January - February 2020
ISSN: 0193 - 4120 Page No. 3016 - 3024

Training Video Data for Semantic Content Extraction in Surveillance using Tensorflow

A. Manju, P.Valarmathie

Article Info
Volume 82
Page Number: 3016 - 3024
Publication Issue:
January-February 2020

Abstract

In recent research, Video-based applications requires techniques to modify the objects in video recordings. The field of computer vision has since quite a while ago and strives to extricate understanding from pictures and recordings sequences. Video information is universal, happening in various ordinary activities, for example, surveillance, traffic, motion pictures, sports, and so on. This enormous measure of video should be investigated and prepared effectively to remove semantic highlights. Street crime is growing in a fast rate, which has requested progressively solid and smart open moderate framework. Such capacities could profit surveillance, video analytics and visually challenged individuals. While watching a long video, people have the uncanny capacity to sidestep pointless data and focus on the imperative events. These key events can be utilized as a more elevated amount depiction or outline of a long video. Propelled by the human visual cortex, this research manages such capacities in computers utilizing neural systems. Helpful or intriguing occasions are first separated from a video and after that profound learning techniques are utilized to extricate natural language summaries for every video sequence. Past methodologies of video portrayal either have been domain specific or utilize a layout based way to deal with fill detected objects, for example, action words or activities to comprise a linguistically right sentence. This work proposes strategies to create visual outlines of long recordings, and furthermore proposes methods to train and produce textual summaries of the videos utilizing recurrent networks like Tensorflow. Fascinating fragments of long video are separated dependent on picture quality just as cinematographic and customer inclination. This tale approach will be a venturing stone for an assortment of inventive applications, for example, video recovery, programmed synopsis for visually impaired persons, automatic movie review generation, video question and answering frameworks.

Article History

Article Received: *14 March 2019*
Revised: *27 May 2019*
Accepted: *16 October 2019*
Publication: *19 January 2020*

Keywords: *Image Processig, Video Analytics, Video Semantic Substance Model.*

I. INTRODUCTION

Immense quantities of recordings are transferred regularly into sites like YouTube, Facebook, and Whatsapp from gadgets like cell phones, PCs and home surveillance cameras. With the new innovation, it is conceivable to mine visual information to get significant bits of knowledge about world. At present, extracting information from video was done physically through human perception. Current innovation use metadata or labels with recordings which are put away with recordings when the video was transferred. Picture annotation has mirrored the semantic gap between video information and unique information. Picture comment was ordered between two standard arrangements: thought based picture recovery and substance based picture recuperation. The past spotlights recovered by picture articles are elevated by thoughts, while they focusing on the lower-level visual component of the picture. Division by region expects to isolate the picture into different areas sharing ordinary characteristics. These techniques register a universal likeness among pictures in light of factual picture properties and fundamental instances of such properties are surface and shading where these techniques are seen to be strong and proficient.

Semantic comprehension of scenes remains a basic research challenge for the image and video

TEST Engineering & Management
January - February 2020
ISSN: 0193 - 4120 Page No. 3016 - 3024

recovery community. Semantic representation of multimedia data have been recognized as initial stage towards progressively successful and retrieval of data in visual media. Despite the fact that new mixed media standards, similar to, MPEG-4 and MPEG-7, give vital methodologies to control and transmission of data and related metadata, the extraction of semantic objects and tags of the moving object with the comparing metadata is out of the expansion. This impels overpowering examination endeavor towards the programmed explanation of annotation of multimedia substance.

The major difference between substance oriented and message oriented recovery a structure is that; the group effort is a key piece of the existing system. Individuals will in general use more elevated amount features like, catchphrases, idea about the content, to understand the pictures and to measure their likeness. While the segments subsequently get isolated using PC vision procedures are generally lower-level components (shading, surface, shape, spatial structure, etc). Learning portrayal and semantic annotation of multimedia substance have been recognized as fundamental stages towards the best control and recovery of visual media. Today, new multimedia benchmarks, similar to, MPEG-4 and MPEG-7, give indispensable functionalities to control and exchange the articles and related metadata. The extraction of semantic substances and comment of the substance with the related metadata, in any case, is out of the degree of these measures and still left to the substance administrator. This convinces overpowering investigation endeavor toward modified annotation of multimedia substance.

