Using SVM and KNN to Evaluate Performance Based on Video Plagiarism Detectors and Descriptors for Global Features

Ekta Thirani[1], Jayshree Jain[2], Vaibhav Narawade[3]

[1]Research Scholar, Faculty of Computer Engineering, Pacific Academy of Higher Education and Research University, Udaipur, India
[2]Department of Computer Engineering, Pacific Academy of Higher Education and Research University, Udaipur, India
[3]Department of Computer Engineering, Ramrao Adik Institute of Technology, Nerul, India

E-mail: [1]ektasarda16@gamil.com, [2]dejayshreejain@gmail.com, [3]vnarawade@gmail.com

Abstract

The detection of video piracy has improved and emerged as a popular issue in the field of digital video copyright protection because a sequence of videos often comprises a huge amount of data. The major difficulty in achieving efficient and simple video copy detection is to identify compressed and exclusionary video characteristics. To do this, we describe a video copy detection strategy that created the properties for a spatial-temporal domain. The first step is to separate each video sequence into the individual video frame, and then extract the boundaries of each video frame by using PCA SIFT and Hessian- Laplace. Next, for each video frame, we have to implement SVM and KNN features in the spatial and temporal domains to measure their performance matrices in the feature extraction. Finally, the global features found in the Video copy detection are accomplished uniquely and efficiently. Experiments arranged a commonly used VCDB 2014 video dataset, showing that result. The proposed approach is based on various copy detection algorithms and shows various features in terms of both accuracy and efficiency.

Keywords: Feature extraction, video copy detection, copyright protection, video security, SVM, KNN, and video plagiarism detection

1. Introduction

The amount of digital multimedia files on the Internet has exploded due to the rapid development of information and communication technology and the widespread usage of the

DOI: https://doi.org/10.36548/jscp.2022.2.004
Received: 25.05.2022, received in revised form: 30.06.2022, accepted: 15.07.2022, published: 25.07.2022

Internet, creating several new security challenges for multimedia data. Signal processing, transmission, or purposely created video distortions all produce a wide range of visual distortions with varying strengths. The robustness of a video detection method is determined by the detection performance against various distortions.

Video detection methods must overcome two major challenges to be effective and efficient: For example, it is difficult to develop a test collection that is representative of real-world video copy detection, and it is impossible to deal with all kinds of video modifications in an effective low-complexity fingerprinting approach, the tradeoff between robustness and efficiency must therefore be balanced by clearly delineating the application requirements. There are three types of transformation where signal processing and decreasing in quality such as Users may easily copy and edit video with a range of adjustments such as compression, re-encoding, resolution reduction – blurring (Gaussian, motion), noise addition (Gaussian, white, random), color changes – changes in contrast/brightness/gamma/luminance, filtering (Gaussian, median, average). Geometric transformations, post-production or edition, and camcording: resizing, scaling, rotation, cropping, zooming, letterbox/pillar box, moving caption (subtitles), insertion of patterns, picture in picture (PiP), flip (horizontal/vertical mirroring), projective transforms, bending, camcording. Desynchronization- time shift, spatial shift, fast/slow motion, change in frame rate, frame loss, frame addition interlaced/progressive conversion, and recompression recognitions to the growing acceptance of numerous video processing software. As a result, there are several video copies available on the Internet. There are several challenges in protecting video copyright in the digital age, and one of the most crucial is being able to recognize illegally copied videos. Many of the features we've included in this paper are to find the global feature-based detection in video.

Local features are based on key points and matching methods between two images and measure their similarity for that images such as SIFT, SURF, LBP, and so on. Similarly, rather than utilizing only limited local features, Global features give invariant descriptions of video frames. This method works effectively for video frames with distinct and distinguishing color values. Even though merits are simple to extract and have a low computing cost, global features failed to distinguish between consecutive frames. The following are the global features: image Gamma Return, get Blur Value of Image, get HSV Value of image, estimate noise, get Rotation of Image, contrast enhancement, dissimilarity, homogeneity, energy, correlation, and entropy so on.

Some approaches are based on machine learning classification algorithms for SVM and KNN and some approaches are applicable for deep CNN models with a track record of picture classification success and a huge amount of training data. To prevent the illegal use of video material, digital video copyright protection has become an important issue, and identifying video copies is a basic requirement. The major difficulty in detecting video copies efficiently and effectively is to identify concise & discrete show features. To that aim, a unique video copy detection method that makes use of the spatial-temporal domain for feature extraction and matching. Our most important contributions are mentioned below.

