iJRASET
International Journal For Research in
Applied Science and Engineering Technology

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 2 **Issue:** IV **Month of publication:** April 2014

**DOI:**

ISSN: 2321-9653

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

# A Literature Survey on Task Specific Image Partitioning

Deepu Baby

*M. Tech Student*
*Department of Computer Science*
*Viswajyothi College Of Engineering And Technology,MG University,Kerala*

**Abstract—** ***Image partitioning is an important preprocessing step for many of the state-of-the-art algorithms used for performing high-level computer vision tasks. The task specific image partitioning method partitions the image by means of correlation clustering, maximizing a linear discriminant function defined over a superpixel graph. Typically, partitioning is made without considering the task in hand. A region-based image representation that made up of by task specific image partitioning will lead to a higher task performance than that reached using any task oblivious partitioning framework. Literature survey had made on various papers which includes the techniques of image segmentation using normalized cuts, mean shift for feature space analysis , graph based image segmentation, quick shift, enforcing label consistency, object class segmentation, single image depth estimation and selecting regions for scene understanding.***

**Keywords—** ***image partitioning,computer vision,correlation clustering,superpixel graph,feature space***

## I. INTRODUCTION

Region based image representations (RBIRs) have been shown to be effective in improving the performance of algorithms for high-level image/scene understanding, which encompasses tasks such as object class segmentation, scene segmentation, surface layout labeling, and single view 3D reconstruction. The effectiveness comes as a result of promoting the following three merits of using the RBIRs. First, the coherent support of a region, commonly assumed to be of a single label, serves as a good prior for many labeling tasks. Second, these coherent regions allow a more consistent feature extraction that can incorporate surrounding contextual information by pooling many feature responses over the region. Third, compared to pixels, a small number of larger homogeneous regions can significantly reduces the computational cost in the successive labeling task.

The image partitioning framework for obtaining RBIRs that realizes these benefits and improves the task specific labeling performance The task-specific image partitioning problem can be described as follows: given an image and labeling task, produce a partitioning of the image into disjoint regions such that each region is homogeneous with respect to the desired labeling of the task, and the labels of its neighboring regions are different. Note that this is different from image labeling in that system aim to produce a partitioning without region-labeling.

There are several general-purpose unsupervised image partitioning algorithms for region based image understanding. For instance, in the superpixel based conditional random fields (CRFs) models , mean-shift , normalized cuts, graph-based local variation algorithm , and their variants such as quick-shift are used to obtain small coherent image regions, called superpixels. These apriori over-segmentations are not related to any task and maybe limited in capturing accurate global information for the successive region-labeling step. To enhance its ability, some recent CRFs are based on either a hierarchy of regions or a set of partitionings. These multiple partitionings are obtained using mean-shift segmentation with different kernel sizes, multiscale normalized cuts , a hierarchical segmentation with increasing edge strength , and a simple region-merging algorithm . These algorithms while empirically successful to a certain extent use task-oblivious

ISSN: 2321-9653

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

partitionings and therefore do not address the task-specific image partitioning problem.The rest of the paper is arranged in such a way that section II contains a detailed literature survey about the image partitioning methods.Section III describes the task specific image partitioning and region based image retrieval system.

## II. Literature Survey

The task specific image partitioning is a method to partition the images to regions according to the tasks for the images are using. So a detailed survey of the current image segmentation techniques and papers are necessary to check the need for a new framework. Literature survey had made on various papers which includes the techniques of image segmentation using normalized cuts, mean shift for feature space analysis , graph based image segmentation, quick shift, enforcing label consistency, object class segmentation, single image depth estimation and selecting regions for scene understanding.

### *A. Class Segmentation and Object Localization with Superpixel Neighbourhoods*

Image level object categorization has led to significant interest on the related fronts of localization and pixel level categorization. Both areas have seen significant progress, through object detection challenges like PASCAL VOC. So far, the most promising techniques seem to be those that consider each pixel of an image. For localization, sliding window classifiers consider a window (or all possible windows) around each pixel of an image and attempt to find the classification which best fits the model. Lately, this model often includes some form of spatial consistency. In this way, sliding window classification as a "top-down" localization technique which tries to fit a coarse global object model to each possible location.

