Analysis of MP3 Bitrate on the Accuracy of Academic Audio Transcription Using Whisper large-v3

Selta Jaya Putra
Informatics Engineering
Faculty of Engineering
Muhammadiyah University of Bengkulu
Bengkulu, Indonesia
seltajaya.16@gmail.com

Ardi Wijaya
Informatics Engineering
Faculty of Engineering
Muhammadiyah University of Bengkulu
Bengkulu, Indonesia
ardiwijaya@umb.ac.id

RG. Guntur Alam
Informatics Engineering
Faculty of Engineering
Muhammadiyah University of Bengkulu
Bengkulu, Indonesia
rggunturalam@umb.ac.id

Abstract— In the digital era, automatic transcription is a crucial solution for converting audio content into text, especially in the context of academic documentation. The main challenge in this process is transcription accuracy, which can be affected by the quality of the audio file, including its bitrate and file size. This study aims to analyze the impact of MP3 bitrate and file size on transcription accuracy using the Whisper large-v3 model. Five academic audio files were converted into five different bitrate levels, ranging from 64 kbps to 320 kbps, and then transcribed automatically using the Whisper model. Evaluation was conducted by calculating the Word Error Rate (WER) as an indicator of transcription accuracy. Additionally, processing time and file size were recorded to assess transcription efficiency. The results show that increasing the bitrate does not always lead to higher accuracy. Bitrates of 128–192 kbps provided the best balance between transcription accuracy, processing efficiency, and file size. This study makes a significant contribution to the development of automatic transcription systems based on ASR models, particularly for audio documentation needs in educational institutions. These findings serve as a technical reference for developing efficient and accurate audio documentation systems in academic environments.

Keywords— Whisper, bitrate, MP3, audio transcription, WER.

I. INTRODUCTION

Learning media have a significant influence on the teaching and learning process. [1], [2]. The industrial and educational revolutions have co-occurred, significantly impacting human life. The Industrial Revolution has had a profound impact on educational changes. With technological developments, a new challenge arises for educators to optimally utilize available technology [3], [4]. The use of technology in education can be implemented through digital learning media as a means to increase the effectiveness of the teaching and learning process [5]. One method that has seen a significant rise in the e-learning trend over the past decade is video-based learning [6]. However, to improve learning effectiveness, video content must focus on core information by eliminating extraneous or irrelevant content [7].

The application of Artificial Intelligence (AI) and machine learning can be an innovative strategic solution to address this challenge. One such approach is Automatic Speech Recognition (ASR), which enables the conversion of speech into text [8], [9]. This technology has become an essential tool for accelerating access to information from audio sources such as instructional videos, seminars, and academic testimonies. The main challenge in this process is ensuring high transcription quality despite the varying technical quality of audio files.

A solution to this issue is the automatic conversion of audio files into text using ASR technology [10], [11]. One of the most advanced ASR models today is Whisper, developed by OpenAI [12]. Whisper is renowned for its capability to handle various languages and audio conditions, utilizing a transformer-based encoder-decoder architecture [13]. Whisper adopts a Large-Scale Weak Supervision approach and is designed to function effectively across diverse languages and sound environments. This model is also equipped with multilingual support, transformer architecture, and fine-tuning processes that enhance transcription quality, even for low-quality audio files.

This study aims to address the problem by focusing on the technical parameters of bitrate and file size in MP3 format. Bitrate is one of the crucial factors that influence audio quality in digital files [14]. However, few academic studies have specifically examined the effect of bitrate on ASR system transcription results, such as those generated by Whisper. Therefore, this research contributes not only to the development of automatic transcription technology but also provides technical guidelines for more efficient and accurate academic audio archiving.

Several previous studies have evaluated Whisper and other transformer-based models for transcription and text summarization purposes. Research has shown Whisper's effectiveness in various languages [15]. For example, the study titled "YouTube Transcript Summarizer using Natural Language Processing" applied NLP for summarizing YouTube videos and emphasized the importance of content filtering [16]. Khoiroh et al. [17] Demonstrated that Whisper, as a speech recognition tool, achieved high accuracy between 87% and 97% in transcribing political debates. Further research discussed Whisper's robustness under high noise conditions [18]. Meanwhile, Sonata [19] Showed that transformer models performed comparably for both Indonesian and English datasets, indicating the strengths of models like Whisper in both local and global contexts.

