

Traffic Sign Recognition System for Imbalanced Dataset

Yildiz Aydin*, Durmus Ozdemir, Gulsah Tumuklu Ozyer

Department of Computer Engineering, Erzincan University, Erzincan, Türkiye.

* Corresponding author. Tel.: 05514190090; email: yciltas@erzincan.edu.tr

Manuscript submitted March 4, 2016; accepted July 4, 2016.

doi: 10.17706/jcp.12.6.543-549

Abstract: In classification problem, the most important factor is training dataset which is effect accuracy rate of classification. However, we encounter with imbalanced data set in real-world applications. In this dataset, the number of images in some classes is rather less than the number of images in other classes. So estimation of classification is tent to majority class and minority classes will be ignored. In this study, an ensemble based method is proposed for increasing accuracy rate of minority class. The results obtained are compared with traditional classifiers (support vector machine (SVM) and k nearest neighbor classifier (KNN)). Bagging based ensemble classifier takes out the issue of inclination toward classifying minority class. As a result, the accuracy of our method result is higher and more efficiency than the other two traditional classifiers.

Key words: Scale invariant feature transform, speed-up robust features, bagging based ensemble, imbalanced dataset, traffic sign recognition.

1. Introduction

Traffic sign recognition system (TSRS) is a very popular issue in nowadays. Especially this system is an essential structure for the future of intelligent vehicle system technologies [1]. Traffic Signs giving the information about the way to the driver so the journey turns into a more secure and easier [2]. It is especially obligatory system for unmanned vehicles which are being worked on it intensively . Traffic sign recognition system is very worthy system because of its advantages mentioned. However, some troubles may be faced in applications of traffic sign recognition system [3]. These troubles are changing weather conditions, corrosion of traffic signs in time, different angles (trees, objects etc.) and different lights or daylight at different angles, wearing off and being damaged of traffic signs (Fig. 1).



Fig. 1. Difficult is recognition of traffic sign.

In the application of TSRS, dataset which was given below are generally used [4].

- German Dataset (German TSR Benchmark (GTSRB))
- Belgium Dataset (KUL Belgium Traffic Signs Data set (KUL Data set))

- Sweden Dataset (Swedish Traffic Signs Data set (STS Data set))
- RUG Dataset (RUG Traffic Sign Image Database (RUG Data set))

Table 1. Standard Traffic Sign Datasets

	GTSRB	KUL	STS	RUG
Number of class	43	100+	7	3
Number of images	50000+	9006+	20000	48

GTSRB, KUL and STS have more picture than RUG, but STS and GTSRB have less classes than KUL.

Applications performed on automatic recognition of traffic signs system basically consist of two stages. These are feature extraction and classification. Features that used in the first stage can be local or global. While local features focus on interested points, global features segment whole image. For histogram of local features extraction, bag of words method must be used. When bag-of-words method is used on imbalanced dataset, this causes purity of clusters will be low, and eventually, no succeeding studies might be done. Imbalanced dataset occurs when the number of samples in one class is more than the other classes. When these datasets are used in applications, classifiers tend to majority class [5].

Traffic sign regions contains interested points because they are visually different regions compared to neighboring areas. Through this property of the sign, advantages of visual attention like elimination of background noise, computation cost can be utilized by using local features. In this study, bagging based ensemble (BBE) classifier that is one of an ensemble method used to solve imbalance learning problem. A strong feature was obtained by combining the histograms of local features. In the study, Histogram of Oriented Gradient feature which is a global feature generally used in TSRS system and classical classifiers (SVM classifier and KNN classifier) was compared with the suggested method. This study consists of four sections. In Section 2 feature extraction was explained in detail. Section 3 includes classifiers and result of the experiment was mentioned in Section 4. Finally, the results were discussed and future studies were indicated.

2. Feature Extraction

The first step in the problems of classification is usually transform the pixel array of image to feature sample which is used to detect objects in the images [6]. For this purpose, Scale-Invariant Feature Transform (SIFT), Speeded up Robust Feature (SURF) and Histogram of Oriented Gradients (HOG) feature were used. In order to extract the histogram of SIFT and SURF features using Bag of Words BoW method is required.

