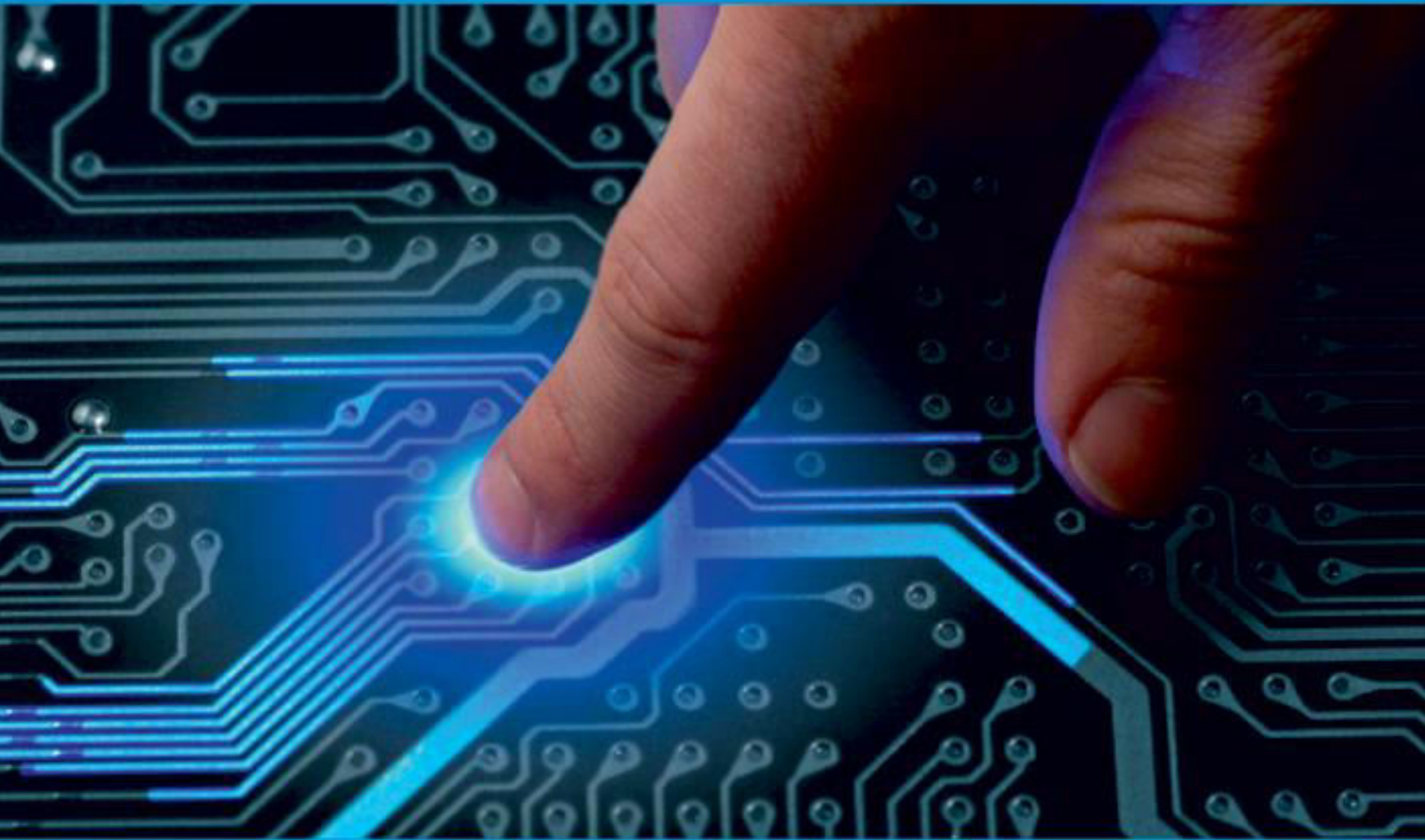
IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 11, November 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.542

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 11, November 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0911013 |

Performance Evaluation of Vehicle Classification using Colour and Deep Features

Mutyala Anusha[1], Sarakanam S Manikanta[2]

