

A Novel Adaboost Regression Classifier for Video Retrieval in Video Sequence

B. Satheesh Kumar^{a,1}, K. Seetharaman^b, B. Sathiyaprasad^a

^a*Research Scholar, Department of CSE, Annamalai University, Tamil Nadu*

^b*Department of Computer and Information Science, Annamalai University, Tamil Nadu*

Abstract. This paper presents a new method for video retrieval, based on machine learning with regression. The proposed classification technique integrates Adaboost and regression classifier for significant retrieval of video frame. The proposed method consists of three stages such as key frames segmentation and gradient of pixels. In this technique, Adaboost classifier is involved in removal of noisy or blurred pixel of the segmented frame. Regression technique converts the video frame pixel either 0's or 1's which eliminates the noises in the frame. For the query video, the adopted classifier evaluates the machine learning system for retrieval of similar frames in the databases using proposed Adaboost Regression (ABR) classifier. Experimental analysis is conducted for video datasets to evaluate the proposed ABR classifier performance evaluation. Results stated that through proposed ABR approach incorporated in machine learning system effectively retrieve video frame for query frame. The proposed ABR classifier technique significantly improves the retrieval rate in terms of accuracy, precision.

Keywords. Video retrieval, Adaboost Regression Classifier, Machine Learning, Key frames, Query video.

1. Introduction

In recent years, repository of digital video has drastically increased which requires effective technique for retrieval of video. Simultaneously, machine learning approach exhibits significant performance pattern of recognition, detection, classification and retrieval for improving performance of neural network [1]. Also, in recent years, several people perform communication through Internet for data transmission and collection. Due to development of information transmission through visual content information are increased. The information visual content is based on video frame diversity, importance, availability and complexity which have been adopted in various fields with respect to unparalleled variation [2]. The evolution of smart phones, people use event for recording which has been performed around them and upload in the network. This leads to flexible transmission of information and sharing, this can be websites of social media, Weibo and Weichat for recorded things happens around them rather than words, which utilizes video frame for sharing of information [3]. This causes drastic growth for video frame resources in the network. However, increased video frame leads to academic research problem with requirement of large video frame

¹ Satheesh Kumar B, Research Scholar, Department of CSE, Annamalai University, India.
Email: vbsatheesh30@gmail.com

data requirement and improves the coverage rate and acquiring of information capability which causes research problem [4-6]. To overcome the limitation related to traditional algorithm deep learning technique has been evolved which is based on the consideration of huge number of programmed designed features [3]. This are all based on the machine learning technique, this uses classifiers are improving the performance of network. The classifiers perform grouping of video or images for semantic categories. The persistence of intra and inter class variability and similarity classification is considered as challenging factor based on machine learning technique [2].

2. Proposed model

The proposed machine learning technique integrates two classifiers for video frame retrieval, that based on Adaboost and regression ABR classifier. The proposed ABR classifier consists of different stages such as pre-processing, feature extraction, classification and retrieval of video sequence based on query frame. The Logistics function value ranges from 0 and 1 in case if value is above 0.5 than automatically it is considered as 0 [11].

$$h_0(x) = g\left(\frac{1}{1 + e^{-\theta T_x}}\right)$$

Above mentioned logistics equation is modified in order to achieve sigmoidal function to cope with adaboost classifier to minimize computational time. The mathematical formulation is mentioned as follows [2]: $y = \omega^T x$

The above equation is basic linear equation model for logistics regression for obtaining sigmoidal function in linear function. To derive a novel ABR Classifier both adaboost classifier and regression classifier are integrated and the basic sigmoidal function with

limit $(-\infty, \infty)$ is [12]: $\frac{1}{1+e^{-x}} = \frac{e^x}{e^x+1}$

Taking probability values for regression

$$P = a_0 + a_1x_1 + a_2x_2 + \dots + a_kx_k$$

$$\left[\frac{P}{(1-P)}\right] = b_0 + b_1x_1 + b_2x_2 + \dots + b_kx_k$$

Taking log on both sides,

$$\log\left(\frac{P}{1-P}\right) = \log(w^T x)$$

After applying natural exponential property,

$$\log\left(\frac{P}{1-P}\right) = \sum_{j=0}^{\infty} b_jx_j$$

Where $P = \sum b_jx_j$ hence for logistic regression equation

$$P = \frac{\exp(b_jx_j)}{[1 + \exp(b_jx_j)]}$$

ABR algorithm's performance can be improved by computational time and accuracy chain rule property and maximum likelihood property is combined; the chain rule property and maximum likelihood property used in this research are stated as follows [7]:

$$F'(x) = F'g(x)g'(x)$$

After applying above property and simplification, we obtained equation as,

$$P = P(k)(1 - P(k))$$

Maximum likelihood estimation for P is,

$$\hat{l}(\theta; x) = \frac{1}{n} \sum_{i=1}^n \ln f(x_i | \theta)$$

$$P = \sum \log P(k_i) + \sum \log (1 - P_i(k_i))$$

Removal of negative term offers,

$$\sum P = \sum P_i$$

Now, $P = (a_0 + a_1x_1 + a_2x_2 + \dots + a_kx_k) \sum P_i$

Hence the final equation for proposed ABR Classifier algorithm is,

$$H = \text{Sigmoid} \left(\sum_{j=1}^N P_i \alpha_i h_i(x) \right)$$

2.1 Preprocessing

In first stage, pre-processing is performed for generated classifier for evaluation of input data sequence in machine learning for extraction of features. This stage minimizes computational processing of video sequences.

2.2 Keyframe selection

Generally, video sequence consists of discrete frame with inclusion of large and redundant data. The computational complexity of the conversion of frame is higher due to large number of frames in the video sequence. In order to overcome this issue most appropriate frame which provides appropriate video information is selected. The proposed classifier estimate keyframe features with consideration of HSV, red, blue, texture feature Gray Level Co-occurrence Matrix and gradient feature HOG and entropy for evaluating data stored in particular frame about video. Almost similar frame are clustered utilizing k-means clustering technique.

2.3 Segmentation

Segmentation is defined as identification or isolation of interested object extracted from the background. After completion of segmentation it stops processing. In this research, we used several video sequence to extract objects in the keyframe from statistical region for merging process. For the segmentation process computational complexity is minimized for learning and training for keyframe deep feature extraction.

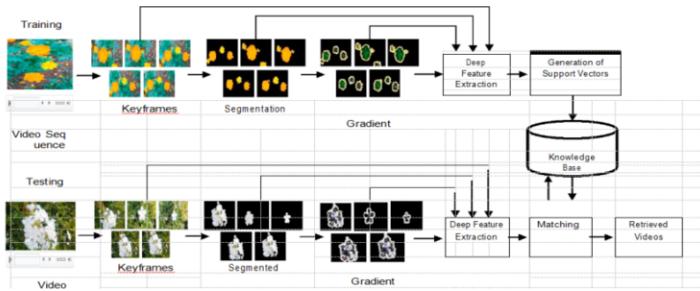


Figure 1. Processing of testing and training

2.4 Gradient

The gradient of video frame are obtained through segmentation of objects using Gaussian filtering. The gradient of features minimal information about the complete image technique for reducing computational complexity of machine learning process through deep extractor features of the keyframe and segmented video frame.

$$G(x) = \frac{1}{\sqrt{2\pi S^2}} e^{\frac{-x^2}{2S^2}}$$

3. Machine learning with proposed ABR Classifier

Machine learning composite of learning features on the video frame through consideration of higher level and hierarchy architecture. Machine learning technique collects features of the video frame keyframe based on the fed input data is major application of computer vision and machine learning. The below architecture represent the architecture adopted for training and testing. Throughput of 1st layer is fed as second layer input, in similar manner entire architecture is performed its operation for video frame retrieval. Output obtained for pooling minimizes amount of specifications in system which need to learn for operating non-linear down sampling. ReLU section by means of conversion of negative values of the frame for mapping zero and perform faster for effective training for maintaining positive values. An image size of 227x227x3 involves convolution operation with inclusion of 96 kernels, second layer throughput uses 256 kernels. 3rd layer of network takes input of 2nd layer and filtering operation with 384 kernels. 4th layer of architecture consists of 384 kernels in which fully connected layer consists of fc6 and fc7 includes 4096 neurons. Softmax layer consists of fc8 with 1000 neurons. The proposed classifier includes adaboost regression classifier for processing machine leaning operation.

