ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14949832

Teaching and Learning Enhancement using Computer Vision and Generative AI

Aravapalli Sri Chaitanya[1] (Ph.D), (Assistant Professor); Ragalahari Akula[2]; Pachava Hemanth[3]; Tanarapu Praveen Kumar[4]; Gurani Durga Suresh[5]

[1](ORCID : 0009-0008-5127-0250) Artificial Intelligence & Machine Learning (CSE) Dhanekula Institute of Engineering and Technology (JNTUK) Vijayawada, India
[2]Artificial Intelligence & Machine Learning (CSE) Dhanekula Institute of Engineering and Technology (JNTUK) Vijayawada, India
[3]Artificial Intelligence & Machine Learning (CSE) Dhanekula Institute of Engineering and Technology (JNTUK) Vijayawada, India
[4]Artificial Intelligence & Machine Learning (CSE) Dhanekula Institute of Engineering and Technology (JNTUK) Vijayawada, India
[5]Artificial Intelligence & Machine Learning (CSE) Dhanekula Institute of Engineering and Technology (JNTUK) Vijayawada, India

Publication Date: 2025/03/01

Abstract: This research introduces a new method for classroom teaching development which combines intelligent AI components with gesture programming solutions and team-based tooling. Users can perform blackboard activities by writing drawings through the Virtual Hand-Controlled Blackboard system using hand gestures. Users can use this function to both record and download sessions to enhance their study capabilities. Through its AI-Powered Quiz Platform that runs on GeniusAI the system turns existing documents into quizzes which both tracks teacher monitoring of student progress and allows students to build customized assessments that boost their autonomous learning abilities. The Collaborative Chat System allows instant query resolutions between educators and learners to enhance their communication efforts. Students can access notas well as recorded classes through the system which makes knowledge consolidation easier to handle. The system delivers Multilingual Support by providing users access to their native language notes in audio and video formats which breaks down language obstacles. This teaching approach includes interactive and accessible and personalized elements to transform traditional education while enabling teachers and students to adapt their educational practices to present-day learning needs.

Keywords: *Virtual Hand-Controlled Black Board, Gesture-Based Technology, AI-Powered Quiz Platform, Genius AI, Multilingual Support, Recorded Classes, Collaborative Chat System, Personalized Learning, Self-Driven Learning, Modern Education Tools.*

How to Cite: Aravapalli Sri Chaitanya; Tanarapu Praveen Kumar; Ragalahari Akula; Gurani Durga Suresh; Pachava Hemanth (2025). Teaching and Learning Enhancement using Computer Vision and Generative AI. *International Journal of Innovative Science and Research Technology*, 10(2), 1109-1117. https://doi.org/10.5281/zenodo.14949832

I. INTRODUCTION

The development of society relies mostly on educational frameworks because emerging technology speeds up educational reforms. Through this initiative teachers can implement an innovative method which uses gestured systems supported by artificial intelligence across multiple languages. Today's technological solutions perform multiple functions that make education more efficient through enhanced resource access and solution of current educational issues.

The educational sector underwent major transitions because digital learning tools entered the classroom. The educational institutions face ongoing difficulties because of linguistic diversity problems alongside physical barrier issues which impact student participation. The project resolves these educational hurdles through a complete organizational structure that establishes advanced technology solutions with actual classroom requirements. The system employs gesture-based interfaces that users can easily understand alongside adaptive content from AI systems and language adjustment features that promote access to all students. Teachers gain the ability to develop interactive learning spaces through this system which allows them to use dynamic educational methods. The use of modern technology allows learners to experience an educational space that integrates conventional

ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14949832

teaching methods into modern teaching practices while providing inclusive learning opportunities.

- *Virtual Black Board* :

Users can utilize gesture-based technology through the Virtual Hand-Controlled Black Board to digitize blackboard writing and diagramming and annotation techniques. This system provides digital communication features between traditional physical and electronic classrooms by maintaining familiar interfaces with innovative technological options. Real-time session recording which students can access whenever they need to review course material is available in this system. Video files from recorded sessions enable students to download the content for both revision purposes and out-of-synchronized studying needs. The combination of flexibility features with interactive capabilities creates an innovative educational tool which strengthens live class participation and builds an adaptable resource for learners to consolidate their knowledge making it important for contemporary classrooms. This tool unlocks a flexible educational environment that provides interactive experiences to students with different learning requirements. Students under both hybrid and virtual learning conditions benefit from the tool because it sustains their active participation as well as supports them through recorded educational content. Schools across modern education benefit from the Virtual Hand-Controlled Black Board because this tool enhances accessibility and acts as an essential educational resource.

- *MCQ Assesment :*

Using the assessment solution developed by GeniusAI students generate quizzes from the documents provided to them. The solution functions to support teachers by minimizing their workload without compromising the important elements of each quiz. Educational institutions use tracking tools to determine which material subjects each student has excelled in along with their specific areas of weakness. The system allows them to monitor data for guiding their instructional approaches. Through the built-in system students may either choose quizzes produced by the platform or generate their own quizzes to enhance their self-preparation and develop autonomy and comprehension abilities. Operational efficiency grows alongside student learning activity because automated systems join forces with individualized teaching tactics through this platform. Students receive improved feedback through adaptive evaluation tests that give educators and learners tools to build their learning potential and challenge established evaluation practices.

