iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** IV **Month of publication:** April 2026

DOI:

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

VidMind AI: An AI-Based System for Automatic YouTube Video Summarization

M. Anitha[1], Nelluri Kalpana[2], Jagarlamudi Abhinaya[3], Ambati Mounika[4]
[1]Assistant Professor, Department of Computer Science and Engineering, Bapatla Women's Engineering College, Bapatla
[2, 3, 4]Department of Computer Science and Engineering, Bapatla Women's Engineering College, Bapatla

Abstract: The rapid rise in the number of online video sharing sites has led to an overwhelming number of multimedia content, making it challenging for users to efficiently access and comprehend the information they need. YouTube is one such popular video sharing site that contains millions of videos on various subjects, which requires a significant amount of time to watch and comprehend the video content. This paper proposes an intelligent system named VidMind AI that can automatically summarize the content of YouTube videos and obtain significant insights from the video transcripts. The proposed system utilizes the YouTube Transcript API to obtain the video transcripts and then applies the Natural Language Processing (NLP) technique to process the video transcripts. The proposed framework utilizes a large language model to perform abstractive summarization, TF-IDF keyword extraction to obtain significant keywords, and lexiconbased sentiment analysis to determine the emotional content of the video transcripts. The proposed system also includes a topic classification module and an interactive question-answering facility that can be used to obtain information from the video transcripts. The architecture is implemented using a Python Flask backend and a web-based interface using HTML, CSS, and JavaScript. The experimental results on different YouTube videos prove the effectiveness of the system in providing concise summaries, identifying important concepts, and increasing the efficiency of information retrieval. The proposed approach can contribute to the improvement of accessibility and productivity in understanding the main concepts of the video content without viewing the whole video.
Keywords: Video Summarization, Natural Language Processing, YouTube Transcript Analysis, Large Language Models, Sentiment Analysis, Information Retrieval.

I. INTRODUCTION

In recent times, online video platforms have emerged as one of the most significant tools for acquiring information, knowledge, and entertainment. However, of all these online video platforms, YouTube is regarded as one of the most popular platforms, which contains billions of videos from different genres, such as technology, education, news, and entertainment, to name a few. However, although this online video platform provides access to immense knowledge for its users, it also poses a major problem for its users, as it requires considerable time to view videos before determining whether the provided information is useful for their requirements or not. Moreover, as the number of videos is increasing at a rapid pace, it is becoming increasingly difficult to extract useful insights from videos. Automatic summarization has been identified as an effective tool in managing large volumes of text information. In the area of Natural Language Processing, summarization tools have been developed to reduce the length of text while preserving the essential concepts. Traditionally, summarization tools have been developed using the traditional method, where important concepts are extracted from the text in the form of sentences. Recently, with the advent of artificial intelligence tools, such as large language models and transformers, abstractive summarization tools have been developed, enabling the creation of concise summaries in natural language. Although several tools are available for summarizing text information, summarizing video information is a complex task, considering the audio, video, and text incorporated in the video. Most videos on YouTube are equipped with text information, also known as captions, enabling the application of Natural Language Processing tools to derive important concepts from the video. Through the analysis of text information, it is possible to develop tools for video summarization, identify important concepts, derive sentiment, and enable users to interact with the video using intelligent tools for answering questions.

To overcome these issues, this paper proposes the VidMind AI, which is an intelligent web-based system that can automatically summarize and analyze the YouTube video content using the video transcript data. The proposed system can take the video URL as input and process the video transcript using the multi-stage NLP pipeline. The proposed framework utilizes the large language model to generate the abstractive summary, TFIDF keyword extraction method to determine the significant keywords, sentiment analysis to determine the emotional tone of the content, and rule-based topic classification mechanism to classify the video content topic.

