OPEN ACCESS JOURNALS
Indiana Journal of Multidisciplinary Research
Abbreviated Tittle- Ind J Multi Res
ISSN *(Online)*- 2583-3820
Journal Homepage Link- https://indianapublications.com/Journals/IJMR
DOI: https://doi.org/10.5281/zenodo.21602324

INDIANA PUBLICATIONS
PRODUCTIVE AND QUALITY RESEARCH

Research Article

Volume-06|Issue-04|2026

Dynamic Subtitle Color Adaptation Using K-Means Clustering

Vighnesha[1], Aryan Tony Johnson[2], Sydel Wivel Afonso[3], Terence Johnson[4]

[123]UG Students, Mangalore Institute of Technology & Engineering, Moodabidre, Mangalore, Karnataka, India.
[4]Professor, Mangalore Institute of Technology & Engineering, Moodabidre, Mangalore, Karnataka, India.

Article History
Received: 05.05.2026
Accepted: 22.06.2026
Published: 25.07.2026

Citation
Vighnesha., Johnson, A. T., Afonso, S. W. & Johnson, T. (2026). Dynamic Subtitle Color Adaptation Using K-Means Clustering. *Indiana Journal of Multidisciplinary Research, 6*(4), 208-215.

Abstract: This work presents a novel approach to improving subtitle readability in video content by dynamically adapting the subtitle text color based on real-time background analysis. Unlike conventional fixed-color or simple contrast-based methods, which often fail in colorful or visually complex scenes, the proposed approach employs the K-Means clustering algorithm to identify the dominant colors in the lower region of each video frame where subtitles are displayed. Based on the extracted color information, the algorithm selects the most suitable subtitle text and outline colors to maximize visibility. The prototype synchronizes background color clustering with frame-by-frame subtitle rendering. Experimental results demonstrate that the proposed method significantly outperforms conventional average color- or brightness-based approaches. The proposed technique is efficient, adaptable, and can be extended to support multi-speaker color attribution and enhanced scene awareness, thereby contributing to the development of more accessible multimedia technologies.

Keywords: Dynamic subtitles, K-Means clustering, subtitle readability, video captioning, color adaptation, background analysis.

INTRODUCTION

Subtitles and captions play a crucial role in improving accessibility and comprehension for viewers watching videos in noisy environments, for individuals who are deaf or hard of hearing, and for non-native speakers. However, the visual complexity and rapidly changing colors commonly found in modern video content can significantly reduce subtitle readability. Traditional subtitle systems typically rely on fixed font colors or simple contrast-based techniques, which often perform poorly against colorful or dynamically changing backgrounds. As a result, subtitles become difficult to read, increasing viewer fatigue and reducing the overall viewing experience.

User feedback and previous studies indicate that rigid subtitle styling not only reduces readability but also causes frustration, particularly when subtitles fail to adapt to changing on-screen content or diverse viewing preferences. A large number of respondents in an online survey expressed a strong preference for accurate, customizable subtitle settings and adaptive display techniques that improve viewer satisfaction and comprehension [1].

To address these challenges, several adaptive subtitle styling techniques, including contrast-based color selection, have been proposed. However, these methods generally consider only average brightness while ignoring complex color distributions and variations within the subtitle background region.

In this work, we propose a novel approach that utilizes the K-Means clustering algorithm to dynamically identify the dominant colors in the subtitle background region [2]. Our prototype video player performs real-time color clustering and synchronizes adaptive subtitle color selection with video playback. Compared with conventional average color- or brightness-based methods, the proposed system provides significantly improved subtitle visibility under visually challenging conditions while integrating seamlessly with web-based and multimedia video players.

Figure 1 illustrates white subtitle text displayed over a white or similarly bright region of the video background [3]. In this situation, the subtitle blends into the background, making it difficult—or even impossible—for viewers to read. This example highlights the fundamental challenge of subtitle readability when subtitle colors are not dynamically adapted to the underlying video content.