2.1. Scale-Invariant Feature Transform (SIFT)

SIFT [7] has particularly localized characteristic and it is not influenced by image adjusting, turning on an axis, being close to distortion, 3D viewpoint, noise and changes in an figurative clarification. It is comprised of two levels, key point identification and specification, which are emerged by two distinctive sub-levels

- Scale-Space Detection: it consist in the computation of scales and images location via potential detection using a Gaussian function.
- Key Point Localization: In this phase keypoint is selected depending on the stability of their measurement
- Orientation Assignment: Assignment using local image gradients in the area of the key point. This steps provide the invariance of those point during the fourth step, which is the Key Point Descriptor
- Key Point Descriptor: Local gradients are measured and they are changed into a representation which allow for local shape distortion and modification of the light.

2.2. Speeded Up Robust Feature (SURF)

SURF [8] identifies interested points in the scale space using full images analysis via box type convolution filters. This feature is constant to rotation and scaling. SURF works differently from SIFT, because it uses the extreme points of a Hessian matrix.

$$H(x, y, \sigma) \begin{bmatrix} L_{xx}(x, y, \sigma) & L_{xy}(x, y, \sigma) \\ L_{xy}(x, y, \sigma) & L_{yy}(x, y, \sigma) \end{bmatrix} \quad (1)$$

The orientation is determined using Haar wavelet responses inside a circular neighborhood of a $6s$ range where s is the scale of the interest point. The image is then processed analyzing the wavelet responses on a $2s$ Gaussian centered on the interest point via a sliding window approach to determine the dominant orientation: This is achieved distributing the answer along the two axis. Image descriptor are determined by selecting a square region of $20s$ along the dominant orientation, afterwards this area is separated in 4×4 squares which are analyzed via Haar wavelet responses along both axis. This leads to a $4 \times 4 \times 4 = 64$ dimensional descriptor.

2.3. Bag-of-Words Method (BoW)

In BoW model need thinking as a document of an image when it is used in the area of image processing. This model consists of steps such as feature determination, feature illustrator and creating of code book. If we need to describe this method, it may be said that it is a vector which shows the number of occurrence of features in picture.

Local features is commonly used with bag-of-words method. After completing feature illustrator step, a lot of ways like this may be used in creating code book step. But clustering method is commonly used. All feature vectors are split into k cluster with this method. Then, center of cluster is described as instructive cluster in describing code book step.

2.4. Histogram of Oriented Gradients (HoG)

Hog is a feature illustrator which is used for perceiving objects in the fields of image processing and computer vision. It creates a feature illustrator by looking to the ways of gradients in the localized parts of image. So, it was used various fields like number plate recognition, object recognition and vehicle recognition. Its use of recognition of objects application was firstly suggested by Dalal [9] and Sashua [10].

HOG feature divides the image into cells and represents the occurrence numbers of gradients in specified directions for each cell. These cells are in 5×5 and 8×8 pixel size. Gradient magnitude in cells are distributed to related parts in histograms in proportion to its angle obtained by interpolation method.

In order to obtain HoG feature from an image, the following steps should be applied.

- Sobel filter is applied in horizontal and vertical directions
- Horizontal and vertical edges of the image is determined
- Gradients and orientation angle of gradients are determined for horizontal and vertical edges of the image

3. Classifiers

Classification is progressed by two steps, regarding to the utilization of pattern and pattern construction. Pattern construction is identified as a predicted class. Labeling of class are performed for each samples and the detection is according to prediction of each class as a consequence of their feature. The other of step is the utilization of pattern that is using for coming next samples or unknown samples classification.



Fig. 2. Processing of classification.

3.1. Support Vector Machines (SVM)

The system use Support Vector Machines (SVM) [11], which are active learning machines: they identify the hyperplane that separate better the training data and it use it to classify new examples.

This method based on structural risk minimization. It changes the classification parameters for the algorithms making it easier to generalize the complex data. Assume that $S = ((\vec{x}_1, y_1), \dots, (\vec{x}_l, y_l))$ is training sample which can linearly separable and the hyperplane is (\vec{w}, b) . This optimization problem is solved by the following equation.

Minimization:

$$y_i[\vec{w} \cdot \vec{x}_i + b] \geq 1, i = 1, \dots, l \quad (2)$$

In order to determine the distance between parameter and hyperplane, a system of equation involving Wolf duality and Lagrange variation method is used: this method enable the use of kernel and a high dimensional feature space.

Assume that linearly separable training set be $S = ((\vec{x}_1, y_1), \dots, (\vec{x}_l, y_l))$ and The measurement of the \vec{a}^* vector might solve the issues such as quadratic optimization of proceeding parameters.