- Input query videos were segregated from each segment of the video sequences after breaking them into multiple segments. Generate feature maps using sampled video frames for matching and extraction of concise and exclusionary properties.
- To identify the content of a query video, the extracted feature extraction from a reference video is saved in a database and used for recognition in offline processing as well as online processing
- For each video frame, two temporal features are retrieved from frames to assess the attributes that complement the SVM and kNN feature but also can increase feature classification ability, resulting in promising detection performance.

The majority of this work is divided into different sections. Section II introduces the related works. Section III goes into great depth about the proposed video copy detection system. Section IV contains the results and analysis, whereas Section V contains the conclusions.

2. Literature Survey

It is proposed in this research to use SVM-KNN to extract global image features and apply them to a method for object recognition. The image's Moment invariant is used to compute global features based on the intensity of the entire image, and global features are generated. A literature study found that most object identification work is based on either local features or global features, with only a few studies using global characteristics for object recognition. There are two sorts of approaches for detecting copies and near-duplicated videos. The usage of a global descriptor is one approach to detecting the video copying features. Jane [1] Speeded-Up Robust Feature, a local feature descriptor, and detector, is used to retrieve related films. In addition to the Hessian matrix and the distribution matrix, the

system's performance can be improved by other techniques. SIFT, PCA-SIFT, and GLOH were found to be less effective by the author than SURF. Basly [7] Clustering video frames using locally SURF features and the KMeans sequential clustering technique has been suggested without any supervision. Its efficiency improved as a result of the elusion requirement of appropriate steps to prevent and the need for classification, as well as the number of nodes, formed increased once the threshold value was significantly lower compared to the other approach.

M.Yeh [2] proposes a multimedia warehousing-based method for video mining The method is split into two stages: construction of the warehouse and video retrieval from its multimedia storage. Euclidian Distance was applied as a similarity measure by the SVM classifier, which examined properties such as Local and Global Color Histogram as well as the RGB algorithms, Eccentricity algorithm, and the K-clustering Means Algorithms.

Hampapur[4] Retrieve video objects by first applying motion, then quantized color, and finally edge density features. Jiang, Q.Y. [5] introduced an SVM-based video search and indexing system. The system employs the state transition rectification approach and transition quality estimation.

Ullah [3] Defines a video object classification scheme in the TRECVID database utilizing offline feature extraction and machine learning to find a solution with a detection accuracy of 98.3 percent or 25 frequency per second. Wang [6] describes video summarization based on dynamic programming to discover the method with the highest tine complexity and total summary extraction of a one-hour film with 25 shots. When it comes to video frames, they were represented in an image by Huang and colleagues using a global image feature, such as the color histogram and texture. A color histogram in HSV color space was used by Wu et al. to identify and delete the bulk of web video fakes.

Each video frame was represented by a global visual feature such as a color histogram or texture. [12] Yeh defines an HSV color histogram was used by Wu et al. for web video duplication detection and removal.

Rajesekaran and Vijayalakshmi Pai [16] showed that moment invariant may be used as a feature extractor for ARTMAP picture categorization. To classify leaf images , Basly [7] employs an SVM with local characteristics. As Liu and colleagues point out, the support vector machine is effective in identifying microparts. Uses Gray le and Chen’s [10] novel approach for co-variance matrix and the moment invariants for battle scene categorization by

Karthika[11]. SVM is used by Ronald and the company to automatically identify deficiencies. le and Chen use a separate classifier for global and local features. Haar-like features are described in the study as local features, while edge features are described as global features. For license plate recognition from a video, the local features play a crucial role [11]. Lowe D.G introduced a scale-invariant feature descriptor, which is effective at detecting previously learned objects in crowded settings with posture changes and illumination changes. [16].

In [6], Leman outlined an innovative approach to GP-script development and evaluation that emphasizes the key components of a GP script. This unique method required direct communication with the actual pixel intensities, eliminating the need for previously established or extracted features to be provided by humans. The proposed method accepts an image as an input and generates a feature vector. A new terminal set, the function set, is used to manage texture image rotation variation.

