IJSES

International Journal of Scientific Engineering and Science
Volume 1, Issue 11, pp. 82-89, 2017. ISSN (Online): 2456-7361

Digital Multimedia Information Retrieval Using Bootstrap Aggregative Learning Classifier

K. Vijayan[1], C. Chandrasekar[2]
[1]Research Scholar, Department of Computer Science, Manomaniam Sundaranar University
[2]Professor, Department of Computer Science, Periyar University

Abstract**—*****A digital multimedia information retrieval plays significant role in the field of image and video retrieval to retrieve the information relevant to a user query. In past decades, many research works have designed for digital multimedia information retrieval. However, classification performance of multimedia information retrieval was not efficient. In order to solve this limitation, Bootstrap Aggregative Learning Classifier (BALC) technique is proposed. The BALC technique is designed with improving the performance of digital multimedia information retrieval using machine learning classifier technique. The BALC technique initially takes video query as input and it applied Kernel-principal component analysis (Kernel-PCA) for extracting the visual features such as shape, color, texture in videos. After that, BALC technique used Bootstrap Aggregation with Support Vector Machine (BA-SVM) Classifier to classify the videos in a given dataset as relevant or irrelevant using video query with improved classification accuracy. At last, the BALC technique retrieves classified relevant videos' based on video query. This process assists BALC technique to improve precision and recall of video retrieval with minimum time. The BALC technique conducts experimental work on parameters such as classification accuracy, time complexity, precision and recall using three datasets with higher classification accuracy and minimum time complexity for multimedia information retrieval as compared to state-of-the-art works.*

Keywords**—** *Bootstrap Aggregation, Kernel-PCA, Video Query, Visual Features, Voting Scheme*.

I. Introduction

With the arrival of broadband networks, high-powered workstations, the digital multimedia information retrieval system attains higher significance. Hence, there is a necessity to index, store and mine the visual information from multimedia database because size of database is too large. With the rapid development in both centralized video archives and distributed WWW video resources, digital multimedia information retrieval gets greater attention. Many research works have been intended for multimedia information retrieval based on visual content. But, classification performance of conventional technique was not at required level. Therefore, there is a need for new digital multimedia information retrieval system to retrieve relevant video from a large dataset, based on visual contents when user input a query video.

A Content-based video retrieval was designed in [1] for retrieving human actions videos from video databases. But, video retrieval performance was poor. An interactive approach was developed in [2] to enhancing the performance of image and video mining. However, true positive rate was higher.

A novel approach was presented in [3] for content based lecture video indexing and retrieval in large lecture video archives with higher precision. But, time taken for retrieving videos was more. Nonparametric Video Retrieval and Frame Classification were designed in [4] to improve classification precision of videos. But, classification accuracy was not adequate.

A Matrix-Based Approach was intended in [5] for categorizing videos in an action video sequences with higher accuracy through performing clustering. However, clustering performance was not efficient for achieving higher accuracy. Multiple Covariance Discriminative Learning (MCDL) technique was developed in [6] for video based object representation and classification with application of multiple covariance matrices. But, precision and recall of video classification was remained unsolved.

A spatiotemporal co-location video pattern mining approach was introduced in [7] to efficient action retrieval in YouTube videos and thereby achieving higher precision and recall. However, the amount of time taken for video retrieval was higher. A hybrid approaches used for content-based video retrieval was analyzed in [8] in order to improve the performance of video retrieval. But, execution time of video mining was more.

In [9], Content-Based Video Recommendation System was designed. However, the system was not effective for achieving higher recall. A multiple-instance learning algorithm was employed in [10] for enhancing video classification. But, video retrieval based on user query was remained unaddressed.

To solve the above said existing issues, Bootstrap Aggregative Learning Classifier (BALC) technique is introduced. The contribution of BALC technique is follows,

- To improve the precision and recall of digital multimedia information retrieval with lesser time complexity, BALC technique is designed and developed with application of Bootstrap Aggregation with Support Vector Machine Classifier.
- To extract the visual features such as shape, color, texture in videos, Kernel-principal component analysis is applied in BALC technique.
- To improve the classification performance of digital multimedia information retrieval, Bootstrap Aggregation with Support Vector Machine Classifier (BA-SVM) is designed in BALC technique. BA-SVM classifier combines the results of all base SVM classifiers using voting scheme in order to achieve higher classification accuracy for digital multimedia information retrieval.

Volume 1, Issue 11, pp. 82-89, 2017.

International Journal of Scientific Engineering and Science
ISSN (Online): 2456-7361

The rest of this paper is organized as follows: In Section 2, Bootstrap Aggregative Learning Classifier (BALC) technique is explained with help of architecture diagram. In Section 3, Simulation settings are presented and the result discussion is explained in Section 4. Section 5 reviews the related works. Section 6 provides the conclusion of the paper.

II. Bootstrap Aggregative Learning Classifier Technique

The Bootstrap Aggregative Learning Classifier (BALC) technique is designed to improve the digital multimedia information retrieval with higher recall and digital multimedia information retrieval is similar to visual content based information retrieval. The BALC technique used Kernel-principal component analysis (Kernel-PCA) for extracting the visual features such shape, color, texture from frames of video. In BALC technique, video retrieval is based on user query video and it designed Bootstrap Aggregation with Support Vector Machine (BA-SVM) Classifier for mining required videos from a given video dataset based on query video.

