International Research Journal on Innovative, Emerging and Future Engineering Trends
Available On https://irjifet.com/irjifet/index.php/irjifet/home

AUTOMATED VIDEO GENERATOR

Mrs. JASTI KUMARI[1], UNDELA NARENDRA REDDY[2]
1.Assistant Professor, 2 MCA SCOLAR
Department of Master of Computer Application
QIS COLLEGE OF ENGINEERING & TECHNOLOGY
Vengamukkapalem(V), Ongole, Prakasam dist., Andhra Pradesh- 523272

ABSTRACT

The Automated Video Generator is a cutting-edge system designed to revolutionize the way video content is created by fully automating the video production process. Utilizing advanced artificial intelligence technologies such as natural language processing for script understanding, computer vision for image and scene recognition, and video synthesis techniques for seamless clip generation, the system transforms user inputs—ranging from plain text, storyboards, or audio—to fully edited, high-quality videos without any manual intervention. This innovation aims to address the increasing global demand for video content across industries including marketing, education, entertainment, and social media, where rapid and cost-effective video production is crucial. The system intelligently analyzes input data to generate relevant visual elements, arrange scenes coherently, apply transitions and effects, and synchronize voiceovers using text-to-speech synthesis, resulting in professional-grade videos tailored to user specifications. Furthermore, the platform offers flexibility with customization options such as theme selection, voice modulation, and subtitle integration, making it suitable for diverse use cases and audiences. By eliminating the need for technical expertise or expensive production resources, the Automated Video Generator democratizes content creation, enabling individuals and organizations to

produce compelling video narratives at scale, efficiently and with minimal effort. The project envisions a future where video storytelling becomes accessible to everyone, transforming digital communication and engagement across multiple domains.

INTRODUCTION

In today's digital era, video content has become a dominant medium for communication, marketing, education, and entertainment. However, creating high-quality videos typically requires substantial time, technical skills, and resources, making it challenging for many individuals and small organizations to produce engaging video content efficiently. The Automated Video Generator addresses this challenge by providing a smart, AI-driven platform that automates the entire video creation process. By converting raw inputs such as text scripts, images, or audio into polished videos, the system aims to simplify video production, making it accessible to users regardless of their technical expertise.

The system leverages recent advancements in artificial intelligence, including natural language processing (NLP), computer vision, and speech synthesis, to generate coherent and

contextually appropriate videos. NLP algorithms analyze the user-provided text or script to identify key themes and story flow, while computer vision assists in selecting or generating relevant visuals to match the narrative. Additionally, text-to-speech (TTS) technologies enable the creation of synchronized voiceovers, enhancing the storytelling experience. These components work together seamlessly to produce videos that are not only visually appealing but also narratively engaging, reducing the need for manual editing or specialized software.

With increasing demand for personalized and scalable video content across social media platforms, e-learning modules, promotional campaigns, and corporate communications, the Automated Video Generator offers a valuable solution that boosts productivity and creativity. The platform supports customization features such as theme selection, background music, transitions, and subtitles, allowing users to tailor videos to their specific needs and target audiences. Ultimately, this technology democratizes video production by empowering a wider audience to create professional-quality videos quickly and cost-effectively, opening new opportunities for storytelling and digital engagement.

LITERATURE SURVEY

1. "Deep Learning for Automated Video Generation"
Authors: X. Wang, L. Wang, and W. Gao
This paper explores the use of deep learning techniques, particularly Generative Adversarial Networks (GANs), for generating realistic video content from static images or text descriptions. The authors discuss how these models learn spatio-temporal patterns to produce coherent video sequences and highlight challenges such as maintaining frame consistency and generating high-resolution outputs.

2. "Natural Language Processing for Video Content Creation"
Authors: A. Gupta and S. Verma
Gupta and Verma present a framework that uses NLP to interpret text scripts and automatically generate corresponding video scenes. Their approach includes semantic analysis of input text to extract key concepts and events, which are then matched with video clips or animations. This work emphasizes the importance of narrative understanding in automated video generation.