This paper proposes a method to identify and localize object classes in images. Instead of operating at the pixel level,it advocate the use of superpixels as the basic unit of a class segmentation or pixel localization scheme. To this end, construct a classifier on the histogram of local features found in each superpixel. Regularize this classifier by aggregating histograms in the neighborhood of each superpixel and then refine our results further by using the classifier in a conditional random field operating on the superpixel graph. This method exceeds the previously published state-of-the-art on two challenging datasets.

It uses quick shift to extract superpixels from our input images. Our model is quite simple: perform quick shift on a five-dimensional vector composed of the LUV colorspace representation of each pixel and its location in the image. Then construct a bag-of-features classifier which operates on the regions defined by the superpixels have found. SIFT descriptors are extracted for each pixel of the image at a fixed scale and orientation using the fast SIFT framework found in. The classifier which results from this is very specific. It finds superpixels which resemble superpixels that were seen in the training data without considering the surrounding region. This means that while a wheel or grille on a car may be correctly identified, the nearby hub of the wheel or the headlight can be detected with lower confidence or missed altogether

Disadvantages

1) Not exploited the use of task-specific training data to produce partitioning that is substantially more beneficial in terms of addressing the task-specific image partitioning problem

### *B. Normalized Cuts and Image Segmentation*

A novel approach for solving the perceptual grouping problem in vision. Rather than focusing on local features and their consistencies in the image data, this approach aims at extracting the global impression of an image. It treats image segmentation as a graph partitioning problem and propose a novel global criterion, the normalized cut, for segmenting the graph. The normalized cut criterion measures both the total dissimilarity between the different groups as well as the total similarity within the groups.an efficient computational technique based on a generalized eigenvalue problem can be used to optimize this criterion. Can apply this approach to segmenting static images, as well as motion sequences

Since there are many possible partitions of the domain of an image into subsets,There are two aspects to be considered here. The first is that there may not be a single correct answer. A Bayesian view is appropriate. There are several possible interpretations in the context of prior world knowledge. The

ISSN: 2321-9653

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

difficulty, of course, is in specifying the prior world knowledge. Some of it is low level, such as coherence of brightness, color, texture, or motion, but equally important is mid- or highlevel knowledge about symmetries of objects or object models. The second aspect is that the partitioning is inherently hierarchical. Therefore, it is more appropriate to think of returning a tree structure corresponding to a hierarchical partition instead of a single flat partition.

This suggests that image segmentation based on lowlevel cues cannot and should not aim to produce a complete final correct segmentation. The objective should instead be to use the low-level coherence of brightness, color, texture, or motion attributes to sequentially come up with hierarchical partitions. Mid- and high-level knowledge can be used to either confirm these groups or select some for further attention. This attention could result in further repartitioning or grouping. The key point is that image partitioning is to be done from the big picture downward, rather like a painter first marking out the major areas and then filling in the details.

Disadvantages

1) The learning criterion does not take into account inference on the full graph
2) Instead, is is based on a local cost by treating each pairwise relations between adjacent nodes as independent samples
3) So it requires a complex and unstable eigenvector approximation which must be differentiable

### C. *Mean shift :A robust approach toward feature space analysis*

A general non-parametric technique is proposed for the analysis of a complex multimodal feature space and to delineate arbitrarily shaped clusters in it. The basic computational module of the technique is an old pattern recognition procedure: the mean shift. For discrete data, prove the convergence of a recursive mean shift procedure to the nearest stationary point of the underlying density function and, thus, its utility in detecting the modes of the density. The relation of the mean shift procedure to the Nadaraya-Watson estimator from kernel regression and the robust M-estimators; of location is also established. Algorithms for two low-level vision tasks discontinuity-preserving smoothing and image segmentation - are described as applications. In these algorithms, the only user-set parameter is the resolution of the analysis, and either gray-level or color images are accepted as input. Extensive experimental results illustrate their excellent performance