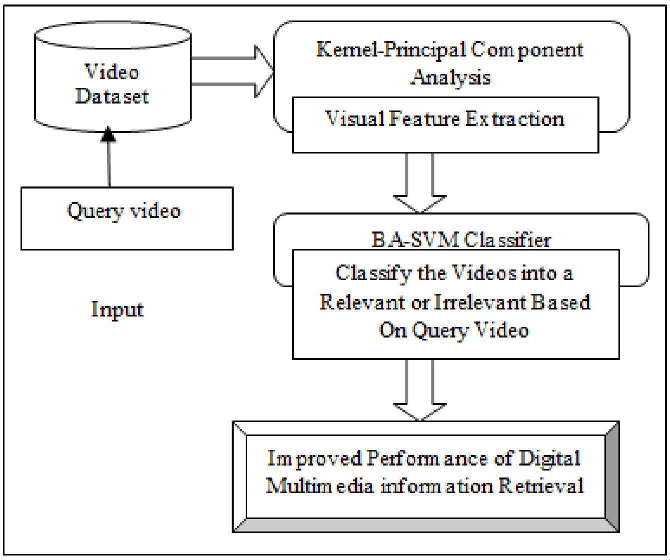

Fig. 1. Architecture Diagram of Bootstrap Aggregative Learning Classifier Technique for Video Retrieval

The BA-SVM classifier aggregates the results of all base SVM classifiers with aid of voting scheme for classifying videos as relevant or irrelevant. Thus, BALC technique efficiently retrieves more relevant videos related to query video. The overall architecture diagram of BALC technique is shown in below Figure 1. From Figure 1, BALC technique initially takes query video as input. Next, BALC technique performs visual feature extraction process with help of Kernel-PCA in order to extract visual features such as shape, color, texture from videos. After visual feature extraction, BALC technique applied BA-SVM Classifier for significantly retrieving relevant videos from a dataset based on query video. This aids for BALC technique to attain higher the precision and recall of video retrieval process and minimum time complexity. The detailed process about BALC technique is explained as follows.

2.1. Kernel Principal Component Analysis Based Visual Feature Extraction

The Kernel-Principal Component Analysis (Kernel-PCA) is used in BALC technique for visual feature extraction from a video in a given dataset. The video includes of collections of video frames. In order to extract visual features of videos, the video sequence is divided into a number of frames $V = \{Fr_1, Fr_2, \ldots Fr_N\}$. In Kernel-PCA, the input video frames with nonlinear structure are transformed into a higher dimensional feature space with linear structure and then linear PCA is performed in high-dimensional space. The Kernel-PCA is a way of observing an arbitrary mapping from the data space (i.e. video frames) into the feature space without compute mapping explicitly, with $i \in [1, N]$. Here; N indicates the number of frames in a video. Combining kernel concepts with PCA obtains a non-linear generalization of PCA called as Kernel-PCA. Kernel-PCA performs linear PCA in a high-dimensional kernel space rather than the original feature space. For a given video $V = \{Fr_i \in V^N \mid i = 1, \ldots, N\}$, Kernel-PCA maps input video V into a feature space FS through a nonlinear mapping φ associated with a given kernel function k which is formulated as,

$$\varphi : V^N \rightarrow FS \quad (1)$$

In transformed feature space, the covariance matrix C is formulated as,

$$C = \frac{1}{N}\sum_{i=1}^{N}\varphi(Fr_i)\varphi(Fr_i)^T \quad (2)$$

From equation (2), φ denotes the non-linear map (i.e. feature space) and T is a transformation matrix whereas N is a total number of frames Fr_i in an input video. Then, linear PCA in FS finds the eigenvectors and the corresponding eigenvalues that satisfy the following expression,

$$\lambda\varepsilon = C\varepsilon \quad (3)$$

From equation (3), ε denotes the eigenvectors and λ is an eigenvalues of video frame in feature space $\varphi(Fr_i)$. The eigenvectors is given by a linear combination of $\varphi(Fr_i)$ which is mathematically represented as,

$$\varepsilon = \sum_{i=1}^{N}\alpha\varphi(Fr_i) \quad (4)$$

By applying equation (2) and (4) into (3) and then multiplying both sides by $\varphi(Fr_i)^T$, the following equation is obtained as,

$$N\lambda\alpha = K\alpha \quad (5)$$

From equation (5), $K \in V^{N\times N}$ is a Gram matrix whereas α denotes the normalized eigen vectors of K. The Gram matrix or the kernel function is represented in the inner product form in order to find the visual features among video frames (Fr_i, Fr_j) as follow

International Journal of Scientific Engineering and Science
Volume 1, Issue 11, pp. 82-89, 2017. ISSN (Online): 2456-7361

$$K_{ij} = k\left(Fr_i, Fr_j\right) = \left(\varphi(Fr_i\right).\left(\varphi(Fr_j\right)) \quad (6)$$

Finally, Kernel-PCA find eigenvectors that related to largest eigenvalues to calculate the projection of visual feature vector onto the principal components $\varepsilon_p \subset \varepsilon$ with aid of the kernel function which is represented as,

$$Y = (\varepsilon_p^k \varphi(Fr_i)) = \sum_{i=1}^{N} \alpha_p^k k(Fr_i, Fr_j) \quad (7)$$

From equation (7), $Y \in V^N$ and $\alpha_p \subset \alpha$ with *n* being the number of principal components extracted (i.e. visual features of video such as shape, color, texture). Thus, the feature space *FS* contains the explicit structure information of frames in a given video that helps for BALC technique to extract visual features such as shape, color, texture in an effective manner.