- *Multilingual Text :*

Multilingual educational assistance systems have minimized learning barriers related to language differences so students from various language backgrounds can study better. This program provides academic materials in learners' native tongues to deliver important educational support to students who do not speak the main class language. The inclusion model under our multilingual support program overcomes language limitations which block students from obtaining equal educational access. Through a multilingual platform every learner can find easy access while the learning environment becomes inclusive for everyone as it supports acceptance of cultural diversity. UUID functions as a language barrier-breaking system which allows for delivering quality educational services that acknowledge students' rights while helping them reach their scholarly goals without depending on their mother tongue.

- *Class Recordance & Notes :*

The educational progress has made it possible for students to find flexible resources via note exchange services and class recording capabilities. Educational resources provided on the platform include audio and video content since it serves different types of learners. Students use recorded content and class notes for revision purposes since these resources provide support for learning technical concepts and relearning forgotten material without time constraints. Students can engage with educational content through an adaptive platform design at times when their individual learning needs are met. The project works to boost educational performance by delivering customized education methods according to individual student learning behaviors.

II. LITERATURE REVIEW

Advanced technologies including AI and ML together with gesture-based systems have improved HCI along with educational tools thus creating more efficient educational experiences that promote better accessibility and engagement.

With Media Pipe [1] tools from Google Akshay Kumar developed hand-tracking technology which enables gesture control without touch operation to create virtual drawings applicable for educational and design work and gaming applications. Following the same approach Subhash Chand Agrawal [2] used Convolutional Neural Networks (CNNs) to analyze motions accurately and integrated drawing capabilities and tool selection within his system. The technological advances show how vision-based systems can enhance digital creativity potential.

Tianyi Zhang [3] dedicated his research to Large Language Model (LLM) optimization and bias reduction for personality assessment within AI. The advanced Google Gemini model alongside OpenAI GPT-3.5 enabled Pratik Pawar [7] to create multiple-choice questions (MCQs) more efficiently through automated generation and precise processing. The testing process developed by Yulia Kumar [8] brings together AI validation and educator review to produce high-quality assessments by maintaining fair testing items.

Systems that utilize gestures have advanced as an important area within communication technology development. Moresh Mukhedkar used gesture sensors alongside natural language processing to transform hand gestures into written text or verbal communication according to his research [4]. Dr. Gladiss Merlin N R [9] designed an advanced gesture-recognition system employing Haar-like features alongside SVM and neural networks and HMM which can be used for virtual reality gaming and assistive technologies.

ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14949832

The research work by Zhehuai Chen and He Huang [6] introduced SALM as a model that unites speech models with language models for simultaneous learning across speech recognition and translation. The researchers at Anagha H M [5] demonstrated the value of RSA encryption for file protection as well as real-time conversation security through their work.

III. PROPOSED SOLUTION

The combination of computer vision technology and gesture recognition systems enables the development of interactive system applications including Virtual Hand-Controlled Black Board and Virtual Painter Application [1]. OpenCV together with MediaPipe enables real-time hand tracking within the application to allow users to draw and form shapes using their fingers. The system helps asynchronous learners through its session recording capability [2]. Through its Virtual Painter application users can change drawing tools and access layer functions with an eraser mode functionality included as an intuitive feature. The selfie segmentation tool within MediaPipe helps users make precise gestures in different environmental settings [3]. The comprehensive solution improves accessibility while supporting interactivity and learning flexibility and drives student participation and inclusive education for upcoming modern classrooms.

The implementation of different enhancements aimed to enhance both the operational capabilities and expandable nature of the AI-Powered MCQ Generator. The extract_text_from_file function now contains try-except error handling blocks that protects successful execution by detecting file issues and producing understandable user alerts [4]. Ordinary prompts provided to the Gemini AI model became direct yet precise which led to faster operations alongside improved accuracy results. Python temp_file functionality enables automatic file deletion reducing both resource consumption and waste. The system performs a validation process that prevents the processing of useless or blank text entries to maintain content quality. The MCQ Generator benefits users through new features which involve MCQ output pagination along with improvements to its PDF generation tool for clear printing. The program handles big text inputs by splitting them into smaller parts to maintain performance [5]. The implementation of dynamic API key management streamlines system deployment and the new logging system supports monitoring purposes and helps with troubleshooting for the MCQ Generator in various educational settings [6].

The solution requires building a web-based multilingual text translation application through the combination of Flask framework with the googletrans library [7]. Flask operates route systems to receive user demands and the LANGUAGES dictionary utilizes ISO 639-1 language codes for displaying selectable translation preferences. Through the translate_text function the application executes translations using the googletrans module. The application operates through two essential routes namely the main routing for input and selection and the /translate pathway for translation submissions. The system checks for input before proceeding with translation according to reference [8]. Best practices for scalability and flexibility together with reliability form part of the implementation plan.

