

Application of K-Means Clustering Algorithm for Video Looping Pattern Identification

Salman Alfarisi Rizwana, Bayu Rohid Habeahan
Politeknik Negeri Medan, Indonesia

Email: salmanalfarisirizwana@gmail.com, bayurohidh@gmail.com

ABSTRACT

Video conferencing has become an important communication tool, especially after the COVID-19 pandemic. However, the use of looping techniques in videos can pose challenges regarding the authenticity of the content. This research aims to develop an automated method for detecting video looping patterns using the K-Means Clustering algorithm. The method involves extracting frames from a video, followed by clustering analysis to identify identical frames. The results showed that from the analysis of 600 frames, the method successfully identified 454 patterns with 544 identical frames, covering 90.67% of the total frames, and achieved a final Confidence Score of 88.4%. Meanwhile, for non-looping videos, only 37.17% identical frames were detected, with a Confidence Score of 40.3%. In conclusion, the K-Means Clustering algorithm proved effective in detecting looping videos, improving the integrity of video communication. This research emphasizes the importance of video authenticity in virtual interaction and recommends the application of detection technologies to improve the quality of video communication. In addition, further research is needed to improve the detection accuracy and explore the application of this method in other multimedia contexts.

Keywords: video conference; looping detection; k-means clustering; frame analysis



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International](https://creativecommons.org/licenses/by-sa/4.0/)

INTRODUCTION

In today's digital age, video conferencing has become an essential communication tool, especially in sectors such as education, business, and government (Haleem et al., 2022; Ninye-Ranor et al., 2022). With the increased adoption of video conferencing, especially after the COVID-19 pandemic, various platforms such as Zoom, Microsoft Teams, and Google Meet have seen a rapid increase in usage (Choukaier, 2024; Nadire & Daniel, 2021; Tochukwu & Nonyelum, 2024).

While this technology has provided flexibility and convenience to its users, there are challenges related to the security and authenticity of video data delivered in conferences (Jang-Jaccard et al., 2016; Mohsin et al., 2019; Patel et al., 2016). One of the

main challenges is the potential misuse of video by using video looping techniques to dishonestly simulate attendance or engagement.

This problem has a significant impact, especially in situations that require live attendance and active interaction (Emahiser et al., 2021; Kaysi et al., 2023; Mentzer et al., 2024). Manual approaches to identifying looping patterns in videos are ineffective, especially due to the large number of participants and time constraints. Therefore, there is a need for an automated method that can accurately detect looping to efficiently ensure video authenticity (Agnisarman et al., 2019; Wang et al., 2021).

The manual approach in identifying looping patterns is not efficient, especially considering the time constraints and the number of participants in a video conference. Therefore, this research proposes the application of K-Means Clustering as the main method to detect video looping (Barcellos et al., 2015; Cai et al., 2020; Lei et al., 2016; Widodo, 2021). K-Means Clustering allows grouping of frames in a video based on visual similarity, so that repetitive frames can be easily detected in the same group.

Using this clustering technique, each frame in the video is analyzed and grouped based on its visual similarity. This process helps identify suspicious repetitive patterns and can indicate the presence of video looping. The advantage of the K-Means Clustering method is its ability to work automatically in detecting and grouping frames that have similar visual patterns, even if small changes or distortions occur in the frames. This approach is expected to provide a fast and accurate solution in detecting video looping in video conferencing, while improving the efficiency and effectiveness of monitoring the authenticity quality of webcam videos (Sarmah et al., 2022; Ukhov et al., 2021).

Previous research has shown various approaches in laboratory management, such as web-based management systems that improve efficiency and data security. However, many existing systems have not been integrated with student portals or features that allow for more granular inventory management (Ghiani et al., 2022). The urgency of this research lies in the need for a more comprehensive solution to address existing problems, especially in the context of digitizing asset management.

The development of a website-based inventory lending system that is directly integrated with the MyITS Portal, so that students do not need to create a new account to access the tool lending service. This system is also equipped with a digitalization feature of laboratory exemption letters that allows automatic verification of the completeness of tool returns.

Through the use of this method, this research aims to provide an effective and efficient tool for automatically detecting video looping, which in turn can help improve the integrity and authenticity of communication in video conferencing.

The benefits of this research include increased efficiency in inventory management, reduced risk of tool loss, and reduced administrative errors. In addition, this system provides convenience for students to track the status of borrowed equipment and obtain a digital laboratory release letter. Thus, this research not only contributes to better laboratory management, but also supports digital transformation in higher education in Indonesia.

RESEARCH METHOD

This research uses a quantitative approach with an experimental design to detect video looping in video conferencing. This research focuses on clustering analysis to detect identical frames in looped videos. The research design involves extracting frames from the looped video and then performing clustering analysis using K-Means algorithm to group the frames.

The population in this study is a video conference that contains recordings of video sessions from various participants. The sample used in this study is a 60-second video that has undergone looping to simulate the presence of participants in a video conference. The data collection technique is done by doing the same simulation by recording a video whose initial duration is 15 seconds, looped into 60 seconds.

The looped video will be compared with a video recorded for the full 60 minutes, without any looping editing. The video will then be extracted frames at regular intervals, and then further analyzed.

In this study the authors used several methods including:

Data mining is the process of using certain approaches to find patterns or certain information in a set of data. Data mining or knowledge discovery in databases (KDD) is a process that involves collecting and analyzing historical data to reveal systems, patterns, and relationships in large data sets. The results of data mining processing can be used to improve the ability to make decisions in the future. Data mining in this research is customized to collect sample frames from videos that have been looped or not. The frames are collected which will be executed in the next method.

Clustering is a technique to differentiate a data set into many groups based on a desired match. Clustering in data mining is a collection of data or objects in clusters (groups) and makes each cluster have data that is almost similar to the original and can be distinguished from objects in other clusters. In this research, I implemented this method by grouping the frames in the extracted video into the same cluster.

