Speech Recognition

Curzon et al. (2021) focused on the privacy problems of AI as a general paradigm of information processing. The general claim is that by nature, AI as a data sharing and processing field is exposed to many privacy risks. They classified five parts that include: computer vision, speech recognition, NLP, knowledge representation, automated reasoning and ML. For the five parts, the authors present their privacy risks and suggest different methods of mitigating them to reduce the inherent vulnerabilities. Li and Zhang (2017) dealt with the security and privacy problems that AI creates, distinctively for the data which is handled by its applications, that involve speech text input and personalized network shopping, to various intelligent answering systems. In this perspective the paper also deals with ethical problems that arise with the usage of AI and its applications, that inherently have vast data to handle. The authors divide the problems into the fields of security, privacy and ethics, and handle each field with explanations about its possible problematic aspects and vulnerabilities. Liu et al. (2021) provided an extensive review of privacy challenges and solutions in the context of ML, focusing on models that were able to extract accent information from trained speech recognition systems. They discuss various ways ML can be a threat to privacy and how ML aids in privacy protection. The study categorizes the interaction between ML and privacy into three types: ML as a target for privacy protection, ML as a tool for enhancing privacy, and ML as a tool for privacy attacks.

The paper also identifies gaps in current research and suggests directions for future studies, highlighting the need for more robust privacy-preserving techniques in ML applications.

Gandeeban et al. (2025) present a novel architecture, SER-EQCNN-ESC, designed to significantly enhance the accuracy and robustness of speech emotion recognition (SER) systems. The proposed framework integrates several advanced components: Bellman Filtering (BF) for noise reduction and normalization, Holistic Dynamic Frequency Transformer (HDFT) for extracting acoustically rich emotional features, and Single Candidate Optimizer (SCO) for optimal feature selection. The classification is performed using an Equivariant Quantum Convolutional Neural Network (EQCNN), whose parameters are further optimized using the Educational Competition Optimization (ECO) algorithm.

References

Curzon, J., Kosa, T. A., Akalu, R., & El-Khatib, K. (2021). Privacy and artificial intelligence. *IEEE Transactions on Artificial Intelligence, 2*(2), 96--108. doi:https://doi.org/10.1109/TAI.2021.3088084

Gandeeban, B., Ranjith, S., Jagan, G., & Chenthil, T. (2025). Advanced Speech Emotion Recognition Utilizing optimized Equivariant quantum convolutional neural network for Accurate Emotional State Classification. *Knowledge-Based Systems*, 113414. doi:https://doi.org/10.1016/j.knosys.2025.113414

Li, X., & Zhang, T. (2017). An exploration on artificial intelligence application: From security, privacy and ethic perspective. *2017 IEEE 2nd international conference on cloud computing and big data analysis (ICCCBDA)* (pp. 416--420). IEEE. doi:https://doi.org/10.1109/ICCCBDA.2017.7951949

Liu, B., Ding, M., Shaham, S., Rahayu, W., Farokhi, F., & Lin, Z. (2021). When machine learning meets privacy: A survey and outlook. *ACM Computing Surveys (CSUR), 54*(2), 1--36. doi:https://doi.org/10.1145/3436755