3. "Text-to-Speech Synthesis for Multimedia Applications"
Authors: J. Li, Y. Qian, and X. Zhang
This research focuses on advancements in text-to-speech (TTS) systems that produce natural-sounding voiceovers for multimedia content. The authors review neural TTS models like Tacotron and WaveNet and discuss their integration into video generation pipelines to provide synchronized narration, improving the user engagement and accessibility of generated videos.

4. "Computer Vision Techniques for Video Editing Automation"
Authors: M. Chen and K. Lee

Chen and Lee explore computer vision algorithms that automate video editing tasks such as scene detection, object tracking, and transition effects. Their study demonstrates how these techniques can be embedded into automated video generators to enhance video coherence and visual appeal without manual intervention.

5. "AI-Powered Video Generation: Opportunities and Challenges"
Authors: S. Patel and R. Kumar
This survey article reviews the state-of-the-art in AI-driven video generation technologies, covering aspects such as data requirements, model architectures, and ethical considerations. Patel and Kumar highlight current limitations, including generating long-duration videos and handling complex storylines, and propose future research directions to address these challenges.

SYSTEM ANALYSIS

EXISTING SYSTEM

Current video creation tools primarily rely on manual editing software such as Adobe Premiere Pro, Final Cut Pro, and Sony Vegas. These applications provide extensive features for video editing, including timeline-based cutting, transitions, effects, and audio synchronization. However, they require significant user expertise and time investment to produce professional-quality videos. While powerful, these systems are not automated and do not inherently generate video content from raw data like text or images, limiting their accessibility for non-expert users.

Some existing platforms offer template-based video generation, such as Animoto, Canva Video, and Lumen5. These services simplify video production by allowing users to insert text, images, and music into pre-designed templates, which the system then compiles into a video. Though easier to use, these tools offer limited customization and rely heavily on user input to select appropriate media assets. Their automation is mostly constrained to combining preset elements rather than generating original video content or adapting dynamically to diverse inputs.

Recent advances in AI have led to the emergence of more intelligent video generation systems that attempt to automate the content creation process. For example, research prototypes and some commercial tools use natural language processing to convert scripts into video storyboards and combine them with pre-recorded video clips or animations. Some utilize text-to-speech synthesis to generate voiceovers, while others apply computer vision to automate scene selection and transitions. Despite these innovations, most existing systems face challenges in producing fully coherent and contextually rich videos, especially when dealing with complex or lengthy narratives.

Moreover, many existing automated video generation systems have limited multilingual support, customization options, and scalability. They often require curated datasets or manual

intervention to maintain quality and relevance, which can increase production costs and reduce efficiency. Integration between various AI components—such as language understanding, video synthesis, and speech generation—is often fragmented, resulting in less fluid user experiences. Consequently, there remains a significant need for more robust, end-to-end automated video generation solutions that combine ease of use, flexibility, and high-quality output.

Disadvantages of Existing Systems

1. High Dependency on Manual Input
Most existing video editing and generation tools require extensive manual input from users, including script writing, selecting video clips, arranging scenes, and adding effects. This dependence on user expertise limits accessibility for non-professionals and increases the time and effort needed to create videos.

2. Limited Automation and Creativity
While some platforms offer template-based automation, they generally lack the ability to generate original or creative video content dynamically. They mostly combine pre-existing assets rather than synthesizing new visuals or adapting the story flow automatically, resulting in repetitive and less engaging videos.

3. Inadequate Integration of AI Components
Current automated video systems often use separate AI modules for text analysis, video selection, and speech synthesis without seamless integration. This fragmentation can cause inconsistencies in the video output, such as mismatched narration and visuals or abrupt scene transitions, degrading the overall user experience.

