ISSN: 0974-5823 Vol. 6 No. 1 January-June, 2021

International Journal of Mechanical Engineering

An Analytical Survey on Texture Based Patterning for Efficient Tracking and Object based Classification in Video Processing

G. Devakumar

Department of Computer Science & Engineering, MS University, Tirunelveli, Tamilnadu, India – 627012

D. Murugan

Department of Computer Science & Engineering, MS University, Tirunelveli, Tamilnadu, India – 627012

Abstract **-** Computer vision has become popular now a day due to its processing and analyzing benefits. Research areas in computer vision are encountered as Frame Conversion, Preprocessing, Key Frame Identification, Shot Detection, Object Detection, Segmentation, Classification and Object Tracking. This literature review concentrates on classifications and Patterning techniques use for classifying video files, frames and objects. The first part covers video classification techniques. Classification techniques are divided into two major sub parts namely Classification on video clips and classification on frames / objects. The second part covers the texture based patterning techniques which increase the accuracy on object tracking. There are three texture based patterns LBP, LEP and LRP have been taken and their performances were analyzed. Finally suggestions are given as conclusion for making the intelligent surveillance system for classifications as well as texture based object tracking.

Index Terms - Video Classification, Object Classification, Frame Classification, Moving Object Tracking, Local Binary Pattern (LBP), Local Extrema Pattern (LEP), Local Rhombus Pattern (LRP).

I. INTRODUCTION

In 1980s the classification had been done in another name as interpretation of objects in captured videos through knowledge based representation methods along with control strategies. Familiar knowledge representation methods were schema, semantic nets, associative memory, declarative schemata, etc, and hypothesie test, query based and perceptual cycle were played as control strategies [1]. The process of dividing scene into parts was called as segmentation. There were techniques used for completing the segmentation task in computer vision namely Thresholding (Image's Histogram value), Edge or Feature Detection (abrupt changes in brightness), Region Growing (based on light changes thresholding) and Textural segmentation (Edges separated by brightness discontinuity) [2]. In 1990s researchers started to expose and classify valuable information from the surveillance videos by the use of image processing techniques. As a result suspicious activity of people, suspicious activity of vehicles in road, suspicious pattern of movement of people near safety critical equipment, Automatic number plate recognition (ANPR), vehicle types detection, crowd density analysis, people tracking by shape models, track cars by 3D models, etc, were proposed through video surveillance CCTV recorded data [5].

Typically, video surveillance system comprises four major phases such as video object identification, video object recognition, video object tracking and video object classification when needed. Classification plays a main role for several functions such as events identification, Intrusions detection, Key frame extraction, Speed control, and so on. When Classification is done with video files, the surveillance systems become an intelligent system of video monitoring. Classification among frames might be based on text, audio and video features [7],[11],[13],[19],[21]. But the techniques differ when the classification has to be done with video clips. By using inherent features of video files like light variance, shapes, audio, visual, size, Object, frame

and speed, the video files can be classified. When it comes to tracking the object in video frames, objects are to be classified based on object's features such as shape, colour and texture. Classification techniques which are surveyed here are suitable only with the dataset generated by normal CCTV cameras not with specialized camera like thermal camera. Processing with thermal camera, a novel feature extraction method based on Target Trait Context was proposed for extracting the boundary of detected objects. Through which objects could be classified [23].

This literature covers two major parts involve in video processing ie., classifications and texture based tracking. First part narrates the classification techniques and their advancements along with merits and demerits in video processing field. Second part covers the texture based patterning techniques (LBP, LEP and LRP) which are used to avoid dilemma and increase accuracy on object tracking. Finally this thorough literature study ends with conclusion and suggestion to make intelligent surveillance system.

II. Classification Study in Computer Vision

There are two varieties of classification in video processing. They are 1. Classification among video files in video dataset and 2. Classification among frames or objects in a video file. The figure1 describes the types of classification in computer vision. This literature surveys both classifications with the help of past research findings.