M.Tech Student, Department of ECE, Chaitanya Institute of Science and Technology Kakinada, India[1]

Assistant Professor, Department of ECE, Chaitanya Institute of Science and Technology Kakinada, India[2]

ABSTRACT: Vehicle similarity identification from a large image database is a critical task. The solution to this problem is the use of avehicle Image Retrieval (VIR) System. The images are described through their content, there is three predominant contentexisting in an image like color, shape, and texture. In this paper, we are evaluating the performance of the VIR systemusing two methods. The first method consists of colour and texture features. The second method consists of Use of CNN. The feature extraction technique is achieved based on an inputquery image from the database and features are saved in a feature dataset. A proposed strategy retrieves similar imagesfrom a database that fulfills the user's desire. The similarity measurement can be done using the Euclidean distance andhashing technique. The overall performance of the retrieval system has been analyzed through the parameters accuracy, Precisionand Mean Average Precision. The experimental result shows encouraging results using CNN which leads to improving accuracy

KEYWORDS: vehicle recognition, classification, CNN, color moments, texture, feature extraction.

1.INTRODUCTION

The explosive increase and ubiquitous accessibility of visual data on the Web have led to the prosperity of research activity in image search or retrieval. With the ignorance of visual content as a ranking clue, methods with text search techniques for visual retrieval may suffer inconsistency between the text words and visual content. Content-based image retrieval (VIR), which makes use of the representation of visual content to identify relevant images, has attracted sustained attention in recent two decades. Such a problem is challenging due to the intention gap and the semantic gap problems. Numerous techniques have been developed for content-based image retrieval in the last decade.With the universal popularity of digital devices embeddedwith cameras and the fast development of Internet technology, billions of people are projected to the Web sharing and browsing photos. The ubiquitous access to bothdigital photos and the Internet sheds bright light on manyemerging applications based on image search. Image searchaims to retrieve relevant visual documents to a textual orvisual query efficiently from a large-scale visual corpus.

Although image search has been extensively explored sincethe early 1990s [1], it still attracts lots of attention fromthe multimedia and computer vision communities in thepast decade, thanks to the attention on scalability challengeand emergence of new techniques. Traditional image searchengines usually index multimedia visual data based on thesurrounding meta data information around images on theWeb, such as titles and tags. Since textual information maybe inconsistent with the visual content, content-based imageretrieval (VIR) is preferred and has been witnessed to makegreat advance in recent years.In content-based visual retrieval, there are two fundamental challenges, i.e., intention gap and semantic gap.The intention gap refers to the difficulty that a user suffers to precisely express the expected visual content bya query at hand, such as an example image or a sketchmap. The semantic gap originates from the difficulty indescribing high-level semantic concept with low-level visualfeature [2] [3] [4].

To narrow those gaps, extensive effortshave been made from both the academia and industry.
From the early 1990s to the early 2000s, there havebeen extensive study on content-based image search. Theprogress in those years has been comprehensively discussedin existing survey papers [5] [6] [7]. Around the early 2000s,the introduction of some new insights and methods triggersanother research trend in VIR. Specially, two pioneeringworks have paved the way to the significant advance incontent-based visual retrieval on large-scale multimediadatabase. The first one is the introduction of invariant localvisual feature SIFT [8]. SIFT is demonstrated with excellentdescriptive and discriminative power to capture visual content in a variety of literature. It can well capture the invariance to rotation and scaling transformation and is robust toillumination change. The second work is the introduction ofthe Bag-of-Visual-Words (BoW) model [9]. Leveraged frominformation retrieval, the BoW model makes a

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 11, November 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0911013 |

compactrepresentation of images based on the quantization of thecontained local features and is readily adapted to the classicinverted file indexing structure for scalable image retrieval.

