

# 01 SOUNDSCAPE

SOUNDSCAPE è una mappa sonora e visiva che permette di raccogliere i suoni a chi ama viaggiare. Ha lo scopo di generare una nuova esperienza di viaggio nel quale si avrà la possibilità di concentrarsi sulle sonorità che rendono unici i luoghi visitati. Inoltre, agli utenti permetterà di raccogliere dei veri e propri album sonori, da poter riascoltare a casa.

**angelica gasperoni**

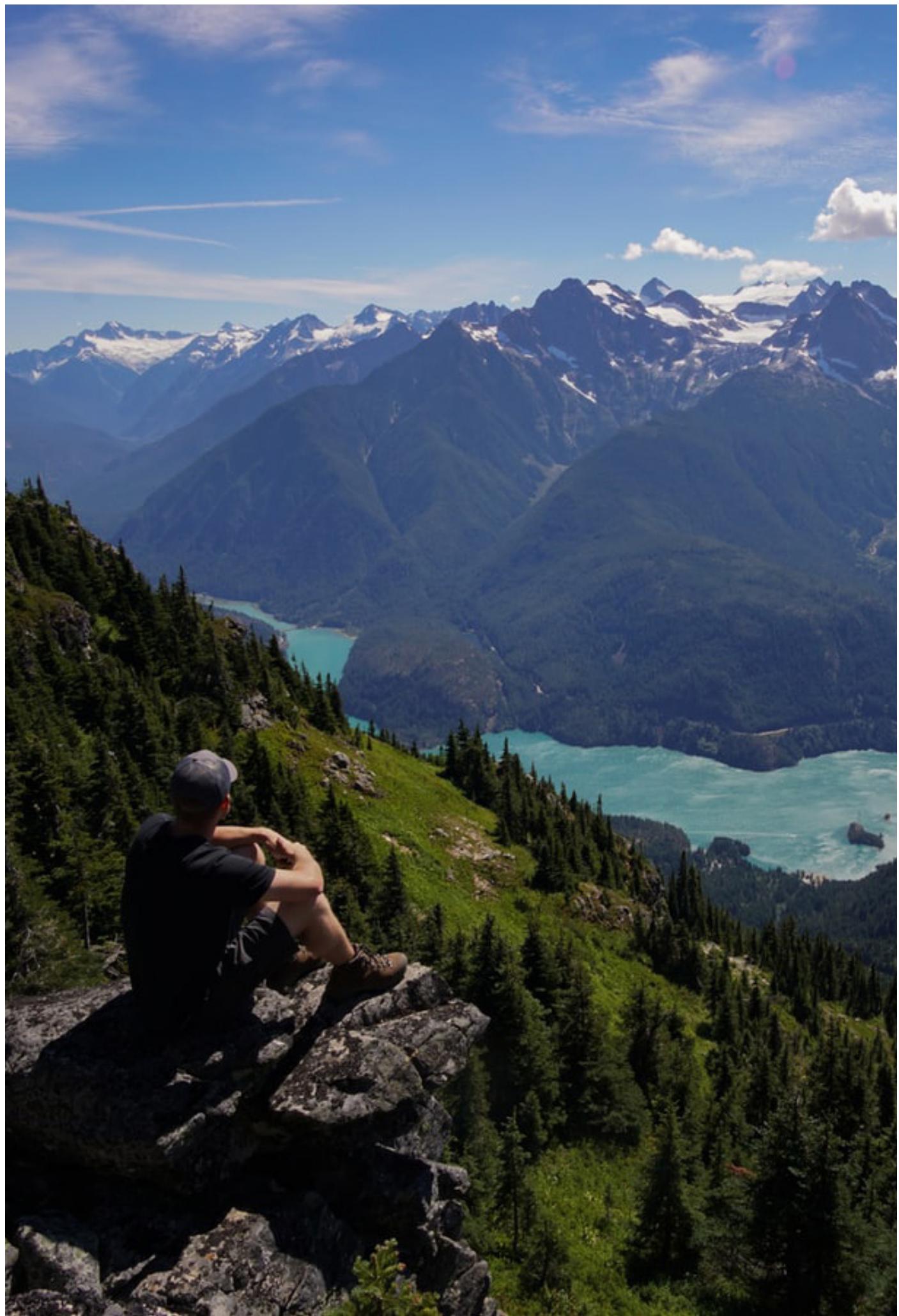


#soundclassifier  
#mapping  
#soundscape  
#travel  
#machinelearning

[github.com/dsii-2020-unirsm](https://github.com/dsii-2020-unirsm)  
[github.com/Angelica147](https://github.com/Angelica147)

a destra

immagine esplicativa di un utente tipo in una situazione d'uso dell'applicazione unsplash.com/



## Making Invisible

Durante le prime settimane del corso, ci è stato chiesto di riflettere su come fosse possibile rendere “Visibile l’invisibile”. In un momento in cui tutta l’Italia si è trovata in una situazione di Lockdown causata dal Covid-19, abbiamo avuto molto tempo per riflettere su ciò che prima davamo per scontato: in particolar modo, io mi sono soffermata su tutti quei suoni che in questo periodo siamo tornati a sentire e ad ascoltare. Da qui è nata la mia riflessione su quello che poteva essere un ambito di ricerca. Questi suoni, nonostante appartenessero alla mia vita quotidiana, erano suoni nuovi, che spesso non riconoscevo, come i versi di alcuni animali o i suoni provenienti dai miei vicini che si dilettavano in nuove attività nei loro giardini. Così mi sono chiesta come avrei potuto mostrare quei suoni che hanno definito le mie giornate chiuse in casa, suoni che probabilmente ho sempre ascoltato, senza rendermene conto perché presa da altro e che spesso non riuscivo a riconoscere. Come avrei potuto rendere progettualmente visibili questi suoni “invisibili”: l’idea iniziale era quella di realizzare una macchina che ascoltasse il suono, lo riconoscesse e tramite esso, ricostruisse delle immagini cercando di ricreare il mondo al quale appartenesse.

