

## The Effectiveness of Audio-Visual Learning Media in Improving Arabic Speaking, Listening, and Vocabulary Skills: Evidence from Indonesian Students

Muh. Fihris Khalik<sup>1</sup> 

<sup>1</sup> Universitas Islam Makassar, Indonesia

### ABSTRACT

**Background.** The increasing demands of global communication and the integration of technology in education have highlighted the importance of mastering foreign languages, including Arabic. However, many university students continue to demonstrate low proficiency in Arabic, particularly in speaking and listening skills. This issue is often associated with limited variation in instructional

**Purpose.** This study aims to evaluate the effectiveness of audio-visual media as a tool for strengthening Arabic language competency, particularly in speaking, listening, and vocabulary mastery.

**Method.** This research employed a quantitative quasi-experimental design involving two groups: an experimental group that utilized audio-visual media and a control group that applied conventional learning methods.

**Results.** The findings revealed that students in the experimental group showed significant improvement in speaking, listening, and vocabulary skills compared to those in the control group. Statistical analysis confirmed a significant difference in the mean scores, indicating that the use of audio-visual media had a positive and measurable impact on enhancing Arabic language competency.

**Conclusion.** The study concludes that audio-visual media is an effective and engaging learning strategy for improving Arabic language skills. Its implementation not only enhances students' comprehension and participation but also supports their readiness to meet the challenges of modern, technology-driven learning environments.

### KEYWORDS

Arabic Language, Learning Media, Vocabulary Acquisition

**Citation:** Khalil, F. M. (2026). The Effectiveness of Audio-Visual Learning Media in Improving Arabic Speaking, Listening, and Vocabulary Skills: Evidence from Indonesian Students. *International Journal of Educational Narrative*, 4(1), 1–10.

<https://doi.org/10.70177/ijen.v4i1.3307>

### Correspondence:

Muh. Fihris Khalik,  
[fihriskhalik@gmail.com](mailto:fihriskhalik@gmail.com)

**Received:** September 2, 2025

**Accepted:** December 10, 2025

**Published:** February 24, 2026

### INTRODUCTION

Arabic is a global language with a strategic role in education, science, and intercultural communication. Mastery of Arabic is essential, particularly for students in religious and linguistic studies, as it encompasses not only reading and writing but also speaking and listening skills that are crucial for both academic and social interaction (Maurya, 2020). However, empirical findings from previous studies indicate that students' proficiency in these productive and receptive skills remains relatively low, especially in speaking and listening competencies. This gap is often associated with limited exposure to authentic language use, insufficient interaction in Arabic during classroom activities, and the dominance of conventional



teaching methods that prioritize memorization over communicative practice. In response to these challenges, this media is expected to facilitate students in learning Arabic in a more effective, interactive, and enjoyable way. The use of audio-visual media also allows for optimal understanding of vocabulary, sentence structure, and intonation. Therefore, audio-visual media has the potential to be used as an innovative and flexible learning strategy.

The integration of audio-visual media has been widely recognized as a promising approach to enhance language learning, as it combines visual and auditory elements to support better understanding of vocabulary, sentence structure, and intonation. Previous studies have demonstrated that such media can improve learning motivation, engagement, and retention. Nevertheless, most existing research tends to focus on isolated language skills such as listening, speaking, or vocabulary without providing a comprehensive evaluation of their combined development. Additionally, limited attention has been given to students' overall responses to the use of audio-visual media in Arabic learning. Therefore, this study seeks to address this gap by examining the effectiveness of audio-visual media in simultaneously improving speaking, listening, and vocabulary skills, while also providing empirical evidence on student engagement and learning outcomes (Hannuksela, 2021).

Despite this, students' Arabic language skills remain low, particularly in speaking and listening. This situation arises from several factors, including limited variety of learning media, low interaction in Arabic, and teaching methods that emphasize memorization and theory over practice. This situation emphasizes the importance of developing learning media that can encourage active student participation. Audio-visual media is considered an effective alternative to address this problem. Through the use of audio-visual media, students are more motivated to participate in the learning process, review vocabulary, and repeatedly practice speaking and listening skills. Thus, this media plays a role not only as a means of delivering material but also as an instrument for developing comprehensive language skills (Cilio, 2023).

