International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 05 | May - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Developing A Software for Dubbing of Videos from English to Other Indian Regional Languages

Anish R Gowda, Student
Department of Computer Science and Engineering, Presidency University,Bangalore
anishrgowda46@gmail.com

Venkata Sreevathsa G, Student
Department of Computer Science and Engineering, Presidency University,Bangalore
gvsreevathsa@gmail.com

Gagan Raam S, Student
Department of Computer Science and Engineering, Presidency University,Bangalore
gaganhb01@gmail.com

Udaya T K, Student
Department of Computer Science and Engineering, Presidency University, Bangalore
udayatk02@gmail.com

Dr.Swapna M
School of Computer Science and Engineering, Presidency University,Bangalore
m.swapna@presidencyuniversity.in

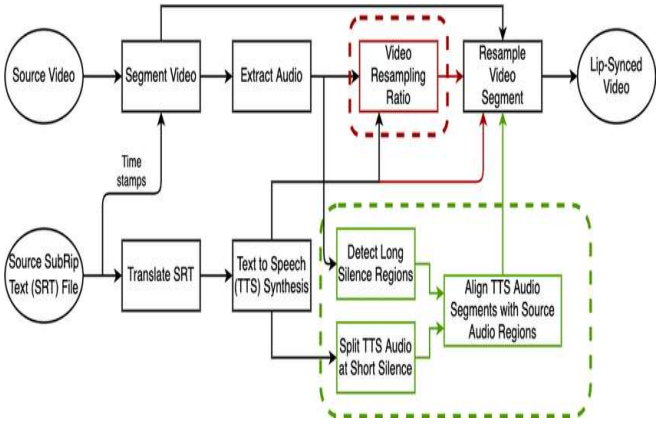
I. Abstract

This project seeks to create an advanced software solution to dub English videos into Indian regional languages, reflecting the growing demand for localized content. With increasing consumption of regional languages, the potential to widen accessibility and viewer engagement with diversified audiences is incredibly high. The software utilizes the latest natural language processing (NLP) and machine learning capabilities to automate the dubbing process, which reduces the time required and the associated costs compared to conventional manual methods. Through the analysis of English audio, speech pattern detection, and accurate translation generation, the system translates into several Indian languages such as Hindi, Kannada, Tamil, Telugu, and Malayalam. The software is designed to transcend literal translations by ensuring cultural and contextual sensitivity through the inclusion of region-specific dialects, idioms, and linguistic details, thus creating a more real and relatable experience for the audience. This method ensures inclusivity and accessibility, where digital content can be consumed in one's own language by people from various linguistic backgrounds without any hindrance. The scalability of this solution gives it worth to most industries, such as entertainment, education, and digital media, as it allows the creators and producers of content to access larger audiences without negatively affecting quality. By highlighting cultural suitability and linguistic accuracy, the software fortifies India's multicultural media ecosystem, bridging language divides and boosting content consumption. Overall, this project enables content producers to engage with local audiences better, making digital content more immersive, interactive, and widely accessible. With its capability to facilitate localization at scale, this software is poised to transform the process of adapting video content for India's linguistically diverse communities to a richer and more inclusive media experience.

Keywords: Dubbing, Accessibility, Linguistic details, Natural Language Processing (NLP), Machine learning, Speech patterns, Translations, Linguistic accuracy, Region-specific dialects, Idioms, Digital media

II. Introduction

Designing a software system for dubbing English content into Indian regional languages involves integrating technologies like Automatic Speech Recognition (ASR), Neural Machine Translation (NMT), and Text-to-Speech (TTS) synthesis in a homogenous form. The linguistic diversity of India poses the issue of high-quality, natural-sounding dubs because of differences in phonetics, syntax, and the intricacies involved in accurate lip synchronization. Nonetheless, improvements in multilingual NMT and speech translation have greatly enhanced the accuracy of translation, especially for low-resource languages spoken in India. Efficient dubbing platforms are dependent on pre-trained deep learning models, real-time speech synthesis, and adaptive timing regulation for correctness and synchronization. Empirical research has revealed that hybrid paradigms integrating statistical models and neural networks can further improve the quality of translation, especially for morphologically complex languages. Moreover, speech-to-speech translation systems are under development for real-time mobile processing, allowing more fluid language adaptation.



