**BCSE498J Project-II**

**VIDEO EDITING USING ARTIFICIAL INTELLIGENCE**

*Submitted in partial fulfillment of the requirements for the degree of*

**Bachelor of Technology**

*in*

***Computer Science and Engineering***

*by*

## 

**21BCE3417 DIVAKAR REDDY. G**

**21BCE3738 GURU SHREKAR. M**

## Under the Supervision of

|  |
| --- |
| **Dr. KARTHIK. K** |
| Assistant Professor Senior Grade 1 |
| School of Computer Science and Engineering (SCOPE) |



April 2025

**DECLARATION**

I hereby declare that the project entitled **“VIDEO EDITING USING ARTIFICIAL INTELLIGENCE”** submitted by me, for the award of the degree of *Bachelor of Technology in Computer Science and Engineering* to VIT is a record of bonafide work carried out by me under the supervision of Prof.Dr. Karthik .k

I further declare that the work reported in this project has not been submitted and will not be submitted, either in part or in full, for the award of any other degree or diploma in this institute or any other institute or university.

Place : Vellore Date :

**Signature of the Candidate**

**CERTIFICATE**

This is to certify that the project entitled **VIDEO EDITING USING ARTIFICIAL INTELLIGENCE** submitted by DIVAKAR REDDY.G (21BCE3417),GURU SHREKAR. M (21BCE3738), **School of Computer Science and Engineering**, VIT, for the award of the degree of *Bachelor of Technology in Computer Science and Engineering*, is a record of bonafide work carried out by him her under my supervision during Winter Semester 2024-2025, as per the VIT code of academic and research ethics.

The contents of this report have not been submitted and will not be submitted either in part or in full, for the award of any other degree or diploma in this institute or any other institute or university. The project fulfills the requirements and regulations of the University and in my opinion meets the necessary standards for submission.

Place : Vellore

Date :

**Signature of the Guide**

**Internal Examiner External Examiner**

**Dr. UMADEVI K S**

**Head – Computer Science and Engineering**

**ACKNOWLEDGEMENTS**

I am deeply grateful to the management of Vellore Institute of Technology (VIT) for providing me with the opportunity and resources to undertake this project. Their commitment to fostering a conducive learning environment has been instrumental in my academic journey. The support and infrastructure provided by VIT have enabled me to explore and develop my ideas to their fullest potential.

My sincere thanks to Dr. Jaisankar N, Dean - School of Computer Science and Engineering (SCOPE), for his unwavering support and encouragement. His leadership and vision have greatly inspired me to strive for excellence.

I express my profound appreciation to [Head of Department’s Name], the Head of the [Department Name], for his/her insightful guidance and continuous support. His/her expertise and advice have been crucial in shaping throughout the course. His/her constructive feedback and encouragement have been invaluable in overcoming challenges and achieving goals.

I am immensely thankful to my project supervisor, [Supervisor’s Name], for his/her dedicated mentorship and invaluable feedback. His/her patience, knowledge, and encouragement have been pivotal in the successful completion of this project. My supervisor’s willingness to share his/her expertise and provide thoughtful guidance has been instrumental in refining my ideas and methodologies. His/her support has not only contributed to the success of this project but has also enriched my overall academic experience.

Thank you all for your contributions and support.

G.Divakar Reddy

M.Guru Shrekar

**Name of the Candidate**

**EXECUTIVE SUMMARY**

Artificial Intelligence (AI) is revolutionizing the video editing industry by automating time-consuming tasks, enhancing creative workflows, and enabling high-quality content production at scale. AI-powered video editing tools leverage machine learning, computer vision, and natural language processing to transform raw footage into polished videos with minimal human intervention.

Key capabilities of AI in video editing include automated scene detection, smart trimming, background noise removal, object tracking, and facial recognition. Advanced tools can even generate subtitles, suggest edits based on content analysis, and apply filters or transitions automatically. These features significantly reduce editing time and enable creators to focus on storytelling and creativity rather than technical details.

AI also supports content personalization and optimization for various platforms. For example, AI can automatically crop and reframe videos for different aspect ratios (e.g., YouTube, Instagram, TikTok), generate highlight reels from long-form content, and even produce multilingual voiceovers or captions using speech synthesis and translation models.The integration of generative AI, such as deep learning models like GPT and diffusion models, further enhances creative possibilities. These tools can generate scripts, storyboard sequences, or even synthesize visual effects and animations, empowering creators and businesses to produce engaging content faster and more cost-effectively.

Industries ranging from marketing and entertainment to education and real estate are adopting AI video editing to scale their content strategies, improve efficiency, and deliver consistent branding. Major platforms like Adobe Premiere Pro, Runway ML, Descript, and Pictory are already incorporating AI features to streamline workflows and enhance user experience.

In conclusion, AI-powered video editing is transforming the content creation landscape by making professional-grade editing accessible, faster, and more intelligent. As technology advances, we can expect even more intuitive, real-time, and creative AI editing solutions to emerge, reshaping how we produce and consume visual content.