Extraction of objects from video and applying semantics can help a wide scope of employments in the photo recuperation territory, which incorporates 1] Improved picture seek through accumulated inquiry semantics; 2] Automated creation of spot and event gazetteer data that can be used; 3] Web pursuit by perceiving relevant spatial areas and time ranges for explicit watchwords; 4] Generation of photo gathering portrayals by region and moreover event/time; 5] Support for label proposals for photos (or distinctive resources) based on region and time of catch.

II. RELATED WORK

AmjadRehman et.al has shown a condition of craftsmanship which review the part extraction for soccer video rundown research. The current systems as for article acknowledgment, video rundown in light of video stream and usage of substance sources in event area have been looked into. Sound, video include extraction procedures and their mix with literary techniques were researched.

K. Karschet.al has made a technique that normally makes possible significance maps from recordings using non-parametric significance examining. The framework is pertinent to single pictures and moreover recordings. Close-by development signals were used to upgrade the deduced significance maps, while optical stream was used to ensure fleeting consistency for videos.

Suet-Peng Yong et.al has started a structure that show semantic settings for key-diagram extraction. Semantic setting of video outlines was separated and its progressive changes were watched with the goal that vital curiosities were found using a one-class classifier. Working with natural life video plots, the framework encounters picture division, feature extraction and coordinating of picture squares, and after that a co-event system of semantic names was worked to demonstrate the semantic setting inside the scene.

Y. Yildirimet.al has depicted a semantic substance extraction structure that enabled the client to address and recover articles, things and musings that are expelled along these lines. A feathery video semantic substance show subject to power was presented that utilizes spatial/common relations in occasion and thought definitions. This meta transcendentalism definition gave a wide-territory material rule with improvement standard that permit to amass reasoning for a provided domain.

AmjadAltadmri et.al has developed a framework for the Automatic Semantic Annotation of unconstrained recordings. The started framework utilizes two non-space specific layers: lower-level visual closeness coordinating, and a comment examination that uses sound judgment learning bases. Sound judgment cosmology was made by joining distinctive sorted out semantic associations.

January - February 2020
ISSN: 0193 - 4120 Page No. 3016 - 3024

N. Inoue et.al has displayed a fast most maximum a posteriori (MAP) change strategy for video semantic requesting that use Gaussian mixture model (GMM) super vectors. In this procedure, a tree-sorted out GMM was utilized to decrease the computational cost, where simply the yield probabilities of mix parts close to a data test were actually determined.

Present day computational models of visual consideration focus on base up data and it overlook view setting. Notwithstanding, perceptions in visual insight demonstrates that people use system to make conceivable item identification in normal scenes by guiding their fixation or eyes to diagnostic locales. It appears, the data of low-level highlights across the scene picture and figures out where a particular ought to be situated. The eye development of an individual demonstrates that zones picked by the best down model have a similar feeling with areas investigated by human eyewitnesses playing out a visual review undertaking for individuals. The outcome favors the recommendation of best down data from visual setting and balances the encyclical of picture locales amid the movement of article recognition. Relative data gives a pathway to efficient item discovery frameworks.

Investigation frameworks that set up together fixation instruments are connected for PC vision as they can propose methodologies for discovering pathway for object detection and recognition. This pathway can be utilized to pick a lot of conceivable hopeful areas of target questions inside a picture. Relative data gives a fundamental pathway to proficient object detection systems. These investigation frameworks are computationally increasingly costly for object recognition. It utilizes a basic awareness mechanism that does not utilize exact data about the presence of the objective. The utilization of visual saliency is insufficient for clarifying human introduction or for structuring object detection procedures.

Traditional video analytics strategies are manual and tedious. These are one-sided because of the relationship of human element. This framework presents a cloud based video analytics structure for versatile and vigorous investigation of video streams. The structure gives the expert to an administrator via robotizing the article recognition and order process from prerecorded video streams. The video streams are then recovered from distributed storage, decoded and dissected on the cloud. The system register thorough pieces of the investigation to GPU fueled servers in the cloud. Ball and face recognition are displayed as two contextual analyses for assessing the structure, with one month of information and a 15 hub cloud. The GPU empowered organization of the structure took 3 hours to perform investigation on a similar number of video streams, in this way making it something like twice as quick than the cloud arrangement without GPUs.

