

# Inclusive Design for Deaf and Hard of Hearing Using AI in a Scalable Way

Swetha Machanavajhala  
Software Engineer 2

Xiaoyong Zhu  
Program Manager 2

---



## True life threatening incident

Carbon monoxide detector was beeping  
for a WEEK at Swetha's house.

Since she is hard of hearing, she was  
unaware until a neighbor informed her.

Swetha Machanavajhala





World population  
7,444,594,497

People with disabilities  
1,000,000,000

---

Deaf or hard of hearing  
360,000,000

**DISABILITY**

**≠**

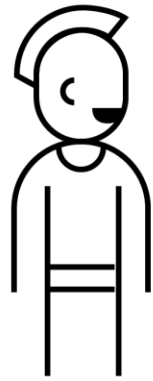
**PERSONAL HEALTH  
CONDITION**

**DISABILITY**

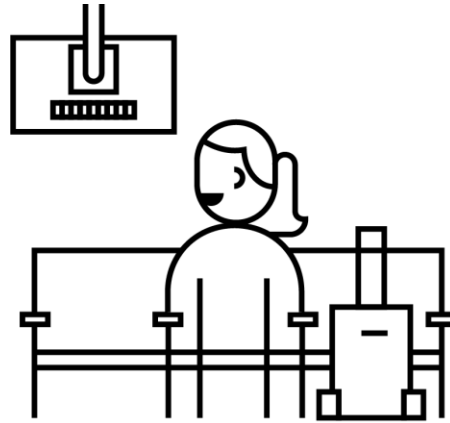
**=**

**MISMATCHED  
HUMAN  
INTERACTIONS**

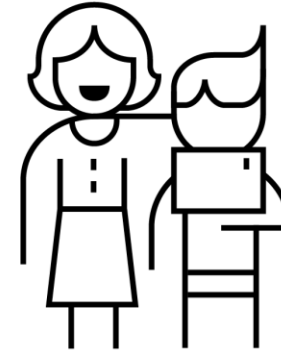
# Inclusive Design



Hard of hearing



Reading in an airport



Teaching a child to read



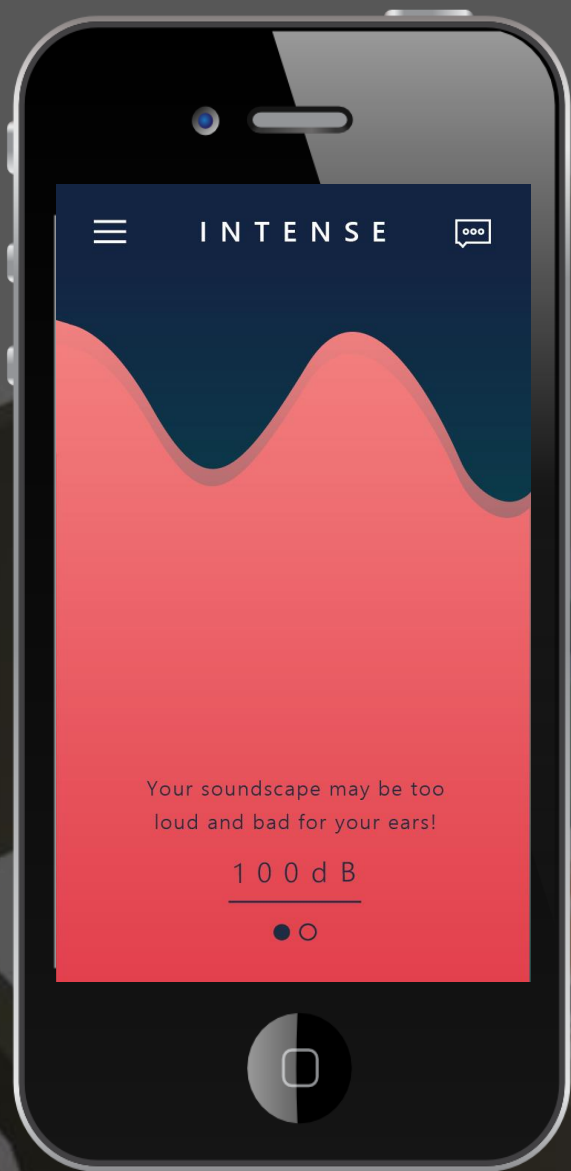
More people benefit from closed captioning

WHY MUST  
(( SOUND ))  
— ONLY BE —  
HEARD?

