INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact factor: 4.295
(Volume 4, Issue 2)
Available online at: www.ijariit.com

Adaptive M-SVM classification model qualified indoor scene images with hybrid feature selection approach

Gagandeep Singh
Shaheed Bhagat Singh State Technical Campus, Ferozepur, Punjab

Sonika Jindal
Shaheed Bhagat Singh State Technical Campus, Ferozepur, Punjab

ABSTRACT

In this thesis, the multi-category dataset has been incorporated with the robust feature descriptor using the scale invariant feature transform (SIFT), SURF and FREAK along with the multi-category enabled support vector machine (mSVM). The multi-category support vector machine (mSVM) has been designed with the iterative phases to make it able to work with the multi-category dataset. The mSVM represents the training samples of main class as the primary class in every iterative phase and all other training samples are categorized as the secondary class for the support vector machine classification. The proposed model is made capable of working with the variations in the indoor scene image dataset, which are noticed in the form of the color, texture, light, image orientation, occlusion and color illuminations. Several experiments have been conducted over the proposed model for the performance evaluation of the indoor scene recognition system in the proposed model. The results of the proposed model have been obtained in the form of the various performance parameters of statistical errors, precision, recall, F1-measure and overall accuracy. The proposed model has clearly outperformed the existing models in the terms of the overall accuracy. The proposed model improvement has been recorded higher than ten percent for all of the evaluated parameters against the existing models based upon SURF, FREAK, etc.

Keywords: *Multi-class Classification, Support Vector Machine, Feature Selection, Indoor Scene Recognition, SIFT.*

1. INTRODUCTION

It is an important task to classify, organize and understand thousands of images efficiently. From an application point of view, scene classification is useful in-content based image retrieval. Scene classification is highly valuable in remote navigation also.

Feature descriptors

- **GIST-** A typical GIST is a computer over a complete image for the scene classification task. It falls in the global image descriptor category.
- **SIFT-** The typical use of SIFT is to match the local regions in two images on the basis of their reconstruction, alignment or other similar. SIFT can be used for the purpose of identification of some specific objects by using BW (Bag of Words) model.
- **HOG-** Histogram of Oriented Gradient (HOG) [10] is used for object-recognition. It is based on computing edge- gradients. Typical HOG works in the sliding window fashion for object detection applications. HOG computes the complete image after dividing it into the smaller cells, called blocks. HOG can be used alongside SVM for feature detection using classification.
- **CENTRIST-** CENTRIST (Census Transform histogram) is a novel visual descriptor, which is more robust to illumination changes, gamma variation etc. as compared to GIST [11] and SIFT [12]. CENTRIST is a histogram of Census Transform (CT) values. CT compares intensity value of a pixel with its neighboring pixels and assigns value 1 or 0 to those pixels. After that, the decimal number corresponding to this sequence of 8 neighboring binary digits is computed and used as CT value of the center pixel. This descriptor retains the local as well as the global structure of the scene. However, there are several limitations of this descriptor. It is not invariant to rotation and scale changes. It also does not consider color information. Further, it cannot be used for precise shape description.

- **Support Vector Machine:** Support Vector Machine (SVM) is a supervised learning based

Classifier

Support vector machines were proposed by Boser et al. in [13]. SVM is supervised machine learning approach specifically designed for pattern matching. SVMs construct a set of hyper-planes that separates the data points into two classes with maximal margin in high dimensional feature space. Mathematically, SVM learns a mapping $\chi \rightarrow \Upsilon$ where x € χ represents the feature vector and y € Υ represents scene category.