3. Proposed Methodology

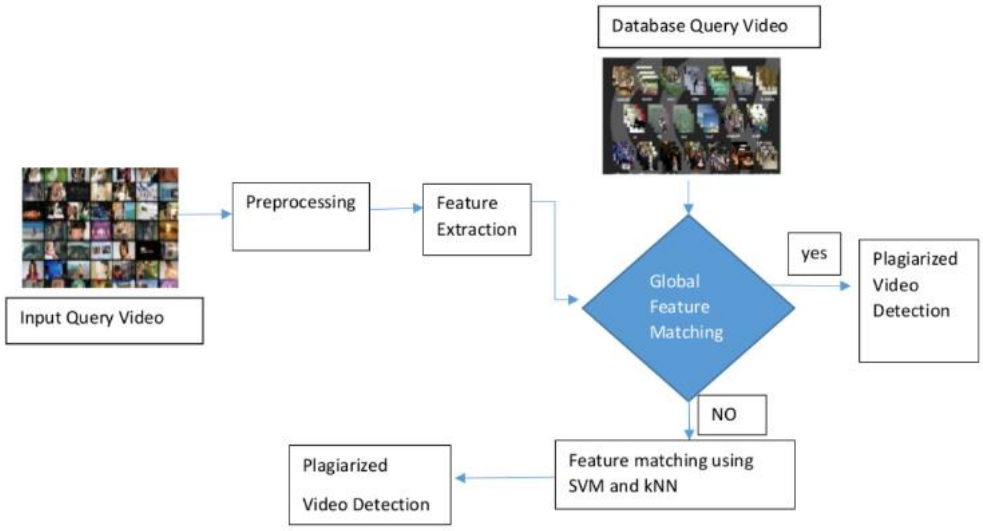

Figure 1. Proposed Methodology diagram for Global feature based on video copy detection system

In figure 1, Video preprocessing is the first step in the offline process (input image), which aims to improve the quality of the video and minimize the effects of transformations. Invariant features are taken from every keyframe that was discarded from the original video. Using these invariant features, it should be possible to distinguish between the original video and its transformed form. To speed up feature retrieval and matching, features are enlisted into a data structure for feature extraction. The online stage: Query videos are evaluated in

this stage of the evaluation process. It is necessary to compare the extracted features to the index structure before performing a feature description on the preprocessed key frames in a query video. After the keyframes on similarity is analyzed by using a Support Vector Machine and KNN classifier. Hu's Moment Invariant and the PCA-SIFT descriptor, which are both invariant to translation, rotation, and scaling, are used to describe the image as a global feature. The KNN is used to locate the closest neighbors to a query image first, and then the local SVM is used to determine the item that belongs to the object set. It is planned to implement the strategy in two stages. Initially, KNN is used to calculate the distance between the query and all training data, and then select the K neighbors that are nearest. In the second stage, support vector machines are used to identify the image. Finally, the system provides a video plagiarized.

3.1 Feature Extraction for Global Features

The moment invariant is important in object and pattern recognition. It has been shown to be the most promising and effective way to represent an image. As soon as the event occurs, it is feasible to reassemble the image. Hu's moment invariants can be rotated, scaled, and translated. An object's distinctive traits must be able to recognize the same object in different sizes and orientations if they are computed. Using Moment Invariants, an item can be identified even though its transformations have changed. preprocessing with Gaussian filtering reduces contrast in the dataset image captured at the start of the process. Get Blur Value of Image, get HSV Value of Image, estimate noise, and get Rotation of Image are used to extract color and texture features from images. Then, novel statistical feature extraction techniques are used to acquire effective features. Because of its efficacy, the k-means clustering method is then used for labeling. and finally, we found whether a given video is plagiarized or not for detection of all global features detectors and descriptors.

In order (a+b)th of a function f(x,y), the two-dimensional mathematical moment q is defined.

$$m_{ab} = \int_{a1}^{a2} \int_{b1}^{b2} x^p y^q f(x, y)\, dx\, dy. \quad (1)$$

Equation 1is defined by the x- and y-axes as a,b = 0,1,2,... and f(x,y) as the intensity value at a certain position in the image. Take note that for the time being, the primary function is the monomial product xpyq. All m ab's for a+b n is included in a set of n moments. This means that the set has 1/2(n+1)(n+2) members, which is equal to n. geometric moments, Hu came up with a collection of non-linear combinations of geometric moments.

moments that are invariant, It has the advantage of being translation, scaling, and rotation invariant. Extracting data using the Moment invariants is really convenient. characteristics derived from flat images Invariants of the moment are features of binary image areas that are related to translation, rotation, and scaling are all invariant.

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}} \tag{2}$$

$$\text{where} \quad \gamma = \frac{p+q}{2} + 1, \; p+q = 2,3,4..... \tag{3}$$

Invariants can be deduced from the second and third normalized central moments based on equation 2 & 3. This set of moment invariants is unaffected by translation, rotation, or scale change from (4) to (8). This is the process of extracting features from an input dataset and converting them into an appropriate set of features. It is at this point that the texture and color of the image are retrieved. [6-10]. Feature extraction mechanism incorporates three various features, including color and texture features, into a single system. These options are better for getting a good idea of what an image looks like. The extracted features are then fed into the k-means algorithm, which uses them to classify the data or classes to train the classifier. Algorithms for extracting global features of an image have been developed in recent years. In contrast to global features, The IPs of an image are extracted using local characteristics and described as a set of vectors before being combined into a single vector. Colors are an important part of the design. Color coherence is a key component (CCFs). Much about an image's content can be gleaned from the color distribution of its pixels. The CCFs aid in the distribution of picture color, which in turn supplies the image's properties. These pixels are evaluated in conjunction with others in the surrounding area to determine the likelihood that the desired color will be produced utilizing their respective color attributes. The intensity of the image is used to extract color features, which improves the system's performance. As a result, the proposed method for extracting color features utilizes an efficient method based on color intensity.

