Object Detection Using Statistical and Textual Model

Fathia Altabouni, Ashraf Huwedi and Kenz Bozed

September 1, 2021

Object Detection using Statistical and textual Model

Fathia Farj Altabouni

Department of sofware endenering, Benghazi University, fathia.altabououni@uob.edu.ly

Ashraf Saad Huwedi

Department of Computer System Design , Benghazi University, ashrafhuwedi@uob.edu.ly

Kenz Amhmed Bozed

Department of Computer System Design , Benghazi University,Kenz.bozed@uob.edu.ly

Abstract— Optical system for humans has the ability to get to know a very large number of objects or classes of object from two-dimensional visual information or three dimensions. Solving this task of detection or recognition by humans without any major effort is due to the ability of humans for very fast parallel processing and their ability for self development, as well as the ability to learn from experience due to the exceptional structure of the human brain. Computer vision systems exist that are able to detecting an object in image, but usually these systems need to be instructed what the object to detected is. And The main difficulty comes from a very large number of objects, each of these objects may appear in an infinite number of several aspects: dimension, shape, color, position, lighting, and shade. In this paper, we present a robust method attempts to deal with the above challenges by using Local Binary Pattern (LBP) and Random Forest techniques(RF),. The proposed method first represents the images using Local Binary pattern (LBP) and then uses the Random Forest algorithm to classify the images based on the LBP histogram. Results show that using the suggested method 95.39% classification rate in the car detection , as well as 94.9%, for cat detection ,and flower detection ,

Keywords: **— Random Forest (RF), Local Binary Pattern (LBP).**

1 INTRODUCTION

Object Detection is one of the most challenging problems in Computer Vision. The difficulty is due to the significant amount of variation between images belonging to the same object category. Other factors, such as changes in viewpoint and scale, illumination, partial occlusions and multiple instances further complicate the problem of object detection [3].Object Detection plays a big part in our lives. We are constantly looking for and detecting objects: people, streets, buildings, hallways, tables, chairs, desks, sofas, beds, automobiles. Yet it remains a mystery how we perceive objects so accurately and with so little apparent effort. One of the main building blocks for computer vision and image analysis is to detect objects in an image and assign them to their corresponding class label. This is referred to as object detection and recognition.

Object detection is an important sub problem in the field of pattern recognition. Object discovery is a classic problem in computer vision and is often described as a difficult task. In many respects, it resembles other computer vision tasks, because it involves creating a fixed solution for distortion and changes in lighting and perspective. What makes object discovery a distinct problem is that it involves both the identification and classification of image areas [2]To detect an object, we need some ideas about where the object is located and how the image is fragmented. This creates a kind of chicken and egg problem, where we need to know its location, to identify the location of an object, to identify the shape of an object, and to identify its location. [1]. The application of this method reaches through many fields, for instance tracking, face detection and video surveillance .

This work includes feature extraction methods for object detection using Local Binary Pattern LBP, and classifier such as Random Forest RF.

2 RELETED WORK

(Juergen Gall1, Nima Razavi1, and Luc Van Gool-)(An Introduction to Random Forests for Multi-class Object Detection).Their research is mainly concerned with describing the general framework of random forests for multi-class object detection in images and give an overview of recent developments and implementation details that are relevant for practitioners. The team has utilized a subset of the PASCAL VOC 2006 database of objects[8].

(Danie Maturana, Domingo Mery, and A´lvaro Soto) (Face Recognition with Optimized Tree-Structured Local Binary Patterns) This group of researcher has deployed a novel method that uses training data to create discriminative LBP descriptors by using decision trees. The algorithm obtains encouraging results on standard databases, and presents good results .The FERET data set and the CAS-PEAL-R1 dataset were used[22].

(Jiri Trefny, Jiri Matas) (Extended Set of Local Binary Patterns for Rapid Object Detection) In this paper, presented researcher Two new encoding schemes were developed for representation of the intensity function in a local neighborhood. The encoding produces binary codes, which are complementary to the standard local binary patterns (LBP).Both new schemes preserve an important property of the LBP, the invariance to monotonic transformations of the intensity. The new LBP encoding schemes were tested on the face detection, car detection and gender recognition problems using the CMU-MIT frontal face dataset, the UIUC Car dataset and the FERET dataset respectively[23].