A study by Fadlilah [20] Evaluated the performance of Whisper model variants and the T5 model for summarizing educational videos, showing that Whisper Turbo proved to be the most efficient with the fastest transcription time and low WER score. This research highlighted the superiority of Whisper Turbo in terms of efficiency and accuracy, but did not analyze the impact of bitrate or MP3 file size used. This GAP forms the basis of the current study. Few studies have investigated the effect of bitrate variation on transcription results using the Whisper model, particularly in the context of academic audio content. Therefore, this study aims to evaluate the impact of MP3 file size and bitrate on transcription accuracy using Whisper large-v3.

The objective of this study is to determine the optimal bitrate that can produce high transcription accuracy without excessively increasing file size. It also aims to provide technical recommendations for the processing of academic audio documents to become more efficient and informative through reliable and accurate speech-to-text conversion.

II. METHODOLOGY

This study aims to analyze the impact of MP3 bitrate variations and file sizes on the accuracy of automatic transcription using the Whisper large-v3 model. The research was conducted systematically, beginning with the collection of academic audio data, followed by bitrate conversion, automated transcription, and concluding with performance evaluation.

A. Research Steps

The research process was designed to ensure each stage was carried out efficiently and systematically. The first step involved collecting audio data from academic documentation at the University of Bengkulu, including national seminars, workshops, leadership testimonials, internal grant monitoring events, and inter-university visits. The dataset was selected to represent a diverse range of educational, technological, and institutional social topics, as well as variations in accents, background noise levels, and speaker delivery styles.

Before the bitrate conversion process, each audio file was analyzed for sound quality, language structure, and pronunciation clarity—this initial analysis aimed to minimize potential recognition errors by Whisper, especially concerning Indonesian phonetics.

The next step was bitrate conversion of the audio files into five variants (64 kbps, 128 kbps, 192 kbps, 256 kbps, and 320 kbps) using the FFMPEG software [21]. The conversion produced MP3 files of varying sizes depending on the bitrate. Following this, automatic transcription was performed using the Whisper large-v3 model, implemented through the faster-whisper library in a Python environment.

The transcription results for each bitrate variant were then compared to manually prepared reference transcripts to calculate the Word Error Rate (WER) using the jiwer library [22]. Additionally, transcription processing time and file size were recorded as supplementary variables to examine correlations between bitrate, file size, processing duration, and transcription error rates.

An overview of the research workflow is presented in Figure 1.

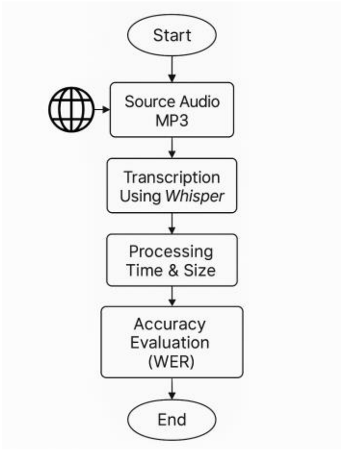

Figure 1. Research Stages Flow

B. Academic Audio Data Sources

The audio files used in this study originated from academic documentation at the University of Bengkulu, including national seminars, leadership testimonials, internal grant monitoring events, and inter-university visits. These audio contents were selected to reflect the diversity of academic contexts and covered topics related to education, technology, and institutional social issues.

Before the transcription process, each audio file was analyzed based on sound quality, language structure, and clarity of pronunciation. This aimed to minimize recognition errors by the Whisper model, particularly in phonetics and Indonesian pronunciation. This curation process was essential to ensure that the resulting transcriptions were accurate and representative.

The selection of audio files was carefully performed to represent a range of noise levels, accents, and speaking styles. This allowed for a more comprehensive and realistic evaluation of the effects of bitrate and MP3 file size on transcription accuracy, reflecting actual academic documentation conditions in higher education environments.

