

# Models for AudioSet: A Large Scale Dataset of Audio Events

This repository provides models and supporting code associated with [AudioSet](#), a dataset of over 2 million human-labeled 10-second YouTube video soundtracks, with labels taken from an ontology of more than 600 audio event classes.

AudioSet was [released](#) in March 2017 by Google's Sound Understanding team to provide a common large-scale evaluation task for audio event detection as well as a starting point for a comprehensive vocabulary of sound events.

For more details about AudioSet and the various models we have trained, please visit the [AudioSet website](#) and read our papers:

- Gemmeke, J. et. al., [AudioSet: An ontology and human-labelled dataset for audio events](#), ICASSP 2017
- Hershey, S. et. al., [CNN Architectures for Large-Scale Audio Classification](#), ICASSP 2017

If you use any of our pre-trained models in your published research, we ask that you cite [CNN Architectures for Large-Scale Audio Classification](#). If you use the AudioSet dataset or the released embeddings of AudioSet segments, please cite [AudioSet: An ontology and human-labelled dataset for audio events](#).

## Contact

For general questions about AudioSet and these models, please use the [audioset-users@googlegroups.com](mailto:audioset-users@googlegroups.com) mailing list.

For technical problems with the released model and code, please open an issue on the [tensorflow/models issue tracker](#) and **assign to @plakal and @dpwe**. Please note that because the issue tracker is shared across all models released by Google, we won't be notified about an issue unless you explicitly @-mention us (@plakal and @dpwe) or assign the issue to us.

## Credits

Original authors and reviewers of the code in this package include (in alphabetical order):

- DAn Ellis
- Shawn Hershey
- Aren Jansen
- Manoj Plakal