***Corresponding Author:** Vighnesha

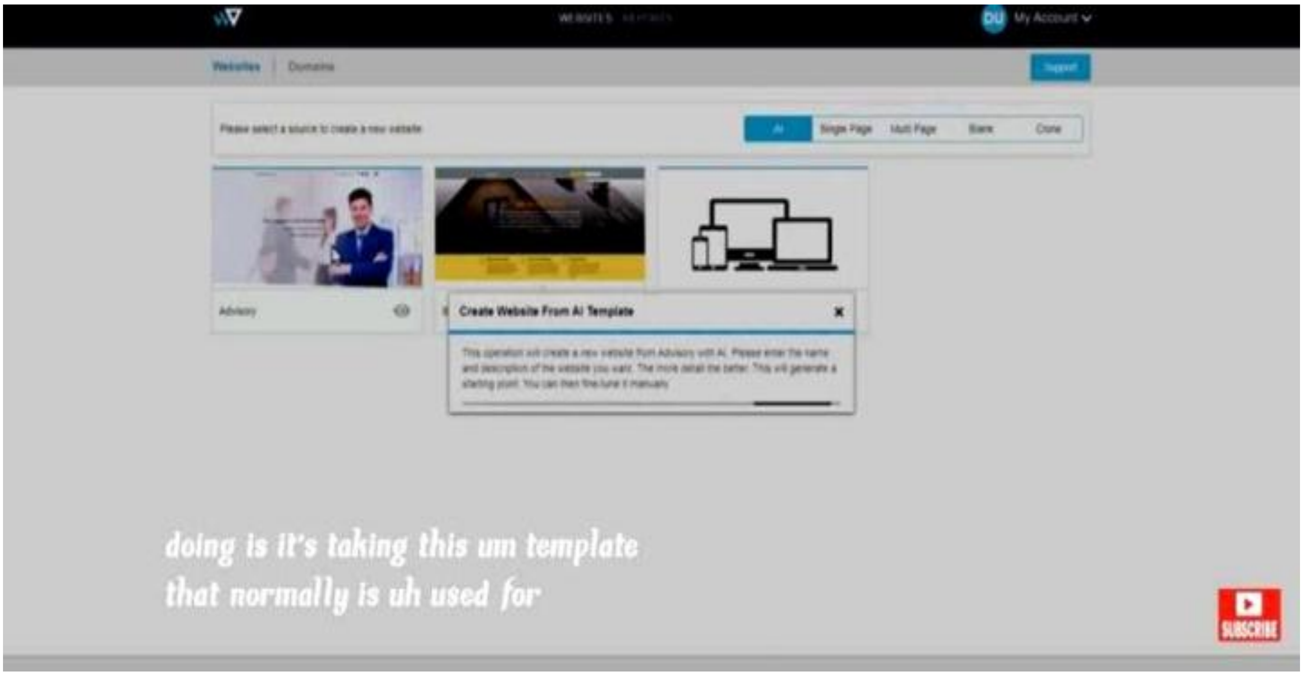

Fig. 1. White subtitle text merging into the video background.

Figure 2 illustrates a scenario in which the user manually changes the subtitle color to a different shade. However, the selected color still overlaps with a background region of similar color or brightness, causing the subtitle to remain difficult to read. This demonstrates the limitations of static or manually selected subtitle colors when dealing with complex or rapidly changing backgrounds [4].