Algorithm

Input: Color Image Filtering

Features: Color

- The first step is to convert RGB to grayscale for future processes. A grey image is created by using the rgb2gray (Filt Color Image) function,

- In the second step, NIBCF is used to extract its color features.
- Image Intensity-Based Color Filter is used to generate the final output image.
- Third step We use the following steps to calculate the Image Intensity-Based Color Filter values.
- There are four levels of intensity: low intensity is equal to lo (intensity of input image); high intensity equals high (intensity of input image); Output J is equal to zeros (higher intensity minus low intensity+1); and finally, there are three levels of brightness: Low, High and Higher
- Then, the values of mov a and mov b are calculated.
- Once we know the values of mov a and mov b, we can calculate Intensity1 and Intensity2.
- The Output Image is generated using the intensity values.
- As the higher and low intensities increase, so do the intensity1 and intensity2 values.
- step four We estimate the input image's color features after performing Image Intensity-Based Color Filter and Output Image calculations.
- J= Final Image for j1 = 1:4: size - K (K,1)
- A(idxr,idxc) = mean(MF(:)) for idxr and idxc
- Color image novel features are represented by the Image Intensity-Based Color Filter Feature values: abs (A (1:28))
- End

For the characteristics of the texture to retrieve textures for feature extraction, we need GLCMs and new statistical features (NSF). Co-occurrence matrix of grey levels GLCM. The correlation coefficients between adjacent pixels are statistical features defined as "novel," or "new," statistical features. Mean, median, skewness (kurtosis), standard deviation, skewness (kurtosis), covariance (covariance), correlation (correlation), entropy (covariance). For an image, these computations yield a set of feature matrices. The use of an NSF is a cost-effective way to gather statistical data.

A standard transformation used in experimental simulations is PiP, which is the superposition of a video in the foreground and another in the background. The most common adjustments in genuine video copies, according to Yah [13], are pattern insertion, scale changes, and camcording.

Camcording is defined by [1] as the unlawful recording of a feature film while it is being projected or shown on a theatre screen, and the resulting recording is known as a

camcorder. Camcording is a serious concern in the film industry. According to industry sources, camcorders are responsible for at least 90% of the earliest available versions of illegally distributed new release films (see [1] for more details).

That is, the most prevalent visual attacks in real copies are camcording and edition attacks. Re-encoding and camcording reduce the video's quality, introduce blur and noise, and introduce projective transformations. The most typical edition or post-production alterations, according to [7], are the addition of a logo, crop, and shift.

The image intensity-based color features are depicted in color; a grayscale image is created from the RGB data. The second step is to determine the characteristics of the color. After that, the intent colorless are used to calculate the Image Intensity-Based Color Filter Consequently, the input image's color features are extracted color new color image features are the IIBCF feature values.

Image Gamma Return means the relationship between a color value and its brightness color described by the gamma scale. There should be a linear relationship between the display device's output brightness and each component's RGB color value for images to app color usually correct. This is not the case with the vast majority of display devices. A technique called gamma correction is used to correct a device's non-linear display characteristics.

A correction function, tailored to the display device's characteristics, is used to map the input values before sending them to the display device. A lookup table is commonly used to implement the mapping function, with a table for each of the RGB color components,

$$Tout = BTcolor \tag{5}$$

In equation 5, where real input that is not negative Tin is raised the owner of inverse function and multiple by the constant B to give the output of the value for Tout.

Spatial domain methods use low-pass filtering, which results in a blurry image. By utilizing the correlation between the pixel intensities of boundary pixels from adjacent blocks, this technique works.

The interest point is detected using the Hessian-Laplace blob detector. The SIFT [16] is used to extract local features once the interest point has been located. Sophisticated picture feature tracking techniques, such as SIFT, have a high feature count of 128 characteristics per

interest point. In order to compress the characteristics to 36 numbers, PCA is used, which simplifies the process. The PCA-SIFT descriptor is utilized in this paper to extract local features. These all are the features we have found in the global feature detector and descriptor in video plagiarized detection.

