CONCEPT-ORIENTED SAMPLE IMAGES SELECTION

D. Djordjevic, A. Dorado, and E. Izquierdo

Electronic Engineering Department
Queen Mary, University of London, UK
E-mail: {divna.djordjevic, andres.dorado, ebroul.izquierdo}@elec.qmul.ac.uk

W. Pedrycz

Electrical and Computer Engineering Department
University of Alberta, Edmonton, Canada
E-mail: pedrycz@ece.ualberta.ca

ABSTRACT

In semantic-based image classification, learning concepts for adding knowledge to the image descriptions is an issue of special interest. This learning increases the capabilities for more "intelligent" image processing. The classifier learns by generalizing specific facts present in a number of design samples. Due to the fact that the learning and classification processes run over image descriptions containing part of the image content, selection of training patterns should take into account relationships among those descriptions. Proposed framework uses clustering mechanisms to support the selection of design samples and annotator's hints to reinforce the classifier learning.

1. INTRODUCTION

Learning concepts from features is an ongoing challenge for researchers and practitioners in different communities such as pattern recognition, machine learning and image analysis, among others.

Although the problem of learning concepts has been studied for decades, it is still an open issue. Focusing on the problem Saitta and Bergadano presented an interesting comparative analysis of results from pattern recognition and theoretical machine learning [1]. Furthermore learning concepts is addressed in the context of semantic-based image classification. Concepts are used to add knowledge to the image descriptions linking human and low-level numerical interpretation of the image content. Augmented descriptions are useful to perform more "intelligent" processing on large-scale image databases.

Bhanu and Dong present a framework for learning concepts based on retrieval experience, which combines partially supervised clustering and probabilistic relevance feedback [2]. In contrast, the introduced approach exploits the capability of support vector classifiers to learn from relatively small number of examples. The approach addresses the selection of design samples without overloading the burden of the professional annotator, which is a shortcoming observed in training strategies as the one presented in [3].

Semantic component of the approach casts the classifier into a supervised learning scope. Using inductive learning the classifier can learn by generalizing specific facts present in a number of design samples (or training patterns).

Taking into account that the learning and classification processes run over low-level descriptions containing only some image content information (e.g. color, texture, shape), there is a clear drawback of selecting design samples looking only at randomly selected images.

This work presents a framework that combines unsupervised clustering and designer hints to assist the learning process in order to refine the classifier model. Conversely, the teaching assistance is given by selecting data-driven (clustering outcomes) and designer-driven (hints) samples. A two-class support vector is used to classify new patterns [4].

Next section introduces the problem of learning concepts. Section 3 describes components and addresses some theoretical issues of the proposed framework. Selected experimental results are presented in Section 4. Concluding remarks are given in Section 5.

2. THE PROBLEM OF LEARNING CONCEPTS

Semantic-based classifiers perform the task of using content-based descriptions (feature vectors) to assign certain objects to a given concept (semantic class or category).

The inductive learning process in learning by examples is carried out by presenting declarative knowledge through a number of labeled samples.

In semantic image classification, concept-wise human subjectivity can be introduced by labeling images as either positive or negative samples of a concept depending on perception of their content.

An image is considered to be a positive sample of a given concept when satisfies a criterion defined by a professional annotator. For instance, a picture is a positive sample of a "building image" if it depicts a visible building object.

Normally, design samples are taken from a large-scale database. Times, relaxations in the selection criteria, subjectivity of the beholder, low quality of the picture due to occlusion, shadows, rotation, etc., and amount of available examples are some of the identified drawbacks in training patterns selection.

In addition, choosing samples based on human perception misses the fact that the classifier will work over descriptions with limited domain knowledge.

Then, the problem that can be stated as follows: "How to assist the learning process in the selection of samples for a given concept?" Accordingly, the following framework is proposed.

3. A FRAMEWORK TO ASSIST CONCEPT LEARNING FROM EXAMPLES

3.1. General Overview

Figure 1 illustrates the proposed framework for training the classifier. In a first step clustering mechanisms are used to assist the professional annotator in the selection of image samples. The second step applies reinforce leaning (likewise relevance feedback) to refine the classifier model.

