International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 05 | May - 2023 SJIF 2023: 8.176 ISSN: 2582-3930

AUTO-BLOGGER

Ayushi Chandra, Kamlesh Kumar, Aryan Sonker, Anuj Gautam

Dept. Of Computer Science & Engineering Babu Banarasi Das Institute Of

Technology And Management Lucknow, Uttar Pradesh, India.

Assistant Professor, Dept. Of Computer Science & Engineering, Babu Banarasi Das

Institute Of Technology and Management Lucknow, Uttar Pradesh, India.

Abstract

Running automatic scripts against well-known content sources or syndicated feeds and assembling the results into blog sites is known as auto-blogging. Unfortunately, a criminal ecosystem is developing to profit from the linked technologies, unlike mashups, which aim to improve the Internet experience by fusing various material sources in creative ways. A rogue site can impersonate a reputable site while running their own ads for profit by mass-scraping content from real ad-supported websites, removing the ads, credit, and by rewriting links. In order to address the issue of traceability when material is copied in bulk from one website and hosted on another, we present a forensic technique.

The content is first minified with a sparse encoding tailored to the user's available steganographic bandwidth, then watermarked repeatedly with different IDs.

Keywords: Automatic blogging, video to text, blogging.

1. Introduction

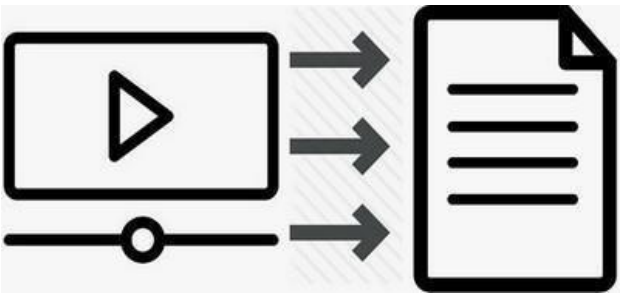
To overcome the issues with the current manual system in use, the "Auto-blogger" was created. The difficulties our current system faces are supported by this program, which aims to eliminate and, in some circumstances, lessen them. Additionally, this system was developed to satisfy the unique needs of the company in order to run operations successfully and efficiently. This program has been as much as possible reduced back to reduce data entry errors. Additionally, when you submit inaccurate data, an error notification is displayed. To utilize this system, the user doesn't need any formal training. The main objective of this program is to make it straightforward to view the uploaded blogs, entries, themes, etc. In addition, blogs can be uploaded, modified, deleted, and utilized for other purposes. Additionally, it is utilized to read and post blog entries from other individuals. Controlling ideas, blogs, entries, material, and viewpoints is tough for any organization, no matter how big or little. Because every online blogging system has different blog needs, we develop customized personnel management systems that are suited to your managerial requirements. This is intended to aid with strategic planning and will ensure that your business is set up to manage your employees wherever you are, whenever you need to. These methods will ultimately help you

manage resources more skillfully.[1]

Abbas Q et al. [1] state that in the blogging world, the idea of auto-blogging is divisive. What does it mean, though, exactly? However, with auto blogging, you don't need to actively create or gather the information. Additionally, there are numerous ways in which you can post the content that is generated automatically. You can decide to become an auto blogger if you don't have enough time to create the content. You will still need some time to get the content from other sources, though. You must comprehend how it operates and all of its benefits and drawbacks in order to decide if it is appropriate for you. The biggest misperception about auto blogging is that it's a brand-new concept. That, however, is inaccurate. People have been using auto blogging since 2000. This type of blogging was formerly done using content aggregator software. This software uses information gathered from several sources to build blogs. Google currently uses an algorithm that labels this information as spam as a result. These products are not displayed in the search results.[2][3]

The main objective of this program is to make it straightforward to view the uploaded blogs, entries, themes, etc. In addition, blogs can be uploaded, modified, deleted, and utilized for other purposes. Additionally, it is utilized to read and post blog entries from other individuals. Controlling ideas, blogs, entries, material, and viewpoints is tough for any organization, no matter how big or little. Because every online blogging system has different blog needs, we develop customized personnel management systems that are suited to your managerial requirements. This can help with strategic planning and make sure that your company is set up so that you can manage your personnel anytime you need to, no matter where you are. Ultimately, these solutions will give you the ability to manage resources.[1]

Fig1.1 Video to Audio



1.1 Types of video to text conversion

Depending on the particular needs and limitations of a project, a variety of video to text conversion approaches may be used. Here are a few typical methods for converting videos to text:

1. Automatic Speech Recognition (ASR): This method of turning spoken language into text is very popular. It entails listening to a video's audio material and turning it into text. To accurately recognise and transcribe speech, ASR systems use statistical models, deep learning algorithms, and language models.