## Casi studio

Definendo i primi passi per approcciarmi al Machine Learning, ho cercato di capire cosa oggi fosse stato realizzato e in che modo sarebbe potuto essere elaborato per realizzare l’idea di progetto iniziale. Sono stati due i casi studio più rilevanti che ho analizzato.

**Bird Sounds** è un progetto di Kyle McDonald, Manny Tan, Yotam Mann per il laboratorio di Ornitologia della Cornell del 2016 che organizza centinaia di versi di uccelli in una visualizzazione interattiva. La macchina è in grado di ascoltare e riconoscere i suoni tramite dei tag che associano a ciascun suono, il nome della specie animale al quale appartiene.

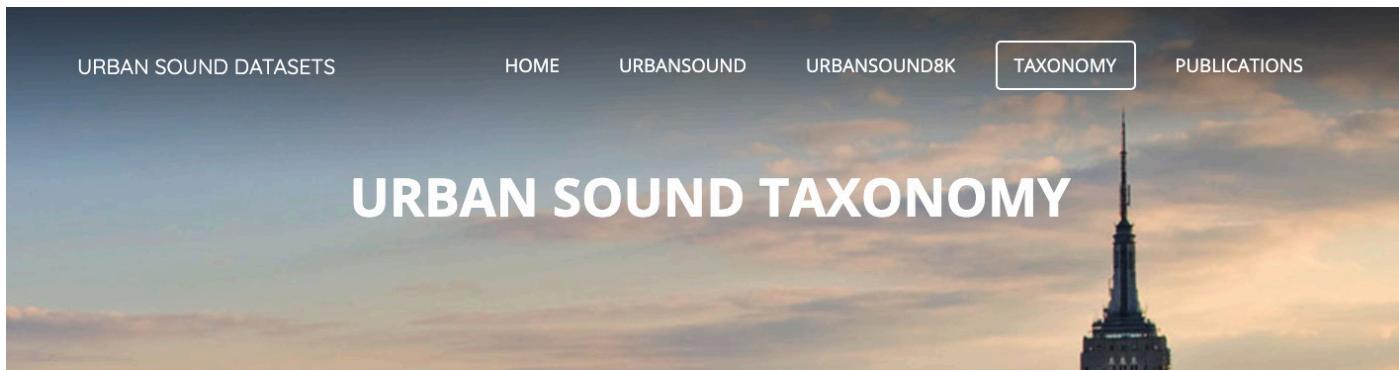
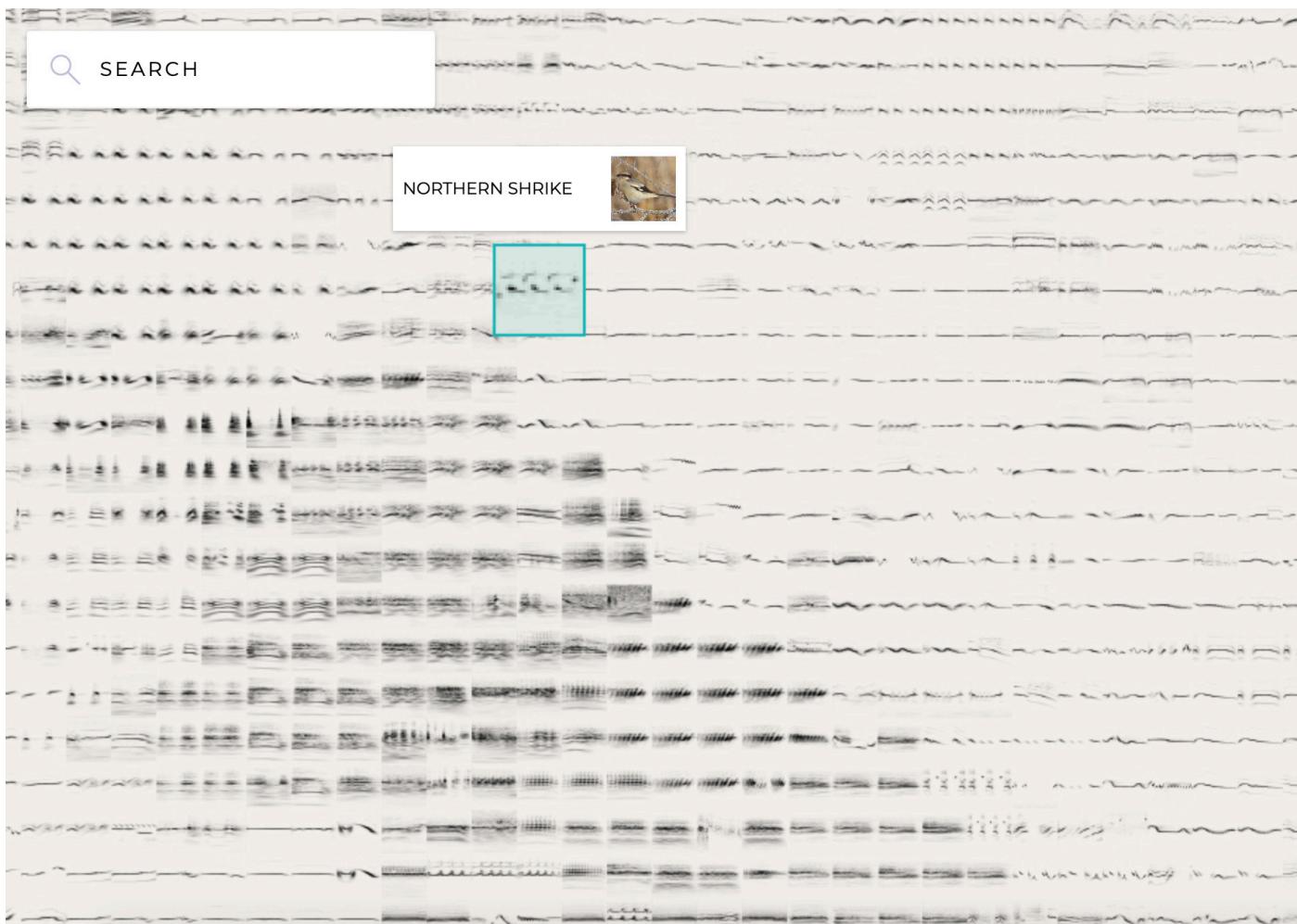
**US8K** è un data set contenente 8732 estratti sonori con etichetta di suoni urbani di 10 classi, tratte dalla tassonomia urbana. Questo progetto è stato realizzato da Justin Salamon, Christopher Jacoby e Juan Pablo Bello nel 2014 per la NYU. I suoni urbani sono stati spesso limitati a sottoinsiemi di tassonomie più am-