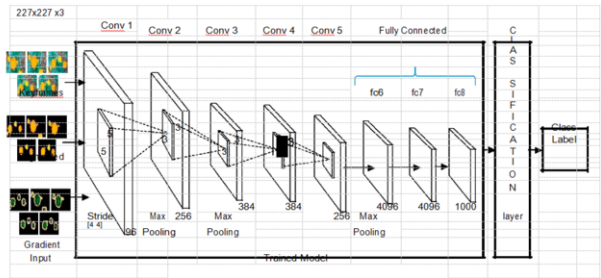


Figure 2. Architecture of Proposed ABR

Datasets: The video related to public domain are collected from available sources and dataset are framed and classified as Dataset - 1 and Dataset - 2. In dataset -1, 2707 video are considered and in dataset -2 2636 videos are considered with 30 various classes in which each classes consists of 35 to160 videos. Those video are in the duration of 4 to 60seconds. The video frame consists of large intra and inters class variations. In real environment video frames are captured with variations of lighting, illumination, partial occlusion and background. The sample video frame are presented below.

VIDEO	FRAME 1	FRAME 2	FRAME 3	FRAME 4
VIDEO 1				
VIDEO 2				

Figure 3. Sample video data

4. Experiments and Results

For experimental analysis of proposed system is considered for processing the video sequence dataset. Initially, keyframes are identified for the obtained video sequence followed by merging of statistical region for segmentation of the keyframes of the video. Then using Gaussian gradient filter frame feature gradient are examined. The features of the video frame are learned and extracted based on the consideration of keyframes and using machine learning technique segmentation and gradient of video frame are measured. The adaboost and regression classifier are used for machine learning process for recognition of query frame video sequence. For the proposed classifier performance in video retrieval metrics like recall, precision, accuracy and F-measures are evaluated. Solutions obtained through changing number of testing and training video. The dataset-1 and dataset - 2 retrieval results are tabulated below from table 1 to table 3. The metrics used for measurement of video retrieval is presented as follows:

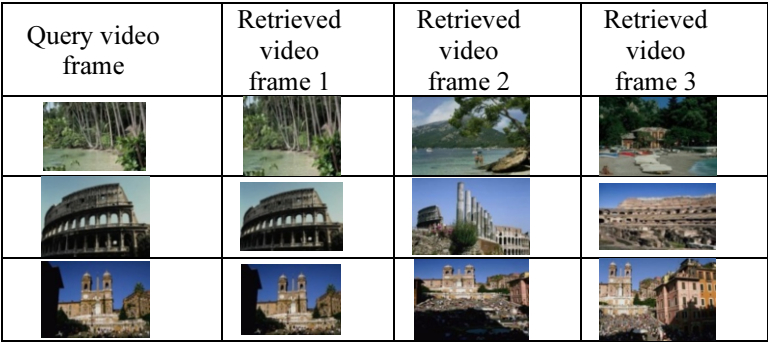


Figure 4. Retrieval of ABR approach

4.1 Comparative Analysis of proposed ABR classifier

The proposed ABR classifier uses machine learning technique for effective feature extraction used to retrieval of video frame using machine learning algorithm. The existing literature related to video retrieval techniques are Gray Level Co-occurrence Matrix [8], Local Binary Pattern [9] and Scale Invariant Feature Transform [10]. Recovered process involved in extraction of Gray Level Co-occurrence Metric, Local Binary Pattern and Scale Invariant Feature Transform. Comparative analysis of proposed classifier with existing conventional approach is exhibited, that the proposed ABR approach achieves good recover efficiency.

Table 1: Two types of key frames without segmentation

Train %-Test%	Precision		Accuracy		F-Measure		Recall	
	DATASE T-1	DATASE T-2	DATASE T-1	DATASE T-2	DATASE T-1	DATASE T-2	DATASE T-1	DATASE T-2
30-70	91.26	85.73	92.17	84.05	90.74	82.19	91.68	83.31
40-60	93.02	87.57	92.79	86.78	91.94	85.80	93.06	86.97
50-50	92.70	89.73	92.50	89.76	91.56	88.57	92.85	89.86
60-40	96.03	90.53	96.53	90.68	95.87	89.56	96.25	90.39
70-30	97.38	92.50	97.37	92.53	96.64	92.33	96.56	92.21
80-20	97.95	95.00	97.72	94.37	96.72	94.69	96.59	94.67

