

## 08 Beyond the mask

Quest'anno, a seguito di una situazione di emergenza sanitaria, le nostre abitudini quotidiane sono state stravolte.

I nuovi dispositivi di sicurezza, che coprono il volto, come le mascherine, si sono inseriti come un oggetto estraneo.

Ciò ha rappresentato un problema enorme per persone con difficoltà uditive, perché limita moltissimo la possibilità di comunicare. Essi hanno bisogno di leggere il labiale, per poter decifrare le parole che non sentono. Beyond the mask è un app o webapp che rendere visibile il linguaggio verbale e non verbale attraverso una sottotitolazione in realtà aumentata.

**Ilenia Balella**



#mask  
#AR  
#p5speech  
#parole  
#silenzio

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

a destra  
immagine evocativa,  
[unsplash.com](https://unsplash.com)



## Ricerca

La mascherina è entrata nella nostra quotidianità, inserendosi come un oggetto estraneo che filtra e nasconde parte della nostra natura espressiva e comunicativa.

A seguito di questo periodo assolutamente particolare, come stanno reagendo le comunità minoritarie? Sono già partite delle iniziative che promuovono la realizzazione e produzione di mascherine trasparenti, che consentono alle persone non udenti di poter vedere il movimento delle labbra. *DHH Project, Deaf and hard of hearing project*, è un progetto partito negli Stati Uniti, che si è trasformato in pochi giorni in una raccolta fondi per finanziare la produzione di queste particolari mascherine. Negli USA non è la prima volta che se ne parla, già nel 2017 era entrata in commercio una mascherina chirurgica trasparente, *The Communicator Clear Window Mask* prodotta da *Safe 'n Clear*. Anche l'Università degli Studi di San Marino ha sviluppato un programma promosso da Massimo Renno, Presidente di Emporio Etico, insieme ad Alessio Abdolahian, Riccardo Varini, Massimo Brignoni, Karen Venturini, Naomi Galavotti, Andreas Sicklinger e Tommaso Lucinato dell'unità di ricerca sul DDC<sup>[1]</sup>. Il gruppo di ricerca, in collaborazione con alcuni Fab-Lab<sup>[2]</sup>, ha realizzato un dispositivo di protezione individuale che permette di riappropriarsi del potere espressivo e comunicativo del nostro volto. Il progetto risponde alle necessità di un'utenza sordumuta, che in questo periodo si è trovata ad affrontare un ulteriore problema, l'impossibilità di comunicare.

All'interno della comunità sorda ci sono diverse sfumature, ed è bene sottolinearne qualcuna. Non tutti conoscono la lingua dei segni, magari perché sono diventati sordi a una certa età, oppure hanno solo un livello lieve di sordità e imparano spesso a leggere solamente il labiale, (definiti "sordi oralisti")<sup>[3]</sup>. Invece il "sordo segnante" (Russo Cardona & Valtorta 2007) usa la lingua dei segni per comunicare. In Italia la perdita di udito è suddivisa in quattro fasce: lieve, media, grave, profonda.<sup>[4]</sup> Nelle prime due

[1] DESIGN DI COMUNITÀ un gruppo di ricerca che "aiuta le comunità nell'individuazione di risorse culturali, sociali, materiali e immateriali e le accompagna durante il processo di progettazione e attuazione". (<https://ddc.unirsm.sm/>)

[2] Fab-Lab di San Marino e del Veneto, in particolare con il Fab-Lab di Mirco Piccin.

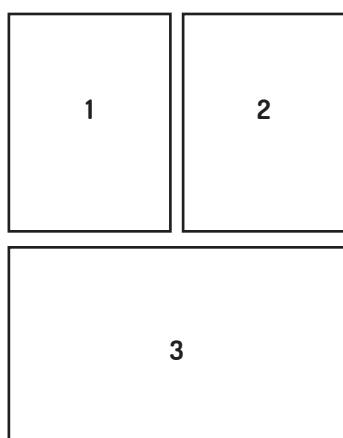
[3] In Italia, il Congresso internazionale di Milano, a partire dal 1880, impose il metodo oralista all'educazione linguistica dei sordi. Per molti anni questa scelta non lasciò spazio a nessun altro metodo educativo.

(Enciclopedia Treccani, n.d.)  
<http://www.treccani.it/enciclopedia/sordita/>

1  
mascherina DHH project, 2020

2  
mascherina del gruppo di ricerca DDC, dell'Università degli Studi di San Marino

3  
The Communicator Clear Window Mask, 2017





fasce la legislazione italiana riconosce un grado più o meno elevato di invalidità, e soltanto da una perdita superiore di 75 decibel si riconosce lo statuto di “sordo”, a coloro che hanno perso l’udito prima dei 12 anni, e per un’ipoacusia pari o superiore a 60 dB all’orecchio migliore (cfr. legge 381/70 pag. 27).