in alto

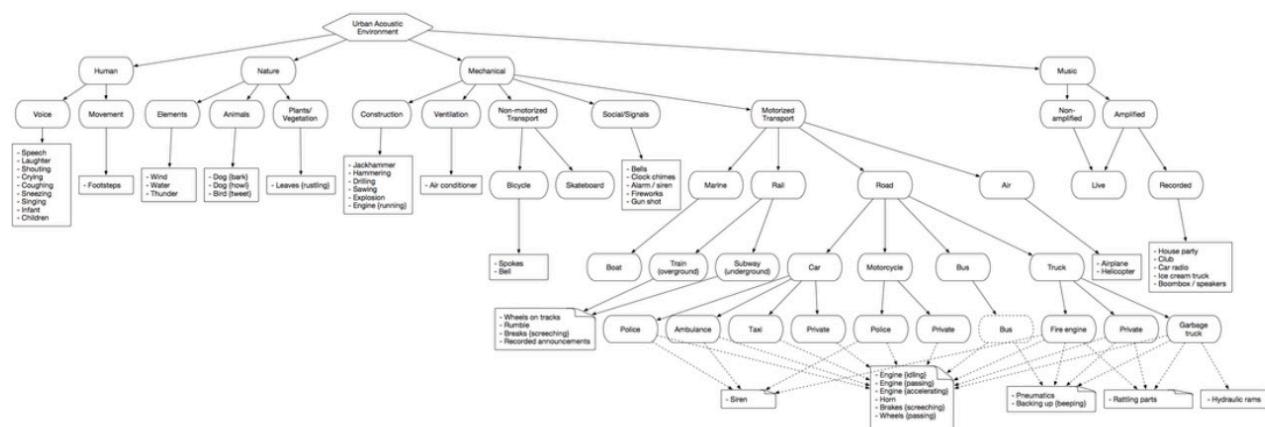
immagine di “Bird Sounds”  
(progetto di Kyle McDonald,  
Manny Tan, Yotam Mann per  
il laboratorio di Ornitologia  
della Cornell del 2016)  
<https://experiments.withgoogle.com/ai/bird-sounds/view/>

in basso

immagine di “US8K”  
(progetto di Justin Salamon,  
Christopher Jacoby e Juan Pablo  
Bello nel 2014 per la NYU)  
<https://urbansounddataset.weebly.com/urbansound8k.html#>



The sound classes in the UrbanSound and UrbanSound8K datasets are taken from the urban sound taxonomy presented below. [Click here to download a scalable PDF version of the taxonomy](#). For further information about the taxonomy please see [our paper](#) and the information provided below.



piè degli ambienti acustici, e quindi affrontano solo parzialmente l'analisi sistematica del suono urbano.

## Concept

Dopo aver analizzato diversi progetti sul suono ed essendo incuriosita da questo approccio, ho deciso di lavorare su un progetto di classificazione sonora. In particolar modo, il mio obiettivo, è quello di realizzare un progetto che riconoscesse non solo i suoni che consideriamo quotidiani, ma anche tutti quei suoni che fanno parte del nostro paesaggio sonoro ma passano spesso in secondo piano. Come secondo passo, la macchina assocerà a ciascun suono riconosciuto delle immagini o dei tag che ne identifichino la provenienza e ne descrivano l'ambiente sonoro di appartenenza.

## Analisi degli algoritmi

Una volta definita l'idea iniziale di concept, sono passata alla prototipazione. Come per il concept, il primo passo è stato quello di analizzare e ricercare quello che già era stato realizzato. In particolar modo, ho analizzato gli algoritmi di Shiffman per il riconoscimento dei suoni e la classificazione di essi.

Sono stati principalmente due gli algoritmi studiati: SoundClassifier18words (un algoritmo che permette al computer di ascoltare e riconoscere 18 parole e definirne il valore di confidence) e Teachable Machine 3: Sound Classification (un algoritmo che permette di riconoscere tre differenti suoni).

Una volta capito come l'analisi dei suoni funzionasse, ho provato ad applicarvi delle piccole modifiche, così da riuscire a capirne al meglio le potenzialità. In particolar modo mi sono concentrata sul modificare l'ambientazione in cui i suoni venivano ascoltati (luoghi più o meno rumorosi, per capirne i limiti) e tipologia del suono ascoltato (verso dello stesso animale, ma riprodotto da sorgenti differenti oppure in contemporanea o meno con altri suoni).