Based on the above pioneering works, the lastdecade has witnessed the emergence of numerouswork on multimedia content-based image retrieval[10] [11] [12] [13] [9] [14] [15] [16] [17] [18] [19] [20] [21][22] [23] [24] [25] [26] [27] [28] [29]. Meanwhile, in industry, some commercial engines on content-based imagesearch have been launched with different focuses, suchas Tineye1, Ditto2, Snap Fashion3, ViSenze4, Cortica5, etc.Tineye is launched as a billion-scale reverse image searchengine in May, 2008. Until January of 2017, the indexedimage database size in Tineye has reached up to 17 billion.Different from Tineye, Ditto is specially focused on brandimages in the wild. It provides an access to uncover thebrands inside the shared photos on the public social mediaweb sites.

Color Features: Images are largely categorized intograyscale images and color images. In a grayscale imagecolor pixel having a solely grayscale area while in acolor image three color intensity ranges are used. Inthe color image red, green and blue intensities are used.Color histogram, color coherence, and color moments areimportant methods used for image retrieval

Texture Features: It measures the homogeneity of apixel over repeated patterns in the image. We can formata retrieval system the use of two tactics particularlystructural and frequency-based approachesShape Features: It gives edges or outlines of an objectexisting in an image. Region and boundary-basedtechniques are used in the retrieval systems based totallyon shape features.

Neural Network: A neural network consists of the inputlayer, hidden layer, and output layer. Convolution NeuralNetwork is used for feature extraction from images

II. LITERATURE SURVEY

Even though Multimedia databases (MMD) is amongthe fastest growing emerging technologies in thefield of database systems. New technologies posenumerous challenges, and MMD has its share ofchallenges. Most of MMD challenges are aroundContent-based Image Retrieval (VIR) systems. VIR isa technique for retrieving images on the basis ofautomatically-derived features such as color, texture andshape. Moreover, multimedia objects contain encodingof raw sensorial data, which compromise the efficientindexing and retrieval. As result of which, Query byImage Content (QBIC) technique using imagedescriptors for indexing and retrieval of multimediaobjects were proposed by various studies to address thisproblem. However, an effective and preciseperformance evaluation benchmarking for thistechnique remains exclusive.

Since the invent of the Internet, and theavailability of image capturing devices such as smartphones, digital cameras, image scanners and geospatialsatellite devices, the size of digital image storage isincreasing rapidly. Efficient image searching, browsingand retrieval tools are required by end users fromvarious domains, including remote sensing, fashiondesign, criminology, publishing, medicine, architecture,etc. It is for this reasons that, many general purposeimage retrieval systems have been developed. Therefore,for the same reasons we explore the in-depth survey ofcontent-based image retrieval technology, descriptorstechnology and performance measure frameworktechnology in order to gain an insight of this domain field.

The main object of a Content-Based Image Retrieval(VIR) system, also known as Query by Image Content(QBIC), is to help users to retrieve relevant images basedon their contents. VIR technologies provide a method tofind images in large databases by using unique descriptorsfrom a trained image. The image descriptors includetexture, color, intensity and shape of the object inside animage. The urgency of efficient image searching, browsingand retrieval techniques by users from large repositoriessuch as the internet, metrological images and geospatialimages is real.

It is reported by [5] that, there are two retrievalframeworks: text-based and content-based. In the textbased approach, the images are manually annotated by textdescriptors, which are then used by a databasemanagement system to perform image retrieval. There aretwo disadvantages with this approach. The first is that ahuman labor at considerable level is required for manualannotation. The second is the inaccuracy in annotation dueto the subjectivity of human perception. To overcomethese disadvantages in text-based retrieval system,content- based image retrieval (VIR) was introduced.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 11, November 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0911013 |

It is asserted by [24], that content-based imageretrieval (VIR), also known as query by image content(QBIC) and content-based visual information retrieval(CBVIR), is the application of computer visiontechniques to the image retrieval problem. It is atechnique which uses visual features of image such ascolor, shape, texture, etc. to search user required imagefrom large image database according to user's requestsin the form of a query image. Images are retrieved onthe basis of similarity in features where features of thequery specification are compared with features from theimage database to determine which images matchsimilarly with given features.