2. LITERATURE REVIEW

Espinace, Pablo and Thomas Kollar[1] have worked on Indoor scene recognition by a mobile robot through adaptive object detection. In this paper authors have proposed a new technique to achieve this goal. As a distinguishing feature, authors used common objects, such as Doors or furniture, as a key intermediate representation to recognize indoor scenes. Authors have framed our method as a generative probabilistic hierarchical model, where they have used object category classifiers to associate low level visual features to objects, and contextual relations to associate objects to scenes. The inherent semantic interpretation of common objects allows us to use rich sources of online data to populate the probabilistic terms of our model. In contrast to alternative computer vision-based methods, authors boost performance by exploiting the embedded and dynamic nature of a mobile robot. In particular, they have increased detection accuracy and efficiency by using a 3D range sensor that allows us to implement a focus of attention mechanism based on geometric and structural information.Giannoulis, Dimitrios and Dan Stowell[2] have worked on a project based upon database and challenge for acoustic scene classification and event detection. In this paper authors have introduced a newly-launched public evaluation challenge dealing with two closely related tasks of the field: acoustic scene classification and event detection. Authors gave an overview of the tasks involved; describe the processes of creating the dataset; and define the evaluation metrics. Finally, illustrations on results for both tasks using baseline methods applied on this dataset are presented, accompanied by open-source code.Antanas, Laura and M. Hoffmann[3] have developed a relational kernel-based approach to scene classification. In this paper authors have shown that relational techniques can also improve scene classification. More specifically, we employ a new relational language for learning with kernels, called kLog. With this language authors defined higher-order spatial relations among semantic objects. When applied to a particular image, they characterize a particular object arrangement and provide discriminative cues for the scene category. The kernel allows us to tractably learn from such complex features. Thus, our contribution is a principled and interpretable approach to learn from symbolic relations how to classify scenes in a statistical framework. Gupta, Saurabh, Pablo Arbelaez, and Jitendra Malik[4] have proposed perceptual organization and recognition of indoor scenes from rgb-d images. The authors have addressed the problems of contour detection, bottom-up grouping and semantic segmentation using RGB-D data. They have focused on the challenging setting of cluttered indoor scenes, and evaluate our approach on the recently introduced NYU-Depth V2 (NYUD2) dataset [27]. They have proposed algorithms for object boundary detection and hierarchical segmentation that generalize the gPb − ucm approach of by making effective use of depth information. They have also shown that our system can label each contour with its type (depth, normal or albedo). We also propose a generic method for long-range amodal completion of surfaces and show its effectiveness in grouping. Juneja, Mayank et. al.[5] have worked on blocks that shout: distinctive parts for scene classification. In this paper, authors have proposed a simple, efficient, and effective method to do so. We address this problem by learning parts incrementally, starting from a single part occurrence with an Exemplar SVM. In this manner, additional part instances are discovered and aligned reliably before being considered as training examples. Authors have also proposed entropy-rank curves as a means of evaluating the distinctiveness of parts shareable between categories and use them to select useful parts out of a set of candidates.

3. EXPERIMENTAL DESIGN

Indoor scene classification system involves various steps like feature extraction, feature selection, feature vector generation, training, and classification.

Feature extraction

The following is the feature matching and classification algorithm for matching the extracted indoor scene image with the different images of the same scene, which are taken at different times, from different viewpoints, or by different sensors.

Algorithm 1: Feature Extraction Algorithm

1. Load the 3-D (colored) test image as test object matrix $\mathbf{T_m}$
2. Convert the test image matrix to grayscale image matrix $\mathbf{G_m}$
3. Define the Gaussian Filter **Gf**
4. Apply Gaussian Filter **Gf** on $\mathbf{G_m}$ to produce the de-noised $\mathbf{G_{md}}$
5. Calculate $\mathbf{G_{md}}$ in the front-ground estimation feature $\mathbf{F_{EF}}$
6. Define the dilation object of adequate shape and size $\mathbf{S_{E1}}$
7. Dilate the image $\mathbf{G_{md}}$ with respect to $\mathbf{S_{E1}}$to produce the image object $\mathbf{A_I}$
8. Perform morphological closing of the image object $\mathbf{A_I}$
9. Subtract the $\mathbf{F_{EF}}$ from $\mathbf{G_m}$ to produce **F(BG)**
10. Return the **F(BG)**

Algorithm 2: Classification with an Appearance-based feature descriptor

Read the source image, and Extract the features from the source indoor scene image. Feature descriptor will be the sub-image, and will describe smaller details than the original Target image.

1. Perform pre-processing step to validate the feature descriptor set and arrange all of the feature descriptors in the single feature sets as the training set.
2. Prepare the group data by adding the group IDs corresponding with all of the samples or feature descriptors in the training set.
3. Run SVM training on the feature descriptor training set and return the weight and bias information for all feature descriptors in the training set.
4. Run SVM classifier by submitting the SVM weight and bias data, group data and the testing feature descriptor vector.
5. Return the matching SVM classification information.
6. Evaluate the SVM classification information and return the decision logic.