(Florian Baumann, Arne Ehlers, Karsten Vogt, Bodo RosenhahnCascaded) (Cascaded Random Forest for Fast Object Detection) They have presented Alternating Regression Forests, a novel Random Forest training procedure for regression tasks, which, in contrast to standard Random Regression Forests, optimizes any differentiable global loss function without sacrificing the computational benefits of Random Forests. ARFs are easy to implement and can be exchanged with standard Random Regression Forests without great efforts. This novel regression gives better performance on machine learning benchmarks compared to Boosted Trees[24].

(Prajowal Manandhar , Zeyar Aung , Wei Lee Woon and Prashanth Reddy Marpu) (Random Forest Ensemble Learning for Object Recognition Using RGB Features Along Object Edge) has presented a simple approach of object classification using ensemble learning with Random Forest with the use of the feature set consisting of RGB color information extracted from the detected edge points. In both, the datasets (similar and distinct objects) of images, the approach is able to produce good results. The distinct objects can be correctly identified in most of the occasions, while the results with the similar objects do not look that bad either[25].

(Bae-Keun Kwon, Jong-Seob Won, and Dong-Joong Kang)(Fast Defect Detection for Various Types of Surfaces using Random Forest with VOV Features) In this paper, Defect detection on an object surface is one of the most important tasks of an automated visual inspection system.

The most modern defect detection systems are required to operate in real-time and handle high- resolution images. One of main difficulties in system applications is that it cannot be used for general inspection of various types of surface without tuning the internal parameters.

The researchers presented, we demonstrate how to solve the problem mentioned above by using simple variance profile values of pixel intensities and applying it to the random-forest-based machine learning algorithm[26].

(Ricardo Acevedo-Ávila ,Miguel González-Mendoza)(A Statistical Background Modeling Algorithm for Real-Time Pixel Classification) An alternative statistical background pixel classifier intended for real-time and low-resource implementation. The algorithm works within a smart video surveillance application aimed to detect unattended objects in images with fixed backgrounds. The algorithm receives an input image and builds an initial background model based on image statistics[19].

(Ziyang Liu,Weihao Li,Dongxiao He, Weixiong Zhang)(Graph Convolutional Networks Meet Markov Random Fields:Semi-Supervised Community Detection in Attribute Networks) this paper proposed an end-to-end deep learning method, namely MRFasGCN, to integrate GCN and MRF for the problem of semi-supervised community detection in attributed net-works. It has architecture of three convolutional layers with GCN as the first two convolution layers and MRF as the third layer. The MRF component utilizes the coarse output of GCN to construct MRF's unary potentials, and then enhances the network-specific pairwise potentials to find better communities[21].

(Jean-Philippe Mercier, Mathieu Garon, Philippe Giguere, Jean-Francois Lalonde)(Deep Template-Based Object Instance Detection)this paper proposed a generic 2D object instance detection approach that uses example viewpoints of the target object at test time to retrieve its 2D location in RGB images, without requiring any additional training (i.e. fine-tuning) step. To this end, we present an end-to-end architecture that extracts global and local information of the object from its viewpoints. The global information is used to tune early filters in the backbone while local viewpoints are correlated with the input image[27].

(Nicolas Carion, Francisco Massa,Gabriel Synnaeve,Nicolas Usunier,Alexander Kirillov,Sergey Zagoruyko)(End-to-End Object Detection with Transformers) this paper proposed a new method that views object detection as a direct set prediction problem. Our approach streamlines the detection pipeline, effectively removing the need for many hand-designed components like a non-maximum suppression procedure or anchor generation that explicitly encode our prior knowledge about the task. The main ingredients of the new framework, called DEtection TRansformer or DETR, are a set-based global loss that forces unique predictions via bipartite matching, and a transformer encoder-decoder architecture. Given a fixed small set of learned object queries, DETR reasons about the relations of the objects and the global image context to directly output the final set of predictions in parallel[28].