2. Manual Transcription: In order to make a textual transcript, manual transcription entails listening to the audio in a movie and manually typing out the spoken words. When compared to automated methods, this method often requires more human labour, takes longer, and costs more money. It can, nevertheless, deliver great accuracy, particularly when working with difficult sounds or niche areas.

3. Hybrid Approaches: Hybrid strategies combine manual editing or correction with automatic speech recognition. An initial transcription of the video's audio is created using ASR techniques, and it is later checked and edited by human transcribers. By utilising the effectiveness of automatic voice recognition, this technique aids in improving the accuracy and quality of the transcriptions.

4. Multimodal Analysis: In a multimodal analysis, the text is produced using both audio and visual data from the video. In addition to voice recognition, visual analysis methods like lipreading, facial expression analysis, and scene analysis can be used to improve the precision and context of the transcriptions. Multimodal methods may enable a more thorough comprehension of the video's content.

5. Domain-Specific Transcription: Techniques for domain-specific transcription can be used in certain fields or domains. These methods concentrate on modifying voice recognition models and language models to match the particular terminologies, jargon, or language patterns common in that field. For applications that are domain-specific, this customisation increases the transcriptions' accuracy and relevance.

The best method for converting videos to text will rely on a number of variables, including the amount of accuracy necessary, the resources available (time, money, experience), the domain-specific needs, and the level of automation sought. To choose the best strategy for a certain video to text conversion project, it's critical to evaluate these aspects.

1.2 Benefits of Auto-blogger project

The benefits of video to text conversion are numerous and apply to many different fields and applications. The following are some major advantages of text to video conversion:

1. Accessibility: By turning video footage into text, it may be viewed by people who have hearing loss or who prefer reading to watching or listening to media. A wider spectrum of consumers can access content equally thanks to text transcripts.

2. Searchability: Text transcripts make it simple to look up specific information within a movie and retrieve it. When compared to manually scrolling through video information, users can quickly identify pertinent sections or keywords by employing text search features.

3. Indexing and Organisation: Video transcripts make it possible to efficiently index and arrange video archives. The ability to categorise and retrieve movies based on their content is made easier by linking searchable text to particular video segments, improving video content management.

4. Textual transcripts of videos make it easier to do content analyses and summaries as well as data mining. Text-based analysis methods can be used for tasks including information extraction, market analysis, sentiment analysis, topic modelling, and the generation of text summaries of video footage.

5. Language Translation and Localization: The conversion of video to text makes it simpler to translate and adapt video material for other languages. Once the video has been converted to text, it is easier to translate the text into several languages, allowing for greater accessibility and reaching audiences across the world.

6. Education and Learning: Textual transcripts of educational or instructional videos serve as useful study aids by enabling viewers to read, annotate, and take notes on the video's subject matter. Additionally, transcripts can be utilised to make closed captions or subtitles, improving learning results.

7. Compliance and Legal standards: In some circumstances, accurate transcriptions of video footage are required by regulatory or legal standards. Converting videos to text assures adherence to accessibility guidelines, copyright laws, or requirements for legal documentation.

8. Content reuse and repurposing: Transcripts offer reusable text that can be used in a variety of new contexts. The scope and worth of the original video footage can be increased by using it to repackage content, produce blog entries, social media snippets, or audio descriptions for visually impaired users.

9. Data Integration and Analysis: Textual transcripts can be combined with other data sources for additional analysis or with machine learning strategies to derive insights, patterns, or conduct sentiment analysis across various data modalities, facilitating greater comprehension and decision-making.