The structure focus on building a versatile and incredible distributed computing stage for performing mechanized examination of thousands of prerecorded video streams with high discovery and grouping exactness. The system decreases latencies in the video investigation process by utilizing GPUs mounted on PC servers in the cloud. A human administrator loses focus from video screens simply following 20 minutes. Manual examination of the recorded video streams is a costly endeavor. It isn't just tedious, yet additionally requires countless, office work spot and assets.

Because of the ascent of violations on all around the globe, video reconnaissance is winding up increasingly basic step by step. This is because of the absence of human belongings to watch this rising number of cameras physically, another computer vision algorithm to do lower and larger level tasks are being created. A creative technique comprise the most praised highlights in Histograms of Oriented Gradients(HOG), a very much upgraded hypothesis of Visual Saliency in video and the high saliency forecast display for Deep Multi-Level Network to distinguish people in video arrangements was created.

Besides, the k – Means calculation to group the HOG include vectors of the decidedly distinguished windows and decided the way pursued by an individual in the video was actualized. Histograms of Oriented Gradients highlights from progressive casings distinguish the arrangement of focuses on the picture looking like a specific individual moving in the video. The HOG highlights to prepare a modern Support Vector Machine classifier for the detection of humans in any frame. Histogram of

January - February 2020
ISSN: 0193 - 4120 Page No. 3016 - 3024

Oriented Gradients is widely utilized in recognition of individuals. Such applications utilize ground-breaking RGB cameras amid night when there is lack of light and when the pictures are not clear. Visual Saliency mapping would bring up higher force where there are humans in the image.

III. VIDEO SEMANTIC SUBSTANCE EXTRACTION FRAMEWORK

The huge progression of online recordings, video thumbnail, as the basic depiction sort of video substance, is winding up being constantly basic to influence client's examining and looking background. Regardless, standard methodology for video thumbnail choice reliably negligence to pass on fulfilling results as they disregard the semantic data (e.g., title, depiction, and question) related with the video. Thus, the picked thumbnail can't generally show the video semantics and the dynamic clicking factor isn't favored that is influenced when the recovered chronicles are associated. Therefore, a semantic substance extraction structure was produced that enables the client to address and recover articles, occasions, and thoughts ordinarily.

In the present days, for the most part video based applications are amazingly enchanting and utilized as a bit of different applications. Consequently, these applications require semantic request and extraction of media substance. Consequently, to audit the video it is principal to perceive the video events. By utilizing unrefined information and lower-level parts client can't accomplish the entire need and to understand the video altogether is progressively crucial. Video Semantic Content Meta Model was exhibited for extraction of articles, occasions and musings regularly through the as of late referenced technique.

The database comprises of the video for working with specific applications. Database was utilized to depict aggregations of videos used as a piece of this initiated work. Heretofore test data was troublesome, yet the creation of current mechanized devices has streamlined securing data.

The object extraction procedure begins with removing the item. Especially, an approach is utilized for the need of extracting and grouping of object. Objects and spatial relationship between articles are removed in every individual frames. At initial stage, have to recognize the video substance and parts to display. In picture design, input search key are the essential video units, which is removed from unprocessed video information that speak well about the video in understandable manner such as theoretical way. A semiautomatic Genetic Algorithm-based object extraction method is used for removing an object in a video.

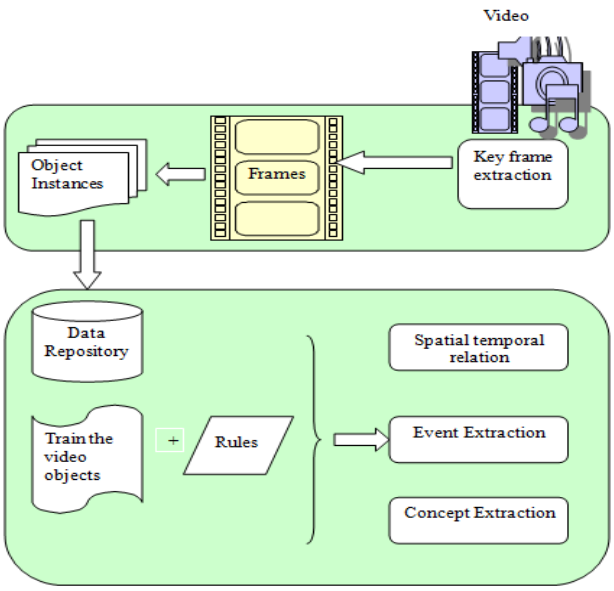

Fig: Video Semantic Substance Extraction Framework

Fuzzy Logic (FL) is a system of thinking that is as near as human thinking. The methodology of FL copies the technique for essential initiative in people that incorporates every single center plausibility between advanced qualities YES and NO. The computational yields of PC on the reason on certifiable or false, relatively individuals pursues if and else criteria.