In pattern recognition and machine learning, a feature vector is an n-dimensional vector of numerical features that represent some object. Many algorithms in machine learning require a numerical representation of objects, since such representations facilitate processing and statistical analysis. When representing images, the feature values might correspond to the pixels of an image. Feature vectors are equivalent to the vectors of explanatory variables used in statistical procedures such as linear regression. Feature vectors are often combined with weights using a dot product in order to construct a linear predictor function that is used to determine a score for making a prediction

For each subset, a parametric representation of feature of interest is obtained andresult is mapped in to a point in multidimensional space. A general tool for feature space analysis.Quality of output is only related to the kernel bandwidth,ie. Resolution of analysis.The vector space associated with these vectors is often called the feature space. In order to reduce the dimensionality of the feature space, a number of dimensionality reduction techniques can be employed.

The mean shift algorithm can be used for visual tracking. The simplest such algorithm would create a confidence map in the new image based on the color histogram of the object in the previous image, and use mean shift to find the peak of a confidence map near the object's old position. The confidence map is a probability density function on the new image, assigning each pixel of the new image a probability, which is the probability of the pixel color occurring in the object in the previous image. A few algorithms, such as ensemble tracking, CAMshift, expand on this idea.

Disadvantages

1) Features with lesser support in the feature space may not be detected inspite of being salient for the task to be executed

ISSN: 2321-9653

INTERNATIONAL JOURNAL FOR RESEARCH IN APPLIED SCIENCE AND ENGINEERING TECHNOLOGY (IJRASET)

*D. Efficient Graph-Based Image Segmentation*

This paper addresses the problem of segmenting an image into regions. It defne a predicate for measuring the evidence for a boundary between two regions using a graph-based representation of the image,then develop an eficient segmentation algorithm based on this predicate, and show that although this algorithm makes greedy decisions it produces segmentations that satisfy global properties. Then apply the algorithm to image segmentation using two different kinds of local neighborhoods in constructing the graph, and illustrate the results with both real and synthetic images. The algorithm runs in time nearly linear in the number of graph edges and is also fast in practice. An important characteristic of the method is its ability to preserve detail in low-variability image regions while ignoring detail in high-variability regions.

A wide range of computational vision problems could in principle make good use of segmented images, were such segmentations reliably and efficiently computable. For instance intermediate-level vision problems such as stereo and motion estimation require an appropriate region of support for correspondence operations. Spatially non-uniform regions of support can be identiified using segmentation techniques. Higher-level problems such as recognition and image indexing can also make use of segmentation results in matching, to address problems such as gure-ground separation and recognition by parts.

To develop computational approaches to image segmentation that are broadly useful, much in the way that other low-level techniques such as edge detection are used in a wide range of computer vision tasks. In order to achieve such broad utility, It is important that a segmentation method have the following properties:

*1)* Capture perceptually important groupings or regions, which often reflect global aspects of the image. Two central issues are to provide precise characterizations of what is perceptually important, and to be able to specify what a given segmentation technique does. There should be precise definitions of the properties of a resulting segmentation, in order to better understand the method as well as to facilitate the comparison of different approaches.

*2)* Be highly efficient, running in time nearly linear in the number of image pixels.In order to be of practical use, segmentation methods should run at speeds similar to edge detection or other low-level visual processing techniques, meaning nearly linear time and with low constant factors. For example, a segmentation technique that runs at several frames per second can be used in video processing applications.

It uses pairwise region comparison method. A new method for image segmentation based on pairwise region comparison. It have shown that the notions of a segmentation being too coarse or too can be defined in terms of a function which measures the evidence for a boundary between a pair of regions. This segmentation algorithm makes simple greedy decisions, and yet produces segmentations that obey the global properties of being not too coarse and not too done according to a particular region comparison function.

