International Journal of Scientific Research & Engineering Trends
Volume 6, Issue 2, Mar-Apr-2020, ISSN (Online): 2395-566X

Review of Image Search using Hybrid Techniques

M. Tech. Scholar Sakina Bohra, Assistant Professor Prof. Lokesh Parashar
Dept. of Computer Science & Engg
Patel College of Science & Technology , Indore , India
sbohra47@gmail.com,lokesh23324@gmail.com

***Abstract* – The big digital image databases are yielded by the widespread of smart devices along with the exponential growth of virtual societies. If not fused with efficient Content-Based Image Retrieval (CBIR) tools, these databases can be counter-productive. The introduction of promising CBIR systems had been witnessed last decade which promoted applications in various fields. In this editorial, ananalysis on state of the art content-based image retrieval which includes theoretical and empirical work, is propounded. This work comprises of publications that cover aspects of research relevant to CBIR area. That is to say, unsupervised and supervised education and combination techniques along with which the low-level image visual descriptors have been reported. Furthermore, challenges and applications that appeared to carry CBIR research have been discussed in this work.**

***Keywords*– Image retrieval; Content-based image retrieval; supervised learning; unsupervised learning.**

I. INTRODUCTION

The significance of digital image databases depends upon how affably and accurately users can retrieve images of their interest. Consequently, retrieval tools and advanced search have been perceived as an urgent need for several image retrieval applications. Text-based image retrieval approaches have been adopted by the earliest search engines. Because of the digital images which are to be mined are either not labeled or annotated using inaccurate keywords are the results that have shown drastic limitations. I.e., text-based retrieval approaches necessitate manual appendix to whole of the image collections. On the other hand, this monotonous manual task is not viable for large image databases.

Tooutshine the challenges met by text-based image retrieval solutions Content-Based Image Retrieval (CBIR) appeared as a promising substitute. As per the fact, digital images, which are mined using CBIR systems, are represented by the use of a set of visual features. As depicted in Figure 1, typical CBIR system consists of an offline phase which targets the extraction and storage of the visual feature vectors which comes from the database images.

En contraire, the online phase permits the user to begin the retrieval task by providing his query image. In the end, a set of images visually relevant to the user query has been returned by the typical CBIR system. Although, its main drawback comprises of the assumptionsthat, the visual similarity imitates the semantic resemblance. Because of the semantic gap [1] between the higher level meaning and the low-level visual features this assumption does not hold anything. In spite of the promising results attained by large-scale applications, such as Yahoo! and Google TM, bridging

the semantic gap remains a difficult task for CBIR researchers. In addition to this, the social network usage, with the widespreadof the low-cost-smart-devices, has re-boosted the research associated to image retrieval. This witnessed a paradigm-shift in the research aims of the novel generation of researchers ofCBIR. Image representation, feature extraction and similarity computation also work as a grave component of archetypal CBIR systems. More specifically, researchers have investigated various components and contributions in order to design successful CBIR system, [2,3,4]. Wide-ranging surveys on CBIR systems have been propounded to report the growth reached by the research community [5,6,7]. Other surveys have been convoluted on highly relevant topics to CBIR systems. Namely, researches on high-dimensional data indexing [8], relevance feedback [9], and medical application of CBIR [10, 11] have been surveyed.

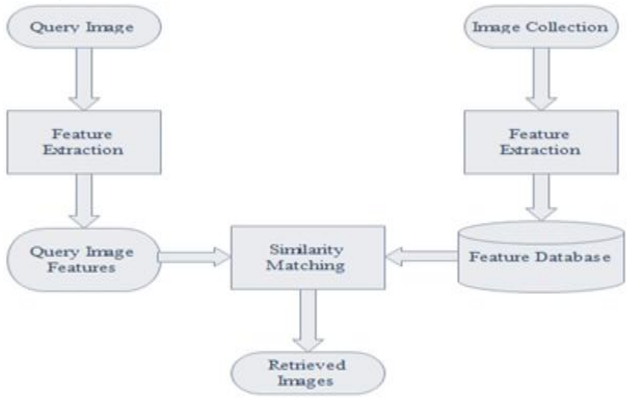

Fig.1. Overview of Typical CBIR System.