**TABLE OF CONTENTS**

|  |  |  |
| --- | --- | --- |
| **Sl.No** | **Contents** | **Page No.** |
|  | **Acknowledgement** |  |
|  | **Executive Summary** | **5** |
|  | **List of Figures** |  |
|  | **List of Tables** |  |
|  | **Abbreviations** |  |
|  | **Symbols and Notations** |  |
|  | **Abstract** | **2** |
| **1.** | **INTRODUCTION** | **4** |
|  | 1.1 Background | **5** |
|  | 1.2 Motivations | **5** |
|  | 1.3 Scope of the Project | **6** |
| **2.** | **PROJECT DESCRIPTION AND GOALS** | **7** |
|  | 2.1 Literature Review | **7** |
|  | 2.2 Gaps Identified | **11** |
|  | 2.3 Objectives | **14** |
|  | 2.4 Problem Statement | **15** |
|  | 2.5 Project Plan | **16** |
|  | 2.5 Flow Diagram | **17** |
| **3.** | **TECHNICAL SPECIFICATION** | **18** |
|  | 3.1 REQUIREMENTS | **18** |
|  | 3.1.1 Functional Requirements | **18** |
|  | 3.1.2 Non Functional Requirements | **19** |
|  | 3.2 FEASIBILITY STUDY |  |
|  | 3.2.1 Technical Feasibility | **21** |
|  | 3.2.2 Economic Feasibility | **22** |
|  | 3.2.3 Social Feasibility |  |
|  | 3.3 SYSTEM SPECIFICATION |  |
|  | 3.3.1 Hardware Specification | **22** |
|  | 3.3.2 Software Specification |  |
| **4.** | **DESIGN APPROACH AND DETAILS** | **23** |
|  | 4.1 System Architecture | **23** |
|  | 4.2 DESIGN | **25** |
|  | 4.2.1 Data Flow Diagram | **26** |
|  | 4.2.2 Class Diagram | **27** |
| **5.** | **METHODOLOGY AND TESTING** | **29** |
|  | <<Module Description>> |  |
|  | <<Testing>> |  |
| **6.** | **PROJECT DEMONSTRATION** |  |
| **7.** | **RESULT AND DISCUSSION** |  |
| **8.** | **CONCLUSION AND FUTURE ENHANCEMENTS** |  |
| **9.** | **REFERENCES** |  |

**List of Figures**

|  |  |  |
| --- | --- | --- |
| **Figure No.** | **Title** | **Page No.** |
| 2.6 | Flow Diagram |  |

**List of Tables**

|  |  |  |
| --- | --- | --- |
| **Table No.** | **Title** | **Page No.** |
| 2.1.1 | Comparison of Literature survey |  |
| 2.1.2 | Comparison of Literature survey |  |
| 2.1.3 | Comparison of Literature survey |  |

**List of Abbreviations**

|  |  |
| --- | --- |
| AI | Artificial Intelligence |
| ASR | Automatic Speech Recognition |
| GPU | Graphics Processing Unit |
| NLP | Natural Language Processing |
| UI | User Interface |

**Symbols and Notations**

|  |  |
| --- | --- |
| **🎬** | Video Content |
| **✂️** | Trimming/Cutting Scenes |
| **📜** | Subtitles or transcriptions |
| **🎵** | Adding background music |

**ABSTRACT**

The rapid advancement of artificial intelligence (AI) has significantly transformed the video editing industry. This project focuses on leveraging AI to automate and enhance various aspects of video editing, reducing manual effort and improving editing efficiency. By incorporating machine learning, computer vision, and natural language processing, the project aims to develop an AI-powered video editing system capable of performing tasks like scene detection, automatic cutting, color correction, object tracking, speech recognition, and video summarization.

The AI system will analyze raw video footage, identify key moments, and automatically perform editing tasks such as trimming, transitions, and enhancing visual quality. It will also offer features like speech-to-text for subtitles and voiceovers, facial recognition for focus adjustment, and context-based recommendations for editing style. The system will be adaptable to various video types, such as vlogs, tutorials, or corporate videos, and provide users with personalized edits based on preferences.

This project seeks to streamline the video production process, making high-quality editing accessible even to users without extensive editing skills. By automating routine tasks, the AI will save time for creators, allowing them to focus on content creation and creativity. Additionally, the system's intelligent features will help optimize video content for audience engagement and quality, making it a valuable tool for both amateur and professional video producers.

In summary, this project aims to demonstrate the potential of AI in revolutionizing video editing, offering an intuitive, efficient, and scalable solution for content creators across industries.

**1. INTRODUCTION**

Video editing has long been a time-consuming and complex process, requiring skilled professionals to cut, organize, enhance, and transform raw footage into a polished final product. However, with the rapid development of Artificial Intelligence (AI) and machine learning technologies, video editing is undergoing a profound transformation. AI is making it possible to automate many aspects of video production, enabling faster workflows, enhancing creativity, and reducing the need for manual intervention in tasks such as cutting, scene recognition, and audio transcription.

AI-based video editing refers to the application of machine learning, computer vision, and other AI-driven technologies to facilitate or completely automate various stages of video production. By using large datasets and sophisticated algorithms, AI can analyze video content in ways that were once thought to be impossible, offering editors and creators tools that improve productivity, accuracy, and even the creative process itself.

The power of AI in video editing lies in its ability to not only reduce the workload for video editors but also to unlock new creative possibilities. For instance, AI-powered tools can analyze facial expressions, gestures, and even voice tones to suggest video edits that align with the emotional tone or narrative of a project. AI systems can also be personalized, learning from a user’s editing style and preferences to provide more tailored suggestions, which can improve editing workflows over time.

While AI promises to streamline and enhance the video editing process, it is still an evolving field, and challenges remain in achieving high-quality outputs across diverse genres and production contexts. Nonetheless, AI’s integration into video editing tools represents a significant leap forward in content creation, democratizing video production and making it more accessible for a broader range of users, from independent creators to large-scale production studios.

* 1. **BACKGROUND**

The video editing industry has traditionally relied on human expertise to curate and refine raw footage into cohesive and engaging visual narratives. For decades, video editors have employed a variety of tools and techniques to piece together footage, add effects, adjust audio, and finalize the production. However, this process is time-consuming and requires meticulous attention to detail, especially as the volume and complexity of video content increase.