Disadvantages

1. Time complexity is high
2. Less efficient for creating partitions based on semantic scene segmentation

*E. Quick Shift and Kernel Methods for Mode Seeking*

The complexity of the recently introduced medoid-shift algorithm in clustering N points is O(N2), with a small constant, if the underlying distance is Euclidean. This makes medoid shift considerably faster than mean shift, contrarily to what previously believed. Then exploit kernel methods to extend both mean shift and the improved medoid shift to a large family of distances, with complexity bounded by the effective rank of the resulting kernel matrix, and with explicit regularization constraints. Finally, It show that, under certain conditions, medoid shift fails to cluster data points belonging to the same mode, resulting in over-fragmentation.

By introducing a novel, simple and extremely eficient clustering algorithm, called quick shift, that explicitly trades of under- and overfragmentation. Like medoid shift, quick shift operates in non-Euclidean spaces in a straightforward manner. It also show that the accelerated medoid shift can be used to initialize mean shift for increased eficiency. This paper also algorithms to clustering data on manifolds, image segmentation, and the automatic discovery of visual categories

The paper address this issue in two ways. First, propose using medoid shift to simplify the data and initialize the more accurate mean shift algorithm. Second, propose an alternative mode seeking algorithm that can trade of mode over and under

ISSN: 2321-9653



fragmentation.This algorithm, related to, is particularly simple and fast, yields surprisingly good segmentations, and returns a one parameter family of segmentations where model selection can be applied.

By introducing a novel, simple and extremely eficient clustering algorithm, called quick shift, that explicitly trades of under- and overfragmentation. Like medoid shift, quick shift operates in non-Euclidean spaces in a straightforward manner. It also show that the accelerated medoid shift can be used to initialize mean shift for increased eficiency. This paper also algorithms to clustering data on manifolds, image segmentation, and the automatic discovery of visual categories

The paper address this issue in two ways. First, propose using medoid shift to simplify the data and initialize the more accurate mean shift algorithm. Second, propose an alternative mode seeking algorithm that can trade of mode over and under fragmentation . This algorithm, related to, is particularly simple and fast, yields surprisingly good segmentations, and returns a one parameter family of segmentations where model selection can be applied

Disadvantages

1) Can't be used for the clustering of binary foreground and background sementation

## III. Task-Specific Image Partitioning for Region Based Image Retrieval

To overcome the above found disadvantages of various image partitioning methods,a new idea of task specific image partitioning,is introduced. A system which can be used for region based image retrieval by task specific image partitioning is proposed.

### A. Introduction

Region Based Image representations (RBIRs) have been shown to be effective in improving the performance of algorithms for high-level image/scene understanding, which encompasses tasks such as object class segmentation, scene segmentation, surface layout labeling, and single view 3D reconstruction . The effectiveness comes as a result of promoting the following three merits of using the RBIRs. First, the coherent support of a region, commonly assumed to be of a single label, serves as a good prior for many labeling tasks. Second, these coherent regions allow a more consistent feature extraction that can incorporate surrounding contextual information by pooling many feature responses over the region. Third, compared to pixels, a small number of larger homogeneous regions can significantly reduces the computational cost in the successive labeling task. In this paper, an image partitioning framework for obtaining RBIRs that realizes these benefits and improves the taskspecific labeling performance.

Up until now, using RBIRs in an image labeling system is a two-stage process: first, a general-purpose image partitioning method is used to obtain the RBIR, and second, this RBIR is used by a model that is trained with task-specific supervision. The first stage is task-oblivious but has direct influence on the performance of the model in the second stage. Ignoring the task at hand during creation of the RBIR is therefore a limitation of current systems that it address.