AI-powered dubbing solutions with emotional prosody modeling and speaker adaptation technologies improve the naturalness and expressiveness of the dubbed material. The advent of generative AI for voice cloning and synthesis has

provided new opportunities for generating high-quality, speaker-consistent dubbing. The focus of future work must be to improve processing efficiency, enhance translation quality in terms of domain specificity, and broaden linguistic coverage for supporting more Indian languages. A computer-based dubbing system adapted to the special nature of Indian languages can significantly contribute to enhancing access to education, entertainment, and information. By bridging language divides, such technology facilitates cultural preservation and integration, so that speakers of local languages have increased access to online media in their own languages. Ultimately, AI-based dubbing solutions can revolutionize content localization, making digital media more engaging, inclusive, and accessible to India's multilingual communities.

III. Literature review

Machine translation (MT) among Indian languages has been a research priority since India is a linguistically diverse country. [1] The author put forward a real-time machine translation system specifically for Indian languages that solves issues such as morphological complexity, syntactic differences, and low resource availability. They build upon past rule-based, statistical, and neural machine translation (NMT) methods and seek to enhance real-time performance and accuracy. Previous methods, including Statistical Machine Translation (SMT), exhibitedlimited success owing to the unavailability of large parallel corpora. Recent improvements in NMT, particularly Transformer-based systems, have shown dramatic improvements in translation quality. The authors capitalize on these improvements while tailoring the system to low-latency usage. Their work also takes into account the distinct grammatical patterns of Indian languages to achieve contextual correctness in translations. The research adds to the body of work by concentrating on real-time application, which is usually not addressed in traditional MT research. The authors incorporate deep learning methods to improve the performance of Indian language translation systems. Their method has the potential for use in education, governance, and media to enable easy communication across linguistic divides. Future research in this area can look into domain-specific optimizations and additional low-resource language translation improvements.

Neural Machine Translation (NMT) has transformed language translation, yet low-resource languages like Indian languages remain in the background due to the lack of parallel corpora. Various research has tried NMT methods to overcome these issues and enhance the quality of translations.[2] Traditional methods like Statistical Machine Translation (SMT) and rule-based techniques failed to deal with the morphological richness and syntactic variations in Indian languages. The advent of deep learning-powered NMT models, especially sequence-to-sequence models with attention, greatly improved translation quality. Nevertheless, base NMT models need large volumes of data to work well, which makes them less ideal for languages with fewer linguistic resources. To address the problem, transfer learning and data augmentation techniques have been explored, such as pre-trained multilingual models and fine-tuning approaches to improve underrepresented language translations. Findings show that utilizing high-resource languages and using domain adaptation techniques result in dramatic translation quality improvements. This work proves that low-resource Indian languages can be assisted with state-of-the-art NMT techniques via pre-training and optimization of linguistic resources. Synthetic data generation and active learning may be areas for further development to further improve translation accuracy and enhance the capabilities of NMT systems.

Multilingual Neural Machine Translation (MNMT) has been essential in improving translation quality for more than one language, especially in linguistically rich Indian regions. [3] MNMT systems for Indic languages have been the primary focus of research to alleviate data sparsity, domain adaptation, and linguistic differences between languages. Earlier research has established that basic NMT models, although competent in high-resource languages, have difficulties with low-resource languages owing to the restricted availability of parallel corpora. To address such problems, some techniques like transfer learning, subword tokenization, and common representations across related languages have been investigated. Multilingual corpora training with language-specific optimization has been found to improve translation quality and fluency. Transformer models have showcased advancements, especially for low-resource Indic languages, by utilizing fine-tuning methods and adaptive training procedures. Research shows that multilingual training aids not only individual language translation but also enhances cross-lingual knowledge transfer. Advanced improvements in the future might involve using large-scale generative language models and reinforcement learning methods to further tune MNMT systems for Indian languages.