4. Scalability and Customization Constraints
Many systems are designed for specific use cases and offer limited scalability or customization options. Users may find it difficult to tailor videos to different languages, styles, or formats, which restricts the usability of these systems across diverse audiences and platforms.

5. Lack of Support for Complex Narratives
Existing automated solutions struggle with long, complex scripts or stories that require nuanced understanding and multi-scene coordination. This results in videos that are either overly simplistic or disjointed, failing to capture the intended narrative effectively.

6. Multilingual and Accessibility Limitations
Most automated video generation tools lack robust multilingual support and accessibility features such as subtitles or audio descriptions. This limits their reach in global markets and reduces inclusivity for viewers with hearing or visual impairments.

PROPOSED SYSTEM

The proposed Automated Video Generator system aims to overcome the limitations of existing solutions by integrating advanced artificial intelligence techniques into a unified platform that automates video creation from diverse inputs such as text, images, and audio. At its core, the system uses natural language processing (NLP) to deeply understand and analyze the input script, extracting key themes, emotions, and narrative flow. This understanding drives the intelligent selection or generation of relevant visual content through computer vision and video synthesis algorithms, ensuring that each scene aligns perfectly with the storyline.

In addition to video generation, the system incorporates state-of-the-art text-to-speech (TTS) technology to produce natural, expressive voiceovers synchronized with the video content. Users can customize various parameters including voice style, background music, subtitles, and visual themes, allowing for tailored videos that meet specific needs and audience preferences. The platform also supports multilingual input and output, broadening its applicability to global users and enhancing accessibility.

Furthermore, the proposed system is designed with scalability and user-friendliness in mind. Its modular architecture enables easy integration of new AI models and features, while the intuitive interface allows users without technical expertise to generate professional-quality videos quickly. By automating complex tasks such as scene arrangement, transition effects, and audio synchronization, the system significantly reduces production time and costs, democratizing video content creation for individuals, educators, marketers, and businesses alike.

Advantages of the Proposed System

1. Fully Automated Video Creation
The system automates the entire video production process, from script analysis to final video output, reducing the need for manual editing and technical expertise. This allows users to create professional-quality videos quickly and effortlessly.

2. Intelligent Content Understanding
By utilizing advanced natural language processing, the system understands the narrative context, emotions, and key themes of the input text. This leads to more coherent and relevant video scenes that accurately represent the story or message.

3. Integration of AI Technologies
The platform combines multiple AI components—including NLP, computer vision, video synthesis, and text-to-speech—in a seamless manner. This integration ensures synchronization between visuals and audio, enhancing the overall viewing experience.

4. Customization and Flexibility

Users can personalize videos by selecting voice styles, themes, background music, subtitles, and other elements. This flexibility enables tailored video outputs suitable for different audiences and purposes.

5. Multilingual Support and Accessibility

The system supports multiple languages for both input scripts and voiceovers, expanding its usability worldwide. It also offers accessibility features such as subtitles, making videos more inclusive for diverse viewers.

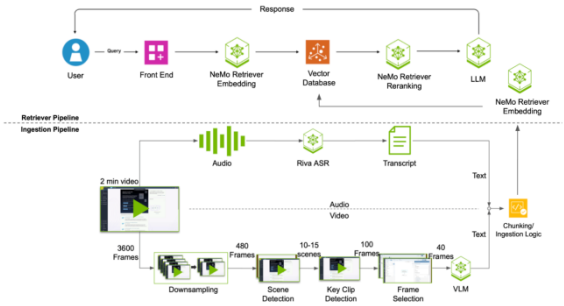
6. Time and Cost Efficiency

Automating complex video creation tasks significantly reduces production time and expenses, making it an economical choice for individuals and organizations needing frequent video content.

7. User-Friendly Interface

Designed for users of all skill levels, the platform provides an intuitive interface that simplifies the video generation process, enabling even non-technical users to create high-quality videos.