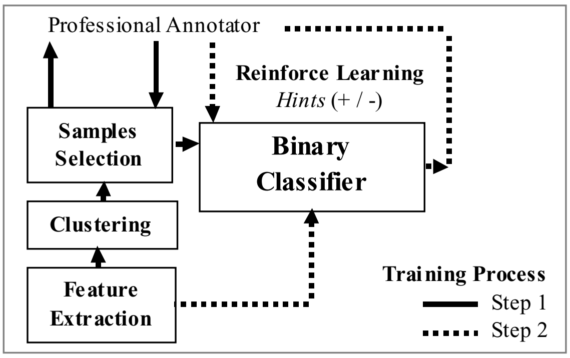

Figure 1. Framework for training the classifier

3.2 Assisting the Training Process through Unsupervised Clustering

As defined in [5], *cluster analysis* is the organization of a collection of patterns into clusters based on similarity. Such a similarity between patterns is quantified or measured using a proximity metric (e.g. Euclidean, Mahalanobis). In this approach clustering outcomes are used to reveal any underlying structure in the feature space and minimize the drawbacks in choosing design samples.

Visual primitives organized into vectors are clustered using the standard fuzzy c-means algorithm [6] based on optimization-minimization of the criterion function:

$$J_m = \sum_{k=1}^{N} \sum_{i=1}^{C} u_{ik}^m \left\| \mathbf{x}_k - \mathbf{v}_i \right\|^2, \quad 1 < m < \infty \tag{1}$$

where u_{ik} is the degree of membership of a low-level feature vector $\mathbf{x}_k$ in the cluster i, $\mathbf{v}_i$ is the p-dimension prototype (center) of the cluster, and $\|\cdot\|$ is any norm expressing the proximity between a given pattern and the corresponding cluster prototype.

As illustrated in Figure 2, the nearest patterns to the cluster prototypes are used as candidates of design samples to train the classifier. The source images for those patterns are presented to a professional annotator who identifies positive and negative examples of the concept.

3.3. Binary Classifier

A support vector classifier (SVM) performs the classification task. SVM achieves good generalization performances over various pattern recognition problems [7]. Besides, as presented in [8], SVM perform well even with very small training data sets.

The class of hyperplanes is denoted as:

$$(\mathbf{w} \cdot \mathbf{x}) + b = 0, \quad w \in \mathbf{R}^p, \quad b \in \mathbf{R} \tag{2}$$

The solution of the classification problem is based on an optimal hyperplane, representing the solution of the following optimization problem:

$$\arg\min \left(\frac{1}{2} \|\mathbf{w}\|^2 \right) \tag{3}$$

under the following condition:

$$y_i((\mathbf{w} \cdot \mathbf{x}_i) + b) \geq 1, \quad i = 1, \ldots, N \tag{4}$$

Where $< \mathbf{x}_i, y_i >$ is a sample and $y_i = \Omega(\mathbf{x}_i) = 1$ if $\mathbf{x}_i$ satisfies the designer-defined criterion regarding to a given concept and $y_i = \Omega(\mathbf{x}_i) = -1$ otherwise.

$\Omega(\cdot)$ denotes the classifier expressed in the simplest case as a function

$$\Omega : \mathbf{R}^p \to \{-1, 1\} \tag{5}$$

This solution should maximize a margin that is the minimal distance from the closest design samples of different classes to the decision surface. The optimization problem (Eq. 3) and (Eq. 4) is a quadratic problem often solved by conversion to Wolfe dual [9].

Introducing kernels enhances the classifier in case of nonlinear feature spaces. The proposed approach uses a Gaussian kernel with a fix kernel width σ:

$$K(\mathbf{x}, \mathbf{x}') = \exp\left({-\|\mathbf{x} - \mathbf{x}'\|^2} \middle/ {2\sigma^2} \right) \tag{6}$$

3.4. Assisting the Training Process through Reinforce Learning

As depicted in Figure 1, the system captures hints of domain knowledge relate to the classification problem.