Clustering is necessary because the unprocessed data is quite large and not easy to analyze or study manually. The purpose of clustering in this scenario is to better understand the data and analyze the quality of the data. The most widely utilized clustering method in scientific and industrial research is the K-Means algorithm. Each cluster is represented by the mean of the data in the cluster.

A method that can be utilized to cluster data is KMeans. Clustering techniques are widely used, and their popularity soars with the increasing amount of data available. K-Means is a simple method for clustering analysis that aims to determine the best way to divide the nth entity into groups called clusters.

K-Means algorithm is a method used to categorize data into groups based on distance, criteria, conditions, or certain characteristics. Data in a group is expected to have the shortest distance and to have similarities or similarities in criteria, conditions, or

characteristics. Thus, the K-Means algorithm allows the grouping of objects that have similarities between them.

K-Means is a heuristic algorithm used to divide data into K groups by minimizing the total squared distance in each group. This method is also one of the non-hierarchical approaches in data clustering that aims to partition data into one or more clusters or groups. In this method, data that has similar or similar characteristics will be grouped together in one cluster, while data with different characteristics will be grouped into different groups.

The instruments used in this research are image processing software, namely OpenCV for frame extraction from video and K-Means clustering from the scikit-learn library to group frames based on their visual similarity. Each extracted video frame is converted into a vector representation, then clustering is performed to identify groups of frames that have similarities. Data analysis techniques were performed by comparing frames belonging to the same cluster to detect if there were identical frames, indicating looping.

In this study, the researcher acts as an operator who manages the video data and frame processing, and evaluates the results of the clustering analysis. The validity of the research results was obtained by comparing the detection results of identical frames appearing in cluster groups with frames that actually have visual similarities due to the looping process. This research took place over several weeks to obtain valid results from various video samples analyzed.

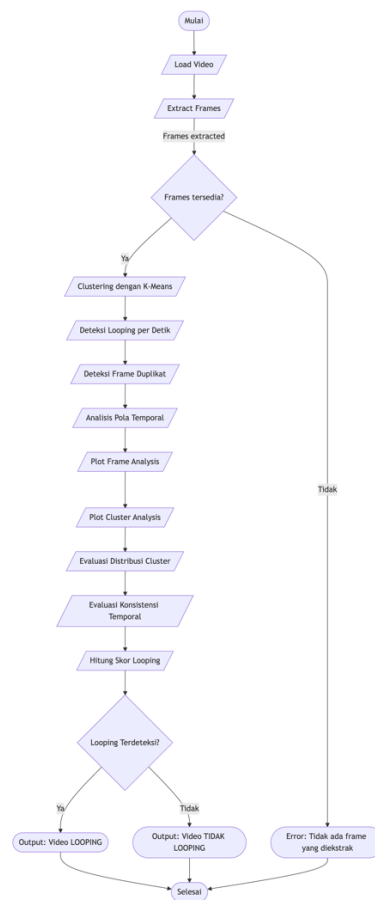


Figure 1. Program Flowchart

The videos analyzed were both looped and unlooped (normal) videos with a duration of 60 seconds. Each frame was extracted individually, resulting in 600 frames used for analysis.

1. Pattern Detection Method

a. Frame Intensity

In this step, the frame intensity is calculated for each frame by taking the average of the pixel values in grayscale format. This process helps detect patterns of visual changes in the video and identify frames with similar characteristics based on their brightness levels. This Mean Intensity becomes the base feature for further analysis.

b. Frame Clustering

Video frames are clustered using the K-Means algorithm based on visual features such as average intensity, color histogram, or edges. By grouping similar frames into similar clusters, this process makes it easier to analyze overall visual patterns and reduce data complexity.

c. Looping Pattern Detection

Video segments that have repetition patterns are further analyzed. Using statistics such as average distance between repetitions and temporal consistency, significant looping patterns can be identified. These patterns are usually sequences of frames or clusters that repeat in the video.

d. Duplicate Frame Detection

At this stage, identical or almost identical frames in the video are identified. The similarity between frames is calculated using methods such as cosine similarity, and frames with high similarity are considered as duplicates. The detection of these duplicate frames provides an indication of significant content repetition, which is often part of a looping pattern.

2. Looping Evaluation

The training data that has been built will be tested which includes the accuracy value and F1 Score value. The accuracy value is obtained from the correct prediction for positive and negative data from all data. F1 Score is a value that indicates if the model built has good precision and recall values. To obtain the accuracy value, precision value, recall value, and f1 score can use Confusion Matrix. [6]

a. Accuracy

Accuracy is a ratio that has a prediction of the correct value (positive value and negative value) based on the entire data. Accuracy can describe the accuracy of the classification model used. The accuracy value can be obtained using the following equation. [6]

$$\text{accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad [6]$$

Description:

TP = True Positive

TN = True Negative

FP = False Positive

FN = False

b. Precision

Precision is the ratio that has a positive true value when compared to the overall positive predicted result. Precision can describe the accuracy of the desired data with the prediction results obtained by the classification model. The precision value can be obtained using the following equation.

$$\text{precision} = \frac{TP}{TP+FP}$$

c. Recall

Recall is the ratio of true positive predictions compared to the overall true positive data. Recall describes the results of the classification model used in retrieving information. The recall value can be obtained using the following equation.

$$\text{recall} = \frac{TP}{TP+FN}$$

d. F1-Score

F1 Score is the ratio of true positive predictions compared to the overall true positive data. Recall describes the results of the classification model used in retrieving information. The recall value is obtained with the following equation.

$$\text{F1 Score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

e. Confidence Score

Confidence score in a clustering algorithm is a metric or value that indicates how confident an algorithm is in grouping data into specific clusters. It helps in evaluating the quality of the clustering and gives an indication of how well the data matches the resulting clusters. Here is the equation to calculate the confidence score.

$$\text{Confidence Score} = 2 - \frac{\text{Jarak ke Centroid}}{\text{Jarak Maksimum ke Centroid}}$$

RESULTS AND DISCUSSION

We experimented with two different types of videos for comparison: one video that has been edited and looped, and one video that has not been looped. By comparing the analysis results of the two videos, we were able to evaluate the effectiveness of the looping detection method applied. There are several indicators that reference the identification conclusion in this research, namely: confidence score, identical frame score, unique cluster, cluster transition, number of loop patterns, temporal pattern score.

