

Recently, the video gaming industry has been growing very fast. The primary objective of a video game is to increase the number of users so that it can be made more engaging by adding sound effects, music backgrounds, and voice actors. The process of hiring voice actors is very demanding and costs time, money, and effort. Our proposed project was to build a commentator voice synthesizer, especially for Fifa. We generated speech with background noise by using advanced deep-learning models that take the text as input. Our model was trained on 15 minutes of speech. We crawled a couple of videos from youtube for a football game commentated by a famous Tunisian commentator. We applied different methods for data preprocessing, including voice analysis detection, segmentation, and vowelization. Our model achieved very high-quality speech from a small amount of data.