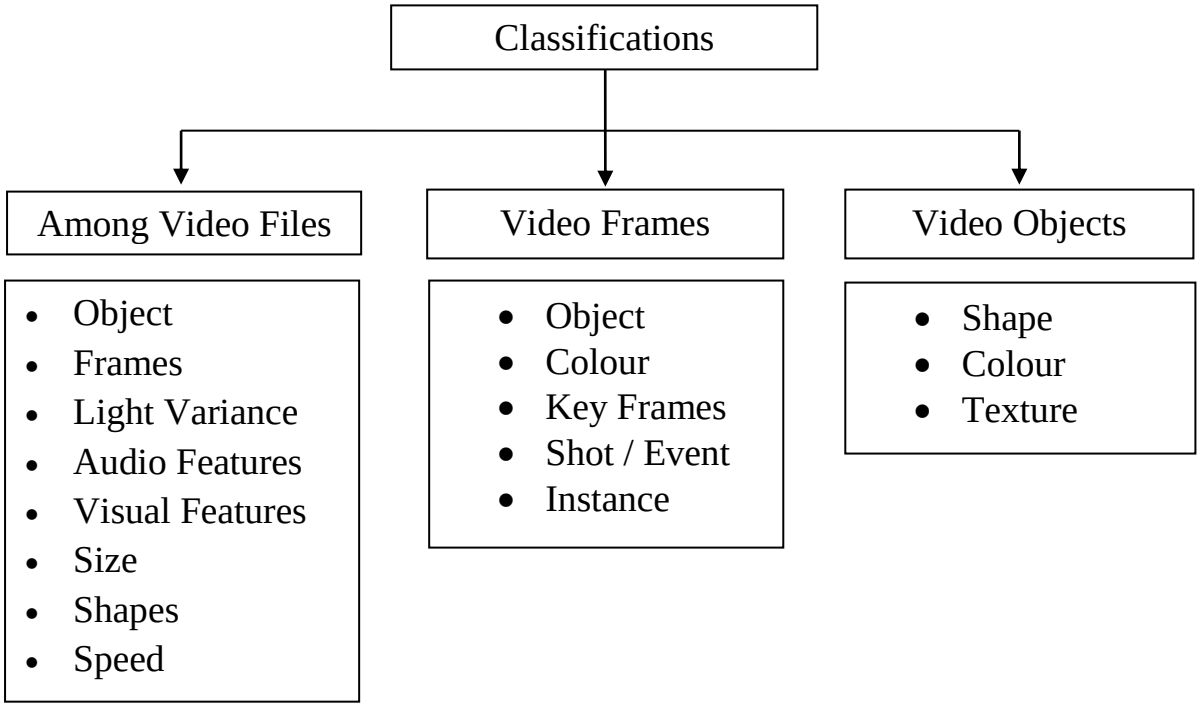

Fig. 1. Classifications type in video processing

a. Classification on Video files

Initially classification in computer vision was based on audio and visual effects of the videos [02] [03]. In 2007 P. Wang et al proposed the techniques based on text classifiers along with the assistance by audio video clues since it was observed that the reliability of audio video confidence is lower than the text in terms of classification. News video clips dataset was taken for this experimental work and the news story classification was done by their proposed multi model features. In which they introduced SVM confidence vector for extracting text features and GMM technique confidence vector for extracting audio visual features, and then Text-Biased decision strategy was used for combining these two multi model confidence vectors. Text based classification and SVM based meta-classification techniques were compared for classifying the video news stories [17]. In mid of 2010 Video clips were classified by another work [20]. Here three clip-level representations were experimented; they are single Gaussian modeling, Gaussian mixture modeling, and probabilistic latent semantic analysis of a Gaussian component histogram. The classification of video clips in

the video set of huge video clips was completed by the well known support vector machine (SVM) classifiers based on the Kullback–Leibler, Bhattacharyya, or Mahalanobis distance measures with some summary features. Until then the text classifiers involved for video files, classification datasets didn't get any modification. But in 2011 March there was one research publication [12] by A Karpenko et al with the technique that a dataset was created with large collection of highly compressed tiny youtube videos. Firstly, large amount of same kind youtube videos had been collected and then the files were compressed as tiny videos for speedy classification. A variety of frame classification tasks performed based on simple nearest neighbor methods and the rapidness while classification was ensured. As a continuation, video files were classified into news, cartoon, movie, song, etc., by audio and video based classification by [13]. Mel frequency cepstral coefficient was used for audio data characterization and colour histogram features were extracted for visual data characterization. This proposed classification was done based on auto associative neural network (AANN).