It is defined by [18] that, in computer vision, visualdescriptors or image descriptors are defined as thedescriptions of the visual features of the contentsin images, videos, or algorithms or applications thatproduce such descriptions. They describe elementarycharacteristics such as the shape, the color, the texture orthe motion, among others.It is describe by [28], that visual descriptors aredivided in two main groups: General informationdescriptors, which they contain low level descriptors whichgive a description about color, shape, regions, textures andmotion, and specific domain information descriptors whichthey give information about objects and events in the scene.

In their book [6] describe the general informationdescriptors as consisting of a set of descriptors that coversdifferent basic and elementary features like: color, texture,shape, motion, location and others. The color descriptor isthe most basic quality of visual content. Five tools aredefined to describe color; Dominant Color Descriptor(DCD), Scalable Color Descriptor (SCD), Color StructureDescriptor (CSD), Color Layout Descriptor (CLD), andGroup of frame (GoF) or Group-of-pictures (GoP). TheTexture descriptors are used to characterize image,textures, or regions. They observe the region homogeneityand the histograms of these region borders. The set ofdescriptors is formed by: Homogeneous TextureDescriptor (HTD), Texture Browsing Descriptor (TBD), andEdge Histogram Descriptor (EHD). The Shape descriptorcontains important semantic information due to human'sability to recognize objects through their shape.

However, this information can only be extracted bymeans of a segmentation similar to the one that the humanvisual system implements. These descriptors describeregions, contours and shapes for 2D images and for 3Dvolumes. The shape descriptors are formed by; Regionbased Shape Descriptor (RSD), Contour-based ShapeDescriptor (CSD) and 3-D Shape Descriptor (3-D SD).While, the Motion descriptors are defined by four differentdescriptors which describe motion in video sequence. Thedescriptor set is formed by; Motion Activity Descriptor(MAD), Camera Motion Descriptor (CMD), MotionTrajectory Descriptor (MTD), and Warping and ParametricMotion Descriptor (WMD and PMD). Finally, the Locationdescriptor element's location in the image is used todescribe elements in the spatial domain.

III. PROPOSED METHOD

The process of Vehicle image retrieval (VIR) system consists of the following sixmain stages of: image acquisition, image pre-processing,feature extraction, similarity matching, resultantretrieval image and user interface and feedback.

3.1Image acquisition

It is the process of acquiring a digital image fromthe image database. The image database consists of thecollection of n number of images depends on the userrange and choice.

3.2 Image pre-processing

It is the process of improving the image in waysthat increases the chances for success of the otherprocesses. The image is first processed in order toextract the features, which describe its contents. Theprocessing involves filtering, normalization,segmentation, and object identification. Imagesegmentation is the process of dividing an image intomultiple parts. The output of this stage is a set ofsignificant regions and objects.

3.3 Feature Extraction

It is the process where features such as shape,texture, colour, etc. are used to describe the content of theimage. The features further can be classified as low-leveland high-level features. In this stage visual information isextracts from the image and saves them as features vectorsin a features database. For each pixel, the image descriptionis found in the form of feature value (or a set of valuecalled a feature vector) by using the feature extraction. These feature vectors are used to compare the query withthe other images and retrieval.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 11, November 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0911013 |

3.4 Similarity Matching
It is a process that entails the information about eachimage is stored in its feature vectors for computationprocess and these feature vectors are matched with thefeature vectors of query image (the image to be search inthe image database whether the same image is present ornot or how many are similar kind images are exist or not)which helps in measuring the similarity. This step involvesthe matching of the above stated features to yield a resultthat is visually similar with the use of similarity measuremethod called as Distance method. There are variousdistances methods available such as Euclidean distance,City Block Distance, and Canberra Distance.

3.5. Resultant Retrieved images
It is the process that searches the previouslymaintained information to find the matched images fromdatabase. The output will be the similar images havingsame or very closest features as that of the query image.

3.6 User interface and feedback
It is the process which governs the display of theoutcomes, their ranking, the type of user interaction withpossibility of refining the search through some automaticor manual preferences scheme etc. The Figure 1 belowdemonstrates the VIR System and its various components.