C. Transcription Using the Whisper Model

Whisper is an Automatic Speech Recognition (ASR) system developed by OpenAI using a Large-Scale Weak Supervision approach [23]. The model was trained on diverse datasets consisting of audio from various recording environments, speakers, and languages, resulting in a robust and flexible system [24]. Whisper utilizes a transformer-based encoder-decoder architecture and is available in multiple sizes, with the largest model containing up to 1.55 billion parameters [25].

Whisper supports over 98 languages and is designed to handle a wide range of multilingual speech recognition tasks. Its training data format was specially crafted to optimize performance under complex real-world conditions. This study employed the Whisper large-v3 variant, which is particularly capable of processing audio with complex sentence structures and background noise [26].

The Whisper model was implemented using the faster-whisper library in a Python environment, without requiring additional fine-tuning, and was therefore evaluated based on its original capabilities. Each audio file, converted into five MP3 bitrate levels, was processed sequentially for accuracy evaluation. These transcription outputs were then compared to manual reference transcripts using the Word Error Rate (WER) metric [27].

D. Transcription Evaluation

The transcribed text from each MP3 audio file was used as the basis for evaluating the accuracy of the Whisper large-v3 model in converting speech to text. The evaluation was conducted by comparing the transcription results to manual reference transcripts using the Word Error Rate (WER) metric [28]. WER measures the percentage of word errors in the transcription process, including deletions, insertions, and substitutions [29].

WER values were calculated automatically using the Jiwer library, allowing for fast and objective evaluations of each transcription result [30]. In addition to WER, transcription time and MP3 audio file size were also recorded as supplementary parameters to assess processing efficiency across different bitrate variations.

By observing the WER values and their correlation with bitrate and file size, the extent to which audio quality influences transcription accuracy can be determined. This evaluation aims to identify the most efficient and accurate bitrate configuration for use in Whisper-based automatic transcription systems.

III. RESULTS AND DISCUSSION

The transcription evaluation was conducted using five audio files with varying durations, representing diverse academic events, including seminars, workshops, and testimonials. Each file was converted into five different bitrate levels (64 kbps to 320 kbps) using FFMPEG, producing a total of 25 audio combinations. These files were transcribed using the Whisper large-v3 model, which is known for its robustness in handling different languages and audio conditions. All audio files had manually prepared reference transcripts, enabling accurate comparison using the Word Error Rate (WER) metric. Additionally, processing time and file size were recorded for each combination to assess the efficiency and technical impact of bitrate variation on transcription performance.

Table 1. Comparison of Bitrate and Size Variations Based on Transcript Processing Time and Word Error Rate (WER) Value

Title	Bitrate	File Size	Processing Time (sec)	Word Error Rate (WER)
Testimoni Rakerpim PT 2024 LLDIKTI II Rektor UM Bengkulu	64 Kbps	1,251 KB	203,54	21,93%
	128 Kbps	1,691 KB	182,37	26,32%
	192 Kbps	2,427 KB	192,22	24,56%
	256 Kbps	3,162 KB	194,87	29,82%
	320 Kbps	3,898 KB	200,89	27,19%
Workshop STEM BASE CODING Prodi Matematika FKIP UM BENGKULU	64 Kbps	1,195 KB	407.86	17,49%
	128 Kbps	2,040 KB	633,03	37,16%
	192 Kbps	2,912 KB	263.69	18,58%

	256 Kbps	3,784 KB	378.72	17,49%
	320 Kbps	4,657 KB	452,54	39,34%
Penerimaan Mahasiswa KKN UM Surakarta ke UM Bengkulu 2024	64 Kbps	1,322 KB	229,17	13,30%
	128 Kbps	1,915 KB	225,10	18,09%
	192 Kbps	2,695 KB	230,07	18,09%
	256 Kbps	3,476 KB	239,16	17,02%
	320 Kbps	4,256 KB	227,53	17,02%
Workshop Nasional Technopreneurship Ignite Your Techopreneurial Spirit	64 Kbps	1,343 KB	190,53	11,24%
	128 Kbps	1,856 KB	196,72	12,43%
	192 Kbps	2,683 KB	199,94	11,83%
	256 Kbps	3,511 KB	197,45	13,02%
	320 Kbps	4,338 KB	195,08	11,83%
Monitoring dan Evaluasi Penelitian dan Pengabdian Hibah Internal UM Bengkulu 2023	64 Kbps	1,525 KB	212,31	25,60%
	128 Kbps	2,140 KB	197,72	27,38%
	192 Kbps	3,016 KB	226,79	33,33%
	256 Kbps	3,892 KB	218,75	29,76%
	320 Kbps	4,768 KB	185,81	30,95%