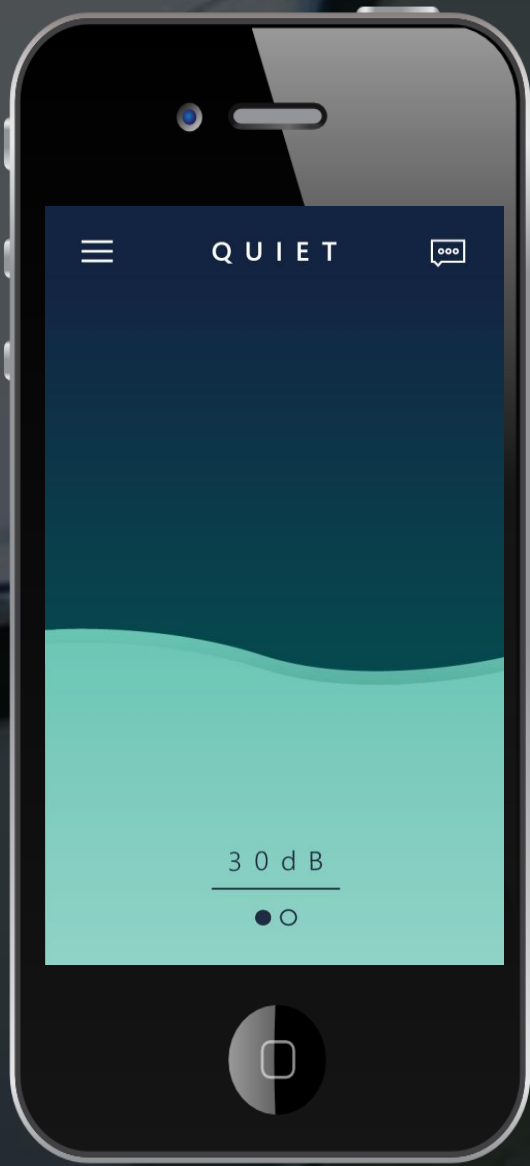


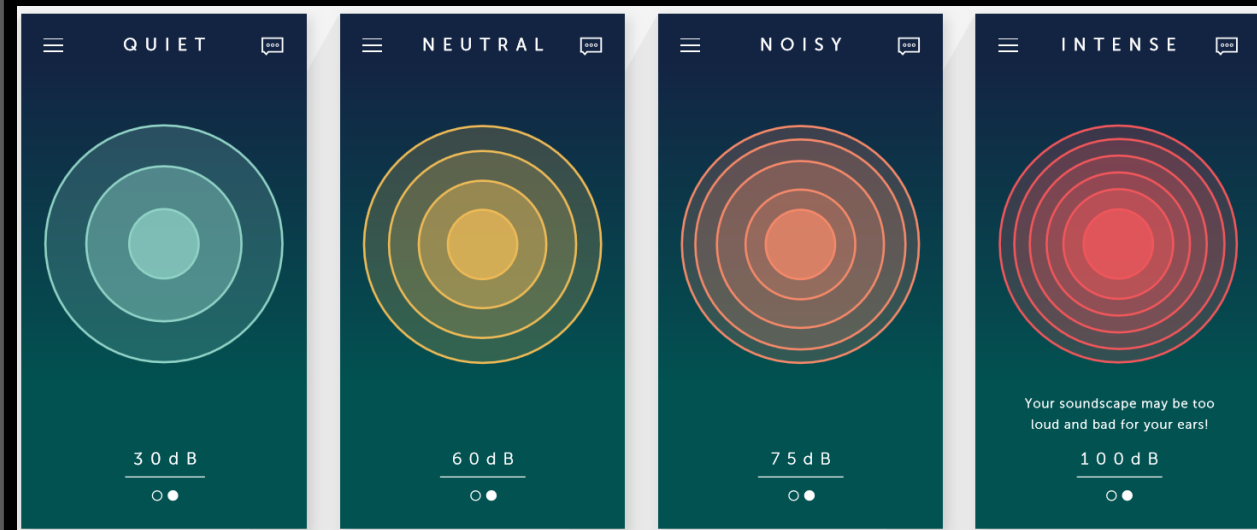
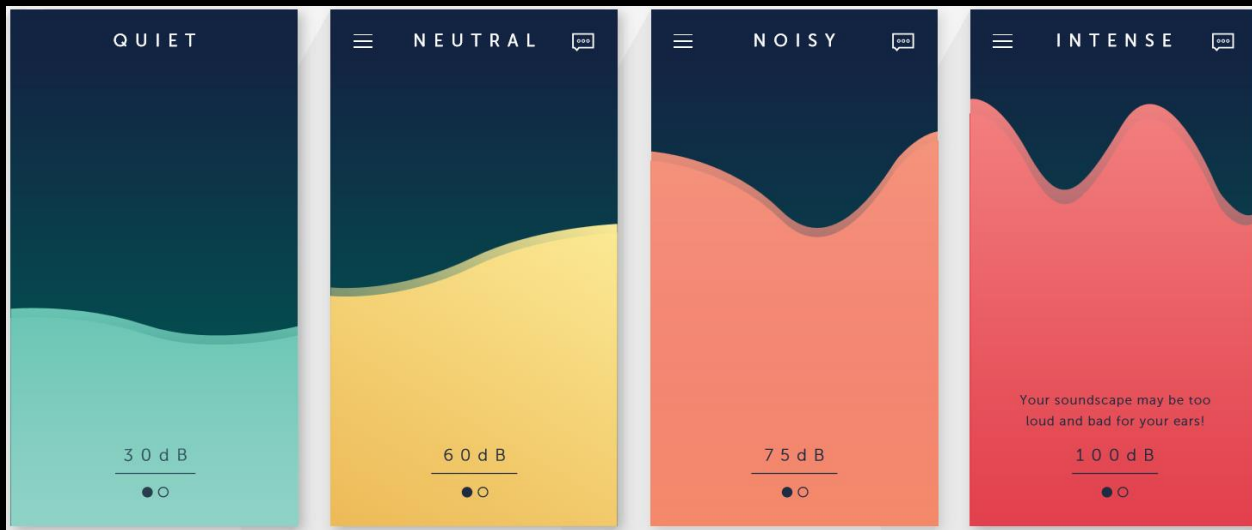