// Kernel-Principal Component Analysis based Visual Feature Extraction Algorithm
Input: Video dataset
Output: Extracts Visual Features from videos
Step 1: Begin
Step 2: **For** each video in a dataset
Step 3: **For** each frames in a video
Step 4: Compute covariance matrix using (2)
Step 5: Compute eigenvectors and eigenvalues of that matrix using (3) and (4) and Find eigenvectors corresponding to largest eigenvalues
Step 6: Extracts visual features of video using (7)
Step 7: **End For**
Step 8: **End for**
Step 9: End

Algorithm 1 Kernel Principal Component Analysis Based Visual Feature Extraction

Algorithm 1 depicts the process of Kernel-PCA for extracting the visual features from a video. As shown in Algorithm 1, Kernel-PCA computes the covariance matrix for a collection of frames in an input video sequence. Subsequently, Kernel-PCA evaluates eigenvectors and eigenvalues of that covariance matrix. Subsequently, Kernel-PCA discovers eigenvectors related to largest eigenvalues and finally projects visual features vectors onto the principal components. This supports for Kernel-PCA to efficiently extract visual features such as shape, color, texture in a frames of video.

2.2 Bootstrap Aggregation with Support Vector Machine Classifier

After extracting the visual features, Bootstrap Aggregation with Support Vector Machine (BA-SVM) Classifier is used in BALC technique for efficient visual content based multimedia information retrieval. The BA-SVM Classifier is a machine learning ensemble meta-algorithm to improve the stability and accuracy of visual content based information retrieval. The BA-SVM Classifier categorizes the videos in a given video database as relevant or irrelevant based on the visual features of query video. The Support Vector Machine (SVM) classifier performs video classification based on query video through determining the optimal separating hyperplane.

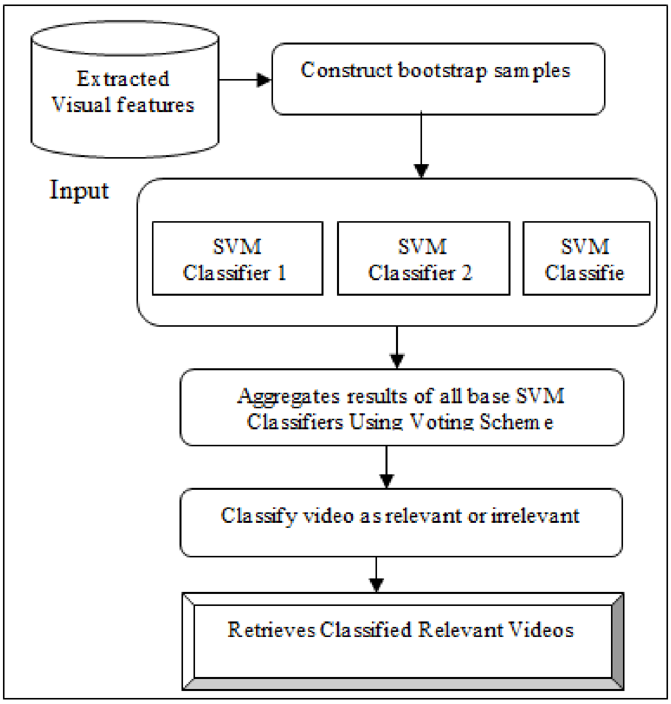

Fig. 2. Process of Bootstrap Aggregation with Support Vector Machine (BA-SVM) Classifier for Digital Multimedia Information Retrieval

But, classification performance and time complexity of SVM classifier is degraded when size of video dataset is large. The above said drawback is overcome through in BALC technique. Figure 2 shows process of BA-SVM Classifier for digital multimedia information retrieval. As shown in figure, BA-SVM Classifier initially constructs k bootstrap samples from a given video dataset.After that, each bootstrap sample is trained by SVM classifier. Consequently, the results of all base SVM classifiers are aggregated in order to efficiently categorize the video as relevant or irrelevant based on visual features of query video with minimum time. This helps for BALC technique to achieve higher classification accuracy and lower time complexity for digital multimedia information retrieval. At last, BA-SVM Classifier efficiently retrieves the classified relevant videos. This process resulting in improved precision and recall of digital multimedia information retrieval in a significant manner. The SVM classifier includes set of training samples $\{(X, Y_1), (X_2, Y_2), \ldots (X_n, Y_n)\}$ where X_i indicates the numerous videos in a given dataset and Y_i denotes the classification output $Y_i \in \{+1, -1\}$.

The classifier result $Y_i = 1$ is said to be a relevant video and $Y_i = -1$ is said to be an irrelevant video. Initially, different training bootstrap samples sets are generated from an input video dataset for attaining higher classification results for efficient visual content based information retrieval.

