Swanirman Sunirmit Publications of Research Volume 4, Issue 3-September 2024 | [2024-25]
ISSN [Online]: 2583-2654

An Improved Video Content Partition System to Automatically Split the Video Frames into Optimised Number of Clusters Using Improved Clustering Technique

Dr Geetha C Megharaj ,Mrs. SophiaS,Sangeetha A, Sonakshi S, Tejashwini V
Department of CSE, Dr.T. Thimmaiah Institute of Technology

Abstract: The improved clustering system called Robotic video frame clustering system. The proposed (RVFC) system aims to automatically split the video frames into optimal number of distinct clusters with higher similarity using clustering technique. The proposed (RVFC) system contains five stages, the Robotic video frame clustering system gets the input videos with different size from different data store. Next stage, the (RVFC) system converts video into sequence of frames using standard video to frame converter tool. In the third stage the proposed (RVFC) system preprocess the video frames by optimizing the size of the frame , grey scale conversion and reducing the noise. In the fourth stage the proposed (RVFC) system extracts the relevant features from the individual frames using Convolutional Neural Network(CNN). Following the previous stage, the fifth stage has the proposed system following the previous stage result and split the featured set of input video frames into optimal number of relative cluster using K-Means clustering technique. The proposed Robotic video frame clustering system is well suitable to robotically separate the video frames into optimal number of clusters.

Keywords: Robotic video frame clustering system(RVFC), Preprocess, relevant features, CNN, K-means clustering technique.

I INTRODUCTION

Video consists of a series of still images known as frames. Each frame captures a moment of the scene, and when played sequentially, these frames generate continuous motion forming the video. serves as a versatile tool across diverse fields. In medicine, it aids surgical procedures by providing real-time visual guidance. In education, it enhances learning through interactive visual content, Surveillance systems rely on video for security monitoring etc.Thus rapid rise of digital video content emphasizes the need for efficient methods in video analysis and processing. In managing video content, a crucial task is dividing video frames into clusters, enabling various applications like editing, summarization, and steganography. The effectiveness of these applications heavily depends on the precision and efficiency of the video content partitioning system.

The aim behind the development of an improved Video Content Partition System lies in the critical role that play in diverse domains, including computer vision, multimedia analysis, and information retrieval. With the increase of online platforms and the exponential growth of video content, there is an escalating demand for automated systems capable of efficiently organizing and processing these vast datasets. The methods for video content partitioning, often reliant on basic clustering algorithms, exhibit shortcomings in adaptability, precision, and scalability.

The current system introduces an improved Video Content Partition System designed to automatically split video frames into an optimized number of clusters. The primary objective of this system is to enhance the efficiency of video content analysis by employing a clustering technique. Many methods often face challenges in determining the optimal number of clusters and suffer from suboptimal performance in terms of precision and efficiency. The proposed system addresses these limitations through the integration of an clustering algorithm, ensuring partitioning results.

Machine learning techniques significantly enhance the effectiveness of video partition systems and clustering processes. These algorithms excel in pattern recognition, enabling the identification of similarities and differences within video frames, a crucial step in the grouping process. Machine learning facilitates feature extraction, automatically extracting information from video frames. Unsupervised learning techniques group similar video frames without the need for labelled data, contributing to the efficiency of the partitioning process. Identifying the optimal number of clusters is a challenge addressed by machine learning. With adaptive learning capabilities, these models continuously improve and refine their performance over time, particularly when exposed to a variety of video content. The real-time processing capabilities of certain algorithms make video partition systems suitable for applications that require quick and responsive video analysis.

Video frame clustering is a qualitative analysis where it refers to grouping individual frames within video based on their similarities in visual content. Video frame clustering is useful in various research field and can be employed in various applications. The unsupervised clustering method is suitable for video frame clustering as it allows the identification of inherent patterns or structures within video sequence without the need for pre-labelled data. This project aims to propose an system called Robotic Video Frame Clustering system (RVFC)thatuses an unsupervised clustering technique to robotically split the video frames into optimal number of clusters. The Robotic Video Frames

ClusteringSystem (RVFC) is a system that addresses the growing demand for efficient video contentpartitioning systems. This system introduces a structured approach to automatically segmentvideo frames into an optimal number of clusters. The RVFC system is effective system that works right from inputting the video to validating the clusters of the video frame.

The proposed RVFC system aims to convert the given input video into sequence of frames, these set of frames are pre-processed for better analysis in subsequent stage. The relevant features are extracted from pre-processed frames. The featured set of frames are separated into optimal number of clusters using an clustering technique and validated to ensure that the frames are well organized into clusters.

II RELATED WORK

Sreedhar Kumar.S et al. In [1], Have reported a system called Automatic Video Frame Separation system (AVFS). It is anIntensified System for Identifying Distinct Clusters over the Video Using Improved Semi-Supervised Clustering Technique. The paper presents an intensified system for identifying distinct clusters in video frames using an improved semi-supervised clustering technique. The AVFS system aims to automatically separate video frames into a finite number of distinct clusters with higher similarity. The system contains five stages, including video frame separation, image preprocessing, feature extraction, optimum N means, and search distinguishable centroid objects. It uses a novel approach to improve the quality of clustering algorithms.

Wenguan Wang et al. [2] have carried out a framework called Semi-Directed Video Item Division with Super-Directions. The paper gives an outline of a semi-directed video division approach in view of an original video portrayal called "super-direction." This portrayal catches reliable movement designs, comparable appearances, and close spatiotemporal connections. The strategy successfully separates target objects from complex foundations and beats other cutting edge techniques. It considers precise spread of beginning comments in the principal outline onto the leftover casings. The thickness tops based grouping calculation is utilized to produce supertrajectories, and broad trial examinations exhibit its reliably better presentation contrasted with other division strategies. Generally, the methodology is compelling in catching rich design data of recordings and precisely sectioning objective forefront objects in testing situationsAdhitya Nugraha et al. [3] Have designed a system called Determining the Senior High School Major Using Agglomerative Hierarchial Clustering Algorithm. The paper discusses the application of the Agglomerative Hierarchical Clustering technique to classify students based on their interests and skills in order to determine their high school majors. The study emphasizes the importance of tailoring high school major selection to the interests, talents, and academic skills of students. The researchers cleaned and reduced the original student data for the study and developed a prototype application for visualizing data processing using the Agglomerative Hierarchical Clustering algorithm. The paper provides valuable insights into the use of data mining techniques for educational decision-making.

Shruti Jadon et al. [4] have detailed a framework called Solo video rundown structure utilizing keyframe extraction and video skimming. The paper proposes an unaided video rundown system to address the difficulties of overseeing enormous scope video-based information. It examines the hardships of deciding and isolating significant substance from immaterial substance in recordings and presents a structure for keyframe extraction and video skimming. The creators try different things with various solo methodologies and assess their presentation utilizing the SumMe dataset. They contrast customary vision-based calculations and profound learning-based highlight extraction and presume that the profound learning-based approach performs better, particularly in unique perspective recordings. The paper's key commitments incorporate proposing another unaided system for video synopsis.

Abiodun M. Ikotun et al. [5] have detailed a framework called K-implies bunching calculations: A far reaching survey, variations examination, and advances in the time of enormous information. The paper gives an extensive survey of the K-implies grouping calculation and its variations, introducing a proposed scientific categorization of late variations and moving application regions. It frames the exploration movements including the K-implies grouping calculation and talks about the enhancements made to the standard K-implies bunching calculation revealed in the writing. The review envelops four viewpoints: methodical survey, proposed scientific categorization, inside and out investigation, and open issues with suggested future patterns. The fundamental commitments incorporate an exhaustive survey of the calculation, ID of open examination issues, and conversations on future degree. The record additionally subtleties the inquiry procedure, screening results, and the choice models for the articles.

Teng Li et al. [6]have detailed a framework called a gathering agglomerative various leveled bunching calculation in view of groups grouping strategy and the clever comparability estimation. The paper presents a Meta-Bunching Gathering plan in light of Model Choice (MCEMS) to address the test of outfit grouping. It centers around working on the quality and variety factors in gathering grouping by using a bi-weighting strategy for model choice. The MCEMS conspire includes making meta-bunches through the re-grouping of essential bunches and deciding the ideal number of groups by consolidating comparable groups and taking into account a limit. The proposed calculation is assessed utilizing recreations on datasets from the UCI storehouse, showing its prevalence over different calculations. The paper likewise presents assessments of underlying descriptors for model determination and last dendrogram age, alongside examinations with different calculations as far as deciding the quantity of ideal groups and bunching precision.

Muhammad Asim et al. [7]have carried out a framework called A Key Edge Based Video Synopsis utilizing Variety Elements. The paper proposes a vigorous and straightforward key casing extraction procedure for video outline. It contrasts the outcomes and best in class video synopsis approaches utilizing the Open Video Dataset. The proposed approach is assessed on 50 recordings and estimated by Review, Accuracy, and Fmeasure. It centers around shot limit location to separate delegate outlines for creating a video summation. The strategy similarly recognizes sharp and progressive video slices and is less delicate to clamor and glimmers.

Theresa Ullmann et al. [8] in have designed a system called Validation of cluster analysis results on validation data: A systematic framework. The paper discusses the validation of clustering results, focusing on the development of a systematic framework to summarize existing literature on validation data in cluster analysis. It distinguishes between internal and external validation techniques and emphasizes the importance of validating and replicating clustering results on a validation dataset. The paper reviews classical validation techniques such as internal and external validation, stability analysis, and visual validation, and formalizes different types of validation of clustering results on a validation dataset. It also discusses method-based validation and result-based validation, providing examples of clustering studies from the applied literature that used a validation dataset. This paper is a valuable resource for understanding and implementing cluster validation techniques in research and practical applications.

Fengsui Wang et al. [9] have implemented a system called Keyframe Generation Method via Improved Clustering and Silhouette Coefficient For Video Summarization. The paper presents a method for video summarization through keyframe generation using improved clustering and silhouette coefficient. It involves feature extraction using colour features and the Local Binary Pattern (LBP) operator, followed by obtaining initial clustering results and optimizing them. The experimental results demonstrate that the proposedmethod outperforms other state-of-the- art algorithms for video summarization. It is evaluated using two publicly available video datasets and compared with other methods for performance evaluation. The precision of the proposed algorithm is higher than several other algorithms, and the generated video summaries are more reasonable and superior to other existing methods. The method is helpful for users to quickly retrieve and understand video content. This paper contributes to the field of video summarization.

Taile Peng et al. [10]have reported a system called Video Classification Based on the Improved K-Means Clustering Algorithm. The paper introduces an enhanced K-Means clustering algorithm for video classification, focusing on the use of multi-visual features to achieve more precise results. It discusses the shortcomings of traditional video classification methods and proposes a novel approach that leverages multi-visual features for improved accuracy. The algorithm is compared to other methods, demonstrating its superior clustering results. Additionally, the paper is supported by Key Natural Science Research Projects in Colleges of Anhui, China.

III METHODOLOGY

In this section, the detail of the proposed (RVFC) system methodology is presented, The clustering system called Robotic video frame clustering (RVFC) system. The proposed (RVFC) system were to automatically split the video frames into optimal number of distinct clusters with higher similarity using clustering technique. The (RVFC) system contains five stages, in this project RVFC system gets the input videos with different size from different data store.Initially, it employs a standard video to frame converter tool to transform the video into a sequence of frames. Subsequently, the system preprocesses these frames by converting coloured frames to grey scale, reducing noise and optimizing their size. In the fourth stage, Convolutional Neural Network (CNN) is utilized to extract relevant features from individual frames. In the final stage, an K-Means clustering technique is applied to split the featured set of frames into an optimal number of clusters, making the proposed RVFC system adapt at efficiently organizing video frames for robotic applications.

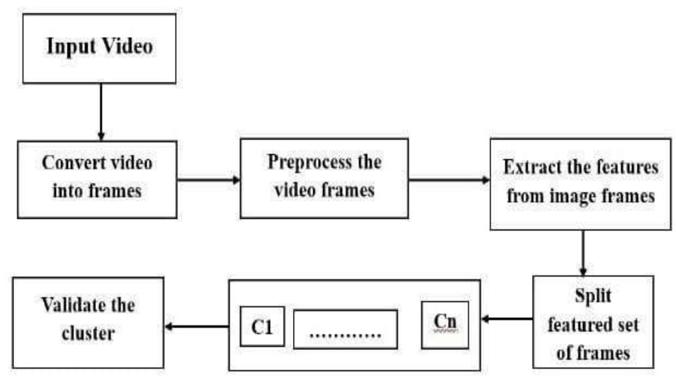


Fig.1. Proposed (RVFC) System Architecture

A data flow diagram (DFD) is realistic portrayal of the "stream" of information through a data framework. An information stream outline can likewise be utilized for the perception of information handling (organized plan). It is normal practice for a creator to draw a setting level DFD first which shows the connection between the framework and outside substances. DFD's show the progression of information from outside substances into the framework, how the information

ISSN [Online]: 2583-2654

moves starting with one interaction then onto the next, as well as its consistent stockpiling. There are just four images: Squares addressing outside substances, which are sources and objections of data entering and leaving the framework. Adjusted square shapes addressing processes, in different strategies, might be called Exercises, Activities, Systems, Subsystems and so on which accept information as info, do handling to it, and result it. Bolts addressing the information streams, which is mathematical information.

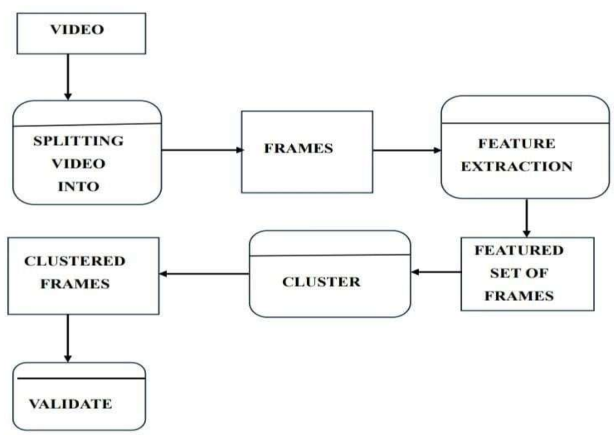


Fig. 2. Data Flow Diagram

A. Input Video

An input video is selected from the various sources such as social media, web server, camera etc. video consists of sequence of frames, Number of frames per second in a video depends on the frame rate defined for a video duration in seconds, FPS = Frame rate per second in a video depends on the frame rate defined for a video and it has been defined as X0 = Xi for, Xi =Xi_{rj} , r =0,1,2,……,h, j =0,1,2,..,w, where Xi indicates the ith reference frame from the frame deposit X with n frames, Xi_{rj} is the rth row and jth Column in the ith frame X

B. Converting video into frames

The continuous video stream is converted into a series of still images or frames. This process allows users to analyse, manipulate, or process individual frames independently, rather than dealing with the video as a continuous sequence.

Video is converted into set of images using video to jpg convertor tool which is deployed in RVFC system weighted average of the RGB channels to preserve luminance information. The resulting grayscale image represents the original in black and white, with pixel values indicating light intensity. This simplifies subsequent analysis and reduces computational complexity in certain applications. $\bar{X}$ is the input given to convert into grayscale. The Technique used here is Luminosity method. In this method, the grayscale intensity is calculated using weighted sum of its Red, Green and Blue values.

To convert $\bar{X}$ matrix into grayscale the mathematical equation is given by: $Xg = 0.2989\, R(\bar{X}) + 0.5870\, G(\bar{X}) + 0.1140\, B(\bar{X})$ Where R, G and B are the red, green and blue channels in $\bar{X}$ matrix respectively.

C. Preprocess the video frames

Preprocessing video frames is essential for improving data quality and facilitating analysis.Preprocessing is essential in machine learning, involving resizing for standardized input dimensions and grayscale conversion to simplify data. Noise reduction to enhance data quality by reducing unwanted distortions. These steps collectively optimize input data, contributing to improved model performance and generalization.

1. Resizing stage

X is the input image and X should be resized to a new size Xresized, the resizing process typically involves interpolation to determine pixel values in the resized image. X should be resized to a new width while maintaining the aspect ratio preservation. The aspect ratio is typically defined as the ratio of the width to the height of an image. The original width is denoted as W0 and the original height as H0. The aspect ratio (AR) is given by:

$$AR = W0 / H0$$

To preserve the aspect ratio when resizing, new width is denoted as Wn and to find the corresponding height Hn the mathematical formula is given by:

$$AR = W0 / H0 = Wn / Hn$$

ISSN [Online]: 2583-2654

The above equation can be rearranged as : Hn or Xresized = Wn / AR

2. *Conversion to grayscale image*

Converting an image to grayscale entails reducing it from RGB color, which includes RED, GREEN, and BLUE channels, to a single gray level. Grayscale images are simpler and require less memory compared to RGB color images. This conversion task is common in image processing. Grayscale images utilize a single channel to represent N= DXFPS Where, N= Total number of frames, D= Video

$$\overline{X_f} = \{X_{f1}, X_{f2}, \ldots\ldots\ldots, X_{fn}\}$$

Begin

1. The Convolution operation is applied to the input image X′ with a filter F by equation (8)

2.Apply the ReLU Activation function a_i to introduce non-linearity by equation (9).

3. Apply max pooling to $\bar{X}$ down sample the spatial features like shape from preprocessed video frames dimensions of the activation map a_i by equation (10).

4. Repeat the above steps with multiple filters and for multiple layers to extract the hierarchical feature

3 Reducing noise

Gaussian blur is a popular image processing technique used for smoothing and noise reduction. It involves convolving the image with a Gaussian filter, which assigns weights to surrounding pixels based on a Gaussian distribution. This results in a softened appearance, reducing high-frequency noise and enhancing overall image quality. The mathematical equation of Gaussian Blur is given by

$$G(x,y) = \frac{1}{2\pi\sigma^2}\, e^{-\frac{x^2+y^2}{2\sigma^2}}$$

Noise is often considered as random or irrelevant information. Reducing noise in the context of image involves the process of minimizing or unwanted or irrelevant variations or disturbances in the image.

The technique used in reducing noise is Gaussian Blur. image to extract the feature from the preprocessed video frame

D. Feature Extraction

In this stage, the proposed RVFC system extracts with the goal of identifying unique patterns using convolution neural network. Shape features extracted via Convolutional Neural Networks (CNNs) play a crucial role in video frame clustering because the y .= $\{X_{f1}, X_{f2}, \ldots\ldots\ldots\ldots, X_{fn}\}$ from the input intensity rather then separate color channels.

Feature Extraction Algorithm:

Input: Preprocessed video frames $X_{i(gn)}' = \{ X1_{(gn)}', X2(gn)', \ldots\ldots\ldots\ldots, Xn(gn)' \}$

Output: The featured frame vector set represents a set of images with extracted features and reduced spatial dimensions.

capture spatial patterns and structural information within each frame. feature extraction is a critical step. CNNs automatically learn relevant features from input images. The preprocessed video frames are given as input to the CNN and output of CNN featured frame vector set

featured set of frames $X_{i(gn)}' = \{ X1_{(gn)}', X2_{(gn)}', \ldots\ldots\ldots\ldots Xn(gn)' \}$. End.

Typically, this conversion involves taking a

$$X' = \max_{(m,n)\epsilon \text{pooling window}} a$$

X′ is the result of the max pooling operation at position (u,v), a_i is the ReLU activated value at position (u,v), and the max operation is taken over the pooling window. The set of frames X′ after feature extraction and pooling can be represented as X′= $\{P_1, P_2, \ldots, P_n\}$. This notation signifies that X′ consists of images obtained after applying the convolutional layer, ReLU activation, and max pooling to each image in X′. The output X′ which is given as input to the next stage. feature extraction is a critical step. CNNs automatically learn relevant features from input images.

$$z_i = \sum_{m=1}^{M} \sum_{n=1}^{N} X_i' \cdot F + b$$

Where, z_i is the result of the convolution operation , X_i' is the pixel value of the preprocessed frame , F(m,n) is the filter weight at position , b is the bias term, and M and N are the dimensionsof the filter. Later, Apply the Rectified Linear Unit (ReLU) activation function element-wise to the convolution result.

$$a_i = max(0,z_i)$$

Next step, Assuming max pooling with a pooling window of size P×Q, Apply the pooling operation for the i-th bunches in view of their component likenesses. When applied to include extricated video outlines, K-implies plans to bunch outlines with comparative spatial attributes into groups. The calculation begins by haphazardly introducing K bunch centroids, which address the focuses of the groups. These centroids are commonly addressed as vectors in a similar element space as the info information. Then, at that point, for each casing in the dataset, the calculation figures the Euclidean distance between the edge's highlights and every centroid.

Mathematically, the distance between the frames X_f and a centroid can be calculated using the Euclidean distance formula:

$$d(x, y) = \sqrt{\sum_{i=1}^{n} (y_i - x_i)^2}$$

After all samples are assigned to clusters, the algorithm updates each centroid by computing the mean of the feature values of all samples assigned to that cluster. This process iterates until the centroids no longer change significantly or a maximum number of iterations is reached.

E. Frame Clustering

In this clustering stage, the proposed RVFC framework iteratively parts the info highlighted outline vector set (Xf) into K divergent bunches in light of K-Means technique.K-implies grouping is a famous unaided AI Assuming a single convolutional layer with a filter (kernel) F and a bias term bcalculation used to parcel information into K p, the convolution operation for the airticular -th imag

Algorithm:
Input**:** Input the featured frame vector set $\overline{X_f}$={X_{f1},X_{f2},……….. ,X_{fn}}
Output**:** Optimal number of clusters, {C1, C2, ……………., Cn}

Begin
Initialize the k centroids randomly or using a predefined method (e.g., k-means++).

1. Assign each frame to the nearest centroid based on the extracted features by computing euclidean distance.
2. Update the centroids by computing the mean of the s features of all frames assigned to each centroid.
3. Repeat steps 3-4 until convergence criteria are met (e.g., no significant change in centroids or maximum number of iterations reached).
4. Return the cluster labels for each video frame.
5. Finally the Optimal number of clusters, {C1, C2, ……………., Cn} are formed .

End

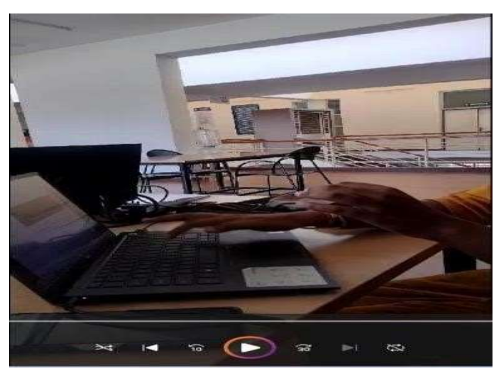
F. Validation

Video frame cluster validation is a critical process in video content analysis and clustering. This stage, the proposed RVFC system measures the quality of the clusters through by using silhouette score.The silhouette score is a metric used to assess the quality of clusters created by a clustering algorithm, such as K-means. It quantifies how similar an object is to its own cluster compared to other clusters. The silhouette score ranges from -1 to 1, where a high value indicates that the object is well matched to its own cluster and poorly matched to neighboring clusters. A score close to 1 suggests dense, well-separated clusters, while a score close to -1 indicates that the samples are assigned to the wrong clusters.

Formula for silhouette score:

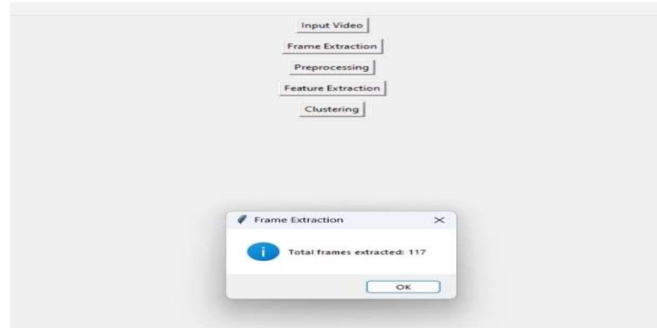
$$S(i) = \frac{b(i)-a(i)}{max(a(i),b(i))}$$

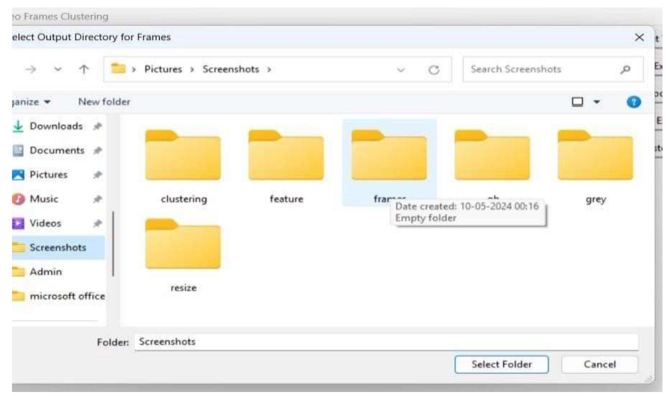
V RESULTS

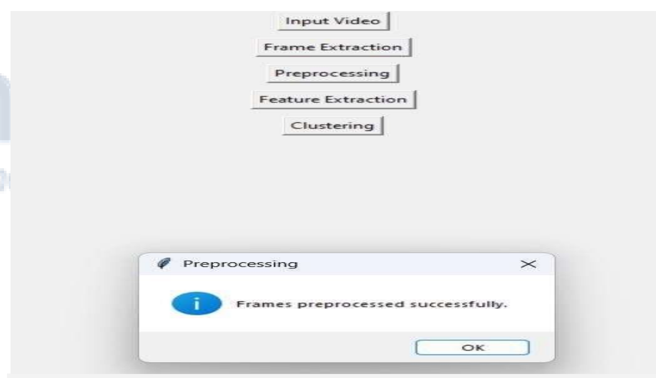


ISSN [Online]: 2583-2654

Fig 3. Input video Figure Fig 4. Extracted Frames Fig 5. Resized Frames Fig 6. Grey scale Converted Frames Fig 7. Noise removed Frames







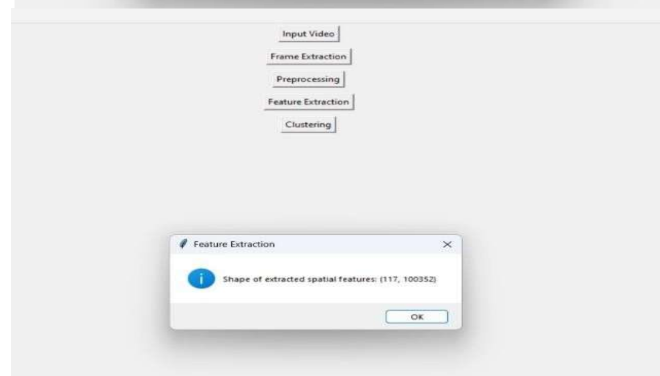
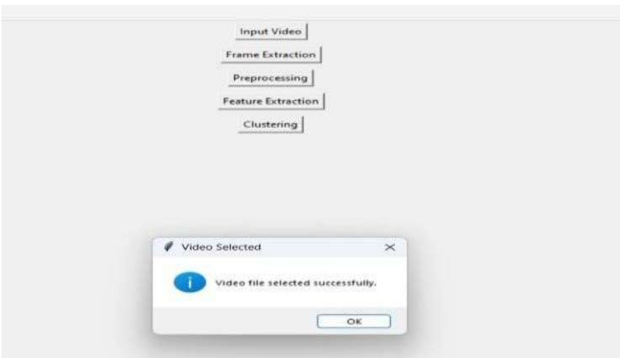
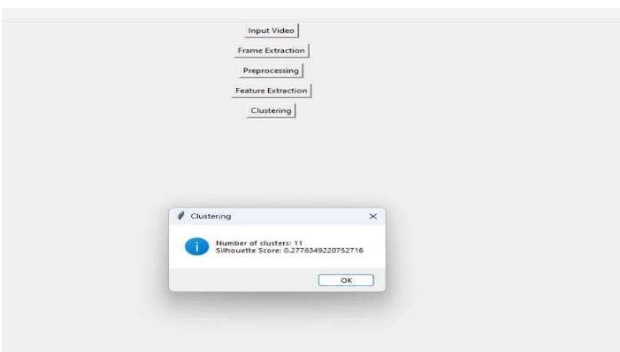


Figure 8. Dialog box for frame extraction
Figure 9. Dialog box to specify the output folder
Figure 10. Dialog box for frame preprocessing
Figure 11. Dialog box for feature extraction