## Prototipazione

Una volta sperimentati questi algoritmi, ho cercato di capire come avrei potuto utilizzarli al meglio per il mio progetto. La prima parte di riconoscimento

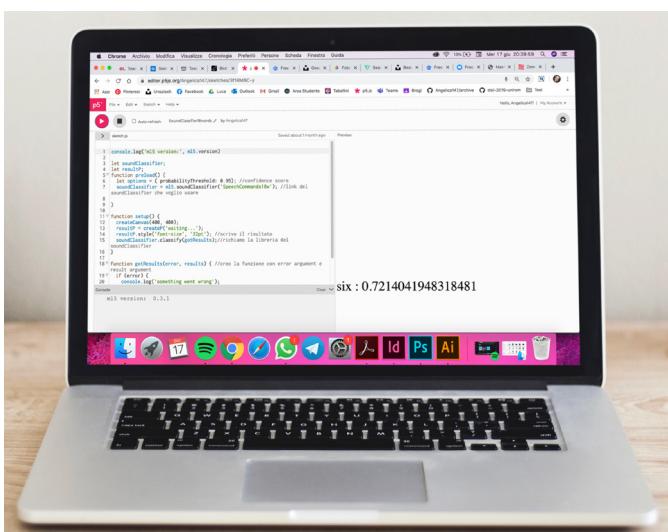
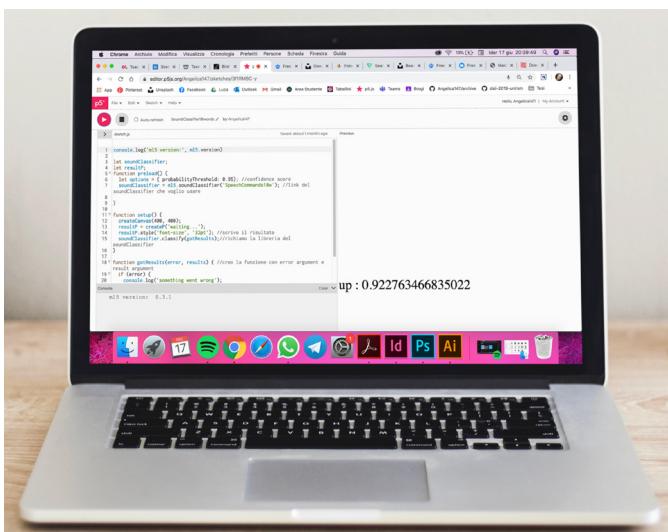
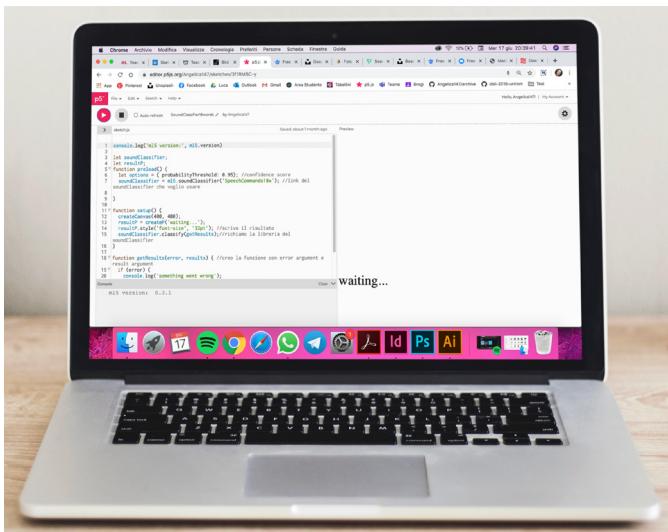
1-3

screenshot algoritmi  
SoundClassifier18words  
prototipati visualizzabile  
per esteso su [github.com/Angelical47](https://github.com/Angelical47). La prima foto in alto rappresenta la fase di ascolto, le consecutive,  
rappresentano la fase di  
riconoscimento della parola.

4-6

screenshot algoritmi  
Teachable Machine 3 prototipati  
visualizzabile per esteso su  
[github.com/Angelical47](https://github.com/Angelical47). La  
prima foto in alto rappresenta la  
fase di ascolto, le consecutive,  
rappresentano la fase di  
riconoscimento del suono.

1	4
2	5
3	6



del suono è stata realizzata utilizzando Teachable Machine (uno strumento che permette all'utente di addestrare un proprio modello di apprendimento automatico e riconoscere così, suoni o pose). In particolar modo, la macchina è stata allenata a riconoscere e distinguere dal paesaggio sonoro, quelli che sono i versi dei cani e dei gatti e indicare a quale animale appartiene il suono ascoltato. Successivamente, è stato unito l'algoritmo di Sketch-RNN (un software in grado di riconoscere un'immagine e completare un disegno partendo da qualunque forma grazie all'analisi di numerosi disegni appresi nel tempo), collegando così la sua libreria di disegni ai suoni. In questo modo, l'algoritmo, dopo aver riconosciuto la tipologia del suono, è in grado di iniziare a disegnarlo, riproducendo gli schizzi di Sketch-RNN. Entrambi gli algoritmi usano gli stessi tag per identificare i suoni e i disegni.

