**A Project Report**

**On**

# Efficient Live Video Segmentation using Machine Learning Algorithm

***Submitted in partial fulfillment of the***

***requirement for the award of the degree of***

**BACHELOR OF TECHNOLOGY**
**In**
**COMPUTER SCIENCE & ENGINEERING**

**Session 2023-24**


Submitted By:
Abhinav Kumar Choudhary (21SCSE1011615)
Kanika Yadav (21SCSE1011084)
Sakshi Srivastava (21SCSE1011610)

Under the guidance of
Dr Janarthanan S
Assistant Professor

**SCHOOL OF COMPUTER SCIENCE AND ENGINEERING**
**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

**INDIA**


**Jan, 2024**

## SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

## CANDIDATE'S DECLARATION

We hereby certify that the work which is being presented in the project, entitled "Efficient Live video segmentation using machine learning algorithm" in partial fulfillment of the requirements for the award of the B. Tech. (Computer Science and Engineering) submitted in the School of Computer Science and Engineering of Galgotias University, Greater Noida, is an original work carried out during the period of October, 2023 to December 2023, under the supervision of Prof. Janarthanan S(Assistant Professor) , Department of Computer Science and Engineering, Department of School of Computer Science and Engineering .

The matter presented in the project has not been submitted by us for the award of any other degree of this or any other places.

Abhinav Kumar Choudhary (21SCSE1011615)
Kanika Yadav (21SCSE1011084)
Sakshi Srivastava (21SCSE1011610)

This is to certify that the above statement made by the candidates is correct to the best of my knowledge.

Dr Janarthanan S
Assistant Professor

# CERTIFICATE

This is to certify that Project Report entitled "**Efficient Live video segmentation using machine learning algorithm**" which is submitted by **[Abhinav Kumar Choudhary(21SCSE1011615) Kanika Yadav (21SCSE1011084) Sakshi Shrivastav (21SCSE1011610)]** in partial fulfillment of the requirement for the award of degree B. Tech. in Department of School of Computer Science and Engineering Department of Computer Science and Engineering . India is a record of the candidate own work carried out by him/them under my supervision. The matter embodied in this thesis is original and has not been submitted for the award of any other degree

**Signature of Examiner(s)** **Signature of Supervisor(s)**

**Signature of Program Chair** **Signature of Dean**

Date: 29 Jan, 2024
Place: Greater Noida

# ACKNOWLEDGEMENT

It gives us a great sense of pleasure to present the report of the B. Tech Project undertaken during B. Tech. Third Year. We owe special debt of gratitude to Professor Dr Janarthanan S Sir (Assistant Professor) of Department of Computer Science & Engineering India for his constant support and guidance throughout the course of our work. His/Her sincerity, thoroughness and perseverance have been a constant source of inspiration for us. It is only his cognizant efforts that our endeavors have seen light of the day.

We also take the opportunity to acknowledge the contribution of Professor Dr. Janarthanan S (Assistant professor), Department of Computer Science & Engineering India for his full support and assistance during the development of the project.

We also do not like to miss the opportunity to acknowledge the contribution of all faculty members of the department for their kind assistance and cooperation during the development of our project. Last but not the least, we acknowledge our friends for their contribution in the completion of the project.

Signature:

Name :

Roll No.:

Date :

# TABLE OF CONTENT

Page No

DECLARATION.......................................................................................... 2
CERTIFICATE…........................................................................................ 3
ACKNOWLEDGEMENTS ............................................................................4
TABLE OF CONTENT..................................................................................5
ABSTRACT .................................................................................................6
CHAPTER 1
INTRODUCTION ..........................................................................................7
CHAPTER 2
LITERATURE SURVEY................................................................................ 8
2.1 Video Partition Segmentation ........................................................... 8
2.2 Video Object Detection.......................................................................8
2.3 K-Means Clustering ........................................................................... 8
CHAPTER 3
SYSTEM DESIGN AND METHODOLOGY ..............................................9
CHAPTER 4
IMPLEMENTATION AND RESULTS.................................................. 11
CHAPTER 5
CONCLUSIONS...........................................................................................13
REFERENCES .........................................................................................14