In this work, building the model ،using the local binary technique to extract the features of the image and then use random forests was also used as a classifier to build the model and use sliding window for scan image , for the improve detected result.

3 PROPOSED METHODOLOGY

In this paper we are working on the object Detection using the local binary technique to extract the features of the image and then use random forests was also used as a classifier to build the model. We have taken several steps to process the data to build the system as follows:

- Preprocessing the data.
- Training the object detector classifier and choosing the final model.
- Test

The initial process or phase before the training begins, is to transform the images training dataset into a format that can be used by the system proposed. It is important to mention that he target format must be a vector of integer numbers.

These steps show how Preprocessing the data is done :

1. All the images for object are resized to 32 × 32 pixels.
2. It is converted to a gray scale image.
3. For each single image, we extract the LBP histogram which is only 59 bins vector.
4. The same steps are applied to non-object images without step one. Also, using real size and sliding window as a technique that scan different image scales.
5. Finally, concatenate all the vectors for object and non-object training data in a single matrix, which is considered as the input to the Random Forest.

3.1 Sliding Window technique

Input Image Exploration for object candidate searching is done by means of a window-sliding technique applied at a different image scale. As a result, object can be detected at different image locations and resolutions. The following parameters describe the scanning method which has been selected as a trade-off between object detector resolution and algorithm speed:

- Block Size: a square or rectangular block that determines the resolution of the face detector, as in the case of the system 32 × 32.
- Moving Scan step: a number of pixels that defines the window-sliding step to obtain the next block to be analyzed, which is 4.
- Down-Sampling Rate: Down scale factor for the scaling technique to reach all locations and scales in an image is 0.2.

3.2 Feature Extraction

features are extracted using basic local binary pattern(LBP). The basic local binary pattern operator, introduced by Ojala et al. and was based on the assumption that texture has locally two complementary aspects, a pattern and its strength It is actually defined as a grey scale invariant texture measure,which was proposed for texture analysis first, then it has proved a simple yet powerful approach to describe local structures. LBP is a simple, efficient, easy to compute, and robust to monotonic illumination and variations operator[6].

The LBP operator [14] is one of the best performing texture descriptors and it has been widely used in various applications.LBP has been successfully used for many different image analysis tasks, such as facial image analysis, biomedical image analysis, aerial image analysis, motion analysis, and image and video retrieval. The LBP method is very efficient due to its easy-to-compute feature extraction operation and simple matching strategy[17].

The original LBP operator labels the pixels of an image with decimal numbers, called Local Binary Patterns or LBP codes, which encode the local structure around each pixel. It proceeds thus, as illustrated in Figure 1.The labels for the image pixels are obtained by thresholding 3X3 neighborhood of each pixel as shown in Figure 1.

Each bit is made zero or one based on the difference in intensities between the corresponding pixel and the central pixel. The string of bits obtained is followed in clockwise or counterclockwise direction to get an 8 digit binary number. The binary number is converted into its decimal equivalent to obtain LBP label for the center pixel. Since the 3x3 neighborhood consists of 8 pixels excluding the center, a total of 28 =256 different labels are possible.

To deal with the texture at different scales, the operator was later generalized to use neighborhoods of different sizes [18]. During the two past decades, the Local Binary Patterns descriptor demonstrated remarkable performance and high robustness in extracting distinguishing features from a given image. Therefore, this feature extraction method has been widely applied in diverse challenging computer vision applications including face recognition [20].

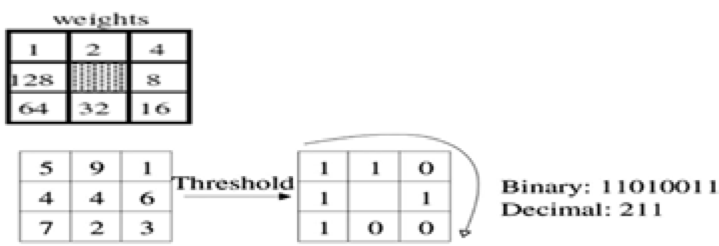

Figure 1: Example of the basic LBP operator.Inserting Content Elements[18].