## SOUNDSCAPE - Cos'è?

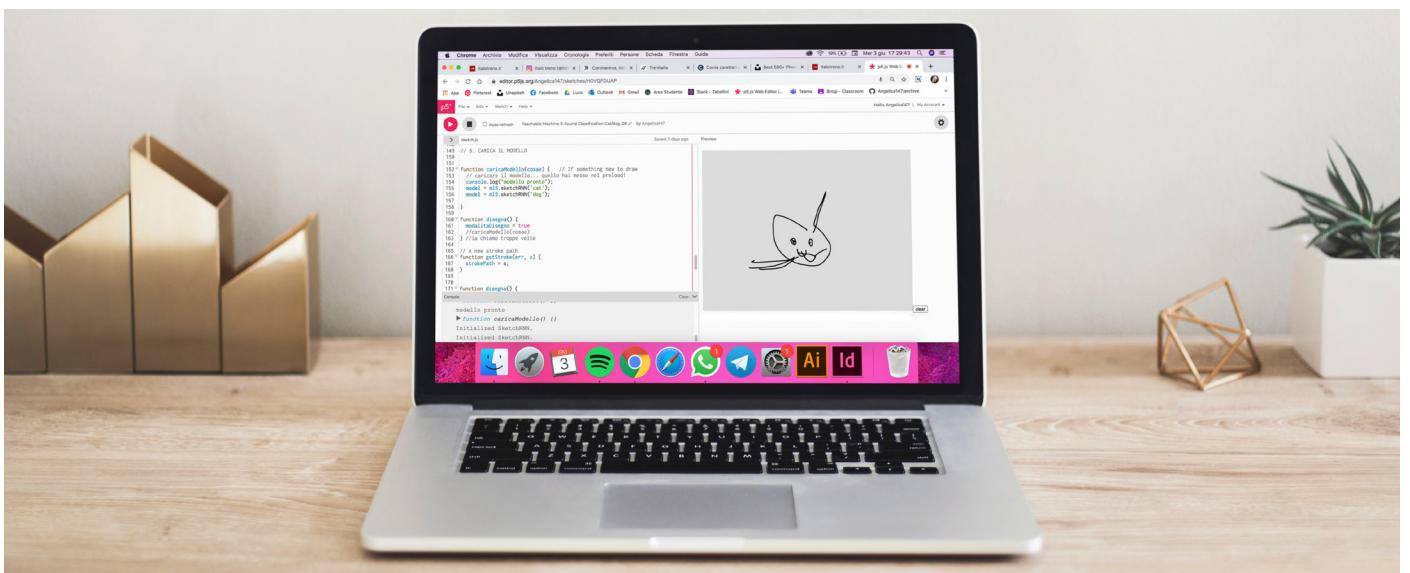
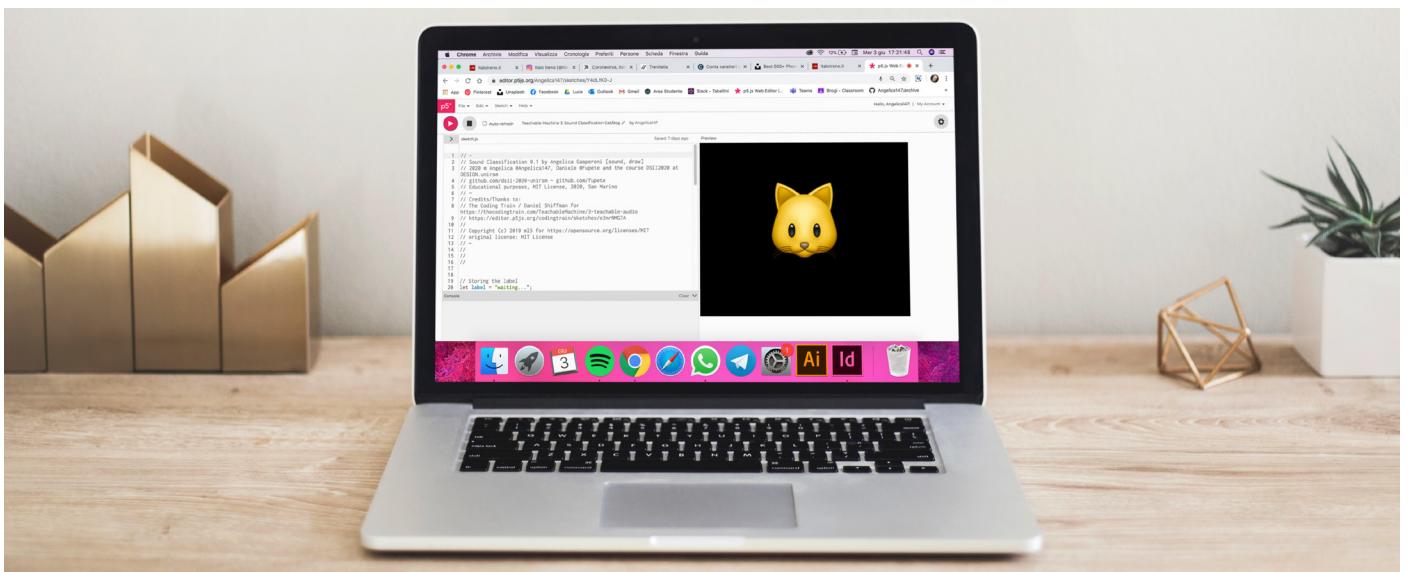
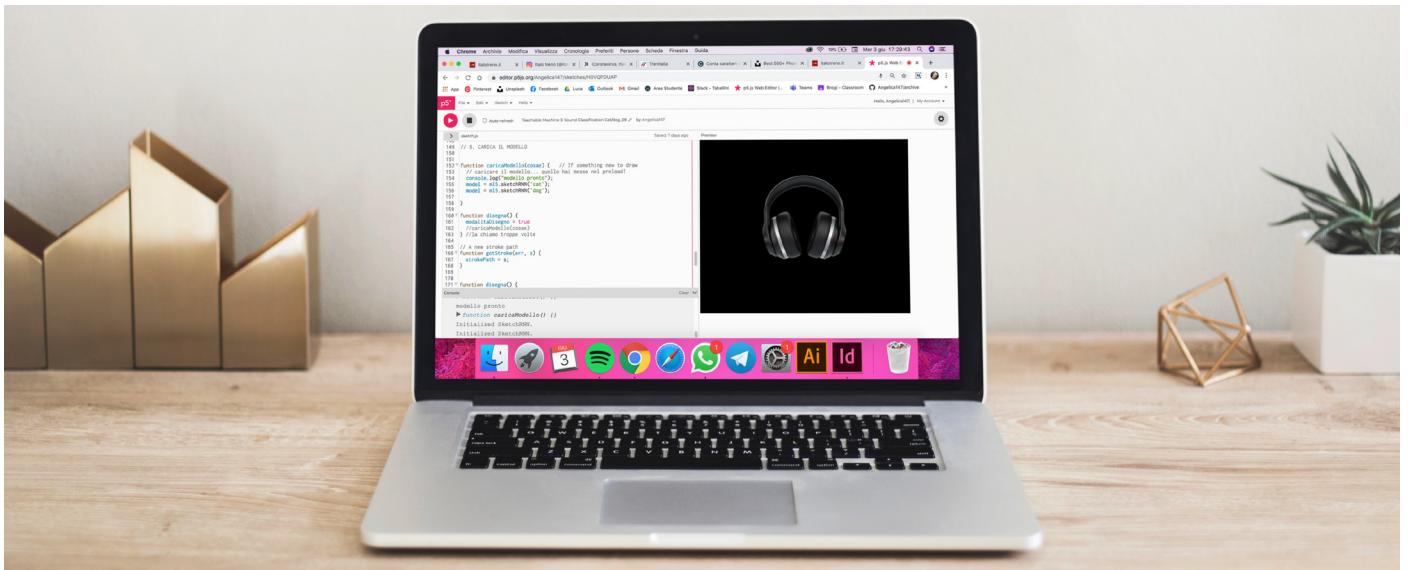
Una volta realizzato il prototipo, sono andata a ridefinire con le nuove conoscenze apprese, quello che era il concept iniziale di progetto. Da qui, è nata l'idea di SOUNDSCAPE: una mappa sonora e visiva che permette di raccogliere i suoni a chi ama viaggiare. Ha la capacità, tramite Machine Learning, di ascoltare quello che è il paesaggio sonoro in cui si trova l'utente e riconoscere i suoni che lo costituiscono. Inoltre, dà la possibilità di memorizzarli e categorizzarli attraverso la geolocalizzazione. In questo modo, i viaggiatori potranno mostrare e rivivere quegli ambienti e quei suoni che hanno reso unici i loro viaggi, rivivendo le emozioni provate e grazie alla categorizzazione dei suoni, potranno trovare nuove mete di viaggio che hanno le stesse sonorità.