3.2 Classification of Feature Extraction

The k-nearest neighbor (KNN) method is one of the most basic in machine learning [13]. The closest neighbor approach is a nonparametric data categorization tool. It attempts to classify how close a source of data is to methods that are divided into two classes involving the collection of predictor variables nearest to it. The uses relevant metrics to determine the average distance between the two comparable points in that space. The algorithm then chooses who among the values in the training set are more similar and should be considered, while selecting the category in identifying a particular hypothesis was done simply by selecting the k closest pieces of data to that observation.

k-means clustering is used in this method of clustering. It is common to employ the k-means clustering algorithm. Color, size, and shape are used to categorize images in a database using clustering techniques. Shinde [14] was used to label every image in the database. When the image is analyzed, a set of extracted features and their associated labeling results are used to determine the image's classification. To enhance the image, a customized genetic algorithm is a retrieval process by optimizing the extracted features. The k-means clustering method is applied when an activation update is necessary. The final step is the retrieval of the captured images.

The following data has been entered: an initialized collection of image pixels. Labeled Clusters are the final product.

Step 1: is to compute the cluster's K centroids. Equation primarily assigns a maximum of K clusters (3).

Step 2: Determine the Euclidean distances between each pixel and each of its centroids.

Step 3: Determine which cluster each pixel belongs to using the Euclidean distance between every pixel and the closest cluster of recently computed centroids

Step 4: Create a new centroid by combining all the pixels in the same cluster into a single point to proceed to step 4, a maximum number of iterations must be met or all clusters must remain unaffected.

Step 5: Finally, processed image pixels are labeled with the most recently allocated cluster centroids.

Stop

Grouping pixels in an image according to the number of groups they belong to, up to a maximum of K to determine the K-centroid, it is assumed that the distance between a pixel and the gravitational center of its group is shorter than the distance between other groups centers of gravitation.

Using the centroid closest to each pixel, the Euclidean distances between each centroid and the pixel may be calculated & compared. Completely pixels in a group are computed and allocated to a group before the new distance is calculated. It is possible to recalculate the new centroids if the number of repeats or the centroids remains unchanged. Pixels in the centroid group associated with the final Euclidean distances are labeled with these final distances when training a naïve Bayes classifier K-means results typically deteriorate with increasing numbers of recursive calculations.

$$Kmeans_{result} = \sum_{j=1}^{k} \sum_{i=1}^{n} \|(x_i^j - c_j)\|^2 \tag{6}$$

$$\text{where } \|(x_i^j - c_j)\|^2$$

In equation 6, Groups or clusters are defined by the distance between the center of a group of pixels (xij) and the center of a cluster of pixels (n). If there are K clusters, then the centroid calculation is k = 1, 2, 3, K

Euclidean distance calculation:

$$D(m,n) = \sqrt{\sum_{j=1}^{m} (mi - ni)\ 2} \tag{7}$$

By using this formula, we have to calculate the distance between two nearest neighbor features.

SVM algorithm generates results after a group for characterized test datasets [9]. Support Vector Machine is a kernel approach for classifying and predicting data. SVM is based on the concept of selection vectors, which define the boundaries of the selection. A decision line classifies a set of inputs. Transformations are made when compared to the input data in order to make them distinct for classification purposes by increasing the number of dimensions. Any information not directly relevant to the intended prediction is discarded in regression because the results are based on the boundaries. SVM is used in a variety of fields, including face recognition, bioinformatics, and image processing.

SVM is a selective classifier defined in hyperplane separation. A method outputs an ideal-type plane that is divided into two halves after receiving properly labeled training data. In two-dimensional planar instances, the hyperplane splits the area into two groups, each of which lies on each side. The SVM finds the best hyper-plane and divides the n-dimensional feature space into two distinct classes. As aj = 1, the first class represents what we're talking about, while the second class represents all of the other notions When the distance between training examples for both classes is maximised, a hyper-plane is said to be optimal. The maximized are used to describe this area. I > 0 support vectors are used to determine the margin of safety determining (λ BΛΛΛλ + D $\sum$ ξj s).

The first class represents the current notion, while the second class represents the remaining concepts, i.e. ai = 1. When the distance between neighbouring training cases for both classes is minimised, a hyper-plane is said to be optimal. This is known to as the margin. The support vectors, I >0, that is acquired via optimizing, define the margin. It's worth noting the significance of this kernel function M(), which translates the distance between feature vectors into a higher dimensional space where the hyper-plane separator and its support vectors are obtained. Again when the SVM are accepted, specifying a conditional probability for an unknown test sample x′ becomes simple. SVM classifiers were built using hue features. SVM classifier and test dataset feature vectors will be built using simple training scaling. The correct kernel function is used, such as RBF or linear kernel function. The optimal B and gamma parameters were determined via cross-validation. Train the entire training set with the appropriate B and gamma parameters, and then predict the class for the sample set. To forecast output, SVM classifiers use global features such as Hue moments and HSV histograms obtained from key-frames. To accurately estimate the output, the proposed system leverages an RBF kernel with optimal C and y values.