*Oliver Sacks* nel libro “Vedere voci”(del 2014) parla delle differenze tra i sordi postlinguistici, ovvero coloro che hanno perso l’udito successivamente all’acquisizione del linguaggio vocale, e i prelinguistici, che sono nati sordi e non hanno avuto la possibilità di udire alcun suono. La differenza principale tra i due è che i primi, nella loro memoria hanno ancora il ricordo del rumore, e grazie ad un fenomeno chiamato “voce fantasma”, la loro mente proietta automaticamente dei suoni, che hanno in memoria, sull’esperienza visiva. Ovvero riescono a sentire per esempio il suono delle foglie mosse dal vento se guardano un albero, oppure il suono di una voce familiare. Ciò non accade per i sordomuti, che vivono in un mondo di totale assenza di suono. Solo in Italia ci sono circa 5 milioni di audi-olesi, di cui 70.000 sordi prelinguali (Giglioli, 2008). Senza la possibilità di vedere il labiale, non hanno strumenti per leggere.

In questo paragrafo analizziamo alcune tecnologie odierne (-2020) utili allo sviluppo del progetto, che utilizzano la *Speech technology*. Quando si parla di *Speech technology* si intende qualsiasi tecnologia che abbia a che fare con la voce. Ci sono diverse categorie tra cui *speech recognition*<sup>[5]</sup>, *speech verification*<sup>[6]</sup>, *real time speech to text conversion*<sup>[7]</sup>, *interactive voice response* (IVR)<sup>[8]</sup>, *speech synthesis*<sup>[9]</sup> e *speech analytics*<sup>[10]</sup>. La tecnologia TTS (*text-to-speech*) converte il testo scritto in una sintesi vocale. È spesso utilizzata nell’ambito delle disabilità visive, da persone con dislessia o altre difficoltà di lettura. Inoltre la sintesi vocale viene impiegata anche nell’intrattenimento come giochi e animazioni. La NLP, *Natural Language Processing*, è la capacità di un elaboratore elettronico di comprendere un linguaggio nella sua complessità. Infatti adotta un

[4] Decreto Ministeriale 5.2.1992

[5] Il riconoscimento vocale è il metodo mediante il quale un computer riconosce ed elabora il linguaggio orale umano. Spesso utilizzato per applicazioni vocali automatizzate, come call center o sistemi di dettatura.

[6] La speech verification utilizza il riconoscimento vocale per verificare la correttezza del discorso pronunciato. Spesso utilizzata per l’insegnamento della pronuncia assistita dal computer.

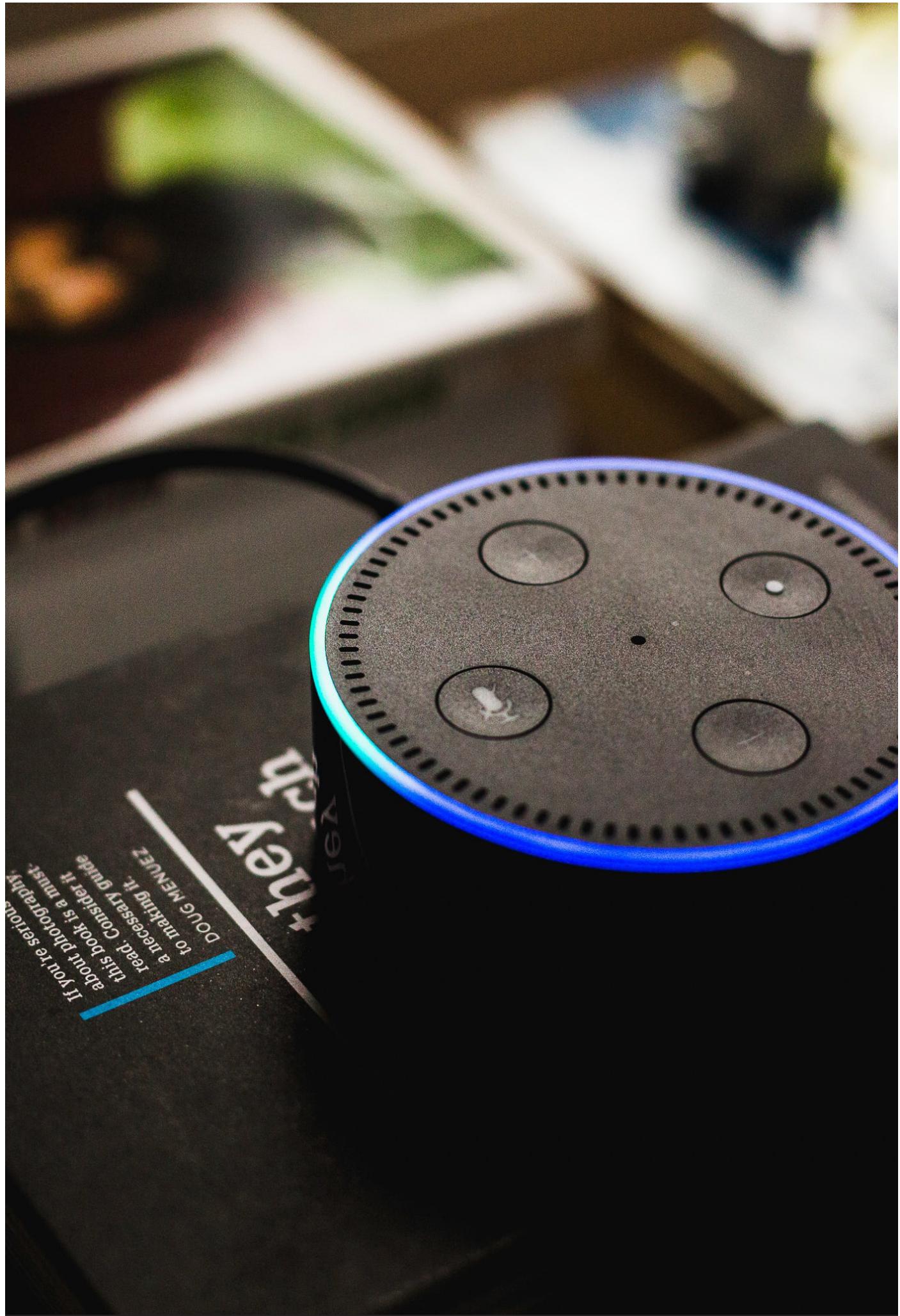
[7] Si tratta di una conversione da voce a testo in real time basata sulla tecnologia di machine learning. La Speech-to-Text di Google consente di convertire l’audio in testo applicando modelli di rete neurale.