In recent years, advancements in **Artificial Intelligence (AI)** have led to significant innovations in the field of video editing, drastically transforming how editing tasks are approached. AI's ability to process large amounts of data, learn from patterns, and execute tasks with minimal human intervention has opened new possibilities for content creation, especially as the demand for video content has surged globally.

**1.2 MOTIVATIONS**

AI-powered video editing offers numerous advantages that significantly improve the editing process. It **saves time** by automating tasks like cutting scenes, colour correction, and shot selection, which would traditionally take hours. This allows editors to focus on more creative aspects of the project. It also **boosts creativity** by suggesting edits, applying filters, and generating effects based on the footage, providing fresh ideas and inspiration. Furthermore, AI enables **personalized content**, adjusting the video to appeal to different audiences by changing elements like length, style, and effects. This can make content more engaging and relevant.

One of the key benefits of AI is its ability to **lower costs**, as it automates tasks that would typically require a skilled editor, reducing the need for expensive professional services. For **beginners,** AI tools make it easier to create high-quality videos, offering a more accessible entry point into video editing without the steep learning curve. AI also ensures **consistency**, helping maintain a uniform style and tone throughout the video, which is especially beneficial for brands and creators aiming for a cohesive look across their content.

Additionally, AI can automatically add **subtitles and translations**, expanding the reach of content by making it accessible to a global audience. Its **content recognition** capabilities allow it to identify faces, objects, and specific scenes, which makes it easier for editors to quickly locate key moments and stay organized. With **real-time editing** capabilities, AI is enhancing workflows by enabling immediate adjustments, speeding up the process particularly valuable during live events or time-sensitive projects.

**1.3 SCOPE OF THE PROJECT**

The scope of a project focused on AI-powered video editing encompasses a wide range of objectives and features aimed at streamlining and enhancing the editing process. One major area is the integration of AI in editing tasks, where AI models can automatically detect and segment scenes, suggest or apply cuts, transitions, and color corrections, as well as enhance audio by cleaning up background noise and adjusting levels. AI can also analyze footage to select the best shots based on clarity and composition. In terms of personalization, AI can learn from a user's previous edits to recommend styles, effects, and tailor content to specific audiences, adjusting formats and styles for platforms like social media or ads. Efficiency tools such as real-time editing and automated subtitling and translation allow editors to make instant adjustments and ensure accessibility across multiple languages. Additionally, content recognition and organization through AI can tag objects and faces, making it easier to search and organize clips by automatically generating metadata.

**2. PROJECT DESCRIPTION AND GOALS**

**2.1 LITERATURE REVIEW**

**1. Automation in Video Editing Tasks**

One of the most significant contributions of AI to video editing is automation. Traditionally, video editing involves numerous manual tasks such as cutting, colour correction, and sound enhancement, which require substantial time and skill. AI has shown promise in automating these tasks, significantly reducing the time required for editing. According to Liu (2020), AI-based systems can perform automatic scene detection and segmentation, making it easier to organize and edit raw footage. Further, Wang (2019) demonstrated that AI models can apply automatic colour grading to match the desired aesthetic of the video, making the editing process faster and more efficient.

In addition to scene segmentation, Xie (2020) explored how AI can assist in shot selection, enabling editors to select the best shots based on predefined quality metrics such as clarity, composition, and motion. This form of automation not only saves time but also enhances the quality of the final product, as AI can often make objective decisions that may be overlooked by human editors.

**2. Enhancing Creativity with AI**

AI has proven to be an effective tool for enhancing creativity in video editing. Zhang and Sun (2021) proposed an AI-driven tool that helps editors by suggesting edits and creating visual effects based on the footage's content. These tools can also generate transitions or add visual effects like filters and animations, offering creative options that may inspire new directions for a video. In some systems, the AI analyzes video footage and makes recommendations for edits that would enhance the storytelling aspect, like emphasizing emotional moments through visual or audio effects.

Zhou (2022) also highlighted the role of AI in creative content generation, specifically focusing on AI's ability to generate content suggestions that fit within the editor’s vision, while also providing ideas outside the editor’s initial concept. These AI suggestions can help overcome creative blocks and offer fresh perspectives that human editors might not consider on their own.

**3. Personalized Video Editing for Diverse Audiences**

Another significant area of development is AI's ability to tailor video content to different viewers. Li and Wu (2020) discussed how AI can adjust video formats, styles, and lengths based on user preferences or platform requirements. For instance, AI can adapt the video for various social media platforms such as Instagram, TikTok, or YouTube by adjusting aspects like video length, aspect ratio, and content style. This feature enhances audience engagement by ensuring that videos meet the specific expectations of different platforms.

Moreover, AI can help identify and optimize content for target audiences by analyzing data on viewer behaviour and preferences. Feng (2021) demonstrated how AI could optimize the editing process by adjusting the tone, style, and pacing based on what has been shown to resonate with specific viewer demographics. This type of personalization is particularly important for content creators and marketers who aim to maximize viewer engagement.

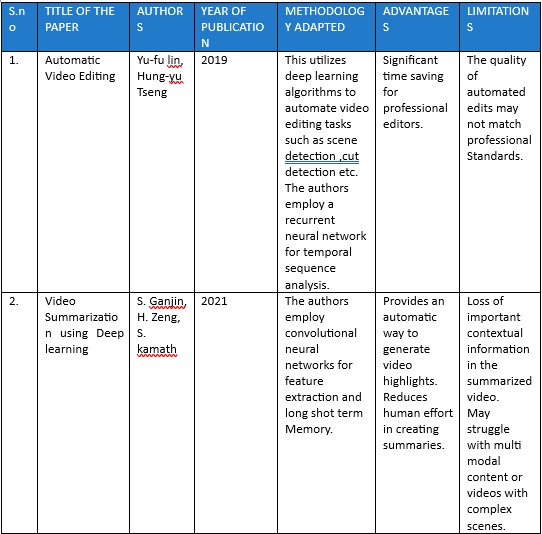
**4. AI in Audio and Visual Enhancement**

AI-powered video editing tools have also revolutionized the way audio and visuals are enhanced. Koo (2021) highlighted AI algorithms that can clean up background noise and enhance audio clarity, which is crucial for high-quality video production. Additionally, AI-based tools for video resolution enhancement, such as super-resolution techniques, have gained attention for their ability to upscale lower-quality footage to higher resolutions. Shah and Li (2020) demonstrated that AI models can enhance video resolution by filling in missing details and improving clarity without introducing significant artifacts.