4. Experimental Results

After the edges are eliminated, the Hu's moment invariants are computed, and the feature vector is generated. With the help of the training images, KNN is used to identify the nearest neighbours of a given image. Otherwise, the SVM is used if a label can be obtained. If not, the algorithm ends. The object was recognized with the help of the new algorithm. One SVM and one KNN are used to compare the results. According to the findings, the suggested classifier outperforms other existing classifiers. The proposed framework was evaluated using data from the VCDB and lecture video. In the system different classes, 800 training videos, and 500 assessment movies are used. This table illustrates the confusion matrix of SVM, kNN and SVM and kNN fused to Tables 3, 4, 5, and 6. Twelve classes are employed in the experiment. With a classification accuracy of 0.89, the SVM classifier's output confusion matrix is shown in Table 3.

Table 1. Details description of the VCBD dataset

Dataset	Frames partitions	Dataset	No. of Key Frames
VCDB 2014	Divide-1	Training Datasets	800
	Divide-2	Testing Datasets	500

Table 2. The dataset has several classes

Category	Training video count	Testing Video count
Sports	150	100
Trailer	135	85
Lecture	110	90

Table 3. Confusion matrix for SVM

Methods	Sports	Trailer	Lecture
Zoom	15	10	5
Blur	0	15	9
Border	11	5	25

Crop	17	0	3
Flip	3	18	6
Noise Addiction	3	25	0
Rotate	1	5	0
Picture in Picture	10	7	3
Low Contrast	4	8	5
High Brightness	15	12	8
Text insertion	0	7	7
High Contrast	12	11	9

Table 4. Confusion matrix for kNN

Methods	**Sports**	**Trailer**	**Lecture**
Zoom	5	0	3
Blur	0	5	2
Border	1	3	25
Crop	7	1	9
Flip	6	14	8
Noise Addition	7	29	1
Rotate	5	7	1
Picture in Picture	1	3	4
Low Contrast	0	5	2
High Brightness	5	2	4
Text insertion	3	2	6
High Contrast	12	10	9

Table 5. Confusion matrix for SVM and kNN

Methods	Sports	Trailer	Lecture
Zoom	24	6	8
Blur	11	9	3
Border	3	13	5
Crop	16	12	12
Flip	7	15	1
Noise Addiction	33	22	11
Rotate	9	16	23
Picture in Picture	17	23	7
Low Contrast	29	31	9
High Brightness	8	24	12
Text insertion	14	3	5
High Contrast	11	1	20

