

Speech Commands: A Dataset for Limited - Vocabulary Speech Recognition

Paper : <https://arxiv.org/abs/1804.03209>

Summary : The paper presents the "Speech Commands" dataset, a specialized audio collection for training and evaluating keyword spotting models. It contains over 105,000 utterances of 35 words from 2,618 speakers, designed for tasks like on-device recognition of keywords (e.g., "Yes", "No", "Stop") in low-resource environments. The dataset addresses challenges like minimizing false positives from background noise and ensuring energy-efficient models for devices with limited computational power. It is intended to standardize testing protocols by providing a benchmark for reproducible and comparable accuracy metrics, with a focus on tasks distinct from general speech recognition, such as keyword detection.