Non-looping (Normal) Video Analysis

1. Video Analysis Based on Intensity and Cluster (Normal)

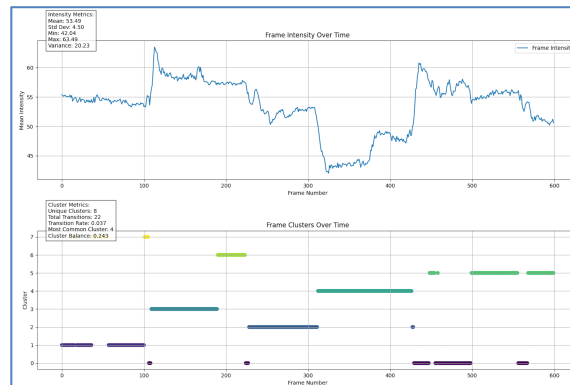


Figure 2. Intensity and Cluster Analysis Chart (Normal)

Top Chart Analysis

Analysis of the "Frame Intensity Over Time" graph revealed fluctuations in the average intensity of the video frames over time. Initial observations showed a period of intensity stability at the beginning of the video, followed by a significant sharp increase around the 100th frame. Subsequently, there is a period of high intensity fluctuation up to the 400th frame, and then a significant drop around the 300th frame.

After this period, intensity tends to stabilize with minimal fluctuation. Statistical analysis of intensity showed a mean value of 53.49, with a standard deviation of 4.50 and a variance of 20.23, indicating a relatively controlled level of variability.

Bottom Graph Analysis

Meanwhile, the "Frame Clusters Over Time" graph visualizes the grouping of video frames into 8 different clusters, with cluster 4 being the most dominant cluster. Transitions between clusters are observed to occur over time, especially around frames 100, 300, and 450, which coincide with the significant changes observed in the intensity graph. Statistical analysis of the clusters showed a total of 32 transitions with a transition rate of 0.057 and a cluster balance of 0.212.

Repeated intensity changes and cluster transitions indicate the presence of an organized structure in the video, even if a looping pattern is not explicitly observed. These two graphs simultaneously provide a visual representation of scene changes and visual conditions, which can be utilized for automated video content analysis, identification of important segments, and anomaly detection. Further research is needed to examine the correlation between intensity changes and cluster transitions with more specific video content contexts.

2. Output Analysis Based on Cluster Transition Graph Every Time

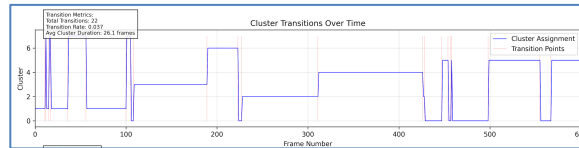


Figure 3. Cluster Transition Over Time Graph (normal)

Chart Analysis

The "Cluster Transitions Over Time" graph visualizes how the cluster of video frames changes over time in a video that is not looping. The horizontal axis shows the sequence of video frames from start to finish, while the vertical axis represents different clusters, from 0 to 7, that group the frames based on their visual characteristics.

The blue line on the graph indicates the active cluster in each frame, and the vertical red line marks the transition points between clusters. From this graph, we can see that the video undergoes some significant visual changes, as indicated by the transitions between clusters. The early part of the video shows higher dynamics with more frequent cluster transitions, indicating the possibility of a rapid change in scene or visual conditions at the beginning of the video. Cluster 4 appears to be the dominant cluster, especially in the latter part of the video, suggesting visual stability in that part. Periods of cluster stability, such as in cluster 2 around frames 150-250 and cluster 4 around frames 450-600, indicate the presence of video segments with consistent visual characteristics.

The transition statistics, 32 transitions with a transition rate of 0.057 and an average cluster duration of 26.5 frames, provide a quantitative overview of the frequency and duration of visual changes in the video. With no looping, this graph directly represents the visual progression from the beginning to the end of the video, which can be used to understand the narrative flow and scene changes in the video.

The main difference in the interpretation of this graph, knowing that the video is not looping, is that we can see the flow of visual change in a linear fashion from the beginning to the end. If the video is looping, we need to consider that the end of the video will return to the beginning, which will affect how we interpret transitions and cluster stability. In the case of non-looping videos, we can assume that each visual change seen on the graph is unique and occurs in the chronological order of the video.

3. Output Analysis Based on Frame Distribution Graph between Clusters and Entropy Graph

a. Frame Distribution Chart Between Clusters (Normal)

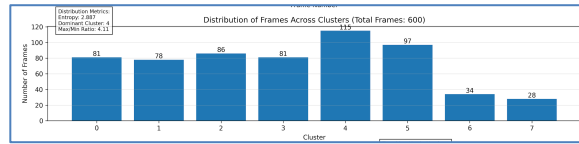


Figure 4. Frame Distribution Chart Between Clusters (Normal)

Graph Analysis

The frame distribution graph of this non-looping video presents an interesting picture of how the frames are distributed among the eight different clusters. This distribution is not entirely even, but shows significant variation between clusters, indicating differences in the duration and visual characteristics of different video segments.

