to calculate the mean value:

$$CL = \frac{1}{k}\sum_{i=1}^{k} CL\_SS_i \qquad (2)$$

Or weighted mean value:

$$CL = \frac{1}{k}\sum_{i=1}^{k} w_i CL\_SS_i \qquad (3)$$

Where w_i is the weight of classification result of sub-space SS_i, and satisfies:

$$\sum_{i=1}^{k} w_i = 1 \qquad (4)$$

The problem is how to identify the optimal weights. In this paper, we suggest to use SVM to identify the suitable weights.

MANN [3, 4] has used Neural Network for identify the weights or importance of the local results. In this research, we suggest that the parameter of the hyper plans of SVM is instead of the weights w_i. Although SVM need to be trained first, the parameter of SVM is adjusted to suitable for the training data in the specific problem.

3.2 Using SVM to aggregate the classify result of all sub-images

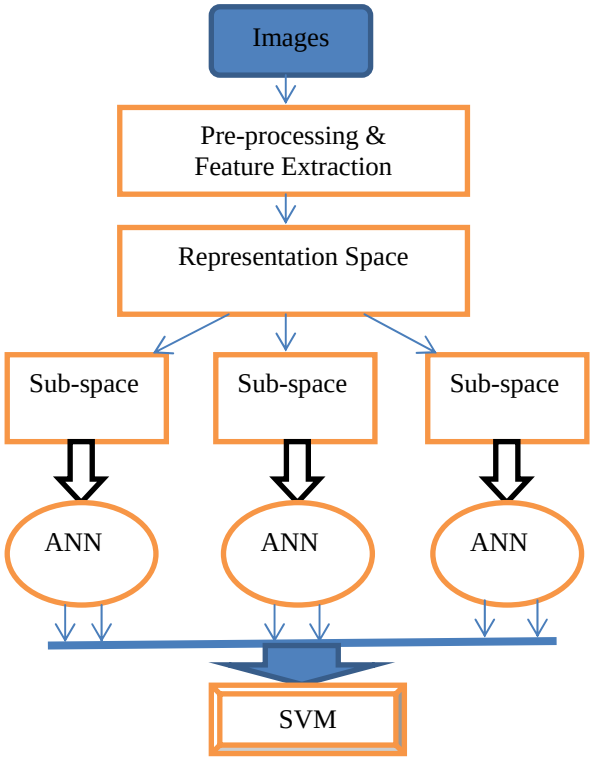

Fig. 6 Image classification using ANN_SVM model

In the above Fig. 6, we use SVM to combine all of ANN's classify results. Here SVM is the solution for identifying the weight of the ANN's result. In our proposal model, there are some parameters as the following:

k: the number of sub-space = the number of ANN

n: the number of classes = the number of output nodes of ANN = the number of hyper plans of SVM

3.3 Using ANN_SVM for Roman numerals recognition application

We develop the system for Roman numerals recognition with k = 3 and n = 10. We have k=3 ANN(s) (corresponding 3 feature vectors) and n=10 classes (corresponding 10 identified classes). It means that a Roman numeral image will be extracted to k=3 feature vectors and need to classify into one of n=10 classes (from I to X).

The input image is preprocessing square image (20x20 pixel), and the output of ANN is the 10-dimensional vector. The ten elements of the output vector is corresponding to the dependence of ten Roman numerals (I, II, III, IV, V, VI, VII, VIII, IX, X). The real value is between 0 (not in the corresponding class) and 1 (in the corresponding class). In this experiment, we just test in ten classes like digital number, but in Roman numerals. We apply our proposal method for Roman numerals classification because the book chapter number

is often Roman numeral. Thus we can apply for fast accessing book chapters in reading book application of mobile device.

The original image is decomposed into a pyramid of image as the following [3]:
4 blocks (16x16 pixels) --> 4 input nodes for ANN_1
16 blocks (8x8 pixels) -->16 input nodes for ANN_2
5 overlap blocks (9x32 pixels) --> 5 input nodes for ANN_3