Machine translation for Indian languages has made great strides, with increasing emphasis on enhancing translation quality and support for all 22 scheduled Indian languages.[4]Recent advances have brought about IndicTrans2, a high-quality multilingual neural machine translation (MNMT) model that is designed to provide more fluent and accurate translations while tackling issues like low-resource languages, data skew, and linguistic variety. Developing on earlier MNMT models such as IndicTrans, the present system applies advanced training strategies, bigger corpora, and enhanced Transformer-based structures. It has been seen through earlier research that multilingual training is of greater benefit; however, retaining uniformity for languages with data sets of dissimilar sizes poses a problem. For this issue, large-scale synthetic data production, advanced tokenization methods, and language-tailored adjustment have been enforced. Besides, various attempts have been made to make IndicTrans2 accessible, enabling broad usage in education, government, and media industries. Future research shall further improve IndicTrans2 by integrating generative AI models and domain-specific fine-tuning for better contextual accuracy.

Translation among Indian languages, especially those that are closely related to each other such as Kannada and Tamil, is

challenging because of variations in script, morphology, and syntax.[7]A readable translation framework has been proposed to translate simple Kannada sentences into Tamil with accuracy and fluency by taking care of linguistic differences. Traditional translation mechanisms were based on rule-based systems, which performed poorly with compound sentence structures and contextual knowledge. Statistical Machine Translation (SMT) subsequently enhanced translation quality but at the expense of huge parallel corpora, which are usually lacking for Indian language pairs. A hybrid approach of rule-based and statistical methods has been followed in an attempt to increase translation reliability. The system uses morphological analyzers and syntactic parsers to better capture language nuances, and grammatically correct translated output is ensured. This study emphasizes the role of language-specific fine-tuning in machine translation and the requirement for high-quality linguistic resources. Future developments can involve Neural Machine Translation (NMT) and deep learning-based models for further enhancing translation accuracy and contextual comprehension.

Speech translation from English to Dravidian language is challenging because of structural, phonetic, and syntactic differences.[8]A speech translation system has been suggested to translate English speech into Dravidian languages to overcome issues in pronunciation and word-order variance. Existing speech-to-text and text-to-text translation systems have been shown to be effective with rule-based and statistical approaches, but these tend to lack fluency and contextuality. To improve performance, machine learning paradigms with automatic speech recognition (ASR) and machine translation (MT) modules have been used. The system uses phoneme mapping and language modeling to enhance translation accuracy. This work contributes to the creation of real-time speech translation software, especially for low-resource Dravidian languages, by providing better alignment between spoken English and the target languages. Future development will possibly include speech translation neural models and deep learning algorithms to better enhance pronunciation modeling and contextual fluency.

Speech-to-speech translation (S2ST) systems are important tools in overcoming the language barrier, especially in multilingual nations such as India.[11]A bidirectional speech translation system has been built to enable local travel information sharing between Indian languages. Conventional text-based rule-based and statistical machine translation (SMT) approaches were confronted with difficulties in managing real-time spoken language variability. Building upon progress in text-to-speech synthesis (TTS) and automatic speech recognition (ASR), this system incorporates an S2ST framework. It is composed of three principal modules: ASR for speech input processing, MT for translating speech that is recognized to the target language, and TTS for producing human-sounding output speech. Linguisitic issues like phonetic differences, code-switching, and domain terms have been addressed to expand practical application in real-world travel situations. This research makes a contribution to speech translation by providing an actual-time multilingual communication solution. Future developments can include deep learning-based neural machine translation (NMT) and end-to-end speech translation models to enhance fluency, minimize latency, and accommodate more Indian languages and dialects.

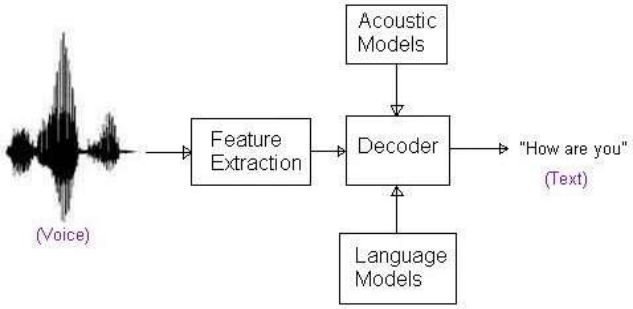
Video dubbing and translation are becoming more crucial for making multimedia content available in various languages. A system of automated video translation and dubbing has been implemented to overcome limitations in synchronizing speech with video content while being natural and understandable.[14]Conventional video translation was based on manual transcription, translation, and voice-over dubbing, which was time-consuming and expensive. Recent developments in artificial intelligence (AI) and deep learning have made it possible to develop computerized systems incorporating speech recognition, machine translation, and text-to-speech synthesis. Using these technologies, an uninterrupted workflow has been developed to provide precise speech translation while maintaining the tone of the speaker and lip-sync precision. The system uses neural machine translation (NMT) methods, speech synthesis models, and adaptive timing algorithms to improve dubbing quality. In comparison to traditional methodologies, the method enhances efficiency through the automation of the whole pipeline, meaning extensive human involvement is eliminated. This study adds to the growing body of AI-based multimedia localization research, as it makes content more accessible to a wide linguistic crowd. Future development is likely to involve enhanced emotional prosody modeling, real-time adaptation in live streaming, and the incorporation of generative AI models for more natural and expressive voice dubbing.