[8] Interactive voice response è un sistema utilizzato spesso per alleggerire il carico di lavoro degli operatori nei call center, che interagisce con gli utenti attraverso l’uso della voce e la selezione multipla di operazioni tramite la tastiera telefonica.

[9] La sintesi vocale è la produzione artificiale della parola umana.

[10] La speech analytics è il processo di analisi delle chiamate registrate solitamente usato per raccogliere informazioni sui clienti, per migliorare la comunicazione e l’interazione futura.

**a destra**  
immagine evocativa,  
[unsplash.com](https://unsplash.com)



metodo di *Deep Learning* per l'analisi semantica e sintassi. Riconosce il contesto d'uso di una parola, e utilizza un database per determinare la semantica. Il sistema di Google *Speech-to-speech translation systems*, rileva in automatico il parlato e lo traduce nella lingua di preferenza. Purtroppo il traduttore di testo di Google è stato criticato per aver fatto spesso traduzioni incomplete o errate, ed è spesso descritto come approssimativo (Chuck, 2019).

L'incapacità di Google Translate di produrre una buona grammatica è un problema informatico, ed un limite tecnologico ancora in via di sviluppo (Chuck, 2019). La lingua parlata ha infinite sfumature, ed è estremamente difficile tenerne conto nella programmazione di un software. I modelli per il *machine learning* dovranno essere perciò sempre più accurati.

## Casi studio.

Di seguito sono riportati alcuni casi studio correttati alla ricerca. *SoundShirt* e il *Toyota Sensitive concert* sono due progetti che esplorano le potenzialità della trasmissione acustica tramite la sensazione aptica. Generano un valore aggiunto trasformando delle onde sonore in vibrazione, per dare la possibilità, anche alle persone non udenti, di poter “sentire” la musica con un altro senso. Aprono il campo ad un nuovo modo di fare esperienza, soprattutto nel mondo virtuale, dove la sensazione aptica accentua la qualità dell'avventura digitale.

Un'altro progetto interessante è *Your Typeface*, che sperimenta la modificabilità di un font attraverso le espressioni facciali, un interessante approccio che esplora l'espressività e la comunicazione non verbale applicata alle caratteristiche di un font. La parola acquista un significato che va oltre la semantica, il glifo è carico di una sua forma espressiva. *Voice-mask* invece è un progetto opensource che utilizza arduino. È una maschera “parlante”, ovvero dispone di luci led che si animano imitando il labiale. È un progetto sperimentale disponibile sulla piattaforma di github, pensato per essere prodotto DIY.<sup>[11]</sup>

[11] il link al tutorial:

<https://medium.com/@tgleiel/how-to-make-a-voice-activated-led-facemask-diy-guide-b300bc146f38>

e un video di riferimento:

<https://twitter.com/TylerGlaiel/status/1265035386109128704>

1

*SoundShirt*, CuteCircuit, 2016

2

*Your Typeface*, Overtone, 2020

3

*Toyota sensitive concert*, Toyota e studio Nebula e NoideaLab, 2015

4

*Vicemask*, Tyler Glaiel, 2020

1

3

2

4



( · - · )

# Your Typeface

A digital experience by Overtone and Set Snail



*SoundShirt* è un progetto del 2016 realizzato da una società di moda, specializzata in abbigliamento tech, CuteCircuit, con sede a Londra.<sup>[12]</sup> Si tratta di una maglia che sfrutta la sensazione aptica, per far esperire la musica con altri sensi. È stata realizzata associando determinati strumenti alle diverse parti del corpo, in modo tale da avere delle vibrazioni distinte a seconda del tipo di suono. Per le persone sordi, e non solo, è un nuovo modo di percepire la musica e di sperimentare fisicamente nuove sensazioni anche nel mondo virtuale.

*Your Typeface*, è un progetto che esplora il mondo dei font variabili, realizzato dall'azienda Overtone nel 2020.<sup>[13]</sup> Per questo progetto è stato creato un sito, che permette all'utente di interagire con la variabilità del font utilizzando la webcam. Il font reagisce alle espressioni del viso in tempo reale, l'utente giocherà contraendo tutti i muscoli facciali per scoprire il funzionamento e cercare il carattere tipografico preferito da poter scaricare. Il font varia in base a diversi parametri: l'apertura in altezza della bocca regola il peso del font, la larghezza invece regola la compattezza.