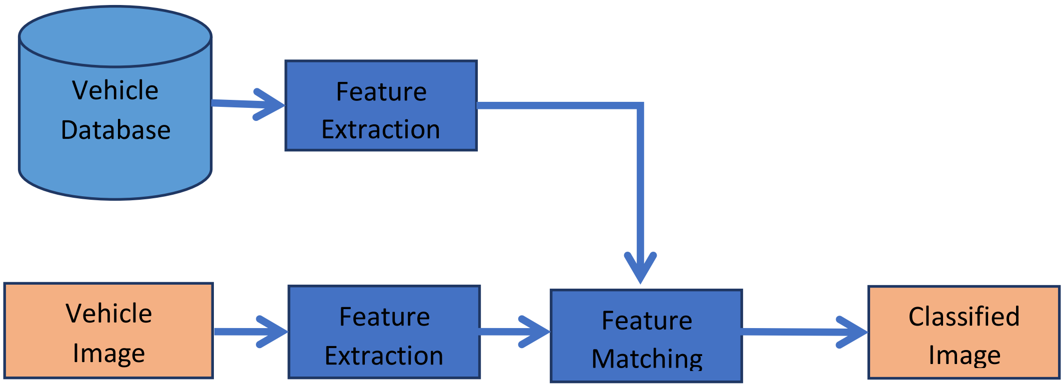


Figure.1 Vehicle Image Retrieval System

IV. FEATURE EXTRACTION

There has been tremendous work on various ways to deal withthe recognition of different kinds of features in images. Thesefeatures can be classified as follows:

4.1 Low Level Features
Features in this category are all application independent, e.g.color, texture, and shape. According to concept level, they canbe further divided into:

- Pixel-level features

features be determined at every pixel,for example color, area, and the first and secondderivatives of gray-scale values at every pixelPixels are extracted and stored in an array. The arraycontains the RGB components of each pixel. Each pixelin the image is then processed to identify the featurevectors of the image. Edges were used as the onlyfeature vector.

- Local features

The local image description isestablished on the reason that images can be described bycharacteristics registered on regions of the image. Can bedetermined over the consequences of image division andedge detection algorithms. Object shape is an example ofsuch feature [14].

- Global features

The global image descriptor iscomposed by color and texture features being computed.

- Texture Feature Extraction

The second element of the new system is the texture feature. For this purpose, EHD algorithm is used. Texture is an important feature of expected images. A variety of techniques have been proposed for estimating texture comparability. These strategies ascertain proportions of image texture, for example, the level of differentiation, coarseness, directionality and consistency [8]; or periodicity, directionality and randomness.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 11, November 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0911013 |

4.2. CNN Based Features
This is our proposed frame work for utilizing features from a pretrained deep CNN. We extract features from Pre-trained VGG16deep CNN model for image retrieval task. A deep CNN modelusually consists of many layers that incrementally calculatefeatures. As outlined in Fig. 2. deep CNN model incrementallylearns the features through layers of convolutions andsubsampling.

