OBJECT RECOGNITION USING TENSORFLOW IN OPENCV

[1]Alfiya Mirza, [2]Minal Surjuse, [3]Nidhi Waghmare, [4]Priyanka Wasnik, [5]Manish Assudani
[1]UG Student, [2]UG Student, [3]UG Student, [4]UG Student, [5]Assistant Professor
[1]Computer Science and Engineering,
[1]Anjuman College of Engineering and Technology, Nagpur, India.

Abstract : [1] Object recognition is one of the most fascinating abilities that humans easily possess since childhood. With a simple glance of an object, humans are able to tell its identity or category despite of the appearance variation due to change in pose, illumination, texture, deformation, and under occlusion. Object recognition is an important part of computer vision because it is closely related to the success of many computer vision applications. A number of object recognition algorithms and systems have been proposed for a long time in order to address this problem.

[1]Object recognition is concerned with determining the identity of an object being observed in the image from a set of known labels. Oftentimes, it is assumed that the object being observed has been detected or there is a single object in the image. Detecting and recognizing objects in unstructured environments is one of the most challenging tasks in computer vision research. Parallel to the role of deformable shape in object recognition, deformable illumination is designed as an object detection technique.

***Index Terms* – Object Recognition.**

I. INTRODUCTION

[1]Object recognition is a computer vision technique for identifying objects in image or videos. Object recognition is a key output of deep learning and machine learning algorithms. When humans look at a photograph or watch a video, we can readily spot people, objects, scenes, and visual details. The goal is to teach a computer to do what comes naturally to humans to gain a level of understanding of what an image contains.

[1]Object recognition is a task of finding 3-dimensional (3D) objects from two dimensional (2D) images and classifying them into one of the many known object types. It is an important part of computer vision because it is closely related to the success of many computer vision applications such as robotics, surveillance, registration and manipulation etc. A number of object recognition algorithms and systems have been proposed for a long time toward this problem.

[6]Object recognition - in computer vision is the task of finding and identifying objects in an image or video sequence. Humans recognize a multitude of objects in images with little effort, despite the fact that the image of the objects may vary somewhat in different viewpoints, in many different sizes / scale or even when they are translated or rotated. Objects can even be recognized when they are partially obstructed from view. This task is still a challenge for computer vision systems. Many approaches to the task have been implemented over multiple decades. It is an important part of computer vision because it is closely related to the success of many computer vision applications such as robotics, surveillance, registration and manipulation etc. A number of object recognition algorithms and systems have been proposed for a long time toward this problem. Yet, a general and comprehensive solution to this problem has not be made. In model-based object recognition, a 3D model of the object being recognized is available. The 3D model contains detailed information about the object, including the shape of its structure, the spatial relationship between its parts and its appearance. This 3D model provides prior knowledge to the problem being solved. There are two common ways to approach this problem. The first approach involves obtaining 3D information of an object from images and then comparing it with the object models. To obtain 3D information, specialized hardware, such as stereo vision camera, is required to provide the 3D information in some forms. The second approach requires less hardware support but is more difficult. It first obtains the 2D representation of the structure of the object and then compares it with the 2D projections of the generative model.

a. [6]Model-Based Object Recognition :-

[6]In model-based object recognition, a 3D model of the object being recognized is available. The 3D model contains detailed information about the object, including the shape of its structure, the spatial relationship between its parts and its appearance. This 3D model provides prior knowledge to the problem being solved. This knowledge, in principle, can be used to resolve the potential confusion caused by structural complexity and provide tolerance to noisy or missing data. There are two common ways to approach this problem. The first approach involves obtaining 3D information of an object from images and then comparing it with the object models.

[6]To obtain 3D information, specialized hardware, such as stereo vision camera, is required to provide the 3D information in some forms. The second approach requires less hardware support but is more difficult. It first obtains the 2D representation of the structure of the object and then compares it with the 2D projections of the generative model .Using 3D model has both the advantages and the disadvantages. On one side, explicit 3D models provide a framework that allows powerful geometric constraints to be used to achieve good effect.