This study was designed to assess the effectiveness of audio-visual media in strengthening Arabic language competency. The primary focus of the study was on the ability of audio-visual media to improve students' speaking, listening, and vocabulary skills. This study aimed to provide empirical evidence regarding the role of audio-visual media in supporting the strengthening of Arabic language competency (Martínez-Sánchez, 2021). This evaluation will determine whether audio-visual media is a viable alternative learning strategy. The research findings are expected to provide practical information for teachers and educational institutions to optimally integrate audio-visual media. This study also emphasizes the importance of media innovation in addressing the challenge of low Arabic language competency among students (Cao, 2023).

Listening and speaking skills are fundamental components of language acquisition, enabling learners to understand meaning and express ideas accurately. With the advancement of educational technology, audio-visual media has emerged as a promising instructional tool that integrates sound and visual elements to support learning (Uppada, 2023). Empirical studies suggest that such media can improve comprehension, retention, and learner engagement. However, despite these advantages, its application in Arabic language classrooms remains inconsistent and under-optimized, indicating a gap between technological potential and actual pedagogical practice.

The main objective of this study is to improve students' speaking, listening, and vocabulary skills in Arabic through the use of audio-visual media, compared to conventional learning approaches. This study is designed to provide a practical contribution to the development of more effective, engaging, and interactive Arabic language learning strategies. Audio-visual media is expected to increase students' learning motivation while facilitating faster and more comprehensive

understanding of the material. Furthermore, this study also serves as a guide for teachers in selecting and implementing appropriate media to strengthen Arabic language competency. By implementing audio-visual media in the classroom, students have the opportunity to learn repeatedly and actively, thereby increasing their participation in the learning process (Akman, 2024).

One of the advantages of this media is its ability to present learning materials in a realistic and contextual manner, making it easier for students to understand Arabic in everyday life. With this approach, students don't just passively receive information but actively practice speaking and listening skills. Furthermore, audio-visual media allows students to adjust their learning pace to suit their individual abilities, making the learning process more adaptive (Xu, 2022). The results of this study are expected to serve as a foundation for developing innovative, engaging learning methods that meet the needs of today's students. The use of this media can be an effective alternative for comprehensively improving Arabic language skills. Thus, audio-visual media not only plays a role in achieving academic competency but also encourages student motivation and active participation in the learning process (Ahirwal, 2020).

Some relevant research that supports this study includes: Xia, (2020) which examines interactive videos to improve Arabic listening skills, Revelli, (2022) which evaluates audio-visual animation in vocabulary mastery, and Zhao, (2023) which examined the effect of audio-visual media on students' speaking skills. All three studies showed positive results in the use of audio-visual media. The main difference in this study lies in the integration of three competency aspects speaking, listening, and vocabulary and a comprehensive assessment of students' responses to the media, thus providing a more comprehensive understanding of the effectiveness of using audio-visual media in Arabic language learning.

In other words, this research is relevant in addressing the problem of low Arabic language competency among students through the development of innovative learning media. Evaluating the effectiveness of audio-visual media is expected to provide a more in-depth understanding of its contribution to improving Arabic language skills (Hu, 2021). The results of this study are expected to serve as a foundation for designing more interactive, innovative, and relevant Arabic language learning strategies to meet the needs of today's students. Furthermore, this study emphasizes the importance of integrating audio-visual media into the curriculum to ensure optimal and sustainable achievement of Arabic language competencies.

## RESEARCH METHODOLOGY

This study employed a quantitative approach with a quasi-experimental design to assess the effectiveness of audio-visual media in improving Arabic language competency. The selection of a quasi-experimental design was considered the most appropriate for this study because it enables comparison between two groups without full randomization, which is often impractical in educational settings. In the university context, administrative and ethical constraints make random assignment difficult; therefore, this design allows for a realistic yet systematic evaluation of learning interventions. By comparing an experimental group receiving audio-visual media treatment and a control group undergoing conventional instruction, this approach provides empirical evidence of differences in learning outcomes while still maintaining internal validity to a reasonable extent (Ilyas, 2023).

The population of this study consisted of students enrolled in Arabic language courses at a specific level in a university, as this group represents learners with relatively comparable academic exposure. The sampling technique used was purposive sampling, with clearly defined criteria to

ensure transparency and consistency. The inclusion criteria included: (1) students actively enrolled in the targeted course, (2) students with a minimum attendance rate of 80% during the study period, and (3) students who agreed to participate fully in all stages of the experiment. The sample was then divided into two groups based on existing class structures: one designated as the experimental group and the other as the control group. The experimental group received instruction using audio-visual media, while the control group continued with conventional teaching methods such as lectures, discussions, and text-based exercises without additional media (Yang, 2023).

The sample size was determined to ensure the feasibility of statistical analysis and to enhance the representativeness of the findings. Prior to the treatment, both groups were assumed to have relatively similar levels of Arabic language competency, which was further confirmed through pre-test results. This procedure was intended to minimize bias and ensure that any observed differences in post-test results could be attributed primarily to the intervention rather than external variables. Moreover, this sampling strategy contributed to data consistency and strengthened the reliability of the findings (Hendriks, 2021).

The research instruments included an Arabic language competency test and a questionnaire. The competency test was administered in the form of pre-test and post-test to measure improvements in speaking, listening, and vocabulary mastery. The questionnaire was designed to capture students' perceptions, motivation, and level of engagement during the learning process using audio-visual media (Röchert, 2022).



Figure 1. Research Flow

Figure 1 the validity and reliability of the instruments were tested prior to data collection to ensure accuracy and consistency. All test items and questionnaire statements were developed based on the research objectives and relevant theoretical frameworks. The data collection procedure followed a structured sequence: pre-test administration, implementation of the treatment for the experimental group, and post-test along with questionnaire distribution. This systematic procedure ensured that the data obtained were comprehensive and suitable for in-depth analysis (Palani, 2022).

Despite its strengths, this study has several limitations. First, the use of a quasi-experimental design without random assignment may introduce potential selection bias. Second, the relatively limited sample size and its restriction to a single institution may affect the generalizability of the

findings. Third, external factors such as students' prior exposure to Arabic or individual learning motivation may not be fully controlled. Therefore, future studies are recommended to employ randomized designs, larger and more diverse samples, and additional control variables to enhance the robustness of the findings.

## RESULT AND DISCUSSION

The findings of this study are presented clearly; however, a deeper interpretation of the data reveals more nuanced insights into students' learning development. The results demonstrate a significant difference in Arabic language proficiency between students exposed to audio-visual media and those engaged in conventional instruction. Notable improvements were observed in speaking, listening, and vocabulary mastery, indicating that audio-visual tools do not merely enhance outcomes but also support the learning process in a more holistic manner. Compared with prior studies, these findings are consistent with existing research highlighting the effectiveness of multimedia-based instruction in language learning, while also reinforcing the idea that interactive media can foster greater engagement and comprehension. Therefore, this study contributes to the growing body of evidence that audio-visual media functions as an effective, adaptive, and student-centered learning strategy in contemporary educational contexts.

Furthermore, the implications of these findings extend beyond test performance, particularly in relation to student engagement and motivation. The experimental group not only achieved higher scores but also demonstrated increased participation, confidence, and enthusiasm during the learning process. These outcomes align with previous research emphasizing the role of technology in promoting active learning environments and supporting individualized learning experiences. The use of audio-visual media enables students to learn at their own pace, enhances their ability to process information, and encourages direct practice of language skills. In this regard, the study underscores the importance of integrating innovative instructional media into Arabic language education, as it not only improves academic achievement but also strengthens learner autonomy and interaction. Consequently, the findings hold significant implications for educators and policymakers in designing more effective, engaging, and relevant language learning frameworks in the modern era.

### Table 1.

The experimental group experienced a higher increase in scores compared to the control group on the Arabic language competency test.