L'applicazione è ideata per chiunque ami viaggiare e visitare luoghi nuovi, per chi, oltre alle classiche foto, vuole riportare a casa una piccola parte di quei luoghi che ha visitato. L'utente infatti ha la possibilità di registrare i suoni degli ambienti in cui si trova e visualizzarli in una mappa, in base al luogo in cui sono stati registrati. Una volta a casa, potrà quindi riascoltarli e immergersi nuovamente nei suoi luoghi preferiti di viaggio.

1  
immagine del prototipo nella fase di ascolto del suono visualizzabile per esteso su [github.com/Angelica147](https://github.com/Angelica147).

2  
immagine del prototipo nella fase di riconoscimento del suono visualizzabile per esteso su [github.com/Angelica147](https://github.com/Angelica147).

3  
immagine del prototipo nella fase di disegno del suono riconosciuto visualizzabile per esteso su [github.com/Angelica147](https://github.com/Angelica147).



## Cosa significa?

L'obiettivo finale di questo progetto è quello di dare la possibilità ai viaggiatori di tornare a casa con non solo i classici album fotografici, ma delle vere e proprie librerie sonore dei propri viaggi, così da poter rivivere in qualsiasi momento quelle che sono state le loro esperienze di viaggio. Gli utenti potranno creare così un personale itinerario sonoro e scoprire durante il viaggio quelle sonorità che spesso rimangono "invisibili".

## Come funziona?

L'applicazione utilizza la tecnologia del Machine Learning e in particolar modo il Sound Classifier. Innanzitutto viene ascoltato l'ambiente sonoro, poi, ciascun suono viene confrontato con una libreria contenente i suoni dal mondo e identificato. Ad ogni suono riconosciuto, verrà associata poi un'immagine che lo identifica. Una volta che tutti i suoni sono stati riconosciuti, l'utente avrà la possibilità di salvare la traccia creando così una sua personale libreria di paesaggi sonori dei luoghi che ha visitato. Ciascuna traccia verrà salvata con il tag delle posizioni in cui è stata registrata: l'utente, potrà così riascoltare quelli che sono stati i suoi salvataggi e sceglierli sulla mappa che verrà visualizzata sul telefono.

## Cosa succederebbe se?

Per quanto riguarda la prototipazione, il passo successivo sarà quello di creare una vera e propria libreria sonora che contenga il maggior numero di suoni che costituiscono gli ambienti sonori. Non solo i versi delle varie specie animali, ma anche i suoni naturali, come l'acqua di un fiume che scorre o il suono delle foglie degli alberi, e i suoni artificiali, come i rumori delle macchine, delle persone e degli ambienti più cittadini. Più vasta sarà la libreria, più sarà preciso il riconoscimento dei suoni che costituiscono il paesaggio sonoro.

Per quanto riguarda invece il concept di progetto, sarebbe interessante renderlo una vera e propria applicazione social, in cui i vari utenti potranno interagire tra loro, curiosare tra le librerie e i viaggi degli amici, provando a ricreare quelle che sono state le emozioni provate.

1

prototipo dell'applicazione SOUNDSCAPE nelle sue funzioni principali.