# ABSTRACT

This project states the need for real-time partition based video segmentation using a laptop camera, focus to reform applications such as object recognition and augmented reality. The Problem Statement involves the challenge of accurately identifying and defining individual objects within live video streams, requiring refined segmentation techniques. The Process uses deep learning approach, specifically the K-Means architecture, for instance video segmentation. The project consists of data collection, pre-processing, model training, and the development of a real-time video processing channel. The Output indicates the system's ability of precise segmentation and identifying individual instances, with performance metrics indicating its effectiveness. The Introduction highlights the significance of the project in changing daily business operations through advanced computer vision applications. The project successfully states the identified problem of real-time instance video segmentation. The Process indicates the extensive steps, from dataset preparation to model training and real-time implementation. Results display the project's achievements, highlighting precision, recall, and F1 score metrics. In Conclusion, the report insist that the project aim have been met, providing a groundwork for future advancements in partition video segmentation. The abstract summarizes the purpose, problem, approach, outcomes, and final assessments of the project in a short and informative manner, encouraging a change in the way video segmentation tasks are done in practical scenarios.

# CHAPTER 7

# INTRODUCTION

Segmenting and detecting the objects of interest in the video is critical for effectively analyzing and using video big data. Computer vision requires two basic tasks – segmenting and detecting video objects. The object segmentation mask is made by dividing the pixels in the video frame into two groups: foreground target and background region. In behaviour recognition and video recovery, this is the main root of the problem.

Object detecting resolves the exact location of the target in the video image and generates the object bounding box, which is needed for intelligent monitoring, big data video analysis, and other applications. The segmentation and detecting of video objects appear to be separate issues, but they are actually interconnected. Despite the many benefits and applications of video segmentation, there are also several challenges and limitations that need to be considered. Some of the key challenges and limitations of video segmentation include the following:

**Variability in video content and quality.** This can include variations in lighting, resolution, frame rate, and other factors that can affect the appearance and characteristics of the video. Different methods have been developed in past years for dealing with large variety of object appearance, including multi-scale features, deep learning-based methods, and domain adaptation techniques.

- **Lack of temporal consistency.** Videos are the sequence of frames, and the contents of the scene can change significantly from frame to frame. Which makes it very difficult to maintain consistency in the segmentation across all the frames.
- **Computational complexity.** Video segmentation are the computationally intensive, especially for the large or high-resolution video datasets. This cause so much challenges in performing real-time or online video segmentation or scaling the segmentation process to extensive video collections.

# CHAPTER 8

# LITERATURE SURVEY

**Video Instance Segmentation:** Instance segmentation are not only the grouping of pixels into different semantic classes, but it is also a grouping them into different object instances. A two-stage model is generally adopted, in which first generate object proposals using a Region Proposal Network (RPN), and then predict object bounding boxes and masks using aggregated RoI features. The proposed video instance segmentation not only requires segmenting object instances in each frame, but also used in determining the comparability of objects across all the frames.

**Video Object Detection:** Video object detection targeting objects detecting in the videos, which is first proposed as the part of ImageNet visual challenge. While object identity information are often utilized for improving robustness of detection algorithms, the evaluation metric is limited to per-frame detection and does not require joint object detection and tracking.

**K-Means Clustering:** An unsupervised machine learning technique called K-means clustering is widely used in many fields, including image and video processing. One of the most important tasks in computer vision is video segmentation, which is dividing a video into areas or segments that make sense based on features like texture, motion, or color. In this technique, K-means clustering is a fundamental tool.

The main focus of our project is to minimize the difference between the ground truth and the hypotheses. In other words, a good video instance segmentation method should be able to have a good detection rate of all instances, track all the instances certainly and locate the instance boundaries accurately. It should be noted that there will be some minor difference between our task and the multi-object tracking problem in that a still object instance is treated as a ground truth, and if an object is stopped or out of scene for several frames then repeat in the following frames, the instance label should be consistent.