[6]Other model features can be predicted from just a few detected features based on the geometric constraints. On the other side, using models sacrifice its generality. The model schemas severely limit the sort of objects that they can represent and it is quite difficult and time-consuming to obtain the models.

b. [6]View-Based Object Recognition :-

[6]In view-based object recognition, 3D model of the object is not available. The only known information is a number of representations of the same object viewed at different angles and distances. The representations of the object can be obtained by taking a series of images of the same object in a panorama fashion. Most of these operate by comparing a 2D, image representation of object appearance against many representations stored in a memory and finding the closest match. Matching of this type of recognition is simpler but the space requirements for representing all the views of the object is large. Again, there are many ways to approach this problem. One of the common way is to extract salient information, such as corner points, edges and region etc, from the image and match to the information obtained from the image database.

[6]Another common approach extracts translation, rotation and scale invariant features, such as SIFT, GLOH and RIFT, from each image and compares them to the features in the feature database View-based object recognition systems have the advantage of greater generality and more easily trainable from visual data. View-based approach is generally a useful technique. However, since matching is done by comparing the entire objects, some methods are more sensitive to background clutter and occlusion. Some methods solve this problem by applying image segmentation on the entire objects so as to divide the image representations into smaller pieces for matching separately. Some other methods avoid using segmentation and solve the problem by employing voting techniques, like Hough transform methods. This technique allows evidence from disconnected parts to be effectively combined.

c. [6]FEATURE BASED OBJECT RECOGNITION

[6]A number of models have been proposed under this approach. These are as follows.

- [6]Local Features And Histogram Based Planar Object Recognition

[6]Visual object recognition system recognizes an object in scene image. Although it has been researched over the past few decades, it‟s still a difficult problem because the object in scene image can be affected by a lot of variations including image pose, illumination changes, clutter background, etc. In general, there are two main approaches for object recognition including global and local features based methods. The global features are extracted from the whole content of object, while the local features are extracted from local pieces of object. Global features based methods commonly use shape and colour features. However, these methods are sensitive to object occlusion and cluttering parts which disturb recognition. Moreover, using the global features for object

recognition requires object segmentation from scene image which is still a challenging problem. Changes, object transformation and object variation, variety local features have been proposed and then they are applied to many applications such as object detection and recognition, image matching, image stitching, object tracking, etc. In 1999, D. Lowe proposed a Scale Invariant Feature Transformation (SIFT) feature detector and descriptor for object recognition. SIFT is invariant to image translation, scaling, rotation and partially invariant to illumination changes and affine transformation but complexity and heavy. This work becomes the original inspiration for the most of local feature descriptors proposed later. Then, a Speed Up Robust Features (SURF) was reported as a more efficient substitution for SIFT by H. Bay[7] in 2006 because it produces smaller descriptor as well as speeds up matching step. However, these feature descriptors are still complicated and take much time for matching. For object recognition system, the goal is to make descriptors faster to compute, more compact while remaining robust. To address these requirements, many researchers tended to build a light weight descriptor based on binary string[8]. In 2010, Calonder proposed a Binary Robust Independent Elementary Features (BRIEF) descriptor which relies on a relatively small number of intensity difference tests to represent an image patch as a binary string. Global approach is simpler and faster than local approach but not robust under illumination, image pose changes and clutter images. Therefore, we propose an approach to combine the local and global feature to make a robust object recognition system. In our method, the object is detected first using local features matching, and then the background is removed. Next, we extract histogram feature on segmented image. To combine the spatial information into histogram feature, a local histogram features is proposed that describes the intensity distribution of different region in image. The goal in this approach is dealing with planar objects such as book cover, cd cover, painting, box, logo image, etc. So, a system is built using start of- art binary features, the robust feature ORB combine with local histograms. Phan D. et al. have

proposed an approach for planar object recognition using binary local invariant feature and local histogram features, in scene image. First, the object is detected based on a homograph which represents transformation of

object in scene image from reference image, estimated from matching pairs of key points between two images. Then, local intensity histograms are computed from blocks inside scene object. In order to locate these blocks, a

reference object is divided into many blocks then the corresponding blocks of the scene object are located based on the homograph with assumption that object is a plane. To make the feature invariant to the illumination, the local histogram is computed from only intensity component (V) of HSV color of image. Similarity of two images is calculated from the similarity of their local histogram features and the matching ratio. The recognized object is the most similar reference object in data set. For evaluation, we experiment our method with a planar dataset from Stanford University which includes book cover, cd cover, painting, business card, etc.