Table 1 presents the comparative results of WER and processing time for each bitrate variation per audio file. To assess the content integrity of the transcriptions qualitatively, Table 2 provides a side-by-side comparison of transcripts based on one selected audio content, "Workshop Nasional Technopreneurship." Figure 2 illustrates the WER trends relative to bitrate across all tested audio files. This analysis aims to determine whether a consistent pattern exists between increasing bitrate and decreasing transcription error, and how such patterns inform technical decisions in automatic transcription systems.

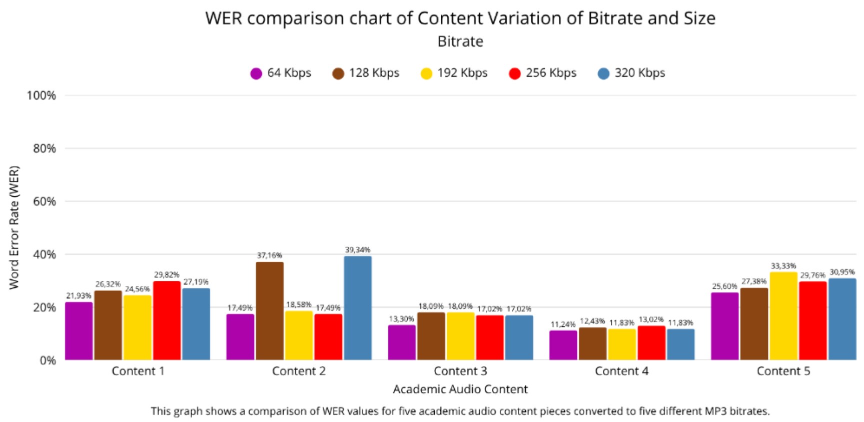

Figure 2. Comparison Chart of WER from Content Variation of Bitrate and Size

The graph in Figure 2 reveals that WER does not consistently decrease as bitrate increases. In some instances, such as at 128 kbps and 320 kbps, a spike in transcription errors is observed. This indicates that improved audio fidelity does not always correspond to better ASR performance. Notably, the highest WER values were observed at 320 kbps in audio files such as "Monitoring Hibah Internal" and "Workshop STEM," suggesting that larger file sizes do not necessarily guarantee higher transcription accuracy. The Whisper model's performance is also heavily influenced by voice clarity, background noise, and speaker pronunciation, as demonstrated by the WER variation across files with the same bitrate.

Generally, bitrate levels between 128 and 192 kbps yield relatively stable performance, with efficient file sizes and processing times. This range offers the best trade-off for automatic transcription applications, especially for academic documentation with varying audio quality. The testing involved five academic audio recordings, each converted into five bitrate levels, resulting in 25 test scenarios. Transcriptions produced by Whisper large-v3 were evaluated against manual references using the WER metric, while processing time and file size were also analyzed. The results showed that although a higher bitrate yields clearer audio, it does not necessarily lead to more accurate transcriptions. Therefore, technical decisions on optimal bitrate should strike a balance between computational efficiency, storage, and content characteristics.