IV. METHODOLOGY

Youtube Video Downloading: Download YouTube videos to be processed offline within a dubbing software framework that allows the unobstructed integration of Automatic Speech Recognition (ASR), Neural Machine Translation (NMT), and Text-to-Speech (TTS) technology. Fast video recall is key to the quality of the extracted audio, which, in turn, is imperative to recognize and translate speech correctly[14]. Studies on multilingual speech-to-speech translation are emphatic regarding the pre-processing of video material for better performance of the system[11][18]. When a video is downloaded, the audio is retrieved and processed using ASR for transcription. The transcribed text is translated to Indian regional languages and synthesized using TTS to produce a fully dubbed version for wider reach.

Audio Extraction from Video:Audio extraction from the video is a crucial process in creating a dubbing software system that allows speech processing for transcription and translation purposes. Audio extracted is used as input for Automatic Speech Recognition (ASR) that translates speech content into text to be processed further[11][14]. Speech-to-speech translation research highlights the need for good quality audio extraction in order to improve the accuracy of ASR and reduce interference from background noise[5][18]. After extracting the audio, ASR produces transcripts, which are then translated into Indian regional languages with the

help of Neural Machine Translation (NMT). The text is then read out using Text-to-Speech (TTS) so that dubbing becomes smooth.



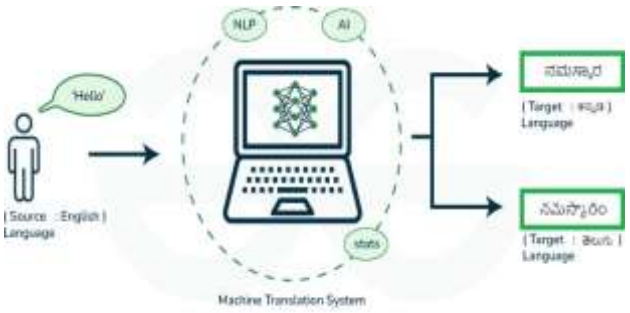
ASR

Automatic Speech Recognition (ASR): It is essential in creating a software system for dubbing videos from English to Indian regional languages. ASR technology translates spoken words into text, which is the basis for machine translation and speech synthesis in dubbing systems. Recent studies have investigated ASR enhancements for low-resource languages using deep learning models, transfer learning, and multilingual training to improve accuracy and flexibility[8][11][18]. Real-time ASR has been adopted in mobile and cloud environments to provide transparent speech recognition across varying linguistic contexts[5][16]. ASR in a dubbing platform interprets the English original speech, which is next fed through NMT for proper text output in the target language. Translated text is generated using Text-to-Speech technology to create the dubbed voice. Through improving Indian language ASR models via domain adaptation and robustness against noise, overall dubbing quality is improved a great deal such that it yields context-sensitive translations in an accurate manner.

$$\hat{W} = \arg\max_{W} P(W|X)$$

The picture has the basic Automatic Speech Recognition (ASR) equation, which relies on Bayes' Theorem. The equation is employed in speech-to-text applications to identify the most probable words spoken from audio inputs.