10. Time and Cost Savings: Automating the process of creating text from films saves time and money by eliminating the need for manual transcribing. Especially when working with high numbers of video content, this increases efficiency, productivity, and cost effectiveness.

These advantages emphasise the usefulness and adaptability of video to text conversion, improving video content's accessibility, searchability, and analytical capabilities across a range of fields, businesses, and user needs.

2. Proposed work

According to Chen J, et al. [4] states that the proposed work for the video-to-text conversion project aims to develop an efficient and accurate system that can automatically generate written descriptions of videos. The key components of the project include:

1. Data Collection and Preprocessing: Gathering a diverse dataset of videos along with their corresponding transcripts or captions. Preprocessing steps involve cleaning, segmenting, and aligning the video and text data.

2. Feature Extraction: Extracting relevant visual and audio features from the video frames, such as frame-level visual features using convolutional neural networks (CNNs) and audio features using techniques like spectrogram analysis.

3. Video Understanding: Employing techniques such as deep neural networks, recurrent neural networks (RNNs), or transformer models to analyze the video content and capture temporal dependencies in order to understand the sequence of events, actions, and objects present in the video.

4. Language Generation: Utilizing natural language processing (NLP) techniques, including recurrent or transformer-based architectures, to generate coherent and descriptive text based on the extracted video features and contextual information.

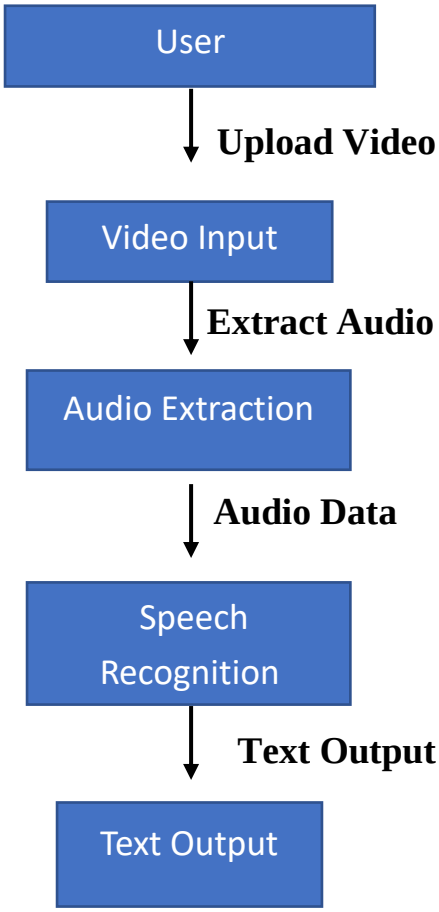
5. Evaluation Metrics: Defining appropriate evaluation metrics, such as BLEU (Bilingual Evaluation Understudy), METEOR (Metric for Evaluation of Translation with Explicit ORdering), or CIDEr (Consensus-based Image Description Evaluation), to assess the quality and accuracy of the generated text descriptions.

6. Model Optimization and Fine-tuning: Iteratively refining the model by optimizing various components, adjusting hyperparameters, and performing fine-tuning using techniques like backpropagation or reinforcement learning to improve the overall performance.

7. Performance Evaluation: Conducting thorough evaluations and comparisons with existing methods on benchmark datasets to demonstrate the effectiveness and superiority of the proposed video-to-text conversion system.

By implementing these steps, the project aims to deliver a robust and state-of-the-art solution for automatically converting video content into textual descriptions, benefiting applications such as video search, content recommendation, and accessibility for the visually impaired.[5]

2.1. Data Flow Diagram



2.2. Application Development

The proposed auto-blogger website will be developed using the following technologies:

Programming Language: The application will be developed using a high-level programming language such as JavaScript.

Web Framework: The application will be built using a web framework such as Express.js.

Database: The application will use a database which is MongoDB to store web traffic data.

3. Methodology

The development of the auto-blogger website involves the following steps:

Data Collection: Start selecting the video you want to convert to text. Make sure it is a high-quality video with clear audio and minimal background noise.

Audio Extraction: Extract the audio from the video by using video processing libraries or tools to separate the audio track from the video file. For that we have used FFmpeg- is a powerful open-source multimedia used to extract audio from the video.

