Image Segmentation and Object Recognition Using Machine Learning

Ashima Sood[1(✉)] and Sahil Sharma[2]

[1] Punjabi University Regional Centre, Mohali, India
soodashima91@gmail.com
[2] Thapar University, Patiala, India
sahil.sharma@thapar.edu

Abstract. The digital image processing is a fast pace growing field which requires pre-processing of the images before their actual usage and experimentation. Image Segmentation and Object Recognition is one of the phase which requires the prior information about the object to give the extracted attributes as output of the images. Here, in the present research work we have already some outdoor segmented images which have been identified as to which category of classification they belong. This work has been accelerated with the help of machine learning that is a novel technique to identify the objects with outstanding accuracy. The machine learning algorithms have aided in this process of identifying the objects like sky, brick, cement, grass etc. This collaboration of image segmentation with machine learning has proved to be accessible in large datasets where after segmentation images can classify themselves into a category provided it has attributes of the images.

Keywords: Image segmentation · Object recognition · Machine learning · Classification · Ensemble learning

1 Introduction

The immeasurable growth in the digital images has given created a new trend of technology in the area of digital image processing and image segmentation is one of the indispensable step in the field of images, videos and other computer applications which has been introduced in the next section. There has been many image segmentation techniques evolved so far which has resulted in many improvements and advancements in this field. These segmented images need to identified so that the work of relating the images and recognizing them is made easier. This comes under the area of object recognition which means finding the images with the help of various methods used so far like binary pattern matching, gradients, edges, Haar wavelets and other feature based techniques. The current work deals with object recognition of already segmented images using machine learning models where images are identified on the basis of attributes like intensity mean, region pixel count, saturation mean, hue mean and many more. On the basis of these attributes a class distribution of randomly selected 7 outdoor images has been done and the objective is to perform practical experimentation to show how accurately the machine learning models

A. Unal et al. (Eds.): SmartCom 2016, CCIS 628, pp. 204–210, 2016.
DOI: 10.1007/978-981-10-3433-6_25

can identify these images with the help of the attributes given in the dataset. This work requires the detail study of image segmentation and object recognition area where we can understand how the images have been pre segmented with the help of the various key features given. The variety of algorithms in machine learning give a vast scope improvement in this field. The models are further subjected to ensemble learning which gives diversification of models as models give better results together with more data points. It is like polling in the parliament where different views can give the better decision patterns rather than one taking the decision as a whole. There are different approaches of ensemble learning such as bagging, boosting, voting based techniques, combination rules, AdaBoost, decision templates, hierarchical mixture of experts and stack generalization [1]. In the present work average of accuracies of models has been taken for the ensemble learning approach. This has proved the consistency of models as the accuracies of respective models chosen best fit for ensemble learning has shown the as better results as it was with the individual models. This technique has been clearly shown in further section in the methodology and results section.

The remainder section follows, Sect. 2 which gives the introduction of image segmentation; Sect. 3 gives the basic idea of object recognition. The Sect. 4 has discussed the related work which has been done previously in this field. The Sect. 5 covers the methodology part in which the subsections secondary dataset, feature reduction, setting target value, machine learning, K-Fold validation and ensemble learning is discussed. At last the work has been concluded with the future scope of the work.

2 Image Segmentation

Image segmentation is breaking down the image into segments in order to obtain area of interest and also differentiating these segments from the background. In this, the objects of the image are separated according to the property of those objects. Image Segmentation is a pre-processing technique used in digital image processing. It is the method in which an image is segregated in different regions of concern based on properties such as texture, intensity, or colour [2]. The extraction of object and replacing it with the some other background is very common method of segmentation and very often used in movie industry. The process of segmentation depends on the extent of problem to be solved, therefore the segmentation process is stopped once the area of interest have been isolated [3]. With the improvement in image processing field the colour image segmentation techniques are more into concern which can be consider as the extension of gray scale segmentation process. But many gray image methods of segmentation are not directly applied to colour image.

Image segmentation is done by using various algorithms and these algorithms are broadly classified on the basis of two basic approaches of intensity values:

Discontinuity: In this segmentation is done on the abrupt changes in an image like edges.
Similarity: This category of partitioning is done on the regions of image which define a specific set of predefined properties.

3 Object Recognition

The object recognition is closely related to segmentation process. They are mutually dependent on each other. Object recognition is used to extract objects from the image in the real world. We have been given the set of objects and the main task is to set labels to these objects in an image. This task of labeling the object is done with the help of various features like color, size, shape, pixel intensity etc. The dataset used contains the objects which have been recognized based on the different features and the study of this concept comes under this area of how features are detected and possibilities are made about the rightly identified object from the image. There is a systematic approach of objet recognition which involves various parameters which are discussed in Fig. 1 below.

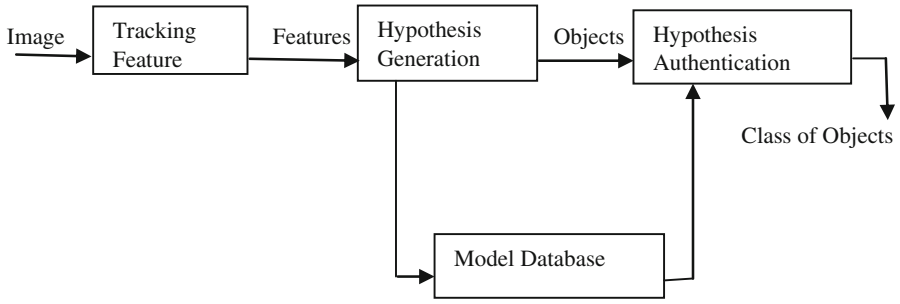

Fig. 1. Object recognition process

In the above system of block diagram various constituents play their specific roles for recognizing the object. The features are attributes of the objects which help in describing the object for example color, shape, boundary, size of the object. Tracking feature is like feature detector which identify the features of the object which further helps in forming the hypothesis. The hypothesis generation phase gives the likelihood to the objects and thus reduces the search space for the system [4]. After this, the model database in which different models of object recognition are present eliminates the objects which are of less likelihood and refines the hypothesis for the verification of the same. At the end after verification of hypothesis using models, the distribution class of objects is formed which is the result of object recognition from the scene of real world.

4 Related Work

The previous work done in image segmentation has used many techniques and algorithms with its diverse applications and explanation with examples. The concerned area is of different image segmentation methods based on the technology of histogram and segmentation on neighborhood [5]. The color images somehow use the enhance technique of gray scale images, however the collaboration of region based algorithm and watershed algorithm has been used for color image segmentation. This combined method of color image segmentation has reduced the time complexity of seed growing

method unlike traditional methods [6]. Likewise many other techniques of image segmentation has been discussed so far like Edge-based image segmentation using neutral networks machine learning which is implemented in MATLAB and C++ with the help of optimal border selection [4]. The image segmentation and object recognition work jointly on the images where selecting and recognizing the images require generative models. These models require features or attributes which are like active parameters to make hypothesis of the objects obtained. In case of missing values like region, area and features a well-known tool has been used known as EM algorithm. This Expectation-Minimization algorithm is used to generate the maximum likelihood of parameters for the probabilistic models. This maximum likelihood is used to figure out the features and obtain the candidate objects out of it [7]. Previously, for pattern recognition, machine learning techniques has been used. The detection of images related to building material is to be done. The approach used in this work is mainly of two steps i.e. feature extraction and classification. After the task of detection feature set, the vectors of features are encapsulated and fed into the classifier for classification of objects and detection task is performed. This has been done with the help of machine learning algorithms which has been termed as classifiers namely Radial Basis Function, Support Vector Machines and Multilayer Perceptron. After using these classifiers, the best performing model has been chosen [8]. This has been the novel approach of how the machine learning algorithms can accurately perform object recognition. Different machine learning algorithms have proved to be successful in various research fields and are able to perform well in different scenarios [9–13].

5 Methodology

The proposed method in this work is the result of various machine learning models which have been implemented on the segmented images for object recognition. The class distribution of objects is done in prior in the dataset on the basis of some features. These features can be reduced to make the dataset with instances and attributes more appropriate which is helpful in large dataset i.e. removing the redundancy from the data. Once the feature reduction is done, the classification of objects is done to make it relevant for machine learning algorithms. The practical implementation is supported by R environment. Once the implantation is done the accuracies of the respective models is generated. The consistency check is made on these accuracies to make the work more authenticated and validated. At the end, ensemble learning is done to take the average result of the output obtained.