Through the application users can submit audio documents to the system for transcription followed by conversion capabilities [9]. The system processes uploaded files through file conversion to MP3 format with the help of the pydub library then uses speech_recognition to generate textual output. The system saves transcription results in a temporary file which then becomes accessible to users [11]. The application delivers smooth user experience through error handling which displays suitable error notifications to users who attempt to process unsupported files [12]. The system enables users to download their MP3 file processing results from a specified endpoint after completion of calculations.

IV. METHODOLOGY

The virtual drawing application allows users to create dynamic drawings with its features that detect hands along with webcam-based interaction. This application relies on OpenCV for image processing and NumPy for matrix operations and it also uses Mediapipe for hand detection as well as gesture recognition functions. The HandTrackingModule observes hand landmarks which enable users to control operations on the virtual canvas. Through selfie segmentation the application separates the user's hand to establish a smooth drawing area.

The second version incorporates functionalities such as freehand drawing, shape creation and erasing ability. The basis of the system stays consistent but mode selection happens through finger elevations while hand shape commands stem from multi-finger sequences. In real-time finger tracking the Euclidean distance formula determines the length between identified hand points. The mathematical formula aids real-time development of shapes with lines on the drawing surface.

- *The Euclidean Distance Formula :*

$$D = \sqrt{((x2 - x1)^2 + (y2 - y1)^2)}$$

This topic encompasses the building of a system that extracts text from diverse file types such as PDF, DOCX, and text files, to then generate multiple-choice questions (MCQs) with the help of an AI model. The system is developed for file uploads, extracts text content from PDF documents with the aid of the pdfplumber library, docx files via the docx library, and from text files with the use of UTF-8 decoding. If any errors result, like unsupported format or decoding failures, the system does output descriptive error messages to guide the users. For MCQ generation, the systems are built upon the Google Gemini AI model, where the system is fed the extracted the text, the AI is then able to formulate the required number of multiple-choice questions each with four options (A, B, C, D) and a designated answer. However, these precisely-defined MCQs are formatted in such a way as to separate their titles from the difficult part. This enables the simplified retrieval and presentation of its content. This helps combines efficient text extraction and AI powered MCQ generation to ease the task of creating education material.

ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14949832

This translation web application is designed for supporting multiple languages while using the Flask framework and the googletrans library to perform translations. Flask takes care of routing and interaction with the users while googletrans uses the API of Google Translate for the translation. A LANGUAGES dictionary supports dynamic selection of the languages through a drop down menus using built-in ISO 639-1 language codes for their descriptions. The most of the translation logic is encapsulated in the translate_text function, which increases modularity and ease of maintenance. The application implements two main routes: the home route where the HTML page is rendered, and the /translate route where translation requests are serviced. Validation and exception handling are performed so that there are no inconsistencies in performance. The application makes structured JSON responses that include the translated texts, which makes the application easier to maintain and open to future improvements.

Flask is used in this web application for routing as well as rendering, pydub for audio conversion, and speech_recognition for the transcription of the audio files. This application enables the user to upload audio files, convert speech to text, save the audio files as mp3, and download the files in a user friendly manner.

$$SNR = \frac{Psignal}{Pnoise}$$

The predefined paths TEMP_TEXT_FILE, TEMP_AUDIO_FILE, and TEMP_MP3_FILE are used for temporary files. Temporary Uploads are made possible through the root route, whereas transcription is processed in the /record route for reliable transcription services. The generated files are more accessible now, for users can download them at the /download route.

This system is integrated in such a way that the users can upload audio files, convert and transcribe speech to text, change it to mp3 format, and download presaved text or the mp3 file. Speech-to-text, converting the audio into text, is conducted through Google Speech Recognition API and for the mp3 conversion pydub is used. For the temporary files, paths are already defined, so it is easy to manage them. The uploaded content is processed, noise is removed, and then the mp3 is saved. The text is stored in the system and a jason file is presented with the data. In case of unresponsive documents, improper API connection, or faulty files, system error management pre-saves these scenarios and returns relevant status codes for further action. The FPS algorithm enhances these problems to ensure faster video processing.

➢ *Where FPS is Calculated as:*

$$FPS = \frac{Number\ of\ Frames}{Elapsed\ Time}$$

V. ALGORITHMS

➢ *Open CV :*

Computer vision (CV) is a powerful and capable-of-platform, for providing algorithms for real-time computer vision, image processing, etc, and widely used in the industry of robot, automobiles, medical, entertainment and securing domain. Since its introduction in 2000, it is an absolute necessity for image and video processing, providing a whole arsenal of algorithms for banal or complex operations. Principal operations are image transformations (scaling, rotation, cropping), filtering (i.e., Gaussian Bluring), and edge detection (i.e., Canny Edge Bluring) to localize objects and perimeters. CV also offers thresholding for binary images, applicable to object detections and segmentations. Due to the strong correlation between keypoints used by feature detection algorithms (e.g., SIFT, ORB) in different views, these can be used to perform image contouring and 3D reconstruction next.