Additionally, the proposed VidMind AI system has the facility to interact with the user through the chatbot facility, which can answer the user questions based on the video content. The proposed VidMind AI system is implemented using the Python Flask backend and the HTML, CSS, and JavaScript frontend technologies. The proposed VidMind AI system can efficiently process the video content and provide the user with the video content summary and insights in an efficient way. The proposed VidMind AI system can efficiently summarize the video content and provide the user with the video content insights without watching the video content completely.

The main contributions of this work are summarized as follows:

1) Development of an automated system for summarizing YouTube videos using transcript analysis.
2) Integration of various NLP methods, such as abstractive summarization, keyword extraction, sentiment analysis, and topic classification.
3) Development of an interactive question-answering system for querying video content.
4) Development of a web-based system for improving accessibility and efficiency in video content analysis.

In conclusion, VidMind AI addresses the challenge of extracting useful information from large volumes of online video content. By applying Natural Language Processing techniques to video transcripts, the system generates concise summaries and analytical insights, enabling users to quickly understand the key ideas without watching the entire video.

II. LITERATURE SURVEY

Kini *et al.* [1] proposed an AI-based video summarization system integrating OpenAI Whisper for speech recognition and BART for abstractive summarization. TF-IDF and Named Entity Recognition identify key highlights and generate quizzes. Implemented with Flask–MongoDB on Render Cloud, it achieves 94–99% transcription accuracy, 5.4× compression, 4 min processing time for 10-min videos, and 99.94

Thanam *et al.* [2] proposed an NLP framework for YouTube video transcription and summarization using ASR with domain-specific language models. It applies extractive and abstractive summarization to enhance content accessibility and information retrieval from large-scale video content.

Kanagaraj *et al.* [3] proposed an AI-based video summarization framework using a Deep Belief Network (DBN) for concise summaries and a Radial Basis Function (RBF) model for improved retrieval. The DBN-RBF integration outperforms traditional methods in retrieval efficiency and informativeness. Nagaraju *et al.* [4] proposed an AI-driven framework for real-time YouTube content analysis, integrating sentiment analysis, trend detection, and predictive insights. Unlike traditional tools, it incorporates competitor benchmarking to support creators and digital marketers in decision-making and performance analysis.

Palivela *et al.* [5] proposed an AI framework transforming YouTube transcripts into educational resources using NLP techniques including summarization, translation, and sentiment analysis. Multilingual capabilities generate prompt sheets highlighting key concepts, improving engagement and accessibility for global learners.

Kalaiarasi *et al.* [6] proposed an AI-based YouTube analytics platform integrating modules for thumbnail generation, script generation, keyword extraction, and performance analysis. Using LLMs and trend analysis tools, it automates insights to help creators optimize content and improve visibility.

Kolhe *et al.* [7] investigated an LLM-based video summarization framework using LangChain, evaluating LLaMA and Lamini. LLaMA achieves higher quality while Lamini offers lower cost. Challenges include speech-to-text errors and model bias, motivating future multimodal research.

Kontostathis *et al.* [8] proposed an AI-based video summarization tool for social media, using summarization models and aspect ratio transformation to recommend key segments. An interactive interface enables user editing, ensuring summaries meet platform format and length requirements.

Anthati *et al.* [9] proposed a multilingual video summarization system integrating transcription, translation, and summarization. Supporting YouTube links and offline downloads, it provides transcripts, summaries, and translations, enhancing accessibility for educators, researchers, and content creators.

Vayadande *et al.* [10] reviewed AI-generated news video creation from text, covering web scraping, authentication, image retrieval, voice-over generation, and publishing. Challenges include source authenticity and visual selection, highlighting AI's potential in digital journalism.

Reddy *et al.* [11] proposed a Smart Video Assistant using Whisper for transcription and sentence embeddings in a vector database for search. It generates topic summaries and includes a chatbot for context-aware Q&A, enabling efficient information extraction from long videos.

Balushi *et al.* [12] proposed an AI-driven multi-modal synthesis system integrating PDF querying, Whisper-based transcription, and LLM-powered summarization with multilingual support. Using embedding models and vector databases, it achieves high accuracy for multi-source information retrieval.