ISSN [Online]: 2583-2654





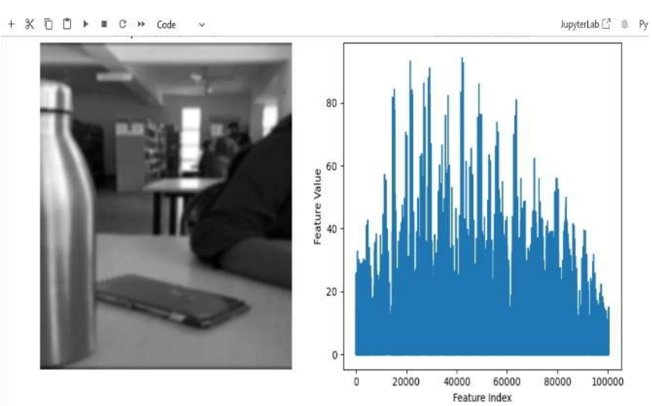


Fig 10. Dialog box for clustering
Fig 11 Dialog box for successful video selection
Fig 12. Output page
Fig 13. Feature extraction graph

VI. CONCLUSION

An Improved Video Content Partition System To Split The Video Frames Into Optimised Number Of Clusters Using Clustering Technique has come to a successful conclusion with the completion of the Introduction, Literature Survey, System Requirement Specification, System Design and Implementation. The requirements for the system have been effectively identified and the design of the system has been formulated. The literature survey conducted for the project provides a clear overview of the existing Machine Learning techniques and how they can be used to cluster the featured set of frames. With the completion of this phase of the project, the Robotic Video Frame Clustering (RVFC) system is designed to automatically categorize video frames into distinct clusters based on their similarities. It operates through five stages: obtaining input videos of various sizes, converting them into frames, preprocessing frames to optimize them, extracting relevant features using Convolutional Neural Networks (CNN), and finally, clustering frames using K-Means. This approach streamlines video analysis, particularly beneficial in robotic applications where understanding visual data autonomously is crucial. By efficiently organizing frames, RVFC enhances the efficiency of robotic systems, offering structured insights for tasks like surveillance, object recognition, etc.

REFERENCES

[1] Sreedhar Kumar.S, Velantina.V, Nishabai M, Sangeetha.V, "An Intensified System for Identifying Distinct Clusters over the Video Using Improved Semi-Supervised Clustering Technique",4th International Conference on Innovative Data Communication Technology and Application, pp 771-780, 2022.

[2] Wenguan Wang, Jianbing Shen, Fatih Porikli, Ruigang Yang, "Semi-Supervised Video Object Segmentation with Super Trajectories", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.41, no.4, pp 1-14,2019.

[3] Adhitya Nugraha, Mahendra Arista Harum Perdana, Heru Agus Santoso, Junta Zeniarja, Ardytha Luthfiarta, Ayu Pertiwi, "Determining The Senior High School Major Using Agglomerative Hierarchial clustering Algorithm", 2018 International seminar on Application For Technology of Information and Communication, pp 225-228,2018.

[4] Shruti Jadon,Mahmood Jasim,"Unsupervised video summarization framework using keyframe extraction and video skimming", 2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA), pp 140-145,2020.

[5] Abiodun M. Ikotun, Absalom E. Ezugwu, Laith Abualigah, Belal Abuhaija, Jia Heming, "K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data", information Sciences, pp 178-210,2022.

[6] Teng Li,Amin Rezaeipanah, ElSayed M.Tag El Din, "An ensemble agglomerative hierarchical clustering algorithm based on clusters clustering technique and the novel similarity measurement", Journal of King Saud University – Computer and Information Sciences, pp 3828-3842,2022.

[7] Muhammad Asim,Noor Almaadeed, Somaya Al-maadeed,Ahmed Bouridane, Azeddine Beghdadi, "A Key Frame Based Video Summarization using Colour Features", IEEE Access 2018 Colour and Visual Computing Symposium (CVCS), pp 1-6,2018.

[8] Theresa Ullmann,Christian Hennig, Anne-Laure Boulesteix, "Validation of cluster analysis results on validation data: A systematic framework", WIREs Data Mining Knowl Discovery, pp 1-19,2021.

[9] Fengsui Wang,Jingang Chen, Furong Liu, "Keyframe Generation Method via Improved Clustering and Silhouette Cofficient for video Summarization", Journal of Web Enginnering, Vol.20, no.1, pp 147-169,2021.

[10] Taile Peng,Zhen Zhang, Ke Shen, Tao Jiang, "Video classification based on the improved K-Means Clustering Algorithm", IOP Conference Series: Earth and Environmental Science, pp 1-7,2020.