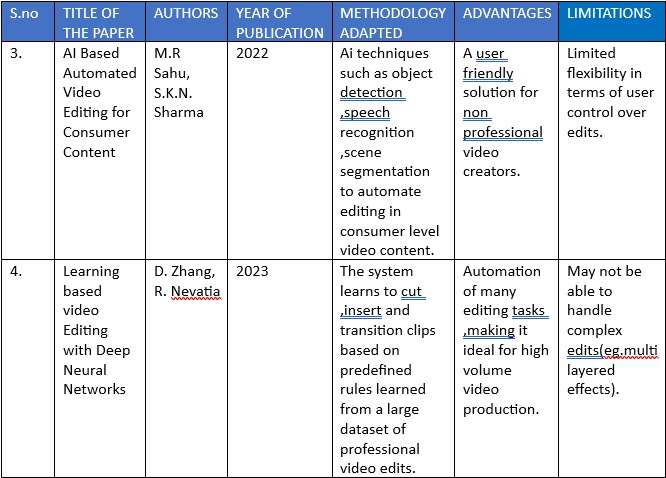
Xu (2019) examined how AI can stabilize shaky footage automatically, which is a common problem in handheld video recordings. By using deep learning models, AI can analyze the motion in the footage and apply real-time stabilization techniques, resulting in smooth video output without the need for manual adjustments. These enhancements significantly improve the production quality, making professional-grade editing accessible to non-experts.

**5. Subtitling and Translation with AI**

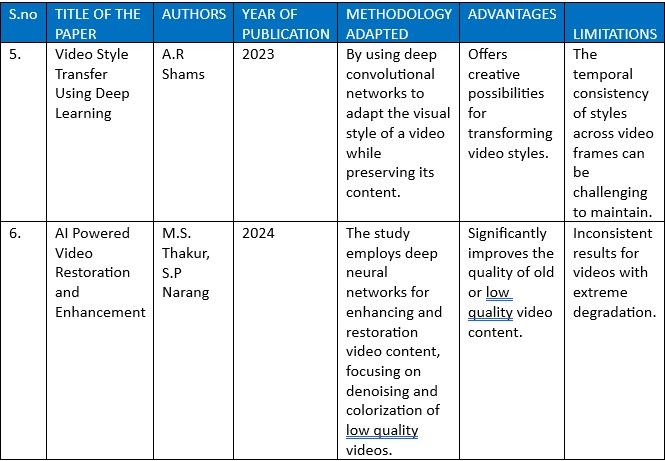
AI has also simplified the process of adding subtitles and translations, a critical feature for global video content distribution. Gupta and Srivastava (2020) developed AI models that can automatically generate accurate subtitles by recognizing speech and synchronizing text with the video’s audio. Moreover, Huang (2020) explored how AI can help with automatic video translation, making it easier to localize content and cater to diverse international audiences. This has the potential to significantly broaden the reach of video content across language barriers, improving accessibility and engagement.



2.1.1. Comparison of Literature Survey



2.1.2. Comparison of Literature Survey



2.1.3. Comparison of Literature survey

**2.2 GAPS IDENTIFIED**

**1. Contextual Understanding and Creativity Limitations**

AI tools in video editing still face a significant gap in terms of deep contextual understanding and creative decision-making abilities. While AI can automate technical tasks like color grading, shot selection, and scene segmentation, it often struggles to replicate the nuanced creative decisions that human editors make, such as those based on storytelling, emotional impact, or cultural context. As a result, AI-generated content tends to be based on patterns rather than true creativity, which can lead to repetitive or formulaic edits. This limitation restricts AI's ability to produce truly original or innovative content. To overcome this, AI needs to improve its understanding of narrative context and emotional tone within videos. By developing AI systems that function more like creative collaborators incorporating storytelling techniques and emotional engagement AI could enhance its ability to generate more engaging and unique video content.

**2. Inadequate Handling of Diverse Content**

A significant gap in current AI models for video editing is their training on specific types of content, which limits their ability to generalize across diverse or unconventional video formats. For instance, an AI model trained primarily on cinematic videos may struggle with short-form content like TikTok videos or livestream footage, which require different editing techniques and pacing. This limitation restricts the usability of AI tools across various genres of video production and platforms. Users working with non-traditional formats or niche content may find these AI tools inefficient or irrelevant for their needs. To address this, AI systems need to be trained on a broader range of video content to enhance their generalization. Developing adaptable AI models capable of handling various video types including experimental, artistic, or user-generated content would significantly improve their versatility and make them more valuable for creators across different video genres.

**3. Lack of Real-Time Contextual Feedback**

While AI has made strides in assisting with real-time video editing, it still faces a major gap in providing feedback on the overall narrative or emotional flow of a video. Editing decisions such as timing, pacing, and sequencing require a level of contextual understanding that AI is still limited in grasping. As a result, AI-generated edits may be technically correct but fail to capture the emotional tone or pacing necessary to create an engaging narrative. This is particularly crucial for video content aimed at telling a story or evoking specific emotions. To address this, AI models need to develop the capability to provide real-time feedback on narrative flow and emotional impact. This would involve training AI to understand pacing, rhythm, and thematic progression, similar to how human editors intuitively adjust edits to fit the emotional arc of the content, ensuring a more cohesive and emotionally resonant video.