During March 2015 Nirav Bhatt carried out a survey on the techniques that used for video classification. In that literature, text, audio and visual features based video files classification techniques were illustrated. The video based approach consists of color histogram changes in sequent frames, shot based features i.e., hard cut, fade-in, fade-out and dissolve, and object based features for classifying the video files were studied and reported thoroughly [19]. Besides, the end of 2015 another survey on video files classification on the basis of text, audio and visual was carried out [21]. In which, the techniques were categorized belongs to three classifications as following.

- Text based Approach
- Closed captions and optical character recognition techniques were used.
- Audio based approach
- Two domain based principles are introduced namely time domain and frequency domain. Root mean square and zero crossing rates are coming under time domain approach and frequency centroid, bandwidth and MFCC are coming under frequency domain approach.
- Visual based Approach
- Colour, shot and object based approaches were used.

This literature study on video files classification concludes that the effective classification could be done based on audio approaches since it requires few computational resources and better classification also possible if as many as techniques are combined [21].

b. Classification based on Frames / Objects

Generally the classification in a video file is based on either object or frames. Whenever the problem arises with respect to tracking the object in frame sequence, the features of the object are to be considered. Whereas, the features of entire frames are important if the problem comes with frames classification such as shot, instance and event detections. Initially the classification had been done over tracked objects for differentiating the particular object from other tracked objects [I6]. During 2003 one video object classification technique was proposed based on computing recurrent motion for each tracked objects. A specific feature vector called `Recurrent Motion Image' (RMI) to calculate repeated motion of object was developed. Based on their RMIs, object's type was justified because all objects do not have same RMI. This proposed technique was efficient both in terms of computational and space criteria [I6]. But the accuracy during tracking was found lacked.

To overcome the accuracy issue of existing, a multilayer hyper plane mechanism was proposed in 2005 with high accuracy in classification though the frame sequence contains difficult contents. In which, Block level classification was proposed to deal with pixel by pixel classification and pyramid boundary refining technique was introduced for obtaining pixel wise object boundary in rapid manner [9]. Meanwhile several techniques with concurrent efficiency were proposed by researchers in the field of video object classification. In 2006 an extraordinary work was proposed by Dedeoğlu Y et al in object classification area. The motivation of the work carried out by Dedeoğlu Y et al was to design a human action recognition system that could be integrated into an ordinary visual surveillance system with real-time moving object detection, classification and activity analysis capabilities. Similarly an intelligent surveillance system was proposed with instance based machine learning algorithm for object classification by Dedeoğlu Y et al. This Intelligent system covered not only classification but also human action recognition with the help of Silhouette model matching. Besides, these template matching based supervised learning was used to make classification among human,

group and vehicles. This was very first time that object silhouette was introduced for recognizing the human actions in scenes such as walking, kicking, boxing, etc, in video processing field[10]. As a research breakthrough, the chaotic series-based feature binding framework was proposed in 2012 October for objects classification. Through which the classification accuracy got improved than before and the objects were also clustered as well [14].

In the starting of 2013 classification in video frames extended to next level as events classification by researchers [15]. The hidden Markov model (HMM) was commonly used as a graphical model, but the HMM had several limitations such as the assumption of observation independence, the form of observation distribution and the Markov chain interaction. However, the video events classification was done by a hidden conditional random field (HCRF) model integrated with independent component analysis (ICA) feature functions [15]. Followed by, the video objects and scenes (events) were detected, tracked and classified through mining semantic context feature information of video frames that was proposed by Tianzhu Zhang et al. By Semantic context information mining, the intelligent video surveillance systems with object detection, tracking, classification and even abnormal event detection tasks have been carried out. For object classification, object specific context information in frame sequence was extracted. It comprises X-Y image coordinates, Area in pixels, Speed, Direction of motion, Aspect ratio and percentage occupancy of object [16].

In addition to feature extraction, the combined features started to influence in classification field. The following are two combine features techniques which increased efficiency that couldn't be attain by using single feature.

1. The combined static and spatiotemporal feature with its cooccurrence of appearances and the local part movements was proposed for classification with accuracy. In which adaptive block-based gradient intensities and histograms of oriented gradients were introduced for extracting the static appearance features and the optical-flow-based entropy values of instantaneous with short-term movements were introduced for extracting the spatiotemporal features [20]

2. The combined speed and area features were used for detecting the object and label them respectively. The support vector machine classifier was used to classify the pedestrians and vehicles present in video frames which was already labeled. This unsupervised label algorithm achieves high efficiency in objects classification [24].