Subtitling and Timestamping: Subtitling and timestamping are essential to align translated text with video content, keeping the dialogue in sync with on-screen action. Speech translation research points to the importance of exact timing to preserve the natural flow of translated content[14][18]. Automatic Speech Recognition (ASR) produces transcripts from audio extracted, which are then translated into Indian regional languages by Neural Machine Translation (NMT)[11]. Timestamping methods aid in breaking speech into correctly synchronized subtitle blocks, enhancing readability and usability. Machine translation research for low-resource languages stresses maximizing subtitle accuracy and timing tweaks for enhanced understanding[2][4]. Timestamped subtitles aid in synchronizing Text-to-Speech (TTS) output with original speech in a dubbing system, maintaining natural pace and lip-sync enhancements. Incorporating AI-based synchronization techniques maximizes efficiency and makes dubbing smooth and accessible to multilingual viewers.

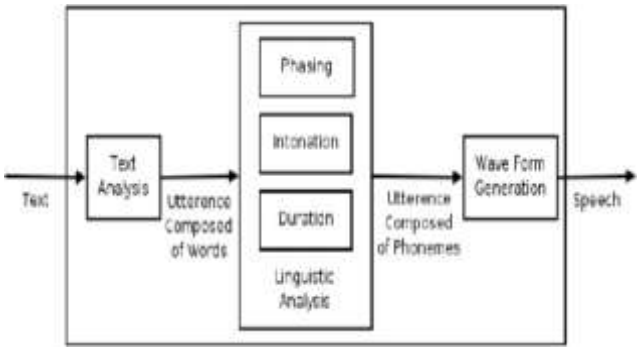


NMT

Neural Machine Translation (NMT): Neural Machine Translation (NMT) is a deep learning approach that improves translation accuracy by capturing linguistic and contextual features, and hence is extremely effective for English to Indian regional language translation. Multilingual NMT research has established better performance, particularly for low-resource Indian languages, by using transfer learning, parallel corpora, and transformer models[2][3][4].State-of-the-art NMT models fine-tune fluency, syntax, and context relevance to deliver high-quality dubbings for dubbing purposes[10][17]. Within a dubbing software framework, NMT has the important function of translating ASR-generated English subtitles into target Indian languages with semantic and syntactic accuracy[11][18]. The translated subtitles are further converted using Text-to-Speech (TTS) technology to produce natural-sounding dubbed speech. By using domain adaptation methods and fine-tuning NMT models for Indian languages, the software is able to deliver more accurate and culturally sensitive translations, which increases accessibility and user experience.

$$\mathcal{L}(\theta) = \frac{1}{N}\sum_{n=1}^{N}\sum_{i=1}^{T_y} logp(y_i^{(n)}|y_{<i}^{(n)}, X^{(n)}, \theta) \quad (4)$$

The equation displayed in the photo is an expression of the loss function applied in Neural Machine Translation (NMT) models, specifically in sequence-to-sequence learning probabilistic modeling. It is a log-likelihood function that optimizes the probability for producing the target sequence correctly.



TTS

Text-to-Speech (TTS): Text-to-Speech (TTS) technology transforms translated text into natural-sounding speech and is a crucial part of video dubbing systems. Speech synthesis research has enhanced pronunciation, prosody, and expressiveness, and improved multilingual TTS performance for Indian languages[8][14][18]. Deep learning-based TTS models provide high-quality voice output without losing tone and emotion[9]. TTS transforms Neural Machine Translation (NMT) output in a dubbing system to produce synchronized speech along with video timestamps[11]. Using AI-powered voice cloning and speaker adaptation, the platform can create realistic and captivating dubbed content, and make videos more accessible to different linguistic communities.

$$\hat{Y} = \arg\max_{Y} P(Y|W)$$

The photo shows the Text-to-Speech (TTS) probabilistic equation, where Y^ is the most likely speech waveform from text input W. It controls speech synthesis by optimizing P(Y| W), which guarantees natural-sounding sound for dubbing and accessibility applications.

Video Processing and Merging Audio/Video: Video processing and audio-video merging are essential processes in the creation of a dubbing system, providing flawless synchronization between translated speech and the visual content. Studies in speech-to-speech translation highlight the need for accurate alignment between synthesized audio and native video timing[11][14][18]. Improved video processing techniques vary playback speed, deal with latency, and optimize lip-sync performance, enhancing viewer experience[9]. Once the voice is created by Text-to-Speech (TTS) from translation, it is combined with the native video while preserving accurate timestamps[8]. Artificial intelligence-based synchronization tools fine-tune timing modifications, providing natural dubbing output for Indian regional language multilingual viewers.