Group	Pre-test Average	Post-test Average	Score Improvement
Experiment	65.2	84.5	19.3
Control	64.8	70.1	5.3
Group	Pre-test Average	Post-test Average	Score Improvement

The findings are presented clearly; however, a deeper interpretation of the data can further enrich the understanding of the results. The table demonstrates that the experimental group achieved a substantially greater improvement in Arabic language proficiency compared to the control group, particularly in speaking, listening, and vocabulary mastery. The mean score increase of 19.3 points in the experimental group, compared to only 5.3 points in the control group, indicates that the use of audio-visual media contributes significantly to learning outcomes. This suggests not only improved comprehension of the material but also a transformation in the learning process, where students become more active, confident, and capable of retaining vocabulary. In contrast, the limited progress observed in the control group highlights the constraints of conventional methods that lack

interactive and practical engagement. Therefore, these findings underscore the pedagogical value of audio-visual media as an effective tool to enhance both language skills and student engagement.

While the results are well described, comparing them with previous studies can provide a broader perspective on their significance. The observed increase in student participation and performance in the experimental group aligns with prior research indicating that audio-visual media enhances learner engagement and facilitates language acquisition. Students exposed to such media demonstrated higher levels of interaction, confidence, and vocabulary comprehension, which is consistent with findings in similar educational contexts. Conversely, the passive behavior observed in the control group reflects patterns reported in studies relying on traditional teaching approaches. These similarities reinforce the reliability of the current findings, while any differences may be attributed to contextual factors such as learning environment, instructional design, or student characteristics. Thus, situating these results within the existing body of research highlights both their relevance and contribution to the field.

Although the study provides relevant results, the implications of these findings for the field require further elaboration. The significant difference revealed by the t-test analysis confirms that the improvement in the experimental group is statistically meaningful and not due to chance, thereby offering strong empirical support for the effectiveness of audio-visual media. This has important implications for Arabic language education, suggesting that integrating such media can enhance not only linguistic competence but also student motivation and participation. Furthermore, the findings indicate the need for educators to adopt more interactive and technology-based approaches in order to meet the demands of modern learning environments. By doing so, educational practices can move beyond traditional methods toward more dynamic and student-centered models.

Overall, the study demonstrates that audio-visual media has a significant positive impact on students' Arabic language skills, as evidenced by both quantitative and qualitative data. Beyond improving academic performance, the use of such media fosters greater engagement, motivation, and deeper understanding of the material. These results imply that audio-visual media can serve as a comprehensive instructional strategy that supports active learning, repeated practice, and contextualized language use. Consequently, this approach not only enhances students' competencies but also contributes to the development of a more interactive and enjoyable learning environment. Therefore, the integration of audio-visual media represents an adaptive and effective solution for advancing Arabic language education in contemporary settings.

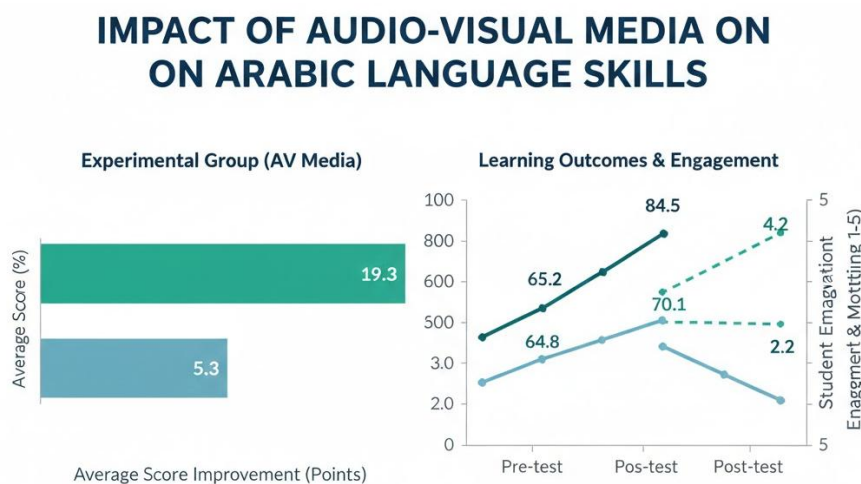


Figure 2. Main Research

Overall, the results of this study indicate that audio-visual media can be utilized as an effective and innovative learning strategy to strengthen Arabic language competency. The application of this media not only contributes to improving students' speaking, listening, and vocabulary skills but also encourages their active participation during the learning process. Furthermore, the use of audio-visual media has been shown to increase students' motivation to learn more intensively and actively engage in classroom activities. These findings serve as an important reference for instructors in systematically integrating audio-visual media into the Arabic language curriculum. With appropriate implementation, these media enable learning to be more creative, interactive, and responsive to students' needs. Audio-visual media also provides opportunities for students to study independently, review material at their own pace, and understand Arabic in real-life contexts. Thus, audio-visual media-based learning strategies not only support the achievement of academic competency but also improve the quality of the overall learning experience and are relevant to the demands of contemporary education.

## CONCLUSION

The research results indicate that the use of audio-visual media has proven effective in improving students' Arabic language competency. Improvements were seen in speaking, listening, and vocabulary mastery, significantly improving compared to conventional learning methods. Audio-visual media presents material in a more realistic and contextual manner, making it easier for students to understand Arabic. Furthermore, the use of this media creates a more engaging and varied learning experience. Thus, the integration of audio-visual media positively contributes to improving the quality of Arabic language learning in higher education.

In addition to improving language competency, the use of audio-visual media also increases student motivation and active participation in the learning process. Students demonstrate higher levels of enthusiasm and are directly involved in Arabic language learning activities. Learning using audio-visual media encourages students to not only understand theory but also apply language skills in more realistic contexts. This confirms that audio-visual media can serve as an interactive, engaging, and relevant learning strategy for today's learners.

This study confirms that the structured and systematic use of audio-visual media in the Arabic language curriculum plays a crucial role in optimizing student competency achievement. The application of audio-visual media not only contributes to improving the quality of learning but also prepares students to face the demands and dynamics of the digital era. Adaptability skills to technology are a crucial element in the contemporary learning process. Therefore, audio-visual media should be used as a strategic component in Arabic language learning to continuously improve the quality of education and student readiness.

## DECLARATION OF AI AND AI ASSISTED TECHNOLOGIES IN THE WRITING PROCESS

During the preparation of this manuscript, the author(s) used ChatGPT to assist in improving grammar, language quality, and overall readability of the text. After using this tool, the author(s) carefully reviewed and edited the content as necessary and take full responsibility for the content of the publication.

## AUTHORS' CONTRIBUTION

Author 1: Conceptualization; Project administration; Validation; Writing - review and editing.

## DECLARATION OF COMPETING INTEREST

The authors declare that they have no known competing financial interests of personal relationships that could have appeared to influence the work reported in this paper.