# Hearing AI









# Visualizing sounds



# Microsoft Translator

Microsoft Corporation Productivity

Everyone

This app is compatible with all of your devices.

Top Developer

★★★★★ 426

Installed



Transcribe  
and Translate  
spoken  
conversations

×

SPEECH TO TEXT ●

Hey There.  
I was wondering  
if you could tell me where  
the local store is?

×

SPEECH TO TEXT ●

Here we are!  
Let me know if you need  
something from me.

Okay?

×

SPEECH TO TEXT ●

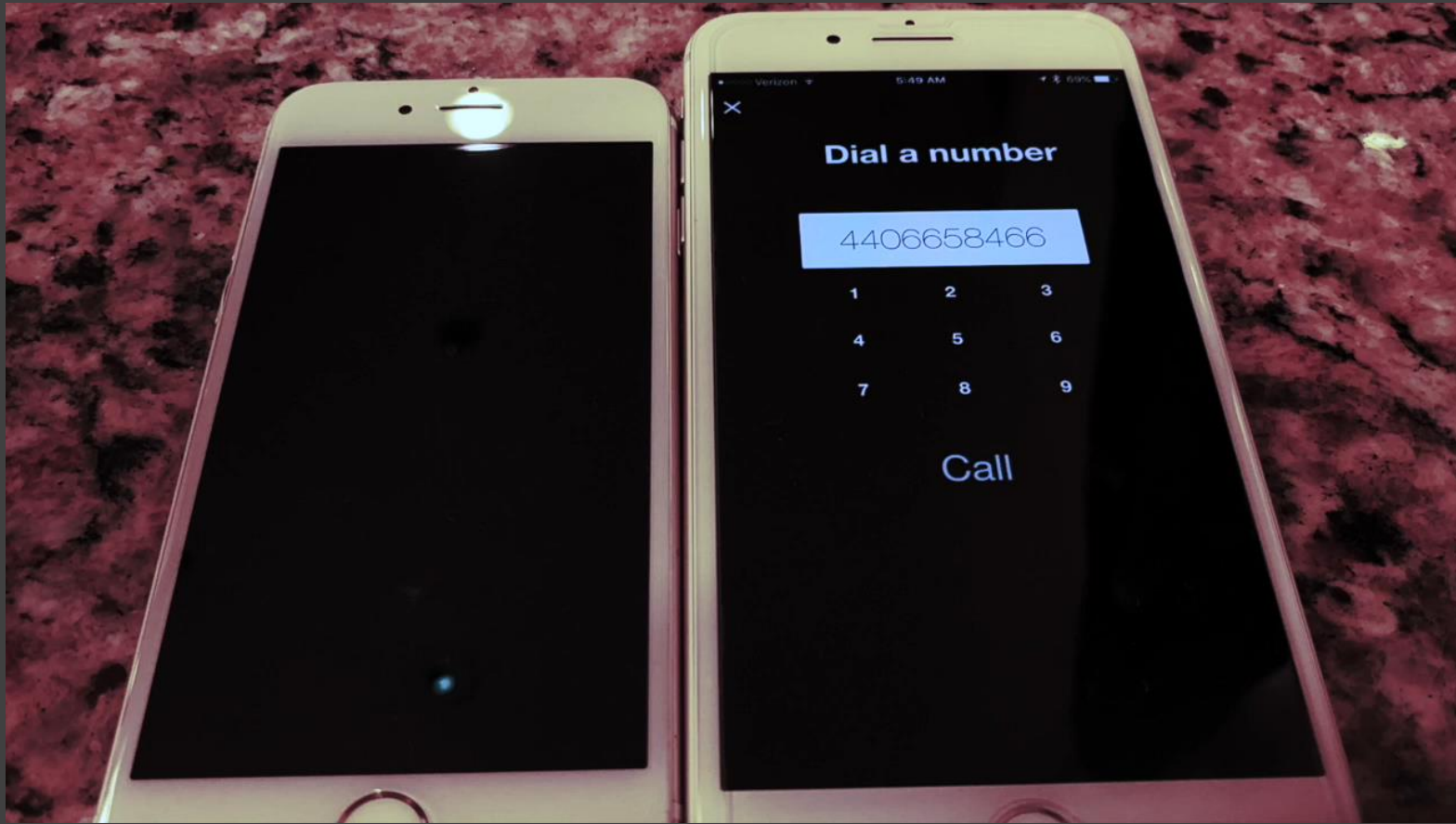
Awesome! Whoever  
came up with that  
must be really smart.