Among several techniques for classification the machine learning methods such as Support Vector Machine, k-Nearest Neighbor and Bayesian Network act as efficient techniques [17]. In the end of 2015 the road lane markings which presented at each frame of video sequence also classified as dashed, solid, dashed solid, solid dashed, or double solid by the proposed Bayesian classifier based mixtures of Gaussians model [22].

Meanwhile researchers had focused on storage area side for enhancing the storing capacity of video files. As a result for solving the issue with storing surveillance video, one transcoding method with low complexity background modeling and adaptive block classification was proposed and this method was also contributed on conference video transmission via different bandwidth also [18]. For compressed videos object classification between person and vehicle, Bag of HEVC (High Efficiency Video Coding) syntax words in HEVC compressed domain was introduced. This proposed object classification has the following steps: (1) Describing each coding block within the moving object region using HEVC syntax features; (2) Constructing a codebook using a clustering method; (3) Representing each moving object using a normalized histogram of codeword from the codebook and (4) Training a binary classifier to classify the moving objects into persons and vehicles [25].

Recently in 2018 a new very large dataset has been introduced for tracing, classifying and localizing the vehicles in road. It overcomes several issues which came from existing datasets. In which Very large MIOvision Traffic Camera Dataset (MIO-TCD) was introduced for doing classification and tracing the vehicles. The multi level deep learning techniques were used for completing the classification and localization tasks. It was decided to move with MIO-TCD after done a survey of existing vehicle detection and localization datasets. The following datasets had been surveyed and their performances were noted as below.

(i) KITTI benchmark Dataset

This dataset used for several functions along with 3D object detection and 3D object tracking. It consists of video frames captured by cars traveling both in rural areas and on highways.

(ii) Cityscapes Dataset

This dataset focuses on the semantic segmentation of urban street scenes including eight different types of vehicles and it has limited to images acquired from 50 urban cities during summer daytime

(iii) Tsinghua-Tencent Traffic-Sign Dataset

It was a large traffic dataset containing 100,000 images with 30,000 traffic-sign instances. It covered various illumination and weather conditions but videos did not contain labeled vehicles.

(iv) Stanford Car Dataset

8,144 training images and 8,041 testing images were in this dataset. In this dataset all pictures were captured with good resolution and good lighting conditions. However, pictures weren't taken in a top down orientation.

(v) Comprehensive Cars Dataset:

This was a still available very large dataset and it had the same limitation as Stanford car dataset had that pictures weren't taken in a top down orientation.

(vi) BIT-Vehicle Dataset

This was most realistic dataset consists of real surveillance data. But data were captured only from top-frontal view during daytime.

(vii) Traffic and Congestions (TRANCOS) Dataset

This dataset was used to count the vehicles on congested highway by surveillance camera system. But the vehicle type was not provided.

(viii) GRAM Road-Traffic Monitoring (GRAM-RTM) Dataset

This dataset was used for tracking multi vehicles. It contains the videos recorded by surveillance cameras under different platform and conditions. Each video in this dataset approximately contains 240 different vehicle objects [26]

Table I

Techniques used for video object / frame based Classification

Sl. No	**Techniques used for Video Classification**
01.	Recurrent Motion Image (RMI)
02.	Pyramid Boundary Refining Technique
03.	Instance Based Machine Learning Algorithm (Template Matching Based Supervised Learning)
04.	Chaotic Series-Based Feature Binding Framework
05.	Hidden Conditional Random Field (HCRF) Model integrated with Independent Component Analysis (ICA)
06.	Support Vector Machine Classifier
07.	K-Nearest Neighbor
08.	Bayesian Network
09.	Bayesian Classifier Based Mixtures Of Gaussians
10.	Combined Intensity Level Distribution (ILD), Complex Wavelet Transform (CWT) and Local Binary Pattern (LBP)

This literature reviews several frame and object based classification techniques as mentioned in Table1. After completing the thorough literature it was observed that the support vector machine classifier machine learning technique was used in several classification problems because of its efficiency in accuracy in frame / object classification [24].