International Journal of Scientific Engineering and Science
Volume 1, Issue 11, pp. 82-89, 2017. ISSN (Online): 2456-7361

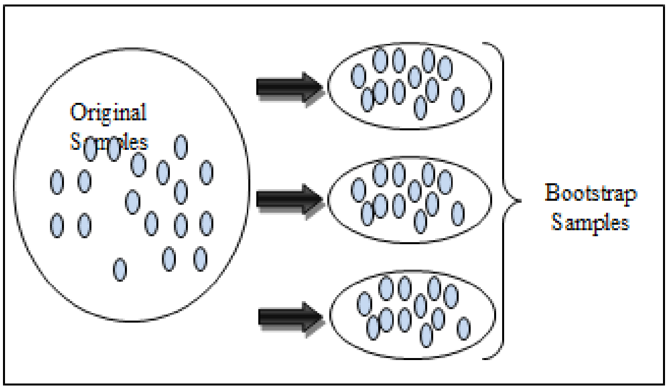

Fig. 3. Constructions of Training Bootstrap Samples for Video classification

Figure 3 demonstrates the creation of training bootstrap samples for effective video classifications. A bootstrap sample is a smaller sample which is "bootstrapped" from a larger video dataset. Bootstrapping is a kind of resampling in which higher numbers of smaller samples of the similar size are frequently drawn with replacement from a single original sample. Generally, original samples are much larger size than bootstrap samples. After building a bootstrap samples, the base SVM classifier is applied in order to separate the videos in a given dataset as relevant or irrelevant with respect to the extracted visual features of query video.

The base SVM classifier is a discriminative classifier for efficiently mining multimedia information's. The SVM classifier segregates the positive (i.e. relevant videos) and negative (i.e. relevant videos) labeled bootstrap samples using optimal separating hyper plane. The optimal hyperplane of SVM classifier for video classification is formulated as,

$$\vec{W}.\vec{x}-\vec{b}=0 \tag{8}$$

From equation (8), $\vec{W}$ denotes the weight vector and $\vec{x}$ indicates collection of videos in a given dataset whereas $\vec{b}$ is a bias. Followed by, the two parallel hyperplanes are chosen to separate the two classes of samples. The region covered by two parallel hyperplanes is known as margin. Any sample (i.e. video) that lies above the separating hyperplane fulfils the following mathematical formula,

$$\vec{W}.x+b>0 \tag{9}$$

Correspondingly, any sample that lies under the separating hyperplane fulfils the following mathematical expression,

$$\vec{W}.x+b<0 \tag{10}$$

Then, distance between two marginal hyperplane is measured as,

$$Dist(h_1,h_2)=\frac{2}{\|\vec{w}\|} \tag{11}$$

From equation (11), $Dist(h_1,h_2)$ indicates a distance between two marginal hyperplane. The videos are classified depends on a similarity function termed as kernel. The similarity between the visual features of video is determined through the kernel. A base SVM classifier evaluates weighted sum of similarities between the visual features of videos in a given dataset and visual features of query video to classify the videos as relevant or irrelevant using below formulation,

$$Y=sign\sum_{i=1}^{n}\vec{W}\beta_i k(x_i,Q) \tag{12}$$

From equation (12), $Y\in\{+1,-1\}$ represents the output of base SVM classifier whereas $k(x_i,Q)$ indicates a kernel function that measures similarity between the visual features of videos Q in a given dataset and visual features of query video Q. Here, "*sign*" represents whether the predicted classification results is positive (i.e. relevant video) or negative (i.e. irrelevant video) and β_i is a Lagrangian multiplier. The classification result of base SVM classifier to classify the positive samples (i.e. relevant video) and negative samples (i.e. irrelevant video) using optimal hyperplane.

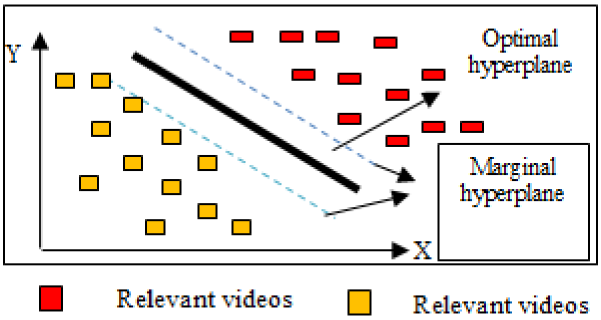

Fig. 4. Classification results of SVM classifier

The classification of videos based on visual features of query video. Based on the extracted visual features, the videos in a dataset are separated from both upper and lower side of the margin through classification. Afterward the output of the all the base SVM classifier is aggregated into a one strong classifier using voting scheme. After training the bootstrap samples with SVM classifier, the majority votes of samples is employed in order to make the strong classifier for efficiently classifying videos as relevant or irrelevant. Thus, the results of the ensemble of Bootstrap aggregation with SVM classifier is determined as,

$$f(x)=\arg\max_{n} N_n \tag{13}$$

From equation (13), $f(x)$ denotes the final decision of the SVM ensemble with majority votes whereas N_n indicates number of video samples in SVMs classifier whose decisions are known to the n^{th} class. With the assists of voting scheme, the BA-SVM Classifier classifies the videos which gets majority vote as relevant videos. Otherwise video is classified as irrelevant video. Therefore, BA-SVM Classifier improves the classification performance of digital multimedia information retrieval with minimal time. The algorithmic process of BA-SVM Classifier is shown in below algorithm 2 for efficiently extracting digital multimedia information.