Clusters 0 to 3, with a relatively balanced number of frames (ranging from 78 to 86), suggest that the initial part of the video has a similar duration and may have consistent visual characteristics. Meanwhile, cluster 4 stands out with the highest number of frames (115), suggesting that the segment of the video it represents has a longer duration and is most likely an important or climactic part of the video. The presence of cluster 5 with an equally significant number of frames (97) indicates the presence of another segment that is also quite long and may have an important role in the video narrative. In contrast, clusters 6 and 7 have a much smaller number of frames (34 and 28), indicating that these segments are shorter and may serve as transitions between scenes or as the final part of the video. The cluster distribution entropy value of 2.887 indicates a high degree of visual diversity in the video, suggesting a variety of different content and scenes. The maximum/minimum ratio of 4.11 confirms the imbalance in frame distribution, with cluster 4 having more than four times the duration of cluster 7.

b. Visual Entropy Graph

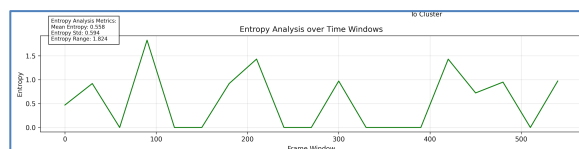


Figure 5. Entropy Visual Analysis Graph

Graph Analysis

The resulting visual entropy graph shows significant fluctuations in entropy values, indicating variations in visual diversity between frame windows. The observed range of entropy values, from 0 to 1.824, reflects the presence of video segments with very low to very high levels of visual diversity.

The average entropy value of 0.556 gives an overview of the moderate level of visual diversity in the entire video. A detailed analysis of the graph reveals some important patterns: (1) the presence of several significant entropy peaks, especially in the frame window around 80, indicating video segments with very high visual diversity; (2) the presence of entropy valleys that reach a value of 0, indicating video segments with very low or homogeneous visual diversity; (3) and the recurring pattern of entropy increases and decreases, suggesting scene changes or visual transitions in the video.

Higher entropy peaks are assumed to represent video segments with more complex details or movements, while entropy valleys are assumed to represent segments with uniform backgrounds or minimal visual changes. This pattern indicates that the video has a clear narrative structure, with distinct segments and varying degrees of visual diversity. These findings support the hypothesis that temporal entropy analysis can provide valuable insights into the structure and visual content of videos, and can be used for video segmentation, video summarization or scene change detection.

4. Output Analysis Based on Identical Frame Graph

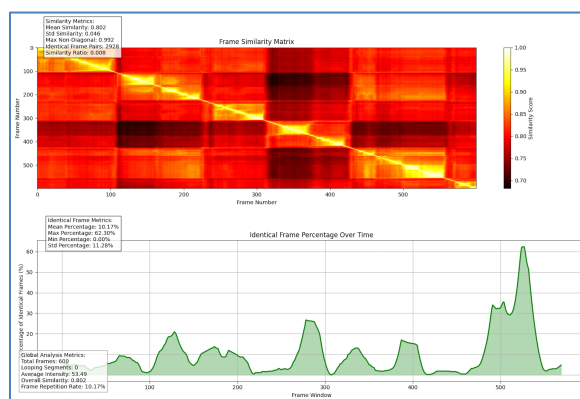


Figure 6. Frame Similarity Matrix & Identical Frame Percentage Graph (Looping)

Top Chart Analysis

Based on the data provided, this video has a total of 600 frames. The average similarity value between frames is 0.802, indicating a fairly high level of similarity in general. The standard deviation of similarity of 0.046 indicates a relatively low variation in similarity between frames.

Further analysis shows that the maximum non-diagonal similarity value reaches 0.992, meaning that there are some frames that are very similar to each other, even outside the main diagonal. The number of identical frame pairs reaching 2,928 also confirmed the presence of a lot of repetition or visual redundancy in this video.

However, the similarity ratio is only 0.008, indicating that the similarity level of the frames in this video is relatively low compared to the reference value. This

indicates that there is significant variation in the visual content throughout the video, even though there are many similar frames.

This graph reveals that the video has a fairly cohesive structure, with many frames similar to each other. However, the low similarity ratio suggests that the video also has considerable visual variety, so it is not overly repetitive or looping.

Table 1. Summary of Frame Similarity Matrix Graph Results

Parameters	Value
Total Frame	600
Average Similarity Score	0.802
Standard Deviation Similarity	0.046
Maximum Non-Diagonal Similarity	0.992
Number of Identical Frame Pairs	2,928
Similarity Ratio	0.008

Bottom Graph Analysis

Based on the data provided, this video has a total of 600 frames. The average percentage of identical frames is 10.17%, indicating that there are quite a few frames that are similar to each other. However, the maximum identical frame percentage reaches 62.30%, indicating that there are segments in the video that have many very similar frames. On the other hand, the minimum identical frame percentage is 0.00%, indicating that there are also segments in the video that have no similar frames.

The standard deviation value of the percentage of identical frames of 11.68% indicates a considerable variation in the degree of similarity between frames. No looping segments were detected in this video. The frame repetition rate of 10.17% indicates that there is a fair amount of repetition or visual redundancy in the video.

Overall, this graph reveals that the video has a fairly varied structure, with segments that have many identical frames, but also segments that are very different. The high level of frame repetition also suggests that there is visual redundancy in the video.

Table 2. Summary of Identical Frame Percentage Graph Results

Parameters	Value
Total Frame	600
Average Percentage of Identical Frames	10.17%
Maximum Percentage of Identical Frames	62.30%
Minimum Identical Frame Percentage	0.00 %

Standard Deviation Percentage	11.68%
Number of Looping Segments	0
Frame Rep Rate	10.17%

Analysis Conclusion

In both graphs, there is a high degree of similarity, but there is no looping pattern in the video. A total of 600 frames form an adequate digital canvas to explore visual complexity, providing enough space to analyze patterns and transformations. Interestingly, although no explicit looping segments were detected, the videos exhibit significant repetition characteristics.

5. Analysis Results Based on Clustering Metric Values.

Table 3. Clustering Metrics Value on Normal Video

Metrics	Value
Accuracy	0.936
Precision	0.910
Recall	0.682
F1-Score	0.780

A precision of 0.91 or 91% indicates that the clustering algorithm rarely gets a frame wrong. Of all the frames predicted to belong to a cluster, 91% of them actually fit and are relevant to that cluster. This high precision value indicates that the error rate in predicting cluster membership is quite low. High precision is essential to ensure that the clustered frames are truly appropriate and relevant to their clusters.