Speech Recognition: For speech recognition we have used deepgram- it is a company that specializes in automatic speech recognition (ASR) technology.

Pre-processing and Post-processing: Here we can remove the filler words and unnecessary information or noise. Spell check, grammar correction and improving the overall readability of the text.

Formatting: Before deployment, make changes as per your need.

Deployment: The application will be deployed on a server and made available to website owners and marketers.

4. Result and discussion

Chen S, et al. [5] states that the process of text-to-video conversion typically involves the following steps:

Text Processing: The input text is analyzed and processed to extract relevant information, such as keywords, entities, or key phrases. Natural Language Processing (NLP) techniques may be used to perform this analysis.

Visual Storyboarding: Based on the processed text, a visual storyboard is created. The storyboard outlines the key scenes, visuals, and transitions that will be included in the video.

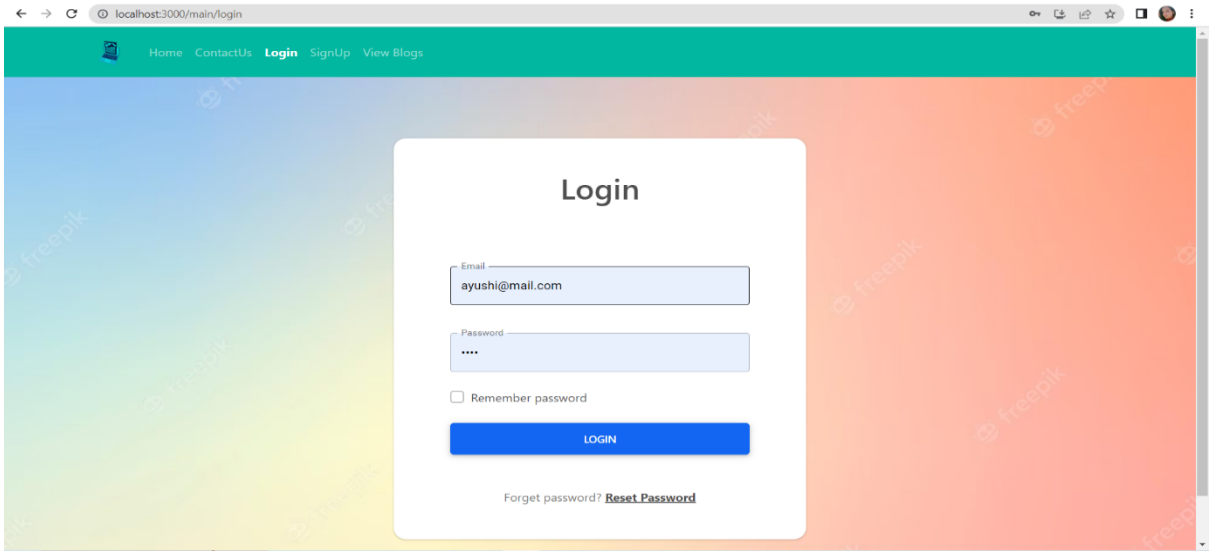
Asset Selection: Visual assets such as images, graphics, animations, and videos are selected to represent the content of the text. These assets can be obtained from stock libraries or custom-created.

Video Generation: Using video editing software or specialized tools, the selected assets are combined and arranged according to the storyboard. Transitions, effects, and animations may be added to enhance the video's visual appeal and convey the message effectively.