Khan *et al.* [13] proposed an AI system detecting misogynistic hate speech in YouTube videos. Using 11,000+ Urdu transcripts transliterated to Roman Urdu, ML/DL models achieved up to 0.96 accuracy, demonstrating AI's potential for ethical content moderation.

Gautham and Somasundaram [14] proposed an AI system generating viral short-form clips from long videos. Integrating transcription, LLM-based engagement segment identification, and video processing for resizing/subtitles, it automates content creation for YouTube Shorts and Instagram Reels.

Varshini *et al.* [15] proposed a deep learning-based video summarization system for LMS using RNNs, BiLSTM, and attention mechanisms to identify keyframes. Trained on human summaries and evaluated with BLEU/ROUGE, it generates concise educational video summaries to improve learning efficiency.

III. PROPOSED SYSTEM / METHODOLOGY

The proposed system, VidMind AI, is designed to automatically summarize and analyze YouTube videos by processing their transcripts through a structured Natural Language Processing (NLP) pipeline. The system aims to reduce the time required for users to understand long video content by generating concise summaries and extracting meaningful insights. The architecture integrates transcript extraction, text processing, summarization, sentiment analysis, keyword extraction, topic detection, and an interactive question-answering module.

A. System Overview

The workflow of the proposed system begins when the user provides a YouTube video URL through the web interface. The system retrieves the transcript associated with the video and processes it using multiple NLP techniques to generate analytical outputs. These outputs include a summarized version of the video content, key insights, sentiment evaluation, keyword identification, and topic classification. Additionally, the system enables users to interact with the transcript using a question-answering interface.

B. Transcript Extraction

The first stage of the methodology involves retrieving the video transcript using the YouTube Transcript API. The API provides timestamped captions that represent the spoken content of the video. These captions are cleaned and combined to form a structured textual transcript. This transcript serves as the primary input for the subsequent NLP processing stages.

C. Text Preprocessing

Before analysis, the retrieved transcript is preprocessed to improve data quality. This step includes removing special characters, normalizing text formats, and eliminating redundant whitespace. The cleaned transcript provides structured input for subsequent NLP processing modules.

D. Abstractive Summarization

The processed transcript is passed to a large language model that generates an abstractive summary of the video content. Unlike extractive methods, abstractive summarization produces new sentences that capture the key ideas of the transcript in a concise and coherent manner. The generated summary provides users with a quick overview of the video's main topics and insights.

E. Keyword Extraction

To identify important concepts within the transcript, the system applies Term Frequency–Inverse Document Frequency (TF-IDF) analysis. This method calculates the importance of words based on their frequency within the transcript and their rarity across the text corpus. The resulting keywords highlight the most relevant terms discussed in the video.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue IV Apr 2026- Available at www.ijraset.com

F. Sentiment Analysis

Sentiment analysis is performed using a lexicon-based approach to determine the emotional tone of the video content. The analysis calculates polarity and subjectivity scores, which indicate whether the overall tone of the transcript is positive, negative, or neutral. This information helps users understand the general sentiment conveyed in the video.

G. Topic Classification

The system also identifies the primary subject area of the video using a rule-based topic classification mechanism. By matching transcript vocabulary with predefined domainspecific keywords, the system categorizes the video into topics such as technology, education, business, or science.

H. Question-Answering Interface

To enhance user interaction, the system includes a chatbotbased question-answering module. Users can ask questions related to the video content, and the system generates responses based on the transcript context. This functionality allows users to retrieve specific information from the video without manually searching through the entire transcript.

I. System Implementation

The proposed system is implemented using a Python-based backend built with the Flask framework. The frontend interface is developed using HTML, CSS, and JavaScript, providing a user-friendly environment for interacting with the system. The architecture integrates external libraries and APIs to perform transcript retrieval, NLP processing, and result visualization.