Recall of 0.682 or 68.2% shows that the algorithm successfully detects most of the frames that should belong to a particular cluster. However, there are still about 31.8% of frames that should belong to a cluster but are not detected by the algorithm. The lower recall value compared to precision indicates that there are some frames that are difficult to identify by the algorithm, so not all relevant frames are successfully clustered correctly.

F1-Score of 0.78 or 78% is a harmonious average between precision and recall. This value indicates that the clustering algorithm has a good balance between precision and the success rate of detecting relevant frames (recall). Although recall is slightly lower than precision, the high F1-Score indicates that the overall algorithm works quite optimally in clustering frames in non-looping videos.

6. Analysis of Results Based on Trust Value

Table 4. Confidence Score Data.

Component	Value
Final Confidence Score (FCS)	40.3%
Raw Confidence Score (RCS)	23.0%
Consistency Score (CS)	79.0%
Median Confidence (MC)	36.1%
Consistency (C)	78.98%
Periodicity (P)	0.00%
Regularity (R)	0.00%
Pattern Stability (PS)	0.0 %

Based on the data provided, the analysis shows that the processed video is a video that does not loop. The Final Confidence Score (FCS) of 40.3% and the Raw Confidence Score (RCS) of only 23.0% indicate the low confidence of the system in assessing the video. This suggests that while there are some elements that may be recognized, the overall reliability in identifying the video as not looping is very limited. The absence of Pattern Stability (PS) at 0.0% indicates that there are no repeating elements in the video, which is a hallmark of looping videos.

In addition, the low Median Confidence (MC) value 36.1% reinforces the conclusion that the majority of detections in the video do not have high confidence. While the Consistency Score (CS) of 79.0% indicates stability in the analysis results, it is not enough to support the detection of video looping. The absence of Temporal Consistency which also reached 0.0% indicates that no elements were detected consistently throughout the duration of the video, further reinforcing that the video did not have a repeating pattern.

Finally, the absence of a detected Pattern Count confirms that the video has no repeating elements. Thus, it can be concluded that a low confidence value indicates that the processed video is a non-looping video. Although there is consistency in the analysis results, the system still needs to be improved to enhance the ability to more accurately detect videos that are not looping.

Looping Video Analysis

1. Video Analysis Based on Intensity and Cluster (Looping)

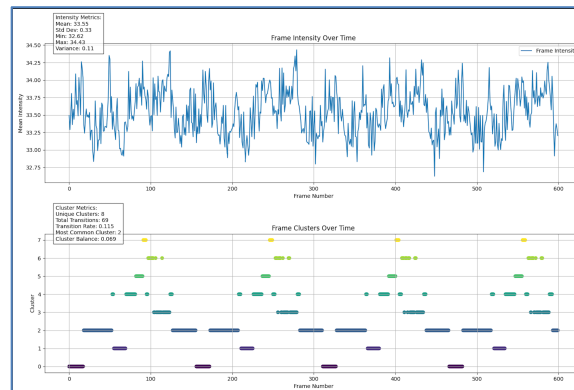


Figure 7. Intensity and Cluster Analysis Graph (looping)

Top Chart Analysis

In the top graph, "Frame Intensity Over Time," we can see the fluctuation of frame intensity against time with values ranging from 30.1 to 34.5. These fluctuations show a consistent periodic pattern, where the intensity rises and falls regularly.

The mean intensity value is 32.33 with a standard deviation of 0.53, indicating that despite variations, the intensity of the frames remains within a relatively narrow range. The temporal consistency value of 0.13 indicates stable changes in intensity between frames, supporting the presence of structured visual patterns and possible looping patterns in the video.

Bottom Chart Analysis

The bottom graph, "Frame Clusters Over Time," shows an even distribution of frames into five clusters (1, 4, 5, 6, and 7). Cluster 2 dominates most of the frames, indicating that the visual patterns in the video often return to this cluster. Clusters 0 and 3 have a sizable number of frames, but remain periodically distributed. The cluster distribution also shows a repetitive pattern, where frames move from one cluster to another within a certain interval.

This supports the indication of looping in the video. With a Cluster Balance value of 0.87, the distribution of clusters is fairly even although there is a dominance of certain clusters, especially Cluster 2 which is the dominant cluster. Overall, this graph shows that the video has a structured visual pattern, with stable intensity fluctuations and repetitive cluster distribution, supporting the hypothesis of a looping pattern in the video.

2. Output Analysis Based on Cluster Transition Graphs at Each Time and Temporal Analysis with Graphs



Figure 8. Cluster Transition Over Time Graph (looping)

The Cluster Transitions Over Time graph shows the cluster transitions over time based on the distribution of frames in the video. The horizontal (x) axis represents the frame number, while the vertical (y) axis shows the cluster assigned to each frame. This graph provides insight into how video frames move from one cluster to another, with the blue line showing the cluster assignment for each frame and the red vertical line showing the transition points between clusters. In general, this graph shows that the transition between clusters does not occur consistently throughout the video. There are segments where the clusters remain stable for some time, followed by segments with many rapid transitions between clusters. This pattern shows that the video has temporal variations, with certain parts being more dynamic (many cluster changes) and other parts being more static (fixed clusters).

In some parts of the graph, such as around frames 0 to 100, 200 to 300, and 400 to 500, the cluster remains stable for a long time. This indicates that the frames in these segments have similar visual characteristics, so all the frames are grouped into the same cluster. These segments likely reflect scenes that are more static or do not have many visual changes, such as a focus on a single object or a consistent background. In contrast, around frames 100 to 200, 300 to 400, and 500 to 600, there are many transitions between clusters in a short period of time. This indicates that the frames in these segments have high visual variation, such as transitions between scenes, fast movements, or significant changes in visual elements. These segments reflect a more dynamic and complex part of the video.

The red vertical lines in the graph show the transition points between clusters. This transition pattern seems to repeat in some video segments, reflecting an organized temporal structure. For example, more frequent transitions in the middle of the video indicate higher dynamics compared to the beginning and end of the video.