*Toyota sensitive concert*, realizzato in collaborazione con l'Ente Nazionale Sordi, con Noidealab e studio Nebula<sup>[14]</sup> per la casa automobilistica, è stato un evento multisensoriale realizzato per avvicinare il mondo musicale alla comunità sorda attraverso la sensazione tattile. Questo evento ha avuto luogo nel 2015 al Torino Jazz Festival e la Toyota ha realizzato delle sfere che connettono la musica a stimoli tattili e visivi. Le sfere erano collegate agli strumenti dei musicisti che suonavano sul palco. Si potevano riprodurre vibrazioni relative alle diverse frequenze degli strumenti musicali e le sfere si illuminavano a tempo di musica.

*Voicemask* è un progetto del 2020 realizzato da Tyler Glaiel<sup>[15]</sup>, un programmatore e game designer statunitense. Ha realizzato una maschera con una matrice led 8x8, comandata da Arduino, che imita il movimento della bocca. In base a determinate variabili, ha assegnato un “disegno led” ad un range di

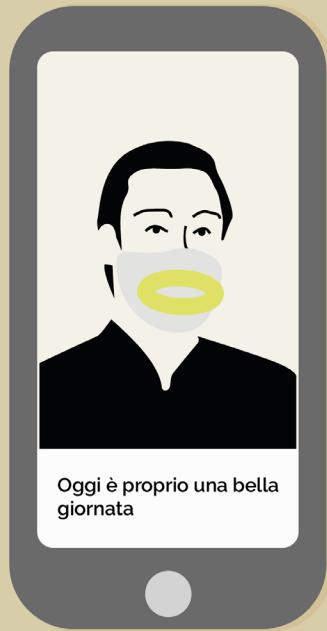
[12] <https://cutecircuit.com/soundshirt/>

[13] <https://www.yourtypeface.com/>

[14] <http://www.studionebula.it/project/toyota-sensitive-spheres/>

[15] <https://github.com/TylerGlaiel/voicemask>

a lato  
immagine esplicativa  
del progetto



volume della voce, in modo tale da avere un movimento fluido dei led. C'è anche la possibilità di far "sorridere" la maschera emettendo il suono "pop".

## Progetto | cos'è

Beyond the mask è una webapp o app progressiva progettata in maniera tale che possa essere prototipata subito con un approccio MVP<sup>[16]</sup>, per ottimizzare i costi di produzione e velocizzare il processo di apprendimento, per poi sviluppare dei prototipi più evoluti.

La webapp sarà utile per le video conferenze, i dibattiti online o le dirette.

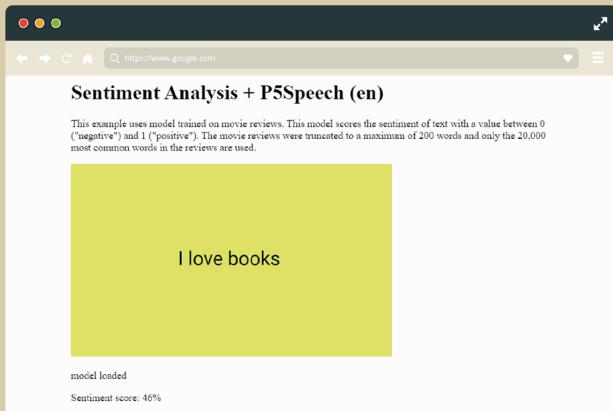
L'app sarà scaricabile ed utilizzabile da smartphone o tablet. Essa sfrutta la telecamera e il microfono del dispositivo, per raffigurare in diretta in realtà aumentata il linguaggio verbale, attraverso una sottotitolazione. Sarà possibile utilizzare un setting attivabile o disattivabile dedicato alla comunicazione non verbale che viene espressa tramite forme e colori che si animano sul volto dell'interlocutore. Questa impostazione è inserita come esperimento e prototipazione per capire le potenzialità dell'espressività veicolata da forme e colori in un contesto comunicativo. L'elemento astratto dovrebbe favorire la trasmissione di valori semantici ed emozionali, celati all'interno della comunicazione, e aiutare a restituire umanità al volto. La componente grafica aggiunta è rappresentata da una forma animata, ovale o rettangolare, che simula il movimento di una bocca che parla, variando la dimensione della sua altezza e larghezza. Se l'emozione veicolata è positiva, il riquadro sarà rappresentato da un'ellisse (gialla se molto positiva, blu se neutra), al contrario se è negativa, il riquadro sarà rettangolare di colore rosso.

L'esigenza progettuale di utilizzare una app nasce dalla necessità e dalla possibilità di usufruire delle sue funzioni più facilmente ed in qualsiasi momento della giornata, per strada, in ufficio, in casa, in un supermercato, al cinema. Attraverso l'utilizzo della sottotitolazione, l'utente potrà comprendere il discorso, anche se l'interlocutore indossa una mascherina, e riuscire ad interagire in sicurezza. Infine, inquadrando il volto della persona

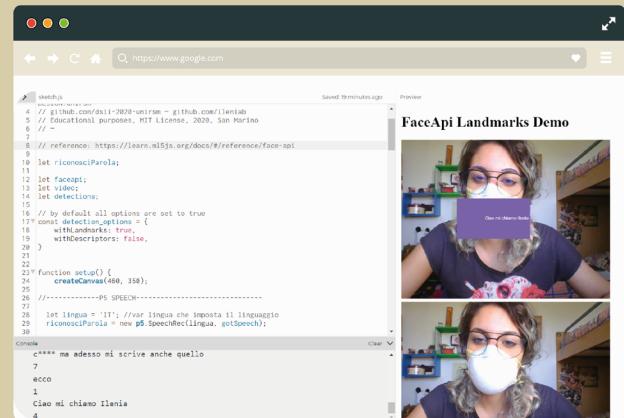
[16] il prodotto minimo funzionante (Minimum Viable Product), termine coniato per la prima volta nel 2001 da Frank Robinson (SyncDev, 2016).

## A lato

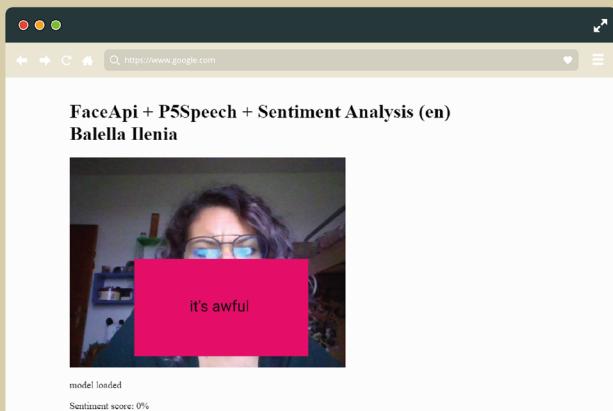
foto in sequenza dei principali prototipi



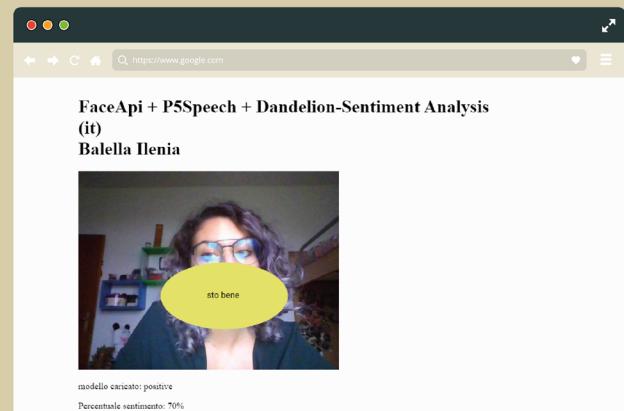
1. Primo prototipo, utilizzo di p5.speech



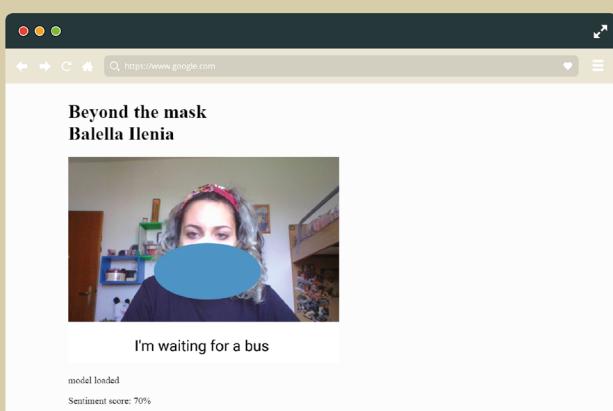
2. Unione di p5.speech con FaceAPI



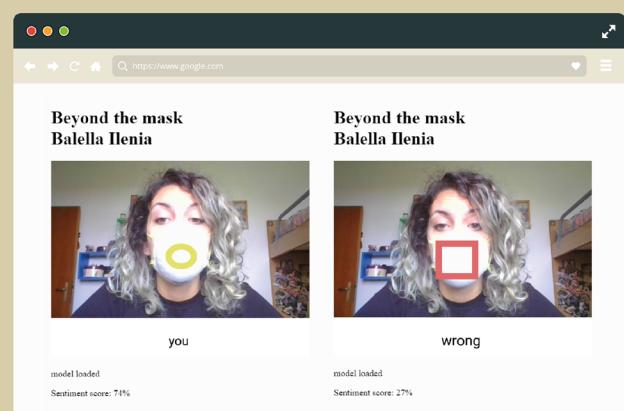
3. Aggiunta del Sentiment Analysis al prototipo



4. Utilizzo della Sentiment Analysis di Dandelion



5. Spostamento del testo come sottotitolo



6. Ottimizzazione della comunicazione non verbale

che parla, c'è la possibilità di mantenere un contatto visivo con l'interlocutore.

## Progetto | come funziona

Le parole non bastano a trasmettere globalmente quel che si vuole comunicare agli altri. Degli studi hanno dimostrato che la comunicazione non verbale ha un'incidenza del 55% sulla trasmissione del messaggio (Mehrabian, 1972). Perciò, per veicolare le informazioni non verbali celate dalla mascherina, beyond the mask si utilizza colori e forme di oggetti, generando pathos ed empatia, per scavalcare il muro inespressivo generato dalla mascherina. Utilizzando la telecamera di un dispositivo mobile si inquadra il volto dell'oratore. L'algoritmo di *FaceApi*<sup>[17]</sup> permette di inquadrare e riconoscere il volto, e la libreria di *P5 speech*<sup>[18]</sup> registra e trascrive il testo. Mentre l'interlocutore parla, si attiva la *Sentiment Analysis*<sup>[19]</sup>, che cambia la forma al riquadro sul volto, in base alla semantica del testo, rispettivamente al range positivo e negativo<sup>[20]</sup>.

