# SoundID

## Powerful AI that goes beyond the human voice to recognize environmental sounds

TrulySecure™ Sound-Identification (SoundID) technology gives developers the opportunity to design rich and interactive products that go beyond sensing of the human voice. SoundID leverages a two stage detection approach. A first-stage processing is optimized for speed and efficiency and robustness detecting candidate events while the optional second-stage revalidation model is optimized for accuracy. When combined, Sensory's solution provides the perfect balance between on-device computation and overall recognition performance.

SoundID includes four core capabilities for sound analysis. Sound Event Detection, the ability to detect fixed, discrete events such as dog barks, coughs, and baby cries. The SDK includes 16 sample sounds with various combinations (home, health, safety, etc.). Sound Scene Tracking, enables listening continuously to audio and reporting relative match to broad sound categories or "scenes". Custom scenes are available by request. Enrolled Sound Detection, Allows end-users to add their own new sounds via a fast & easy enrollment process. Enrolled sounds may be discrete (e.g., a door bell) or continuous (e.g., running water). Alarm Detection, Detects residential temporal-3 (Smoke), temporal-4 (Carbon Monoxide) alarms, continuous alarms and low-battery "chirp" with low computational resources. Sound ID is available as a component of TrulySecure™ Speaker Verification (TSSV), **contact sales** for demos and details,

### Listening on the Edge

Listens & analyzes sounds on device without the security risk of sending audio recordings to the cloud.

### Industrial Applications

Sound ID may be developed for industrial sound detection of specific product issues or malfunctions.

### Low Power, High Accuracy

Multi-stage shallow and deep learning approach results in low power consumption and high accuracy.

### Extensive Sound Library

A library of 16 sounds ($CO_2$ & smoke alarms, glass break, baby cry, snoring, door knock/bell, yelling, coughing, horns, sirens) organized by various categories (home, health, safety, etc).

### User-Enrolled Sounds

Sound ID enables the end user to enroll unique sounds with only 8 seconds of the target sound.

### Multiple Platform Flexiblity

Runs embedded on various OS systems (Linux, Android, Windows, etc.) or can be run on low-powered Application Processors like ARM Cortex A11.

### Noise Robust

Using Sensory's propietary approach, Sound ID works in noisy environments.

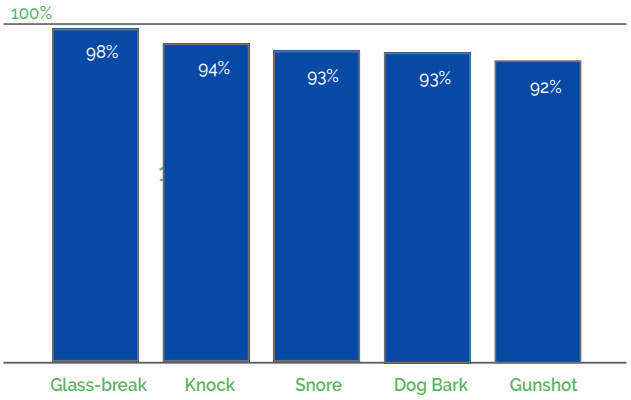
### Dynamic Sound Scenes

Sound ID includes present "sound-scenes" with preset levels for continuous events.

### Accessibility

SoundID may be used to improve accessibility for the hearing impaired, and remote health & safety monitoring.

Mobile | Automotive | IoT | Medical Equipment | Biometric Security




## Performance for Various Sounds

The highlighted detection rates were taken at a threshold of 1 False Alarm in 24 hours on a balanced mix of noise data. Dectection rates here are for single instances, but rates improve with multiple instances. For example, dectection rates for Dog barks improve from 93% in a single instance to 99.9% with three instances. The general effects of multiple instances hold true across SoundID categories, enabling devices with SoundID parametric controls of the number of instances to cause a notification.

## Technical Data & Specifications

Sensory's sound identification technology provides a first-stage detector that can be combined with an optional second-stage revalidation model. The SDK provides 16 base sounds that have been optionally combined into sound packs (all, home, health, safety sounds). Approximate MIPS, Memory and storage requirements are provided as a function of the number of modeled sounds (N) as well as first vs. first and second stage processing.

Enrolled sound-triggers (User-defined sounds) are created through an enrollment process. A minimum of 4 repetions of the event or 8 seconds of continuous audio is required to register a new sound. Typically this requires a fixed storage of 132 kB for the background model and approximately 34 kB of additional on-device storage per enrolled sound. Note that memory and MIPS requirements shown below are provided for a single enrolled sound and there are no SDK limitations as to the number of enrolled sounds that can be processed with the **TSSV SDK**.

| Base Sounds | |
|---|---|
| 1st -stage model size | ~80 kB per sound |
| 2nd-stage model size | 4.5 MB |
| Recognition Memory 1st-stage | .4 + 0.74(N) MB |
| Recognition Memory 2nd-stage | 9.3 + 0.36(N) MB |
| Recognition MIPS 1st-stage | 15.3 + 13.4(N) |
| Recognition MIPS 2nd-stage | 14.3 + 19.8(N) |
| Recognition Spin-Up Time | < 15 msec |

| Enrolled Sounds | |
|---|---|
| Enrollment Requirement | 4+ repetitions or ~ 8 seconds of sound |
| Recognition Requirement | 0.4 to 3.0 seconds |
| Background Model Size | 132 kB |
| Enrollment Model Size | 34kB per sound |
| Enrollment Memory | 2.7 MB |
| Enrollment MIPS | 127 |
| Recognition Memory | 1.4 MB |
| Recognition MIPS | 25 |
| Recognition Spin-Up Time | <15 msec |

Sensory, Inc., 4701 Patrick Henry Drive, BDLG 7, Santa Clara, CA 95054, United States, 408.625.3300, **sensory.com**