- *Gussian Function Formula :*

$$f(x) = ae^{\frac{-(x-b)^2}{2c^2}}$$

➢ *Media Pipeline :*

MediaPipe, an open-source software package for real-time computer vision processes, combines traditional computer vision techniques with current deep learning models. This enables a superior level of accuracy and performance. With a focus on facial area detection, pose detection, hand area detection, and object detection, the pipeline architecture of MediaPipe is unrivaled. It You uses models with less complexity like MobileNet to keep the processing load as low as possible. The framework has many processing steps that start with image or video frame processing. At this stage, the input is an image or video and the output are features, models are used to get important features such as facial landmarks and body joints. After this, noise is reduced and smoothing is done to the features. The final step is post-processing lightens the load, and for applications that require real-time feedback such as AR/VR, the results are shown as fast as possible.

➢ *Image Processing :*

Image processing algorithms are used in object recognition and image segmentation analysis and image editing. The pipeline consists of preprocessing (e.g., Gaussian smoothing and convert to black white), edge extraction (algorithms-e.g., Canny Edge Detection), thresholding (global and adaptive), feature extraction and post processing (dilations and erosions). Such steps result in the fact that the image quality is also improved, as are the characteristics of the image, with the quality of the results being greater. On the other hand, image processing has already been implemented in medical imaging, for example, as well as in face recognition or driverless cars. For example, the Canny Edge Detection algorithm computes gradients, uses non-maximum suppression, and hysteresis, in a way that can achieve very high edge accuracy.

ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14949832

- *Canny Edge Detection Formula:*

$$Edge\ Strength = \sqrt{(Gx^2 + Gy^2)}$$

➢ *File Detection and Text Extraction :*

Detection and text extraction algorithms for files are a key part of creating useful representations of a large number of different file formats so an intelligent analysis can extract meaningful information. File detection algorithms first recognize the type of file under consideration, e.g. The PDF, DOCX or TXT file by analyzing its signature, metadata, etc. At the moment, the accurate extraction protocol is applied to every file extension.

Once identification of the file format is performed, methods of extraction of text, including optical character recognition (OCR) on the image or on text and the natively parsing of file content in textual form, are employed. Extraction of document text from page papers is done using pdfplumber or PyPDF2 in case of PDF pages and for DOCX files it is activated by python-docx to interpret the text inside a given document. For images and for pdfs with embedded photos, the optical character recognition (OCR) utilities like Tesseract may be employed to convert the OCRed text obtained from images into machine readable text.

The text is then sanitized and processed by getting rid of any irrelevant parts in order to clean it up for subsequent steps, with the aim of maintaining the sanctity of the information Retrieved from different file types. Error processing aids in the adequate interaction by using suitable error reports for files that cannot be processed. When everything has been carried out, they are able to download the file in MP3 format through an allocated endpoint with the aim of facilitating the use of the processing of the audio file.

VI. RESULTS

Project deployment makes use of a digital classroom board and portal in the user interface to combine user elements. This is depicted in Figures 2 and 3. The portal allows educators to come up with their own course topic, create custom test instruments, or create topic-specific performance management appraisals while the online whiteboard facilitates direct entry to the class. These tools improve the instructional design management system, motor assessment workstreams, and provide collaborative learning in both students and educators.

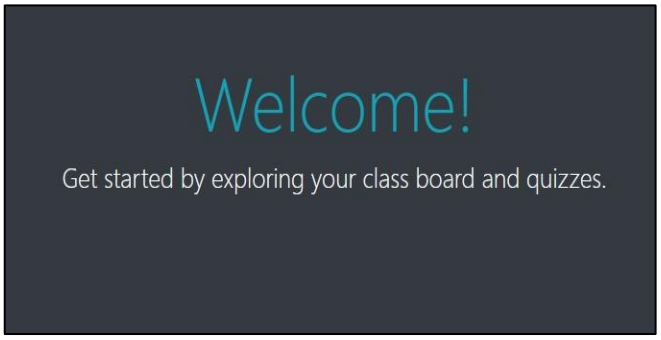

Fig 1 Deploy View

The hand gesture drawing application allows one to draw in real-time upon a blank canvas merely using hand gestures. It uses Mediapipe for hand tracking, OpenCV for the application and screen rendering, and FFmpeg for screen and audio capture at the same time. The program uses landmarks detected on users' hands to switch between drawing modes, which include color options, line/shape creation (circle, square, triangle), and an eraser.

This is a more global system, as gesture controls do not necessitate any other input devices. Gesture recognition for drawing-related hand gestures is similarly evaluated in terms of performance and is highly accurate with low latency and real-time processing. Gesture-based selection for color and brush size is also time-efficient—low latency/reaction time. Render and audio capture for the render and the tutorial is achieved with an FFmpeg plug-in.