III. Texture Based Pattern Study for Object Tracking

Today's Surveillance systems mostly depend on video cameras. The visual data capture by video cameras need to be processed online or offline in order to expose the valuable information. Tracking an object in surveillance video frames is essential after or before the classification if visual data processing is needed. In recent days several applications came into market which evidences the importance of object tracking in captured videos such as Video Motion Detection (VMD), Intelligent Scene Monitoring (ISM), Automatic Number Plate Recognition (ANPR), Autonomous Surveillance systems (CCTV), Automotive driving and safety, Automatic check-in and check-out lanes, Video Teleconferencing, etc., [27]. In general, pedestrians and cars are mostly considered as major occurrences of moving objects. Object tracking techniques can be categorized as four as follows

1. **Tracking based on a moving object region.** This method identifies and tracks a bounding box, which is calculated for connected components of moving objects in 2D space. The method relies on properties of these bounding boxes such as size, color, centroid, shape, or velocity. This works well for moving objects in small numbers. Kalman filters are used to estimate pedestrian parameters. Merging and splitting are allowed.
2. **Tracking based on a moving object model.** A parametric 3D geometry of a moving object was proposed to solve partial occlusion problems, but the main disadvantages are that it consumes time and produces high accuracy for a small number of moving objects only.
3. **Tracking based on an active contour of a moving object.** The contour of a moving object means the boundary curves of the moving object. For example the pedestrians are selecting by the contour of their heads. This method improves the time complexity of a system, but it cannot solve the problem of partial occlusion, and if there is overlapping on two moving objects, this will cause errors in tracking
4. **Tracking based on selected features of moving objects.** Common features are selected for moving objects and tracking these features continuously on further frames. Several papers have been published with the selection of corner features for tracking the vehicles [8].

As a recent advancement in tracking area, a system had been proposed by Sachin U Sharma et al in 2017 which was able to detect the animals within 30 meters in front of the vehicle and avoid the collisions in highways [37]. In supermarkets, this technique extends the possibility of counting the products automatically integrated with the existing surveillance systems [31].

The steps involved for tracking the objects have been depicted as extracting the features such as colour, shape, motion, texture, etc., and matching the objects with one another for identifying the correct object in sequence [29]. In this literature texture based feature extraction which provides high efficiency in accuracy with patterns are studied and familiar patterning techniques have been taken for making comparison to conclude the best. Normally Feature extraction and identification of objects are carried out by several proposed patterning techniques. Local Binary Pattern(LBP), k-Means Clustering LBP, Block Based LBP(BLK-LBP), Centre Symmetric LBP(CS-LBP), Directional Local Extrema Pattern(DLEP), Local Rhombus Pattern(LRP) are the familiar texture based pattern techniques which help to extract the texture feature from the video frames. These patterning techniques will be working along with colour models such as RGB (Red Green Blue), HSV (hue, saturation, brightness) and HSL (hue, saturation, lightness /luminance) [30], [32], [33], [35], [36].

Earlier in image processing field, the Local Binary Pattern (LBP) was used to identify the facial expressions because of its ability to recognize the dynamic texture in the images [28] Later this LBP became a familiar tool for some other purposes such as boundary detection, comparison between objects, creating foreground masks, tracking the selected objects, etc, with low accuracy and high loop space utilization. Further, the Local Extrema Pattern (LEP) and Local Rhombus Pattern (LRP) were introduced by different researchers based on LBP for overcoming these issues. The Directional Local Extrema Pattern was introduced by S. Murala et al in 2012 for extracting the features in four directions (0^0,45^0,90^0 & 135^0) of centre pixel with respect to its

neighbors and comparison was made with existing BLK-LBP (Block Based LBP) and CS-LBP (Centre Symmetric LBP) [30]. In 2013 the combine LEP and RGB Colour model for feature extraction was proposed and this also was compared with LBP to prove its efficiency but in terms of space it takes more than the LBP takes [32]. But with LRP these drawbacks have been overcome and assured efficiency in accuracy has been attained. Besides online object tracking had also been proposed with a new integrated technique which has LBP as an important tool along with colour RGB model and Sobel Edge Detector. The integration was done by choquet integral method [33]. The joined colour - texture and joined colour – Local Rhombus Pattern (LRP) were proposed for patterning the features by separate two researchers in the year of 2015. The selected object was identified without computational complexity and the accuracy in avoiding the false detection and duplication was attained by the proposed joined colour - texture and joined colour – LRP respectively [35], [36]. Recently k-Means Clustering LBP was proposed by K. Shanmugapriya et al in 2017, in which the shape feature extraction was done by k-means clustering algorithm and the Local Binary Pattern was used to extract the text features from the video frames and object tracking was done by Rough Set Theory (RST) [38].

The tracking will be over with mapping the histograms of objects in subsequent video frames. There are several methods have been proposed for mapping the histograms in frame sequences. Amidst Kalman filer estimation with template matching [34], Bayesian tracking method which contains the position of the objects, velocity and the elliptical bounding [31], Particle Filtering [33] and Mean Shift keying technique [30], [32], [33], [35], [36] the Mean shift Keying algorithm is often used due to its efficiency. Several algorithms have been proposed with this mean shift key as a tracker of objects in video frames. This second part of literature reviews three patterns LBP, LEP and LRP are studied. From this literature it is observed that the texture based Local Rhombus Pattern (LRP) along with HSV color provides efficiency in accuracy for tracking with minimum space requirement than other two LBP and LEP provide.

CONCLUSION:

The objective of this survey was to give thorough illustration about past finding techniques in terms of classification and texture based object tracking in video processing. Two major areas such as classification and texture based object tracking had been gone through by this literature. Classification was divided into two subparts as video files classification and frame/object classification with respect to video processing. As a first part of this work, the classification techniques for video files were reviewed and it was observed that the effective classification could be done based on audio approach since audio feature requires few computational resources only and better classification also possible if techniques are combined with it. Besides, the techniques were reviewed for frame and object based classification under second subpart. Among existing frame/object based classification techniques, around ten techniques were noted and it was found that the support vector machine classifier was used in several classification problems because of its efficiency in classification. In second part of this work, Review of texture based patterning techniques which avoid false detecton on object tracking was done. Three texture based patterns LBP, LEP and LRP with RGB and HSV colour features were taken and their performances were analyzed meticulously. Though all three patterns avoid false detection, the LRP with HSV colour provides efficiency in accuracy. Hence, after this literature review, it is suggested that for video files classification alone, any audio based technique can be used for attaining consistency. Otherwise, the support vector machine classifier integrated with LRP-HSV pattern can be used for making the intelligent surveillance system for efficient frame/object classification and tracking.

REFERENCES

[01] A.R. Rao and R. Jain, “Knowledge representation and control in computer vision systems”, *IEEE Expert*, Volume: 3 , Issue: 1 , Spring 1988.

[02] A. Rosenfeld, “Computer vision: a source of models for biological visual processes?”, *IEEE Transactions on Biomedical Engineering*, Volume: 36 , Issue: 1 , Jan.1989.

[03] Kyu-Phil Han, Young-Sik Park, Seong-Gyu Jeon, Gwang-Choon Lee and Yeong- H, “Genre classification system of TV sound signals based on a spectrogram analysis”, *IEEE Transactions on Consumer Electronics*, Volume: 44 , Issue: 1 , Feb 1998.

[04] K. Minami, A. Akutsu, H. Hamada and Y. Tonomura, "Video handling with music and speech detection", *IEEE MultiMedia*, Volume: 5 , Issue: 3, Jul-Sep 1998.

[05] K. Sage and S. Young, "Security applications of computer vision", *IEEE Aerospace and Electronic Systems Magazine*, Volume: 14 , Issue: 4 , April 1999.

[06] Javed O., Shah M. Tracking and Object Classification for Automated Surveillance. In: Heyden A., Sparr G., Nielsen M., Johansen P. (eds) *Computer Vision - ECCV 2002*. Lecture Notes in Computer Science, vol 2353. Springer, Berlin, Heidelberg, 2002.

[07] P. Wang, R. Cai, and S.Q. Yang, "A hybrid approach to news video classification multimodal features," in Proceedings of the Joint Conference of the Fourth International Conference on *Information, Communications and Signal Processing and the Fourth Pacific Rim Conference on Multimedia*, vol. 2, 2003, pp. 787–791.