The chief inspiration for the survey is the progressive growth of associated research traversing several domains all through the last decade and the increase in the number of researchers investigating CBIR. This literary compilation is a sincere appraisal tostate of the art research and future facet of CBIR systems. The rest of this article is arranged as follows: Section 2 focuses on state of the art methods in order to bridge down the semantic gap. Low-level features proposed to capture high-level query semantic are highlighted in Section 3. Emerging research issues related to CBIR systems are introduced in Section 4. Finally, Section 5gives the conclusion.

II. BRIDGING THE SEMANTIC GAP

Depending on the adopted angle of view, the approach in the direction of bridging the semantic gap can be diversely classified. Web image retrieval, art image retrieval, scenery image retrieval are some perceived art techniques. This excerpt discusses approaches that were employed with reference to high level semantics [12]. The techniques are slotted on the basis of:

- Supervised and unsupervised learning methods to study the connection between low level descriptors and high level semantic based CBIR
- Image retrieval based on fusion

1. Supervised and Unsupervised Learning:

For a strong and perceptually relevant image ranking, the drawbacks of single similarity measures have been frequently explained by the researchers. For combating this limit, solutions that rely on learning have been proposed time and again. For a large collection, image classification has been framed as a pre-processing phase to speedup image retrieval. Visualization performance and retrieval process has been enhanced by adopting unsupervised learning. This happens when the images are not explicated. The classification approaches form the fundamentals of the retrieval process [13].

For both supervised and unsupervised learning and their utility in diverse terrains, recent developments have been very progressive. The upcoming section is an elaborated excerpt that deals with the techniques and utility related to CBIR. The idea is to promote brisk classification techniques and discover hacks to counter every limitation associated with the approach. The process of fragmenting patterns into uniform categories in an unsupervised manner is known as clustering. The notion of clustering is to facilitate the visualization and retrieval potential of the system. However, the system still has a lot many challenges to face. Diverse taxonomies of clustering methods have been introduced by different authors. A binary membership value assignment is undertaken for hard clustering, irrespective of the situation of the data instance. Partitioned clustering depends on hard objective function optimization. The real world applications have clusters overlapping one another. Thus, it is not really possible to differentiate between instances laying on the superimposing boundaries. A popular fuzzy clustering algorithm is the Fuzzy C-Means (FCM) algorithm. This enables slow assessment of the instances within a group. These algorithms, however fail to explore the ground truth distribution of data in case it contains asymmetric clusters. An alternative to fuzzy clustering is probabilistic modeling. Mixture modeling assumes the inheritance of clusters and work towards parameter distribution approximation. A recent proposal enables issuing of the data instances from diverse density functions. Such an approach can be classed into: statistical modeling, relational and objective function based paradigm [14].

Each cluster is taken as a restrictively distributes pattern in clustering that depends on statistical modeling. The absolute dataset is thus modeled as a distribution mixture. To approximate the parameters of the mixture components with respect to the cluster properties, the expectation maximization algorithm is employed. The benefit of such an approach is the information provided by it as per the data densities. It is not obligatory to model the mixture components as multivariate distribution. Conventionally, this technique denotes dataset for precise classification and not clustering. Relational approaches do not have a critical mathematical denotation of data points. This is the reason of its wide application in terrains with complex contemplation of image signatures. The relational methods cost a good time due to their prolix computation course. A spectral clustering algorithm is propounded by the researchers in order to group identical images into uniform clusters. The obtained information on partition is used to boost the retrieval process [15].

The sum of intra cluster distance is reduced to a minimum value by the K-means algorithm. A mandatory specification of the number of clusters is the limitation of this algorithm. This is countered by gradual increase in the number of clusters until the mean distance between an instance and its corresponding cluster centre reaches a predefined threshold. To find the number of image clusters, the competitive agglomeration algorithm is used [16].

From an application perspective, researchers from the multimedia community dedicated more attention to Web image clustering. The unsupervised learning techniques are valuable when meta-data is collected in addition to visual descriptors. Unsupervised learning generally serves to recognize new images and assign them to some preset categories before continuing with the retrieval phase. Identically, classification techniques can be grouped into two major categories. The generative modeling based approaches and the discriminative modeling approaches such as decision trees and SVM classifiers where the class boundaries and the posterior chances are learned. The generative modeling uses Bayes formula along with the densities of data instances within each class to approximate the posterior probabilities. Bayesian classification was adopted to propose an image retrieval system.It is also used for system that aimed to capture high-level concepts of natural scenes using low-level features. Images were automatically classified into outdoor or indoor images. Bayesian network was adopted for indoor/outdoor image classification [17].