Overall, the proposed methodology enables efficient analysis and summarization of YouTube videos by combining multiple artificial intelligence techniques within a unified framework.

IV. SYSTEM ARCHITECTURE

The architecture of the proposed VidMind AI system is designed to efficiently process YouTube videos and generate meaningful insights from their transcript data. The system follows a modular pipeline structure in which each component performs a specific task in the overall processing workflow. This modular design improves scalability, maintainability, and flexibility while enabling the integration of multiple Natural Language Processing (NLP) techniques within a unified framework.

As illustrated in Fig. 1, the system begins with user input and progresses through several processing stages, including metadata extraction, transcript retrieval, text preprocessing, NLP-based analysis, and result visualization. Each stage contributes to transforming raw video transcript data into structured information that can be easily understood by users.

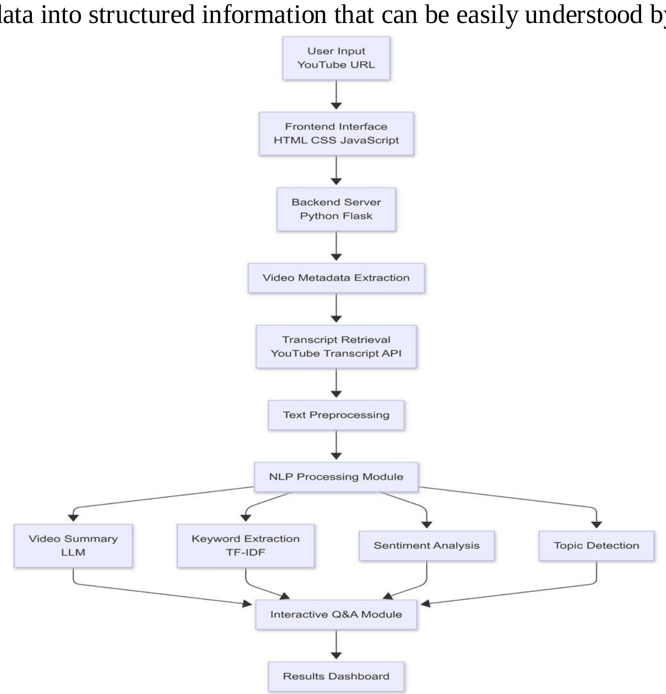

Fig. 1. Architecture and workflow of the proposed VidMind AI system.

A. User Input Module

The process starts when a user provides a YouTube video URL through the web-based interface. This URL serves as the primary input for the system. The interface validates the input to ensure that it corresponds to a valid YouTube link before initiating further processing. Once validated, the URL is forwarded to the backend server for analysis.

B. Video Metadata Extraction

After receiving the video URL, the system extracts basic metadata associated with the video. This information includes the video title, channel name, duration, and other descriptive attributes. Metadata extraction helps the system identify the video and provide contextual information in the final results dashboard.

C. Transcript Retrieval

The next stage involves retrieving the video transcript using the YouTube Transcript API. The API provides caption data that represents the spoken content of the video in textual form. The retrieved captions are timestamped segments that collectively form the complete transcript of the video. This transcript serves as the primary data source for all subsequent NLP processing tasks.

D. Text Preprocessing

Before performing advanced analysis, the transcript undergoes preprocessing to improve its quality and consistency. This step includes removing unnecessary characters, eliminating redundant spaces, and normalizing the text format. The preprocessing stage ensures that the transcript is structured and suitable for further NLP operations.

E. NLP Processing Pipeline

The core component of the system is the NLP processing pipeline, which analyzes the transcript using multiple techniques. This stage consists of several independent modules that operate on the processed text.

- Abstractive Summarization: A large language model is used to generate a concise summary of the transcript. The summary captures the main ideas of the video while reducing the overall length of the content.
- Keyword Extraction: The system applies the TF-IDF technique to identify important keywords and phrases that represent the central topics discussed in the video.
- Sentiment Analysis: A lexicon-based sentiment analysis method is employed to determine the emotional tone of the transcript, indicating whether the content expresses positive, negative, or neutral sentiment.
- Topic Classification: A rule-based classification approach categorizes the video into predefined subject areas such as technology, education, business, or science based on the vocabulary present in the transcript.