**4. Ethical Concerns and Content Manipulation**

As AI continues to rise in video editing, there are increasing concerns about the ethical implications, particularly regarding content manipulation. AI has the potential to create "deepfake" videos or alter content in ways that can mislead audiences or contribute to the spread of misinformation. The misuse of AI in video editing poses serious risks to privacy, security, and the credibility of video content, especially in an era where fake videos can easily go viral on social media. To address these concerns, it is essential to prioritize the ethical use of AI and focus on the detection of manipulated content. Developing AI tools capable of identifying and flagging altered videos, along with ensuring transparency regarding AI-generated edits, will be critical in preventing the misuse of AI technology and maintaining trust in video content.

**5. Real-Time Processing and High-Quality Output**

Despite significant progress in AI for real-time video editing, the computational power required for high-quality, real-time edits, particularly with high-definition or 4K content, remains a major challenge. Many AI tools still struggle with handling large video files or demanding editing tasks, such as applying complex visual effects. This issue becomes especially problematic for editors working with high-quality content or in time-sensitive environments like live streaming or news broadcasting, where lag, delays, or reduced-quality results can disrupt the production process. To overcome this gap, it is crucial to optimize AI algorithms for real-time processing without compromising the output quality. AI systems should be designed to handle high-resolution video efficiently, ensuring faster processing speeds and minimal delay to meet the needs of demanding editing scenarios.

**2.3 OBJECTIVES**

The objectives of a project focused on AI-powered video editing typically aim to improve efficiency, creativity, and accessibility within the video editing process. One key objective is to automate repetitive editing tasks, allowing AI tools to handle time-consuming actions such as cutting scenes, colour grading, audio enhancement, and shot selection, thus reducing the workload for video editors and speeding up the process. Another goal is to enhance creativity by creating AI systems that assist editors in exploring new ideas, suggesting edits, applying filters, or generating visual effects that they may not have initially considered. Improving efficiency is also crucial, with AI tools designed to enable real-time editing with minimal delays, helping editors instantly preview changes and make adjustments efficiently, especially in fast-paced environments. AI can also support personalized content creation by adapting videos based on audience preferences, adjusting factors like length, style, or format to suit different demographics or platforms such as YouTube or Instagram.

Moreover, AI-powered tools can facilitate accessibility for non-professionals, enabling individuals without professional skills to create high-quality content, thereby democratizing video creation. Real-time feedback is another important objective, with AI systems that offer insights on emotional tone, pacing, and narrative flow to ensure the video conveys the intended message effectively. Consistency and branding can be maintained by AI tools that ensure a uniform look and style throughout a video project, which is especially beneficial for brands or content creators.

**2.4 PROBLEM STATEMENT**

The problem statement for a project focused on AI-powered video editing revolves around addressing the challenges and inefficiencies in traditional video editing processes while leveraging AI to enhance creativity, efficiency, and accessibility. Video editing is a time-consuming, process that requires skilled professionals to perform repetitive tasks such as scene cuts, colour correction, audio enhancement, and visual effects. These tasks take considerable time, especially when dealing with large volumes of footage, high-definition videos, and complex editing requirements. Furthermore, traditional editing workflows are often limited in terms of creativity, as human editors may struggle to keep up with the demand for quick edits, while AI tools lack deep contextual understanding and creative decision-making abilities that match human judgment.

Moreover, the growing need for personalized content and cross-platform optimization in today’s digital age presents additional challenges. Creators and brands require video content tailored to specific audiences, platforms, and devices, which can be difficult to achieve manually. Additionally, non-professional video editors often find it difficult to navigate the complexities of high-quality editing tools, making the process inaccessible to many potential content creators.

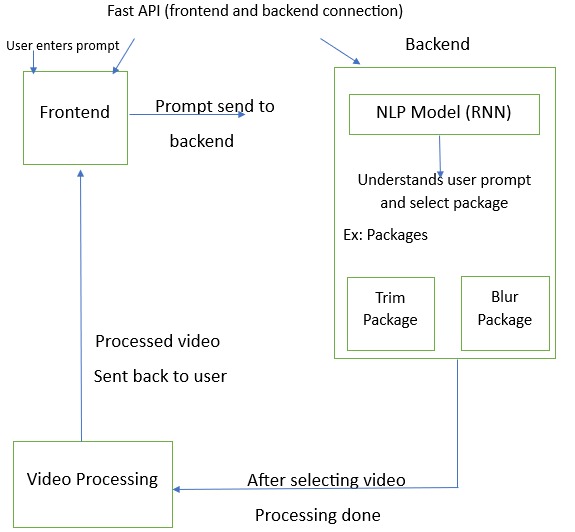
Thus, the problem lies in the need to develop AI-driven solutions that automate repetitive tasks, optimize the video editing process, provide creative assistance, and ensure accessibility and ethical standards, all while maintaining high-quality results and meeting the diverse needs of modern content creators across various platforms.

**2.5 PROJECT PLAN**

The project plan for an AI-powered video editing tool aims to enhance creativity, efficiency, and accessibility throughout the video editing process. The project begins with initialization and planning, where the scope, resources, timeline, and objectives are defined. Key tasks in this phase include stakeholder meetings to gather requirements, determining the project budget and technology stack, and setting clear milestones. The research and development phase follows, involving a thorough literature review, data collection, and the development of AI models for scene segmentation, color grading, audio enhancement, and shot selection. The AI models will be trained using diverse video data to ensure generalization across various formats.