// Bootstrap Aggregating with SVM Classifier Algorithm
Input : $\{(X_1,Y_1),(X_2,Y_2),\ldots(X_n,Y_n)\}$ is a set of training video

samples
Output: Improved classification accuracy and minimum time complexity for digital multimedia information retrieval
Step 1:**Begin**
Step 2: **For** each video samples
Step 3:Construct bootstrap samples with resampling
Step 4: Train base SVM classifier on bootstrap video samples
Step 5: Bootstrap video samples are classified using (12)
Step 6: The results of all base SVM classifiers are aggregated with aid of voting scheme using(13)
Step 7: The majority vote of positive video samples of base SVM classifier is classified as relevant videos
Step 8: The majority vote of negative video samples of base SVM classifier is classified as irrevalant videos
Step 9: **End for**
Step 10: Mines Classified relevant videos using query video
Step 11:End

Algorithm 2 Bootstrap aggregating with SVM classifier for Digital Multimedia Information Retrieval

As shown in algorithm, each training videos, bootstrap samples are generated with re-sampling approach. After that, these bootstrap samples are trained with base SVM classifiers. Next, the output of all SVM classifiers is combined using Aggregating approach and voting scheme. Thus, the majority vote of positive samples of base SVM classifier is classified as relevant videos. In the same way, the majority vote of negative samples of base SVM classifier is classified as irrelevant videos. As a result, BALC technique attains higher classification accuracy with minimum time. Finally, BA-SVM Classifier significantly mines the classified relevant videos based on query video. This process assists for BALC technique to increase the precision and recall of digital multimedia information retrieval in an effectual manner.

III. Experimental Setting

Experimental evaluation of Bootstrap Aggregative Learning Classifier (BALC) technique is implemented in Java Language using three dataset. The BALC technique conducts the experimental works using three data sets namely VIRAT Video Dataset [22] and UCF Sports Action Data Set [23] and INRIA Holidays dataset [24]. The VIRAT Video Dataset used in BALC technique contains many videos recorded from 11 scenes (1080p or 720p). The UCF Sports Action Data Set utilized in BALC technique has collection of actions videos obtained from lots of sports. A number of video present in a UCF Sports Action Data Set is 150 with the resolution of 720 x 480. Moreover, INRIA Holidays dataset includes 500 image collections. The each image group indicates a dissimilar scene or object.

The BALC technique takes 100 videos and 100 images from above three datasets to perform experimental evaluations. The experimental evaluations of BALC technique is conducted for numerous instances with diverse number of videos and images and averagely tens results are depicted in graph and tables for performance analysis. The efficiency of BALC technique is compared with two existing [1] [2]. The effectiveness of BALC technique is evaluated in terms of classification accuracy, time complexity, precision and recall.

IV. Results and Discussions

In this section, the result of BALC technique is discussed. The performance of BALC technique is compared with existing [1] [2] respectively. The performance of BALC technique is analyzed along with the with the help of tables and graphs.

4.1 Impact of Classification Accuracy

In BALC technique, Classification Accuracy (CA) determines the ratio of number of videos that are correctly classified to the total number of videos. It evaluated in terms of percentages (%) and expressed as follows,

$$CA = \frac{number\ of\ videos\ are\ correctly\ classifed}{N} * 100 \qquad (14)$$

From equation (14), the classification accuracy of digital multimedia information retrieval is calculated based on diverse number of videos. Here, N represents the total number of videos. When the classification accuracy of digital multimedia information retrieval is higher, the method is said to be more effectual.

TABLE 1 a) Tabulation of Classification Accuracy for Video Retrieval

Number of videos (N)	Classification Accuracy (%)					
	VIRAT Video Dataset			UCF Sports Action Data Set		
	Content based Video Retrieval	Interactive Approach	BALC technique	Content based Video Retrieval	Interactive Approach	BALC technique
10	65	72	84	69	75	87
20	67	73	85	71	76	88
30	70	75	87	73	78	89
40	72	76	88	74	79	90
50	73	77	89	76	80	91
60	74	79	90	78	82	92
70	76	80	92	79	83	93
80	77	81	93	81	84	94
90	79	83	94	83	86	95
100	80	84	96	85	87	97

TABLE 1 b) Tabulation of Classification Accuracy for Image Retrieval

Number of Images (N)	Classification Accuracy (%)		
	Content based Video Retrieval	Interactive Approach	BALC technique
10	66	72	88
20	68	74	89
30	71	75	90
40	73	76	91
50	75	78	92
60	76	80	94
70	77	81	95
80	79	82	96
90	80	83	97
100	81	87	98

Table 1 a) and b) depicts the tabulation results of classification accuracy for both video and image retrieval using three methods. The BALC technique considers the number of videos and images in the range of 10-100. The proposed BALC technique is presents higher classification accuracy for both video and image retrieval as compared to existing [1] [2]. This is owing to application of BA-SVM Classifier in BALC technique where output of all SVM classifiers is combined using voting scheme to classify the videos in dataset into a relevant or irrelevant. This supports for BALC technique to improve classification accuracy for both

International Journal of Scientific Engineering and Science
Volume 1, Issue 11, pp. 82-89, 2017. ISSN (Online): 2456-7361

video and image retrieval. As a result, proposed BALC technique increase the classification accuracy of video retrieval by 23 % and 15% while using VIRAT video dataset as compared to existing [1], [2] respectively. Similarly for UCF sports action data set, the classification accuracy of video retrieval is improved by 19 % and 13% when compared to existing [1], [2] respectively. Further, proposed BALC technique enhances classification accuracy of image retrieval by 21 % and 13% as compared to existing [1], [2] respectively while using INRIA Holidays dataset.