3.2.1 Uniform Patterns

Certain patterns contain more information than others [15] It is possible to use only a subset of 2p binary patterns to describe the texture of images. Ojala et al. named these patterns uniform patterns .

An LBP is called uniform, if it contains at most two bitwise transitions from 0 to 1 or vice versa when the corresponding bit string is considered circular. For instance, 00000000 (0 transitions) and 01110000 (2 transitions) are both uniform, whereas 11001001 (4 transitions) and 01010011 (6 transitions) are not,Uniform patterns were used to reduce the length of LBP histograms, In uniform LBP mapping ,there is a separate output label for each uniform pattern and all the non-uniform patterns are assigned to a single label. Thus, the number of different output labels for mapping for patterns of P bits is $P(P - 1) + 3$.

For instance, the uniform mapping produces 59 output labels for neighborhoods of 8 sampling points, and 243 labels for neighborhoods of16 sampling points.The reasons for omitting the non-uniform patterns are twofold. First, most of the local binary patterns in natural images are uniform.

Ojala et al. noticed that in their experiments with texture images, uniform patterns account for a bit less than90% of all patterns when using the (8, 1) neighborhood and for around 70% in the (16, 2) neighborhood.

In experiments with facial images, it was found that90.6% of the patterns in the (8, 1) neighborhood and 85.2% of the patterns in the(8, 2) neighborhood are uniform[6].

The second reason for considering uniform patterns is the statistical robustness. Using uniform patterns instead of all the possible patterns has produced better recognition results in many application

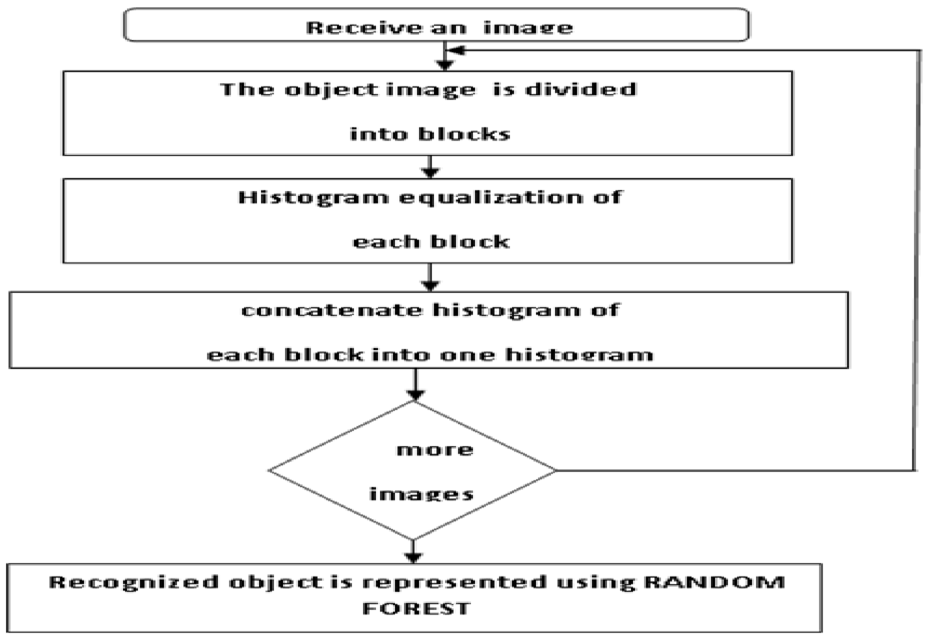

Figure 2:Flow chart shows how LBP works for Object Detection.

3.3 Classification Method

Random forests are ensembles of randomized decision trees that can be applied for regression classification tasks and even both at the same time [7][12]. A random forest consists of a set of trees Tt where each tree consists of split nodes and leaves as illustrated in Figure 3. The split nodes evaluate each arriving image patch and, depending on the appearance of the patch, pass it to the left or right child[8].

The term came from random decision forests that was first proposed by Tin Kam Ho of Bell Labs in 1995. The method combines Breiman's "bagging" idea and the random selection of features, introduced independently by Ho and Amit and Geman in order to construct a collection of decision trees with controlled variation.