Consider Fig. 3.1 that shows ground-truth labeling for different labeling tasks for a given image. The ideal partitioning for the object-specific segmentation task would group each boat into one region and the remaining part into one background . Forthe task of semantic scene segmentation , the preference is to segment each distinct class – sky, tree, water, boat – into a region of its own. In surface layout labeling which does not distinguish between object classes, the preference is to segment the image into regions of coherent surface normals

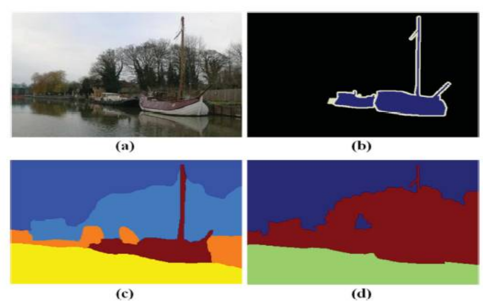



Fig. 3.1 Task Specific Image Partitioning Example

From the example explained above, it is obvious that a task-specific image partitioning algorithm would lead to partitioning that would be more conducive to the particular labeling task in hand than the general-purpose partitioning algorithm. With this goal in mind, explicitly address the task-specific image partitioning problem as follows: given an image and labeling task, produce a partitioning of the image into disjoint regions such that each region is homogeneous

ISSN: 2321-9653



with respect to the desired labeling of the task, and the labels of its neighboring regions are different.

There are several general-purpose unsupervised image partitioning algorithms for region-based image understanding. For instance, in the superpixel-based conditional random fields (CRFs) models , mean-shift , normalized cuts , graph-based local variation algorithm , and their variants such as quick-shift are used to obtain small coherent image regions, called superpixels. These a priori over-segmentations are not related to any task and maybe limited in capturing accurate global information for the successive region-labeling step. To enhance its ability, some recent CRFs are based on either a hierarchy of regions or a set of partitionings. These multiple partitioning are obtained using mean-shift segmentation with different kernel sizes, multiscale normalized cuts , a hierarchical segmentation with increasing edge strength , and a simple region-merging algorithm . These algorithms – while empirically successful to a certain extent – use task-oblivious partitionings and therefore do not address the task-specific image partitioning problem.

*B. System Overview*

It addresses the task-specific image partitioning problem using correlation clustering which is a graph partitioning algorithm that simultaneously maximizes intracluster similarity and inter-cluster dissimilarity. Here, the similarity and dissimilarity must be defined differently according to the task, and this is achieved by learning parameters using task specific training data. Since correlation clustering assigns a label to each edge, in contrast to other image partitioning algorithms, correlation clustering does not require a pre specified number of clusters and distance threshold for clustering. Furthermore, correlation clustering leads to linear discriminant functions which allow for large margin training based on structured support vector machine (S-SVM)

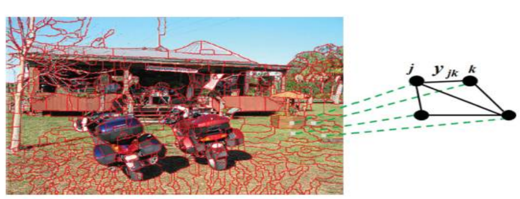



Fig 3.2 Illustration of a part of the graph built on superpixels.

Correlation clustering algorithm starts from a fine superpixel graph to reduce computational cost and also to extract a more meaningful discriminative features from larger consistent regions. To start with a fine superpixelization is typically not a limitation in practice as the number of fine superpixels is much larger (hundreds) than the final number of regions (tens). A rich pairwise feature vector on neighboring superpixels based on several visual cues is defined, and the correlation clustering problem is approximately solved using a linear programming (LP) relaxation technique. Correlation clustering is in general NP-hard, and therefore, the relaxation provides a polynomial-time approximation to its maximum a posteriori (MAP) solution.

For supervised training of the parameter vector, apply a decomposable structured loss function to handle imbalanced classes. incorporate this loss function into the cutting plane procedure for S-SVM training To summarize, main contributions are: *1)* a study on task-specific image partitioning that is suitable for any particular labeling problem at hand, *2)* a supervised correlation clustering on a superpixel graph for task-specific image partitioning; a rich feature vector is taken for robust partitioning, the LP relaxation is used for fast inference, and the SSVM with a modified label loss is used for task-specific training of the parameter vector, and *3)* an empirical validation of the proposed task-specific image partitioning that is more conducive to the successive labeling task in comparison to existing state-of-the-art partitioning algorithms.