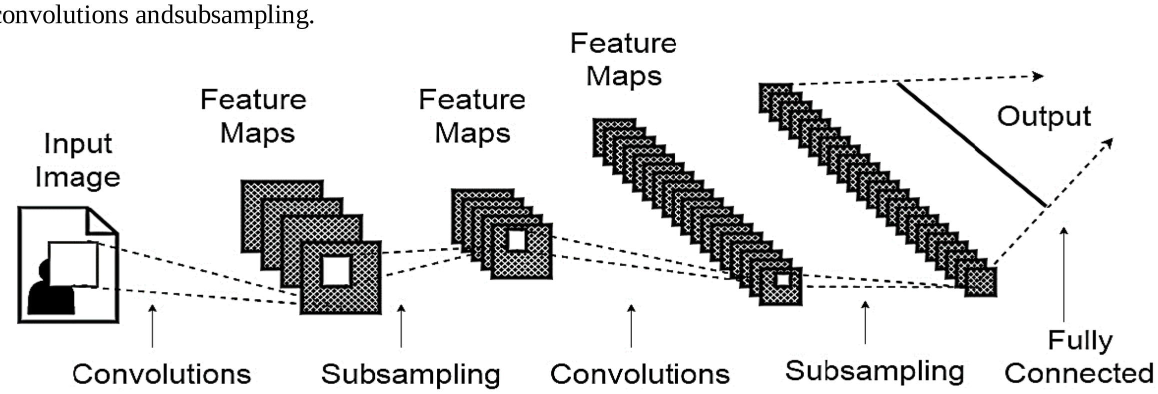


Figure.2 Architecture of Convolution neural networks

In this work, VGG16 deep CNN model is implemented fromPython Keras package. It is a 16-layer deep CNN created by theVisual Geometry Group from University of Oxford [29]. VGG16model is trained on ImageNet, which is a very large-scale datasetcontaining 3+ million digital images distributed across 5000+categories. VGG16 model consists of 5 convolution blocks andeach convolution block contains two convolution layers (size 3X3)and one maxpooling layer (size 2X2). The final classification stepof the model consists of fully connected (FC) layers. Ouralgorithm extracts 4096 features from fully connected FC2. This is output of second and penultimatefully connected layer of the pre-trained VGG16 CNN model. Thefeature extraction is done for each image in the dataset and queryimages.

V. EXPERIMENTAL RESULTS

Our experiment on baseline VIR with handcrafted features(Colour, texture and shape) yields us an average precision of73.25% across all classes of the weather images dataset. Theproposed VIR method which uses pre-trained VGG16 deepCNN features achieves an average precision of 86.73% acrossall classes of the dataset. The improvement in precision rate isobserved across all image classes. Fig. 3. depicts theimprovement in precision as recorded across different retrievalsizes. The improvement in precision rate in Clearimage class is lower as compared to other three classes(Cloudy, Rain and Sunrise), where precision improvement isprofound. Experimental results show that our proposed VIRframe work using features from pre-trained VGG16 CNN modelperforms better than traditional VIR using handcrafted features(Color, texture and shape). The improvement in performance isseen across the fetch sizes and image classes.

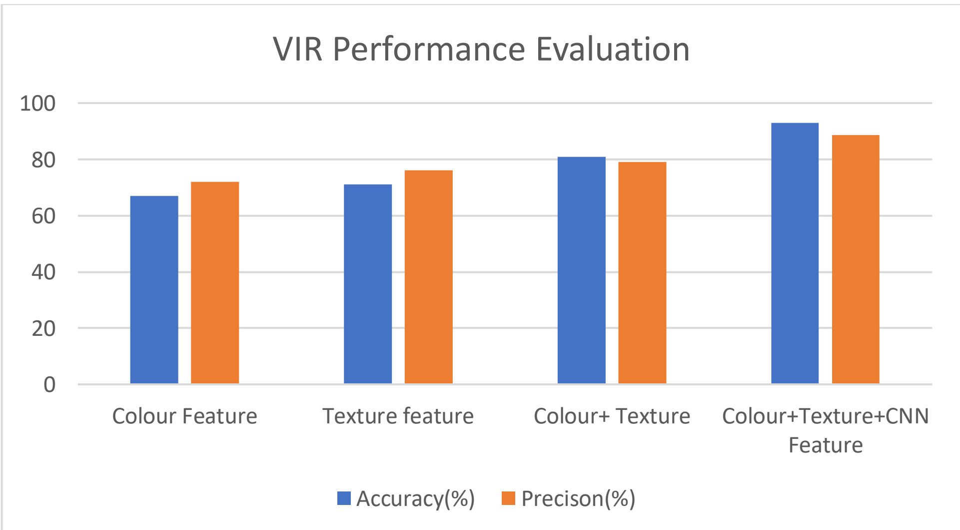


Figure.3 Accuracy and precision plots of VIR system

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542
|| Volume 9, Issue 11, November 2021 ||
| DOI: 10.15680/IJIRCCE.2021.0911013 |