# CHAPTER 9

# SYSTEM DESIGN AND METHODOLOGY

The system design surrounded in the architecture and components are required for the instance video segmentation. The chosen procedure involves a deep learning approach, involving a pre-trained neural network for video segmentation. The K-Means clustering architecture was selected for its effectiveness in instance segmentation tasks. The methodology includes the following steps:

**Data Collection**: Collecting a dataset for training and testing the model, including annotated video sequences with labeled instances.

**Pre-processing**: Arranging the dataset by resizing, normalizing, and augmenting the video to enhance the model's robustness.

**Model Training**: The training of data are done by updating the centroid of this model and fine-tuning it on the dataset to changing it into the specific instance video segmentation task.

**Real-time Video Processing**: Executing the real-time video processing pipeline using the laptop camera feed, integrating the trained model for performing the instance segmentation.

**Performance Matrix:** Performance matrices, also known as evaluation metrics or assessment criteria, are essential tools used to quantify the effectiveness, accuracy, and efficiency of machine learning algorithms, models, or systems.

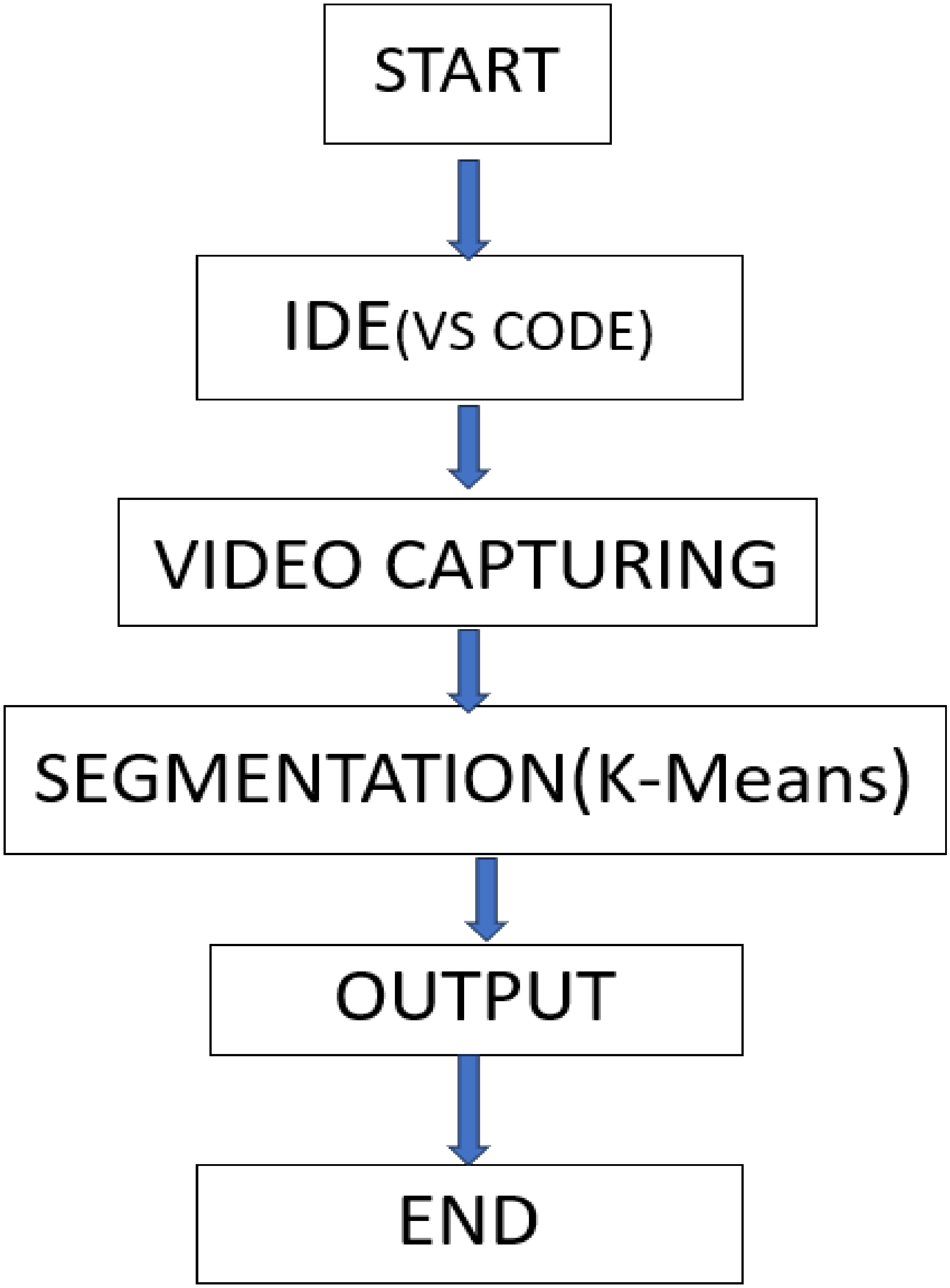
START
IDE(VS CODE)
VIDEO CAPTURING
SEGMENTATION(K-Means)
OUTPUT
END