After partitioning the image into regions, a region based image retrieval is performed from a database of images. For that a query image is partitioned to regions by using any particular partitioning method and the region based image retrieval is performed from a database where the images stored in the DB is also partitioned by similar partitioning method

A Database of partitioned images is created by any partitioning method, such that any task specific partitioning method can be used to make the database. A Query image is inputing to the system. It is partitioned to regions. Then similarity between query image and images in database is calculated. Then the maximum similar images are retrieved. This similarity measurement includes regionwise similarity and total similarity measurement.The image is taken, then

ISSN: 2321-9653



partition it then store it in database. Similarly a number of images are stored in the database.

## IV. Conclusions

A Literature survey had made on various papers which includes the techniques of image segmentation using normalized cuts, mean shift for feature space analysis , graph based image segmentation, quick shift, enforcing label consistency, object class segmentation, single image depth estimation and selecting regions for scene understanding. The proposed system addressing the problem of task-specific image partitioning by supervised training and a region based image retrieval by creating a database of partitioned images and inputing a query image.

The proposed correlation clustering model aims to merge superpixels into regions of homogeneity with respect to the solution of any particular image labeling problem. The LP relaxation was used to approximately solve the correlation clustering over a superpixel graph where a rich pairwise feature vector was defined based on several visual cues. The S-SVM was used for supervised training of parameters in correlation clustering, and the cutting plane algorithm with LP-relaxed inference was applied to solve the optimization problem of S-SVM.

After partitioning, a database of images is created. When a user is inputing a query image, by calculating region similarity and total image similarity the 'n' number of most similar images can be retrieved. The proposed partitioning framework is applicable to a broad variety of other high-level vision tasks.


## Acknowledgment

It is my privilege to render my heartfelt thanks and gratitude to my HOD,group tutor and all the staff members of the Computer Science Department of viswajyothi college of engineering for providing their assistance and support.


## References


[1] Sungwoong Kim, Sebastian Nowozin, Pushmeet Kohli and Chang D. Yoo, "Task-Specific Image Partitioning", in IEEE Transactions On Image Processing,vol. 22, no. 2, february 2013.

[2] Brian Fulkerson, Andrea Vedaldi and Stefano Soatto, "Class Segmentation and Object Localization with Superpixel Neighborhoods".

[3] J. Shi and J. Malik, "Normalized cuts and image segmentation", in IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 8, pp. 888–905, Aug.2000.

[4] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis", in IEEE Trans. Pattern Anal. Mach. Intell., vol. 24, no. 5, pp. 603–619, May 2002.

[5] P. Felzenszwalb and D. Huttenlocher, "Efficient graph-based image segmentation", in Int. J. Comput. Vis, 59, no. 2, pp. 167–181, 2004

[6] A. Vedaldi and S. Soatt, "Quick shift and kernel methods for mode seeking", in Proc. Eur. Conf. Comput. Vis, 2008, pp. 1–14

[7] P. Kohli, L. Ladick`y, and P. H. S. Torr, "Robust higher order potentials for enforcing label consistency", in Proc. IEEE Conf. Comput. Vis. Pattern Recognit, Jun. 2008, pp. 1–8