The designer of fuzzy logic, LotfiZadeh, watched that disparate PCs, the human basic leadership fuses an extent of conceivable outcomes among YES and NO. It tends to be executed in structures with various sizes and limits reaching out from miniaturized scale controllers to vast, arranged, workstation-based control systems. The learning capacity of AND was totally controlled for modified IF-THEN standards time and parameter advancement of fluffy structure, which generally observes as the fundamental issue of fluffy system. Three obstruction structures were openly made for

TEST Engineering & Management

January - February 2020
ISSN: 0193 - 4120 Page No. 3016 - 3024

the normal moving part classes (weight lifter, football, and vehicle).

IV. TRAINING IMAGE WITH TENSORFLOW

A. Segmentation

Segmentation is the way toward creating pixel-wise segments giving the class of the object visible at every pixel. For instance, we could be recognizing the area and limits of individuals inside a picture or distinguishing cell cores from a picture. Formally, picture division alludes to the way toward apportioning a picture into a lot of pixels that we want to distinguish (our objective) and the foundation. The dataset contains a substantial number of car pictures, with every car taken from various points. Moreover, for every car picture, we have a related physically pattern cover; our undertaking will be to naturally make these pattern covers for concealed information. To prepare the picture the accompanying work process should be followed:1) Visualize information/play out some exploratory information examination. 2) Set up information pipeline and preprocessing. 3) Build display. 4) Train demonstrate. 5) Evaluate display. 6) Repeat.

B. Build Our Input Pipeline

It is important to construct a vigorous and adaptable information pipeline that will play pleasantly with our model. Our info pipeline will comprise of the following advances: First, read the bytes of the document in from the filename - for both the picture and the mark. Review that our marks are really pictures with every pixel commented on as vehicle or foundation (1, 0). Second, decipher the bytes into a picture position. Apply image transformations utilizing the resize parameter to resize our pictures to a standard size. The motivation behind why this is discretionary is that U-Net is a completely convolutional network and is along these lines not subject to the information measure. In any case, on the off chance that you decide to not resize the pictures, you should utilize a cluster size of 1, since you can't clump variable picture measure together. Then again, you could likewise basin your pictures together and resize them per smaller than normal group to abstain from resizing pictures to such an extent, as resizing may influence your execution through introduction, and so forth.

- hue_delta - Adjusts the hue of a RGB picture by an arbitrary factor. This is just connected to the actual image (not our mark picture). The hue_delta must be in the range [0, 0.5]
- horizontal_flip - flip the picture on a level plane along the focal hub with a 0.5 likelihood. This change must be connected to both the name and the genuine picture.
- width_shift_range and height_shift_range are ranges (as a small amount of all out width or stature) inside which to haphazardly decipher the picture either on a level plane or vertically. This change must be connected to both the name and the real picture.
- rescale - rescale the picture by a specific factor, for example 1/255.

At that point, mix the data, repeat the information (so we can repeat over it on various occasions crosswise over ages), group the information, at that point prefetch a bunch (for effectiveness). Note that these changes that happen in your information pipeline must be emblematic changes. Set up train and validation datasets as pursues:

```
trg_cf=
{
'resize':[img_shape[0],img_shape[1]],
'scale':1/250.,
'hue_deltavalue':0.1,
'horizontal_flipvalue':True,
'width_shift_rangevalue':0.1,
'height_shift_rangevalue':0.1
}
tr_preprocessing_fn=functools.partial(_augment,**trg_cf)
value_cfg=
{
'resize':[img_shape[0],img_shape[1]],
'scale':1/250.,
}
val_preprocessing_fn=functools.partial(_augment,**value_cfg)
```

C. Build the model

U-Net is particularly great with segmentation tasks since it can restrict well to give high goals

January - February 2020
ISSN: 0193 - 4120 Page No. 3016 - 3024

division veils. What's more, it functions admirably with little datasets and is moderately powerful against overfitting as the preparation information is regarding the quantity of patches inside a picture, which is a lot bigger than the quantity of preparing pictures itself. In contrast to the first model, we will add cluster standardization to every one of our squares.

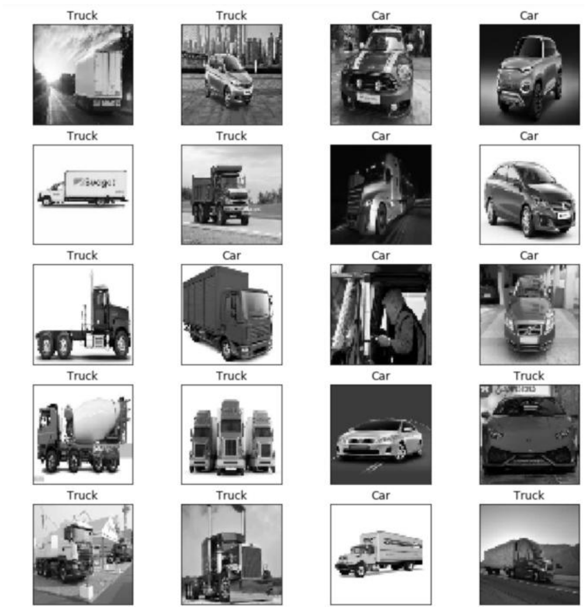

Fig: Classified Data

V. RESULT AND DISCUSSION

The proposed multi-component tracking mechanism is actualized in Python. The datasets utilized for testing the tracking system are football, auto, bar ball video arrangements. The datasets are exceptionally challenging due to the overwhelming inter-person objects and poor picture differentiate amongst components and foundation. The algorithm was accessed on its tracking execution and it was noticed that the detection execution has contrasted our outcomes and existing strategy. Multi-component tracking for the most part confronts three difficulties: component switch among overlapping, new component introduction and re-acknowledgment of re-entering objects. In the accompanying part, it quickly presents two videos and after that discuss about the outcomes as far as previously mentioned challenges.

The existing multi-component tracking mechanism was actualized in MATLAB 7.11.0(R2010b) with i5 processor and 4GB RAM. The datasets utilized for testing the tracking system are football, auto, bar ball video arrangements. The

datasets are extraordinarily demanding due to the devastating inter-person objects and poor picture differentiate amongst components and foundation.

The algorithm was accessed on its tracking execution and it was noticed that the detection execution has contrasted our outcomes and existing strategy. Multi-component tracking for the most part confronts three difficulties: component switch among overlapping, new component introduction and re-acknowledgment of re-entering objects. In the accompanying part, it quickly presents two videos and after that discuss about the outcomes as far as previously mentioned challenges.

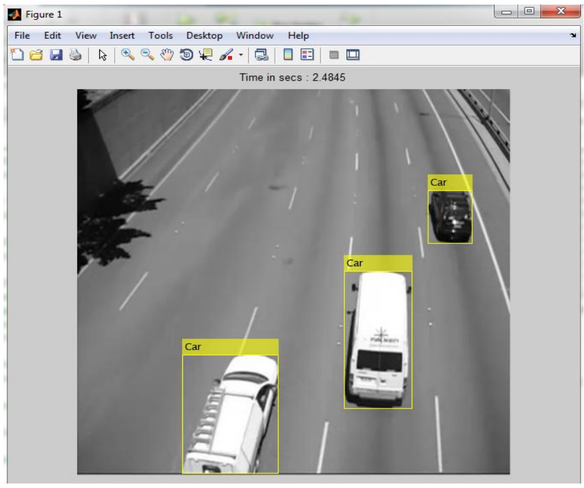

Fig: The tracking outcome on car

The proposed Video analytics for semantic content extraction is an effort to use real time, publicly available data to improve prediction of moving object from streaming video. It was actualized using Tensorflow tool using python.

Experimental Settings:

To track all items all through the benchmark successions, the initiated tracking algorithm depends on a few intuitive parameters. Specifically, the accompanying default parameter settings for our investigations were utilized.

Metrics:

The broadly acknowledged execution metrics Multiple Object Tracking Accuracy (MOTA) and Precision (MOTP) method was utilized. The precision metric MOTP assesses the arrangement of

January - February 2020
ISSN: 0193 - 4120 Page No. 3016 - 3024

genuine positive directions as for the ground truth, though the precision metric MOTA joins 3 error proportions, in particular false positives, false negatives (i.e., missed components), and identity switches. Give s_n^i be the distance between the assessed outcome and the ground truth for component i at time n and b_n the quantity of matches discovered, further then, MOTP was represented as:

$$MOTP = \frac{\sum_{i,n} s_n^i}{\sum_n b_n}$$

The distance d_n^i is really the covering between the evaluated bounding box and the ground truth. Subsequently, higher estimations of MOTP demonstrate better outcomes.

For the MOTA, let w_n be the quantity of items that exist at moment n. Let likewise k_n, ht_n and kke_n be the quantity of misses, false positives, and jumbles, separately. At that point, the metric can be gotten by,

$$MOTA = 1 - \frac{\sum_n (k_n + ht_n + kke_n)}{\sum_n w_n}$$

The precision metric correlation using initiated framework using trained video and existing neural system is represented in Table1. Precision value acquired for car picture utilizing existing technique is 0.897237while initiated has 0.945792 comparatively for Football, Bar ball pictures precision value computed utilizing existing strategy is 0.916324, 0.907507 yet initiated has 0.975404, 0.964889 individually. It can be clear from the above discussion that, the initiated technique performs powerfully and it has ability to deliver the outcomes which are near the ground truth, regardless of the identities that the components are swapped.

Precision		
Dataset	Neural (MATLAB)	Trained video (Python)
Car	0.897237	0.945792
Football	0.916324	0.975404
Bar ball	0.907507	0.964889

Table1: Performance measures of MOTP

The Accuracy metric comparison using initiated framework using trained video and existing neural system is represented in Table2. Accuracy value got for car picture utilizing existing strategy is 0.912699 while initiated has 0.963376 likewise for Football, Bar ball pictures Accuracy value figured utilizing existing technique is 0.92785, 0.915761but initiated has 0.954688, 0.970593 respectively. Our initiated technique is better in assigning tracks to the right component, without taking into account how near it really is from the right position.

Accuracy		
Dataset	Neural (MATLAB)	Trained video (Python)
Car	0.912699	0.963376
Football	0.92785	0.954688
Bar ball	0.915761	0.970593

Table2: Performance measures of MOTA

VI. CONCLUSION

In this paper, a system with trained video was processed using Tensorflow and its features to new summarization were displayed. Video semantics has been removed utilizing trained video in light of its principles. In this framework, the layers of trained video to give the flow to extraction procedures were represented. The target, recognition of moving component has been performed by two systems, the initiated system and neural system. The results have been analyzed for both of the methods, where the initiated Video Semantic Substance Extraction Framework have accomplished preferable outcome over existing procedure.

V.REFERENCES

1. A.P. Pons. Object Prefetching Using Semantic Links. ACM SIGMIS Database, 37(1):97-109, 2006.
2. AmjadAltadmri,andAmr Ahmed, “A framework for automatic semantic video annotation” Journal of Multimedia Tools and Applications, Vol.72,pp. 1167-1191 2014
3. AmjadRehman, and TanzilaSaba, “Features extraction for soccer video semantic analysis: current achievements and remaining issues” Journal of Medical Imaging and Health Informatics, Vol.41, pp. 451-461, 2014
4. Anil Aksay, Alptekin Temizel, and A. Enis Cetin. Camera tamper detection using wavelet analysis for video surveillance. In IEEE Int.