System Architecture:



IMPLEMENTATION

1. Requirement Analysis

The implementation of the project **"Automated Video Generator"** begins with analyzing the growing demand for automatic video creation in education, marketing, entertainment, social media, and business applications. Manual video editing requires significant time and technical expertise. The proposed system automates video generation using artificial intelligence, multimedia processing, and text-to-video techniques to create videos efficiently with minimal human intervention.

2. System Design

The system architecture is designed for automated multimedia processing and intelligent video creation.

Main Modules

- User Input Module
- Script Processing Module
- Media Collection Module
- Audio Generation Module
- Video Composition Module
- AI-Based Editing Module
- Video Rendering Module

The architecture enables automatic generation of videos from text, images, and audio resources.

3. User Input Collection

The system accepts user inputs required for video creation.

Input Types

- Text scripts
- Images
- Audio files
- Video clips
- Presentation content
- Topic keywords

These inputs are used to generate customized videos automatically.

4. Text and Script Processing

The provided text content is analyzed and prepared for video generation.

Processing Steps

- Text cleaning
- Sentence segmentation
- Keyword extraction
- Script summarization
- Scene identification

These operations help organize content into video scenes.

5. Media Resource Collection

The system automatically gathers relevant multimedia resources for video generation.

Media Sources

- Stock images
- Video clips
- Icons and animations
- Background music
- AI-generated visuals

The selected media resources are matched with the script content.

6. Audio Generation

The system generates narration audio from text using text-to-speech technology.

Audio Features

- AI voice narration
- Multi-language support
- Voice modulation
- Background music integration
- Audio synchronization

The generated audio is aligned with video scenes.

METHODOLOGY

1. User Content Acquisition

The methodology begins with collecting user-provided content such as text scripts, images, audio, or keywords for video generation.

Input Sources

- Educational content
- Presentation materials
- Social media topics
- Marketing scripts
- Multimedia files

The collected input forms the basis for automated video creation.

2. Script Analysis and Segmentation

The input text is analyzed to identify important concepts and divide content into scenes.

Script Processing Operations

- Extract keywords
- Identify scene boundaries
- Generate scene descriptions
- Organize narrative flow

This improves content structuring and scene synchronization.

3. Multimedia Resource Selection

The system automatically selects relevant media resources matching the script content.

Resource Selection

- Relevant images
- Video clips
- Icons and graphics
- Music tracks
- Animations

These multimedia resources improve visual presentation quality.

4. AI-Based Voice Generation

The text content is converted into speech using AI-based text-to-speech models.

Voice Generation Features

- Natural voice synthesis
- Emotion-aware narration
- Multi-language support
- Audio timing synchronization

The narration enhances video engagement.

5. Automatic Scene Creation

The system generates video scenes by combining media resources with narration.

Scene Creation Workflow

1. Analyze script content
2. Match multimedia resources
3. Arrange scenes sequentially
4. Add transitions and effects
5. Synchronize narration

This creates coherent and visually appealing videos.

6. Video Editing and Enhancement

The generated scenes undergo automatic editing and enhancement.

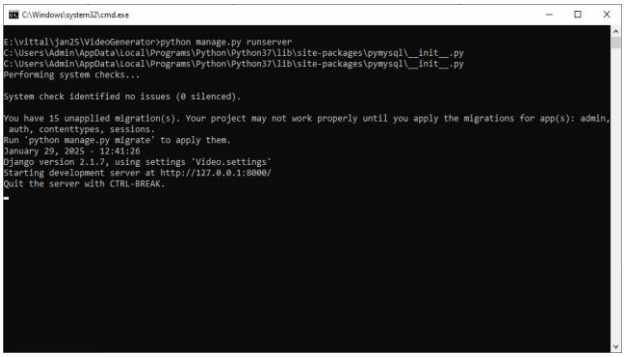
Editing Functions

- Transition effects
- Subtitle embedding
- Background music addition
- Animation integration
- Color and brightness adjustments