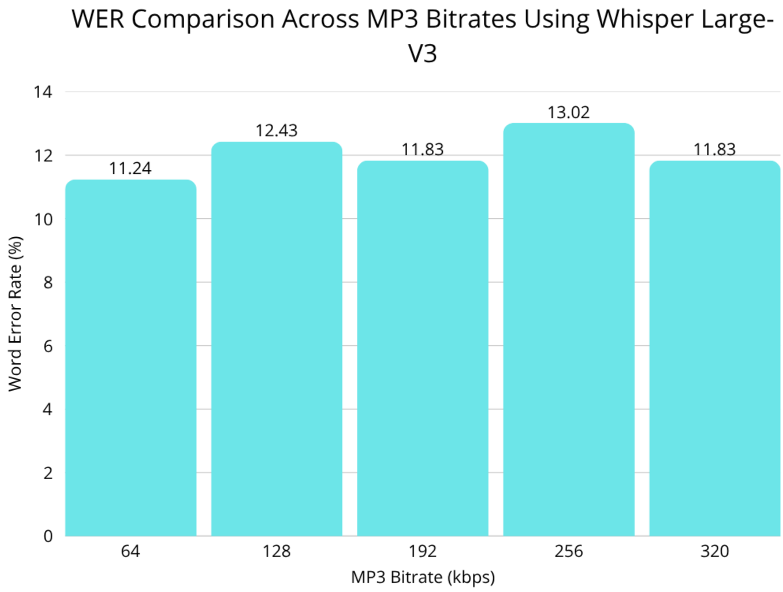

Figure 3. WER Comparison for Academic Audio Transcription Using Whisper large-v3 at Different MP3 Bitrates

Based on Tables 1 and Figure 3, bitrate influences not only numerical accuracy (WER) but also content completeness. For instance, at 64 kbps, the transcript closely resembles the reference with minor changes in speaker names. However, at 128 kbps and above, discrepancies in institutional names or TV station mentions become apparent. This demonstrates that even small phonetic changes at certain bitrates can significantly impact Whisper's word recognition. No absolute correlation between higher bitrate and lower WER was found. For example, in the audio file "Workshop STEM BASE CODING," a bitrate of 64 kbps yielded a lower WER than 128 kbps and 320 kbps. Similar patterns were observed in other files as well.

Regarding processing time, the relationship with file size was also nonlinear. At 128 kbps for "Workshop STEM BASE CODING," the processing time was longer than at 192 or 256 kbps. This suggests that transcription speed is also affected by content complexity, articulation, and background noise. Overall, bitrate levels between 128 and 192 kbps provide a reasonable balance between file size, processing time, and transcription accuracy. However, optimizing a transcription system should consider audio content characteristics, not merely bitrate.

IV. Conclusion

This study evaluated the impact of bitrate variation and MP3 file size on the accuracy of audio-to-text transcription using the Whisper large-v3 model. Based on tests conducted on five academic audio contents, each converted into five different bitrate levels (64 kbps to 320 kbps), it was found that increasing the bitrate does not always correlate directly with improved transcription accuracy. The evaluation results showed that bitrates of 128–192 kbps offered the most balanced Word Error Rate (WER) alongside efficient processing time and file size. For example, in the "Workshop Nasional Technopreneurship" file, the 64 kbps and 192 kbps bitrates recorded the lowest WERs of approximately 11.24% and 11.83%, respectively, indicating that low bitrate efficiency can still compete in terms of accuracy. Additionally, in some cases, the processing time at 128 kbps was higher than at higher bitrates, as observed in the "Workshop STEM BASE CODING" file.

This study concludes that selecting the optimal bitrate should take into account the complexity of the content, the audio quality, and the intended use of the transcript—not solely the bitrate value—because content complexity affects transcription performance more significantly than bitrate alone. The contribution of this study lies in providing a technical guideline applicable to academic documentation systems, particularly for educational institutions aiming to implement efficient and accurate automatic transcription using Whisper. Future studies may enhance the evaluation by including additional parameters, such as sampling frequency and audio file format, and by comparing Whisper with other ASR models under more diverse acoustic conditions.

REFERENCES

[1] S. Zakiya, K. Nisa, and D. Darmawan, “Pengaruh Media Pembelajaran Terhadap Hasil Belajar Siswa Setingkat Sekolah Dasar,” vol. 04, no. 01, p. 2025, 2025, doi: 10.9000/jpt.v4i1.2087.g519.

[2] M. I. Rizal, M. Yusron, M. El-Yunusi, and D. Darmawan, “Literasi Digital, Pemanfaatan Media Pembelajaran dan Kemandirian Belajar: Kontribusinya terhadap Prestasi Akademik di SMA Intensif Taruna Pembangunan Surabaya,” *Journal of Basic Educational Studies*, vol. 4, no. 2, p. 905, 2024.

