Advanced Science and Technology Letters
Vol.113 (Art, Culture, Game, Graphics, Broadcasting and Digital Contents 2015), pp.146-150
http://dx.doi.org/10.14257/astl.2015.113.30

Stochastic Classifier Integration Model and Its Application to Satellite Image Classification

Miso Jang[1] and Dong-Chul Park[1],

[1] Dept. of Electronics Engineering, Myong Ji University, Rep. of KOREA,
parkd@mju.ac.kr

Abstract. A classifier called SCIM (Stochastic Classifier Integration Model) is proposed in this paper for achieving higher accuracy in classification tasks. The proposed SCIM utilizes Naive Bayes Classifier as its local classifiers and performs the classification task optimally by combining the decision results from local classifiers while minimizing computational costs required for the classification task. In order to evaluate the performance of the proposed SCIM, experiments on a satellite image classification task is conducted in this paper. The results show that the proposed SCIM outperforms individual classifiers based on individual features and the conventional CIM in terms of classification accuracy.

Keywords: Image Classifier, Probability, Multiple feature vector, Integration model, Naive Bayes

1 Introduction

Rapid increase in the size of digital images has been witnessed recently. Automatic processing of these image contents requires efficient pattern classification techniques. The best strategy in designing a pattern classifier may be the one that utilizes all of the available features efficiently because different features can help to describe objects more precisely. Various approaches to utilize different feature vectors are proposed including Ensemble Classifier [1]. A classifier model, called PFC (Partitioned Feature-based Classifier), that efficiently uses various available features extracted through various feature extraction tools and enhances the classification performance was proposed [2]. In order to improve the classification accuracy, CIM (the classifier integration model) was later proposed as a fusion method for multiple classifiers [3]. The local classifiers used in CIM utilizes either unsupervised learning algorithms including Centroid Neural Networks [4] or supervised learning algorithms. In order to achieve more accurate classification performance and higher classification speed over conventional learning algorithms, Naive Bayes algorithm [5] based on stochastic process is adopted for local classifiers in this paper. Since the Naive Bayes algorithm does not require any iterative procedure for its training process, its training process is quite simple and ~~it~~ requires a small amount of training data to estimate the parameters while it has shown very competitive accuracy in classification results.

The remainder of this paper is organized as follows. Section 2 summarizes the Naive Bayes algorithm and classifier integration model. In Section 3, the stochastic

ISSN: 2287-1233 ASTL

classifier integration model is proposed. In Section 4, the proposed stochastic classifier integration model is applied to a satellite image classification problem in order to evaluate the performance of the proposed algorithm. Concluding remarks are given in Section 5.

2 Related Researches

2.1 The Naive Bayes Classifier

The Naive Bayes Classifier is based on the *Bayes*' theorem of probability [6]. In *Bayes*' theorem, the conditional probability that the given data $\boldsymbol{x}$ belongs to a class k can be calculated from the following equation:

$$P(c_k|\boldsymbol{x}) = P(c_k)\frac{P(x|c_k)}{P(x)}. \tag{1}$$

The Eq. (1) shows that we can decide the optimum class by choosing the class with highest probability among all possible classes, C, which can minimizes the classification error. In doing so, we need to estimate $P(\boldsymbol{x}|c_k)$ and it requires an assumption that any particular value of vector $\boldsymbol{x}$ conditional on c_k is statistically independent on each dimension [6]. This assumption results in simpler calculation cost and efficient data processing. Now, the Naive Bayes classifier can estimate the results of classification as the following equation:

$$k = argmax_k\, P(c_k)\prod_{i=0}^{n} P(x_i|c_k). \tag{2}$$

Note that the assumption of statistically independence in each feature sometimes does not hold in some cases and causes problems in some practical cases [5].

2.2 The Classifier Integration Model

The classifier integration model shown in Fig. 1 considers how each local classifier performs on each class of training data [3]. First, assume that each local classifier, k, is given with the following accuracy table, called Expertise Table, on the classification decision, Q_k :

$$Q_k = \begin{bmatrix} q_{11}^k & \cdots & q_{1N}^k \\ \vdots & \ddots & \vdots \\ q_{M1}^k & \cdots & q_{MN}^k \end{bmatrix}. \tag{3}$$

where q_{ij}^k represents the probability that the local classifier, k, classifies the data of class i with class j through the training stage with the given training data.