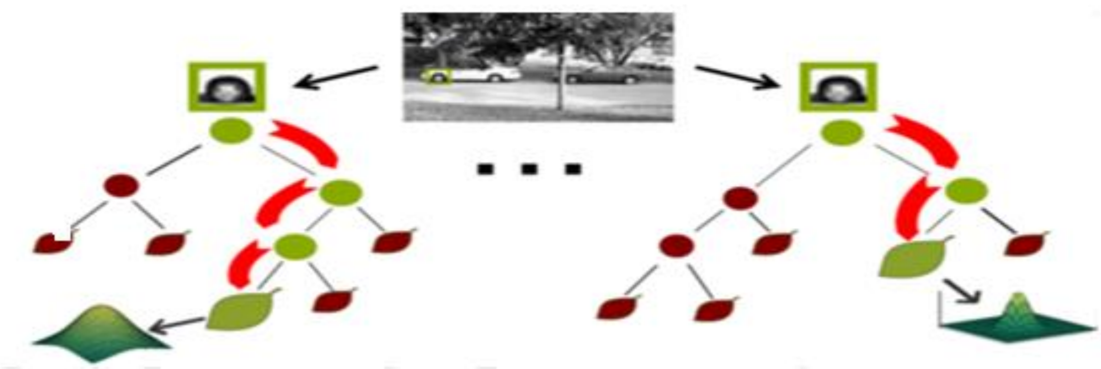

Figure 3: random forest consists of a set of trees that map an image patch to a distribution stored at each leaf[8].

The disks indicate split nodes that evaluate the appearance of a patch and pass it to the right or left child until a leaf is reached.

Random Forests [9] have recently become a popular approach in Computer Vision. They have been used for a large number of classification [10],[11][12], [16], and regression tasks [12],[16].

Local Binary Pattern (LBP) was chosen for extracting features, and Random Forest (RF) was used as a classifier to build the model. As illustrated in Figure 3, the research plan can be described as:

- Training the object detector classifier and choosing the final model.
- Testing and conclusion.

3.2.1 Training the object detector classifier and choosing the final model.

After getting the right input format, thein this stage is doneTraining the object detector classifier and choosing the final model,after getting the right input format, the classifier must be trained to get the appropriate model for the testing phase.

The experimental results in the Object Detection System are first concerned with choosing block size for LBP , where before implementing the object detection method proposed in the literature. First, the binary pattern operator was applied with the change in the size of the neighborhood until we obtained a standard that gives the best results with the random forest classification.

Training phase starts by entering the input training data extracted from the LBP features into RF classifier and changing the number of trees along with the number of object/Non-object samples, where the model selection depends largely on the amount of training data, and the number of trees in the forest (classifier). Having the proper values of these parameters is essential to obtain good classification results. The model with the best TP (True Positive) value is the appropriate one.

3.2.2 OBJECT DETECTION TESTING FRAMEWORK.

For the purpose of the testing classifier, entering the input testing data extracted from the LBP features and also entered the database model that has been trained into RF predict classifier with changing the number of trees (100 / 50) to compare the results, thus obtaining the required classification either object or Non_object. It should report the test results including accuracy and error rate.