A determinare il cambiamento del colore, è l'espressione facciale, con l'emozione che verrà rilevata dalla versione completa di *FaceApi*, allenata sul volto coperto dalla mascherina. La percezione del colore infatti, influenza in parte il comportamento umano. L'effetto che un colore può avere nella nostra psiche è molto personale ed è influenzato anche dalla cultura e società in cui si vive (Goethe, 1810). Ci sono stati diversi studi sulla teoria del colore, secondo i quali, ad esempio, i colori venivano abbinati a quattro caratteri dell'uomo (collerico, flemmatico, sanguigno e melanconico; Goethe e Schiller 1815 "rosa dei temperamenti). Goethe indica infatti il giallo come il colore della luce, quindi sereno ed eccitante, il blu è invece associato al "comune" (Goethe, 1810). Perciò è difficile determinare un colore in assoluto che possa essere associato ad una emozione. Per questo sono stati scelti tre colori, giallo (positivo), il blu (neutro) ed il rosso (negativo). Ma esistono 7 emozioni primarie: paura, tristezza, gioia, disprezzo, disgusto, sorpresa, rabbia (Ekman, 1972)<sup>[21]</sup>. Quindi sarebbe interessante lasciare la possibilità all'utente di settare i

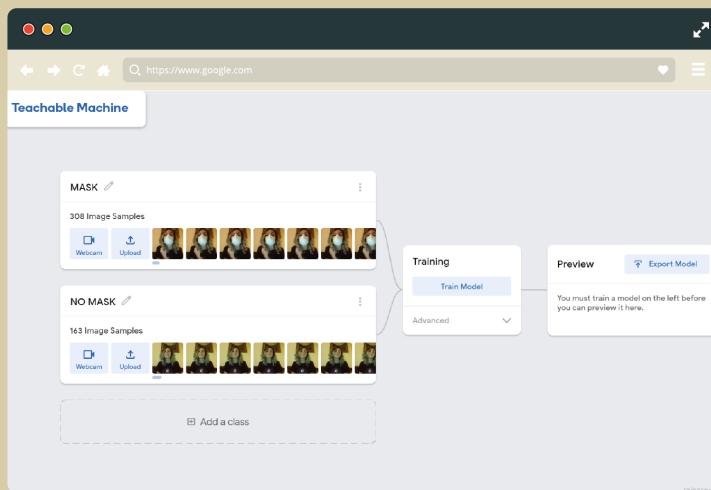
[17] face-api.js è un modulo javascript, che implementa diverse reti neurali per rilevare un volto, e riconoscere i suoi punti nello spazio.

[18] p5.speech è una estensione di p5 che aggiunge funzionalità di Speech Synthesis e Speech Recognition. Il sistema può ascoltare ed elaborare la voce di un interlocutore, ed è in grado anche di trasformare un testo scritto in una sintesi vocale.

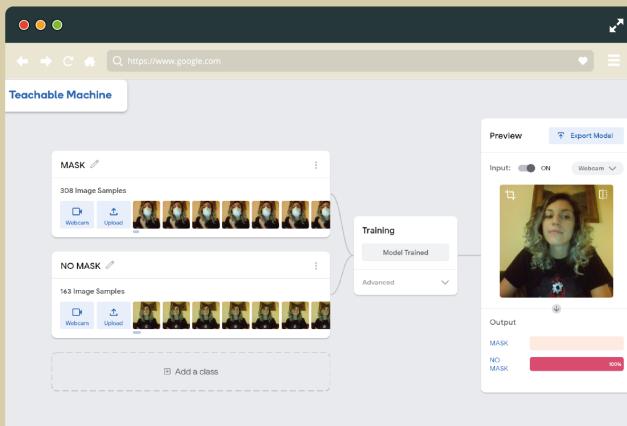
[19] La Sentiment Analysis è un sistema allenato ad analizzare un testo per identificare e classificare le informazioni presenti in esso. Il modello utilizzato da ml5 è allenato sulle recensioni IMDB.

[20] range che va da 0 a 100 (da 0 a 33 negativo, da 34 a 66 neutro, da 67 a 100 positivo)

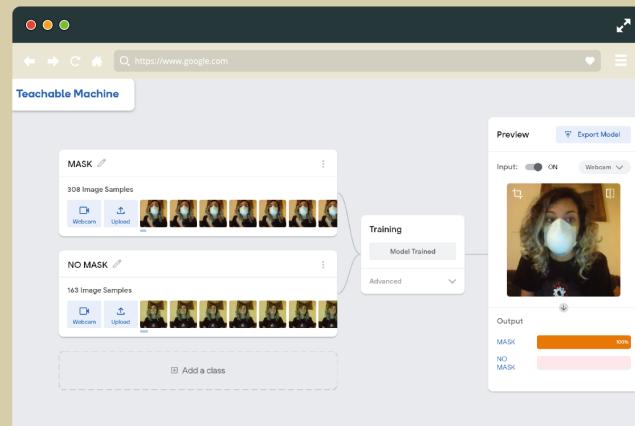
a destra  
allenamento di un modello che riconosce la mascherina su Teachable Machine di Google