This is the solution to effective gesture drawing in a digital age—not for personal use, virtual whiteboard classes, or a digitally focused active learning classroom. Future versions could have more refined gestures and options for other artistic applications and customization.

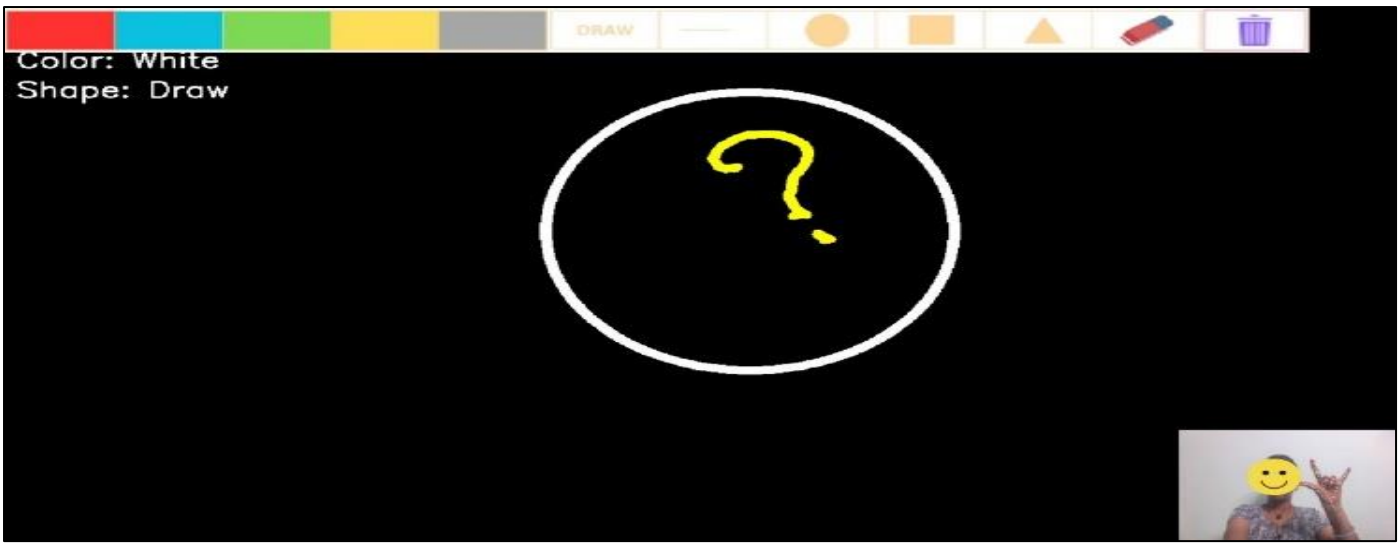

Fig 2 Virtual Black Board

ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14949832

To create a test is to go to the software's login page. The teacher logs in and fills in the blanks to create a test (test title, upload of created test document, start/stop date/time, length of the test, number of questions). The teacher logs back in afterward to view responses and grades. Nothing but the teacher can access anything once time is up—seemingly automatically. Results are stored in the software's cloud, meaning they're secure and not subject to any integrity concerns. Results are only found on the teacher's personal login portal, so all grade results are confidential and secure. Furthermore, the teacher can link relevant resources to the assessment as well, adding classroom notes to the assessment under a category she's created, yet another way to professionally link lessons and resources in one location.

Fig 3 Test Creation

To access and attempt the practice test, students are required to create a login. Each test has a unique test ID, along with designated opening and closing timings, ensuring that students can take the test within a specified window.

After completing the test, students are encouraged to post any doubts or questions they have related to the test. These doubts will be collected and reviewed by the teacher, who will address them through a dedicated teacher's dashboard. This ensures that students can clarify their concerns and enhance their understanding of the test material after completion.

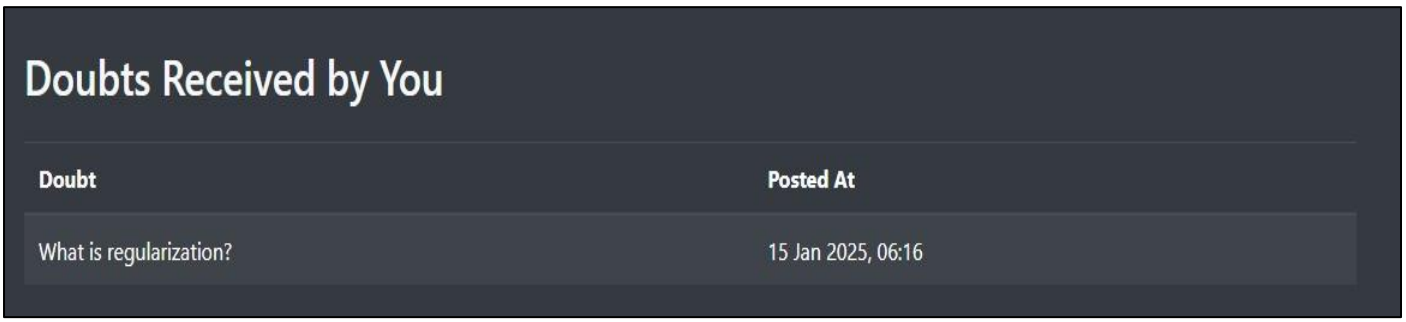

Fig 4 Doubt Posting

Every student receives Figure 4, so questioning is not only encouraged whenever—from after class, after the quiz, during study—and it offers an integrated opportunity to ask a question needed for immediate comprehension. For the teacher, she has the portal, so although it's posed by different students at different times, she will see the question posed from within her portal associated with the student's ID and the question posed. Thus, question-answer facilitated communication is seamless for the teacher to contextualize for students and facilitate clearer comprehension.

ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14949832

If Fig-5 is difficult to understand, users can fill in the "Text Input" field for linguistic aid to translate the source text. This is the first step of encoding a complicated or confusing expression. Next, users choose from a dropdown menu of 56 different languages to select the target language, allowing them to choose which translation language is appropriate. Once the target language is chosen, the "Translate" button becomes active, triggering the translation process.

Upon doing so, the system's natural language processing (NLP) engine is activated, accepting this input before generating the associated translated text for the chosen language. It allows a precise, automated process for fast, accurate transformative, surmounting language hurdles as well as enabling cross-linguistic communication and understanding for a wide variety of user populations. It allows for smooth communication in different languages, granting a better comprehension and teamwork worldwide.

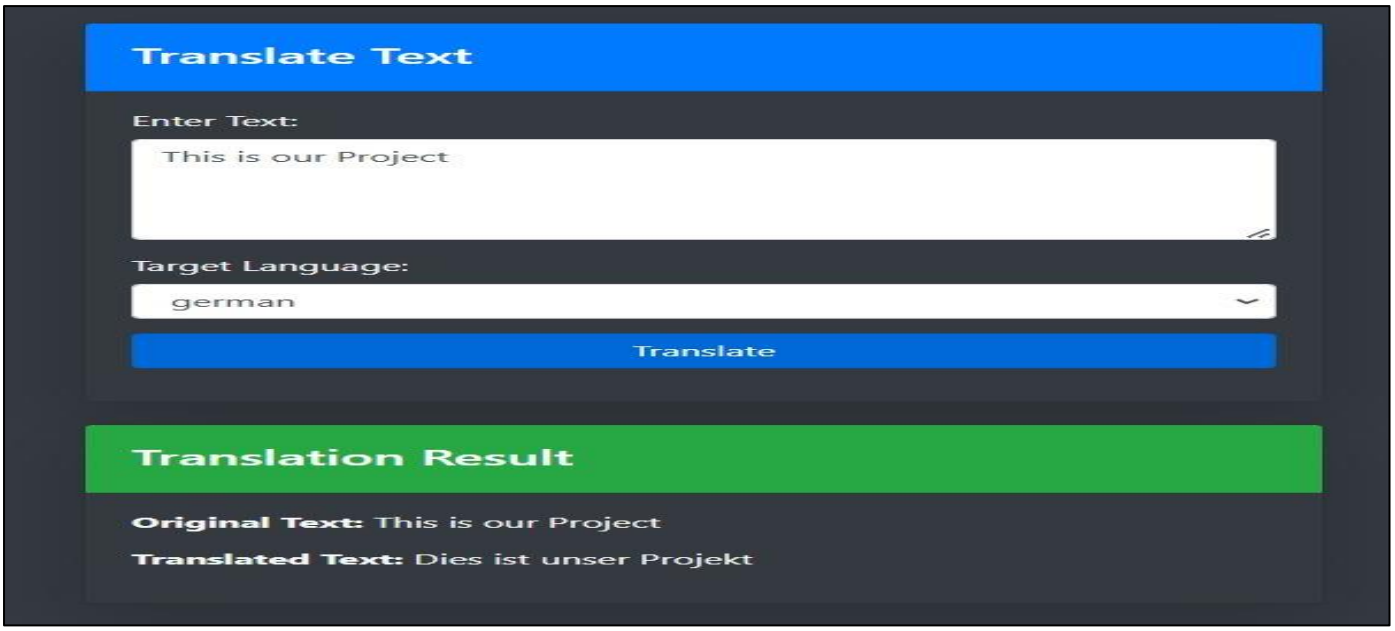

Fig 5 Multi-Lang Text Conversion

Modern technologies form the foundation of a system which extracts and processes information from spoken and video content. The system uses ffmpeg together with pydub to extract audio with no temporary file creation that enhances performance while decreasing unnecessary I/O operations and data management requirements. The automatic speech recognition (ASR) model Whisper serves as a sophisticated text transcribing system that operates well with various accents and challenging acoustic settings to maintain accuracy levels. The transcribed text undergoes condensation through the Hugging Face model facebook/bart-large-cnn because it demonstrates exceptional performance in generating and summarizing texts. BART reduces lengthy transcriptions into brief yet meaningful summaries that maintain key information to provide high-quality context-based processed material for spoken data analysis.