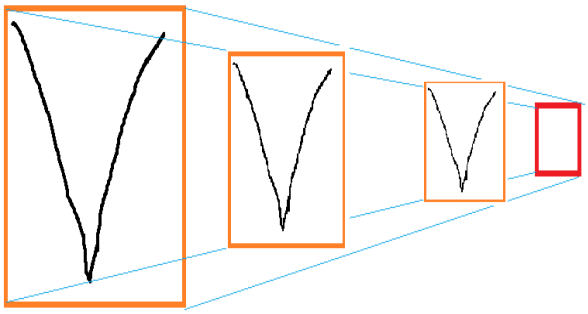

Fig. 7 Roman numerals image decomposition

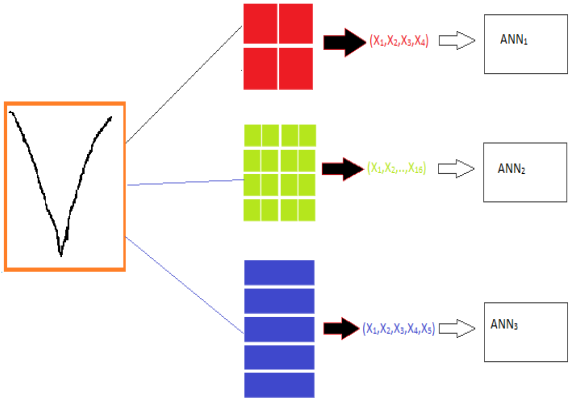

Fig. 8 Classifying on k=3 sub-spaces with k=3 ANN(s)

In the above Fig. 8, the feature vector of decomposing level 1, 4 red blocks, are the input of ANN_1, the feature vector of decomposing level 2, 16 green blocks, are the input of ANN_2, and the feature vector of overlap level , 5 blue blocks, are the input of ANN_3.

In this experiment, k = 3 is the number of feature vectors of the image. Each feature vector will be processed by an ANN. Thus k is also equal to the number of ANN(s). The dimension of ANN's output vector is n, the number of classes. The first node of the ANN's output is the probability of class "I". The second node of the ANN's output is the probability of class "II"… The 10^{th} node of the ANN's output is the probability of class "X". All ANN(s) create k output vectors and every output vector has ten dimensions.

The output of all ANN(s) will be integrated by SVM as follow:

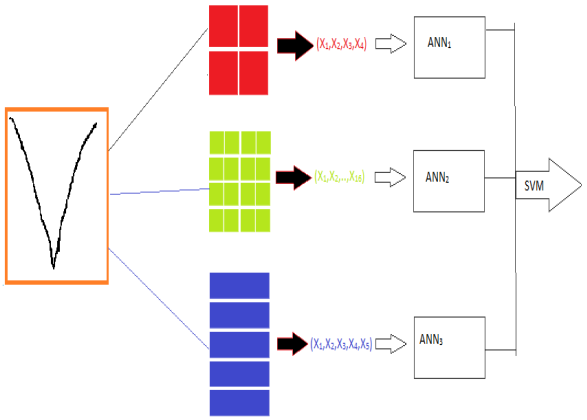

Fig. 9 ANN_SVM model for Roman numerals recognition

In the above Fig. 9, we use ANN_SVM model with k=3 and n=10 to apply for Roman numerals recognition from I to X.

4. Experiment and Analysis

We use Fast Artificial Neural Network (FANN) library, applying for developing the Artificial Neural Network component, and Accord.NET, applying for developing Support Vector Machine (SVM) component, to develop ANN_SVM model. The precision ratio = (correct classifying samples) / (sum of testing samples).

Our training dataset contains 322 matrixes of images of Roman numerals. A Roman numeral image is encoded a shape matrix like below:

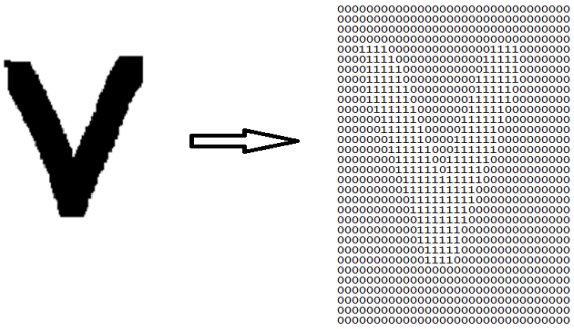

Fig. 10 Roman numeral to shape matrix

The precision recognition is tested directly in our application by drawing the Roman numeral in the lower-left drawing canvas and the result is displayed in the upper-left classification canvas. The right diagram shows the detail of the integration result of SVM, classifying the Roman numeral image as follow:

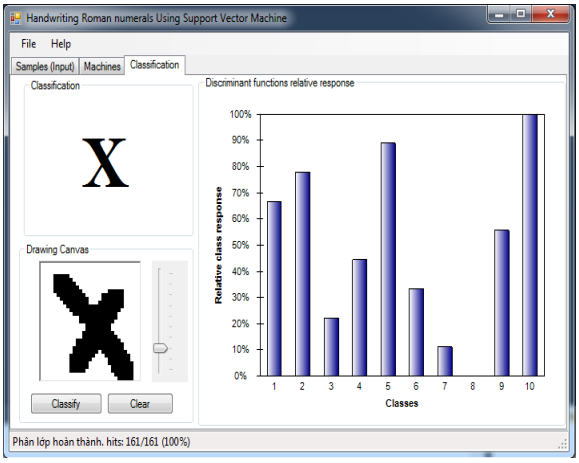

Fig. 11 The GUI of our application demo

The average classification rate is 86% and the detail results of Roman numerals recognition are shown in the table below:

Table 1. Roman numerals classification

Testing Times	Class	Precision
10	I	100%
10	II	90%
10	III	90%
10	IV	80%
10	V	100%
10	VI	80%
10	VII	70%
10	VIII	70%
10	IX	80%
10	X	100%

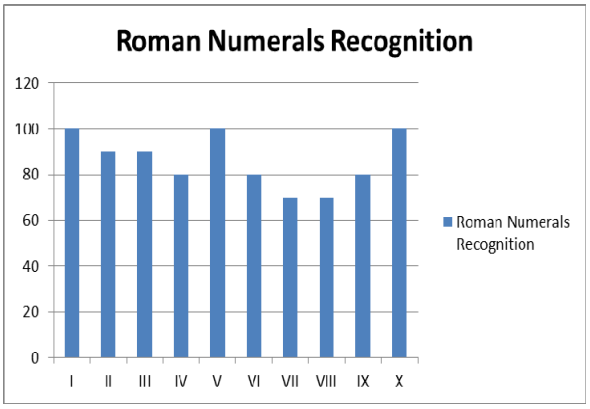

Fig. 12 Roman Numerals Recognition Precision

In the above Fig 12, the precision of class I, V and X is high because these classes do not need to separate for classification. While the classes IV, VI or IX are multi classes and must be separate I and V, V and I, or I and X, before we classify them into the correct classes. In order to improve the precision of classification, we need to develop the relation of multi classes.

5. Conclusion and future work

In this research, we develop an integrated model of ANN and SVM with two parameters (k and n) to apply for image classification, called ANN_SVM. Where

n = the number of classes
= the number of output nodes of an ANN
= the number of hyper plans of SVM
k = the number of an image's feature vectors
= the number of ANN(s).

ANN_SVM is the integrating model of two kinds of soft computing technique in image classification. It is a two layers classifier.

The first layer contains k ANN(s), and this layer give the classifying result based on one by one image's feature vector. The second layer contains a SVM classifier, and its purpose is to integrate all results of the first layer.

ANN_SVM is easy to design and deploy for the specific classification problem. The precision is high, but the performance of processing time need to improve, especially we apply for complex image classification such as facial image. The training time of ANN_SVM is also a problem in the large dataset. Finally, we must redesign and rework all ANN_SVM model when the number of classes increases.

References

[1] D. Lu, Q. WENG, *A survey of image classification methods and techniques for improving classification performance*, International Journal of Remote Sensing, 2007,
Vol. 28, No. 5, pp.823-870.

[2] Qing Chen, *Real-time Vision-based Hand Gesture Recognition Using Haar-like Features*, Instrumentation and Measurement Technology Conference Proceedings, 2007. IMTC 2007. IEEE, 2007, pp.1-6

[3] Thai Hoang Le, *Applications of Artificial Neural Networks to Facial Image Processing*, Artificial Neural Networks - Application, Dr. Chi Leung Patrick Hui (Ed.), ISBN: 978-953-307-188-6, InTech, Available from: http://www.intechopen.com/books/artificial-neural-networks-application/applications-of-artificial-neural-networks-to-facial-image-processing, 2011, pp. 213-240

[4] Thai Le, Phat Tat, Hai Tran, *Facial Expression Classification based on Multi Artificial Neural Network and Two Dimensional Principal Component Analysis*, International Journal of Computer Science Issue (IJCSI), 2011, Vol. 8, No. 3, pp. 19-26.