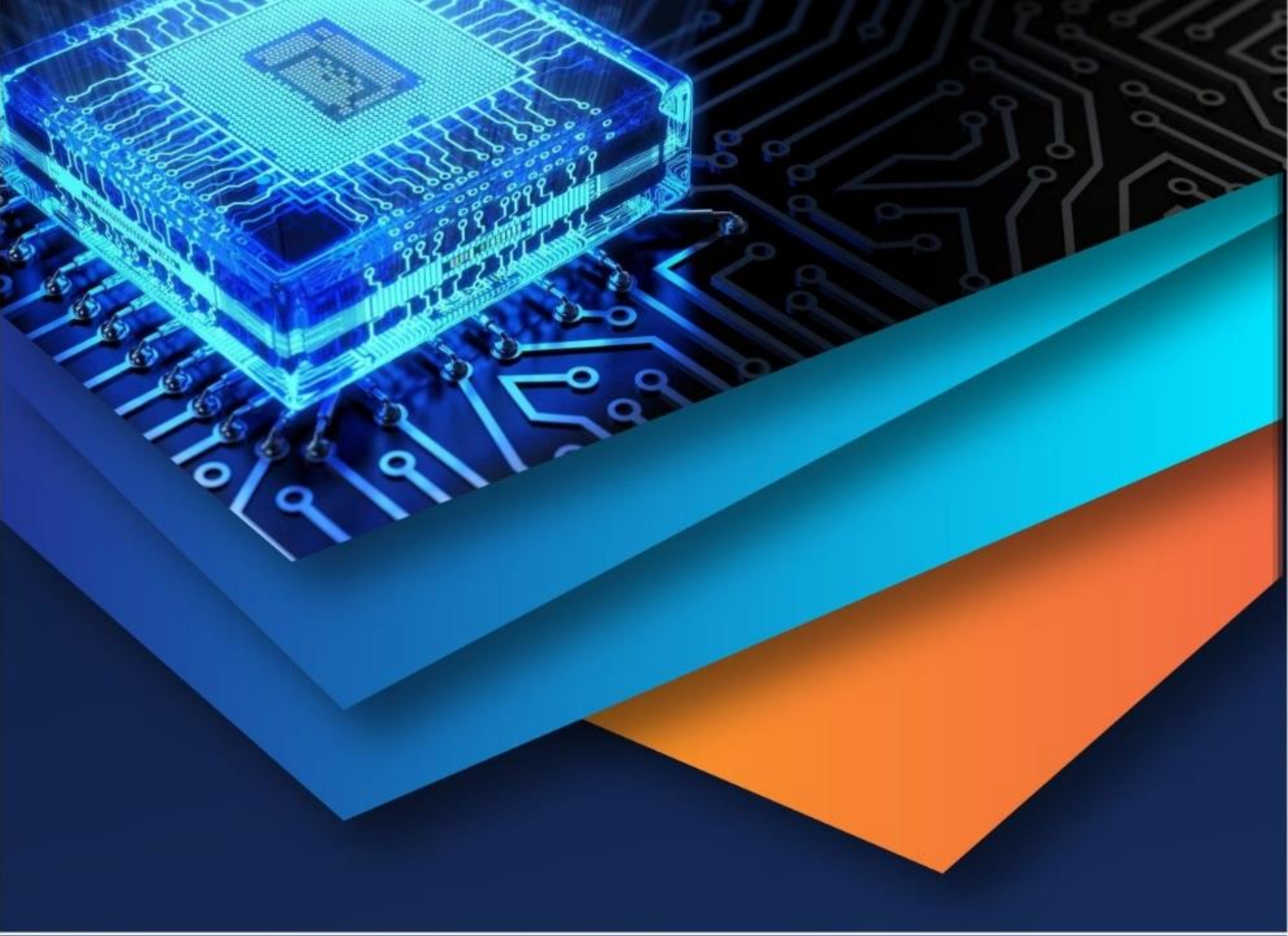
[8] L. Ladick`y, C. Russell, P. Kohli, and P. H. S. Torr, "Associative hierarchical CRFs for object class image segmentation", in Proc. IEEE Int. Conf. Comput. Vis,Sep.–Oct. 2009, pp. 739–746

[9] S. Gould, R. Fulton, and D. Koller, "Decomposing a scene into geometric and semantically consistent regions", in Proc. IEEE Int. Conf. Comput. Vis, Sep.–Oct. 2009, pp. 1–8.

[10] Stefania Ardizzoni, Ilaria Bartolini and Marco Patella, "Windsurf: Region-Based Image Retrieval Using Wavelets"

[11] B. Liu, S. Gould, and D. Koller, "Single image depth estimation from predicted semantic labels", in IEEE Conf. Comput. Vis. Pattern Recognit Jun. 2010, pp. 1253–1260.

[12] M. P. Kumar and D. Koller, "Efficiently selecting regions for scene understanding", in Proc. IEEE Conf. Comput. Vis. Pattern Recognit, Dec. 2010, pp. 1–8.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY