Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 6, June : 2023

ZAAN EFFECTIVE CONTENT BASED VIDEO RETRIEVAL USING DEEP LEARNING

[1]**M.Ajay Kumar**, PG Scholar, Dept. of Computer Science Engineering, Malla Reddy University, Hyderabad
[2]**G.S. Naveen Kumar**, Associate Professor & HoD, Department of CSE (Data Science), Malla Reddy University, Hyderabad
[1]ajayajju126@gmail.com, [2]gsnaveen@mallareddyuniversity.ac.in

Abstract

In this research work, we analyse the challenges and objection of image to video retrieval, thatutilizes the query image to explore the relevant-frames from vast collection of the videos. A new framework depending on CNNs (convolutional neural networks) is recommended to perform the content based video retrieval with the less storage cost also with higher search capability. The recommended framework subsistsof extraction algorithm with respect to key-frame and the feature collection strategies. Particularly, the extraction algorithm of key-frame takes benefit of clustering idea; so in that case excessive information is taken out from video-data and also the storage cost is highly shortened. The feature collection scheme adapts the average-pooling to encrypt the features of deep-convolutionalpatterns followed by the pursuetofine retrieval, that admits the speedy retrieval in the content based video-database. The outcomes from the considerable experiments on the two publicly accessible datasets determine that the recommended method attains the preferable accuracy and efficiency over the other advanced visual search approaches.

Keywords – Video Retrieval, Deep Convolutional Features, Query frames,Key-Frame Extraction.

I. Introduction

Massive images &videos are induced and uploaded to the online networks and web. With a wide amount of publicly accessible data, the visual search become a vital frontier concept in field of data retrieval. Here, exist various types of visual-search tasks, including the video-to-video (V2V) search, image-to-image (I2I) searchmoreover the image-to- video (I2V) search. Specifically, the well-known I2I visual search can be used for product search, in which relevant images are retrieved by the query image,The V2V search is commonly used for copyright protection, in which video clips are found via a relevant video. The I2V search addresses the problem of retrieving relevant video frames or specific timestamps from a large database via the query image. This technology is relevant for numerous applications, such as brand monitoring, searching film using slides, and searching lecture videos using screenshots [6]. In this work, we study the specific task of I2V search, which is especially challenging because of the asymmetry between the query image and the video data. Video data can be divided into four hierarchical structures: video, scene, shot, and frame. When considering only the visual content, a video is a sequence of frames displayed at a certain rate. For example, a video with a frame rate of 30 fps is equivalent to 30 images in one second,The structure of a video means that adjacent frames are highly correlated with each other. To perform large-scale retrieval, we should select representative frames of a video frame sequence to reduce redundant information for further processes [19]. Key- frame extraction, which could represent the salient content and information of the video, is the technique employed to remove redundant or duplicate frames. In this work, we propose a cluster-based key-frame extraction algorithm to summarize the video sequences [26].

Inspired by the advances in content-based image retrieval (CBIR), we propose to take advantage of the image retrieval techniques to image-to-video search. In CBIR, one of the most challenging issues is the association of pixel-level information with human-perceived semantics. Although some hand-crafted features have been proposed to represent images, the performance of these descriptors is notNsatisfactory [12]. Recently, the CNN-based descriptors have shown excellent performance on various computer vision tasks, such as image classification, instant search, and target tracking. Encouraged by the advances in the deep convolutional neural network, our works share similarities with other CNN-based methods extracting features of the frame via pretrained CNNs.

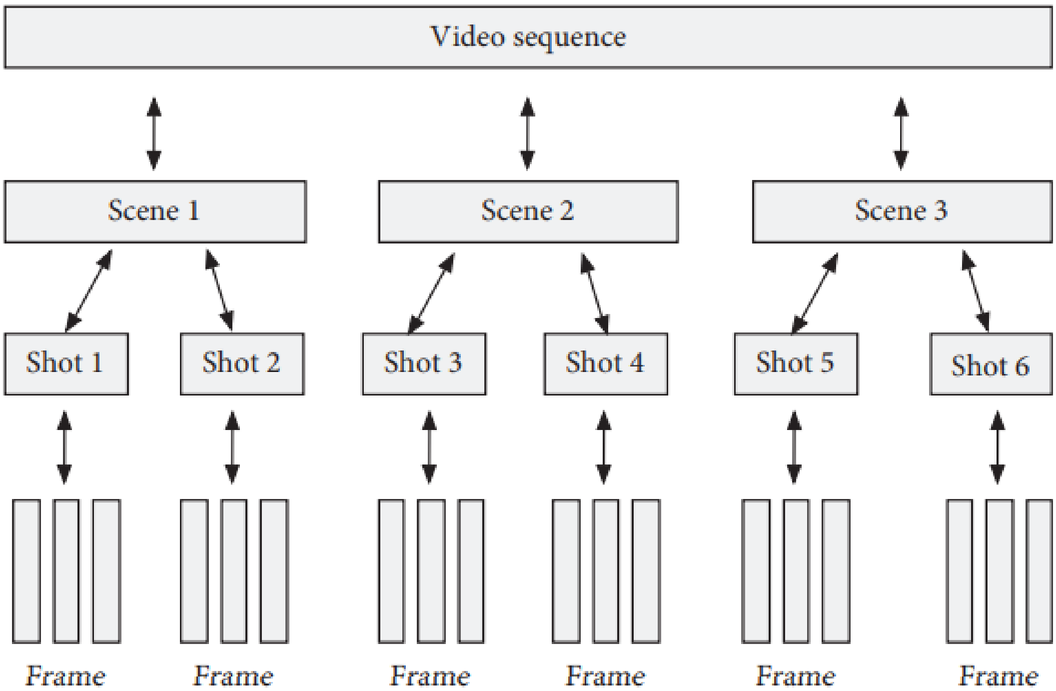

Fig 1: Structure of video data

This research work mainly presents two contributions:

i) We proposed a cluster-based key-frame extraction algorithm to remove a large amount of redundant information in the video, which could greatly reduce storage cost.

ii)We took advantage of an aggregation method based on average pooling to encode deep local convolutional features, which allows rapid retrieval in the large-scale video database.

II. Approach

Our method involves two main elements: key-frame extraction, frame representation, the first element is a key pre-process to recapitulate the video data. Subsequently, the emphasize representation of the key frame is learned by the pretrained broad convolutional neural networks. Basically, relevant frames to the objection image are retrieved afterwards feature aggregation [24]. Here, in this process of indexing and extracting the description for an image, and sign that the length of the ratio is much smaller than that of the descriptor. For large-scale retrieval assignments, it is very significant to quickly precise down the exploration using the image index. In the coarse-level search, the query image's ratio is compared to the indices of key frames (DB of the index) that are extracted from video frames to develop m candidates, Then, the caption of the query image, that contains more instruction than the index, is related to the descriptors (DB of the descriptor) of mnominees in the fine-level search using Euclidean radius [11][15]. The smaller the Euclidean distance is, the above the level of affinity of the two portraits is. Each candidate is graded in an ascending order by similarity so, top n ranked frames are preferred as the final outcome.

a) **Key-Frame Extraction**. Key-frame extraction is the support of video analysis and content-based video retrieval. As specified in the previous segment, a video is arrangement of frames displayed at a convinced rate, and adjacent frames are highly enforced with each other. Key-frame extraction adopt frames to compile the video while removing unnecessary information. In this task, we adopt the cluster-based innovation to extract representative frames. The major idea of the cluster-based design is to split the frame sequences into different clusters according to the frame features, and then the frame closest facing the cluster center would be preferred as a key frame. However, this algorithm requires a pre-specified experimental parameter, the number of clusters, that directly affects the outcome of key-frame extraction. It is very problematic to compute the number of arrays in the case where the video content is unclear. To address this problem, we propose an enhanced key-frame extraction method.

b) **Frame Representation**. Our path is similar to former effort which extracted convolutional features taken away pretrained CNNs. However, they discard the softmaxas well as fully connected layers of the initial network while keeping convolutional layers to get back local features [23]. Our work target on local features scheduled to the problem that overall descriptors may fail the invariance assumption to image changes. In this performance, we choose the suitable deep neural network extracts frame features, that was trained on the ILSVRC dataset. The structure consists of a stacked 3 × 3 convolutional kernel and max-pooling layers, tracked by three fully linked and softmax layers [4].

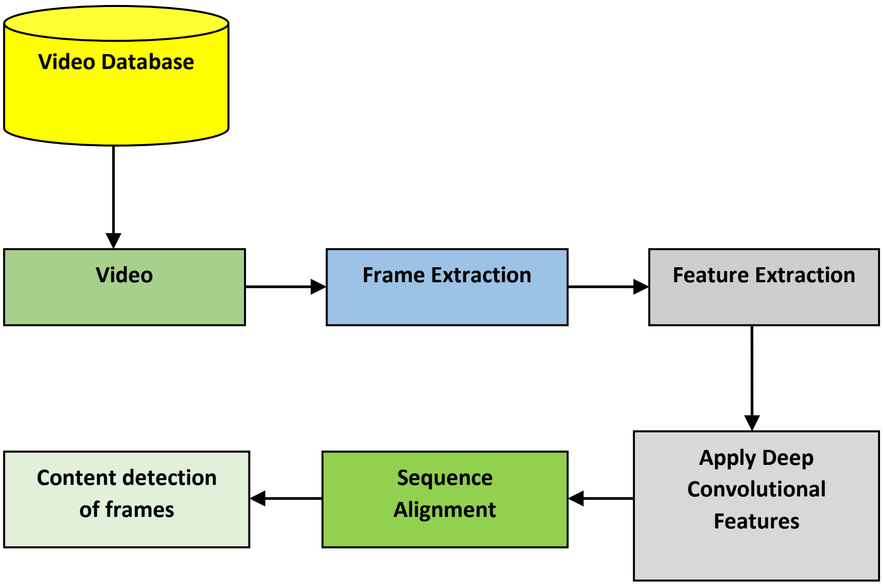

Fig 2: Flow Mechanism

III. Experiment

In this portion, we demonstrate the advantages of our method. We introduce the datasets, evaluation metrics, and parameter setting. Then, we represent our experimental outcomes with performance comparison with several existing search approaches.

a) EvaluationMetric. Query images for retrieval are captured by OpenCV, an open source library for computer vision. For evaluation, it is considered a visual match on condition that the query image and the retrieved frame are from the same video clip. Performance is measured in terms of accuracy:

$$\text{Acc} = \frac{\text{no.(visual matches)}}{\text{no.(retrieved frames)}}$$

In order to show the performance variation, we test different parameter settings for our key-frame extraction algorithm. The compression ratio is used to measure the compactness of the extracted key-frame sequence, which is defined as:

$$\text{Compression ratio} = 1 - \frac{\text{no.(key frames)}}{\text{no.(frames)}}$$

The Varience in pixel extraction of a frame is subdivided into

- Min Pixel Extraction
- Min-Max Pixel Extraction
- Max Pixel Extraction

IV. Simulation Results and Discussions

Fundamentally, the retrieval of frame from video using deep convolutional features is operated with **minimum, minimum – maximum, maximum pixel extraction**.
Here the **query image** is given as the **input image** to find the similar frame using three pixel extraction method is studied

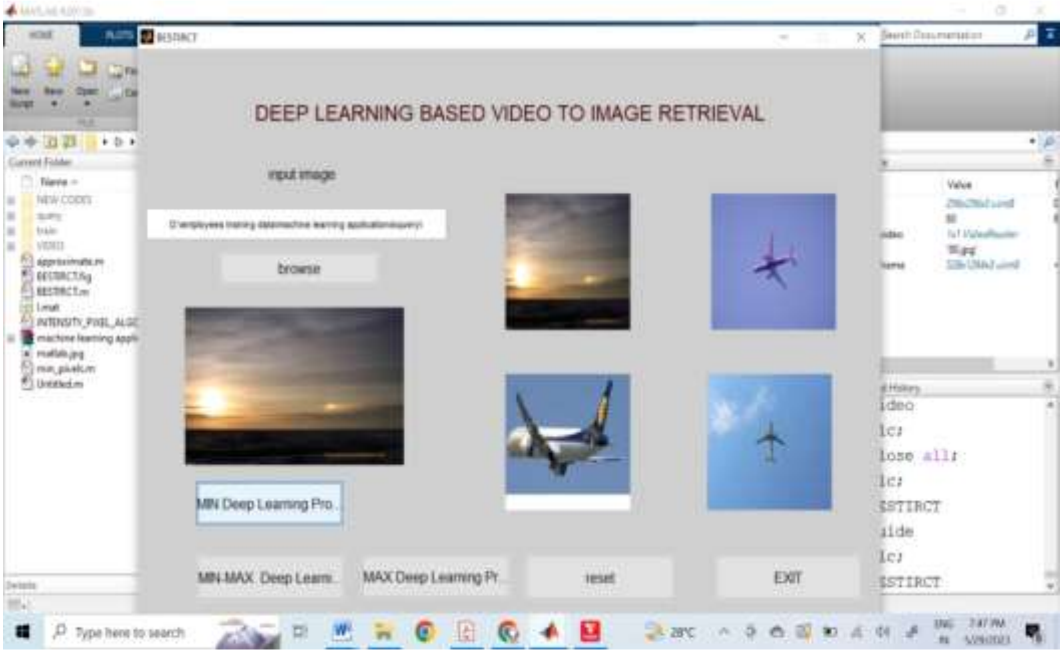

Fig 3: Min Pixel Extraction of query image

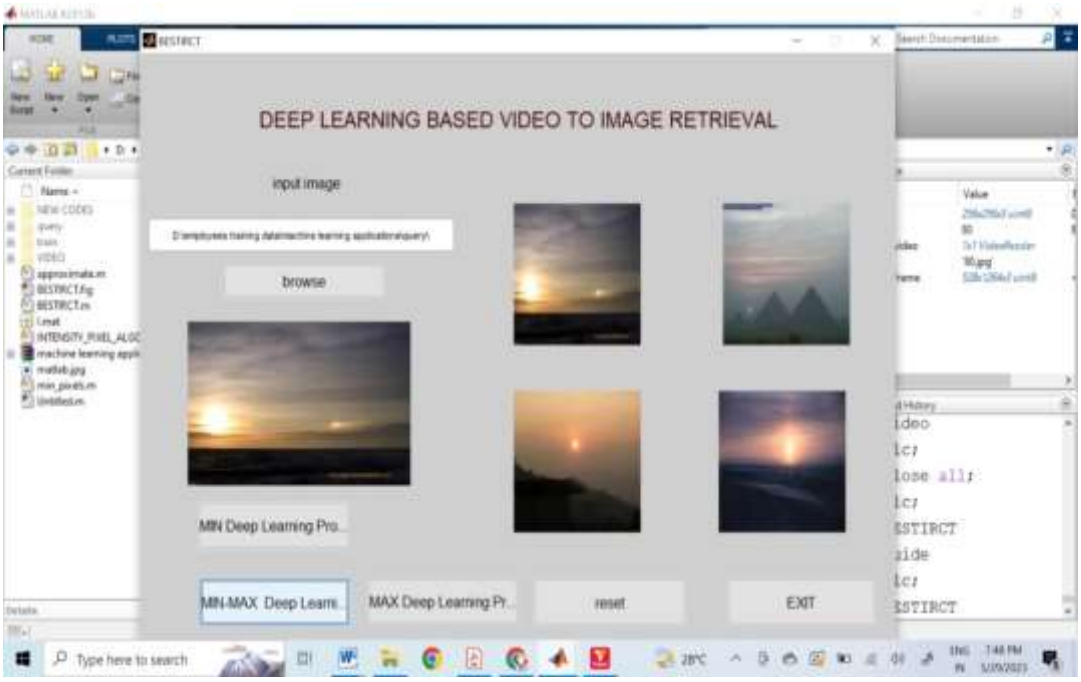

Fig 4: Min – Max Pixel Extraction of query image

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 6, June : 2023

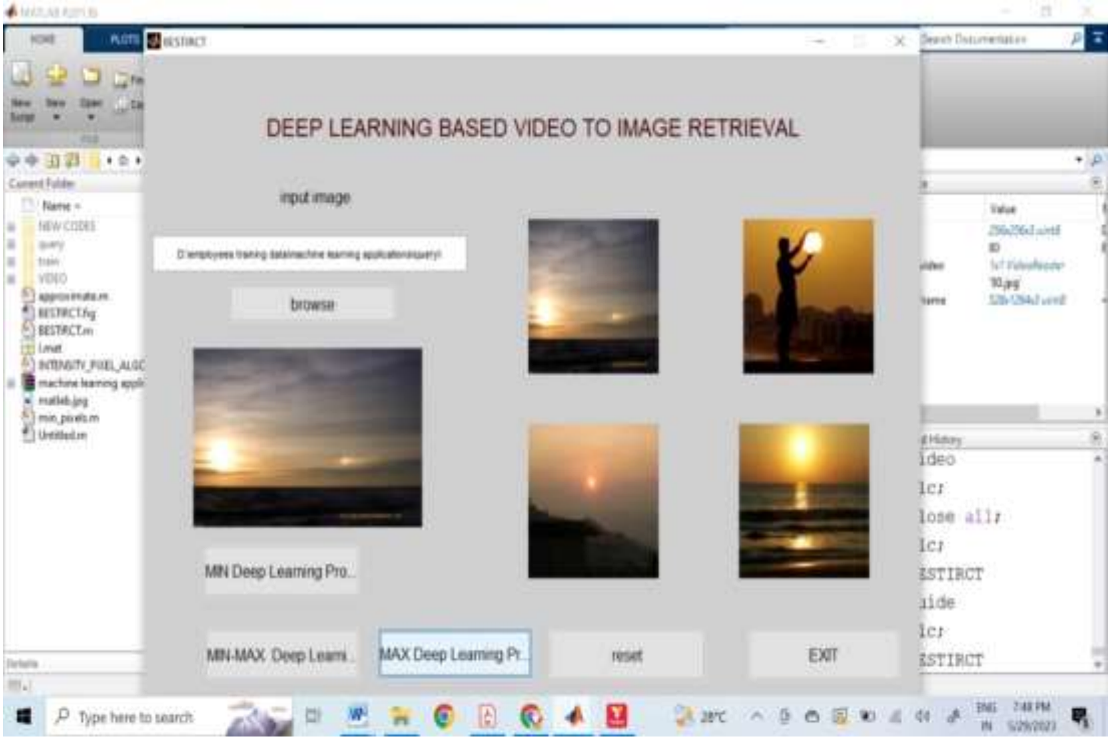

Fig 5: Max Pixel Extraction of query image

Here, using pixel methods, the similar images are obtained for a query image from the data sets, the suggested algorithm achieves the effective results in terms of accuracy and outperforms other methods by large margins.

The V2I retrieval frames are obtained of a research oriented video, https://youtu.be/VDbp-WxmYCo, using the deep learning and convolutional models.

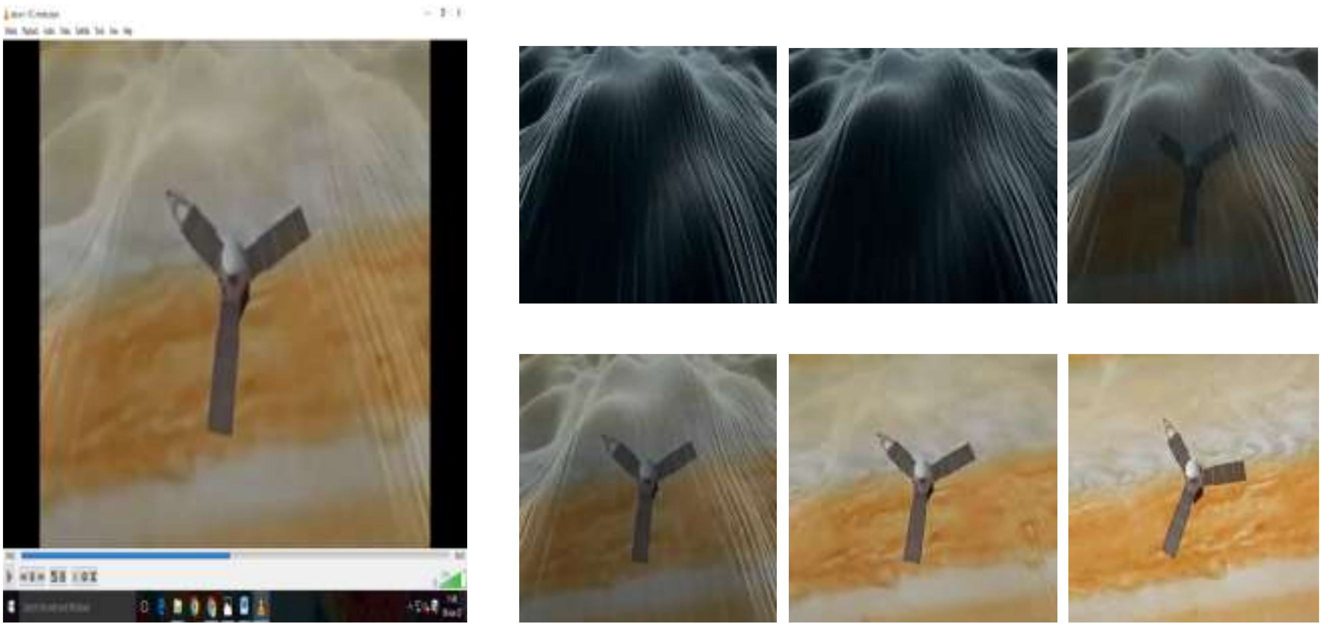

Fig 6: Extraction of Frames from video data

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 6, June : 2023

V. Conclusion

In this research paper, we suggested a model dependent on the deep convolutional features to solve the issue of the image-to-video retrieval. The models which presented in this work are based on the feature representation and key-frame extraction. The empirical results determined that our method attained competitive performance with respect to other CNN-based representations, and performed excellent in search time and cost of indexing. However, the proposed method appears to be more appropriate for tasks in which query images are from the original video frames. The quality problem of the query image caused by geometric transformations and occlusion might affect the search accuracy. In future work, we aim to explore an effective method to reduce the impact of image quality issues.

References

1) A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,”Advances in neural information processing systems, vol. 1, pp. 1097–1105, 2012.
2) M. Ajay Kumar, N.Rajesha, AnupamaDeshpande, “Cognitive Radio Network based on Energy Spotting Method with Enhanced Sensing Accuracy”, IJITEE (International Journal of Innovative Technology and Exploring Engineering), (ISSN 2278-3075) Volume-8 Issue-12 (October, 2019)pp. 4257-4261.
3) M. Ajay Kumar, Dr.N.Rajesha, “Effective Analysis on Matched Filter Technique in Cognitive Radio”, IJAER, Volume 14, Number 3 (2019) pp. 840-844.
4) Naveen Kumar, G. S., and V. S. K. Reddy."High Performance Algorithm for Content-Based Video Retrieval Using Multiple Features." Intelligent Systems and Sustainable Computing.Springer, Singapore, 2022.637-646.
5) M. Ajay Kumar, Dr.N.Rajesha, Dr.AnupamaDespande, “Attainable Throughput Determination for Cognitive Radio”, IJRAT April 2019 (ISSN-2321-9637) Volume 7, No 4, pp. 459-462.
6) M. Ajay Kumar, N.SravanGoud, R.Sreeram, R.GnanaPrasuna,, “Image Processing based on Adaptive Morphological Techniques”, 2019 International Conference on Emerging Trends in Science and Engineering (ICESE) , 2019.
7) N.Manisha Reddy, P.SaiPoojitha, M.Ajay Kumar, K.Ramya, “Reducing the Sensing Errors by Adopting the Effective Matched Filter Threshold Estimation in Lower SNR Conditions”, 2019 International Conference on Emerging Trends in Science and Engineering (ICESE) , 2019.
8)“Deeply Learned Classifiers for Age and Gender Predictions of Unfiltered Faces”, OlatunbosunAgbo-Ajala and SerestinaViriri, April 2020.
9) M. Ashok, N.Rajesha, M.Ajay Kumar, “Performance Analysis of Energy Spotting and Cyclostationary Methods in Cognitive Radio”, 2019 International Conference on Emerging Trends in Science and Engineering (ICESE) , 2019.
10) M. Ajay Kumar, N.Rajesha, AnupamaDeshpande, “Optimal Energy Resources and Sub-Carrier Allocation for OFDMA Based Dynamic Cognitive Radio”, 2019 International Conference on Emerging Trends in Science and Engineering (ICESE), 2019.
11) Kumar, G. N., & Reddy, V. S. K. (2019). Key frame extraction using rough set theory for video retrieval. In Soft Computing and Signal Processing (pp. 751-757).Springer, Singapore.
12) R. S. Vidhya, P. Ashritha, M. A. Kumar et al., "Image Noise Declining Approaches by Adopting Effective Filters", 2019 International Conference on Emerging Trends in Science and Engineering(ICESE), 2019.
13) M. Ajay Kumar, N.Rajesha, AnupamaDeshpande “Declining of Sensing Errors in Cognitive Radio by Threshold Determination of Energy Spotting Method”, IJRTE (International Journal of Recent Technology and Engineering), (ISSN 2277-3878) Volume 8, Number 2S3 (July-2019) pp. 895-898.
14). M. Ajay Kumar, Dr.N.Rajesha, “Performance Enhancement of Effective Spectrum Sensing Approaches in Cognitive Radio”, JETIR March 2019 (ISSN-2349-5162) Volume 6, Issue 3, pp.438-44.
15) GS Naveen Kumar, VSK Reddy (2020).Detection of shot boundaries and extraction of key frames for video retrieval. International Journal of Knowledge-based and Intelligent Engineering Systems, 24(1), 11-17.
16) M. Ajay Kumar, Sahebgoud H Karaddi, G Haritha, G Harish, “COMPREHENSIVE ANALYSIS OF CYCLO-STATIONARY FEATURE DETECTION TECHNIQUE FOR EFFICIENT SPECTRUM USAGE: FUTURE RESEARCH AND RECENT ADVANTAGES”, “Turkish Journal of Physiotherapy and Rehabilitation”, ISSN: 2651- 446X, Volume: 32, Issue: 3, pp. 34514-34523.

17) G Susrutha, K Mallikarjun, M Ajay Kumar, M Ashok, “Analysis on FFT and DWT Transformations in Image Processing”, 2019 International Conference on Emerging Trends in Science and Engineering (ICESE) , 2019.
18) M. AJAY KUMAR, RAJESHA.N, ANUPAMA DESHPANDE, “OPTIMIZATION OF ENERGY SPOTTING METHOD WITH ADAPTIVE-SENSING ALGORITHM IN THE COGNITIVE RADIO AT LOWER SNR RANGE USING AWGN CHANNEL” , JCR. 2020; Volume 7, Issue-14: 2578-2584, doi: 10.31838/jcr.07.14.492.
19) G. Vishali, H. Meghana, M.Ajay Kumar, N.Rajesha “An Investigation on Image Conversions and Edge Detection Patterns in Image Processing”, IEEE International Conference on Emerging Trends in Science and Engineering, held at Hyderabad on 18th& 19th September 2019 (ISBN 978-1-7281-1871-0).
20) E. Mohedano, S. Amaia, M. Kevin, M. Ferran, N. E. O’Connor, and N. Xavier Giro-i, “Bags of local convolutional features for scalable instance search,” in Proceedings of the 2016 ACM on International Conference on Multimedia Retrieval, ACM, New York City, NY, USA, April 2016.
21) M Ajay Kumar, “Automatic Prediction of Age and Gender using Artificial Neural Networks”, “ZKG International”, ISSN: 2366-1313, Volume VII, Issue II, 2022.
22) G. Csurka, R. D. Christopher, F. Lixin, W. Jutta, and B. C´edric, “Visual categorization with bags of keypoints,” in Proceedings of the Workshop on Statistical Learning in Computer Vision, Meylan, France, 2004.
23) Ganesh Miriyala, V. V. Mani, “A nonlinear modelled low-complex ADO-OFDM for Visible light communication systems”, Optik- International Journal for Light and Electron Optics (2021), https://doi.org/10.1016/j.ijleo.2021.167831.
24) Dr. M Ajay Kumar, Dr.SanthiChebiyyam, Dr. G Archana Devi “Encountering and Noticing of Malicious URLS using Deep Learning Methods”, “Industrial Engineering Journal”, ISSN: 0970-2555, Volume 52, Issue 4, April 2023.
25) J. Wang, H. Li, X. Zhang, R. Wu,” VLC-Based Indoor Positioning Algorithm Combined with OFDM and Particle Filter”, China Communications, vol. 16, no. 1, 2019, pp. 86-96.
26) Naveen Kumar, G. S., and V. S. K. Reddy."High Performance Algorithm for Content-Based Video Retrieval Using Multiple Features." Intelligent Systems and Sustainable Computing.Springer, Singapore, 2022.637-646.
27) Chen Zhang et al, “Large-Scale Video Retrieval via Deep Local Convolutional Features”, Hindawi, Volume 2020, https://doi.org/10.1155/2020/7862894.
28) K. David and H. Berndt, “6G vision and requirements: Is there any need for beyond 5G?” IEEE Vehicular Technology Magazine, vol. 13, no. 3, 2018, pp. 72–80.