Simulation Environment

The details of implementation of the proposed model have been discussed under this chapter. Firstly, the proposed model has been developed using the MATLAB simulator. The results have been obtained from various aspects.

4. RESULTS

Performance Parameters

The performance of the proposed indoor scene recognition model has been verified under this chapter. The section explains the performance parameters utilized for the purpose of evaluation of the results proposed a model. The statistical parameters to measure the statistical errors (Type 1 and Type 2) are measured in order to evaluate the overall performance of the proposed model by evaluating the samples by the means of the programming or the manual binary classification. The proposed model evaluation is entirely based on this statistical analysis. The following table explains the significance of type 1 and types 2 statistical errors for the evaluation of the hypothesis.

Table 1: Hypothesis decision parameter entities in type 1 and 2 errors

	Doesn't Have The Condition (Satisfies Null Hypothesis)	Has The Condition (Does Not Satisfy Null Hypothesis)
Tests Negative (Null Accepted)	True Negative TN or n_{00}	False Negative FN or n_{10}
Tests Positive (Null Rejected)	False Positive FP or n_{01}	True Positive TP or n_{11}

Analysis of 100 test cases

The first experiment has been conducted over the 100 test samples, which has been randomly selected out of the given image set. The randomizer module generates the random index containing the hundred image ids, which are acquired from the given dataset. Such randomly selected samples are further processed and analyzed under the proposed model for the result evaluation.

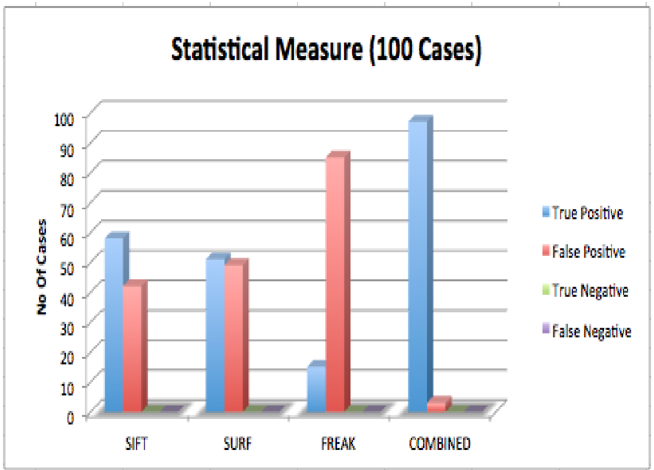

Figure 1: Type 1 and type 2 errors for 100 test cases

The type 1 and type 2 errors have been evaluated from the testing of the input test samples. The graph obtained from the values of the statistical type 1 and 2 errors have been presented in the figure1. The figure1 has been obtained from the hundred testing samples and all of the samples show the equal statistical errors from the first

The table2 shows the values obtained for the figure 1.

Table 2: Type 1 and type 2 errors for 100 test cases

	SIFT	SURF	FREAK	COMBINED
True Positive	58	51	15	97
False Positive	42	49	85	3
True Negative	0	0	0	0
False Negative	0	0	0	0

Table 2 defines the values for the figure 1. Table 2 contains the statistical errors obtained from the 100 randomly selected samples.

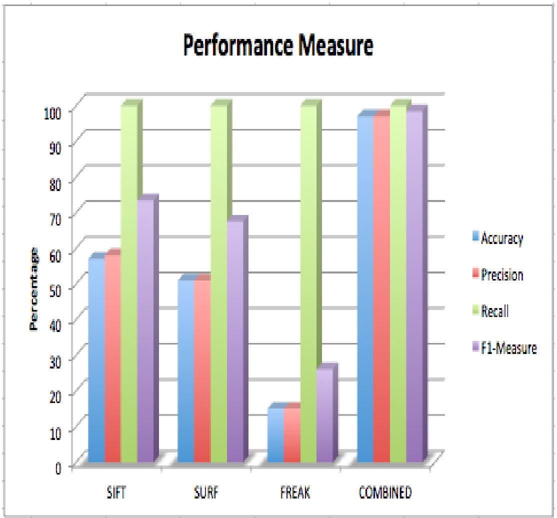

Figure 2: Performance measures for 100 test cases

The figure2 shows the performance measures calculated over the obtained 100 samples. The proposed model and other models based on SIFT, SURF and FREAK have been recorded with the variable performance measure over the given statistical errors.