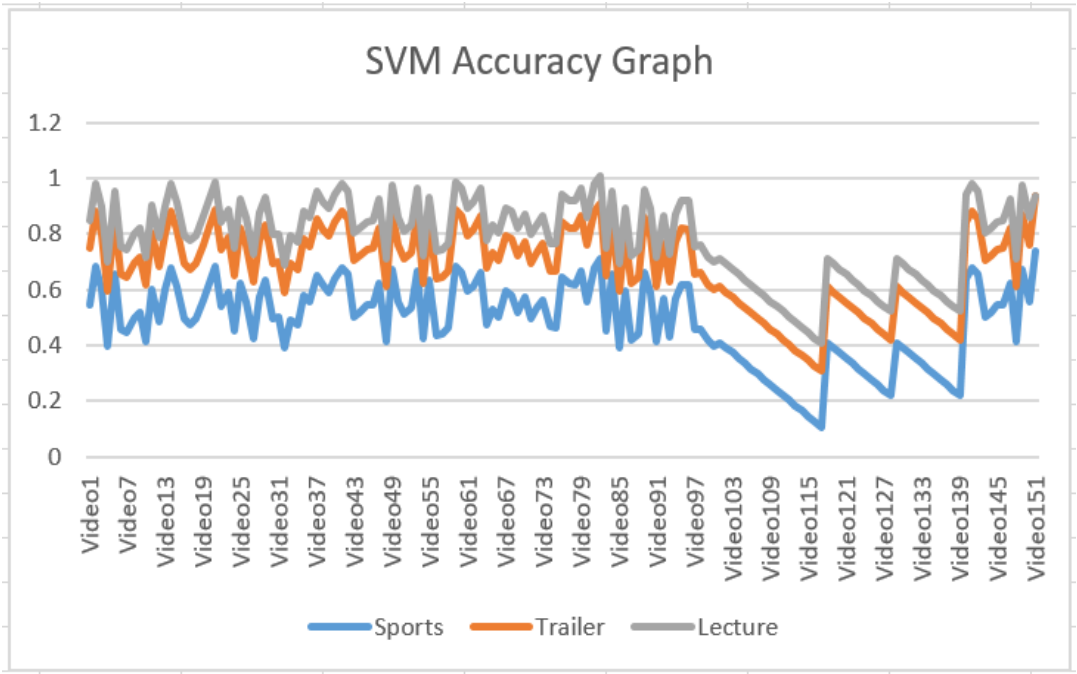

Figure 2. SVM Accuracy Graph

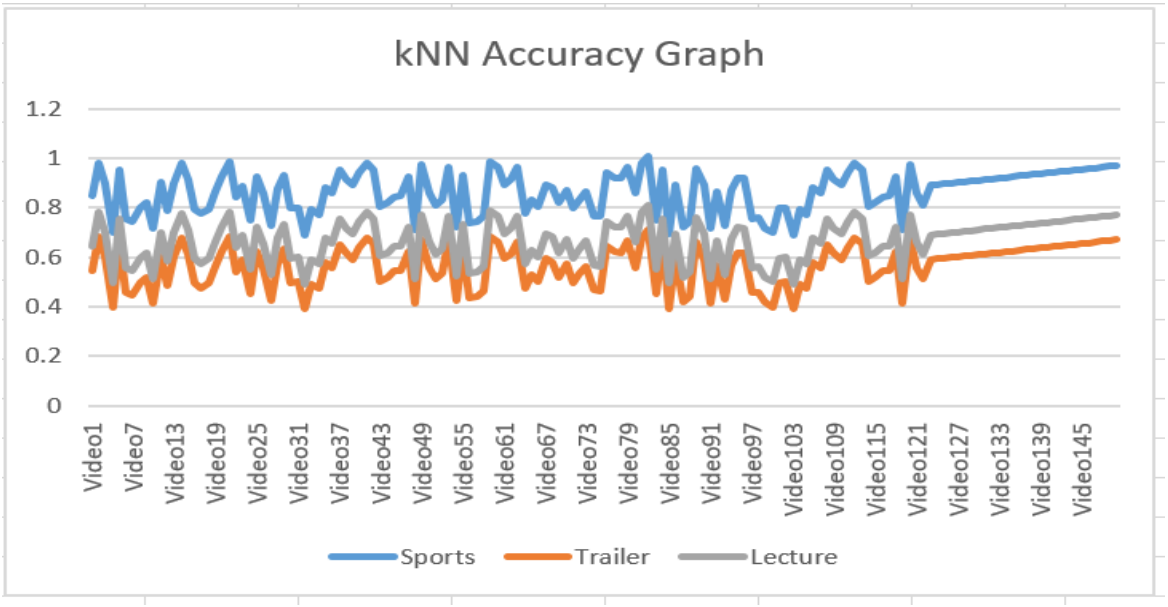

Figure 3. kNN Accuracy Graph

Table 4. Accuracy for SVM & kNN

Accuracy	Sports	Trailer	Lecture
SVM	80%	94%	92%
kNN	91%	90%	93%

Table 4 presents the accuracy for classifier SVM and kNN with obtained kNN has the highest accuracy in Lecture 93% as compared to other class.

5. Conclusion

Within the focus of this research, we have suggested a global feature-based object detection method that takes advantage of the hu's moment invariant during the processes of rotation, scaling and translation. In an initial step, picture gamma return, get Blur Value of Image, get HSV Value of Image, estimate noise, and so on using feature extraction techniques. Based on the feature extraction, the classifier KNN-SVM is utilized to determine the object. If there is no match, the KNN classifier is used first to locate the closest item from the nearest features, and to identify the object, a support vector machine is used. The proposed method has a higher degree of accuracy when it comes to object recognition. Even if just a portion of the object is known, the global characteristics can be used to locate it with a high degree of certainty. Based on the findings of the experiment shown in the table. In our

view, it is quite clear that combining SVM and KNN with global characteristics can produce superior outcomes that output says that the video is being plagiarized. A method that uses polynomial function as the kernel function was presented, and the majority of the findings are even better than those obtained using conventional methods such as KNN and SVM. Kernel principal component analysis simplifies the handling of high-dimensional data by reducing the feature vector. To overcome the challenges posed by the severe content variances, future research on copy detection should focus specifically on the frame matching step.

References

[1] Janwe, Nitin J., and Kishor K. Bhoyar. "Neural network-based multi-label semantic video concept detection using the novel mixed-hybrid-fusion approach." Proceedings of the 2nd International Conference on Communication and Information Processing. ACM, 2016.

[2] M. Yeh, K. T. Cheng, "Video copy detection by fast sequence matching", in ACM Proc. of Int.Conf. on Image, Video Retrieval, 2009.