×

SPEECH TO TEXT ●

We should  
leave. Fire  
alarm just  
went off

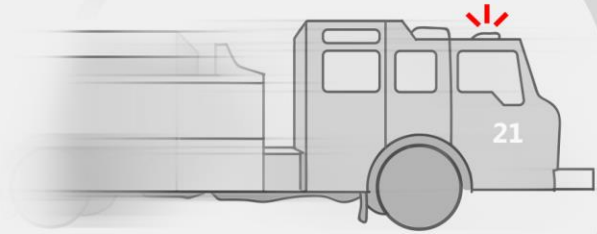
# Capturing loudness real-time



Transcribing phone calls real-time



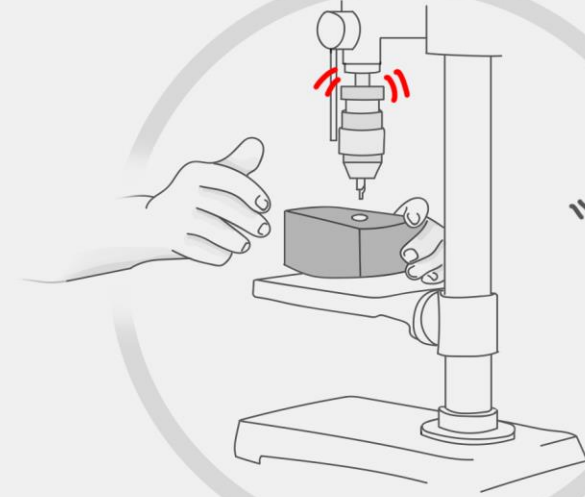
Hearing AI can do more...



"Siren"



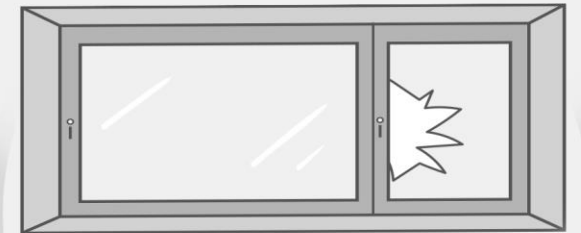
"coffee shop"



"Manufacturing"



"Music"



"broken glass!"

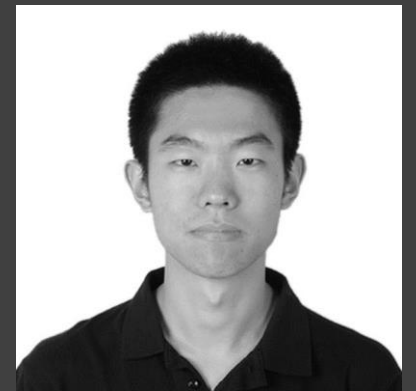
Can we scale up?

---

# Deep Learning on audio in Azure

---

Xiaoyong Zhu





# URBAN SOUND DATASETS

TWO DATASETS AND A TAXONOMY FOR URBAN SOUND RESEARCH

[READ MORE](#)





Is  
everywhere  
you are...!

CNN on  
audio? How?