When given trained local classifiers $\{C_1, C_2, \cdots, C_N\}$ and the expertise table Q_k assume that our objective is to assign a label among M labels to a data point $\boldsymbol{x}$. The data $\boldsymbol{x}$ consists of N feature vectors and the cluster center $\boldsymbol{P}_j$, $j = 1, 2, \cdots, M$, on i_{th} classifier is given by p_{ij}, $i = 1, 2, \cdots, N$. When a data point $\boldsymbol{x}$ passes through a

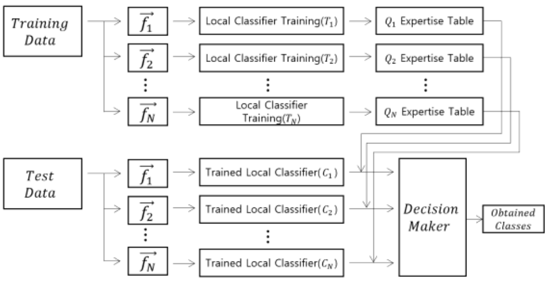

Fig. 1. Classifier Integration Model [3]

local classifier C_i, $i = 1, 2, \cdots, N$, the distance, $D(\boldsymbol{x}, \boldsymbol{P}_j)$, between the data point $\boldsymbol{x}$ and a cluster center $\boldsymbol{P}_j$, $j = 1, 2, \cdots, M$, projected on N different classifier spaces consists of the following N elements :

$$D(\boldsymbol{x}_1, \boldsymbol{p}_{1j}), D(\boldsymbol{x}_2, \boldsymbol{p}_{2j}), \cdots, D(\boldsymbol{x}_N, \boldsymbol{p}_{Nj}). \quad \textbf{(4)}$$

By considering the expertise table defined in Eq. (3) of each classifier, we now compute the distance given as follows:

$$\frac{1}{D(\boldsymbol{x}, \boldsymbol{P}_i)} = \sum_{i=1}^{N} \frac{q_{jj}^{i}}{\|D(\boldsymbol{x}, \boldsymbol{P}_{ij})\|}. \quad \textbf{(5)}$$

The classification results for the data point $\boldsymbol{x}$, $Class(\boldsymbol{x})$, is now obtained by the following equation[3]:

$$Class(\boldsymbol{x}) = argmax_k\, D(\boldsymbol{x}, \boldsymbol{P}_j), \quad \forall j = 1,2, \cdots, M. \quad \textbf{(6)}$$

3 The Stochastic Classifier Integration Model

Conventional CIM calculates the local classification decision probability and uses for parameters in global classification decision. However, the proposed SCIM (Stochastic Classifier Integration Model) calculates the classification probability by using Eq. (1) directly when SCIM adopts Naive Bayes classifier for its local classifier. When the feature value is a continuous value, the proposed SCIM estimates the probability that a feature vector component is classified as its class with the following probability density function:

$$P(x_i|c_k) = \frac{1}{\sqrt{2\pi}\sigma_{c_k}} e^{-\left(x_i - \mu_{c_k}\right)^2 / 2\sigma_{c_k}^2}. \quad (7)$$

where the probability density function is formed during the training stage of local classifiers with the mean, μ_{c_k}, and standard deviation, σ_{c_k}, of each j_{th} class data for each feature vector component x_k.

During training procedure, the probability density function, $P(x_i|c_k)$ for each

(a) Bridge (b) Harbor (c) Pond (d) Urban

Fig. 2. Examples of Satellite Image dataset

dimension of feature vector on each local classifier is first calculated and the $P(\boldsymbol{x}|c_k)$ for each local classifier is found. Once the probability density function is found for each local classifier, the training procedure for the global classifier is terminated. The decision making procedure for a given data is that the feature vectors are pass through corresponding local classifiers and the class for the given data is found by the following equation:

$$Class(\boldsymbol{x}) = argmax_k \frac{1}{N} \sum_i^n \frac{P(x|c_{ik})}{P(x)}. \quad \textbf{(8)}$$