[3] Ullah, F., Wang, J., Farhan, M., Jabbar, S., Wu, Z., & Khalid, S. (2020). Plagiarism detection in students' programming assignments based on semantics: multimedia e-learning based smart assessment methodology. Multimedia tools and applications, 79(13), 8581-8598.

[4] Hampapur, A., Hyun, K., & Bolle, R. M. (2017, December). Comparison of sequence matching techniques for video copy detection. In Storage and Retrieval for Media Databases 2002 (Vol. 4676, pp. 194-201). International Society for Optics and Photonics.

[5] Jiang, Q. Y., He, Y., Li, G., Lin, J., Li, L., & Li, W. J. (2019). SVD: A large-scale short video dataset for near-duplicate video retrieval. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 5281-5289).

[6] Wang, Y., Hou, Z., Leman, K., Pham, N. T., Chua, T., & Chang, R. (2011, January). Combination of local and global features for near-duplicate detection. In International Conference on Multimedia Modeling (pp. 328-338). Springer, Berlin, Heidelberg.

[7] Basly, H., Ouarda, W., Sayadi, F. E., Ouni, B., & Alimi, A. M. (2020, June). CNN-SVM learning approach is based on human activity recognition. In International Conference on Image and Signal Processing (pp. 271-281). Springer, Cham.

[8] Muralidharan, R., & Chandrasekar, C. (2012, March). Combining local and global features for object recognition using SVM-KNN. In International Conference on Pattern Recognition, Informatics and Medical Engineering (PRIME-2012) (pp. 1-7). IEEE.

[9] Jiang, Y. G., Jiang, Y., & Wang, J. (2014, September). VCDB: a large-scale database for partial copy detection in videos. In European conference on computer vision (pp. 357-371). Springer, Cham.

[10] Liu, H., Fan, Z., Chen, Q., & Zhang, X. (2022). Enhancing face detection in video sequences by video segmentation preprocessing. Applied Intelligence, 1-11.

[11] Karthika, P., & VidhyaSaraswathi, P. (2018, May). Digital video copy detection using steganography frame-based fusion techniques. In International Conference on ISMAC in Computational Vision and bio-engineering (pp. 61-68). Springer, Cham.

[12] Zhou, Z., Yang, C. N., Chen, B., Sun, X., Liu, Q., & QM, J. (2016). Effective and efficient image copy detection with resistance to arbitrary rotation. IEICE Transactions on Information and systems, 99(6), 1531-1540.

[13] Yeh, M. C., & Cheng, K. T. (2009, July). Video copy detection by fast sequence matching. In Proceedings of the ACM International Conference on Image and Video Retrieval (pp. 1-7).

[14] Shinde, S. R., & Chiddarwar, G. G. (2015, January). Recent advances in content-based video copy detection. In 2015 International conference on pervasive computing (ICPC) (pp. 1-6). IEEE.

[15] Lian, S., Nikolaidis, N., & Sencar, H. T. (2010). Content-based video copy detection–a survey. Intelligent Multimedia Analysis for Security Applications, 253-273.

[16] Rajesekaran and Vijayalakshmi Pai. (2011). Object recognition using SVM-KNN based on geometric moment invariant. International Journal of Computer Trends and Technology, 1(1), 215-220.

Author's biography

Ekta Sarda is currently working as an assistant professor in Computer engineering department, RAIT. She is having 15 years of teaching experience. She has completed her ME in computer science and engineering from Rajasthan Technical university kota and pursuing PHD in computer engineering from PAHAR University. Her area of specialization is Pattern Recognition. She has published paper in 6 journals and 3 international conference and 1 book chapter. She has also published one book in Computer Network.

Jayshree Jain is a professor at Pacific University. After graduating in IT, she did MTech, MBA and Ph.D. Her research area in her Doctorate was & quot; Potential of e-Learning in Indian Education System & quot;. She is a life member of CSI, IIIE, ISTD, and many other professional bodies. He has been a member of the organizing committee for various national and international seminars/workshops. Dr. Jain has also completed SCJP Certification.

Vaibhav Narawade is a Professor at the Ramrao Adik Institute of Technology at Nerul affiliated to the University of Mumbai where he has been a faculty member since 2019. Professor Narawade area of expertise includes Wireless Sensor Network, Image Processing, and Data Science. He has having and Administrative He has served University of Mumbai as Member 2010) and currently also served as Member of Management Council and Member of Senate, University of Mumbai (2017 to 2022), In Academics he is a n Computer Science and Engineering, Information Technology fraternity, and Faculty Narawade has author / coauthored in various International conference / journal publications and book chapter. He is a fellow of ISTE and ACM.