[08] Zang Q., Klette R. "Object Classification and Tracking in Video Surveillance. In: Petkov N., Westenberg M.A. (eds) *Computer Analysis of Images and Patterns. CAIP 2003*. Lecture Notes in Computer Science, vol 2756. Springer, Berlin, Heidelberg, 2003.

[09] Yi Liu, and Yuan F. Zheng, "Video Object Segmentation and Tracking Using -Learning Classification", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 15, No. 7, July 2005.

[10] Dedeoğlu Y., Töreyin B.U., Güdükbay U., Çetin A.E. "Silhouette-Based Method for Object Classification and Human Action Recognition in Video". *Computer Vision in Human-Computer Interaction. ECCV 2006*. Lecture Notes in Computer Science, vol 3979. Springer, Berlin, Heidelberg, 2006.

[11] Keansub Lee, and Daniel P. W. Ellis, "Audio-Based Semantic Concept Classification for Consumer Video", *IEEE Transactions on Audio, Speech, and Language Processing* Volume: 18, Issue: 6, Aug. 2010.

[12] 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*, Volume: 33, Issue: 3, March 2011.

[13] K. Subashini, S. Palanivel and V. Ramaligam, "Audio-Video Based Segmentation and Classification using AANN", *IJCAT*, Volume 1- Issue 2, 2012, 53-56

[14] H. Azhar and A. Amer, "Classification of surveillance video objects using chaotic series", *IET Image Processing*, Volume: 6, Issue: 7, October 2012.

[15] Xiaofeng Wang and Xiao-Ping Zhang, "An ICA Mixture Hidden Conditional Random Field Model for Video Event Classification", *IEEE Transactions on Circuits and Systems for Video Technology*, Volume: 23, Issue: 1, Jan. 2013.

[16] Tianzhu Zhang, Si Liu, Changsheng Xu and Hanqing Lu, "Mining Semantic Context Information for Intelligent Video Surveillance of Traffic Scenes", *IEEE Transactions on Industrial Informatics*, Volume: 9, Issue: 1, Feb. 2013.

[17] Insaf Setitra, "Object Classification in Videos - An Overview*", Journal of Automation and Control Engineering*, Vol. 1, No. 2, June 2013.

[18] Xianguo Zhang, Tiejun Huang, Yonghong Tian, Mingchao Geng, Siwei Ma and Wen Gao, "Fast and Efficient Transcoding Based on Low-Complexity Background Modeling and Adaptive Block Classification", *IEEE Transactions on Multimedia*, Volume: 15, Issue: 8, Dec. 2013.

[19] Nirav Bhatt, "A survey on video classification techniques", *JETIR (ISSN-2349-5162)*, Volume 2, Issue 3, March 2015.

[20] Chung-Wei Liang and Chia-Feng Juang, "Moving Object Classification Using a Combination of Static Appearance Features and Spatial and Temporal Entropy Values of Optical Flows", *IEEE Transactions on Intelligent Transportation Systems*, Volume: 16, Issue: 6, Dec. 2015.

[21] Maurício Braga de Paula and Cláudio Rosito Jung, "Automatic Detection and Classification of Road Lane Markings Using Onboard Vehicular Cameras", *IEEE Transactions on Intelligent Transportation Systems*, Volume: 16, Issue: 6, Dec. 2015.

[22] Maurício Braga de Paula and Cláudio Rosito Jung, "Automatic Detection and Classification of Road Lane Markings Using Onboard Vehicular Cameras", *IEEE Transactions on Intelligent Transportation Systems*, Volume: 16 , Issue: 6 , Dec. 2015.

[23] DongWon Yang and HyunWook Park, "A New Shape Feature for Vehicle Classification in Thermal Video Sequences", *IEEE Transactions on Circuits and Systems for Video Technology*, Volume: 26, Issue: 7, July 2016.

[24] HE Yi, SANG Nong, GAO Changxin and HAN Jun, "Online Unsupervised Learning Classification of Pedestrian and Vehicle for Video Surveillance", *Chinese Journal of Electronics*, Vol.26, No.1, Jan. 2017.

[25] Liang Zhao, Zhihai He, Wenming Cao, and Debin Zhao, "Real-Time Moving Object Segmentation and Classification From HEVC Compressed Surveillance Video", *IEEE Transactions on Circuits and Systems for Video Technology*, Volume: 28, Issue: 6, June 2018.

