

REPORT

source ("<https://arxiv.org/abs/1804.03209>")

The "Speech Commands" dataset, as described in the paper "<https://arxiv.org/abs/1804.03209>," is a collection of short spoken words designed for keyword spotting tasks. It differs from other speech recognition datasets in several ways. First, it focuses on individual words rather than entire sentences or phrases, making it more suitable for specific keyword identification applications. Second, the dataset is relatively small, comprising only 35 words, which can be beneficial for training smaller models and reducing computational overhead.

The data collection process involved recording participants speaking each word into a microphone. To ensure diversity, the dataset includes recordings from a variety of speakers with different accents, genders, and ages. Additionally, the recordings were made in various noise environments to simulate real-world conditions and improve the model's robustness to background noise. The authors believe that this dataset can be a valuable resource for researchers and developers working on keyword spotting and other speech-related tasks.

Date : - 11-09-2024