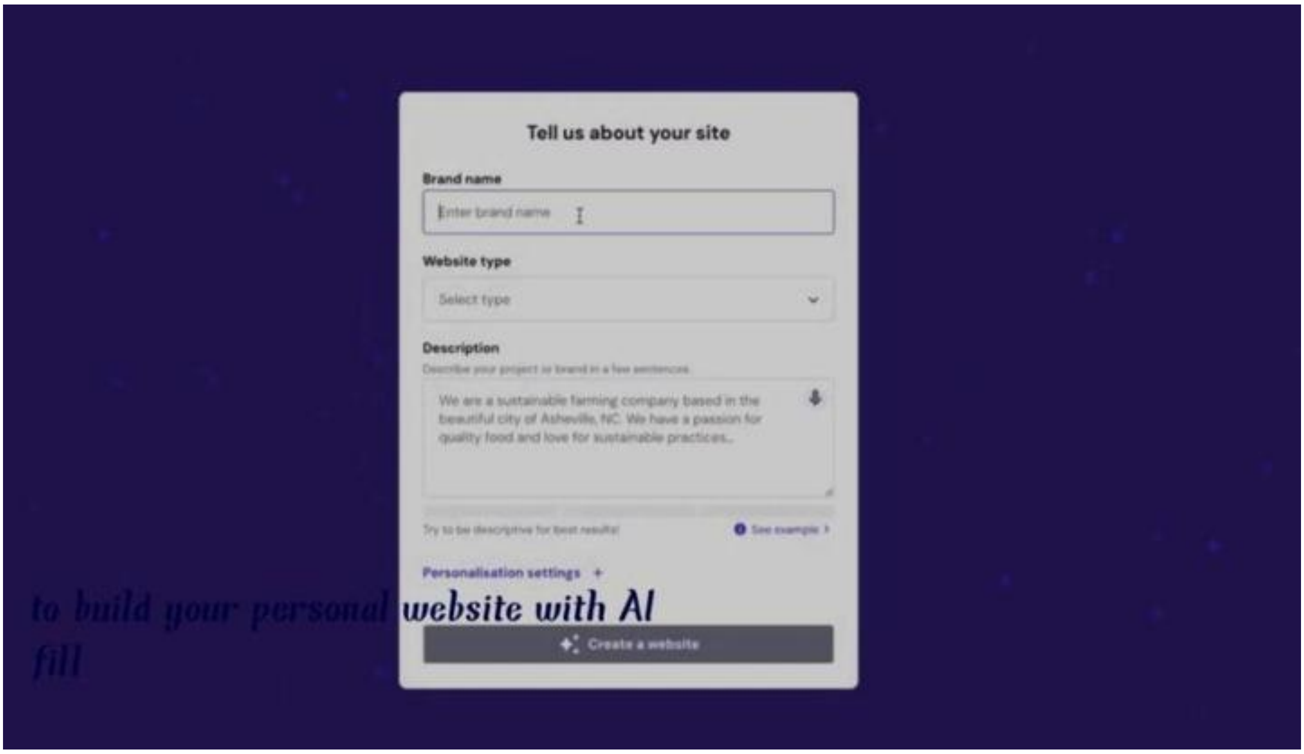

Fig. 2. User-selected subtitle color failing against a similar background.

Figure 3 shows white subtitle text displayed over a solid black background box to improve contrast. Although this approach enhances readability, the opaque background box can be visually intrusive because it obscures portions of the video content. This example illustrates the trade-off between subtitle visibility and viewer experience when using fixed background boxes instead of adaptive subtitle styling techniques [5].

***Corresponding Author:** Vighnesha

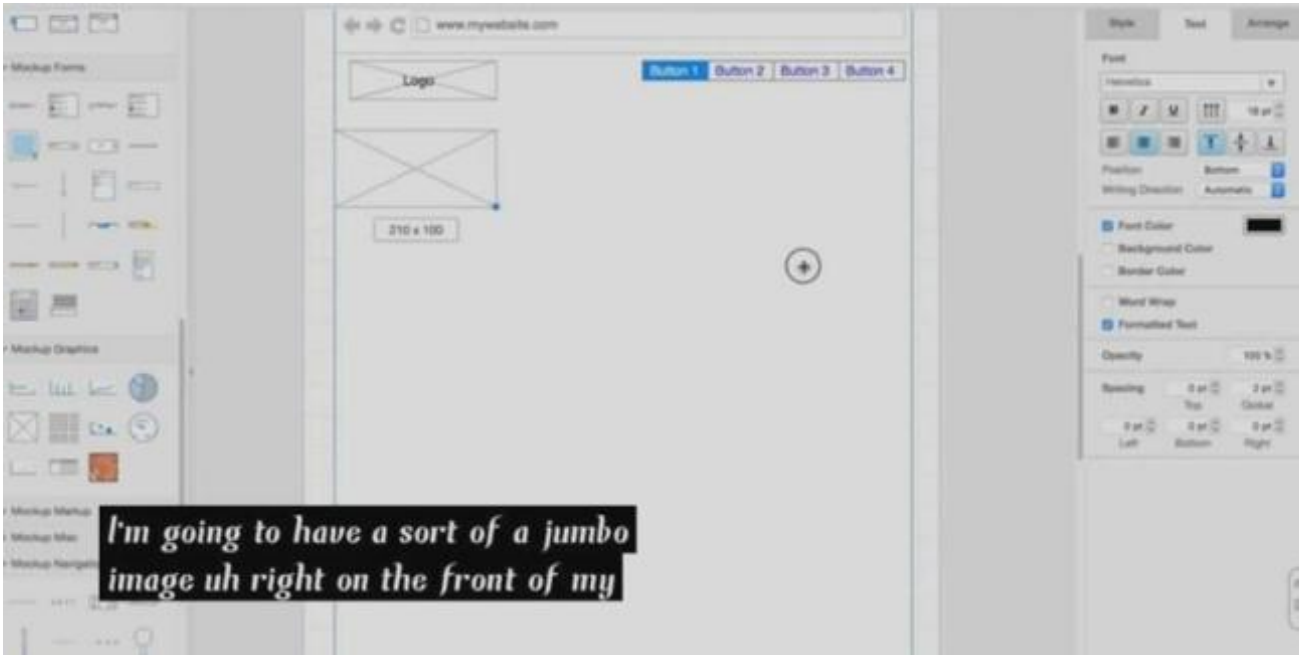

Fig. 3. White subtitle text over a solid black background box.

LITERATURE REVIEW

Montagud et al. present a web-based framework that enables extensive subtitle customization and synchronization across single- and multi-screen environments. Users can modify the font, size, color, position, number of lines, language, and timing of subtitles, along with synchronized display on companion devices. This level of customization enhances accessibility and improves the overall Quality of Experience (QoE) by adapting subtitles to users' contexts and preferences [6].

Rodriguez-Alsina et al. present a video subtitling system for Ambient Intelligence environments that enables subtitle customization, modification, and synchronization across multiple displays and devices. Their context-aware approach, based on HTML5 standards, provides personalized subtitles with flexible synchronization and seamless distribution, thereby improving accessibility, usability, and the overall user experience across various viewing platforms [7].