- [6]Semantic-Feature-Based Object Recognition By Using Internet Data Mining

There are many successful algorithms that promise a high accuracy in recognizing objects, e.g. SIFT[9], etc. However, unfortunately they are only on image feature base, therefore could not establish relationships among images and the words inside. Moreover, the recognition itself is based on the given objects which may cause inconveniences in certain fields such as robotics. In the other hand, it is very difficult for the traditional image- processing-based methods to cluster objects, which function similarly but appear differently, correctly, because the image-based methods could not retrieve hidden property information that is beneath these objects‟ appearances.

[6]To overcome these problems, in this paper a method is introduced that describes, identifies and clusters objects by measuring their semantic similarities. In order to detect text regions in an image, a text region classifier is trained with a SVM module in advance. Each image is processed by the SVM module to detect text regions . These text regions are then split out , and then binarized with thresholds calculated from their histograms. The binarized images are then passed into an Optical-Character-Recognition (OCR) module to read text labels out from images into text strings that are comprehensible to computers. The process is mentioned as Text Extraction . These text strings, been considered as keywords that are related to the image itself and the object that is inside, are searched one by one with a Google Web Search API to look for logically related information. With a number of web-search results in hand, these results are splitted into single words, which are considered as component words. These component words are definitely related to the corresponding keywords. This process is mentioned as Web Search and Component Words Extraction.

[6]Xu J. et al. have considered a problem of automated object description and clustering. Because traditional image-processing based object recognition algorithms can only cluster objects in image-base, they have proposed a method to describe an object in human language and group similar objects together in text processing way. This paper describes a system that recognizes objects with text labels printed on the surface of objects themselves or their packing cases. By analyzing them, objects could be described in English words, and then be clustered into corresponding groups. we presented an algorithm that automatically read text labels that are printed on the surface of objects, and then search them through Internet, mining results to describe the object in human languages as feature words. Objects with feature words description are calculated similarities between each other to get them grouped into several groups.

[6]DISCUSSION

[6]In this paper, a survey of the recent researches on object recognition was given. Several recently proposed object recognition approaches are described and discussed. From the survey, we know that creating distinctive key point descriptors for object matching is a critical step to the whole object matching process because it is the invariance property of the intelligently made descriptors that make the object matching invariant to the scale. While view-based object recognition approach is generally a bottom-up approach, it is possible to incorporate top-down approach technique into it to make the object recognition faster and more accurate.

REFERENCES

1. http://in.mathworks.com/solutions/deep-learning/object-recognition.html
2. Andrew Gleibman Sampletalk Research, POB 7141, Yokneam-Illit 20692, Israel www.sampletalk.com
3. Chih-Wei Hsu, Chih-Chung Chang, and Chih-Jen Lin. "A Practical Guide to Support Vector Classification". Department of Computer Science, National Taiwan University, Taipei 106, Taiwan. http://www.csie.ntu.edu.tw/~cjlin
4. Bernt Schiele and James L. Crowley. Transinformation for active object recognition. In International Conference on Computer Vision, New Delhi, India, 1998.
5. R.C. Veltkamp, M. Tanase, "Content-Based Image Retrieval Systems: A Survey", Technical Report UU-CS-2000-34 (extended version), 2002.
6. https://www.researchgate.net/publication/272503547_A_REVIEW_ON_OBJECT_RECOGNITION_TECHNIQUES_FOR_PRACTICAL_APPLICATION_Index_Terms_Object_Recognition_Model_Based_Object_Recognition_Approach_View_Based_Object_Recognition_Approach_Feature_Based_Object_R
7. H. Bay, T. Tuytelaars, L.V. Gool, "SURF: Speeded Up Robust Features", European Conference on Computer Vision, May 2006.
8. M. Calonder, V. Lepetit, C. Strecha, and P. Fua, "Brief:Binary robust independent elementary features"European Conference on Computer Vision, 2010.
9. David G. Lowe, Distinctive Image Features from Scale- Invariant Keypoints. The Proceedings of the Seventh IEEE International Conference on Computer Vision: pp.1150-1157 vol.2, 1999.