Figure.4. Proposed method Experimental Results

V. CONCLUSION

This system works for searching and retrieving images.Regarding the "huge" size of the database, our systemprovided good results. Using more performance measures fineadjustments can be made with more features and possiblyprovide the users with the best options of retrieval as defaultparameters, the system attempts to present a hybrid techniquefor VIR, which uses the combination of Feature extractionwith better image retrieval accuracy. The proposed systemmatches the images if the dominant color is similar. Thislimitation can be resolved by using more than one feature options to represent the image. In the next section, some ideasto enhance the system have been stated.The present work can be extended by improving therecognition rate by increasing the feature vectors and using acombined approach to retrieve similar images. The presentimplementation has an application in lot of fields such asmilitary, medicine, crime detection, etc. An embedding ofnumber plate recognition program with this method will helpto identify vehicles automatically, which will help in findingstolen vehicles. Given an image database of vehicles, theprogram can retrieve similar images of cars from database inaccordance to input image. Furthermore using number platerecognition program the user can search for number plates bygiving the number as query, and retrieve information aboutthe vehicle. Further studies regarding measuring theperformance of more options and 3D visualization of thesearch results are currently being investigated.

REFERENCES

[1] Y. Rui, T. S. Huang, M. Ortega, and S. Mehrotra, "Relevance feedback: a power tool for interactive content-based image retrieval,"IEEE Transactions on Circuits and Systems for Video Technology,vol. 8, no. 5, pp. 644–655, 1998.
[2] A. Alzubi, A. Amira, and N. Ramzan, "Semantic content-basedimage retrieval: A comprehensive study," Journal of Visual Communication and Image Representation, vol. 32, pp. 20–54, 2015.21. http://acmmm13.org/submissions/call-for-multimedia-grandchallenge-solutions/msr-bing-grand-challenge-on-image-retrievalscientific-track22.tp://tianchi.liyun.com/competition/introduction.htm?spm=5176.100069.5678.1.SmufkG&raceId=231510&lang=en US18
[3] X. Li, T. Uricchio, L. Ballan, M. Bertini, C. G. Snoek, and A. D.Bimbo, "Socializing the semantic gap: A comparative survey onimage tag assignment, refinement, and retrieval," ACM Computing Surveys (CSUR), vol. 49, no. 1, p. 14, 2016.
[4] Z. Lin, G. Ding, M. Hu, and J. Wang, "Semantics-preservinghashing for cross-view retrieval," in IEEE Conference on ComputerVision and Pattern Recognition (CVPR), 2015, pp. 3864–3872.
[5] A. W. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain,"Content-based image retrieval at the end of the early years,"IEEE Transactions on Pattern Analysis and Machine Intelligence,vol. 22, no. 12, pp. 1349–1380, 2000.
[6] M. S. Lew, N. Sebe, C. Djeraba, and R. Jain, "Content-based multimedia information retrieval: State of the art and challenges,"ACM Transactions on Multimedia Computing, Communications, andApplications (TOMM), vol. 2, no. 1, pp. 1–19, 2006.
7] Y. Liu, D. Zhang, G. Lu, and W.-Y. Ma, "A survey of contentbased image retrieval with high-level semantics," Pattern Recognition, vol. 40, no. 1, pp. 262–282, 2007.
[8] D. G. Lowe, "Distinctive image features from scale invariantkeypoints," International Journal of Computer Vision, vol. 60, no. 2,pp. 91–110, 2004.
[9] J. Sivic and A. Zisserman, "Video Google: A text retrieval approach to object matching in videos," in IEEE Conference onComputer Vision and Pattern Recognition, 2003, pp. 1470–1477.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.542

|| Volume 9, Issue 11, November 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0911013 |