Maximization:

$$W(\vec{a}) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j=1}^n y_i y_j \alpha_i \alpha_j \vec{x}_i \cdot \vec{x}_j \quad (3)$$

3.2. K Nearest Neighbour Classifiers (KNN)

In an example based method k nearest neighbor classifiers, the sample in the test cluster classifies with respect to the distance between samples in the training cluster. The samples in training cluster are weighted by the distance to the test sample. The closest sample has the maximum weight. In this classifier all samples in the training data are used.

3.3. Bagging Based Ensemble Classifiers (BBE)

Bootstrap add additional training clusters which help the learning system to build a classifier. Bootstrap is occurred by two steps of processes. In the first one multinomial experiments are performed N times to create a size N training set; one experiment belonging to the set is selected and the samples have 1/N probability to hold. During the second step, the process is repeated from a casual number r to N times, the r training sample is then selected and added to the original set. Some of the original tests may not be selected, but some of them may be embedded in the training set. Those bootstrap training steps a help the development of the classifier and the class which fulfill more patterns is considered the optimal output.

4. Experimental Results

This section shows the result of the experiments. The effect of imbalanced database on the performance of global (hog) and local (SIFT, SURF, SIFT+SURF) features was analyzed with 3 classifier: SVM KNN and BBE. When svm and knn is used, some parameters must be determined. Performance of svm depends on

choosing kernel functions and performance of knn depends on choosing k which shows the number of neighbour.

In this study, Belgium Traffic Sign Recognition (BTSR) dataset was used because it has more samples and classes than the other database. BTSR dataset consist of 62 classes divided into two part: one is a training set included 4591 images and the other is test dataset included 2534 images.



Fig. 3. Steps of process of application.

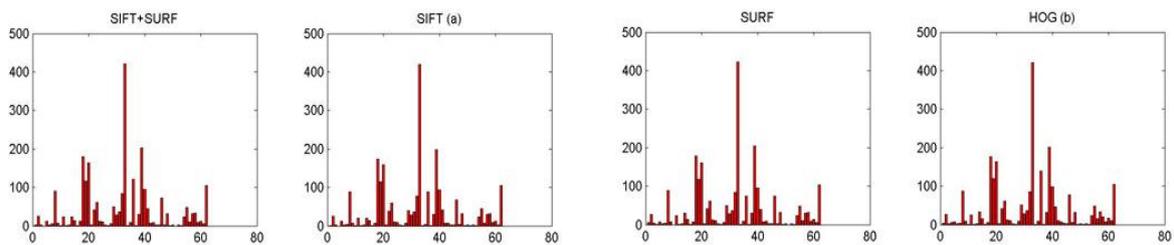


Fig. 4. (a) Classification result of SIFT+SURF feature and BBE classifier, (b) Classification result of SIFT feature and BBE classifier, (c) Classification result of SURF feature and BBE classifier, (d) Classification result of HOG feature and BBE classifier.

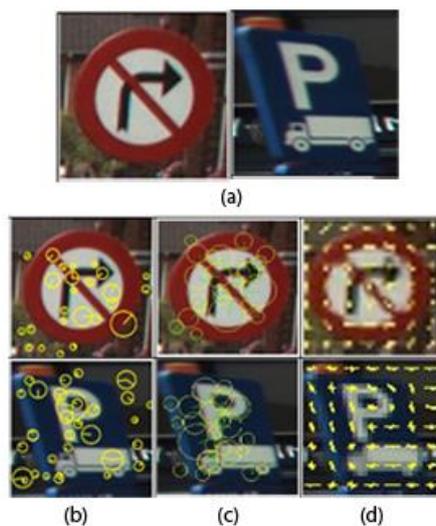


Fig. 5. a) Original images, (b) SIFT keypoints, (c) SURF keypoints, (d) HOG features.

Columns of SIFT and SURF features were combined to obtain SIFT+SURF feature in order to increase the obtained performance on BTSR. In Fig. 4, the highest classification performances obtained was presented.

The highest performance were achieved when SIFT, SURF, SIFT+SURF are used with BBE and hog is used with svm.

As it understands from Table 2 the performance of the experiments, which use local feature, differ from in respect to varying numbers of samples in the classes.