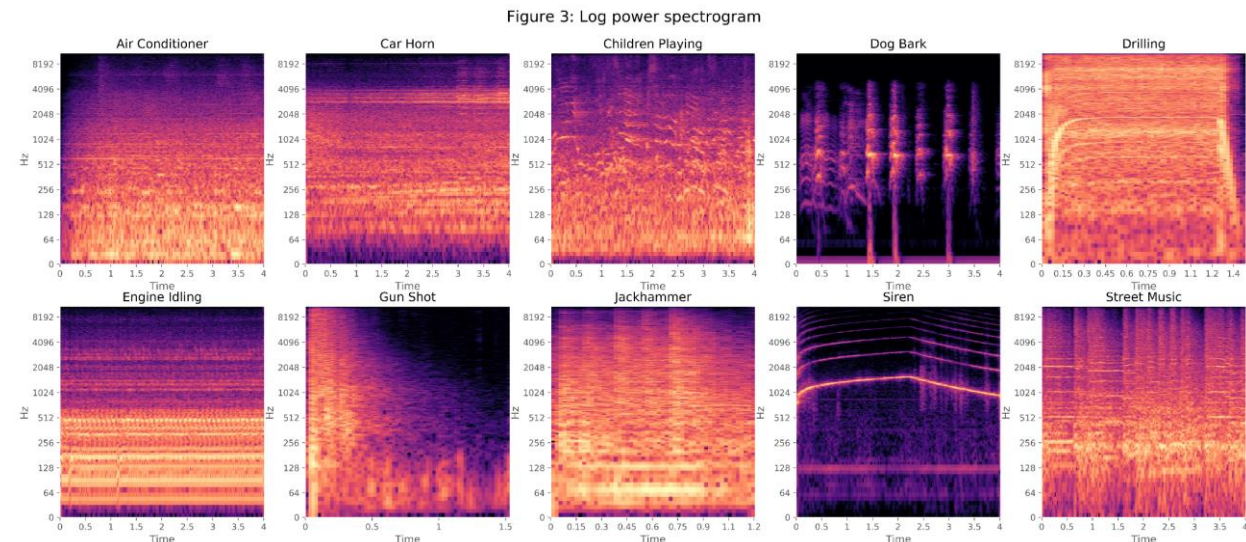
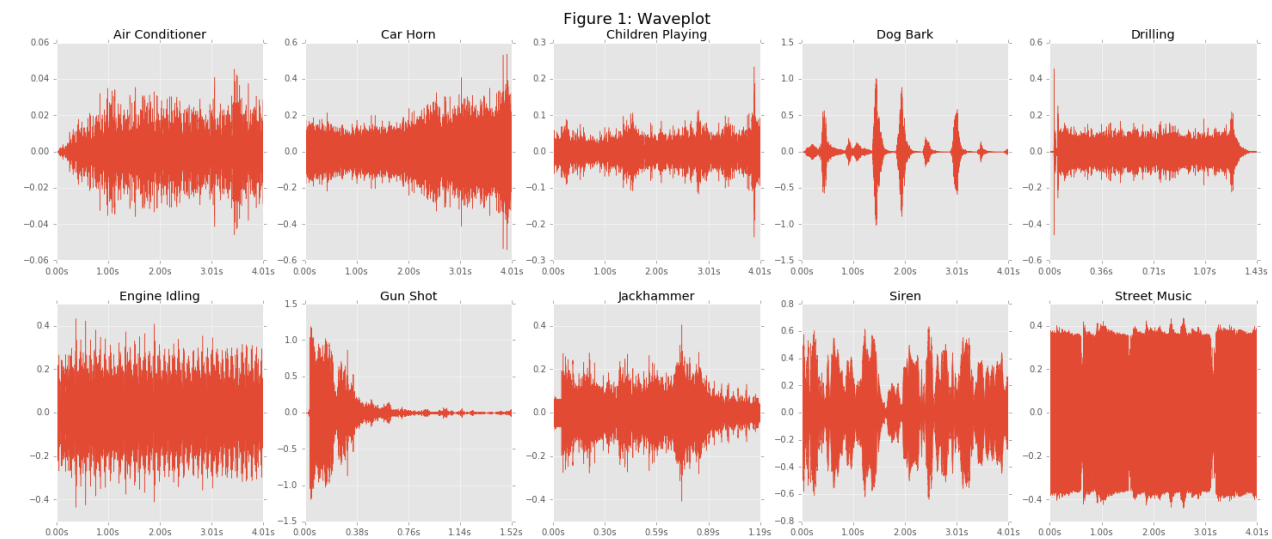