Rodriguez-Alsina et al. also propose a video subtitling system for Ambient Intelligence environments that enables subtitles to be synchronized, personalized, and modified across multiple displays and devices. By leveraging HTML5 standards, the system offers context-aware subtitles through flexible distribution and synchronization, thereby enhancing accessibility and user experience across multiple viewing devices [8].

Akahori et al. propose a dynamic subtitle placement system that utilizes speaker localization, Region of Interest (ROI) analysis, and eye-tracking data to minimize unnecessary eye movements and improve viewer comfort. By positioning subtitles close to the active speaker and just below the ROI, the system enhances subtitle recognition while allowing viewers to maintain focus more naturally and comfortably [9].

Hong et al. propose a dynamic captioning approach that improves accessibility for viewers with hearing impairments by enhancing caption synchronization, positioning, and emphasis. Using face recognition, visual saliency, and speech-to-text alignment, captions are placed near the active speaker while emphasizing spoken words, enabling viewers to follow conversations more effectively and recognize emotional cues [10].

Kurzhals et al. introduce gaze-adaptive captions that dynamically adjust subtitle placement based on viewers' eye gaze in real time. By positioning subtitles near the viewer's focal region, the system minimizes distraction and aligns with natural viewing behavior. User studies demonstrate improved reading efficiency and more natural gaze patterns, particularly among novice viewers [11].

Durairaj proposes an interactive color image segmentation technique based on the HSV (Hue, Saturation, Value) color space, which has been shown to outperform several alternative methods. A Python-based graphical user interface allows users to adjust the Hue, Saturation, and Value parameters to isolate foreground objects effectively for object recognition applications [12].

Chitade and Katiyar employ K-Means clustering for color-based image segmentation, primarily using satellite imagery. Their approach improves color separation before classification using color attributes. Among the evaluated color spaces, the Lab color space performs better than RGB, YCbCr, and HSV. K-Means

***Corresponding Author:** Vighnesha

effectively segments images into distinct color regions without requiring labeled training data [13].

Yamini and Rajagopal propose an adaptive color image segmentation technique for steganalysis based on Otsu's thresholding method. Color stego images are first converted into grayscale before segmentation to evaluate the effectiveness of pixel-neighbor clustering. Experimental results demonstrate that Otsu's method identifies adaptive regions containing hidden information more accurately than Roberts edge detection and particle swarm optimization-based clustering techniques [14].

Cai and Zhang propose an average color vector method in the RGB color space for distinguishing and identifying monochromatic objects. The method computes average color values, standard deviations, and Euclidean distances between pixels and reference vectors. When evaluated on tomato color classification using Canny edge detection, the approach achieved good performance but exhibited limitations when distinguishing between similar color shades [15].

Rajinikanth and Couceiro propose a color image segmentation technique that combines Otsu's between-class variance with the Firefly Algorithm (FA) to optimize RGB histogram thresholds for both bi-level and multi-level segmentation. The effectiveness and quality of the segmentation are evaluated by comparing Brownian motion, Lévy Flight, and Gaussian randomization strategies using performance metrics such as CPU time, Structural Similarity Index Measure (SSIM), and Peak Signal-to-Noise Ratio (PSNR) [16].

Kim proposes the use of Gaussian Mixture Models (GMMs) together with the Expectation-Maximization (EM) algorithm for detecting color patches in images and video sequences. GMM parameters are generated from a reference image using CbCr chrominance features, and posterior probability estimation enables pixel-wise detection and segmentation of objects in input images and video frames [17].

Thiruvenkadam and Sakthivelu propose a technique for enhancing the contrast and brightness of low-light images by combining mathematical morphology with a modified version of Weber's Law. Morphological operators first improve background illumination before Weber's Law is applied, resulting in superior enhancement of both general and MRI images compared with conventional histogram equalization. Performance evaluation demonstrates significant improvements in Average Mean Brightness Error (AMBE) and SSIM [18].

Wang and Chen propose a two-stage background estimation method for removing shadows from single document images to improve visual quality and digitization. Their method performs iterative local background color estimation for shadow detection, followed by global background color-based shadow removal and tone normalization. Although effective for static document images, the method is not intended for dynamic subtitle adaptation in video content [19].

Although existing research has made significant progress in subtitle customization, subtitle placement, and static image processing, real-time, context-aware subtitle color adaptation during video playback remains largely unexplored. Most previous studies focus on static images, subtitle positioning, or user customization rather than dynamically adapting subtitle colors according to changing video backgrounds.

To address this limitation, the proposed method employs K-Means clustering to analyze the dominant background colors within the subtitle region on a frame-by-frame basis. Based on the detected dominant colors, the system automatically adjusts the subtitle text and outline colors to maintain optimal contrast and readability. This adaptive approach significantly improves subtitle visibility in visually complex and dynamically changing scenes, thereby addressing an important gap in multimedia accessibility.