Figure. 1

# CHAPTER 4

# IMPLEMENTATION AND RESULTS

The implementation phase involves coding the designed system, integrating the necessary libraries and its frameworks, and optimizing it for real-time performance. The main achievement of this project is its ability to perform instance segmentation in real-time. Instance segmentation involves identifying and defining each objects within an image, and the code extends its capability for live video streams.

Every video frame can be thought of as a collection of pixels or features when it comes to video segmentation. These attributes can be used to group similar pixels together using K-means clustering, which efficiently segments the video into regions of interest. The segmentation task and the video's properties influence the features that are selected. Color intensities, texture features, and even motion vectors are examples of common features.

The executed project holds significant potential for various applications, including object recognition, tracking, and augmented reality. Real-time instance segmentation is a critical component in computer vision systems deployed in scenarios where dynamic object identification is necessary.

The project's flexibility to live video streams from a laptop camera makes it versatile for different environments and use scenarios. It increases the possibilities for applications such as video editing, content creation, and interactive systems.

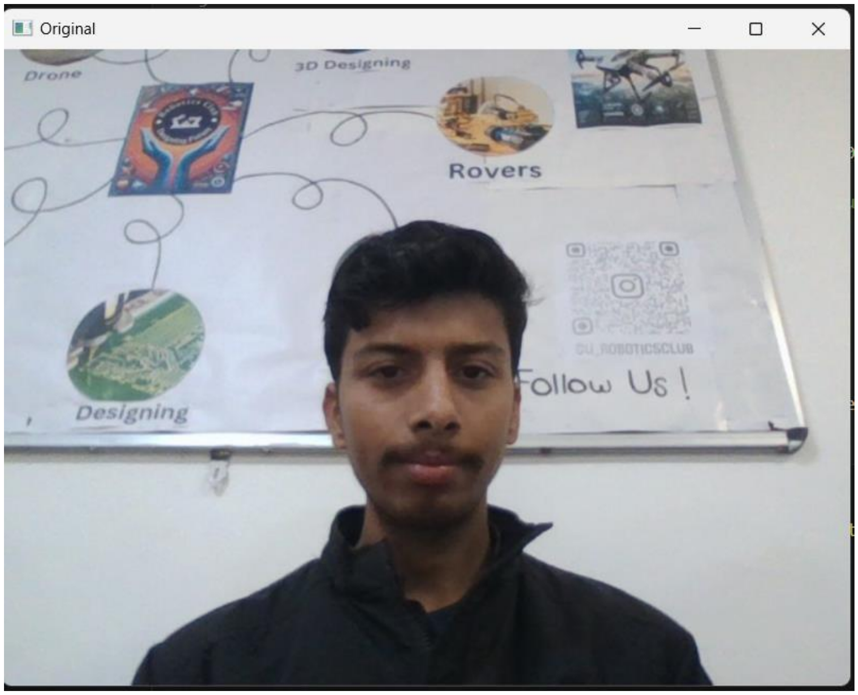


Figure 2 Before Segmentation

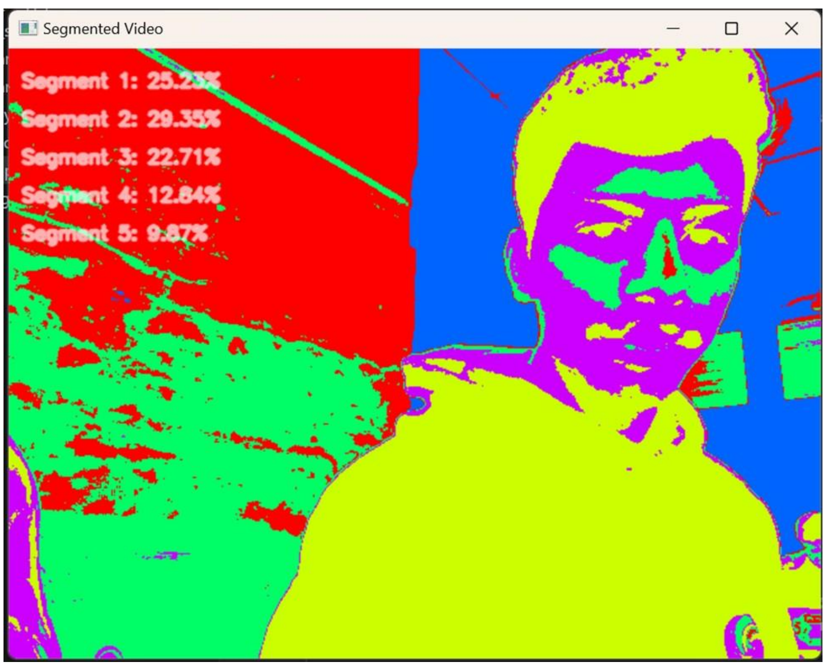


Figure 3 After Segmentation

# CHAPTER 5

# CONCLUSION

The live video segmentation system using Machine learning demonstrates promising results, showcasing the potential for real-time applications in object recognition, tracking, and augmented reality. The choice of the K-Means architecture, combined with a well-prepared dataset, increases the system's accuracy. Further some more improvements may include optimizing for