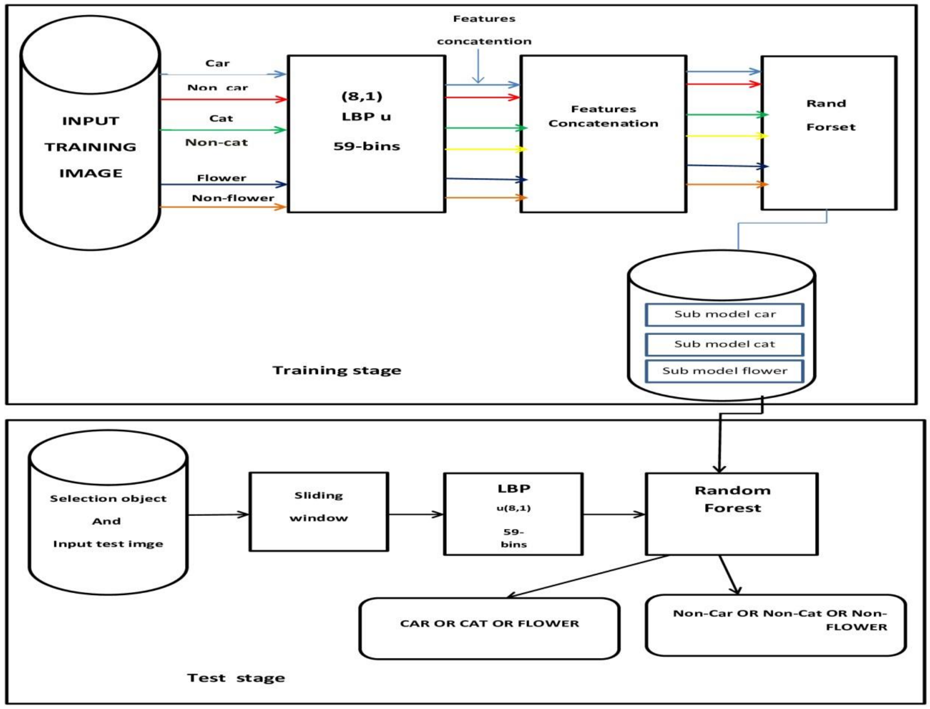

Figure 4:object Detection scheme.

4 EXPERIMENTAL RESULTS

The main research goal of this paper was to find out whether a system that learns from examples can be used to object detection and eliminate false detections, while maintaining the true detections.In this paper, the best suggestions were offered to detect a object The technique of extracting LBP features and RF classification was used, as well as the technique of sliding window to reveal the object, get the results and compare them.

These techniques were applied to three databases: CAR database, CAT database, FLOWER database. We obtained different results from TP, FN, the results were as follows: The CAR database gave the best accuracy at 95.39% and the lowest error rate is 0.04.

The CAT database gave the best accuracy at 94.9% and the lowest error rate is 0.04. The FLOWER database gave the best accuracy at 95.9% and the lowest error rate is 0.05.

4.1 Car Database

The training set is Labeled Cars in the Stanford cars dataset. Database contains 16,185 images of 196 classes of cars. The data is split into 8,144 training images and 8,041 testing images, where each class has been split roughly in a 50-50 sp, as illustrated in Figure 5

Figure 5: car datasbase sample.

4.3 Cat Database

The training set is Labeled cats in the cats dataset containing 10,000 cat images, We randomly divide the data into three sets: 5,000 images for training, 2,000 images for validation and 3000 images for testing. as illustrated in Figure 6

Figure 6: cat database sample.

4.3 Flower database

The training set is Labeled cats in the Flower dataset containing 4500 Flower images, We randomly divide the data into three sets: 2,000 images for training, 500 images for validation and 200images for testing. as illustrated in Figure 7

Figure 7: flower database sample.

After finding the appropriate LBP feature type and block size, a great deal time was spent on testing the system to discover the amount of training and testing data for both object and non-object data. The table 6.2 below shows the results for different sizes of car and non-car data and their TP and TN. And the table 6.3 shows the results for different sizes of cat and non-cat data and their TP and TN .The table 6.4 shows the results for different sizes of flower /non- flower data and their TP and TN.

Table 1: Test results for different cars /non- car samples

Model name	Training data		Testing data				
	No. of car	No.of non car	No. of car	No. of non car	No. Of Tree	**TP**	**FN**
Car_M1	3000	10,000	1000	30000	100	95.3	0.12
Car_M2	3000	10,000	1000	30000	50	95.36	0.08
Car_M3	3000	25,000	1000	30000	50	93.9	0.20
Car_M4	3000	25000	1000	30000	100	93.3	0.26
Car_M5	3000	15000	1000	30000	50	95.69	0.08
Car_M6	3000	15000	1000	30000	100	95.59	0.13
Car_M7	2500	10000	2000	30000	100	95.39	0.04
Car_M8	2500	10000	2000	30000	50	93.8	0.18
Car_M9	2000	10000	2500	30000	100	93.1	0.12