5.1 Secondary Dataset

The dataset collected is of secondary type involving the data of image segmentation. It includes the instances of seven outdoor images which are drawn randomly. These images are already hand segmented so that classification of images can be done. In total there are 2310 instances and 19 attributes giving the vast features of the images obtained. The classes of objects are distributed as 1 = brick face, 2 = sky, 3 = foliage, 4 = cement, 5 = window, 6 = path, 7 = grass [14]. The dataset is of classification type. The attributes

of this dataset represent the features of the images collectively like vedge-mean, hedge-sd, value-mean, saturation-mean, hue-mean, exblue-mean, exgreen-mean, exred-mean, rawred-mean, rawblue mean, rawgreen-mean and many other which has been used as attributes in the dataset.

5.2 Feature Engineering

In the process of feature engineering we opted for reducing the number of features using Pearson Correlation technique. So, the original 19 features has been reduced to 14 features for improving the time complexity for building the machine learning models. The 19 features also contain one numeric constant feature which does not affect the overall performance of the output so that feature can be ignored. Table 1 below shows the accuracy comparisons when models were built with 18 features and 14 features respectively.

Table 1. Feature engineering

S. No.	Model name	18 features accuracy	14 features accuracy
1.	CART	87.52381	89.14286
2.	Random Forest	95.76190	94.80952
3.	Gaussian SVM	86.76190	86.47619
4.	Linear Model	91.76190	91.47619

5.3 Training and Testing

After the feature engineering phase, it was clear that the models can be trained and tested just using 14 features selected using Pearson Correlation method. Thus machine learning models were tested for their performances. Total of 13 models were trained to do the correct classification based on the training and testing data set. Table 2 shows the accuracy, confusion matrix error and the time taken to build each model.

Table 2. Results of training and testing

S. No.	Model name	Method, Package	Accuracy	Confusion matrix error	Time taken (s)
1.	CART	rpart, rpart	89.14286	10.85	0.04
2.	Conditional Inference Tree	cpart, party	53.00000	47.00	0.05
3.	Random Forest	rf, randomForest	94.80952	5.19	0.33
4.	Conditional Inference Random Forest	cforest, party	90.38095	9.61	2.54
5.	Gaussian SVM	rbfdot, kernlab	86.47619	13.52	0.43
6.	Generalized Linear Model	lm, glm	91.47619	8.52	1.60
7.	Polynomial SVM	polydot, kernlab	91.00000	9.00	0.42
8.	Linear SVM	vanilladot, kernlab	91.00000	9.00	0.39
9.	Hyperbolic Tangent SVM	tanhdot, kernlab	42.09524	57.90	0.43
10.	Laplacian SVM	laplacedot, kernlab	86.61905	13.38	0.48
11.	Bessel SVM	besseldot, kernlab	85.71429	14.28	0.61
12.	ANOVA RBF SVM	anovadot, kernlab	93.04762	6.95	0.79
13.	Spline SVM	splinedot, kernlab	88.90476	11.09	1.46

5.4 Ensemble Learning

Ensemble learning means mixing up the results of multiple machine learning models to get more reliable outputs. This in a way gives versatility to the prediction modelling. Hence, top 3 models namely Random Forest, ANOVA RBF SVM and Generalized Linear Model from the training and testing table based on their accuracy were selected for the ensemble modelling using average method. The accuracy obtained by using average ensemble method is 93.11111 which is neither maximum nor minimum making performance more reliable.

5.5 Cross Validation

This phase has its own importance for checking the consistency of performances given by selected machine learning models. A 5-fold cross validation has been applied using the three top models mentioned in previous section along with the consistency of ensemble model. Figure 2 below shows the accuracy of 5 folded cross validation in the form of histogram.