F. Interactive Question-Answering Module

To further enhance user interaction, the system incorporates a question-answering module that allows users to ask queries related to the video content. The module analyzes the transcript and generates responses based on the relevant context, enabling users to quickly retrieve specific information without manually reviewing the entire transcript.

G. Results Dashboard

The final stage of the architecture is the results dashboard, which presents the processed information in a structured and user-friendly format. The dashboard displays the generated summary, extracted keywords, sentiment analysis results, detected topics, and responses from the question-answering module. By presenting these insights in a consolidated view, the system enables users to efficiently understand the key aspects of the video content.

Overall, the proposed system architecture provides an integrated framework for analyzing YouTube videos using transcript-based NLP techniques. The modular design ensures that each component performs a specialized function while collectively contributing to the overall objective of efficient video content summarization and analysis.

V. ALGORITHMS AND MODEL DESIGN

The VidMind AI system utilizes multiple Natural Language Processing (NLP) techniques to analyze and summarize YouTube video transcripts. Each module within the system is designed to perform a specific analytical task that contributes to the overall functionality of the framework.

The primary algorithms employed in the system include abstractive summarization using a large language model, TF-IDF based keyword extraction, sentiment analysis using a lexicon-based approach, and rule-based topic classification.
Table I summarizes the algorithms used in different modules of the VidMind AI system.

TABLE I
ALGORITHMS USED IN VIDMIND AI MODULES

Module	Algorithm	Purpose
Summarization	Large Language Model	Generates concise summary of transcript
Keyword Extraction	TF-IDF	Identifies important terms in transcript
Sentiment Analysis	Lexicon-based approach	Determines emotional tone of content
Topic Classification	Rule-based classification	Categorizes video domain
Question Answering	Context-based retrieval	Answers user queries from transcript

A. *Abstractive Summarization Model*

The summarization component of the system is responsible for generating concise descriptions of the video content. In this module, the transcript obtained from the YouTube video is provided as input to a large language model capable of performing abstractive summarization. Unlike extractive methods that simply select sentences from the original text, abstractive summarization generates new sentences that represent the essential meaning of the content.
The model processes the transcript and identifies the most relevant semantic information using contextual understanding of the text. The output is a short summary that highlights the key points of the video while preserving the original meaning. This approach enables users to quickly understand the central ideas of long video transcripts.

B. *Keyword Extraction using TF-IDF*

Keyword extraction is implemented using the Term Frequency–Inverse Document Frequency (TF-IDF) technique. TFIDF is a statistical method used to evaluate the importance of a word within a document relative to a collection of documents.
The term frequency (TF) measures how frequently a word appears in the transcript, while the inverse document frequency (IDF) reduces the weight of common words that appear frequently across multiple documents. The TF-IDF score for a term is calculated as:
$TF\text{-}IDF(t, d) = TF(t, d) \times IDF(t)$ (1) where $TF(t, d)$ represents the frequency of term t in document d, and $IDF(t)$ represents the inverse document frequency of the term. Words with higher TF-IDF scores are considered more significant and are selected as keywords representing the main topics of the transcript.

C. *Sentiment Analysis*

The system performs sentiment analysis to determine the emotional tone expressed in the video transcript. This analysis is implemented using a lexicon-based approach, where each word in the transcript is associated with a predefined sentiment score.
The sentiment polarity is calculated by aggregating the sentiment values of words within the text. The resulting score indicates whether the overall sentiment of the content is positive, negative, or neutral. This information helps users understand the tone of the video and identify whether the speaker expresses opinions, criticism, or neutral information.