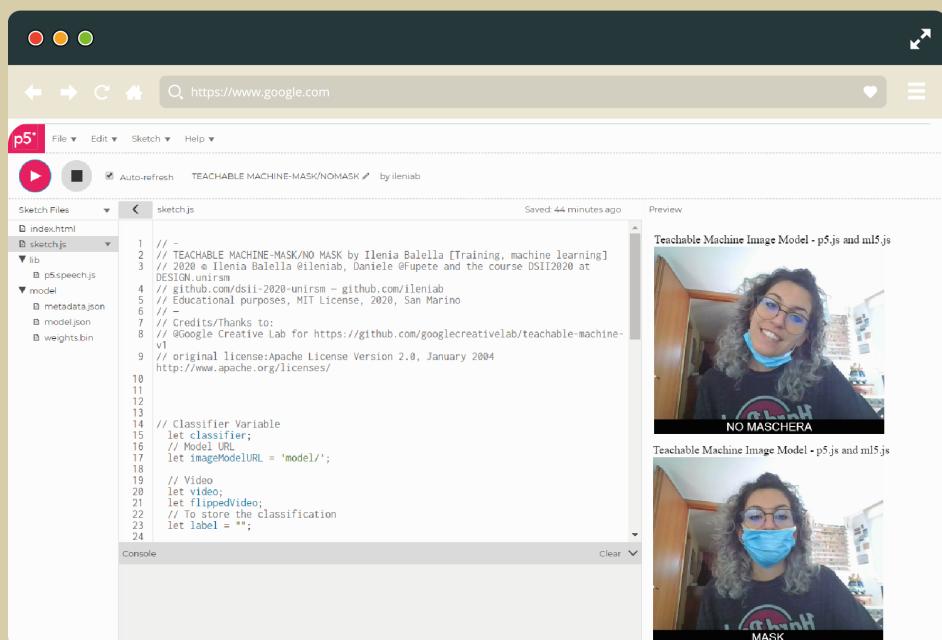
## 1. Preparare il proprio modello e allenarlo su Google Teachable Machine



## 2. Verificare il risultato



## 3. Verificare il risultato



## 4. Caricare il modello sull'editor di p5

colori abbinati alle emozioni rilevate dall'algoritmo. Infine il tono e l'intensità della voce potrà variare la dimensione del font. Esiste un progetto chiamato *Empath Web API*<sup>[22]</sup>, che utilizza il linguaggio python e riesce a determinare l'emozione in base al tono e alla frequenza della voce, come rabbia, gioia e tristezza. Per il momento queste tecnologie citate sono state quelle risultate funzionali nella fase di ricerca e prototipazione. Per quanto riguarda il sistema speech-to-text saranno da individuare dei framework e librerie simili ma più performanti rispetto a P5 speech. Così come il sistema di Face Recognition che dovrà essere allenato a riconoscere le emozioni dei volti con la mascherina.