Table 2 Test results for different cat /non- cat samples

Model name	Training data		Testing data				
	No. of cat	No.of non cat	No. of cat	No. of non cat	No. Of Tree	TP	FN
Cat_M1	2500	10000	3000	30000	50	93.35	0.21
Cat_M2	2500	10000	3000	30000	100	93.3	0.24
Cat_M3	2500	25000	3000	30000	100	90.5	1.13
Cat_M4	2500	25000	3000	30000	50	90	1.20
Cat_M5	2500	15000	3000	30000	100	94.8	0.18
Cat_M6	2500	15000	3000	30000	50	95	0.14
Cat_M7	2000	10000	3500	30000	100	94.9	0.04
Cat_M8	2000	10000	3500	30000	50	93.1	0.10

Table 3 Test results for different flower /non- flower sample.

Style Tag	Training data		Testing data				
	No. of flower	No.of non flower	No. of flower	No.of non flower	No.		No.
Flower_M1	1500	10000	3000	30000	50	93.35	1.13
Flower_M2	1500	10000	3000	30000	100	93.3	0.53
Flower_M3	1500	25000	3000	30000	100	90.5	0.24
Flower_M4	1500	25000	3000	30000	50	90	0.24
Flower_M5	2500	15000	3000	30000	100	94.8	0.18
Flower_M6	2500	15000	3000	30000	50	95	0.20
Flower_M7	2000	10000	3500	30000	100	94.9	0.05
Flower_M8	2000	10000	3500	30000	50	93.1	0.35

5 CONCLUSION AND FUTURE Work

The empirical part of this paper could successfully produce satisfactory results for Object Detection . As a new application that uses object detection system with LBP and RF and the sliding window to reveal the object, get the results and compare them.

These techniques were applied to different databases: car database, cat database, flower database, . We obtained different results from TP, FN. The CAR database gave the best accuracy at 95.39% and the lowest error rate is 0.04.

The CAT database gave the best accuracy at 94.9% and the lowest error rate is 0.04. The FLOWER database gave the best accuracy at 95.9% and the lowest error rate is 0.05.In future work, it is hopeful that the system employed and tested here would motivate other researchers to explore more detection types .Certainly, other contributions can improve and complement to the present techniques for much better results ,as well as implementing similar detection techniques on video data.

REFERENCES

[1] Walther, D., Itti, L., Riesenhuber, M., Poggio, T., & Koch, C. (2002, November). Attentional selection for object recognition—a gentle way. In International workshop on biologically motivated computer vision (pp. 472-479). Springer, Berlin, Heidelberg.

[2] Girshick, R. (2015). Fast r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 1440-1448).

[3] Fahad Shahbaz Khan1., Rao Muhammad Anwer., Joost van de Weijer., Andrew D. Bagdanov., Maria Vanrell., Antonio M. Lopez. in Proceedings / CVPR, IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society Conference on Computer Vision and Pattern Recognition • June 2012

[4] J. Komala Lakshmi, and Dr.M.Punithavalli conducted "A Survey on Performance Evaluation of Object Detection Techniques in Digital Image Processing"November 2010,International Journal of Computer Science

[5] Sanjivani Shantaiya , Kamal Mehta , Keshri Verma A Survey on Approaches of Object Detection, International Journal of Computer Applications (0975 – 8887) Volume 65– No.18, March 2013

[6] Ojala, T., Pietikäinen, M., & Mäenpää, T. (2001, March). A generalized local binary pattern operator for multiresolution gray scale and rotation invariant texture classification. In International Conference on Advances in Pattern Recognition (pp. 399-408). Springer, Berlin, Heidelberg

[7] Criminisi, A., Shotton, J., & Konukoglu, E. (2011). Decision forests for classification, regression, density estimation, manifold learning and semi-supervised learning. Microsoft Research Cambridge. Rep. MSRTR-2011-114, 5, 12.

[8] Razavi, N., Gall, J., & Van Gool, L. (2011, June). Scalable multi-class object detection. In CVPR 2011 (pp. 1505-1512). IEEE.

[9] Breiman, L. (2001). Random forests. Machine learning, 45(1), 5-32.

[10] Huang, C., Ding, X., & Fang, C. (2010, August). Head pose estimation based on random forests for multiclass classification. In 2010 20th International Conference on Pattern Recognition(pp. 934-937). IEEE.