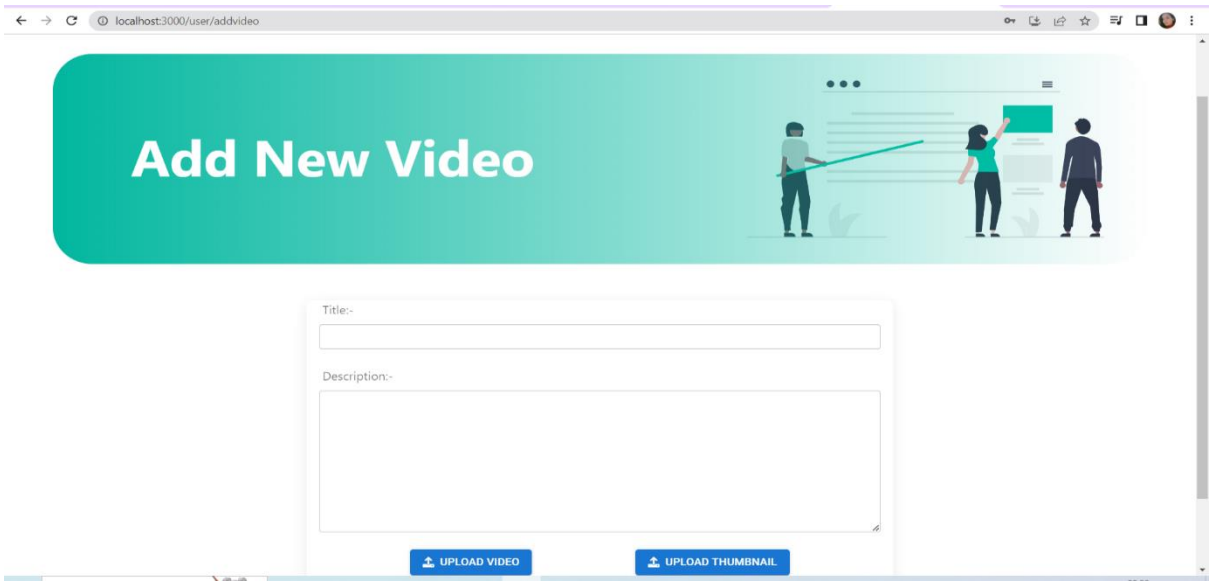
Audio Integration: Background music, voice-over narration, or sound effects can be added to enhance the video and provide a more engaging experience.

Rendering and Export: The video is rendered or exported into a desired format, such as MP4 or MOV, with suitable resolution and settings.[5]