[5] Mohammadmehdi Bozorgi, Mohd Aizaini Maarof, and Lee Zhi Sam, *Multi-classifier Scheme with Low-Level Visual Feature for Adult Image Classification*, Springer, Communications in Computer and Information Science, 2011, Vol. 181, No. 6, pp. 793-802

[6] D. Zhang, Z.-H. Zhou, and S. Chen, *Diagonal principal component analysis for face recognition*, Pattern Recognition, 2006, Vol. 39, pp. 140-142.

[7] Thai Hoang Le, Nguyen Thai Do Nguyen, Hai Son Tran, *Facial Expression Classification System Integrating Canny, Principal Component Analysis and Artificial Neural Network*, 3rd International Conference on Machine Learning and Computing, ICMLC Proceedings, 2011, Vol. 4, pp. 306-309.

[8] V.H. Nguyen, *Facial Feature Extraction Based on Wavelet Transform*, Artificial Intelligence and Computational Intelligence, Lecture Notes in Computer Science, 2009, Vol. 5855/2009, pp. 330-339, DOI: 10.1007/978-3-642-05253-8_37

[9] Bishop, C.: Pattern Recognition and Machine Learning. Springer Press, 2006

[10] Hao Jiang, Wai-Ki Ching,Zeyu Zheng,*"Kernel Techniques in Support Vector Machines for Classification of Biological Data"*, IJITCS, 2011, Vol.3, No.2, pp.1-8.

[11] Haiyan Li,Guo Lei,Zhang Yufeng,Xinling Shi,Chen Jianhua, *A Novel Method for Grayscale Image Segmentation by Using GIT-PCANN*, IJITCS, 2011, Vol.3, No.5, pp.12-18, DOI:10.5815/ijitcs.2011.05.02

[12] Nilima Kulkarni, *Color Thresholding Method for Image Segmentation of Natural Images*, IJIGSP, 2012, Vol.4, No.1, pp.28-34, DOI: 10.5815/ijigsp.2012.01.04

[13] Le Hoang Thai, Tran Son Hai, *Facial Expression Classification Based on Multi Artificial Neural Network*, International conference on Advance Computing and Applications, 2010, Volume of Extended Abstract, pp. 125-133

[14] Thai Hoang Le., Nguyen Do Thai Nguyen, Hai Son Tran, *Landscape Image of Regional Tourism Classification using Neural Network,* the 3rd International Conference on Communications and Electronics, ICCE 2010. Proceedings

[15] G.M. Foody, A. Mathur, *A relative evaluation of multiclass image classification by support vector machines*, Geoscience and Remote Sensing, IEEE Transactions, 2004, Vol. 42, No. 6, pp.1335-1343

[16] Yang Mingqiang, Kpalma Kidiyo, Ronsin Joseph, *A survey of shape feature extraction techniques*, Pattern Recognition, Peng-Yeng Yin (Ed.), 2008, pp.43-90

First author profile:

Dr Le Hoang Thai received B.S degree and M.S degree in Computer Science from Hanoi University of Technology, Vietnam, in 1995 and 1997. He received Ph.D. degree in Computer Science from Ho Chi Minh University of Sciences, Vietnam, in 2004. Since 1999, he has been a senior lecturer at Faculty of Information Technology, Ho Chi Minh University of Natural Sciences, Vietnam. His research interests include soft computing pattern recognition, image processing, biometric and computer vision. Dr. Le Hoang Thai is the author and co-author over thirty five papers in international journals and international conferences.

Second author profile:

Tran Son Hai is a member of IACSIT and received B.S degree and M.S degree in Ho Chi Minh University of Natural Sciences, Vietnam in 2003 and 2007. From 2007-2010, he has been a senior lecturer at Faculty of Mathematics and Computer Science in University of Pedagogy, Ho Chi Minh city, Vietnam. Since 2010, he has been the dean of Information System department of Informatics Technology Faculty and a member of Science committee of Informatics Technology Faculty. His research interests include soft computing pattern recognition, and computer vision. Mr. Tran Son Hai is co-author of six papers in the international conferences and national conferences.

Third author profile:

Prof. PhD. Nguyen Thanh Thuy received B.S degree in Mathematics, and Ph.D. degree in Computer Science from Hanoi University of Technology, Vietnam, in 1982 and 1987. He has been the professor of Vietnam since 2010. Now he is a Vice Rector of VNU University of Engineering and Technology, Ha Noi city, Vietnam. He majors in Machine Learning, Intelligence Computing and Computer Science.