4.2 Impact of Time Complexity

In BALC technique, Time Complexity (TC) computes the total amount of time required for mining relevant videos from a given a dataset. It evaluated in terms of milliseconds (ms) and formulated as,

$$TC = N * Time(Retrieving\ one\ video) \quad (15)$$

From equation (15), the time complexity of digital multimedia information retrieval is evaluated in which N indicates the number of videos taken for experimental evaluations. Lower time complexity of digital multimedia information retrieval, more efficient the method is said to be. Table 2 a) and b) illustrates results of time complexity for both video and image retrieval using three methods and number of videos and images in the range of 10-100. The proposed BALC technique is presents minimum time complexity for the both video and image retrieval when compared to existing [1] [2] respectively.

TABLE 2 a) Tabulation of Time Complexity for Video Retrieval

Number of videos (N)	Time Complexity (ms)					
	VIRAT Video Dataset			UCF Sports Action Data Set		
	Content based Video Retrieval	Interactive Approach	BALC technique	Content based Video Retrieval	Interactive Approach	BALC technique
10	32.7	26.1	10.2	28.8	17.2	8.5
20	35.2	29.7	12.8	31.2	20.4	11.3
30	37.1	31.5	15.5	33.4	22.9	14.5
40	40.5	33.8	18.7	36.7	25.3	17.8
50	43.3	36.3	20.3	39.6	28.7	20.6
60	46.8	39.4	23.8	42.8	31.5	22.9
70	48.6	41.7	27.5	45.4	33.2	25.1
80	51.4	44.3	29.4	48.3	35.7	27.4
90	53.9	47.1	32.9	51.2	38.6	29.8
100	56.6	50.8	35.3	53.1	41.4	31.4

TABLE 2 b) Tabulation of Time Complexity for Image Retrieval

Number of Images (N)	Time Complexity (ms)		
	Content based Video Retrieval	Interactive Approach	BALC technique
10	27.2	19.5	6.9
20	29.8	21.3	9.2
30	32.6	23.8	11.8
40	35.7	25.6	15.1
50	38.5	27.5	17.0
60	41.3	30.2	19.8
70	44.4	33.7	21.4
80	46.8	35.6	22.9
90	48.1	38.5	24.5
100	51.5	41.4	26.7

This is due to application of BA-SVM Classifier in BALC technique in which results of all base SVM classifiers is combined together using voting scheme in order to categorize the videos as relevant or irrelevant based on query video with minimum time. This helps for BALC technique to reduce time complexity for digital multimedia information retrieval. Therefore, proposed BALC technique lessens the time complexity of video retrieval by 51 % and 43% while using VIRAT video dataset as compared to existing [1], [2] respectively. Likewise for UCF sports action data set, the time complexity of video retrieval using proposed BALC technique is minimized by 50 % and 31% as compared to existing [1], [2] respectively. Moreover, proposed BALC technique decreases time complexity of image retrieval by 57 % and 43% as compared to existing [1], [2] respectively while using INRIA Holidays dataset.

4.3 Impact of Precision

In BALC technique, Precision (P) computes the ratio of number of relevant videos retrieved based on query video to total number of videos N. It is evaluated in terms of percentages (%) and represented as,

$$P = \frac{Number\ of\ relevant\ videos\ retrieved}{N} * 100 \quad (16)$$

From equation (16), the precision rate of digital multimedia information retrieval is measured with different number of videos. Higher precision of digital multimedia information retrieval, more efficient the method is said to be.

Figure 5 a) and 5b) illustrates the precision for video and image retrieval using three methods. From figure, it is descriptive that the precision using proposed BALC technique is higher for both video and image retrieval as compared to existing [1] [2].

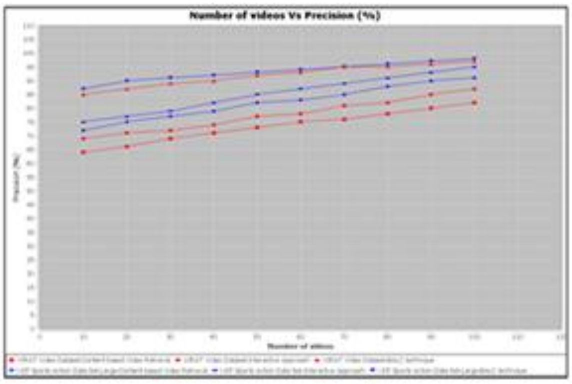

Fig. 5 a) Measurement of Precision for Video Retrieval

This is because of the BA-SVM Classifier developed in BALC technique for improving the performance of digital multimedia information retrieval. In BA-SVM Classifier, classification of videos is based on votes of base SVM classifier. Thus, the majority vote of positive samples of base SVM classifier is categorized as relevant videos whereas majority vote of negative samples of base SVM classifier is classified as irrelevant videos. This assists for BALC technique to extract the relevant videos related to query video efficiently.

Volume 1, Issue 11, pp. 82-89, 2017. ISSN (Online): 2456-7361

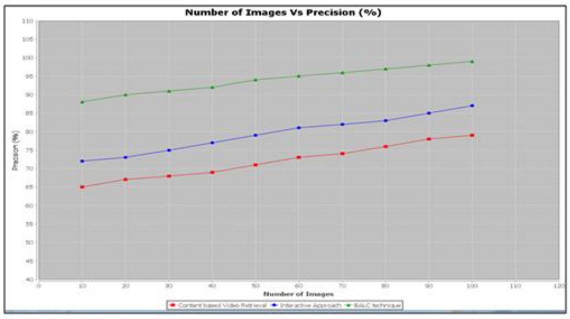

Fig. 5 b) Measurement of Precision for Image Retrieval

This process helps for BALC technique to increase the precision for digital multimedia information retrieval. For that reason, proposed BALC technique attains precision of video retrieval by 26 % and 19% while using VIRAT video dataset as compared to existing [1], [2] respectively. In the same way for UCF sports action data set, the precision of video retrieval using proposed BALC technique is enhanced by 14 % and 10% as compared to existing [1], [2] respectively. In addition, proposed BALC technique achieves 31% and 19 % precision for image retrieval as compared to existing [1], [2] using INRIA Holidays dataset.