Trick

Convert 1D Audio to 2D image

2D image = mel-spectrogram

Input to CNN = mel-spectrogram

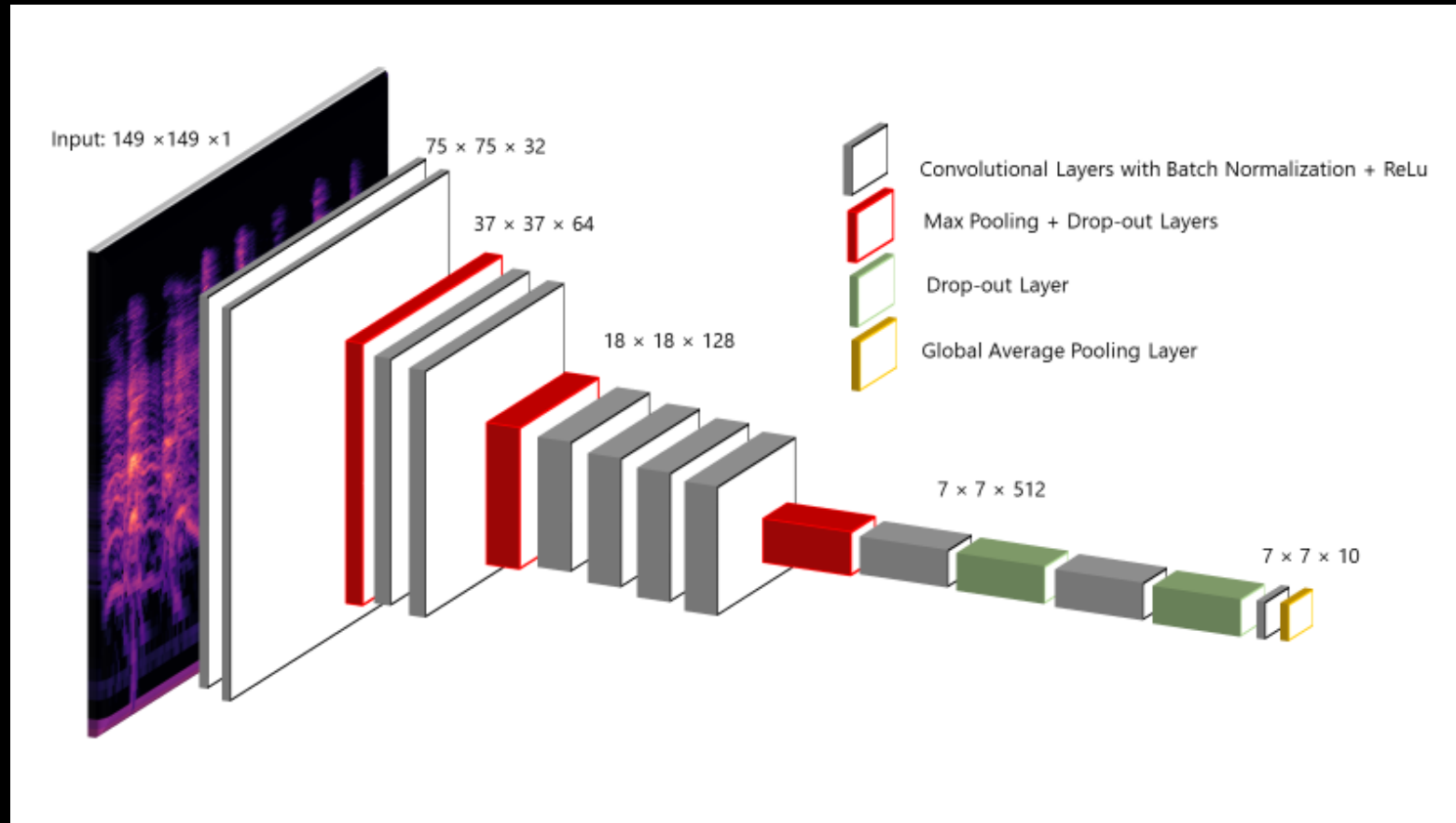


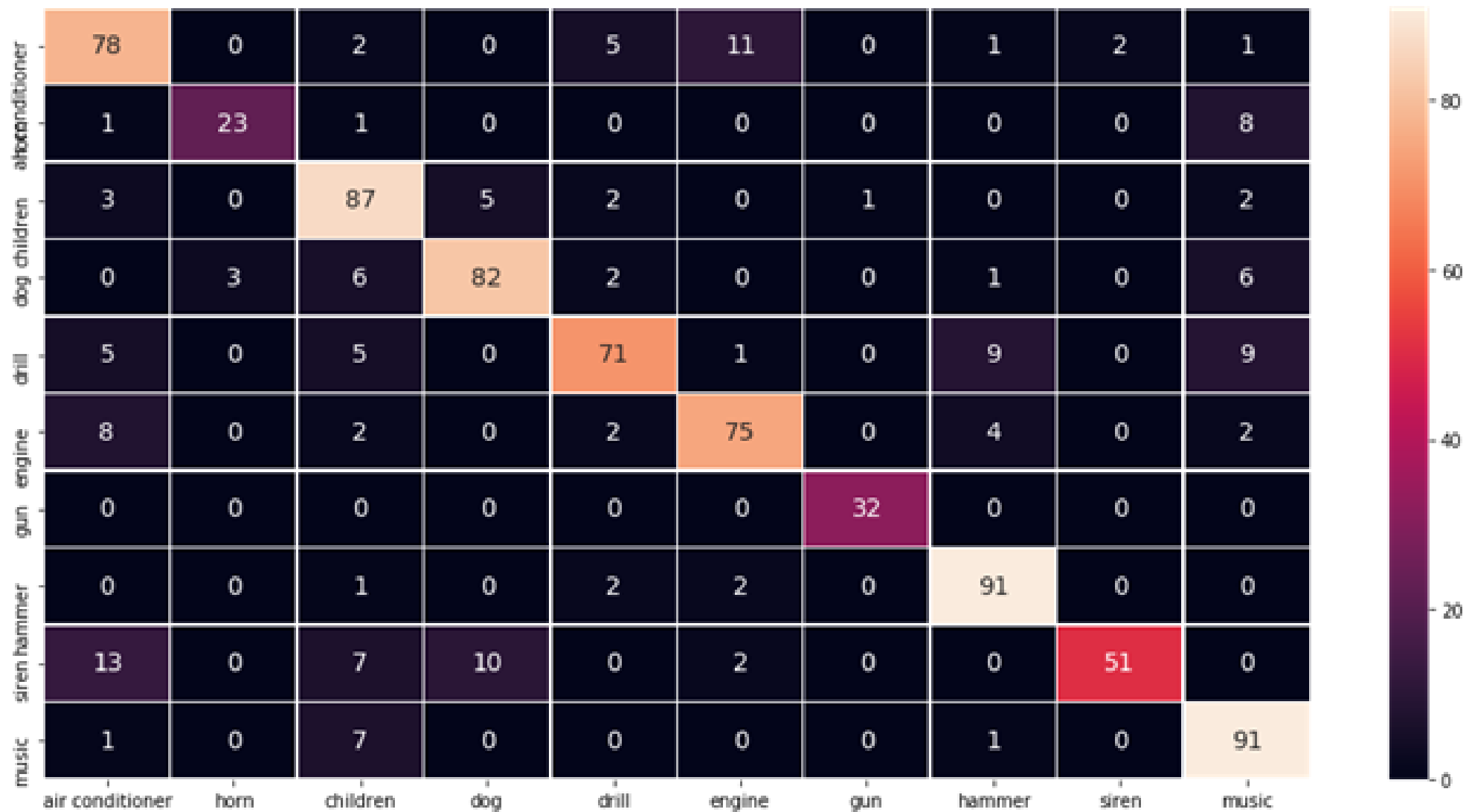
Convert audio amplitude to what human ears hear