2-3

immagini esplicative di situazioni d'uso nella fase di registrazione del suono e nella fase di aggiunta della foto.

4

immagine esplicativa di situazione d'uso: nella fase di visualizzazione e confronto della libreria

1

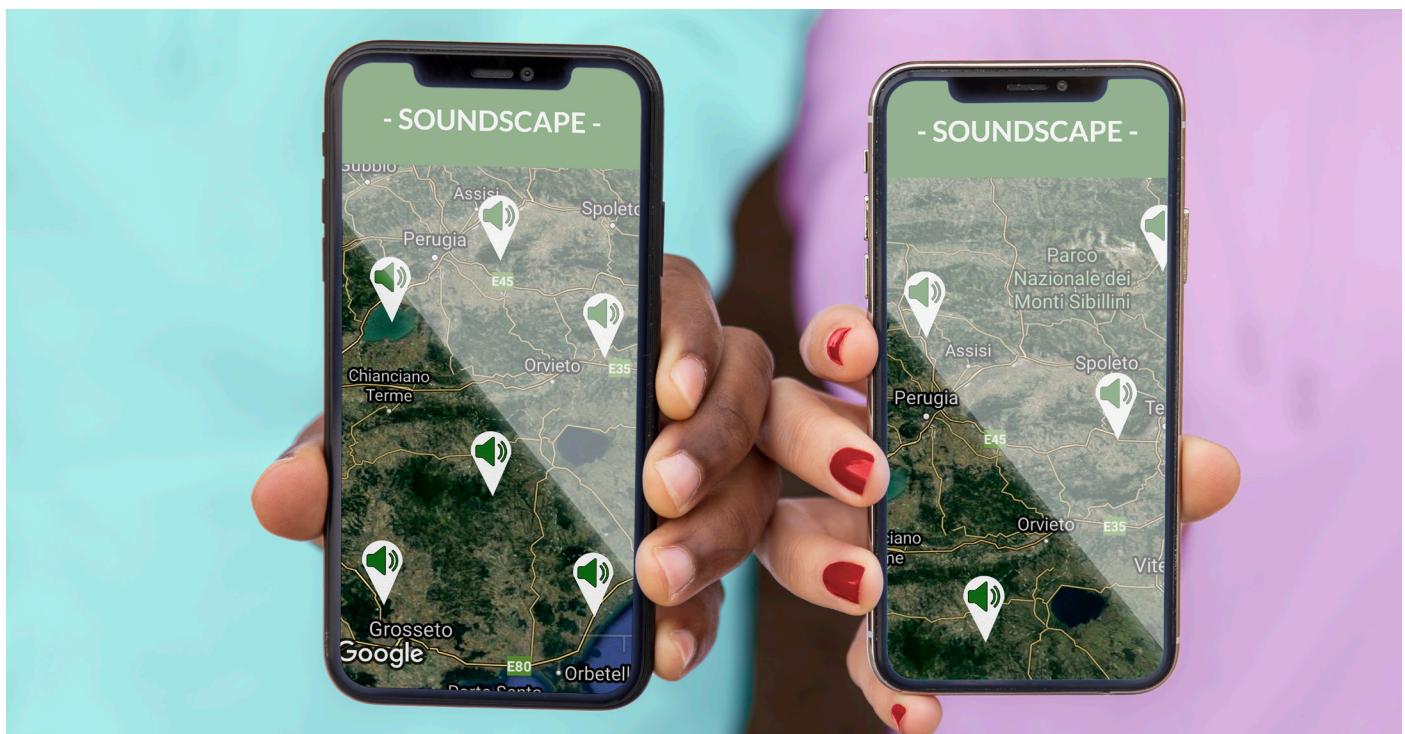
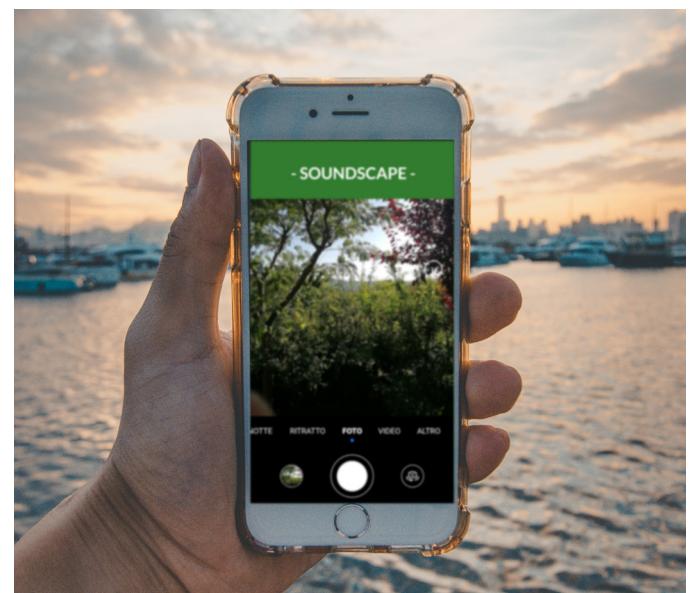
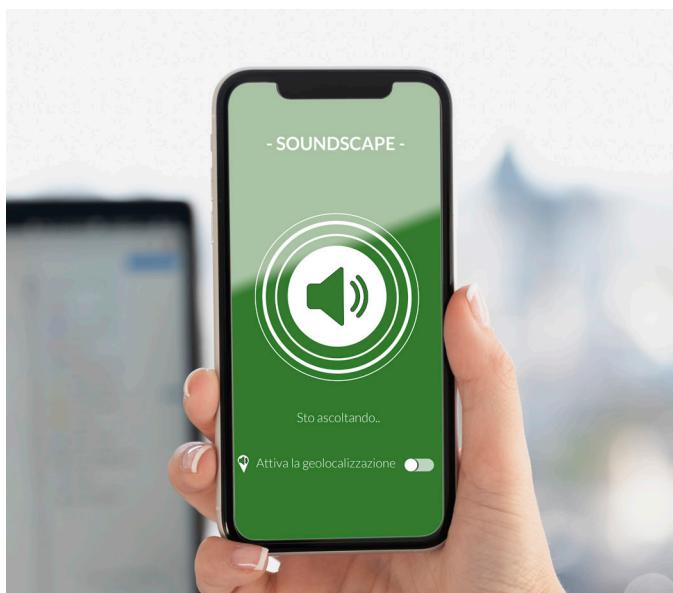
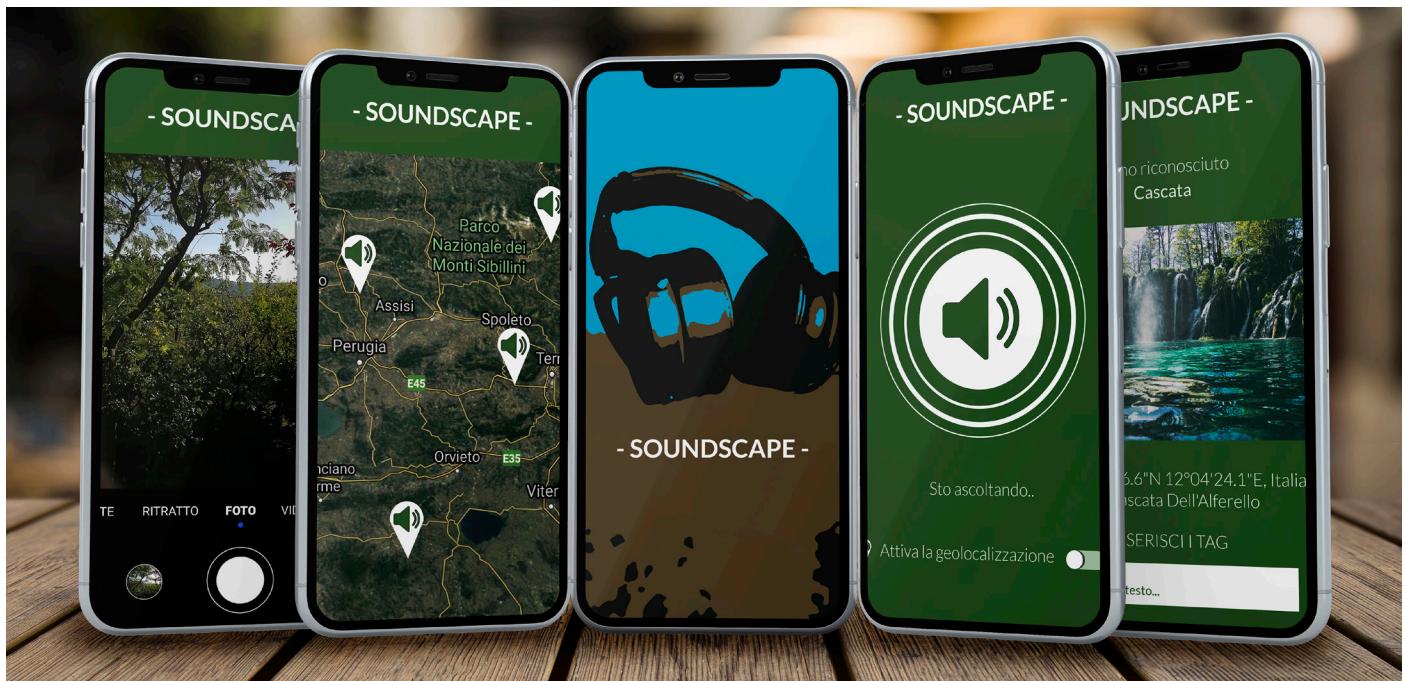