D. *Topic Classification*

Topic classification is used to categorize the video content into predefined subject areas. The system employs a rule-based classification approach in which the transcript vocabulary is compared with domain-specific keyword lists. These keyword lists correspond to different subject categories such as technology, education, business, science, and entertainment.
When words from the transcript match the vocabulary of a particular category, the system assigns the corresponding topic label. The topic with the highest number of matching keywords is selected as the primary category of the video.

E. Question-Answering Model

To enable interactive exploration of video content, the system includes a question-answering module. Users can submit natural language questions related to the video, and the system retrieves relevant information from the transcript to generate a response.

The question-answering module uses contextual analysis to identify relevant sections of the transcript that correspond to the user's query. The response is then generated based on the identified information, allowing users to obtain specific insights without manually reviewing the entire transcript.

Overall, the combination of these algorithms enables the VidMind AI system to transform raw transcript data into structured knowledge, improving the efficiency and accessibility of video content analysis.

VI. RESULTS AND DISCUSSION

The experimental evaluation demonstrates that the proposed VidMind AI system effectively analyzes YouTube video transcripts and generates meaningful insights. The abstractive summarization module produces concise summaries that capture the essential ideas of the video content. The TF-IDF based keyword extraction method successfully highlights the main concepts discussed in the transcript.

Sentiment analysis provides useful information about the emotional tone of the video, while the topic classification module categorizes videos into relevant domains such as technology and education. Furthermore, the interactive questionanswering module allows users to retrieve specific information directly from the transcript, improving the usability of the system.

TABLE II
PERFORMANCE OF VIDMIND AI MODULES

Module	Function	Observation
Summarization	Generates concise video summary	Captures key ideas clearly
Keyword Extraction	Identifies important terms	Highlights main topics
Sentiment Analysis	Detects emotional tone	Mostly neutral for tutorial videos
Topic Classification	Classifies video domain	Correctly identifies content category
Question Answering	Retrieves specific information from transcript	Provides context-aware responses

Table II summarizes the performance of the major modules in the proposed VidMind AI system. The results indicate that each module effectively contributes to transcript-based video analysis by generating concise summaries, identifying important keywords, detecting sentiment, classifying topics, and providing context-aware responses.

Overall, the results indicate that the proposed framework can efficiently transform lengthy video transcripts into structured insights that improve accessibility and reduce the time required to understand video content.

VII. CONCLUSION AND FUTURE WORK

In this paper, we presented VidMind AI, an intelligent system for automatically summarizing and analyzing YouTube videos using transcript-based Natural Language Processing techniques. The proposed framework integrates multiple components, including abstractive summarization, TF-IDF based keyword extraction, sentiment analysis, topic classification, and an interactive question-answering module. By processing video transcripts, the system generates concise summaries and extracts meaningful insights that help users quickly understand the core content of videos.

The results demonstrate that the system effectively converts lengthy transcripts into structured information such as summaries, keywords, sentiment indicators, and topic categories, thereby reducing the time required to identify relevant information in video content. The inclusion of a question-answering module further enhances usability by enabling users to retrieve specific details directly from the transcript.

However, the system's performance depends on the availability and accuracy of video transcripts. Future work will focus on improving the system by incorporating multimodal analysis techniques that combine audio, visual, and textual information, as well as integrating more advanced language models and expanding multilingual support.