The integration and testing phase involves combining the AI models with the platform interface and conducting rigorous functionality, usability, and performance testing. Cross-platform compatibility is also tested to ensure smooth operation on various devices. Once the system is fully functional, the deployment and launch phase begins, where the tool is made available to end-users through digital channels, with customer support systems established to assist users.

The project budget will account for development costs (AI model training, software development, user interface design), technology costs (cloud infrastructure, storage), marketing and launch expenses, and ongoing maintenance. The goal is to deliver a comprehensive AI-powered video editing solution that automates repetitive tasks, enhances creative capabilities, optimizes video quality, and provides an accessible and ethical platform for a wide range of users.



2.6 Flow Diagram

**3. TECHNICAL SPECIFICATION**

Requirement analysis for an AI-powered video editing tool is a critical phase in the development process. It involves understanding the needs of stakeholders, identifying the functional and non-functional requirements, and determining the technical specifications for the project.

**3.1. REQUIREMENTS**

**3.1.1 Functional Requirements**

**a) Automation of Editing Tasks**

* **Scene Detection & Segmentation**: AI should be capable of detecting and segmenting different scenes automatically based on visual and audio cues, allowing editors to quickly navigate raw footage.
* **Colour Grading & Correction**: The tool should apply automatic colour correction and grading to enhance the visual appeal of videos, maintaining consistency and improving aesthetics.
* **Audio Enhancement**: AI should be able to clean up background noise, adjust levels, and improve audio quality without manual intervention.
* **Shot Selection**: AI should analyze footage and automatically select the best shots based on predefined criteria such as visual quality, clarity, and composition.
* **Real-Time Editing**: The tool should allow real-time video editing with instant previews and changes, ensuring efficiency in workflows, especially for live events or time-sensitive projects.

**b) Creative Assistance**

* **AI-Driven Suggestions**: The system should provide intelligent suggestions for cuts, transitions, filters, and effects based on the content of the video.
* **Personalized Content Adjustments**: AI should adapt the content for specific audiences by adjusting video length, format, or style. For example, AI should help reformat videos for platforms like YouTube, Instagram, or TikTok.

**c) Multilingual Support**

* **Subtitles and Translation**: The tool should be able to generate accurate subtitles and translations for videos in multiple languages, increasing accessibility and global reach.

**d) Content Recognition and Categorization**

* **Face, Object, and Scene Recognition**: AI must recognize faces, objects, and scenes to automatically tag and organize video content. This will facilitate faster searching and organization within the tool.

**3.1.2. Non-Functional Requirements**

**a) Performance**

* **Real-Time Editing with Low Latency**: The tool should support fast processing, especially for high-definition (HD) and 4K content, ensuring that edits can be made in real time without noticeable delays.
* **Scalability**: The system should be able to handle large volumes of video data, including high-definition and long-form content, while maintaining performance.

**b) Usability**

* **User-Friendly Interface**: The system should provide an intuitive interface for both professional editors and non-experts, ensuring that complex editing tasks can be done with minimal effort.
* **Customization Options**: The tool should allow users to customize settings, workflows, and editing preferences based on their specific needs, offering flexibility for different types of projects.

**c) Reliability**

* **Error-Free Editing**: The AI-powered system must be reliable and deliver high-quality results without errors, especially in critical tasks such as color grading and audio enhancement.
* **High Availability**: The tool should be available with minimal downtime, especially for cloud-based platforms, to ensure that users can rely on it for their ongoing projects.

**d) Security and Privacy**

* **Data Protection**: As video content can contain sensitive information, the system should have strong data encryption, secure file storage, and privacy policies to protect user data and prevent unauthorized access.
* **Content Integrity**: The tool should prevent unauthorized alterations of video content and provide traceability for changes made, especially in contexts where ethical concerns (such as deepfakes) are involved.

**3.2. FEASIBILITY STUDY**

**3.2.1. Technical Feasibility**

The integration of Artificial Intelligence (AI) into video editing is technically highly feasible, largely due to advancements in core AI technologies such as natural language processing (NLP). These fields collectively enable automation, accuracy, and scalability in video editing processes that were traditionally labor-intensive and time-consuming.

Modern AI-based video editing platforms use deep learning models to analyze raw video footage and identify critical components such as faces, objects, and distinct scenes. These models facilitate automated tasks including scene detection, background blurring, object tracking, and motion stabilization. These platforms offer technically complex features such as automatic captioning, background noise reduction, lip-syncing, colour correction, and summarization, all presented in a user-friendly way for both beginners and professionals.

Natural language processing plays an equally important role by allowing AI systems to transcribe speech, generate subtitles, summarize spoken content, and even assist in creating scripts. Text-to-speech (TTS) and speech-to-text (STT) technologies further automate tasks such as voiceovers and closed captioning, drastically reducing manual input.

Compatibility with existing video editing ecosystems enhances the practicality of AI tools. Many AI solutions are available as plug-ins or integrations with popular non-linear editing systems (NLEs) like Final Cut Pro. This means that professional editors can incorporate AI features into their established workflows without the need for significant retraining or infrastructure changes.

The technical environment for AI-driven video editing is well-developed and continues to advance. The combination of powerful cloud infrastructure, mature AI models, and practical software tools demonstrates that AI is not only technically feasible but also ready for widespread use in the video editing industry. As the technology evolves, it is expected to unlock even greater efficiency, accuracy, and creative potential, making high-quality video production more accessible and scalable across industries and skill levels.

**3.2.2. Economic Feasibility**

The economic feasibility of integrating Artificial Intelligence (AI) into video editing is highly advantageous, offering notable benefits in terms of cost reduction, operational efficiency, and scalability for both individuals and organizations. AI-powered video editing solutions automate repetitive and time-consuming tasks such as trimming scenes, generating subtitles, removing background noise, and applying visual effects or transitions. This automation significantly reduces the need for manual labour and post-production efforts, leading to lower overall production costs.