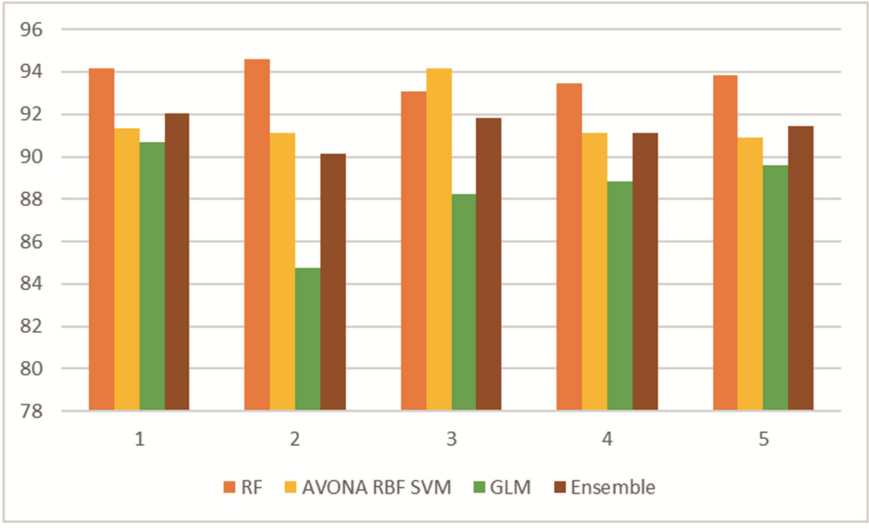

Fig. 2. 5-fold cross validation

6 Conclusion and Future Scope

The result of implementation shows that for the object recognition in the discussed problem can be best solved with the ensemble model of Random Forest, ANOVA RBF SVM and Generalized Linear Model, giving the accuracy of 93.11111% as discussed in Sect. 5. This work can be extended on large datasets with millions of images of real

world with features being extracted. Many other machine learning models can also be explored and built according to the type of datasets of related problems.

References

1. Polikar, R.: Ensemble based systems in decision making. IEEE Circuits Syst. Mag. **6**(3), 21–45 (2006). Third Quarter
2. Bali, A., Shailendra, N.S.: A review on the strategies and techniques of image segmentation. In: 2015 Fifth International Conference on Advanced Computing & Communication Technologies (ACCT). IEEE (2015)
3. Rother, C., Kolmogorov, V., Blake, A.: Grabcut: interactive foreground extraction using iterated graph cuts. ACM Trans. Graph. (TOG) **23**(3), 309–314 (2004). ACM
4. Brejl, M., Sonka, M.: Edge-based image segmentation: machine learning from examples. In: International Joint Conference on World Congress on Computational Intelligence on Neural Networks Proceedings, vol. 2. IEEE (1998)
5. Yan-Li, A.: Introduction to digital image pre-processing and segmentation. In: Seventh International Conference on Measuring Technology and Mechatronics Automation (ICMTMA). IEEE (2015)
6. Tang, J.: A color image segmentation algorithm based on region growing. In: 2nd International Conference on Computer Engineering and Technology (ICCET), vol. 6. IEEE (2010)
7. Ramadevi, Y., Kalyani, B., Sridevi, T.: Synergy between object recognition and image segmentation. Int. J. Comput. Sci. Eng. **2**, 2767–2772 (2010)
8. Rashidi, A., Sigari, M.H., Maghiar, M., Citrin, D.: An analogy between various machine-learning techniques for detecting construction materials in digital images. KSCE J. Civ. Eng. **20**, 1178–1188 (2015). Springer
9. Alpaydin, E.: Introduction to Machine Learning. The MIT Press, Cambridge, Massachusetts London, England (2010)
10. Liaw, A., Wiener, M.: Classification and regression by randomForest. R News. **2**(3), 18–22 (2002)
11. Burges, J.C.: A tutorial on support vector machines for pattern recognition. Bell laboratories. Lucent technologies. Data Min. Knowl. Discov. **2**, 121–167 (1998)
12. An, T.K., Kim, M.H.: A new diverse adaboost classifier. In: International Conference on Artificial Intelligence and Computational Intelligence, pp. 359–363 (2010)
13. Wu, X., Kumar, V.: The Top Ten Algorithms in Data Mining. Data Mining and Knowledge Discovery Series. Chapman & Hall/CRC, Minneapolis, Minnesota, USA (2009)
14. https://goo.gl/mTetx0