Table 2. Results

Classifier	Feature	Accuracy Rate
KNN	HOG	0.88
KNN	SIFT	0.63
KNN	SURF	0.30
KNN	SURF+SIFT	0.34
SVM	HOG	0.95
SVM	SIFT	0.36
SVM	SURF	0.68
SVM	SURF+SIFT	0.81
BBE	HOG	0.94
BBE	SIFT	0.88
BBE	SURF	0.88
BBE	SURF+SIFT	0.91

5. Conclusion

In this research show that Hog features and svm classifier methods gave the similar results with [7]. But different results were obtained from [4] in SIFT features and SVM classifier method. Because HOG feature generates features by dividing the whole image. Therefore, it does not affected by the sample numbers in database. However in TSR system, local descriptor should be used if taking advantage of visual attention is desired, performance of this local descriptors changed depending on the varying sample numbers in classical classifiers (SVM, KNN). It caused a bias problem by eliminating the samples in minority class. Through BBE classifier used to solve this bias problem, performance of the classifier was improved by classifying the minority class properly. Also SIFT+SURF feature robust than sift so proposed method's accuracy rate increased. In the future studies, we plan to develop a local feature based application more successful than the global feature based applications developed for TSR system.

References

- [1] Mathias, M., Timofte, R., Benenson, R., & Van Gool, L. (2013, August). Traffic sign recognition — How far are we from the solution? *Proceedings of the 2013 International Joint Conference on Neural Networks* (pp. 1-8).
- [2] Won, W. J., Lee, M., & Son, J. W. (2008, June). Implementation of road traffic signs detection based on saliency map model. *Proceedings of 2008 IEEE Symposium on Intelligent Vehicles*, (pp. 542-547). IEEE.
- [3] Hua, X., Zhua, X., Lia, D., & Li, H. (2010). Traffic sign recognition using Scale invariant feature transform and SVM. *Proceedings of A Special Joint Symposium of ISPRS Technical Commission IV & AutoCarto in conjunction with ASPRS/CaGIS Fall Specialty Conference November* (pp.15-19).
- [4] Mogelmoose, A., Trivedi, M. M., & Moeslund, T. B. (2012). Vision-based traffic sign detection and analysis for intelligent driver assistance systems: Perspectives and survey. *IEEE Transactions on Intelligent Transportation Systems*, 13(4), 1484-1497.
- [5] Kotsiantis, S., Kanellopoulos, D., & Pintelas, P. (2006). Handling imbalanced datasets: A review. *GESTS International Transactions on Computer Science and Engineering*, 30(1), 25-36.
- [6] Chavez, A. J. (2012). Image classification with dense SIFT sampling: An exploration of optimal parameters. Doctoral dissertation, Kansas State University.

- [7] Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2), 91-110.
- [8] Bay, H., Tuytelaars, T., & Van Gool, L. (2006). Surf: Speeded up robust features. *Proceedings of ECCV 2006 on Computer Vision* (pp. 404-417). Springer Berlin Heidelberg.
- [9] Dalal, N., & Triggs, B. (2005, June). Histograms of oriented gradients for human detection. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition: Vol. 1* (pp. 886-893). IEEE.
- [10] Shashua, A., Gdalyahu, Y., & Hayun, G. (2004, June). Pedestrian detection for driving assistance systems: Single-frame classification and system level performance *Proceedings of 2004 IEEE Symposium on Intelligent Vehicles* (pp. 1-6). IEEE.
- [11] Vapnik, V. (1995). *The Nature of Statistical Learning Theory*. Springer, New York.



Yildiz Aydın was born in Erzincan, Turkey in 1988. She received the B.Sc. degree in computer engineering from the University of Suleyman Demirel, Isparta, Turkey, in 2011. She is currently an M.S. student at Department of Computer Engineering of Ataturk University

In 2011, she joined the Department of Information Technologies, University of Erzincan, as an expert. His current research interests include pattern recognition, artificial intelligent and images processing.



Gulsah Tumuklu Ozyer received her B.Sc. degree from Erciyes University in Computer Engineering Department in 2001. She was a visitor researcher at Penn State University, USA in James Z. Wang Research Group between February 2007 and March 2008. She is currently a Ph.D. student at the Department of Computer Engineering of Middle East Technical University, Turkey. Her research interests include computer vision, image processing and pattern recognition.



Durmus Ozdemir was born in Kutahya, Turkey in 1981. He has completed his Ph.D. degree in the Department of Computer & Instructional Technology from Ataturk University in 2015, Erzurum, Turkey. He received the M.S. degree in computer engineering from Karadeniz Technical University in 2009, Trabzon, Turkey and He completed B.Sc. degree in computer engineering from European University of Lefke in 2004, Turkish Republic of Northern Cyprus. He is currently an Assistant Professor at the Department of Computer Engineering in Erzincan University, Erzincan, Turkey. Dr. Ozdemir's research has been in various areas of robotics systems. His research interests are robotics in education, autonomous and mobile robots, robot programming and pattern recognition.