Table 3: Performance measures for 100 test cases

	SIFT	SURF	FREAK	COMBINED
Accuracy	57	51	15	97
Precision	58	51	15	97
Recall	100	100	100	100
F1-Measure	73.417	67.54967	26.08696	98.477

5. CONCLUSION

The proposed model has been proposed on the basis of the various feature descriptors for the indoor scene recognition. The image feature extraction methods play the major role in the indoor scene recognition by extracting the useful features, which are evaluated to analyze the category of the input query image. The low-level feature descriptor reduces the visual properties of the images, which makes it easier to find the stronger regions within the image data. The stronger regions are inter-matched by using the efficient classification method in order to assess the indoor scene recognition in the given image data. In this thesis, the combined approach of SIFT,SURF, and FREAK along with the multi-class support vector machine (mSVM) has been proposed for the indoor scene recognition. The probabilistic classification with the multi-class support vector machine has been utilized for the robustness of the classification. The proposed model has undergone several experiments and has been found better than the previous option in the terms of precision, recall, f1-measure and overall accuracy and combined model has outperformed the other descriptors with mSVM classification.

6. REFERENCES

[1] Espinace, Pablo, Thomas Kollar, Nicholas Roy, and Alvaro Soto. "Indoor scene recognition by a mobile robot through adaptive object detection." Robotics and Autonomous Systems 61, no. 9 (2013): 932-947.
[2] Giannoulis, Dimitrios, Dan Stowell, Emmanouil Benetos, Mathias Rossignol, Mathieu Lagrange, and Mark D. Plumbley. "A database and challenge for acoustic scene classification and event detection." In Signal Processing Conference (EUSIPCO), 2013 Proceedings of the 21st European, pp. 1-5. IEEE, 2013.
[3] Antanas, Laura, M. Hoffmann, Paolo Frasconi, Tinne Tuytelaars, and Luc De Raedt. "A relational kernel-based approach to scene classification." InApplications of Computer Vision (WACV), 2013 IEEE Workshop on, pp. 133-139. IEEE, 2013.
[4] Gupta, Saurabh, Pablo Arbelaez, and Jitendra Malik. "Perceptual organization and recognition of indoor scenes from rgb-d images." In Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on, pp. 564-571. IEEE, 2013.
[5] Juneja, Mayank, Andrea Vedaldi, C. V. Jawahar, and Andrew Zisserman. "Blocks that shout: Distinctive parts for scene classification." In Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on, pp. 923-930. IEEE, 2013.
[6] Monadjemi, Amir, B. T. Thomas, and Majid Mirmehdi. Experiments on high resolution images towards outdoor scene classification. Technical report, University of Bristol, Department of Computer Science, 2002.
[7] Duda, Richard O., Peter E. Hart, and David G. Stork. Pattern classification. John Wiley & Sons,, 1999.
[8] Fitzpatrick, Paul. "Indoor/outdoor scene classification project." Pattern Recognition and Analysis.
[9] Quattoni, Ariadna, and Antonio Torralba. "Recognizing indoor scenes." In*Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pp. 413-420. IEEE, 2009.
[10] P. Dollar, Z. Tu, P. Perona, S. Belongie, Integral channel features, in: British Machine Vision Conference, 2009.
[11] S. Vasudevan, R. Siegwart, Bayesian space conceptualization and place classification for semantic maps in mobile robotics, Robotics and Autonomous Systems 56 (2008) 522–537.
[12] P. Espinace, T. Kollar, A. Soto, N. Roy, Indoor scene recognition through object detection, in: IEEE International Conference on Robotics and Automation, 2010.
[13] Mesnil, Grégoire, Salah Rifai, Antoine Bordes, Xavier Glorot, Yoshua Bengio, and Pascal Vincent. "Unsupervised and Transfer Learning under Uncertainty-From Object Detections to Scene Categorization." In *ICPRAM*, pp. 345-354. 2013.