REFERENCES

[1] M. Kini, S. A. Vaturi, and S. V. Khanaganvi, "AI-Based Video Summarizer System," in Proc. 2025 Modern Electronics Devices and Intelligent Communication Systems (MEDCOM), 2025, pp. 892–898.
[2] A. Thanam, E. Kamalanaban, M. Y. Devi, C. Surekha, M. S. Keerthi, and T. R. Manasvi, "An NLP-Driven Approach for Automated Transcription and Summarization of YouTube Videos," in Proc. 2025 8th Int. Conf. Circuit, Power & Computing Technologies (ICCPCT), 2025, pp. 330– 334.
[3] K. Kanagaraj, S. Abhang, J. S. Kumar, R. K. Gnanamurthy, and V. Balaji, "AI-based video summarization for efficient content retrieval," ICTACT Journal on Image & Video Processing, vol. 14, no. 2, 2023.
[4] K. Nagaraju, P. S. Yadav, S. Khandate, V. Kharate, and J. Syed, "AIBased Real-Time Monitoring of YouTube Content," in Proc. 2025 3rd DMIHER Int. Conf. Artificial Intelligence in Healthcare, Education and Industry (IDICAIHEI), 2025, pp. 1–6.
[5] L. H. Palivela, J. Karthik, D. P. S. Raju, and P. Jayanth, "AI-Powered Innovative Method for YouTube Education Using Next-Gen Prompt Sheets," in Driving Quality Education Through AI and Data Science, IGI Global Scientific Publishing, 2025, pp. 557–582.
[6] P. Kalaiarasi, M. Shaikshavali, M. V. Praneeth, N. S. L. Lavanya, J. V. J. Reddy, and S. Junaid, "AI Youtube Analytics Tools—Smarter Growth Insights!," in Proc. 2025 Int. Conf. Sustainable Communication Networks and Application (ICSCN), 2025, pp. 1364–1369.
[7] S. P. Kolhe, S. Gupta, and L. Upadhyay, "AI-Driven Video Summarization Using LangChain and LLMs," in Proc. Int. Conf. Computer Vision and Robotics, Springer Nature Switzerland, 2025, pp. 474–487.
[8] I. Kontostathis, E. Apostolidis, K. Apostolidis, and V. Mezaris, "Enhancing User Control in AI-Based Video Summarization for Social Media," in Proc. Int. Conf. Multimedia Modeling, Springer Nature Singapore, 2025, pp. 119–126.
[9] Y. Anthati, K. Mudde, H. Damuluri, and V. R. Thupakula, "Multilingual Video Summarizer," in Proc. 2025 Int. Conf. Computational Robotics, Testing and Engineering Evaluation (ICCRTEE), 2025, pp. 1–6.
[10] K. Vayadande, M. Bohri, M. Chawala, A. M. Kulkarni, and A. Mursal, "The Rise of AI-Generated News Videos: A Detailed Review," in How Machine Learning is Innovating Today's World: A Concise Technical Guide, pp. 423–451, 2024.
[11] P. A. Reddy, G. Prasanna, S. Sulthana, P. Vaishnavi, S. Srilekha, and G. Arora, "Smart Video Assistant: A Video Summarizer and Interactive Chatbot System," in Proc. 2025 5th Int. Conf. Advancement in Electronics & Communication Engineering (AECE), 2025, pp. 812–816.
[12] A. A. Balushi, A. S. Al-Bemani, S. Al Araimi, G. Balaji, U. Suresh, and A. Najeeb, "AI-Driven Multi-Modal Information Synthesis: Integrating PDF Querying, Speech Summarization, and Cross-Language Text Summarization," Procedia Computer Science, vol. 258, pp. 2996–3018, 2025.
[13] Z. A. Khan, Y. Xia, F. Khaliq, W. Jiang, and M. S. Anwar, "AIBased Women Hate Speech Detection System Using YouTube URLs for Effective Content Moderation," IEEE Intelligent Systems, 2025.
[14] B. Gautham and K. Somasundaram, "Social Media Content Generation: An AI-Based System for Automated Viral Short-Form Video Creation," in Proc. 2025 IEEE 2nd Int. Conf. Information Technology, Electronics and Intelligent Communication Systems (ICITEICS), 2025, pp. 1–9.
[15] S. R. Varshini, G. Akshaya, M. Sivarathinabala, R. J. Chitra, and J. Sneha, "Real-Time Video Summarizer for LMS With Temporal Modeling and Attention Mechanism Using Deep Learning," in Examining AI Innovations in Gaming and Entertainment, IGI Global Scientific Publishing, 2026, pp. 245–276.