In this graph, it can be seen that some clusters appear more frequently than others. For example, cluster 1 and cluster 2 seem to dominate most of the video segments. This suggests that most frames in the video have visual characteristics similar to these clusters, while other clusters may only appear occasionally. Transitions between clusters reflect significant changes in the visual characteristics of the frames. For example, a move from cluster 1 to cluster 3 may reflect a scene transition, a change in the main object in the video, or a change in visual conditions such as lighting or color.

The transition patterns seen in this graph show that the video has an organized temporal structure, with more static and more dynamic segments alternating. This pattern can be used to understand the flow of the video, such as when scenes change or when certain patterns reappear.

3. Output Analysis Based on Frame Distribution Graph between Clusters and Entropy Graph

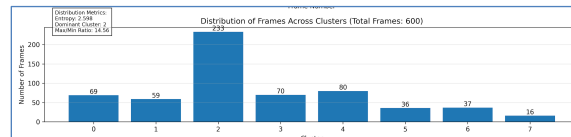


Figure 8. Frame Distribution Graph between Clusters and Entropy Analysis (Looping)

Chart Analysis

The frame distribution graph shows a very uneven pattern, with 600 video frames divided into 8 clusters (0 to 7). Cluster 2 dominates the distribution with 233 frames, which is equivalent to 38.8% of the total frames. This indicates that most of the frames in the video have very similar characteristics, reflecting the pattern of intense repetition in certain segments, which is typical of looping videos. In contrast, cluster 7 has the least number of frames, with only 16 frames or 2.7% of the total. The frames in this cluster most likely have unique or infrequent characteristics, such as transitions between scenes or moments that are visually distinct from the rest of the video.

This imbalance in distribution is reinforced by the maximum/minimum ratio of 14.56, which indicates that the number of frames in the dominant cluster (cluster 2) is much larger than that in the minority cluster (cluster 7). In addition, the entropy value of 2.598 indicates a low level of diversity in the distribution, where most of the frames are concentrated in certain clusters, especially cluster 2, while other clusters have fewer frames. The intermediate clusters (0, 1, 3, 4, 5, 6) have a more balanced number of frames, ranging from 36 to 80 frames (6%-13.3%), which likely represents minor variations in the looping pattern, such as changes in camera angle, color variations, or minor movements that remain repetitive.

This pattern provides deeper insights into the structure of the video. The dominance of cluster 2 indicates the presence of intense repetition in certain parts of the video, which may be the main focus or core segment of the content. On the other hand, the presence of minority clusters such as cluster 7 indicates the presence of frames with different characteristics that may represent transitions or outliers in the video. This distribution can also be used for further analysis, such as temporal segmentation to understand how these clusters appear in time, or outlier analysis to understand the different characteristics of frames in the minority clusters. Overall, this distribution pattern shows that the video has a highly focused structure in certain parts, with little variation in others, reflecting the looping nature of the video.

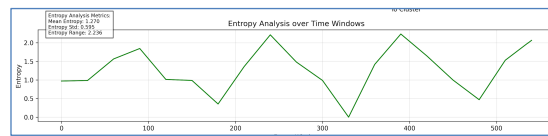


Figure 9. Entropy Analysis (Looping) Chart

Chart Analysis

The entropy analysis graph shows how the degree of diversity of the video frame distribution changes over time. The entropy values in this graph fluctuate between 0.5 and close to 2.5, with an average of about 1.59. This range of values reflects the variation in the frame distribution pattern in each time window. When the entropy value is high, the frame distribution is more evenly distributed among the clusters, indicating greater variation in video content. Conversely, low entropy values indicate a concentration of frames in a particular cluster, reflecting a strong repetition pattern in that video segment.

The entropy fluctuations seen in the graph show a repetitive up-and-down pattern. The entropy peaks occur around the 100, 300 and 500 time windows, with values close to 2.4. These segments reflect parts of the video with dynamic visual changes, such as transitions between scenes or variations in visual elements. In contrast, entropy valleys appear around the 50, 200, and 400 time windows, with values close to 0.5. In these segments, the frame distribution is highly concentrated, reflecting looping patterns or repetitive static scenes. This fluctuation pattern indicates a transition between dynamic and static segments in the video.

The entropy fluctuation pattern also reflects the repetitive temporal structure of the video. Segments with high entropy indicate greater content variation, such as transitions or scene changes, while segments with low entropy indicate concentration in certain clusters, reflecting focus on specific objects or repetitive visual patterns.

4. Output Analysis Based on Identical Frame Graph

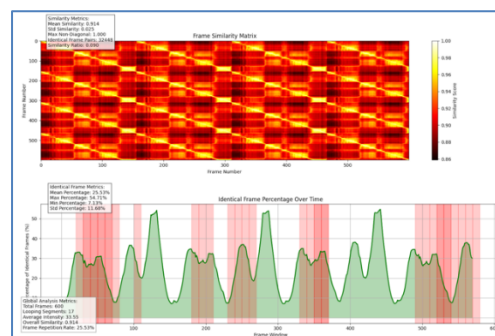


Figure 9. Frame Similarity Matrix & Identical Frame Percentage Graph (Looping)

First Graph Analysis

The first graph, the Frame Similarity Matrix, shows the degree of similarity between frames in the video. Lighter colors (yellow) indicate high similarity, while darker colors (dark red to black) indicate low similarity. It shows a repeating diagonal pattern, indicating a consistent visual structure and looping in the video.

The average value of similarity between frames is 0.914 with a standard deviation of 0.025, which indicates most frames have a high degree of similarity. In addition, there were 32,448 identical frame pairs, with a similarity ratio of 0.90. This indicates that 90% of the frames in the video are significantly similar, supporting a strong visual repetition pattern.