METHODOLOGY

K-Means Clustering Algorithm

Algorithm

1. Arbitrarily initialize **K** cluster centroids ($mean_1$, $mean_2$, ..., $mean_k$).
2. Assign each data point to the cluster whose centroid is nearest.
3. Compute the new centroid (mean) for each cluster.
4. Repeat Steps 2 and 3 until convergence.

The K-Means algorithm is one of the most widely used clustering algorithms for partitioning data into K distinct clusters based on similarity. After selecting the desired number of clusters (K), K data points are randomly chosen as the initial cluster centroids.

The algorithm then iteratively performs two main steps. In the first step, each data point is assigned to the cluster whose centroid is closest according to a distance metric, typically Euclidean distance. Once all data points have been assigned, the second step recalculates the centroid of each cluster by computing the arithmetic mean of all points belonging to that cluster.

The newly computed centroids are then compared with those from the previous iteration. If the centroids remain unchanged—or equivalently, if the cluster assignments no longer change—the algorithm is considered to have converged, and the current clusters are treated as the final clustering result.

If the centroids differ, the assignment and centroid-update steps are repeated until convergence is achieved.

*Corresponding Author: Vighnesha

For high-dimensional or very large datasets, convergence may require many iterations. In some cases, the algorithm may oscillate or require excessive computation before stabilizing. To prevent an infinite or excessively long execution, a maximum iteration limit, denoted by **T**, is commonly specified. If convergence has not been achieved after **T** iterations, the algorithm terminates, and the clusters obtained during the **T**th iteration are considered the final clusters [20].

K-Means Clustering Algorithm
Arbitrarily select **K** pixels (pix_1, pix_2, pix_3, …, pix_k) from the input image as the initial cluster centroids (means).

Repeat:
1. Assign each 3D RGB pixel in the input image to the cluster with the nearest centroid.
2. Compute the new mean RGB value for each cluster.
3. Repeat until convergence.

Initially, **K** pixels from the bottom 20% of the pixel colors in the 3D RGB space are randomly selected as the initial cluster centroids. Thereafter, two iterative steps are performed until the pixels converge into **K** clusters.

In the first step, each pixel in the 3D RGB color space is assigned to the cluster whose centroid is closest, using the **Euclidean distance** as the similarity measure. In the second step, the mean RGB value of each cluster is computed by averaging the RGB values of all pixels assigned to that cluster. The cluster centroid is then updated using this newly computed average.

These assignment and update steps are repeated until one of the following conditions is satisfied:
- The cluster centroids no longer change between successive iterations.
- The cluster assignments remain unchanged from the previous iteration.
- The predefined maximum number of iterations (**T**) is reached.

The lowest 20% of the pixel intensities in the 3D RGB space are clustered into **three clusters** using the K-Means algorithm. The cluster containing the largest number of pixels is considered to represent the dominant background color behind the subtitles. The luminance of this dominant background color is then calculated to determine whether **gray level 0 (black)** or **gray level 255 (white)** should be used for the subtitle text and outline, thereby maximizing contrast and improving readability.

IMPLEMENTATION

To evaluate the proposed real-time adaptive subtitle color algorithm, we developed a **web-based platform** that allows users to upload video clips together with subtitle (**SRT**) files. The system analyzes the subtitle background region on a frame-by-frame basis using the **K-Means clustering algorithm** to identify the dominant background colors. Based on the detected colors, the system automatically adjusts both the subtitle text color and the outline color to ensure optimal contrast and readability throughout video playback, thereby addressing the common problem of subtitles blending into complex or dynamically changing backgrounds.