[3] J. Mutaqin, I. Amirudin, R. Rizky, S. Fauziah, and far Amirudin, “PENDIDIKAN DI ERA REVOLUSI INDUSTRI 4.0: TANTANGAN DAN SOLUSI EDUCATION IN THE ERA OF THE FOURTH INDUSTRIAL REVOLUTION: CHALLENGES AND SOLUTIONS”, [Online]. Available: https://jicnusantara.com/index.php/jiic

[4] Z. Chik, A. H. Abdullah, A. Zawawi, M. Noor, and S. Ismail, “Journal of Economics, Finance and Management Studies Impact of Industrial Revolution 4.0 (IR4.0) Knowledge, Application Learning, University Policy, Commitment to Study and Motivation on Assimilate IR4.0 in Education”, doi: 10.47191/jefms/v7-i7-06.

[5] H. Maulida, E. Putry, V. Nuzulul ’adila, R. Sholeha, and D. Hilmi, “VIDEO BASED LEARNING SEBAGAI TREN MEDIA PEMBELAJARAN DI ERA 4.0.”

[6] D. Akademi, K. Pelamonia, B. Rahmat, and] Darmiati, “Pengembangan Media Pembelajaran dengan Video Based Learning,” 2021.

[7] S. Osmani and D. Tartari, “The Impact of Digital Technology on Learning and Teaching: A Case Study of Schools in Durrës, Albania,” *Journal of Educational and Social Research*, vol. 14, no. 6, pp. 193–209, Nov. 2024, doi: 10.36941/jesr-2024-0165.

[8] K. Amin, L. Elvitaria, and L. Trisnawati, “Jurnal Politeknik Caltex Riau Artificial Intelligence Automatic Speech Recognition (ASR) untuk pencarian potongan ayat Al-Qu’ran,” 2022. [Online]. Available: https://jurnal.pcr.ac.id/index.php/jkt/

[9] B. H. Juang and L. R. Rabiner, “Automatic Speech Recognition-A Brief History of the Technology Development,” 2004. [Online]. Available: http://www.recording-history.org/

[10] A. K. Katuri, S. Salugu, G. Tharuni, and C. S. Gouri, “Conversion of Acoustic Signal (Speech) Into Text By Digital Filter using Natural Language Processing,” *Int J Eng Adv Technol*, vol. 12, no. 1, pp. 14–18, Oct. 2022, doi: 10.35940/ijeat.A3802.1012122.

[11] H. Wijaya, “Teknologi Pengenalan Suara tentang Metode, Bahasa dan Tantangan: Systematic Literature Review,” *bit-Tech*, vol. 7, no. 2, pp. 533–544, Dec. 2024, doi: 10.32877/bt.v7i2.1888.

[12] V. Chemudupati *et al.*, “On the Transferability of Whisper-based Representations for ‘In-the-Wild’ Cross-Task Downstream Speech Applications,” May 2023, [Online]. Available: http://arxiv.org/abs/2305.14546

[13] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. Mcleavey, and I. Sutskever, “Robust Speech Recognition via Large-Scale Weak Supervision,” in *Proceedings of the 40th International Conference on Machine Learning*, A. Krause, E. Brunskill, K. Cho, B. Engelhardt, S. Sabato, and J. Scarlett, Eds., in Proceedings of Machine Learning Research, vol. 202. PMLR, May 2023, pp. 28492–28518. [Online]. Available: https://proceedings.mlr.press/v202/radford23a.html