A major economic advantage lies in the considerable reduction in editing time. Traditional video editing is resource-intensive and demands skilled professionals, especially for complex or high-resolution projects. With AI, much of this process can be completed automatically and quickly. This speeds up content delivery and enables editors to focus on higher-level creative decisions. As a result, industries like digital marketing, education, media, and social media benefit from faster, more affordable content generation, helping them keep up with high production demands.

Although initial costs such as licensing fees, training, or custom AI development may be required, these are generally offset by long-term savings. Increased productivity, reduced labour, and faster turnaround ensure a high return on investment (ROI) over time.

AI in video editing offers a highly economical solution for content creation. By minimizing labour, reducing costs, and streamlining production, AI enables users from individual creators to large corporations to produce high-quality video content efficiently and affordably. This makes AI a viable and smart economic choice in the evolving digital media landscape.

**3.2.3. Social Feasibility**

The social feasibility of using AI in video editing is highly favourable, given the increasing reliance on digital content across communication, education, entertainment, and marketing sectors. As society becomes more digitally engaged, there is a growing need for tools that support fast, accessible, and creative content creation needs that AI driven video editing tools are well-equipped to meet.

AI in video editing democratizes content creation by removing technical and financial barriers. Traditionally, professional video editing required costly software and skilled professionals, limiting access to high-quality production. This opens opportunities for a broader range of individuals including students, freelancers, educators, and small business owners to participate in digital storytelling and online content creation.

Inclusivity is another major benefit. AI tools offer features like automatic subtitle generation, text-to-speech (TTS), and speech-to-text (STT), which make video content accessible to people with hearing or visual impairments and to those speaking different languages. This promotes digital inclusiveness and aligns with global efforts toward equal media access for all audiences.

Additionally, AI supports the fast-paced nature of social media by enabling rapid editing and content production. This helps creators respond swiftly to current trends, events, and audience feedback, increasing engagement and interaction. Influencers, digital marketers, and educators benefit from the ability to regularly produce high-quality, relevant content without extensive editing knowledge or time investment.

AI-enhanced editing also improves the overall quality of user-generated content. Cleaner visuals, better audio, and coherent storytelling elevate the standard of online communication, making information more compelling and professional. This is particularly impactful as online content plays a growing role in shaping public opinion, education, and cultural narratives.

However, the adoption of AI in video editing is not without social concerns. Potential job displacement for traditional editors, misuse of AI-generated content, and ethical issues regarding transparency must be addressed.

Overall, the social impact of AI in video editing is largely positive. It fosters creativity, expands accessibility, and enhances engagement in the digital space. As digital literacy increases and AI becomes more integrated into daily life, the societal acceptance of AI video editing is expected to grow, making it a socially viable and valuable innovation for diverse communities.

**3.3. SYSTEM SPECIFICATION**

**3.3.1 Hardware Specification**

For a video editing project using AI, having the right hardware is crucial to handle the demanding tasks involved. A fast processor (CPU), such as an Intel i5 or AMD Ryzen 7, is essential for managing both video editing and AI computations efficiently. A graphics card (GPU) like the Intel Iris Xe is necessary to speed up rendering and boost AI performance, especially for tasks like real-time video processing. RAM (Memory) is also important, as video editing with high-resolution files can be very memory-intensive. A minimum of 8GB of RAM is recommended, with 32GB being ideal for smooth performance. For storage, a solid-state drive (SSD) with at least 512GB is recommended to store video files and software, and adding an extra 1TB of storage will help with larger video files and AI models. A high-resolution display (1080p ) is vital for clear video details during editing, and a larger screen or dual-monitor setup can help improve workflow efficiency. Lastly, since video editing and AI tasks can generate a lot of heat, a cooling system is necessary to prevent the system from overheating during extended use.

**3.3.2. Software Specification**

The libraries you mentioned MoviePy, Pygame, Librosa, Pillow (PIL), and OpenCV (cv2)—are essential tools for audio and video manipulation in Python. MoviePy is a versatile video editing library that allows you to cut, concatenate, resize, add effects, and manipulate both video and audio. It can also create animations, add text, and handle multiple formats. Pygame, primarily used for game development, also supports video playback, sound mixing, and media manipulation for interactive applications. Librosa is a specialized library for audio analysis and manipulation, perfect for tasks like pitch shifting, tempo changes, and spectral analysis. Pillow (formerly PIL) focuses on image processing and can be used to manipulate video frames, apply filters, resize, crop, and convert image formats. Lastly, OpenCV is a powerful computer vision library, ideal for video analysis, real-time frame manipulation, object detection, motion tracking, and video capture, making it widely used in video processing. Together, these libraries provide a comprehensive toolkit for multimedia manipulation in Python.

**4. DESIGN APPROACH AND DETAILS**

The system design for an AI-powered video editing tool focuses on creating a robust and efficient platform that automates video editing tasks, enhances creativity, and ensures high-quality output. The design encompasses the architecture, components, and interactions required to integrate AI functionalities with traditional video editing workflows. Below is an outline of the system design for such a tool:

**4.1. System Architecture Overview**

The system will be composed of multiple layers that work seamlessly to provide a smooth user experience. The architecture can be broken into the following main components:

**a) Frontend (User Interface)**

* **Dashboard**: Users will interact with an intuitive dashboard where they can upload video files, select editing templates, and view AI-driven suggestions. This interface will also include controls for video preview, editing tools, and AI recommendations.
* **Editing Interface**: A timeline-based interface for traditional video editing tasks such as cutting, trimming, and applying transitions. AI-driven suggestions will be visible as additional options alongside manual editing controls.
* **Real-Time Preview**: Users will have the ability to preview AI-generated edits in real time, allowing for fast adjustments and instant feedback.
* **AI Recommendations**: The system will provide suggestions for color correction, audio enhancement, and scene cuts based on the video’s content. AI models will offer suggestions, but users can either accept or modify them as needed.
* **Multilingual Support**: Users will be able to add AI-generated subtitles and translations to reach a global audience.