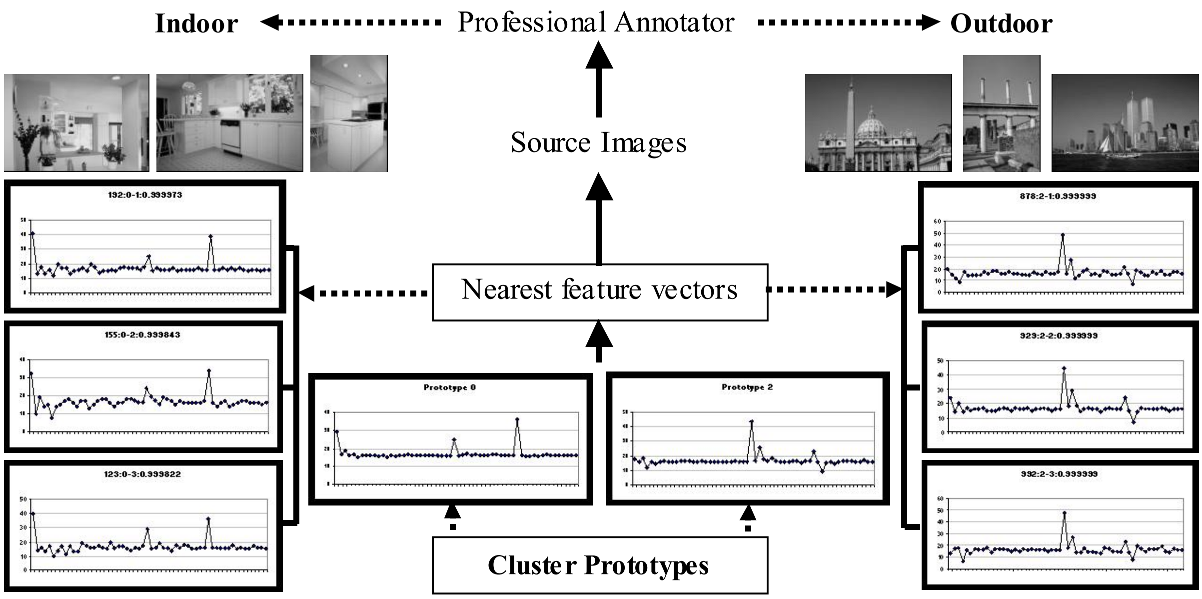

Figure 2. Selecting sample images. Candidates for training the classifier are chosen from the nearest vectors to the cluster prototypes. The selection is based on low-level similarity between vectors (color layout descriptions) and cluster centers.

During the second step of the training process, a professional annotator provides hints indicating to the classifier whether its decisions were either right (positive hint) or wrong (negative hint).

The classifier uses those hints to adjust the boundaries between patterns containing (or not) the concept. These boundaries are defined by the hyperplane based on support vectors.

The idea of this supervised learning step is not to estimate distributions of the known/unknown patterns but to learn the support vectors. These vectors define the optimal nonlinear decision hyperplane and are determined from the known training set.

4. EXPERIMENTAL RESULTS

Experiments were conveyed with imagery selected from Corel stock gallery [ref. corel.com]. More than 1000 pictures were collected from the Corel categories containing indoor images from collections as office, interior, bathroom-kitchen and outdoor images from collections as building, Rome, New York City, etc.

The feature space consists of vectors containing color layout descriptions (cf. [10]). The MPEG-7 color layout is a histogram-based descriptor, and it was set up to 58 coefficients (28 luminance and 15 chrominance coefficients, each). The matching procedures in the test use the basic L2 norm. While aware that based only on low-lever color content, hints provided by professional annotator's would not be able to infer semantics, in our experiment setting we have chosen only one descriptor to study the possible advantage of the approach and later on extend the experiments to spaces built with combined descriptors.

Representative samples of positive and negative images in relation to indoor/outdoor categories can be observed in Figure 3.

The three training approaches summarized in Table 1 are used to assess the performance of the classifier within the proposed framework.

Table 1. Training approaches applied to assess the classifier performance within the framework.

SVM+FCM	SVM classifier assisted with hints provided by a professional annotator governed by clustering (FCM) results during the training phase. Samples are selected from the nearest patterns (see Figure 2 and 3) to the cluster prototypes.
SVM+RL	SVM classifier using only reinforcing learning (RL). The classifier is trained with hints provided by a professional annotator.
SVM+FCM+RL	SVM classifier is trained combining both clustering results and reinforce learning.

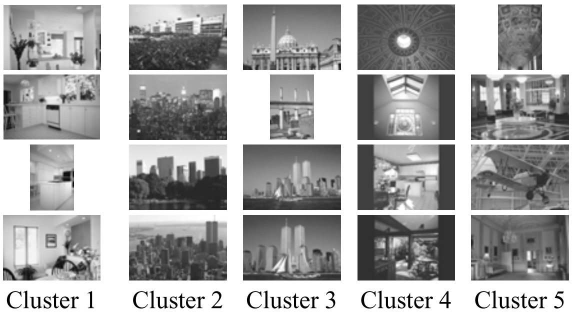

Figure 3. Ranking images by highest memberships in the clusters.

Mean accuracies obtained in the experimental studies are presented in Figure 4. The lowest accuracy is obtained when the support vector classifier learns only from clustering outcomes; the classifier performance is improved when using reinforce learning (RL) and even more accuracy is achieved when selection of samples from clusters is followed by RL.