[11] Shotton, J., Johnson, M., & Cipolla, R. (2008, June). Semantic texton forests for image categorization and segmentation. In 2008 IEEE Conference on Computer Vision and Pattern Recognition (pp. 1-8). IEEE.

[12] Criminisi, A., Shotton, J., Robertson, D., & Konukoglu, E. (2010, September). Regression forests for efficient anatomy detection and localization in CT studies. In International MICCAI Workshop on Medical Computer Vision (pp. 106-117). Springer, Berlin, Heidelberg.

[13] Ahonen, T., Hadid, A., & Pietikainen, M. (2006). Face description with local binary patterns: Application to face recognition. IEEE Transactions on Pattern Analysis & Machine Intelligence, (12), 2037-2041.

[14] Pietikäinen, M., Hadid, A., Zhao, G., & Ahonen, T. (2011). Computer vision using local binary patterns (Vol. 40). Springer Science & Business Media.

[15] Gall, J., & Lempitsky, V. (2009). Class-specific hough forests for object detection. In Decision forests for computer vision and medical image analysis (pp. 143-157). Springer, London.

[16] Girshick, R., Shotton, J., Kohli, P., Criminisi, A., & Fitzgibbon, A. (2011, November). Efficient regression of general-activity human poses from depth images. In 2011 International Conference on Computer Vision (pp. 415-422). IEEE.

[17] Huang, D., Shan, C., Ardabilian, M., Wang, Y., & Chen, L. (2011). Local binary patterns and its application to facial image analysis: a survey. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 41(6), 765-781

[18] Zhou, X., Gong, W., Fu, W., & Du, F. (2017, May). Application of deep learning in object detection. In 2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS) (pp. 631-634). IEEE.

[19] Ricardo Acevedo-Ávila1 ,Miguel González-Mendoza1,Andrés García-García1, A Statistical Background Modeling Algorithm for Real-Time Pixel Classification, Comp. y Sist. vol.22 no.3 México jul./sep. 2018

[20] Mohamed Kas,Youssef El Merabet,Yassine Ruichek,Rochdi Messoussi, Survey on Local Binary Pattern Descriptors for Face Recognition, March 2019, Acts of the Second Conference of the Moroccan Classification.

[21] Di Jin,Ziyang Liu,Weihao Li, Dongxiao He, Weixiong Zhang,Graph Convolutional Networks Meet Markov Random Fields: Semi-Supervised Community Detection in Attribute Networks,2019, The Thirty-Second AAAI Conference on Artificial Intelligence (AAAI-18).

[22] Daniel Maturana, Domingo Mery and Alvaro Soto, Face Recognition with Optimized Tree-Structured Local Binary Patterns,2012 .

[23] Trefný, J., & Matas, J. (2010, February). Extended set of local binary patterns for rapid object detection. In Computer vision winter workshop (pp. 1-7).

[24] Baumann, F., Ehlers, A., Vogt, K., & Rosenhahn, B. (2013, June). Cascaded random forest for fast object detection. In Scandinavian Conference on Image Analysis (pp. 131-142). Springer, Berlin, Heidelberg.

[25] Manandhar, P., Aung, Z., Woon, W. L., & Marpu, P. (2015). Random forest ensemble learning for object recognition using rgb features along object edge. In ICCIT 2015 The Fourth International Conference on Communications and Information Technology.

[26] Bae-Keun Kwon, Jong-Seob Won & Dong-Joong KangFast defect detection for various types of surfaces using random forest with VOV features,2015,International Journal of Precision Engineering and Manufacturing,

[27] Jean-Philippe Mercier, Mathieu Garon, Philippe Giguère, Jean-François Lalonde, Nov 2020(v3),, Deep Template-based Object Instance Detection, IEEE/CVF Winter Conference on Applications of Computer Vision (WACV).

[28] Nicolas Carion, Francisco Massa,Gabriel Synnaeve,Nicolas Usunier,Alexander Kirillov,Sergey Zagoruyko,2021,End-to-End Object Detection with Transformers European Conference on Computer Visio.,