Conf. on Advanced Video and Signal Based Surveillance, AVVS 2007 , pages 558-562
5. Cavallaro, O. Steiger, and T. Ebrahimi, “Semantic video analysis for adaptive content delivery and automatic description” In the proceedings of:IEEE Transactions on Circuits and Systems for Video Technolog, Vol.15, pp.1200-1209, 2005
6. Chuanping Hu, Zheng Xu, Yunhuai Liu, Lin Mei, Lan Chen, and Xiangfeng Luo, “Semantic Link Network based Model for Organizing Multimedia Big Data” , IEEE Transactions on Emerging Topics in Computing, 2013
7. G.G. Medioni, I. Cohen, F. Bre´mond, S. Hongeng, and R. Nevatia, “Event Detection and Analysis from Video Streams,” IEEE Trans. Pattern Analysis Machine Intelligence, vol. 23, no. 8, pp. 873-889, Aug. 2001.
8. H. Yu, C. Pedrinaci, S. Dietze, and J. Domingue. Using linked data to annotate and search educational video resources for supporting distance learning. IEEE Transactions on Learning Technologies, 5(2):130-142, 2012.
9. H. Zhuge, X. Chen, X. Sun, and E. Yao. HRing: A structured P2P overlay based on harmonic series. IEEE Transactions on Parallel and Distributed Systems, 19(2):145-158, 2008.
10. H. Zhuge, X. Chen, X. Sun, and E. Yao. HRing: A structured P2P overlay based on harmonic series. IEEE Transactions on Parallel and Distributed Systems, 19(2):145-158, 2008.
11. H. Zhuge, X. Chen, X. Sun, and E. Yao. HRing: A structured P2P overlay based on harmonic series. IEEE Transactions on Parallel and Distributed Systems, 19(2):145-158, 2008.
12. H. Zhuge. Communities and Emerging Semantics in Semantic Link Network: Discovery and Learning. IEEE Transactions on Knowledge and Data Engineering, 21(6):785-799, 2009.
13. H. Zhuge. The Knowledge Grid -- Toward Cyber-Physical Society, World Scientific Publishing Co., Singapore, 2012. 2nd Edition.
14. Hakeem and M. Shah,“Multiple Agent Event Detection and Representation in Videos,” Proc. 20th Nat’l Conf. Artificial Intelligence (AAAI), pp. 89-94, 2005.
15. Haritaoglu, D. Harwood, and L. S. Davis, “W : Real-time surveillance of people and their activities,” IEEE Trans. Pattern Anal. Machine Intell., vol. 22, pp. 809–830, Aug. 2000.
16. Huitao Luo, alexandros Eleftheriadis, An interactive authoring system for video object segmentation and annotation, Elsevier, Signal processing:Image Communication, pp 559-572, 2002
17. Irena Koprinska, Sergio Carrato, Temporal video segmentation: A survey, Elsevier, Signal processing:Image Communication, pp 477-500, 2001
18. J. Steffens, E. Elagin, and H. Neven, “Person spotter-fast and robust system for human detection, tracking and recognition,” in Proc. IEEE Int. Conf. Automatic Face and Gesture Recognition, 1998, pp. 516–521.
19. Julia Vogel and BerntSchiele, “Semantic Modeling of Natural Scenes for Content-Based Image Retrieval” International Journal of Computer Vision, Vol. 72, pp. 133-157, 2007
20. K. Karsch, C. Liu, and S. B. Kang, “Depth Extraction from Video Using Non-parametric Sampling ” In the proceedings of: IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 36, pp. 2144-2158, 2014
21. K. Petridis, S. Bloehdorn, C. Saathoff,and N. Simou, “Knowledge representation and semantic annotation of multimedia content” In the proceedings of: IEE Proceedings - Vision, Image and Signal Processing, Vol.153, pp. 255-262, 2006
22. L. Wang and S. Khan. Review of performance metrics for green data centers: a taxonomy study. The journal of supercomputing, 63(3):639-656, 2013.
23. L. Wang, J. Tao, et al. G-Hadoop: MapReduce across distributed data centers for data-intensive computing. Future Generation Computer Sys-tems, 29(3):739-750, 2013.
24. L. Wu and Y. Wang. The process of criminal investigation based on grey hazy set. 2010 IEEE International Conference on System Man and Cybernetics, pp.26-28, 2010.
25. M. Petkovic and W. Jonker, “An Overview of Data Models and Query Languages for Content-Based Video Retrieval,” Proc. Int’l Conf. Advances in Infrastructure for E-Business, Science, and Education on the Internet, Aug. 2000.
26. M. Petkovic and W. Jonker,“Content-Based Video Retrieval by Integrating Spatio- Temporal and Stochastic Recognition of Events,” Proc. IEEE Int’l Workshop Detection and Recognition of Events in Video, pp. 75-82, 2001.
27. M. Steinbach, G. Karypis and V. Kumar. A Comparison of Document Clustering Techniques. KDD Workshop on Text Mining, 2000.
28. Moeslund, Thomas B, Introduction to video and image processing, Springer, 2012
29. N. Inoue and K. Shinoda, “A Fast and Accurate Video Semantic-Indexing System Using Fast MAP Adaptation and GMM Supervectors” In the