The Frame Similarity Matrix graph provides important information about the similarity pattern between frames in a video, especially at the maximum non-diagonal similarity. A maximum non-diagonal similarity value of 1,000 indicates the presence of frames that are identical even though they are not in the same position in the video sequence. The pattern of bright lines that appear periodically outside the main diagonal of the matrix indicates the repetition or looping of certain scenes. This means that certain frames have a very high degree of similarity with other frames at a certain distance in time, which is often found in videos with repetitive patterns such as animations or static scenes.

The fundamental difference between the average similarity value and the similarity ratio is also an important aspect in this analysis. The average similarity value of 0.914 reflects the global level of similarity between frames in the video, taking into account all partial and perfect similarity values. On the other hand, the similarity ratio of 0.090 only counts pairs of frames that are completely identical (similarity = 1.0), making it more specific in detecting explicit repetition or looping patterns. This difference indicates that although the frames as a whole have a high degree of similarity, only a small percentage of frame pairs are truly identical.

Below is a table summarizing the results of the Frame Similarity Matrix graph above:

Table 5. Summary of Frame Similarity Matrix Graph Results

Parameters	Value
Total Frame	600
Average Similarity Score	0.914
Standard Deviation Similarity	0.025
Maximum Non-Diagonal Similarity	1.00
Number of Identical Frame Pairs	32.448
Similarity Ratio	0.90

Second Graph Analysis

The second graph, Identical Frame Percentage Over Time, shows the percentage of identical frames within a specific time window throughout the video. This graph shows clear fluctuations, with peaks signifying segments where frame repetition is very high. The average percentage of identical frames is 25.53%, with the maximum value reaching 54.71% and the minimum 7.13%. The standard deviation of 11.68% indicates that there is significant variation in the repetition rate across segments.

The graph also identifies 17 looping segments marked with red areas, where intense repetition of frames occurs. The overall frame repetition rate is 25.53%, which means about a quarter of the total frames in the video are identical frames.

Below is a table summarizing the results of the Identical Frame Percentage graph above:

Table 6. Summary of Identical Frame Percentage Graph Results

Parameters	Value
Total Frame	600
Average Percentage of Identical Frames	25.53%
Maximum Percentage of Identical Frames	54.71%
Minimum Identical Frame Percentage	7.13%
Standard Deviation Percentage	11.68%
Number of Looping Segments	17
Frame Rep Rate	25.53%

5. Analysis of Results Based on Clustering Metrics Values

Table 7. Clustering Metric (Looping) Table

Metrics	Value
Accuracy	0.959
Precision	0.910
Recall	0.794
F1-Score	0.848

Precision measures how many frames predicted to belong to a particular cluster actually do, i.e. the ratio of True Positives to the total frames predicted to belong to that cluster (True Positive + False Positive). A high precision, such as 91% in this result, indicates that the algorithm rarely gets the frames wrong, so the error rate in

predicting the members of a particular cluster is quite low. Precision is very important in ensuring that the clustered frames are actually relevant to the cluster.

Recall (sensitivity) is a metric that measures how many frames that should belong to a particular cluster were successfully detected by the algorithm. Recall is calculated as the ratio between True Positives and the total frames that should belong to that cluster (True Positive + False Negative). In this result, a recall of 79.4% indicates that the algorithm successfully detected most of the relevant frames, although there were about 20.6% of frames that should have gone into the cluster but were not detected. Recall is important when the focus is on ensuring all relevant frames are identified, although this may come at the expense of precision.

F1-Score is the harmonic mean between precision and recall, which provides a balance between the two. This metric is particularly useful when there is an imbalance between precision and recall, as F1-Score gives an overall picture of the algorithm's performance. In this result, an F1-Score of 84.8% indicates that the algorithm has a good balance between the level of precision and the success rate of detecting relevant frames (recall). A high F1-Score value indicates that the algorithm is performing optimally, even though recall is slightly lower than precision. Overall, the combination of these four metrics shows that the clustering algorithm used is quite reliable in detecting looping patterns in videos.

6. Analysis of Results Based on Trust Value

Table 8. Confidence Score Data

Component	Value
Final Confidence Score (FCS)	88.4%
Raw Confidence Score (RCS)	61.3%
Consistency Score (CS)	70.0%
Median Confidence (MC)	34.3%
Consistency Stability (C)	69.96%
Periodicity (P)	91%
Regularity (R)	76%
Pattern Stability (PS)	78%

The Final Confidence Score (FCS) value of 88.4% indicates that the system has a high level of confidence in judging the video as looping. This reflects the system's good ability to identify repetitive elements in the video. Moreover, the Raw Confidence Score (RCS) of 61.3% also supports this, indicating that the system can recognize some important features that signify looping patterns.

The Consistency Score (CS) value of 70.0% indicates that despite variations in detection, the system is able to provide fairly stable results in video analysis. This is important because consistency in detection is an indicator that the system does not produce random results. The Median Confidence (MC) of 34.3% shows that even

though there are some low-confidence detections, overall, the system can still provide reliable results.

Furthermore, the Consistency Stability (C) value of 69.96% indicates that the system can maintain stability in the detection results, despite some variations. The very high Periodicity (P) of 91% indicates that there are clear repeating elements in the video, which is typical of looping videos. Similarly, Regularity (R) is 76%, indicating that the repetition pattern in the video is quite consistent.

Finally, the Pattern Stability (PS) value of 78% indicates that there is a stable pattern in the video, further reinforcing that the video is looping. Overall, the results of this analysis show that the analyzed video is indeed a looping video, with the system having a good ability to detect repetitive and stable elements.

CONCLUSION

This research successfully develops and implements an automated method for detecting video looping using the K-Means Clustering algorithm. Through the analysis of 600 frames from various videos, the proposed method shows high effectiveness in identifying repetitive patterns and identical frames. The analysis results show that looping videos are able to identify 454 patterns with 544 identical frames, reaching 90.67% of the total frames, and resulting in a final Confidence Score of 88.4%. In contrast, the non-looping video produced only 37.17% identical frames with a Confidence Score of 40.3%. The method also demonstrated superior performance in terms of pattern stability and detection confidence, with 91% precision and 79.4% recall for looping videos. Although the detection accuracy for non-looping videos reached 93.6%, these results confirm that the K-Means Clustering algorithm is highly effective in detecting looping videos. Thus, this research successfully achieved the set goal of automatically identifying and detecting video looping. These findings highlight the importance of video authenticity in virtual interactions and demonstrate that the application of detection technologies such as K-Means Clustering can improve the integrity of video communications. Further research is recommended to deepen the analysis and improve the detection accuracy, as well as explore the application of this method in other contexts in the field of multimedia and communication.