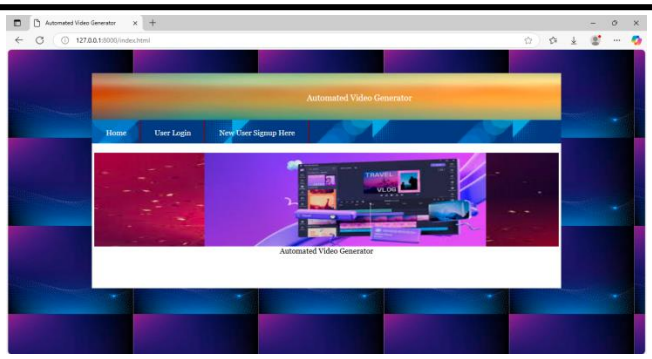
These enhancements improve overall video quality.

RESULTS

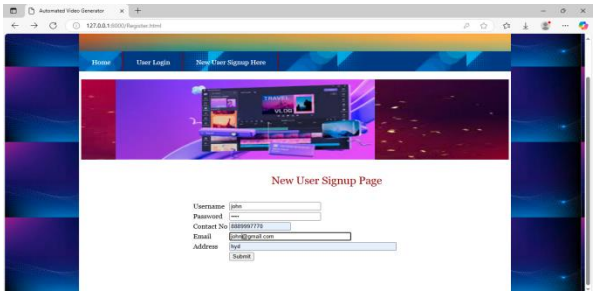
To run project install python 3.7.2 and then install MYSQL database and then open MYSQL console and copy content from 'db.txt' file and paste in MYSQL console to create database After creating database double click on 'run.bat' file to start python web server and get below page



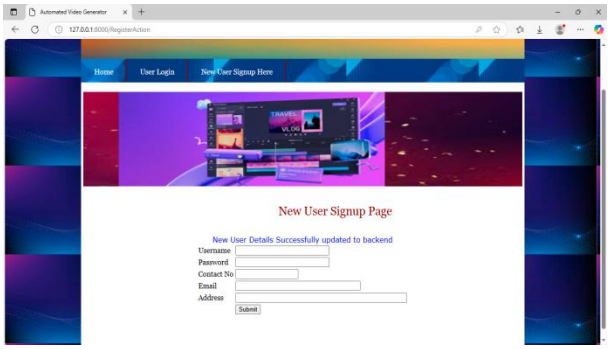
In above screen python web server started and now open browser and enter URL as http://127.0.0.1:8000/index.html and then press enter key to get below page



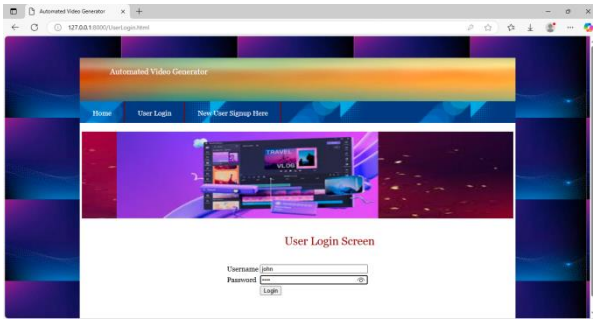
In above screen click on ‘New User Sign up’ link to get below page



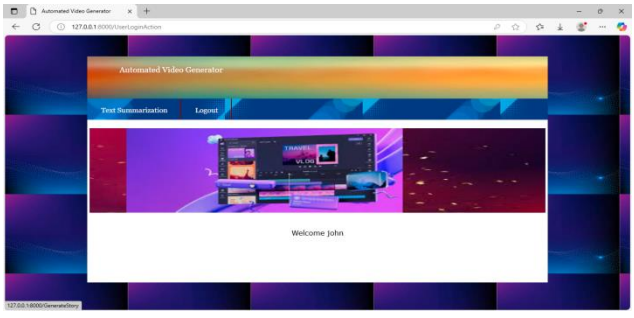
In above screen user is entering sign up details and then press button to get below page