## REFERENCES

- Ahirwal, M.K. (2020). Audio-visual stimulation based emotion classification by correlated EEG channels. *Health and Technology*, 10(1), 7–23. <https://doi.org/10.1007/s12553-019-00394-5>
- Akman, A. (2024). Audio Explainable Artificial Intelligence: A Review. *Intelligent Computing*, 3(Query date: 2026-01-24 23:34:59). <https://doi.org/10.34133/icomputing.0074>
- Arayata, PAE (2022). Chyilax: A 3D Video Game as a Marketing Tool for Mental Breakdown Awareness Campaign. 2022 IEEE 14th International Conference on Humanoid Nanotechnology Information Technology Communication and Control Environment and Management Hnicem 2022, Query date: 2026-01-24 23:34:59. <https://doi.org/10.1109/HNICEM57413.2022.10109535>
- Cao, Y. (2023). Attention-Guided Neural Networks for Full-Reference and No-Reference Audio-Visual Quality Assessment. *IEEE Transactions on Image Processing*, 32(Query date: 2026-01-24 23:34:59), 1882–1896. <https://doi.org/10.1109/TIP.2023.3251695>
- Cilio, S. (2023). Analysis of quality information provided by “Dr. YouTube™” on Phimosis. *International Journal of Impotence Research*, 35(4), 398–403. <https://doi.org/10.1038/s41443-022-00557-5>
- Cortivo, D.D. (2021). CNN-based multi-modal camera model identification on video sequences. *Journal of Imaging*, 7(8). <https://doi.org/10.3390/jimaging7080135>
- Cover, R. (2022). Deepfake culture: The emergence of audio-video deception as an object of social anxiety and regulation. *Continuum*, 36(4), 609–621. <https://doi.org/10.1080/10304312.2022.2084039>
- Gupta, R. (2023). Class Prototypes based Contrastive Learning for Classifying Multi-Label and Fine-Grained Educational Videos. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2023*(Query date: 2026-01-24 23:34:59), 19923–19933. <https://doi.org/10.1109/CVPR52729.2023.01908>
- Hannuksela, MM (2021). An Overview of Omnidirectional Media Format (OMAF). *Proceedings of the IEEE*, 109(9), 1590–1606. <https://doi.org/10.1109/JPROC.2021.3063544>
- Hendriks, S. (2021). Azalea: Co-experience in remote dialogue through diminished reality and somaesthetic interaction design. *Conference on Human Factors in Computing Systems Proceedings*, Query date: 2026-01-24 23:34:59. <https://doi.org/10.1145/3411764.3445052>
- Hu, R. (2021). AVMSN: An Audio-Visual Two Stream Crowd Counting Framework under Low-Quality Conditions. *IEEE Access*, 9(Query date: 2026-01-24 23:34:59), 80500–80510. <https://doi.org/10.1109/ACCESS.2021.3074797>
- Ilyas, H. (2023). AVFakeNet: A unified end-to-end Dense Swin Transformer deep learning model for audio–visual deepfakes detection. *Applied Soft Computing*, 136(Query date: 2026-01-24 23:34:59). <https://doi.org/10.1016/j.asoc.2023.110124>
- Madugalla, A. (2020). Creating Accessible Online Floor Plans for Visually Impaired Readers. *ACM Transactions on Accessible Computing*, 13(4). <https://doi.org/10.1145/3410446>
- Martínez-Sánchez, M.E. (2021). Analysis of the social media strategy of audio-visual OTTs in Spain: The case study of Netflix, HBO and Amazon Prime during the implementation of Disney+. *Technological Forecasting and Social Change*, 173(Query date: 2026-01-24 23:34:59). <https://doi.org/10.1016/j.techfore.2021.121178>
- Matta, G. (2022). Communicating Water, Sanitation, and Hygiene under Sustainable Development Goals 3, 4, and 6 as the Panacea for Epidemics and Pandemics Referencing the Succession of COVID-19 Surges. *ACS Es and T Water*, 2(5), 667–689. <https://doi.org/10.1021/ACSESTWATER.1C00366>
- Maurya, R. (2020). An Extended Visual Cryptography Technique for Medical Image Security. 2nd International Conference on Innovative Mechanisms for Industry Applications Icimia 2020