BIBLIOGRAPHY

- Agnisarman, S., Lopes, S., Madathil, K. C., Piratla, K., & Gramopadhye, A. (2019). A survey of automation-enabled human-in-the-loop systems for infrastructure visual inspection. *Automation in Construction*, 97, 52–76. <https://doi.org/10.1016/j.autcon.2018.10.019>
- Barcellos, P., Bouvié, C., Escouto, F. L., & Scharcanski, J. (2015). A novel video based system for detecting and counting vehicles at user-defined virtual loops. *Expert Systems with Applications*, 42(4), 1845–1856.
- Cai, W., Zhao, J., & Zhu, M. (2020). A real time methodology of cluster-system theory-based reliability estimation using k-means clustering. *Reliability Engineering & System Safety*, 202, 107045. <https://doi.org/10.1016/j.res.2020.107045>

- Choukaier, D. (2024). Enhancing English As A Foreign Language (EFL) Instruction Through Digital Teaching Platforms: Analyzing The Impact Of Microsoft Teams, Zoom, And Google Meet On Communication And Participation. *Educational Administration: Theory and Practice*, 30(6), 2404–2418. <https://doi.org/10.53555/kuey.v30i6.5748>
- Emahiser, J., Nguyen, J., Vanier, C., & Sadik, A. (2021). Study of live lecture attendance, student perceptions and expectations. *Medical Science Educator*, 31, 697–707.
- Ghiani, G., Laporte, G., & Musmanno, R. (2022). *Introduction to Logistics Systems Management: With Microsoft Excel and Python Examples*. John Wiley & Sons.
- Haleem, A., Javaid, M., Qadri, M. A., & Suman, R. (2022). Understanding the role of digital technologies in education: A review. *Sustainable Operations and Computers*, 3, 275–285. <https://doi.org/10.1016/j.susoc.2022.05.004>
- Jang-Jaccard, J., Nepal, S., Celler, B., & Yan, B. (2016). WebRTC-based video conferencing service for telehealth. *Computing*, 98(1), 169–193. <https://doi.org/10.17705/1CAIS.01213>
- Kaysi, F., Aydemir, E., & Surucu, Y. I. (2023). An Artificial Intelligence-Based Approach to Assess and Classify University Students' Attendance in Live Sessions. *International Online Journal of Educational Sciences*, 15(4).
- Lei, J., Jiang, T., Wu, K., Du, H., Zhu, G., & Wang, Z. (2016). Robust K-means algorithm with automatically splitting and merging clusters and its applications for surveillance data. *Multimedia Tools and Applications*, 75, 12043–12059.
- Mentzer, N. J., Isabell, T. M., & Mohandas, L. (2024). The impact of interactive synchronous HyFlex model on student academic performance in a large active learning introductory college design course. *Journal of Computing in Higher Education*, 36(3), 619–646.
- Mohsin, A. H., Zaidan, A. A., Zaidan, B. B., Albahri, O. S., Albahri, A. S., Alsalem, M. A., & Mohammed, K. I. (2019). Blockchain authentication of network applications: Taxonomy, classification, capabilities, open challenges, motivations, recommendations and future directions. *Computer Standards & Interfaces*, 64, 41–60. <https://doi.org/10.1016/j.csi.2018.12.002>
- Nadire, C., & Daniel, S.-A. (2021). A Comparison of Online Video Conference Platforms: Their Contributions to Education during COVID-19 Pandemic. *World Journal on Educational Technology: Current Issues*, 13(4), 1162–1173.
- Ninye-Ranor, L. C., Igbinedinon, V. I., & Nwadiani, C. O. (2022). OFFICE MANAGERS'UTILIZATION OF DIGITAL COMMUNICATION TOOLS FOR EFFECTIVE DISSEMINATION OF INFORMATION IN THE 21ST CENTURY BUSINESS ORGANISATIONS. *Technical and Vocational Education Journal (TAVEJ)*, 8, 176–184.
- Patel, V. M., Chellappa, R., Chandra, D., & Barbello, B. (2016). Continuous user authentication on mobile devices: Recent progress and remaining challenges. *IEEE Signal Processing Magazine*, 33(4), 49–61.
- Sarmah, P., Das, R., Dhamija, S., Bilgaiyan, S., & Mishra, B. S. P. (2022). Facial identification expression-based attendance monitoring and emotion detection—A deep CNN approach. In *Machine Learning for Biometrics* (pp. 155–176). Elsevier.
- Tochukwu, I. C., & Nonyelum, O. F. (2024). Effectiveness of Electronic Meeting and Video Conferencing Tools and Techniques. *IUP Journal of Computer Sciences*, 18(4), 29–55.
- Ukhov, P. A., Dmitrochenko, B. A., & Ryapukhin, A. V. (2021). The practice of

- technological deception in videoconferencing systems for distance learning and ways to counter it. *Amazonia Investiga*, 10(40), 153–168.
- Wang, Z., Zhang, Q., Yang, B., Wu, T., Lei, K., Zhang, B., & Fang, T. (2021). Vision-based framework for automatic progress monitoring of precast walls by using surveillance videos during the construction phase. *Journal of Computing in Civil Engineering*, 35(1), 4020056. [https://doi.org/10.61/\(ASCE\)CP.1943-5487.0000933](https://doi.org/10.61/(ASCE)CP.1943-5487.0000933)
- Widodo, J. (2021). Analisis kebijakan publik: Konsep dan aplikasi analisis proses kebijakan publik. Media Nusa Creative (MNC Publishing).