Table 2: Two types with segmentation

Train %-Test%	Precision		Accuracy		F-Measure		Recall	
	DATASE T-1	DATASE T-2	DATASE T-1	DATASE T-2	DATASE T-1	DATASE T-2	DATASE T-1	DATASE T-2
30-70	93.25	87.18	92.61	86.46	92.27	84.11	92.39	85.26
40-60	94.65	88.97	94.30	89.45	94.23	88.36	94.62	89.35
50-50	94.09	90.83	93.77	92.13	93.29	90.96	93.69	92.22
60-40	95.40	94.70	95.14	95.39	95.17	94.66	95.91	95.07
70-30	96.75	96.40	97.25	96.65	96.82	96.46	97.20	96.70
80-20	97.08	97.08	97.79	97.09	97.08	96.76	97.26	96.72

Table 3: Two types of with Gradient

Train %-Test%	Precision		Accuracy		F-Measure		Recall	
	DATASE T-1	DATASE T-2	DATASE T-1	DATASE T-2	DATASE T-1	DATASE T-2	DATASE T-1	DATASE T-2
30-70	87.35	76.23	88.40	79.63	86.55	75.78	87.00	77.77
40-60	89.59	82.36	90.51	83.15	88.91	81.10	89.15	82.59
50-50	90.64	85.49	90.28	86.48	89.42	84.55	89.99	85.81
60-40	92.39	89.51	92.62	89.82	92.04	88.58	93.03	89.26
70-30	94.02	91.68	94.63	91.37	93.79	91.26	94.25	91.61
80-20	94.93	92.70	95.27	91.84	94.30	91.44	94.62	91.27

5. Conclusion

In this paper proposed a classifier technique for video retrieval for query video frame sequence are presented. The presented technique is based on the keyframe for clustering analysis for video retrieval. The proposed ABR approach adopted in three modalities keyframe, segmentation and gradient for feature extraction using machine learning technique. Results stated that proposed ABR classifier for improving video retrieval efficiency rather than existing technique.

References

- [1] Zhang, Chengyuan, et al. "CNN-VWII: An efficient approach for large-scale video retrieval by image queries." *Pattern Recognition Letters* **123** (2019): 82-88.
- [2] He, Ran, et al. "Learning structured ordinal measures for video based face recognition." *Pattern Recognition* **75** (2018): 4-14.
- [3] Min, Min, et al. "Retrieval of cloud top properties from advanced geostationary satellite imager measurements based on machine learning algorithms." *Remote Sensing of Environment* **239** (2020): 111616.
- [4] Zhou, Lina, Yin Xiao, and Wen Chen. "Vulnerability to machine learning attacks of optical encryption based on diffractive imaging." *Optics and Lasers in Engineering* **125** (2020): 105858.
- [5] Bressan, Rafael S., Pedro H. Bugatti, and Priscila TM Saito. "Breast cancer diagnosis through active learning in content-based image retrieval." *Neurocomputing* **357** (2019): 1-10.
- [6] Passalis, Nikolaos, et al. "Variance-preserving deep metric learning for content-based image retrieval." *Pattern Recognition Letters* **131** (2020): 8-14.
- [7] Zhang, Zhenwei, and Ervin Sejdić. "Radiological images and machine learning: trends, perspectives, and prospects." *Computers in biology and medicine* (2019).
- [8] Pedronette, Daniel Carlos Guimarães, et al. "Semi-supervised and active learning through Manifold Reciprocal kNN Graph for image retrieval." *Neurocomputing* **340** (2019): 19-31.
- [9] Wang, Yanfei, et al. "Deep cascaded cross-modal correlation learning for fine-grained sketch-based image retrieval." *Pattern Recognition* **100** (2020): 107148.
- [10] Tzelepi, Maria, and Anastasios Tefas. "Deep convolutional learning for content based image retrieval." *Neurocomputing* **275** (2018): 2467-2478.
- [11] Tang, Jie, et al. " π -Hub: Large-scale video learning, storage, and retrieval on heterogeneous hardware platforms." *Future Generation Computer Systems* **102** (2020): 514-523.
- [12] Alzu'bi, Ahmad, and Abdelrahman Abuarqoub. "Deep learning model with low-dimensional random projection for large-scale image search." *Engineering Science and Technology, an International Journal* (2020).