Image classification using SVM as supervised learning technique has been propounded.Lately, advanced multimedia query processing systems using SVM based MIL framework has been proposed .MIL structure considers l training images as labeled bags where the bag has a set of instances represents a region i extracted from a training image i, and indicates a positive or negative example for a given class value. The mapping of these bags to a new feature space, where supervised learning technique can be trained to classify unlabelled instances, is the key component of MIL [18].

An image classification system has been proposed as a key element of an image retrieval system. Such classification techniques along with new information theory based clustering have boosted the integration of clustering and classification components into typical image retrieval systems. Various supervised learning techniques, such as neural network, were also considered for high-level concept learning. Specifically, the authors used 11 concepts. Namely, they considered water, fur, cloud, ice, grass, rock, road, sand, tree, skin, and brick.

A large training set including low-level region descriptors is then used as input for neural network classifier. This aims to learn the association between high-level semantic (concept labels) and low-level descriptors. The main limitation of this approach is its high computational cost and the relatively large data required for training. Besides these learning techniques, decision trees methods such as ID3, C4.5 and CART are used to predict high-level categories particularly, the authors used CART algorithm to derive decision rules that associate image color features to keywords such as Marine, Sunset, and Nocturne. In [19], a two-class (relevant and irrelevant) categorization model is solved using a C4.5 decision tree. Despite their strength to noise and handling of missing data, decision trees exhibit a lack of modularity.

Multimodal Fusion and Retrieval:
Distinct approaches with novel ideas have been proposed for the purpose of image retrieval in the last decade. These approaches relied highly on image and text modalities. Multimedia and speech retrieval solutions have also been proposed. Only the text and image modalities are used for image retrieval purpose. In such approaches, aggregation is considered to be a typical hack that greatly works in the direction of enhancing CBIR precision. This inquiry fusion can also be counter-productive for that matter. In such cases, optimal mode is studied for aggregation of different modalities. Recently, some fusion techniques have been devised by the researchers for application in image retrieval and image annotation. The following section comprises an overview of a survey that is related to image retrieval applications. To fuse multiple classifier outputs, traditional future approach is assumed. For validation of the attained rules, some ground truth data must be available to the process. Yet another alternative of fusion depends on the retraining of distinct classifiers for optimization of fusion rules. The offline performance of fusion learning makes the computational application inexpensive, which enhanced the modality usage. Over fitting still remains an unsettled challenge. Despite all efforts to reduce the over fitting of data, this has paved way for a relatively new domain of research in order to recognize the pattern and process images. An efficient fusion pattern is required to combat the real world issues [20].

For combination of varied learners, global and local approaches are necessary. Local approach provides a degree of confidence to every learner on the basis of a training set. Global approach gives an average degree of confidence. By making use of optimal data based weights, a more precise classification performance can be attained. It is obligatory for local fusion technique to uniformly cluster the input data. Appointment of unlabelled instances to regions is an element of supervised learning. In the testing stage, outlining of the dynamic data classification is undertaken. an ordinary vicinity in the feature space local regions can be used to obtain classifier accuracy. Another local fusion approach called Context-Dependent Fusion (CDF) initially groups the training samples into uniform context clusters. the sequentially independent elements of CDF are local expert mode selection phases and clustering. A generic context reliable fusion approach was then propounded by the researchers. This proposal categorizes feature space and combines the outputs of individual expert models at the same time. To predict the aggregation weights for individual classifier models, a simple linear aggregation is employed. Even so, the weights sometimes fail to reflect integration between distinct learners [21].

The fusion decision regions are discovered by the unmonitored clustering of samples of training. this follows the selection of highest performance classification on every local region. A novel clustering approach had been propounded. This was done to part the samples of training into correct and incorrect classified samples. This is succeeded by appointment of the most precise classifier in the test. This is done to provide the fusion decision. Lately, an approach was introduced which fragmented the data instances into uniform groups and used their low level features. The inferred clusters were used for the aggregation of individual classifier decisions. The relative precision of these classifiers was reflected by the weights. To address the immunity of this proposal in response to noise and outliners, another probabilistic approach was employed by the researchers. This approach adapted the fusion technique of sub-regions of the space of feature. a probabilistic membership was produced by this approach algorithm that reflected the conventionality of the data instances for reduction of noise impact. Expert learners are then tied up with the resultant clusters. For all the classifiers, the cumulative weights are studies simultaneously. At last, individual confidence values are produced for aggregation weights that correspond to the nearest cluster. This approach still remains vulnerable to the local minima, despite of working efficiently for some other applications [21].