Document Creation: Document generation is necessary for transcript, translation, and subtitle storage and management in a video dubbing framework. Multilingual translation and speech-to-text processing research identifies the necessity of formatted text types to enable efficient workflow[2][4][11].

Automatic Speech Recognition (ASR) produces transcripts, which are translated via Neural Machine Translation (NMT) and saved for subsequent processing[3][18]. These documents enable precise subtitle synthesis and Text-to-Speech (TTS) generation[8][14]. Through arranging data in orderly forms such as SRT or XML, dubbing software makes the right synchronizations and is available for multilingual content dissemination of Indian regional languages.

V. Results & Discussions

This video dubbing program easily dubs English videos into Indian regional languages. The user can upload a video file or insert a YouTube link, select target languages, and enable auto-dubbing. The system captures audio, translates it, and creates subtitles and voiceovers, providing a quick, efficient, and accurate dubbing experience for multilingual audiences.

This is the UI of the dubbing software that is used to translate English videos into Indian regional languages. A test video is uploaded, and an alternative input of a YouTube link is given. Some target languages such as Kannada, Telugu, Hindi, Tamil, and so on are chosen. After clicking on the "Translate & Subtitle" button, the system captures audio, translates it, and automatically creates dubbed voiceovers and subtitles.

This picture indicates a user choosing a video file from the system storage to upload into the dubbing software. The software interface provides users with an option to select a file to translate and dub into Indian regional languages. Upon selection, the user can choose a target language and start the translation process, making content more accessible across linguistic regions.

The image here depicts the video dubbing software successfully translating content. A test video was uploaded by the user and the target language chosen was Kannada ("kn"). The software finished the translation and dubbing, showing confirmation messages. The user is now able to download the translated video and the document, which makes creating multilingual content more efficient and convenient.


Translations

Original Text:

the team we have this time around is amazing when the auction started I think all of us were a bit you know nervous as the auction brings that energy on to the table and there was a lot of uncertainty on day one what can happen who is going where and all that kind of stuff but I think the way are management and everyone maintain their composure to get exactly what we wanted was amazing to see in the end we created A beautifully balance squad with a lot of match winners a lot of leaders who are you know amazing players in their own right it will be a great help to someone like Rajat as well to have such experience such matters around him for him to just focus on what needs to be done and I'm sure all the players will show their full support to Rajat and also allow him to grow into this into this role of responsibility that he has been given I'm looking for to play with all of them and as I said create a new energy and create some amazing

Translation in kn:

The screenshot shows the output of video dubbing software that translates English video into Indian regional languages. The original English transcript and its Kannada translation are shown, illustrating how the system captures and translates spoken content into contextually correct regional language text for localization and accessibility.

VI. Conclusion

Building a software system to dub English videos into Indian regional languages is an important step in bringing digital content more within reach and inclusive. By combining Automatic Speech Recognition (ASR), Neural Machine Translation (NMT), and Text-to-Speech (TTS), the project facilitates smooth, high-quality dubbing with respect to linguistic and cultural sensitivities. Advanced video processing methods add extra strength to synchronization, ensuring a natural watching experience. The potential for the future of this project is huge, particularly with the exponential growth in AI-based speech synthesis and real-time language processing. With further developments in deep learning models, speech recognition and translation accuracy will be higher, which will result in more natural and contextually relevant dubbed content. Furthermore, integrating personalized voice synthesis and emotion-sensing TTS will make the user experience more advanced, where AI-generated voices become more expressive and human. With growing digital consumption, this kind of system can be helpful for educational material, entertainment industries, and differently-abled individuals' accessibility solutions. Future enhancement can also extend to real-time dubbing facilities, enabling live translations for international conferences, web lectures, and streaming services. By enhancing language models, enhancing processing efficiency, and increasing linguistic coverage, this project can change the face of content localization, making linguistic diversity not a hurdle to information and entertainment.

VII. References

[1] Patil, A. H., Patil, S. S., Patil, S. M., & Nagarhalli, T. P. (2022). Real Time Machine Translation System between Indian Languages. In Proceedings of the 6th International Conference on Computing Methodologies and Communication (ICCMC) (pp. 1-5). IEEE.