4 Experiments

For evaluating the proposed pattern classifier based on Stochastic Classifier Integration Model, satellite image data sets collected and used for experiments. Fig. 2 shows examples of bridge area, harbor area, pond area, and urban area. Each class consists of 50 images (512×512) with different views resulting in a total of 200 images in the data set. The features employed in this experiments are DCT (Discrete Cosine Transform coefficients) [7], ULBP (Uniform Local Binary Pattern) [8], and CD (Covariance Descriptor) [9]. The dimensions of each feature vector obtained from DCT, ULBP, and CD are 36, 59, and 36, respectively. In order to evaluate and compare the proposed SCIM with conventional classifier models, experiments are carried out in four ways: 1) use of the above three features as individual features for each classifier, 2) use of the entire 131 dimensional feature which is the concatenated version of the above three feature vectors for a conventional classifier model, 3) use of three local classifiers, where each of local classifier uses only one of three feature vectors and conventional CIM classifier combines the three local classifiers, and 4) the proposed SCIM combines the three local classifiers. As can be seen from Fig. 3, the classification accuracies for conventional classifier, CIM, and SCIM are 58%, 61%, and 78.5% on average, respectively.

5 Conclusions

The SCIM (Stochastic Classifier Integration Model) is proposed and applied to a satellite image classification problem. The proposed SCIM utilizes Naive Bayes

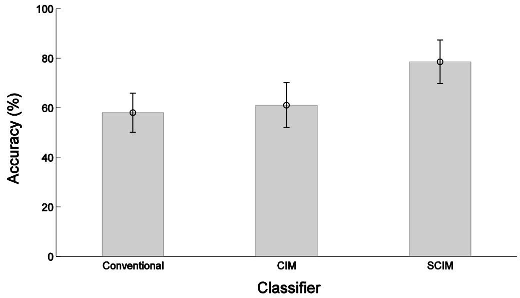

Fig. 3. Classification accuracies among different classifiers

classifier used as its local classifier and adopts the decision procedure used in Naive Bayes classifier scheme. Since Naive Bayes classifier does not require any excessive training procedure commonly used in most of artificial neural networks architecture, the resulting SCIM can yield proper classification decision with very limited computational efforts. The proposed SCIM is applied to a satellite image classification problem for performance evaluation purpose. The features used in experiments are DCT coefficients, ULBP, and CD. The proposed SCIM is compared with the conventional classifier and CIM in terms of training time and classification accuracy. The results show that the proposed SCIM outperforms the conventional classifier and CIM in terms of both training speed and classification accuracy.

Acknowledgments. This work was supported by the IT R&D program of MKE/KEIT (1004019, the development of automotive synchronous Ethernet combined IVN/OVN and safety control system for 1 Gbps class).

References

1. Peng, Y.: A novel ensemble machine learning for robust microarray data classification. Computers in Biology and Medicine. 36, 553--573 (2006)
2. Park, D.C.: Partitioned Feature-based Classier Model. In: Proc. ISSPIT Proc. IEEE Int. Symp. on ISSPIT, pp.412--417 (2009)
3. Park, D.C. et al.: Satellite Image Classification Using a Classier Integration Model. In : 9th IEEE/ACS Int. Conf. on AICCSA, pp. 90--94 (2011)
4. Park D.C.: Centroid Neural Network for Unsupervised Competitive Learning. IEEE Trans on Neural Networks. 11, 520--528 (2000)
5. Lewis, D.: Naive Bayes at forty: The independence assumption in information retrieval. In: Proc. European Conf. Machine Learning, pp.4--15 (1998)
6. Lowd, D., Domingos, P.: Naive Bayes models for probability estimation. In: Proc. ICML (2005)
7. Strang, G.: The discrete cosine transform. SIAM Review 41 (1999)
8. Guo, Z. et al.: A completed modeling of local binary pattern operator for texture classification. IEEE Trans. Image Processing. 19, pp.1657--1663 (2010)
9. Tuzel, O., Porikli, F., Meer, P.: Region covariance: A fast descriptor for detection and classification. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) Computer Vision-ECCV 2006, LNCS, vol. 3952, pp. 589—600, Springer, Heidelberg (2006)