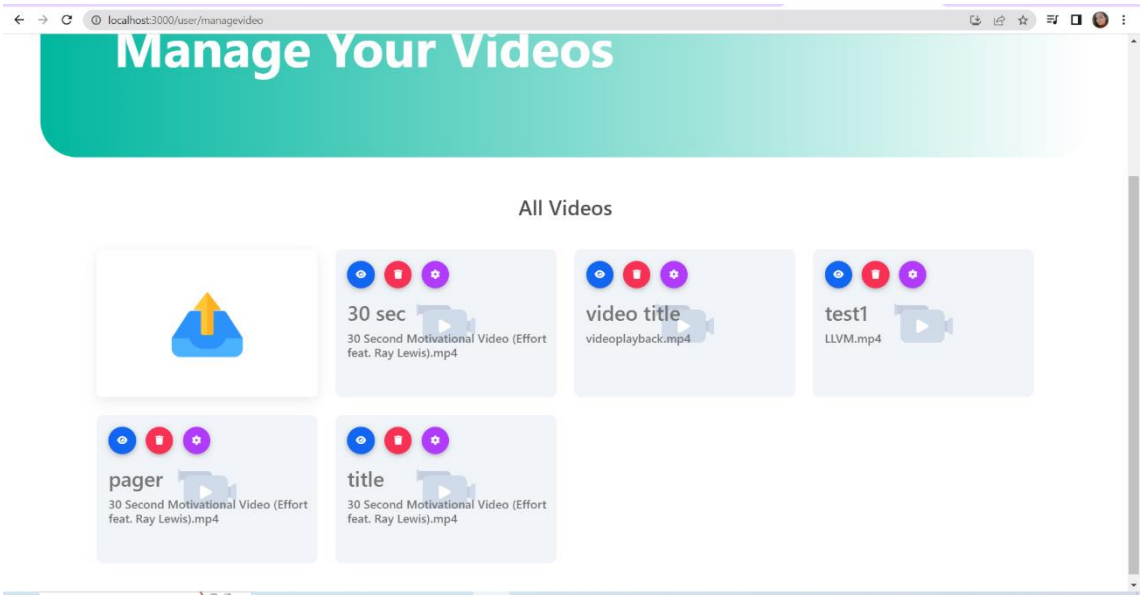
1. Login Page



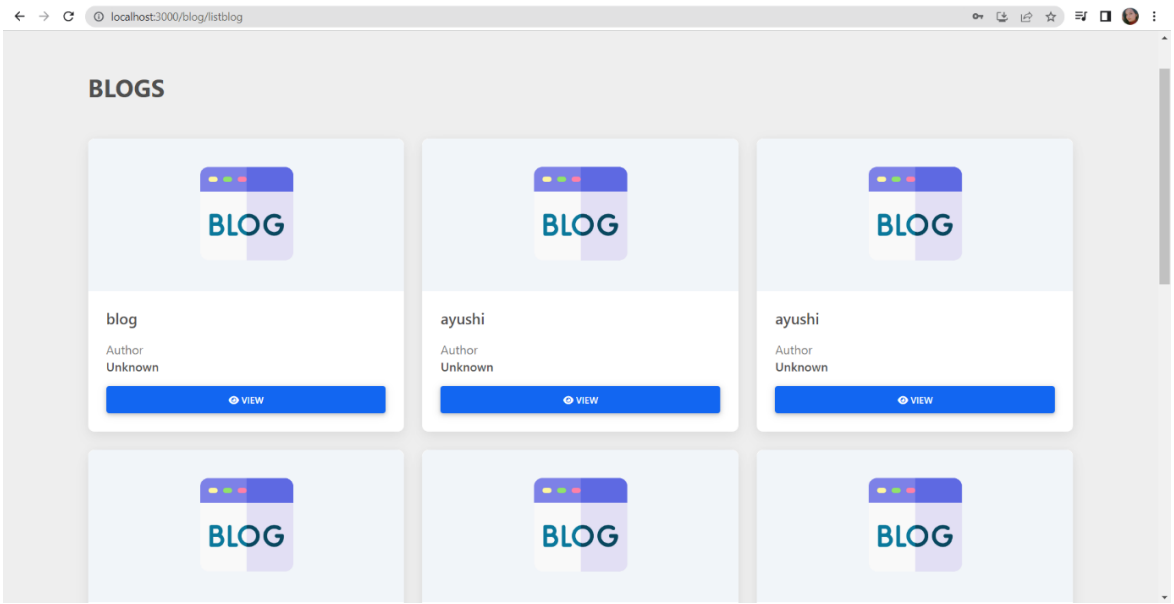
2. Add video page



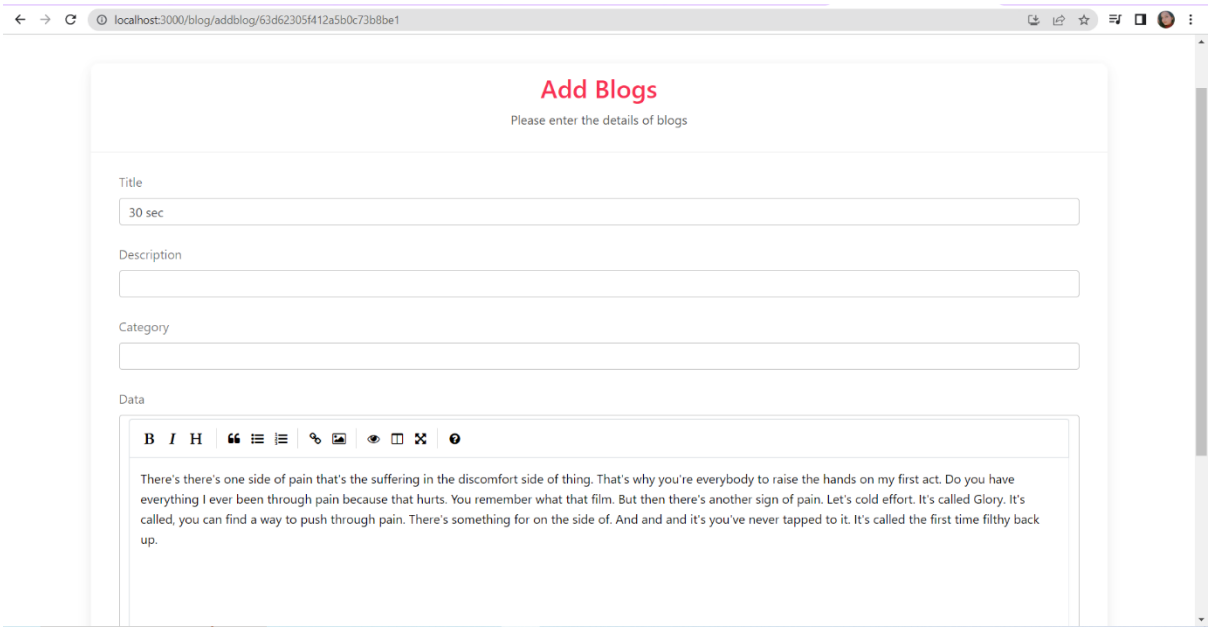
3. Manage video page



4. List blogs



5. Graph for Contact Counter



5. Future Scope

Venkatasubramanian, et al. [6] Future video to text conversion projects provides a wide range of potential for new developments and uses. The following areas present promising prospects for future growth:

1. Enhanced Accuracy: One continuing area of research is enhancing the precision of speech recognition models and algorithms. Higher accuracy and better handling of difficult audio situations, accents, and languages can be achieved with the use of deep learning techniques advancements as transformer models and self-supervised learning.

2. Multilingual and Multimodal Support: The accuracy and contextual understanding of the transcriptions can be improved by extending the capabilities of video to text conversion projects to support a wider range of languages and manage multimodal inputs (audio, visual, and textual). More thorough and precise conversions will be possible with the development of models and methodologies that can handle many languages and incorporate visual clues.

3. Real-time Transcription: Recent developments in low-latency processing and real-time voice recognition algorithms may make real-time transcription possible in a variety of contexts, including video conferences and live streaming events. Possibilities for real-time transcribing will improve interaction, collaboration, and accessibility across a range of fields.

4. Contextual Understanding: By using context-aware methods for video to text conversion, transcriptions' comprehension and applicability can be increased. Contextual data like speaker knowledge, subject modelling, sentiment analysis, or scene understanding can improve the generated text's quality and usefulness, allowing for more complex applications.