2

3

4





## Sitografia

<https://github.com/tensorflow/tfjs-models/tree/master/speech-commands>

<https://github.com/justadudewhohacks/face-api.js/>

<https://github.com/mikesmales/Udacity-ML-Capstone>

<https://github.com/phillipi/pix2pix>

<https://github.com/googlecreativelab/aiexperiments-bird-sounds>

<https://github.com/shiffman>

<https://www.youtube.com/watch?v=cO4UP2dX944>

<https://www.youtube.com/watch?v=TOrVsLklltM>

<https://ml5js.org/reference/api-soundClassifier/>

<https://ml5js.org/reference/api-PitchDetection/>

<https://experiments.withgoogle.com/ai/bird-sounds/view/>

<https://www.lifetrends.it/sketch-rnn/>

<https://teachablemachine.withgoogle.com/train/audio>

<https://birbs.glitch.me/>

<https://magenta.tensorflow.org/get-started>

<https://www.01net.it/teachable-machine-2-0-il-ma-chine-learning-di-google-accessibile-a-tutti/>

<http://ojack.xyz/PIXELSYNTH/>

[http://www.justinsalamon.com/uploads/4/3/9/4/4394963/salamon\\_urbansound\\_acmmm14.pdf](http://www.justinsalamon.com/uploads/4/3/9/4/4394963/salamon_urbansound_acmmm14.pdf)

<https://labs.fluuu.id/lines/>

<http://www.because-recollection.com/breakbot>

<http://tweetflight.wearebrightly.com/>

<https://creativity.withgoogle.com/sound-canvas/>

<https://urbansounddataset.weebly.com/urbansound8k.html#>

**a destra**  
immagine esplicativa di un  
utente tipo in una situazione  
d'uso dell'applicazione  
[unsplash.com/](https://unsplash.com/)