III. LOW-LEVEL FEATURES

To decipher the image content for the CBIR, low level features have been described. Their utility in enhancing the system precision is explained in the following section:

1.Color Features:

Commonest low level feature of an image CBIR is the color feature. RGB, LUV, HSV, HMMD, YCrCb, and LAB are believed to be the closest color spaces to human perception. Color histogram, color moments, color-covariance matrix, and color coherence vector etc are CBIR proposed color features. Ideal MPEG-7 color features include dominant color, scalable color, color structure and color layout. For expression of high level semantics, these still pose limitations. For countering this, an averaging color for all pixels in a region has been put forward as a color feature by the researchers. Image segmentation quality affects this feature. Perceptual color, as described by the authors, was the dominant color in HSV. The largest bin of histogram of colors (10 * 4 * 4 bins) takes the dominant color into account. It then corresponds to average HSV of all pixels. When applied to non-uniform color regions, definitive color feature is not produced by average color. To boost the quality of segmentation and reduce noise, CBIR has adopted image processing as its main component [22, 23].

2.Texture Features:

For effective reduction of gap between high level semantics and image content in a CBIR system, texture features play a significant role. They help the representation of real world image content. These include skin, nature, and fabric etcetera. CBIR has extensively adopted Gabor filtering and spectral feature extraction. Wold and Tamura texture features are propounded to denote visual content in an effective manner, eventually raising the accuracy of the CBIR. Some statistical measures have been lately adopted by MPEG-7. These include regularity, directionality and coarseness. But these are not immune to orientation and scale variation.

The best human vision was attained by the Gabor and wavelet based texture features. Even so, these are affected by shape of image region. The inconsistently shaped and rectangular regions are well dealt by these systems though. To combat this limit, transform application and padding for reshaping the arbitrary regions was done. This reduced the constancy of the extracted texture feature. Projection onto convex sets (POCS) is yet another approach for accurate texture feature extraction. The Edge Histogram Descriptor (EHD) helped efficient representation of natural images by encoding spatial distribution of image edges. It is however affected by object and scene distortions. Gradient vector was extracted from sub-band images obtained with the help of wavelet transform [24 , 25].

3.Shape Feature:

The attributes of shape include circularity, Fourier descriptor, boundary segment, moment invariant, aspect ratio, eccentricity and so on. They have been used extensively in a CBIR approach. Extraction of shape descriptors has been done with the use of area and second order moments. Three shape descriptors have been included by MPEG-7. This has been done for object based image retrieval. A descriptor on the basis of curvature scale space (CSS) which is strong at scaling, translation and rotational variation; a region based feature extracted with the help of Zernik moments, and a 3-D shape descriptor based 3-D meshes of shape surface have been expressed as MPEG-7 standard shape features. The distortions have, however shown some drawbacks with image representation. This limitation has been addressed by some authors in our references, propounding a stronger variant of CSS [26].

4.Spatial Location:

This is another low level feature of an image. The spatial locations act as a parameter of discrimination if the objects show an identical color and texture properties. For representation purposes, a minimum bounding box and spatial centroid are used. When compared to a relative spatial relationship, the intrinsic spatial location does not effectively reflect the semantic information. For this, a two-dimensional string is used and directional relationships are contemplated by its derivatives. This is highly enhanced by topographical relationship. A spatial context modeling algorithm was designed that relied on 6 pair wise spatial region relationships. Composite region template (CRT) was an assuring approach towards capturing semantic classes and spatially arranging regions [27].

IV. RESEARCH ISSUES

1. Query Formulation:

To bridge the semantic gap, the contemplation of a query is a must. OQUEL is an inquiry script or language that has been introduced to support keyword combinations. Fundamental semantic indicators are included in the vocabulary of the novel language. The semantic content is expressed by image regions. To decipher the semantic code, a multi-scale color coherent descriptor and wavelet transform texture features are employed. Even after all these efforts, the language needs more attention [28].