[10] D. Nister and H. Stewenius, “Scalable recognition with a vocabulary tree,” in IEEE Conference on Computer Vision and PatternRecognition, vol. 2, 2006, pp. 2161–2168.
[11] J. Philbin, O. Chum, M. Isard, J. Sivic, and A. Zisserman, “Objectretrieval with large vocabularies and fast spatial matching,” inIEEE Conference on Computer Vision and Pattern Recognition, 2007,pp. 1–8.
[12] H. Jegou, M. Douze, and C. Schmid, “Hamming embedding andweak geometric consistency for large scale image search,” inEuropean Conference on Computer Vision, 2008, pp. 304–317.
[13] W. Zhou, H. Li, Y. Lu, and Q. Tian, “Large scale image search withgeometric coding,” in ACM International Conference on Multimedia,2011, pp. 1349–1352.
[14] O. Chum, J. Philbin, J. Sivic, M. Isard, and A. Zisserman, “Totalrecall: Automatic query expansion with a generative featuremodel for object retrieval,” in International Conference on ComputerVision, 2007, pp. 1–8.
[15] J. Philbin, O. Chum, M. Isard, J. Sivic, and A. Zisserman, “Lostin quantization: Improving particular object retrieval in largescale image databases,” in IEEE Conference on Computer Visionand Pattern Recognition, 2008, pp. 1–8.
[16] O. Chum, J. Philbin, and A. Zisserman, “Near duplicate imagedetection: min-hash and tf-idf weighting,” in British MachineVision Conference, vol. 3, 2008, p. 4.
[17] Z. Wu, Q. Ke, M. Isard, and J. Sun, “Bundling features for largescale partial-duplicate web image search,” in IEEE Conference onComputer Vision and Pattern Recognition, 2009, pp. 25–32.
[18] W. Zhou, Y. Lu, H. Li, Y. Song, and Q. Tian, “Spatial codingfor large scale partial-duplicate web image search,” in ACMInternational Conference on Multimedia, 2010, pp. 511–520.
[19] O. Chum, A. Mikulik, M. Perdoch, and J. Matas, “Total recallII: Query expansion revisited,” in IEEE Conference on ComputerVision and Pattern Recognition, 2011, pp. 889–896.
[20] Y. Zhang, Z. Jia, and T. Chen, “Image retrieval with geometrypreserving visual phrases,” in IEEE Conference on Computer Visionand Pattern Recognition, 2011, pp. 809–816.
[21] X. Zhang, L. Zhang, and H.-Y. Shum, “Qsrank: Query-sensitivehash code ranking for efficient ǫ-neighbor search,” in IEEE Conference on Computer Vision and Pattern Recognition, 2012, pp. 2058–2065.
[22] J. He, J. Feng, X. Liu, T. Cheng, T.-H. Lin, H. Chung, and S.-F. Chang, “Mobile product search with bag of hash bits andboundary reranking,” in IEEE Conference on Computer Vision andPattern Recognition, 2012, pp. 3005–3012.
[23] R. Arandjelovic and A. Zisserman, “Three things everyoneshould know to improve object retrieval,” in IEEE Conference onComputer Vision and Pattern Recognition, 2012, pp. 2911–2918.
[24] S. Zhang, M. Yang, T. Cour, K. Yu, and D. N. Metaxas, “Queryspecific fusion for image retrieval,” in European Conference onComputer Vision (ECCV), 2012.
[25] Q. Tian, S. Zhang, W. Zhou, R. Ji, B. Ni, and N. Sebe, “Buildingdescriptive and discriminative visual codebook for large-scaleimage applications,” Multimedia Tools and Applications, vol. 51,no. 2, pp. 441–477, 2011.
[26] W. Zhou, H. Li, Y. Lu, and Q. Tian, “Large scale partial-duplicateimage retrieval with bi-space quantization and geometric consistency,” in IEEE International Conference Acoustics Speech and SignalProcessing, 2010, pp. 2394–2397.
[27] S. Zhang, Q. Tian, G. Hua, Q. Huang, and S. Li, “Descriptivevisual words and visual phrases for image applications,” in ACMInternational Conference on Multimedia, 2009, pp. 75–84.
[28] S. Zhang, Q. Huang, G. Hua, S. Jiang, W. Gao, and Q. Tian,“Building contextual visual vocabulary for large-scale image applications,” in ACM International Conference on Multimedia, 2010,pp. 501–510.
[29] W. Zhou, Q. Tian, Y. Lu, L. Yang, and H. Li, “Latent visual contextlearning for web image applications,” Pattern Recognition, vol. 44,no. 10, pp. 2263–2273, 2011.