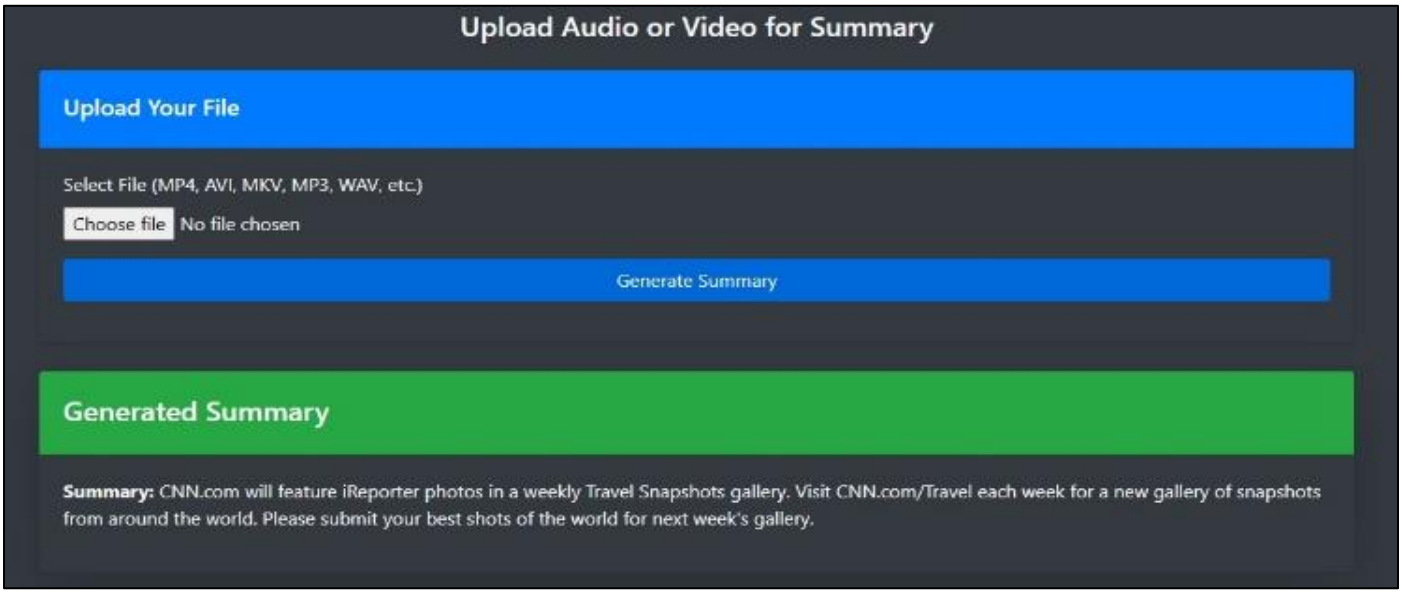

Fig 6 Notes Extraction

ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14949832

The process which merges sound extraction methods with transcription platforms and summarization software results in automatic summary generation of lengthy spoken content. Such an efficient procedure simplifies the management of substantial audio content including meeting recordings and lecture notes and multimedia document reviews. Artificial intelligence working with deep learning technology enables unattended multimedia data processing which allows organizations to improve their summarization functions in academic and business applications. The next development stage targets better audio extraction precision for different file types while implementing improved information summarization techniques to support legal and medical and scientific information fields. The system provides scalable methods that enhance the retrieval of multimedia data with increased efficiency.

The system applies modern technologies to extract while also transcribe and summarize contents within audio and video files. The system uses ffmpeg along with pydub, Whisper and BART tools to automate the conversion of long spoken content into reduced summaries which enhances both accuracy and speed. This approach delivers valuable results in three vital applications that focus on meeting transcription and lecture and multimedia analysis. The system achieves scalability through automation which eliminates manual work to benefit various fields including educational and corporate sectors. It stands as a user-friendly solution that delivers efficient multimedia data processing while extracting important information thus offering practical and effective operation in the current context.

VII. CONCLUSION

The initiative strengthens educational performance through technological solutions which fix limitations present in conventional educational techniques. Through gesture recognition technology and advanced AI learning resources and multilingual features the Virtual Lesson Management System enables personalization of educational experiences. Through the Virtual Hand-Controlled Black Board system teachers can perform digital classroom assistance with hand gestures thereby boosting engagement levels and helping with hybrid instruction. Students benefit from recorded sessions through the system because these sessions enable students to access class materials whenever real-time access becomes unavailable. Students experience an easy development process of quizzes enabled by GeniusAI technology but teachers get additional help from automation that generates test assessment responses. Students obtain better understanding and remain focused during learning because the platform enables them to develop their own quizzes for self-assessment.

Through the Collaborative Chat System students can interact live to find quick answers while developing relationships for peer learning thus forming an inclusive online classroom community. A multilingual capability of this system enables education accessibility for students with different language backgrounds. Students with different learning styles can understand lessons better because the system offers audio and video recordings. These features of the project develop an extensive educational framework which combines digital and physical classrooms through personalized and collaborative learning experiences for students.