5. Domain-specific Adaptation: Accuracy and relevance can be increased by customising video to text conversion models and algorithms for certain domains or sectors. The quality of transcriptions for specialised

applications, such as those in the medical, legal, or technical domains, will be improved by customising models to adapt to domain-specific terminology, jargon, or accents.

6. Integration with Natural Language Processing (NLP): Combining video to text conversion with NLP methods makes it possible to analyse and comprehend the transcriptions at a deeper level. Tasks like entity recognition, sentiment analysis, summarization, question-answering, or information extraction can be made possible by natural language processing, leading to deeper insights and more useful applications.

7. Voice Control and Interaction: Adding voice control and interaction features to video to text conversion projects can improve accessibility and user experience. A more user-friendly and hands-free experience can be achieved by incorporating voice commands to navigate and interact with video material utilising the created transcriptions.

8. Privacy and Security Issues: It's important to take privacy and security issues into account when converting videos to text. Future work should concentrate on assuring secure handling of video data, putting privacy-preserving measures in place, and adhering to data protection laws.

9. Multimodal Visualisation: Creating methods to simultaneously visualise the video content and its text transcriptions helps improve understanding and analysis of the video content. Visualisations that draw attention to significant words, speaker turns, or scene context can help viewers comprehend and interpret video content more effectively.

10. Integration with Smart Assistants and Applications: Using video to text conversion in conjunction with smart assistants and applications can create new opportunities for interactive video experiences, voice-controlled video playback, and personalised content recommendations. Through this integration, video material may be more easily accessed and used on a variety of platforms and devices.