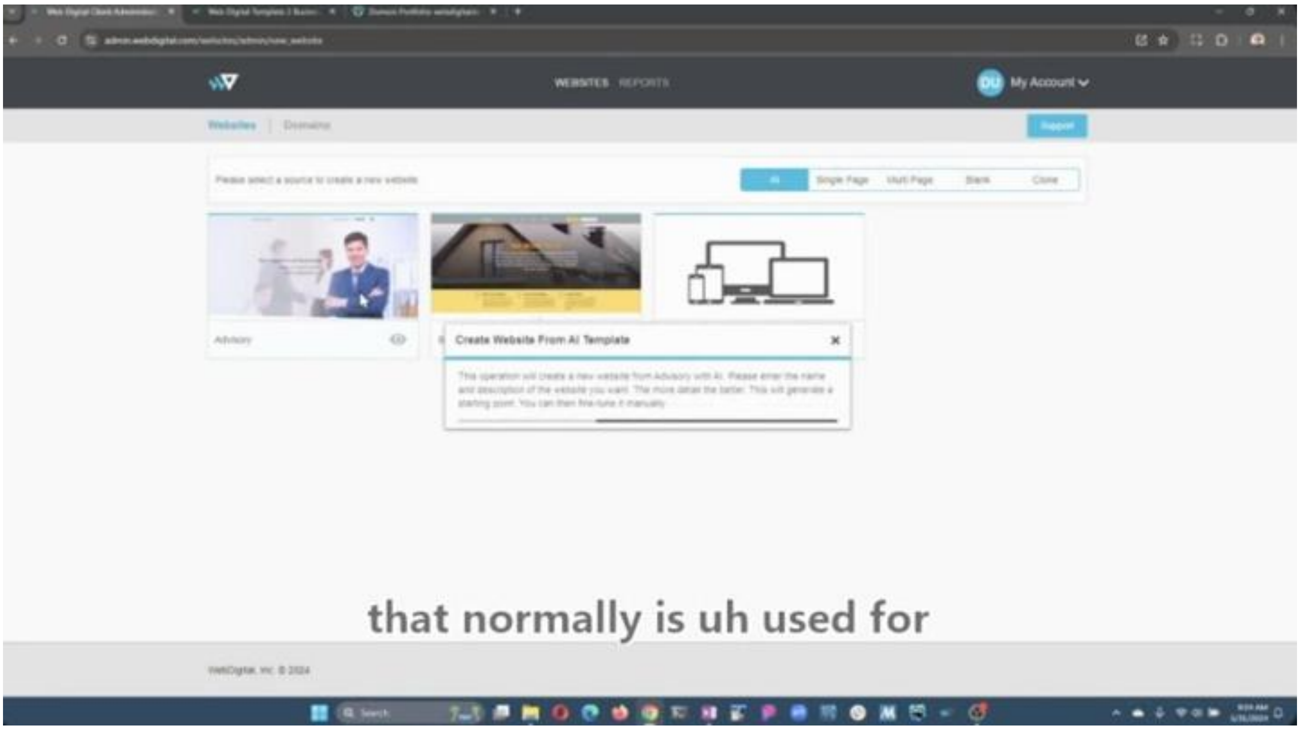

Fig. 4. Dynamic subtitle color adjustment using the K-Means clustering algorithm.

Figure 4 [3] illustrates a screenshot of the proposed system automatically adapting the subtitle color to maximize contrast with the background. This dynamic adjustment significantly enhances subtitle visibility and improves the overall viewing experience. The results demonstrate that the proposed method effectively addresses subtitle readability through real-time, context-aware color adaptation.

RESULTS AND DISCUSSION

The implementation of the proposed K-Means clustering method demonstrates that the system can dynamically identify the dominant colors within the subtitle background region. In situations where conventional static subtitles would normally blend into the background, the automatic adjustment of subtitle text and outline colors substantially improves visibility. Furthermore, the proposed approach eliminates the need for intrusive opaque background boxes or manual subtitle color adjustments, thereby enhancing the overall user experience.

To evaluate the effectiveness of the proposed approach, we compared the **K-Means clustering-based dynamic subtitle color adaptation** method with a conventional **Baseline (Average Invert)** subtitle adaptation technique. The comparative results are illustrated in **Fig. 5** [3] and **Fig. 6** [21].

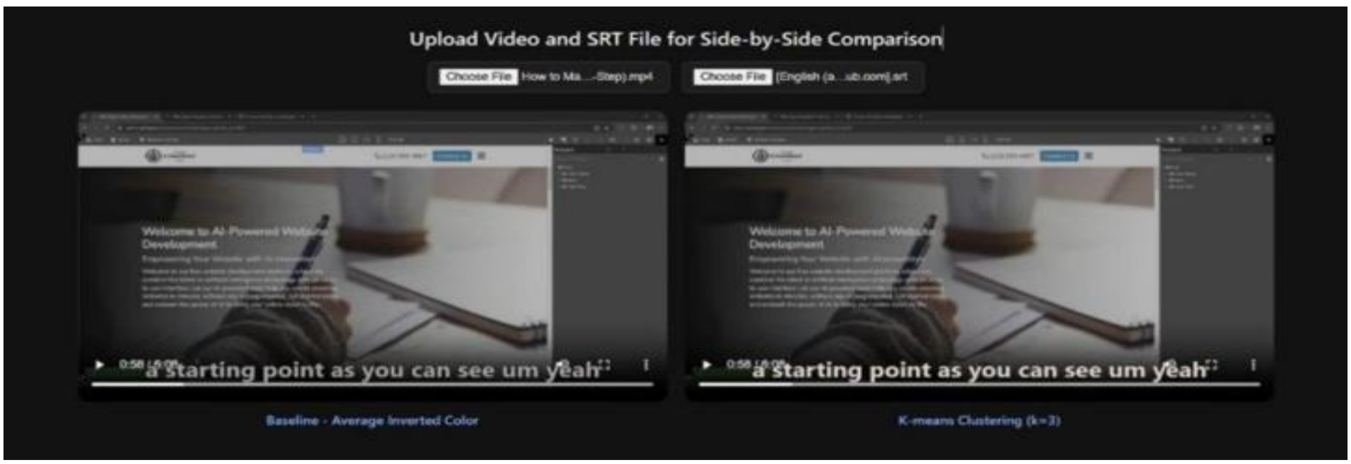

Fig. 5. Comparison between the Baseline and K-Means algorithms.