4.4 Impact of Recall

In BALC technique, recall determines ratio of number of relevant videos that are correctly retrieved according to query video to total number of videos N. The recall is evaluated in terms of percentages (%) and formulated as,

$$Recall = \frac{Number\ of\ correctly\ retrieved\ similar\ videos}{N} * 100 \qquad (17)$$

From equation (17), the recall of digital multimedia information retrieval is evaluated with respect to varied number of videos. Higher recall of digital multimedia information retrieval, more efficient the method is said to be. From the figure, it is clear that the recall using proposed BALC technique is higher for both video and image retrieval as compared to existing [1] [2].This is due to application of BA-SVM Classifier in BALC technique. The BA-SVM applied in BALC technique for attaining higher classification performance for digital multimedia information retrieval. Therefore BA-SVM Classifier supports for BALC technique to retrieve the more relevant videos interrelated to query video efficiently. This process results in improved the recall for digital multimedia information retrieval.

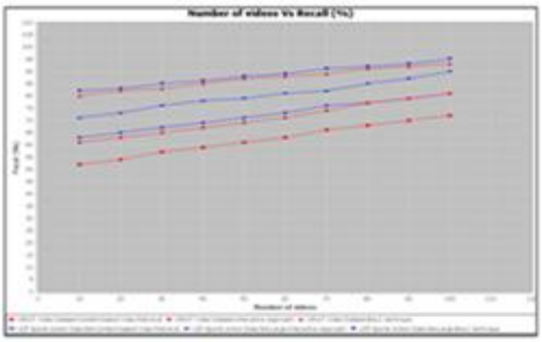

Fig. 6 a) Measurement of Recall for Video Retrieval

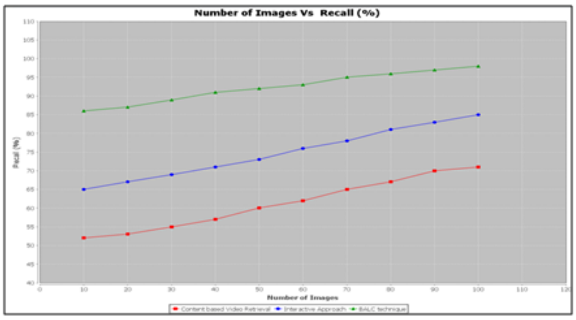

Fig. 6 b) Measurement of Recall for Image Retrieval

Hence, proposed BALC technique obtains 41 % and 24% recall for video retrieval while using VIRAT video dataset as compared to existing [1], [2]. Similarly for UCF sports action data set, the recall of video retrieval using proposed BALC technique is increased by 23 % and 10% as compared to existing [1], [2]. As well, proposed BALC technique acquires 52% and 24 % recall for image retrieval as compared to existing [1], [2] while using INRIA Holidays dataset.

V. RELATED WORKS

In [11], A review of different researches of content based indexing and extraction of visual information was analyzed. Web video categorization was presented in [12] with application of category-predictive classifiers and category-specific concept classifiers. However, the video classification performance was poor.

Multi-layer multi-view topic model was presented in [13]. But, classification performance was not effective. A survey of different techniques designed for feature selection in multimedia was presented in [14] to improve performance of multimedia data mining.

Content based video retrieval using enhanced feature extraction was intended in [15] to retrieve similar videos based on local feature detector. However, the precision of content based video retrieval was poor. Support Vector Machine (SVM) classifier was developed in [16] for content based image and video retrieval. But performance of retrieval system was poor

Video retrieval for shot cluster and classification was presented in [17] based on key feature set with minimum time. However, time complexity was higher. Shrinkage Optimized Direction information Assessment (SODA) was designed in [18] for multimodal video indexing and retrieval with aid of audio and video features. But, processing time taken for video retrieval was higher. Content Based Video Retrieval with Frequency Domain Analysis Using 2-D Correlation Algorithm was intended in [19] to retrieve all the videos objects that matched with user's query image. However, efficiency of video retrieval was lower. A video retrieval system was introduced in [20] to mine required videos through the inputs of its trajectory and/or appearance with higher precision. But, time complexity of video retrieval remained addressed. In [21], A Multi feature content based Video Retrieval was presented. But, time and space complexity of video retrieval was high.

International Journal of Scientific Engineering and Science
Volume 1, Issue 11, pp. 82-89, 2017. ISSN (Online): 2456-7361

VI. CONCLUSION

An efficient BALC technique is developed to achieve higher recall and minimum time complexity for digital multimedia information retrieval by using bootstrap aggregation classifier. Initially, BALC technique takes video query as input and it uses Kernel-PCA for extracting the visual features such as shape, color, texture in videos. Then BALC technique applied BA-SVM Classifier to classify the videos in a given dataset as relevant or irrelevant using video query and it resulting in increased classification accuracy. Finally, the BALC technique mines classified relevant videos' based on video query. Therefore, BALC technique increases precision and recall of video retrieval and also minimizes time complexity. With the experimental performed for BALC technique, it is observed that the recall presents more precise results for digital multimedia information retrieval with an improvement of classification and reduction of time complexity as compared to state-of-the-art works.