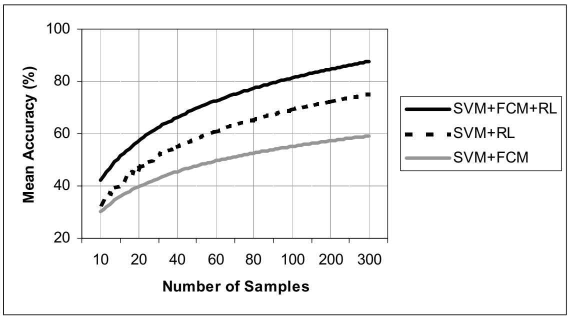

Figure 4. Mean accuracies achieved in the indoor/outdoor classification problem using the training approaches detailed in Table 1.

Accuracy in the first approach (SVM+FCM) decreases rapidly though it is expected due to the sensible reduction on the required supervision. The professional annotator needs only to indicate the class label of each cluster. This lightens the burden of annotation while introducing some noise at the same time.

The second approach (SVM+RL) depends entirely on the images shown to the user. A shortcoming here is the overall subjectivity due to the fact that selection of sample relies completely on the images ignoring any relationship (low-level similarity) between the image descriptions.

The third approach (SVM+FCM+RL), corresponding to the proposed method, shows a higher performance. It has also the advantage of taking into account the underlying low-level structures (revealed by the clusters). It minimizes the required supervision and partially exploits semantic information provided from the professional annotator.

5. CONCLUSIONS

A framework to assist a professional annotator in choosing image samples to train a semantic classifier was presented. The approach uses clustering mechanisms to reveal the underlying structure in training data in order to shift low-level features towards high-level information.

The training process applies reinforce learning to capture hints from the annotator. This training mode reduces the burden of selecting samples randomly as well as improves the quality of the chosen ones taking into account low-level similarity. The reinforcing learning is also a practical way to introduce system's adaptation and can be extended onto the generalization stage in the form of relevance feedback.

6. ACKNOWLEDGMENTS

Support from the Natural Sciences and Engineering Research Council (NSERC) and Canada Research Chair (W. Pedrycz) is gratefully acknowledged.

Part of the research leading to this paper was also done within the framework of the project aceMedia – Integrating knowledge, semantics and content for user-centred intelligent media services (www.acemedia.org).

7. REFERENCES

[1] L. Saitta and F. Bergadano, "Pattern recognition and Valiant's learning framework," *IEEE Trans. Pattern. Analysis Machine Intelligence*, vol. 15, no. 2, pp. 145-155, Feb. 1993.

[2] B. Bhanu and A. Dong, "Concepts learning with fuzzy clustering and relevance feedback," *Engineering Applications of Artificial Intelligence*, vol. 15, pp. 123-138, 2002.

[3] M.R. Boutell, J. Luo, X. Shen, and C. Brown, "Learning multi-label scene classification," *Pattern Recognition*, vol. 37, pp. 1757-1771, 2004.

[4] A.K. Jain, P.W. Duin, and J. Mao, "Statistical Pattern Recognition: A Review," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 4-37, Jan. 2000.

[5] A.K. Jain, M.N. Murty, and P.J. Flynn, "Data Clustering: A Review," *ACM Computing Surveys*, vol. 31, no. 3, pp. 264-323, Sep 1999.

[6] J. Bezdek, *Pattern Recognition With Fuzzy Objective Function Algorithms*, Plenum Press, New York, NY, 1981.

[7] K.-R. Müller, S. Mika, G. Rätsch, K. Tsuda, and B. Schölkopf, "An introduction to kernel-based learning algorithms," *IEEE Neural Networks*, vol. 12, no. 2, pp. 181-201, May 2001.

[8] R. Duin, "Classifiers in almost empty spaces," in *Proc. of 15th Int'l Conf. on Pattern Recognition*, vol. 2, pp. 1-7, Barcelona, Spain, 2000.

[9] S.S. Keerthi, S.K. Shevade, C. Bhattacharyya, and K.R.K. Murthy, "Improvements to Platt's SMO algorithm for SVM classifier design," *Technical report, Dept of CSA*, IISc, Bangalore, India, 1999.

[10] B.S. Manjunath, J.-R. Ohm, V.V. Vasudevan, and A. Yamada, "Color and Texture Descriptors," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 11, no. 6, pp. 703-715, Jun 2001.