With ongoing developments in speech recognition, language processing, and multimodal analysis, the future potential for video to text conversion projects is enormous. These developments will help make video to text conversions more precise, effective, and contextually rich, opening up a wide range of applications across industries, education, accessibility, and content management.[7][8]

6. Conclusion

To sum up, a video to text conversion project is a worthwhile endeavour that has multiple advantages and uses in many different fields. It improves accessibility, searchability, organisation, and analysis of the video content by turning it into text transcripts. Support for various video file formats, audio extraction, reliable speech recognition, speaker diarization, integration of language models, post-processing techniques, user-friendly interface, scalability, real-time processing (if necessary), customization options, and export/integration capabilities are some of the key characteristics of such a project.[10][11]

The project makes video content accessible to people who have hearing difficulties or who prefer text-based content. It provides effective content analysis and summarization, indexing and organisation of video libraries, and efficient search and retrieval of specific information inside videos. Converting videos to text also helps with localization and language translation, regulatory compliance, and content reuse and repurposing.[12]

N. Xu, et al. [13] states that it's crucial to pick the right methods, such as speaker diarization, hybrid approaches, multimodal analysis, and automatic voice recognition, when performing a video to text conversion project. Post-processing methods and customization for particular domains further improve the precision and calibre of the transcriptions.

Overall, a video to text conversion project enhances the effectiveness, usability, and accessibility of video material by giving viewers access to text transcripts that are quickly searchable, interpretable, and analysable.[14][15]

REFERENCES

[1]. Abbas Q, Ibrahim ME, Jaffar MA (2019) A comprehensive review of recent advances on deep vision systems. Artif Intell Rev 52(1):39–76

[2]. Bin Y, Yang Y, Shen F, Xie N, Shen HT, Li X (2019) Describing video with attention-based bidirectional LSTM. IEEE Trans Cybernet 49(7):2631–2641

[3]. Wang, D., Yang, S., Su, D., Liu, X., Yu, D., & Meng, H. (2022, May). VCVTS: Multi-speaker Video-to-Speech synthesis via cross-modal knowledge transfer from voice conversion. In ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 7252-7256). IEEE.

[4]. Chen J, Pan Y, Li Y, Yao T, Chao H, Mei T (2019) Temporal deformable convolutional encoder-decoder networks for video captioning. AAAI 33:8167–8174

[5]. Chen S, Jin Q, Chen J, Hauptmann A (2019) Generating video descriptions with latent topic guidance. IEEE Trans Multimedia 21:2407–2418

[6]. Venkatasubramanian, S., & Mohankumar, R. (2022). A Deep Convolutional Neural Network-Based Speech-to-Text Conversion for Multilingual Languages. In Computational Vision and Bio-Inspired Computing (pp. 617-633). Springer, Singapore

[7]. Soe, T. H., & Slavkovik, M. (2022, June). A content-aware tool for converting videos to narrower aspect ratios. In ACM International Conference on Interactive Media Experiences (pp. 109-120).

[8]. Hemalatha, B., Karthik, B., Balaji, S., Vijayalakshmi, G., & Shaw, R. N. (2022). A Novel Approach for Blind-Image to Audio Conversion in Regional Language. In International Conference on Electrical and Electronics Engineering (pp. 662-668). Springer, Singapore.

[9]. Dong J, Li X, Snoek CGM (2018) Predicting visual features from text for image and video caption retrieval. IEEE Trans Multimedia 20(12):3377–3388

[10]. Abduljabbar, A., Gupta, N., Healy, L., Kumar, Y., Li, J. J., & Morreale, P. (2022, March). A SelfServed AI Tutor for Growth Mindset Teaching. In 2022 5th International Conference on Information and Computer Technologies (ICICT) (pp. 55-59). IEEE

[11]. X. Wang, W. Chen, J. Wu, Y.-F. Wang and W. Y. Wang, "Video captioning via hierarchical reinforcement learning", The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018.

[12]. L. Gao, X. Wang, J. Song and Y. Liu, "Fused GRU with semantic-temporal attention for video captioning", Neurocomputing, 2019.

[13]. N. Xu, A.-A. Liu, Y. Wong, Y. Zhang, W. Nie, Y. Su, et al., "Dual-stream recurrent neural network for video captioning", IEEE Transactions on Circuits and Systems for Video Technology (TCSVT), 2018.
[14]. B. Wang, L. Ma, W. Zhang and W. Liu, "Reconstruction network for video captioning", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7622-7631, 2018.
[15]. A.-A. Liu, N. Xu, H. Zhang, W. Nie, Y. Su and Y. Zhang, "Multi-level policy and reward reinforcement learning for image captioning", Proc. IJCAI, pp. 821-827, 2018.