**Mini Project Report**

**on**

**Video Summarization Generating Movie Trailer:**

Submitted by

## Student Names Reg. Numbers

## M. Navadeep(22bds040), Kamal Das(22bds035), A.Tulasinarayana(22bds004), Om pratham(22bcs133)

Under the guidance of

**Krishnendu Gosh**

**Designation**



**DEPARTMENT OF DATA SCIENCE AND ARTIFICIAL INTELLIGENCE**

**INDIAN INSTITUTE OF INFORMATION TECHNOLOGY DHARWAD**

11/04/2024

## *Certificate*

This is to certify that the project, entitled **“AI-based Movie Trailer Generator”**, is a bonafide record of the Mini Project coursework presented by the students whose names are given below during **Academic Year 2025–25** in partial fulfilment of the requirements of the degree of Bachelor of Technology in Data Science and Artificial Intelligence.

|  |  |
| --- | --- |
| Roll No | Names of Students |
| 22bds040 | M. Navadeep |
| 22bds035    22bds004 | Kamal Das  A. Tulasi Narayana |
| 22bcs133 | U. Om Pratham |

**Krishnendu Gosh**

(Project Supervisor )

**Contents:**

[1.Introduction: 1](#_Toc195277211)

[2. Related Work: 2](#_Toc195277212)

[3. Data and Methods 3](#_Toc195277213)

[4.Results and Discussions 4](#_Toc195277214)

5.Conclusion 5

6.Refrences 6

# 1.Introduction:

In the modern digital era, movie trailers play a crucial role in attracting audiences and creating initial impressions about a film. A trailer must encapsulate the mood, story, and excitement of a movie in just a few minutes. However, traditional trailer creation is a manual, labor-intensive, and time-consuming process involving hours of video inspection, scene selection, and editing.

With the rise of artificial intelligence and machine learning, there is a growing interest in automating creative tasks such as video editing and content summarization. Our project, titled **AI-based Movie Trailer Generator**, leverages modern AI techniques to automate the trailer generation process by intelligently analyzing and selecting the most impactful scenes from a movie.

This system combines multiple AI modules—including emotion recognition, scene detection, audio analysis, and text summarization—to create a concise, engaging trailer with minimal human intervention. It aims to bridge the gap between technology and creative media by delivering a cost-effective and scalable solution to video production workflows. in attracting viewers. Traditional trailer creation is manual, time-consuming, and expensive. This project proposes an **AI-powered movie trailer generator** that intelligently selects impactful scenes based on visual and auditory features, dynamically edits video segments, and assembles a coherent trailer with AI-generated narration, subtitles, and music overlays.

The primary goal of the system is to reduce human effort while maintaining cinematic quality in automatically generated trailers.

# 2. Related Work:

Numerous works have explored video summarization, scene detection, and AI in film editing. These studies have laid the groundwork for developing more sophisticated systems that can analyze and interpret multimedia content:

* **Video Summarization Techniques**: Traditional techniques involve visual features like color histograms, edge detection, and optical flow for identifying keyframes or segments. Recent advancements use deep learning models to identify semantically meaningful clips.
* **Emotion Recognition**: Emotion detection has advanced with the help of convolutional neural networks (CNNs), recurrent neural networks (RNNs), and attention-based transformer models. These models help classify user or character emotions from audio, text, and visual cues.
* **Multimodal Analysis**: Advanced systems now integrate audio, visual, and textual data to derive richer insights about content. This includes voice tone analysis, facial emotion recognition, and natural language processing (NLP) for dialogue interpretation.

Our project builds upon these ideas by combining them into a unified pipeline. Unlike previous works that focus on one modality or task, our system merges multiple AI capabilities to handle real-world complexity in movie trailer generation., scene detection, and AI in film editing:

* **Video Summarization Techniques**: Shot boundary detection, keyframe extraction, and feature-based video summarization have been used in previous research.
* **Emotion Recognition**: Emotional intelligence in content selection is influenced by CNN, RNN, and Transformer-based models.
* **Multimodal Analysis**: Combines audio, visual, and textual features for understanding movie segments.

These works laid the foundation for our system which integrates multiple modalities using deep learning and OpenCV.

# 3. Data and Methods

**Dataset Used**

 **Audio Clips**: From various movie genres to study speech energy and silence patterns.

**Methodology:**

*  **Scene Detection**: Implemented using OpenCV for detecting hard cuts via frame differences and color histograms.
*  **Emotion Recognition**: A DistilBERT-based transformer model was fine-tuned to classify dialogue emotional tone from subtitles and script.
*  **Audio Analysis**: Short-Time Energy (STE) and Zero-Crossing Rate (ZCR) methods were used to detect speech-heavy, silent, or intense background music segments.
*  **Subtitle & Tagline Generation**: Extractive summarization techniques generated short, impactful lines overlaid as subtitles or end tags.
*  **Video Editing & Composition**: MoviePy library stitched together selected clips, added fade transitions, subtitles, voiceover, and music to produce the final trailer.

# 4.Results and Discussions

**Outputs Generated:**

 Automatically selected scenes matched intended emotions.

 Smooth integration of AI-generated narration and subtitles.

 Final trailers were coherent, cinematic, and engaging.

* .

**Performance:**

The effectiveness of each component was evaluated on sample data:

* **Scene detection accuracy**: ~87% using manual annotations for validation.
* **Emotion classification F1-score**: ~78% across multiple emotion categories.
* **Trailer generation time**: Less than 5 minutes to generate a 2-minute trailer using average processing hardware.

**Challenges Faced:**

* **Audio-visual mismatches**: Sometimes the emotion from text did not align with the scene’s visual tone.
* **Pacing and transitions**: Ensuring smooth transitions between scenes without manual fine-tuning was challenging.
* **Data limitations**: Public datasets lacked diversity in scene structures and genres, making generalization difficult.

Despite these, our system performed well across various genres like drama, action, and thriller. Informal testing with peers indicated high engagement and interest in the generated trailers. and aligned music/speech.

* Automated addition of taglines and subtitles.
* Emotionally resonant scenes selected using AI.

**5. Conclusion:**

The **AI-based Movie Trailer Generator** successfully automates trailer creation using deep learning, NLP, and multimedia processing. It showcases how AI can assist in creative media production, reducing human effort and increasing scalability.

Future work may include:

* Genre-specific fine-tuning,
* Adaptive trailer pacing based on content,
* Real-time previews and customization tools.

This project demonstrates how the integration of computer vision, audio processing, and natural language understanding can enhance the automation of storytelling in multimedia content.

# 6. References:

 Gygli, M., Grabner, H., Riemenschneider, H., & Van Gool, L. (2014). *Creating summaries from user videos*. ECCV.

 Potapov, D., et al. (2014). *Category-specific video summarization*. ECCV.

 Sutskever, I., Vinyals, O., & Le, Q. V. (2014). *Sequence to sequence learning with neural networks*. NeurIPS.

 OpenCV Documentation – <https://docs.opencv.org/>

 MoviePy Library – <https://zulko.github.io/moviepy/>