proceedings of: IEEE Transactions on Multimedia, Vol.14, pp. 1196 - 1205, 2012
30. R. C. F. Wong and C. H. C. Leung, “Automatic Semantic Annotation of Real-World Web Images ” In the proceedings of: IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 30, 2008
31. R. Firth. A synopsis of linguistic theory 1930-1955. In Studies in Lin-guistic Analysis. Philological Society: Oxford, 1957.
32. R. T. Collins, A. J. Lipton, and T. Kanade, “Introduction to the special section on video surveillance,” IEEE Trans. Pattern Anal. Machine Intell., vol. 22, pp. 745–746, Aug. 2000.
33. R. T. Collins, A. J. Lipton, T. Kanade, H. Fujiyoshi, D. Duggins, Y. Tsin, D. Tolliver, N. Enomoto, O. Hasegawa, P. Burt, and L.Wixson, “A system for video surveillance and monitoring,” Carnegie Mellon Univ., Pittsburgh, PA, Tech. Rep., CMU-RI-TR-00-12, 2000.
34. Stephan Bloehdorn, Kosmas Petridis, CarstenSaathoff, Nikos Simou, VassilisTzouvaras, YannisAvrithis, SiegfrieHandschuh, YiannisKompatsiaris, and Steffen Staab, “Semantic Annotation of Images and Videos for Multimedia Analysis ” In the proceedings of: Semantic Web Conference, Vol. 3532, pp. 592-607, 2005
35. Suet-Peng Yong, Jeremiah D. Deng, and Martin K. Purvis, “Wildlife video key-frame extraction based on novelty detection in semantic context” Journal of Multimedia Tools and Applications, Vol.62, pp.359-376, 2013
36. T. Berners-Lee, J. Hendler, and O. Lassila. “The Semantic Web”. Scientific American, 284(5):34-43, 2001.
37. TyeRattenbury, Nathaniel Good, and MorNaaman, “Towards Automatic Extraction of Event and Place Semantics from Flickr Tags” In the proceedings of:conference on Research and development in information retrieval, pp. 103-110, 2007
38. W. Liu, T. Mei, Y. Zhang and C. Che, “Multi-Task Deep Visual-Semantic Embedding for Video Thumbnail Selection” In the proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015 ,pp.3707-3715,2014
39. X. Lu, Li Song, Songyu Yu, Nam Ling,’ Object Contour Tracking Using Multi-feature Fusion based Particle Filter’, IEEE Conference on Industrial Electronics and Applications (ICIEA), pp. 237 –242, 2012.
40. X. Luo, Z. Xu, J. Yu, and X. Chen. Building Association Link Network for Semantic Link on Web Resources. IEEE transactions on automation science and engineering, 8(3):482-494, 2011.
41. Xu, Y. Zhang, G. Zhu, Y. Rui, H. Lu, and Q. Huang. Using webcast text for semantic event detection in broadcast sports video. IEEE Transactions on Multimedia, 10(7):1342-1355, 2008
42. Y. Yildirim , Ankara, Turkey, A. Yazici, and T. Yilmaz, “Automatic Semantic Content Extraction in Videos Using a Fuzzy Ontology and Rule-Based Model” In the proceedings of IEEE Transactions on Knowledge and Data Engineering, Vol.25, pp.47-61,2013
43. Ying Liua, DengshengZhanga, GuojunLua,and Wei-Ying Mab, “A survey of content-based image retrieval with high-level semantics” Journal of Magnetic Resonance in Medicine, Vol.40, pp. 262-282, 2007

AUTHORS PROFILE

A. Manju received Master degree from Anna University in 2008. She is currently working towards Ph. D degree at Saveetha University. She is currently working as Assistant Professor in Saveetha Engineering College, Chennai, India. She has 12 years of teaching experience with good programming skills. She has certified EMC Academic Associate in Data Science &BigData Analytics and also certified Programming for Everybody (Python), Python Data Structures, Using Python to Access Web Data and R Programming from Coursera. She completed the online training course Big Data Fundamentals from Big Data University.

Dr. P. Valarmathie received Doctor of Philosophy in Data Mining from Anna University in the year 2013. She is currently working as Professor in Saveetha Engineering College, Chennai, India. Her main research area includes Data Mining, Big Data Analytics, Cloud Computing, Semantic Web Services and Network Security. She has 18 years of teaching experience including 6 years of research. She has published 26 research articles in reputed journals. She has coordinated in NPTEL Certification programs and DSIR Research Recognition. She has certified EMC Academic Associate in Data Science &BigData Analytics and also certified Programming for Everybody (Python), Python Data Structures from Coursera.