[26] Zhiming Luo, Frederic B.-Charron, Carl Lemaire, Janusz Konrad, Shaozi Li, Akshaya Mishra, Andrew Achkar, Justin Eichel, and Pierre-Marc Jodoin, "MIO-TCD: A new benchmark dataset for vehicle classification and localization", *IEEE Transactions on Image Processing*, Volume: 27, Issue: 10, Oct. 2018.

[27] K. Sage and S. Young, "Security applications of computer vision", *IEEE Aerospace and Electronic Systems Magazine*, Volume: 14, Issue: 4, Apr 1999, Pages. 19 – 29.

[28] Guoying Zhao and Matti Pietikainen, "Dynamic Texture Recognition Using Local Binary Patterns with an Application to Facial Expressions", *IEEE Transactions on Pattern Analysis and Machine Intelligence* , Volume: 29, Issue: 6, June 2007 , Pages. 915 – 928

[29] Aishy Amer, "Voting-Based Simultaneous Tracking of Multiple Video Objects", *IEEE Transactions on circuits and systems for video technology*, vol. 15, no. 11, November 2005

[30] "Directional local extrema patterns: a new descriptor for content based image retrieval", Subrahmanyam Murala, R. P. Maheshwari and R. Balasubramanian, *Springer - International Journal of Multimedia Information Retrieval*, Volume. 1, Issue 3, October 2012, pp 191–203

[31] Carlos R. del-Blanco, Fernando Jaureguizar, and Narciso García, "An Efficient Multiple Object Detection and Tracking Framework for Automatic Counting and Video Surveillance Applications", *IEEE Transactions on Consumer Electronics*, Vol. 58, No. 3, August 2012

[32] Subrahmanyam Murala, Q. M. Jonathan Wu, R. Balasubramanian & R. P. Maheshwari, "Joint Histogram Between Color and Local Extrema Patterns for Object Tracking", Proc. SPIE 8663, *Video Surveillance and Transportation Imaging Applications*, 86630T, 19th March 2013

[33] Maha M. Azab, Howida A. Shedeed and Ashraf S. Hussein, "New technique for online object tracking by detection in video", *IET Image Process., 2014,* Vol. 8, Iss. 12, Pages. 794–803

[34] G. Mallikarjuna Rao and Dr.Ch.Satyanarayana, "Object Tracking System Using Approximate Median Filter, Kalman Filter and Dynamic Template Matching", *I.J. Intelligent Systems and Applications*, 2014, Vol. 05, Pages. 83-89

[35] Manisha Verma and Balasubramanian Raman, "Object Tracking using Joint Histogram of Color and Local Rhombus Pattern", *IEEE - Signal and Image Processing Applications* (ICSIPA), 2015

[36] Pallavi C. Hingane and S. A. Shirsat, "Object Tracking Using Joint Color-Texture Histogram", *International Journal of Science and Research*, 2015, Volume 4, Issue 6, Pages. 2603-2606

[37] Sachin umesh sharma & dharmesh j. Shah, "A Practical Animal Detection and Collision Avoidance System Using Computer Vision Technique*", IEEE-Special section on innovations in electrical and computer engineering education*, volume 5, 2017, Pages: 347 - 358

[38] K. Shanmugapriya & R. Suja Mani Malar, "A multi-balanced hybrid optimization technique to track objects using rough set theory", *Springer Verlag - Signal, Image and Video Processing*, Volume. 11, No. 3, Year. 2017, pages 415-421

[39] Jong Sun Kim, Dong Hae Yeom, and Young Hoon Joo, "Fast and Robust Algorithm of Tracking Multiple Moving Objects for Intelligent Video Surveillance Systems", *IEEE Transactions on Consumer Electronics*, Volume: 57, Issue: 3, August 2011

[40] Shuai Zhang, Chong Wang, S. C. Chan, X. G. Wei and C. H. Ho, "New Object Detection, Tracking and Recognition Approaches for Video Surveillance over Camera Network ", *IEEE Sensors Journal*, Volume: 15, Issue: 5, May 2015

[41] Junda Zhu, Yuanwei Lao, and Yuan F. Zheng, "Object Tracking in Structured Environments for Video Surveillance Applications", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 20, No. 2, February 2010