REFERENCES

[1] Simon Jones and Ling Shao, "Content-based retrieval of human actions from realistic video databases", *Information Sciences*, vol. 236, pp. 56–65, 2013.
[2] Andrew Gilbert and Richard Bowden, "Image and video mining through online learning", *Computer Vision and Image Understanding*, Elsevier, vol. 158, pp. 72-84, May 2017.
[3] Haojin Yang and Christoph Meinel, "Content based lecture video retrieval using speech and video text information", *IEEE Transactions on Learning Technologies*, vol. 7, issue 2, pp. 142–154, April-June 2014.
[4] Alexandre Karpenko and Parham Aarabi, "Tiny videos: A large data set for nonparametric video retrieval and frame classification", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, issue 3, pp. 618–630, March 2011.
[5] Peng Cui, Fei Wang, Li-Feng Sun, Jian-Wei Zhang, and Shi-Qiang Yang, "A matrix-based approach to unsupervised human action categorization", *IEEE Transactions on Multimedia*, vol. 14, issue 1, pp. 102–110, 2012.
[6] Yukon Zhang and Quan Liu, "Video based object representation and classification using multiple covariance matrices", *PLoS ONE*, vol. 12, issue 6, pp. 1-10, 2017.
[7] Liujuan Cao, Rongrong Ji, Yue Gao, Wei Liu, and Qi Tian, "Mining spatiotemporal video patterns towards robust action retrieval", *Neurocomputing*, Elsevier, vol. 105, pp. 61-69, April 2013.
[8] Hrishikesh Bhaumik, Siddhartha Bhattacharyya, Mausumi Das Nath, and Susanta Chakraborty, "Hybrid soft computing approaches to content based video retrieval: A brief review", *Applied Soft Computing*, vol. 46, Elsevier, pp. 1008-1029, September 2016.
[9] Yashar Deldjoo, Mehdi Elahi, Paolo Cremonesi, Franca Garzotto, Pietro Piazzolla, and Massimo Quadrana, "Content-Based video recommendation system based on stylistic visual features", *Journal on Data Semantics*, Springer, vol. 5, issue 2, pp. 99–113, June 2016.
[10] Jingjing Liu, Chao Chen, Yan Zhu, Wei Liu, and Dimitris N. Metaxa, "Video Classification via Weakly Supervised Sequence Modeling", *Computer Vision and Image Understanding*, Elsevier, vol. 152, pp. 79-87, November 2016.
[11] R. Priya and T. N. Shanmugam, "A comprehensive review of significant researches on content based indexing and retrieval of visual information", *Frontiers of Computer Science*, Springer, vol. 7, issue 5, pp. 782–799, October 2013.
[12] Mehtab Afzal, Xiao Wu, Honghan Chen, Yu-Gang Jiang, and Qiang Peng, "Web video categorization using category- Predictive classifiers and category-specific concept classifiers", *Neurocomputing*, Elsevier, vol. 214, pp. 175-190, November 2016.
[13] Sujuan Hou, Ling Chen, Dacheng Tao, Shangbo Zhou, Wenjie Liu, and Yuanjie Zheng, "Multi-Layer multi-view topic model for classifying advertising video", *Pattern Recognition*, Elsevier, vol. 68, pp. 66-81, August 2017.
[14] Pui Yi Lee, Wei Ping Loh, and Jeng Feng Chin, "Feature selection in multimedia: The state-of-the-art review", *Image and Vision Computing*, Elsevier, vol. 67, pp. 29-42, November 2017.
[15] Dipika H. Patel, "Content based Video Retrieval using Enhance Feature Extraction", *International Journal of Computer Applications*, vol. 119, issue 19, pp. 4-8, June 2015.
[16] Milan R. Shetake and Sanjay. B. Waikar, "Content based image & video retrieval", *International Journal of Advances in Electronics and Computer Science*, vol. 2, issue 9, pp. 84-87, 2015.
[17] Chuen-Horng Lin and Li-Jung Fu, "Video retrieval for shot cluster and classification based on key feature set", *The Imaging Science Journal*, pp. 1-20, 2017.
[18] Xu Chen, Alfred O. Hero, and Silvio Savarese, "Multimodal video indexing and retrieval using directed information", *IEEE Transactions on Multimedia*, vol. 14, issue 1, pp. 3–16, 2012.
[19] Navdeep Kaur and Mandeep Singh, "Content based video retrieval with frequency domain analysis using 2-D correlation algorithm", vol. 4, issue 9, pp. 388-393, September 2014.
[20] Yuan-Hao Lai and Chuan-Kai Yang, "Video object retrieval by trajectory and appearance", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, issue 6, pp. 1027-1037, June 2015.
[21] Hatim G. Zaini and T. Frag, "Multi feature content based Video Retrieval Using High Level Semantic Concepts", *International Journal of Computer Science*, vol. 2, issue 9, pp. 15-27, September 2014.
[22] VIRAT Video Dataset: http://www.viratdata.org/
[23] UCF Sports Action Data Set: http://crcv.ucf.edu/data/UCF_Sports_Action.php
[24] INRIA Holidays dataset: http://lear.inrialpes.fr/~jegou/data.php