[2] Choudhary, H., Rao, S., & Rohilla, R. (2020). Neural Machine Translation for Low-Resourced Indian Languages. In Proceedings of the Twelfth Language Resources and Evaluation Conference (pp. 3610-3615). European Language Resources Association.

[3] Das, S. B., Biradar, A., Mishra, T. K., & Patra, B. K. (2022). Improving Multilingual Neural Machine Translation System for Indic Languages.

[4] Gala, J., Chitale, P. A., Raghavan, A. K., Gumma, V., Doddapaneni, S., Kumar, A., Nawale, J., Sujatha, A., Puduppully, R., Raghavan, V., Kumar, P., Khapra, M. M., Dabre, R., & Kunchukuttan, A. (2023). IndicTrans2: Towards High-Quality and Accessible Machine Translation Models for all 22 Scheduled Indian Languages.

[5] Guan Y, Zheng L, Tian J (2010) Real-time speaker adapted speech to speech translation system in mobile

environment. In: 10th international conference on signal processing (ICSP). IEEE, pp 577–580
[6] Condon S, Arehart M, Parvaz D, Sanders G, Doran C, Aberdeen J (2012) Evaluation of 2-way iraqi arabic - english speech translation systems using automated metrics. Mach Transl Springer 26(1):159–176
[7] Rajaram BSR, Ramakrishnan AG, Kumar HRS (2013) An accessible translation system between simple kannada and tamil sentences. In: Proceedings of 6th language and technology conference
[8] Sangeetha, J., Jothilakshmi, S. Speech translation system for english to dravidian languages. Appl Intell 46, 534–550 (2017)
[9] Hema Priya, K., Akhilan, N., Aravindh, R., & Janardhana, K. (2024). An In-Depth Investigation into Automatic Dubbing Leveraging ASR, Machine Translation and Deep Voice 3. In S. Manoharan, A. Tugui, & Z. Baig (Eds.), Proceedings of 4th International Conference on Artificial Intelligence and Smart Energy (ICAIS 2024) (pp. 1–12).
[10]Patel, R.N., Pimpale, P.B., Sasikumar, M.: Machine translation in Indian languages: challenges and resolution. J. Intell. Syst.Intell. Syst. 28(3), 437–445 (2019).
[11] Vemula, V. V. B., Narne, P. K., Kudaravalli, M., Tharimela, P., & Prahallad, K. (2010). ANUVAADHAK: A Two-way, Indian Language Speech-to-Speech Translation System for Local Travel Information Assistance. International Journal of Engineering Science and Technology, 2(8), 3865-3873.
[12] Kasthuri, M., Kumar, S.B.R.: Rule based machine translation system from English to Tamil. In: 2014 World Congress on Computing and Communication Technologies, Trichirappalli, India, pp. 158–163 (2014).
[13] Aasha, V.C., Ganesh, A.: Machine translation from English to Malayalam using transfer approach. In: 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Kochi, India, pp. 1565–1570 (2015).
[14] Kumar, K. S., Aravindhan, S., Pavankumar, K., & Veeramuthuselvan, T. (2023). Autodubs: Translating and Dubbing Videos. In V. S. Rathore, V. Piuri, R. Babo, & M. C. Ferreira (Eds.), Emerging Trends in Expert Applications and Security (ICETEAS 2023) (pp. 1–10).
[15] Nimbalkar S, Baghele T, Quraishi S, Mahalle S, Junghare M (2020) Personalized speech translation using google speech API and Microsoft translation API. In: Proceedings of international research journal of engineering and technology (IRJET)
[16] Yun S, Lee Y-J, Kim S-H (2014) Multilingual speech-to-speech translation system for mobile consumer devices. IEEE Trans Consum Electron 60(3):508–516.
[17] Naveen Arivazhagan, Ankur Bapna, Orhan Firat, Dmitry Lepikhin, M. Johnson, M. Krikun, M. Chen, Yuan Cao, G. Foster, Colin Cherry, Wolfgang Macherey, Z. Chen, and Y. Wu. 2019. Massively multilingual neural machine translation in the wild: Findings and challenges
[18] Mhaskar, S., Bhat, V., Batheja, A., Deoghare, S., Choudhary, P., & Bhattacharyya, P. (2023). VAKTA-SETU: A Speech-to-Speech Machine Translation Service in Select Indic Languages.