2. Image Benchmark and Performance Measures :

To thoroughly evaluate CBIR performance, we make use of the Corel image collection subsets. This gives us subjective inferences that are query-dependent. It has been proved time and again that diverse retrieval conclusions can be drawn with the same image cluster. Sans specificationand apt reporting, these results can not be turned to objective. Hence, for a precise assessment of the system, we still require better performance measures [29].

V. CONCLUSIONS

CBIR has been a major attention seeking evolution in terms of visual descriptor extraction, learning approaches and processing the image digitally. Although the visual descriptors have not been able to efficiently bridge the semantic gap. High level semantics have still not been dealt with, despite of all the propounded work in this area. The objective assessment and CBIR system comparison has not been contemplated till now. The need of the hour is the employment of some yet novel approaches that would capture high level semantics. Moreover, some ultra-efficient methodology for visual description is required.

REFERENCES

[1]. A. W. Smeulders, M. Warring, Santini, S., Gupta, A., , and R. Jain, Content-based image retrieval at the end of the early years. IEEE Trans. Pattern Anal. Mach. Intel. 22 (12) (2000), 1349-1380.

[2]. P. Aigrain, H. Zhang, and D. Petrovic, Content-based representation and retrieval of visual media: A review of the state-of-the-art. Multimed. Tools Appl. 3(3) (1996), 179-202.

[3]. Y. Rui, T. Huang, M. Ortega, and S. Mehrptra, Relevance feedback: A power tool in interactive content-based image retrieval. IEEE Trans. Circ. Syst. Video Technol. 8(5)(1998), 644-655.

[4]. Y. Rui, T. Huang, Optimizing learning in image retrieval. In Proceedings of the IEEE Computer Vision and Pattern Recognition (CVPR), 2000.

[5]. Y. Rui, T. Huang, and S. MEHROTRA, Content-based image retrieval with relevance feedback in Mars. In Proc of the IEEE International Conference on Image Processing, 1997.

[6]. Y. Rui, T. Huang, and S. F. Chang, Image retrieval: Current techniques, promising directions and open issues. J. Visual Commun. Image Represent. 10(1)(1999), 3962.

[7]. C. Sonek and M. Warring, Multimodal video indexing: A review of the state-of-the-art. Multimed. Tools Appl. 25(1) (2005), 535.

[8]. M. Rehman, M. Iqbal, M. Sharif and M. Raza. Content Based Image Re-trieval: Survey, World Applied Sciences Journal 19 (3) (2012): 404-412.

[9]. N. Singhai, S K. Shandilya, A Survey On: Content Based Image Retrieval Systems, International Journal of Computer Applications 4(2) (2010).

[10]. D. Zhou, J. Weston, A. Gretton, O. Bouquet and B. Scholkopf, Ranking on data manifolds. In Proc of the Conference on Advances in Neural Information Processing Systems (NIPS), 2003.

[11]. C. Bohm, S. Berthold, and D. A. Keim, Searching in high-dimensional space index structures for improving the performance of multimedia databases. ACM Computes. Surv. 33,(3) (2001).

[12]. H. Muller, N. Michoux, D. Bandon, and A. Geissbuhler, A review of content-based image retrieval systems in medical applications Clinical benefits and future directions. Int. J. Medical Inf. 73(1) (2003), 1-23.

[13]. H.Alraqibah, M. M. Ben Ismail and O. Bchir, "Empirical Comparison of Visual Descriptors for Content based X-ray Image Retrieval", International Conference on Image and Signal Processing (ICISP'14), Cherbourg, June 2014.

[14]. H. Alraqibah, O. Bchir and M. M. Ben Ismail, "X-Ray Image Retrieval System Based on Visual Feature Discrimination", Proceeding SPIE, vol. 9159, International Conference on Digital Image Processing (ICDIP), Athens, April 2014.

[15]. G. Carneiro and N. Vasconcelos, Minimum Bayes error features for visual recognition by sequential feature selection and extraction. In Proc of the Canadian Conference on Computer and Robot Vision, 2005.

[16]. Y. Chen and J. Z. Wang, A Region-Based Fuzzy Feature Matching Approach to Content-Based Image Retrieval, IEEE Trans. Pattern Analysis and Machine Intelligence, 24(9) (2002),2521267.

[17]. R. Shi, H. Feng, T.-S. Chua, C.-H. Lee, An adaptive image content representation and segmentation approach to automatic image annotation, International Conference on Image and Video Retrieval (CIVR), 2004, pp. 545-554.