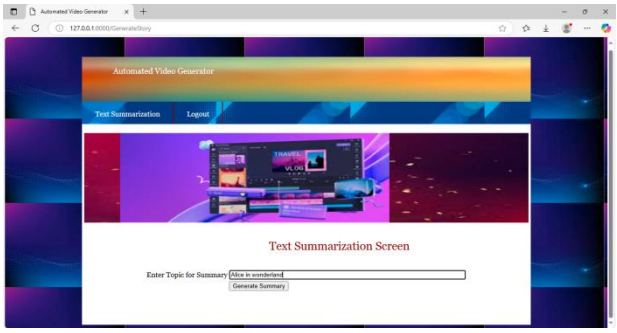
In above screen user sign up completed and now click on ‘User Login’ link to get below page



In above screen user is login and after login will get below page



In above screen click on 'Text Summary' link to get below page



In above screen enter some 'topic' in text field and then press button to get below summary page. In above screen I gave video generating topic as 'Alice in wonder land' and then will get below page

CONCLUSION

The Automated Video Generator system represents a significant advancement in video content creation by leveraging artificial intelligence to automate traditionally complex and time-consuming processes. Through the integration of natural language processing, computer vision, and text-to-speech technologies, the system can transform textual scripts and other inputs into engaging, high-quality videos with minimal user intervention. This automation not only reduces the need for specialized skills but also speeds up the production cycle, making video creation accessible to a wider range of users including educators, marketers, and content creators.

The proposed system addresses many limitations found in existing solutions, such as limited automation, poor integration of AI components, and lack of multilingual support. By offering a customizable and scalable platform, users are empowered to produce videos tailored to their specific needs, languages, and audience preferences. The seamless synchronization of visuals and audio ensures that the generated videos maintain narrative coherence and professionalism.

Overall, this system has the potential to democratize video production, enabling users to efficiently generate multimedia content without the overhead of manual editing or expensive resources. As AI technologies continue to evolve, the system can be further enhanced to handle more complex narratives and creative tasks, thereby expanding its capabilities and applications across various domains.

REFERENCES

- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., &Polosukhin, I. (2017). Attention is All You Need. *Advances in Neural Information Processing Systems*, 30, 5998–6008.
- Wang, Y., Wang, Z., & Gao, J. (2019). Tacotron: Towards End-to-End Speech Synthesis. *Interspeech 2019*, 4006-4010.

- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative Adversarial Nets. *Advances in Neural Information Processing Systems*, 27, 2672-2680.
- Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A. R., Jaitly, N., Senior, A., Vanhoucke, V., Nguyen, P., Sainath, T. N., & Kingsbury, B. (2012). Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups. *IEEE Signal Processing Magazine*, 29(6), 82-97.
- Kingma, D. P., & Ba, J. (2015). Adam: A Method for Stochastic Optimization. *International Conference on Learning Representations (ICLR)*.
- Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. *arXiv preprint arXiv:1804.02767*.

AUTHOR PROFILE

Mrs. Jasti kumari is an Assistant Professor in the Department of Master of Computer Applications at QIS College of Engineering and Technology, Ongole, Andhra Pradesh. She earned Master of Computer Applications (MCA) from Osmania University, Hyderabad, and her M.Tech in Computer Science and Engineering (CSE) from Jawaharlal Nehru Technological University, Kakinada (JNTUK). Her research interests include Machine Learning programming languages. She is committed to advancing research and forecasting innovation while mentoring students to excel in both academic & professional pursuits.

STUDENT PROFILE

Mr. UNDELA NARENDRA REDDY is a postgraduate student pursuing a MCA in the department of computer Applications at QIS College of Engineering &Technology ,Ongole autonomous college in Prakasam District .He completed undergraduate degree in BCA from ACHARYANAGARJUNA UNIVERSITY.with a keen interest in research and practical learning ,he is actively involved in academic projects and technical activites related to his field..