- Conference Proceedings, Query date: 2026-01-24 23:34:59, 415–421. <https://doi.org/10.1109/ICIMIA48430.2020.9074910>
- Middya, AI (2022). Deep learning based multimodal emotion recognition using model-level fusion of audio–visual modalities. *Knowledge Based Systems*, 244(Query date: 2026-01-24 23:34:59). <https://doi.org/10.1016/j.knosys.2022.108580>
- Palani, B. (2022). CB-Fake: A multimodal deep learning framework for automatic fake news detection using capsule neural network and BERT. *Multimedia Tools and Applications*, 81(4), 5587–5620. <https://doi.org/10.1007/s11042-021-11782-3>
- Patel, Y. (2023). Deepfake Generation and Detection: Case Study and Challenges. *IEEE Access*, 11(Query date: 2026-01-24 23:34:59), 143296–143323. <https://doi.org/10.1109/ACCESS.2023.3342107>
- Prindani, A., Syauqillah, M., & Anam, K. (2025). Humanist Intelligence in the Context of Modern Security: Between Operational Effectiveness and Respect for Human Values. *Journal of Cultural Analysis and Social Change*, 10(3), 1396–1412. Scopus. <https://doi.org/10.64753/jcasc.v10i3.2612>
- Ros, G., Fraile Rey, A. F., Calonge, A., & López-Carrillo, M. D. (2022). The Design of a Teaching-Learning Sequence on Simple Machines in Elementary Education and its Benefit on Creativity and Self-Regulation. *Eurasia Journal of Mathematics, Science and Technology Education*, 18(1), 1–22. Scopus. <https://doi.org/10.29333/EJMSTE/11487>
- Revelli, VP (2022). Automate extraction of braille text to speech from an image. *Advances in Software Engineering*, 172(Query date: 2026-01-24 23:34:59). <https://doi.org/10.1016/j.advengsoft.2022.103180>
- Röchert, D. (2022). Caught in a networked collusion? Homogeneity in conspiracy-related discussion networks on YouTube. *Information Systems*, 103(Query date: 2026-01-24 23:34:59). <https://doi.org/10.1016/j.is.2021.101866>
- Shih-En Leu, G. (2020). My Life the Way I See It: Reconstructing Minoritized Youth With Disabilities as Critical Thinkers. *Journal of Adolescent and Adult Literacy*, 64(2), 181–190. Scopus. <https://doi.org/10.1002/jaal.1083>
- Shittu, E., Dashiell-Shoffner, J., & Kim, H. N. (2021). Examining Black Diaspora Participation in Engineering using Narrative Inquiry. *CoNECD*. Scopus. <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85124516789&partnerID=40&md5=e65ef9fb4d969ebb09bed2dca615ac02>
- Suwa, M. (2019). Methodological and Philosophical Issues in Studies on Embodied Knowledge. *New Generation Computing*, 37(2), 167–184. Scopus. <https://doi.org/10.1007/s00354-019-00055-1>
- Teeple, S. K., & Benolken, A. (2023). Exploring the relationship between social-emotional competencies and student performance in online learning environments. *E-Learning and Digital Media*, 20(5), 460–472. Scopus. <https://doi.org/10.1177/20427530221117328>
- Tran, T., Ternov, N. K., Weber, J., Barata, C., Berry, E. G., Doan, H. Q., Marghoob, A. A., Seiverling, E. V., Sinclair, S., Stein, J. A., Stoos, E. R., Tolsgaard, M. G., Wolfensperger, M., Braun, R. P., & Nelson, K. C. (2022). Instructional Strategies to Enhance Dermoscopic Image Interpretation Education: A Review of the Literature. *Dermatology Practical and Conceptual*, 12(4). Scopus. <https://doi.org/10.5826/dpc.1204a189>
- Uppada, S. K. (2023). An image and text-based multimodal model for detecting fake news in OSN's. *Journal of Intelligent Information Systems*, 61(2), 367–393. <https://doi.org/10.1007/s10844-022-00764-y>
- Xia, L. (2020). Audiovisual speech recognition: A review and forecast. *International Journal of Advanced Robotic Systems*, 17(6). <https://doi.org/10.1177/1729881420976082>
- Xu, H. (2022). Audio-Visual Autoencoding for Privacy-Preserving Video Streaming. *IEEE Internet of Things Journal*, 9(3), 1749–1761. <https://doi.org/10.1109/JIOT.2021.3089080>

- Xuan, H. (2020). Cross-modal attention network for temporal inconsistent audio-visual event localization. Aaai 2020 34th Aaai Conference on Artificial Intelligence, Query date: 2026-01-24 23:34:59, 279–286.
- Yang, W. (2023). AVoid-DF: Audio-Visual Joint Learning for Detecting Deepfakes. IEEE Transactions on Information Forensics and Security, 18(Query date: 2026-01-24 23:34:59), 2015–2029. <https://doi.org/10.1109/TIFS.2023.3262148>
- Zhang, X. (2020). DAVD-net: Deep audio-aided video decompression of talking heads. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Query date: 2026-01-24 23:34:59, 12332–12341. <https://doi.org/10.1109/CVPR42600.2020.01235>
- Zhao, B. (2023). AudioVisual Video Summarization. IEEE Transactions on Neural Networks and Learning Systems, 34(8), 5181–5188. <https://doi.org/10.1109/TNNLS.2021.3119969>

---

**Copyright Holder :**

© Muh. Fihris Khalik (2026).

**First Publication Right :**

© International Journal of Educational Narratives

**This article is under:**