# AI Algorithm

11 Layer CNN

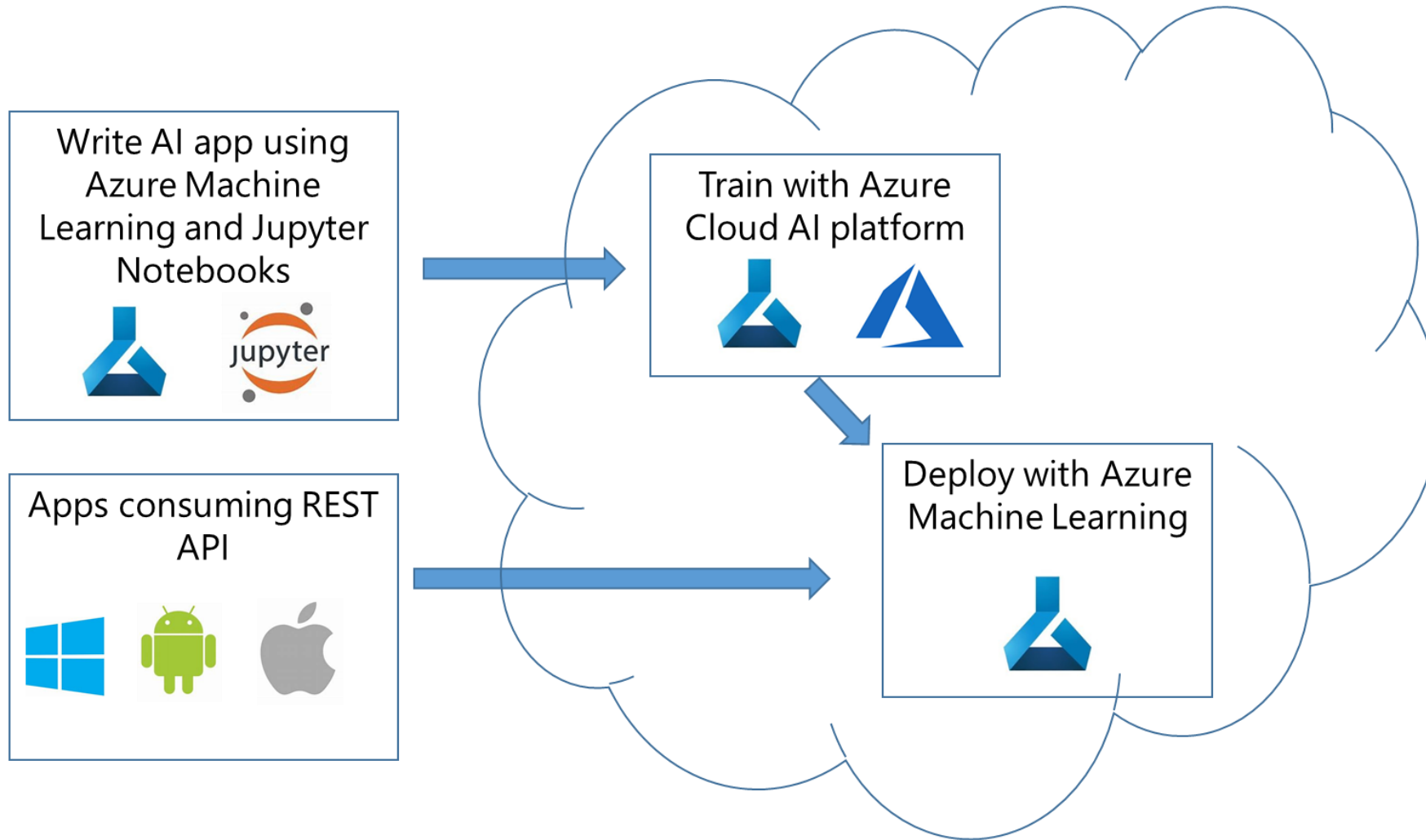
Batch normalization layers



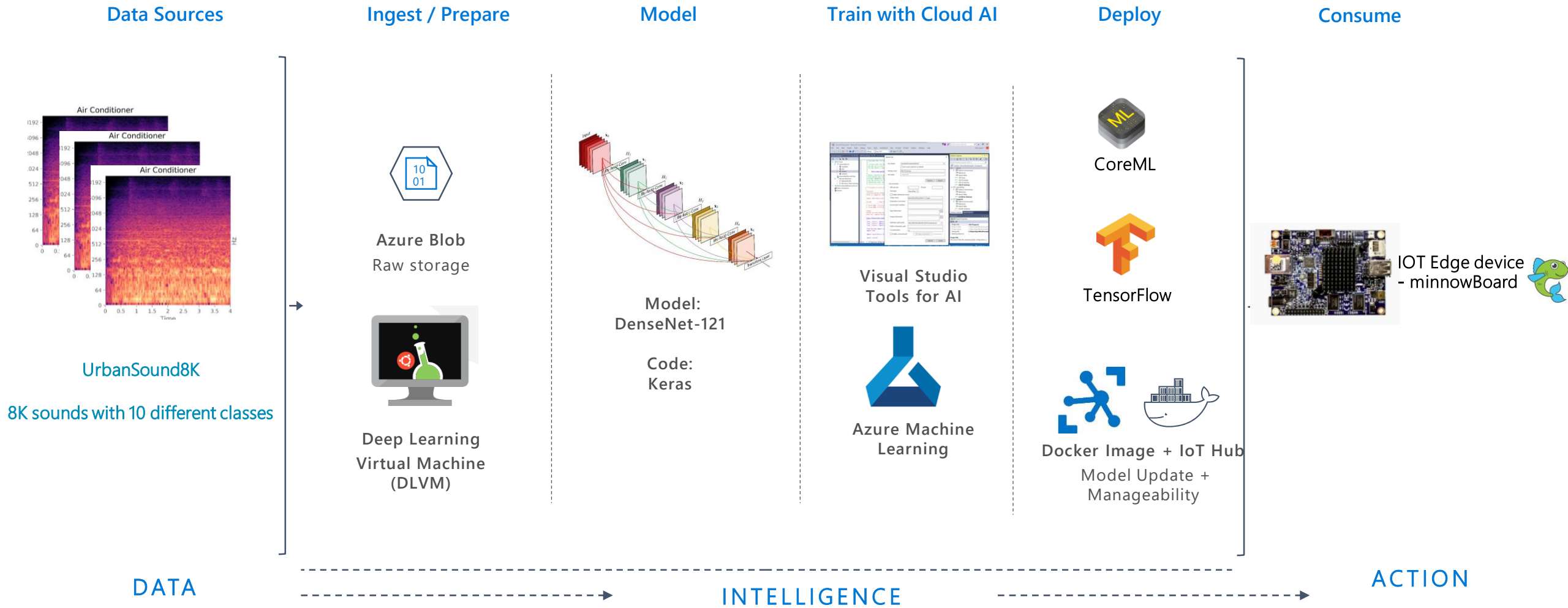


# Performance





Deployment: can be utilized by any application



# Intelligent Sound Prediction - Architecture

<https://aka.ms/hearingaiblog>

Open source

Deploy a backend to identify  
sounds in few minutes!

# Resource Links

Microsoft at CSUN  
[Aka.ms/CSUN2018](https://aka.ms/CSUN2018)

Answer Desk for Accessibility Support  
[aka.ms/eDAD](https://aka.ms/eDAD)

Microsoft Accessibility Overview  
[Microsoft.com/Accessibility](https://Microsoft.com/Accessibility)

# We want to hear from you.



Connect with us in person today or anytime through our support and feedback channels.

Disability Answer Desk for phone/chat/ASL Help:  
[Aka.ms/AccessibilitySupport](https://aka.ms/AccessibilitySupport)

User voice Forum for Accessibility Feedback:  
[Aka.ms/AccessibilityFeedback](https://aka.ms/AccessibilityFeedback)