Fig. 6. Comparison between the Baseline and K-Means algorithms.

A web-based demonstration was developed in which two panels were displayed side by side: one implementing the **Baseline (Average Invert)** method and the other implementing the proposed **K-Means** approach. Both panels displayed identical rapidly changing synthetic backgrounds consisting of a wide range of colors. Subtitle colors, contrast ratios, and processing times were recorded in real time for every video frame, enabling a direct comparison of subtitle adaptation performance and readability under challenging visual conditions.

Tables **1** and **2** summarize the comparative performance of the two methods.

***Corresponding Author:** Vighnesha

Table 1. Performance of the Baseline (Average Invert) Algorithm

Time (s)	BG Color (RGB)	Baseline Subtitle Color	Baseline Contrast	Baseline Adaptation (ms)
0.1	202, 114, 114	61, 137, 137	1.20	150.70
1.9	107, 207, 207	144, 56, 56	3.80	4.70
3.2	209, 98, 98	54, 151, 152	1.07	4.90
4.5	90, 207, 207	159, 56, 56	3.39	4.90
5.9	201, 84, 84	61, 164, 164	1.43	4.70
7.2	81, 191, 191	166, 70, 70	2.53	4.90
8.5	182, 81, 81	78, 166, 166	1.65	4.90
9.8	82, 174, 174	165, 85, 85	1.93	5.50

Table 2. Performance of the K-Means Clustering Algorithm

Time (s)	BG Color (RGB)	K-Means Subtitle Color	K-Means Contrast	K-Means Adaptation (ms)
0.1	202, 114, 114	53, 141, 141	1.15	42.80
1.9	107, 207, 207	148, 48, 48	4.20	15.30
3.2	209, 98, 98	46, 157, 157	1.14	15.20
4.5	90, 207, 207	165, 48, 48	3.67	15.60
5.9	201, 84, 84	54, 171, 171	1.55	15.40
7.2	81, 191, 191	174, 64, 64	2.65	15.80
8.5	182, 81, 81	73, 174, 174	1.86	15.20
9.8	82, 174, 174	173, 81, 81	1.97	18.10

From Tables 1 and 2, it can be observed that the proposed K-Means method consistently achieves contrast ratios that are equal to or higher than those produced by the Baseline method, particularly under dynamically changing backgrounds, while maintaining adaptation times suitable for real-time applications.

Although the Baseline method generally performs faster after the initial frame, it frequently fails to achieve the recommended accessibility contrast ratio of **4.5:1**. In contrast, the proposed K-Means approach successfully adapts to changing backgrounds and reaches this accessibility threshold at **1.9 seconds**. Overall, while both methods adapt subtitle colors according to background changes, the K-Means algorithm provides superior stability and improved contrast across a wider range of visual scenarios.

These results clearly demonstrate the effectiveness of the proposed K-Means clustering approach over conventional subtitle adaptation techniques, particularly in situations where subtitle readability and visual complexity are critical considerations.

Despite its advantages, the proposed method has several limitations. The processing time is higher than that of the Baseline algorithm, and performance may decrease in scenes containing highly fragmented colors, such as rapidly flashing or visually noisy backgrounds. Future work will focus on improving computational efficiency through optimized dynamic region selection, faster clustering techniques, and robust fallback strategies for challenging visual conditions.

During experimentation, different numbers of clusters and subtitle region sizes were evaluated to investigate the sensitivity of the proposed approach to K-Means parameters. The results indicate that using **two to four clusters** provides reliable color identification and optimal subtitle contrast, whereas larger numbers of clusters tend to reduce adaptation accuracy due to excessive color fragmentation.

CONCLUSION

The proposed method was directly compared with the Average Invert algorithm, a baseline adaptive technique that adapts subtitle colors by inverting the average color of the subtitle background region. Compared with brightness-only approaches, the proposed K-Means clustering method identifies and adapts to the locally dominant background colors, resulting in higher and more consistent subtitle contrast ratios—a critical factor for readability in visually complex scenes.

This work addresses the persistent challenge of subtitle readability in videos, particularly when static subtitle colors blend into complex or dynamically changing backgrounds. The proposed methodology performs real-time analysis of the subtitle background using the K-Means clustering algorithm and dynamically adjusts the subtitle text and outline colors to maximize contrast and readability.

