|  |  |  |  |
| --- | --- | --- | --- |
| Title | Text to Speech Synthesis: A Systematic Review, Deep Learning Based Architecture and Future Research Direction | | |
| Author(s) Name | Fahima Khanam, Farha Akhter Munmun, Nadia Afrin Ritu, Aloke Kumar Saha, Muhammad Firoz | | |
| Contact Email(s) | firoz.mridha@aiub.edu | | |
| Published Journal Name | Journal of Advances in Information Technology Vol | | |
| Type of Publication | Review | | |
| Volume | 13 | Issue | 5 |
| Publisher | Advances in Information Technology | | |
| Publication Date | 2022/10 | | |
| ISSN | 17982340 | | |
| DOI | 10.12720/jait.13.5.398-412 | | |
| URL | http://jait.us/uploadfile/2022/0831/20220831054604906.pdf | | |
| Other Related Info. |  | | |
|  | | | |

|  |  |
| --- | --- |
| Abstract |  |
| Text to Speech (TTS) synthesis is a process of translating natural language text into speech. Pieces of recorded speech generate synthesized speech and a database is maintained for storing this synthesized speech. A speech  synthesizer’s output is determined through its resemblance to the person utter and its capacity to be implied. In recent years between the two main subsections: machine learning and deep learning of Artificial Intelligence (AI), deep  learning has achieved huge success in the domain of text to speech synthesis. In this literature, a taxonomy is introduced which represents some of the deep learning-based architectures and models popularly used in speech synthesis.  Different datasets that are used in TTS have also been discussed. Further, for evaluating the quality of the synthesized speech some of the widely used evaluation matrices are described. Finally, the paper concludes with the  challenges and future directions of the text-to-speech synthesis system. | |