



Key words: general audio data, multi-modal data, self-learning, transfer learning

Key words: task-specific data, supervised learning, label scarcity, **active learning**