The experimental results demonstrate significant improvements in subtitle visibility and overall user experience while eliminating the need for intrusive opaque background boxes or manual subtitle color

***Corresponding Author:** Vighnesha

adjustments. The proposed approach effectively addresses an important accessibility gap identified in previous studies by providing reliable, automated subtitle adaptation across a wide variety of video scenarios.

Future work will focus on extending the proposed system to live-streaming environments, incorporating additional accessibility features, and providing greater user customization. Furthermore, future versions may support **multi-speaker subtitle color attribution** by identifying active speakers and assigning unique, high-contrast subtitle colors using the same region-specific K-Means clustering approach synchronized with speech segmentation techniques.

REFERENCES

1. Gorman, B. M., Crabb, M., & Armstrong, M. (2021). Adaptive subtitles: Preferences and trade-offs in real-time media adaptation. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, Article 733, 1–11.
2. MacQueen, J. B. (1967). Some methods for classification and analysis of multivariate observations. In L. M. Le Cam & J. Neyman (Eds.), *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability* (Vol. 1, pp. 281–297). University of California Press.
3. WebDigital. (2025). *How to make a free website with AI in minutes (Step-by-step)* [Video]. YouTube. https://youtu.be/jm6aHoGwDlg
4. Today Web. (2025). *Make a personal website with Hostinger website builder: Step-by-step tutorial* [Video]. YouTube. https://youtu.be/fVeNxW_dKJo
5. Mr Dimmick's Computing Channel. (2025). *Creating a website wireframe in draw.io* [Video]. YouTube. https://youtu.be/GDPDIPj5XWY
6. Montagud, M., Boronat, F., González, J., & Pastor, J. (2017). Web-based platform for subtitle customization and synchronization in multi-screen scenarios. *Proceedings of CONFERENCE 2017*, 81–82.
7. Rodriguez-Alsina, A., Talavera, G., Orero, P., & Carrabina, J. (2012). Subtitle synchronization across multiple screens and devices. *Sensors, 12*(7), 8710–8731.
8. Rodríguez Alsina, A., Talavera Velilla, G., Orero, P., & Carrabina Bordoll, J. (2011). Analysis and design of a subtitling system for ambient intelligence environments. *Sensors*. Multidisciplinary Digital Publishing Institute (MDPI).
9. Akahori, W., Hirai, T., & Morishima, S. (2017). Dynamic subtitle placement considering the region of interest and speaker location. In *Proceedings of VISIGRAPP* (pp. 102–109).
10. Hong, R., Wang, M., Xu, M., Yan, S., & Chua, T.-S. (2010). Dynamic captioning: Video accessibility enhancement for hearing impairment. In *Proceedings* (pp. 421–430).
11. Kurzhals, K., Göbel, F., Angerbauer, K., Sedlmair, M., & Martin, R. (2020). A view on the viewer: Gaze-adaptive captions for videos. In *Proceedings*.
12. Durairaj, H. (2020). Interactive color image segmentation using HSV color space. *Science & Technology Journal, 7*, 37–41.
13. Chitade, A., & Katiyar, S. (2010). Color-based image segmentation using K-means clustering. *International Journal of Engineering Science and Technology, 2*.
14. Yamini, B., & Rajagopal, S. (2017). Image steganalysis: Adaptive color image segmentation using Otsu's method. *Journal of Computational and Theoretical Nanoscience, 14*, 4502–4507.
15. Cai, Y., & Zhang, L. (2012). Average color vector algorithm in color recognition based on an RGB space. In *2012 IEEE 14th International Conference on Communication Technology* (pp. 1043–1047).
16. Rajinikanth, V., & Couceiro, M. S. (2015). RGB histogram-based color image segmentation using the Firefly algorithm. *Procedia Computer Science, 46*, 1449–1457.
17. Kim, D. K. (2017). Color detection using Gaussian mixture models. *Journal of Theoretical and Applied Information Technology, 95*, 4313–4320.
18. Thiruvenkadam, K., & Sakthivelu, K. (2015). Image contrast enhancement using modified Weber's law. In *Proceedings* (June 2015).
19. Wang, B., & Chen, C. (2019). An effective background estimation method for shadow removal in document images. In *2019 IEEE International Conference on Image Processing (ICIP)* (pp. 3611–3615).
20. Jin, X., & Han, J. (2011). K-means clustering. In C. Sammut & G. I. Webb (Eds.), *Encyclopedia of Machine Learning* (pp. 563–564). Springer. https://doi.org/10.1007/978-0-387-30164-8_425
21. FuseSchool – Global Education. (2025). *Biological molecules | Cells | Biology* [Video]. YouTube. https://youtu.be/FgfknBZaVTI

***Corresponding Author:** Vighnesha