real-time performance, enhancing the model's generalization capabilities, and exploring applications in diverse scenarios. In conclusion, the implemented system lays the foundation for advancements in instance live video segmentation, with implications for a wide range of computer vision applications. Future work may focus on refining the model, exploring additional datasets, and extending the system's capabilities to address specific use cases.

# References


[1] Image Segmentation Using Deep Learning: A Survey , Shervin Minaee, Yuri Boykov, Fatih Porikli, Antonio Plaza, Nasser Kehtarnavaz, and Demetri Terzopoulos (2001). https://arxiv.org/abs/2001.05566v1

[2] Real time and video segmentation using machine learning : Vighnesh Tendolkar , Pradeep Muthaye, Aniket Parab, Prof. Shashank Tolye (2022) .

[3] Deep learning for video object segmentation: Mingqi Gao, Feng Zheng, James J. Q. Yu, Caifeng Shan, Guiguang Ding, Jungong Han (2016)

[4] A Study Analysis on the Different Image Segmentation Techniques ,Sujata Saini and Komal Arora. International Journal of Information & Computation Technology. ISSN 0974-2239 Volume 4, Number 14 (2014), pp. 1445-1452

[5] Yasuda, K., Naemura, T., Harashima, H.: Thermo-key: human region segmentation from video. Computer Graphics and Applications, 26–30 (January-February 2004).

[6] Rother, C., Kolmogorov, V., Blake, A.: Grabcut: interactive foreground extraction using iterated graph cuts. In: SIGGRAPH, vol. 23, pp. 309–314 (2004).

[7] Kolmogorov, V., Criminisi, A., Blake, A., Cross, G., Rother, C.: Bi-layer segmentation of binocular stereo video. In: Computer Vision and Pattern Recognition, pp. 407–414 (2005).

[8] DUDA R O, HART P E. Pattern classification and scene analysis[M].New York:John Wiley & Sons,1973.

[9] Youguo Li, Haiyan Wu Department of Computer Science Xinyang Agriculture College Xinyang, China : A Clustering Method Based on K-Means Algorithm: International Conference on Solid State Devices and Materials Science 2012