[18]. X. Zhang, W.K. Cheung, C.H. Li, Learning latent variable models from distributed and abstracted data, Inf. Sci. Online (2013).

[19]. DO, M. N. AND VETTERLI,M., Wavelet-Based texture retrieval using generalized Gaussian density and Kullback-Leibler distance. IEEE Trans. Image Process. 11(2), (2002) 146-158.
[20]. Joshi, D., Naphade, M., And Natsev, A., A greedy performance driven algorithm for decision fusion learning. In Proc of the IEEE International Conference on Image Processing (ICIP), 2007.
[21]. M. M. Ben Ismail, and O. Bchir, , Automatic Image Annotation based on Semi-supervised Clustering and Membership based Cross Media Relevance Model. International Journal of Pattern Recognition and Artificial Intelligence 26(6)(2012).
[22]. YANG, C., DONG, M., and FOTOUHI, F. Region based image annotation through multiple- instance learning. In Proc of the ACM International Conference on Multimedia, 2005.
[23]. M. M. Ben Ismail, Image Annotation and Retrieval Based on Multi-Modal Feature Clustering and Similarity Propagation. PhD thesis packaged, produced and published by BiblioLabs under license by ProQuest UMI (available online for sale). 2011.
[24]. Shereena V.B. and Julie M. David, CONTENT BASED IMAGE RETRIEVAL: CLASSIFICATION USING NEURAL NETWORKS, The International Journal of Multimedia & Its Applications (IJMA) 6(5)(2014), 1-14.
[25]. De Oliveira, J. E., Araujo A. A., and Deserno T. M., Content-based image retrieval applied to BI-RADS tissue classification in screening mammography, World Journal of Radiology 3(1)(2011), 24-31.
[26]. BLEI, D. M. and JORDAN, M. I., Modeling annotated data. In Proc of the Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, 2003. [80] Ben Ismail M. M. and Frigui H., Unsupervised Clustering and Feature Weighting based on Generalized Dirichlet Mixture Modeling". Information Sciences, 274(2014), 35-54.
[27]. Ben Ismail M. M. and Bchir O., Content based Video Categorization using Relational Clustering with Local Scale Parameter, International Journal of Computer Science and Information Technology 8(1)(2016), 27-46.
[28]. Guo J. M.; Prasetyo H.; Chen J. H., Content-Based Image Retrieval Using Error Diusion Block Truncation Coding Features, IEEE Transactions on Circuits and Systems for Video Technology, 25 (3)(2015), 466-481.
[29]. Thepade S.p D.; Shinde Y. D., Robust CBIR using sectorisation of hybrid wavelet transforms with Cosine-Walsh, Cosine-Kekre, Cosine-Hartley combinations, in International Conf on Pervasive Computing (ICPC), 2015.
[30]. Pooja Kewat , Roopesh Sharma , Upendra Singh , Ravikant Itare, "Support Vector Machines Through Financial Time Series Forecasting ", Electronics, Communication and Aerospace Technology (ICECA), 2017 International conference of IEEE , 20-22 April 2017 ,pp. 1-7.
[31]. Yashika Mathur ,Pritesh Jain ,Upendra Singh, "Foremost Section Study And Kernel Support Vector Machine Through Brain Images Classifier ", Electronics, Communication and Aerospace Technology (ICECA), 2017 International conference of IEEE , 20-22 April 2017 ,pp.1-4.
[32]. Vineeta Prakaulya ,Roopesh Sharma ,Upendra Singh, "Railway Passenger Forecasting Using Time Series Decomposition Model ", Electronics, Communication and Aerospace Technology (ICECA), 2017 International conference of IEEE , 20-22 April 2017 ,pp.1-5.
[33]. Sonal Sable ,Ankita Porwal ,Upendra Singh , "Stock Price Prediction Using Genetic Algorithms And Evolution Strategies ", Electronics, Communication and Aerospace Technology (ICECA), 2017 International conference of IEEE , 20-22 April 2017 ,pp.1-5.
[34]. Rohit Verma ,Pkumar Choure ,Upendra Singh , "Neural Networks Through Stock Market Data Prediction" , Electronics, Communication and Aerospace Technology (ICECA), 2017 International conference of IEEE , 20-22 April 2017 ,pp.1-6.