[14] A. Pratiwi, "Perancangan Aplikasi Kompresi File Audio Dengan Menerapkan Algoritma Additive Code," *Journal Global Tecnology Computer*, vol. 1, no. 3, pp. 92–100, 2022.
[15] R. S. A. Pratama and A. Amrullah, "ANALYSIS OF WHISPER AUTOMATIC SPEECH RECOGNITION PERFORMANCE ON LOW RESOURCE LANGUAGE," *Jurnal Pilar Nusa Mandiri*, vol. 20, no. 1, pp. 1–8, Mar. 2024, doi: 10.33480/pilar.v20i1.4633.
[16] Prof. S. H. Chaflekar *et al.*, "YouTube Transcript Summarizer using Natural Language Processing," *International Journal of Advanced Research in Science, Communication and Technology*, no. 1, pp. 108–113, Apr. 2022, doi: 10.48175/ijarsct-3034.
[17] R. F. Khoiroh *et al.*, "Implementasi Speech Recognition Whisper pada Debat Calon Wakil Presiden Republik Indonesia," Jul. 2024.
[18] Y. Gong, S. Khurana, L. Karlinsky, and J. Glass, "Whisper-AT: Noise-Robust Automatic Speech Recognizers are Also Strong General Audio Event Taggers," Jul. 2023, doi: 10.21437/Interspeech.2023-2193.
[19] I. Sonata, "Automatic Speech Recognition in Indonesian Using the Transformer Model," in *2023 International Conference on Informatics, Multimedia, Cyber and Informations System (ICIMCIS)*, 2023, pp. 263–266. doi: 10.1109/ICIMCIS60089.2023.10349042.
[20] M. F. Fadlilah, A. R. Atmadja, and M. D. Firdaus, "Pemanfaatan Transformer untuk Peringkasan Teks: Studi Kasus pada Transkripsi Video Pembelajaran," *Building of Informatics, Technology and Science (BITS)*, vol. 6, no. 3, pp. 2111–2119, Dec. 2024, doi: 10.47065/bits.v6i3.6342.
[21] D. Macháček, R. Dabre, and O. Bojar, "Turning Whisper into Real-Time Transcription System," Jul. 2023, [Online]. Available: http://arxiv.org/abs/2307.14743
[22] H. Sajati and A. Pujiastuti, "THE AUDIO VIDEO OF WEB-BASED COMPRESSION WITH FFMPEG," 2018.
[23] S. Wang, C.-H. H. Yang, J. Wu, and C. Zhang, "Can Whisper perform speech-based in-context learning?," Sep. 2023, [Online]. Available: http://arxiv.org/abs/2309.07081
[24] D. K. Gete *et al.*, "Whispering in Amharic: Fine-tuning Whisper for Low-resource Language," Mar. 2025, [Online]. Available: http://arxiv.org/abs/2503.18485
[25] D. Ferdiansyah, C. Sri Kusuma Aditya, J. Raya Tlogomas No, K. Lowokwaru, K. Malang, and J. Timur, "Implementasi Automatic Speech Recognition Bacaan Al-Qur'an Menggunakan Metode Wav2Vec 2.0 dan OpenAI-Whisper." [Online]. Available: https://journal.trunojoyo.ac.id/triac
[26] J. Bellver-Soler *et al.*, "Multimodal Audio-Language Model for Speech Emotion Recognition."
[27] A. Loubser, P. De Villiers, and A. De Freitas, "End-to-end automated speech recognition using a character based small scale transformer architecture," *Expert Syst Appl*, vol. 252, Oct. 2024, doi: 10.1016/j.eswa.2024.124119.
[28] M. Ihsanudin Syaifullah, "Penerapan Teknologi Automatic Speech Recognition Menggunakan Model Wav2vec2.0 Sebagai Alat Bantu Untuk Mendeteksi Kesalahan Dalam Membaca Al-Qur'an Berbasis Mobile," *JSI : Jurnal Sistem Informasi (E-Journal*, vol. 16, no. 2, p. 2024.
[29] A. NOERCHOLIS, T. DWIANDINI, and F. S. MUKTI, "Optimasi Teknologi WAV2Vec 2.0 menggunakan Spectral Masking untuk meningkatkan Kualitas Transkripsi Teks Video bagi Tuna Rungu," *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, vol. 12, no. 4, p. 877, Dec. 2024, doi: 10.26760/elkomika.v12i4.877.
[30] S. Seyedi *et al.*, "Using HIPAA (Health Insurance Portability and Accountability Act)–Compliant Transcription Services for Virtual Psychiatric Interviews: Pilot Comparison Study," *JMIR Ment Health*, vol. 10, no. 1, 2023, doi: 10.2196/48517.