**b) Backend (Server-Side)**

* **Video Processing Engine**: This component will handle all video processing tasks, such as cutting, encoding, transcoding, and rendering. It will manage the communication between the frontend and the AI models.
* **AI Models**: Machine learning models will power the automation of video editing tasks. These include:
  + **Computer Vision Models**: For scene segmentation, object detection, facial recognition, and visual effects.
  + **Natural Language Processing (NLP)**: For automatic subtitle generation and sentiment analysis.
  + **Audio Enhancement Models**: To clean up background noise and optimize audio levels.
  + **Creative Models**: For suggesting creative edits based on style and user preferences.
* **Cloud Infrastructure**: AI models and video processing tasks will run on cloud platforms like AWS, Google Cloud, or Azure to handle high-demand tasks and ensure scalability. This also enables high-speed video rendering and storage.

**c) Database Layer**

* **User Data**: Store user profiles, editing preferences, and project metadata such as video clips, timestamps, edits made, and AI recommendations.
* **Video Files and Metadata**: Store raw and processed video files along with metadata, including time codes, scene changes, and detected objects.
* **AI Model Storage**: Store trained AI models and user-specific learning data (such as user preferences and edits) for future optimizations.

**4.2. AI Driven Features**

The AI models in the system will handle several critical aspects of video editing, making the process faster, more efficient, and creative. Some of the key AI functionalities include:

**a) Automated Editing Tasks**

* **Scene Detection**: Automatically segment the video into scenes based on visual cues (e.g., scene changes or object transitions).
* **Shot Selection**: AI can recommend the best shots based on composition, lighting, focus, and framing.
* **Colour Grading**: AI models will analyze the footage and apply automatic colour corrections based on industry standards or predefined styles.
* **Audio Enhancement**: Background noise reduction, volume levelling, and audio equalization for improved sound quality.
* **Stabilization**: AI algorithms will detect and correct shaky footage, ensuring smooth transitions.

**b) Creative Assistance**

* **Visual Effects**: Based on the style of the video and user preferences, AI can suggest or apply visual effects like filters, transitions, and motion graphics.
* **Personalized Recommendations**: The AI will analyze past edits and suggest improvements or creative effects based on user behavior and preferences.
* **Content Personalization**: The AI can tailor video edits based on target platforms (e.g., YouTube, Instagram, TikTok) by adjusting length, format, and aspect ratio.

**c) Multilingual and Accessibility Features**

* **Automatic Subtitles**: AI will automatically generate subtitles in the video’s original language using speech recognition and natural language processing.
* **Translation**: The AI will offer video translation services, enabling subtitles or voiceover translation into multiple languages.
* **Accessibility Features**: The tool can provide features for hearing-impaired viewers, such as sign language integration or automated captioning.

**4.3. Real Time Editing and Feedback**

**a) Real Time Editing:**

* AI models will be optimized for real-time processing, enabling fast video editing, especially for live-streamed events or time-sensitive productions.
* **Editing Feedback**: AI will provide suggestions and real-time feedback on video pacing, emotional tone, and narrative flow, helping the editor to refine the video for better viewer engagement.

**b) Instant Video Rendering:**

* With AI, video rendering will be accelerated, allowing editors to see immediate changes as they apply edits.
* High-definition and 4K videos will be handled effectively, maintaining quality while rendering faster than traditional methods.

**4.4. Video Workflow**

The video editing workflow involves multiple stages, each of which leverages AI capabilities to streamline tasks:

**a) Video Ingestion**

* **Upload**: Users upload raw video files to the system. The AI analyzes the footage for scene detection and basic editing suggestions.
* **Metadata Extraction**: AI will extract metadata, including objects, scenes, audio, and spoken words, which will later be used for editing and recommendations.

**b) Initial Editing**

* **Automatic Cuts**: The system will offer automatic scene segmentation and shot selection, which users can modify or fine-tune based on their preferences.
* **AI-Generated Enhancements**: The AI applies color correction, audio improvement, and stabilization to the raw footage.

**c) Creative Exploration**

* **Effect Suggestions**: AI recommends visual effects, transitions, and color grading options based on the video’s content, style, and intended emotional tone.
* **User Collaboration**: Users can either accept or reject AI-driven edits and further customize them using traditional editing tools.

**d) Content Personalization**

* **Platform Adaptation**: AI automatically adjusts the video’s format, aspect ratio, and resolution according to specific platform requirements (e.g., Instagram Stories vs. YouTube videos).

**e) Rendering and Exporting**

* The final video is rendered, ensuring high quality while optimizing for file size and platform-specific specifications.
* The AI provides options for generating multiple versions of the video optimized for different platforms (YouTube, Facebook, TikTok, etc.).

**5. REFERENCES**

[1] <https://arxiv.org/abs/1605.08110>

[2]<https://www.researchgate.net/publication/375833832_Video_Editing_Tools_with_Artificial_Intelligence>

[3]<https://blog.eklipse.gg/tools/7-free-ai-video-editing-software-no-watermark.html>

[4] <https://www.coursera.org/articles/how-to-use-ai-to-create-videos>

[5] <https://penji.co/ai-for-video-editing/>

[6] <https://teachable.com/blog/ai-video-editing-tools>

[7] <https://vitrina.ai/blog/ai-in-video-editing/>