REFERENCES

[1]. Akshay Kumar, Harishchander Anandaram, Richa Pandey, Khursheed Alam, Sumit Chaudhary& Kapil Joshi. “Hand Gesture based AI Controller for Presentation, Virtual Drawing and System Volume Management”, (2024), D.O.I:10.1109/ICIPCN63822

[2]. Subhash Chand Agrawal, Rajesh Kumar Tripathi, Neeraj Bhardwaj, & Prashun Parashar. “Virtual Drawing: An Air Paint Application”, (2024), D.O.I:10.1109/ICECAA58104

[3]. Tianyi Zhang, Antonis Koutsoumpis, Janneke K. Oostrom, Djurre Holtrop, Sina Ghassemi, Reinout E. de Vries.”Can Large Language Models Assess Personality From Asynchronous Video Interviews? A Comprehensive Evaluation of Validity, Reliability, Fairness, and Rating Patterns”, (2024), D.O.I: 10.1109/TAFFC.2024.3374875

[4]. Moresh Mukhedkar, Dilip P. Deshmukh, Om Chaudhari, Harshala Shinde, Amruta Shinde, Abhijeet Kadam. “Efficient Development of Gesture Language Translation System using CNN”,(2024),D.O.I: 10.1109/ICCCNT61001.2024.10725835

[5]. Anagha H M, Kartika Nair, Maitreyi P, Sushma E. “Providing Security for a Chat Application using RSA Encryption with GPG”, (2024), D.O.I: 10.1109/INOCON60754.2024.10511348

[6]. Zhehuai Chen, He Huang, Andrei Andrusenko, Oleksii Hrinchuk, Krishna C. Puvvada, Jason Li, Subhankar Ghosh, Jagadeesh Balam, Boris Ginsburg. “SALM: Speech-Augmented Language Model with In-Context Learning for Speech Recognition and Translation”, (2024), DOI: 10.1109/ICASSP48485.2024.10447553

[7]. Pratik Pawar, Raghav Dube, Amogh Joshi, Zinee Gulhane, Ratna Patil. “Automated Generation and Evaluation of Multiple-Choice Quizzes using Langchain and Gemini LLM”, (2024), DOI: 10.1109/ICEECT61758.2024.10739326

[8]. Yulia Kumar, Anjana Manikandan, J. Jenny Li, Patricia Morreale. “Optimizing Large Language Models for Auto-Generation of Programming Quizzes”, (2024), DOI: 10.1109/ISEC61299.2024.10665141

[9]. Dr. Gladiss Merlin N R, Tamizh Malar S G, Prathiksha K, Nivedha P, Pavithra K. “PICTOQUIZEZZ: Engaging Cognitive_Learning”, (2023), DOI: 10.1109/RMKMATE59243.2023.10369158

[10]. Yingzhe Tang, Mingzhang Pan, Hongqi Li, Xinxin Cao. "A Convolutional-Transformer-Based Approach for Dynamic Gesture Recognition of Data Gloves",(2024), DOI: 10.1109/TIM.2024.3400361

[11]. Xuehao Zhou, Mingyang Zhang, Yi Zhou, Zhizheng Wu, Haizhou Li. "Accented Text-to-Speech Synthesis With Limited Data", (2024), DOI: 10.1109/TASLP.2024.3363414

[12]. Tianrui Wang, Long Zhou, Ziqiang Zhang, Yu Wu, Shujie Liu, Yashesh Gaur, Zhuo Chen, Jinyu Li, Furu Wei. "VioLA: Conditional Language Models for

ISSN No:-2456-2165 https://doi.org/10.5281/zenodo.14949832

Speech Recognition, Synthesis, and Translation", (2024), DOI:10.1109/TASLP.2024.3434425

[13]. Wen-Chin Huang, Tomoki Hayashi, Yi-Chiao Wu, Hirokazu Kameoka, Tomoki Toda. "Pretraining Techniques for Sequence-to-Sequence Voice Conversion", (2021), DOI: 10.1109/TASLP.2021.3049336

[14]. Mingyang Zhang, Yi Zhou, Yi Ren, Chen Zhang, Xiang Yin, and Haizhou Li. "RefXVC: Cross-Lingual Voice Conversion With Enhanced Reference Leveraging", (2024), DOI: 10.1109/TASLP.2024.3439996

[15]. Mingyang Zhang, Yi Zhou, Li Zhao, Haizhou Li. "Transfer Learning From Speech Synthesis to Voice Conversion With Non-Parallel Training Data", (2021), DOI: 10.1109/TASLP.2021.3066047

[16]. Tianrui Wang, Long Zhou, Ziqiang Zhang, Yu Wu, Shujie Liu, Yashesh Gaur, Zhuo Chen, Jinyu Li, Furu Wei. "VioLA: Conditional Language Models for Speech Recognition, Synthesis, and Translation", (2024), DOI:10.1109/TASLP.2024.3434425