## Progetto | prototipi

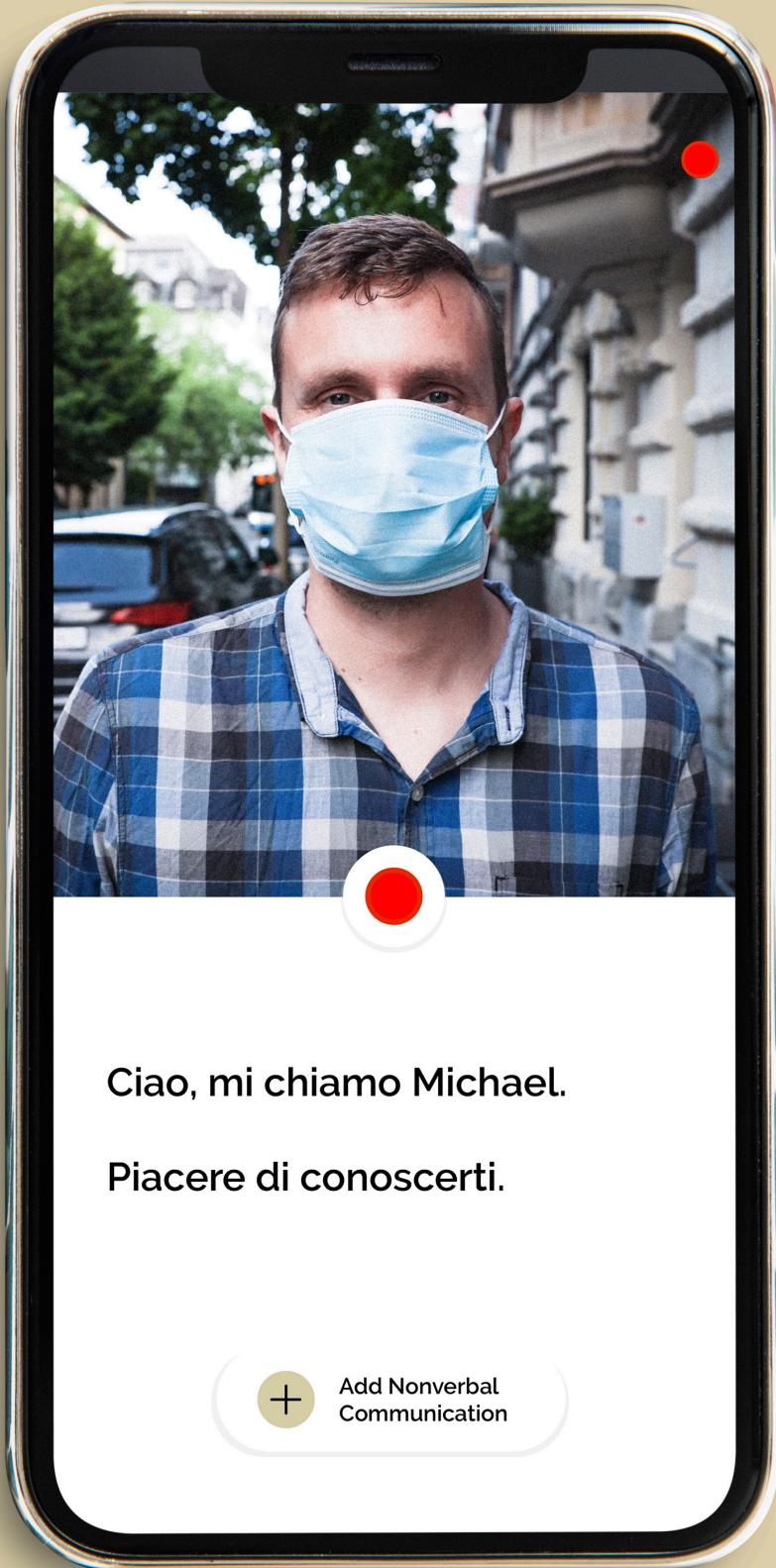
Il prototipo è composto da quattro moduli autonomi tra loro collegati. Inizialmente è stata utilizzata la libreria di p5 speech, che ascolta e registra la voce elaborandola in una stringa di testo, per trascrivere il discorso sulla canvas. Per il passo successivo è stato usato FaceApi di ml5, che rileva le coordinate del volto, per raffigurare un riquadro sul volto. In seguito, adoperando l'algoritmo di Sentiment Analysis, è stato analizzato il testo registrato. La Sentiment Analysis di ml5 indica una scala di predizione che va da 0 a 100, dal quale sono stati determinati 3 parametri di valore positivo, neutro e negativo. È stata sperimentata anche la Emotion recognition, con il face tracking per esplorare le capacità dell'algoritmo di rilevare le emozioni sul un volto che indossa la mascherina. Purtroppo le informazioni non sono sufficienti per una buona analisi, infatti riesce a rilevare con molta fatica due emozioni, la rabbia e la sorpresa. Infine è stato allenato, con il Teachable machine di Google<sup>[23]</sup>, un algoritmo in grado di riconoscere se si indossa una mascherina, per il momento il modello è allenato sulla mia figura, e sarebbe da implementare. Inizialmente era previsto di utilizzare il riquadro sul volto come un contenitore del testo, ma dopo diverse prove, per problemi di leggibilità, è stato spostato in basso, come un tradizionale sottotitolo.

[21] Alcune emozioni sono universali. Paul Ekman nel 1967 si recò sull'isola di Papua Nuova Guinea, per studiare una tribù indigena che non fosse venuta in contatto con nessuna forma di cultura occidentalizzata: i Fore. Scoprì che alcune espressioni facciali legate a precise emozioni, risultavano essere le stesse dei soggetti occidentali.

[22] <https://webempath.net/lp-eng/>

[23] Teachable Machine è uno strumento web che rende facile e veloce la creazione di modelli di machine learning

**a destra**  
immagine mockup di Beyond the mask in un contesto d'uso



Il riquadro si modificherà in base alla comunicazione non verbale, cambiando colore e forma a seconda del sentimento analizzato e dall'emozione registrata. Nel prototipo per il momento ai sentimenti sono stati associati dei colori arbitrari, per il negativo, il colore rosso, solitamente collegato alla rabbia. Invece per il sentimento positivo il giallo, un colore solare e allegro. A metà tra i due un colore neutro, il blu, che trasmette calma e consente di focalizzare l'attenzione. Sono stati sviluppati due prototipi paralleli, uno che utilizza la libreria di *P5 Speech* in lingua inglese, e l'altro invece che utilizza la lingua italiana. Per la versione italiana è stata utilizzata la *Sentiment Analysis* di Dandelion<sup>[24]</sup>. Purtroppo, la demo di Dandelion, utilizzata per lo sviluppo del prototipo, non è molto affidabile. Non comprende esattamente il significato delle frasi, e il suo risultato è quasi sempre neutro. L'alternativa più performante che si è portata avanti è quindi la versione in lingua inglese.

### Progetto | che cosa significa

Significa abbattere un muro comunicativo che esclude tutte quelle persone che hanno difficoltà uditive. Le persone sordi infatti, comunicano e comprendono il mondo in modo diverso. Avere un mezzo comunicativo, sempre disponibile nella loro quotidianità, aiuta l'utente con ridotta capacità uditiva, ad essere incluso all'interno della società senza sentirsi isolato, in un contesto alterato a causa dell'attuale emergenza.

### Progetto | sviluppi futuri

La sottotitolazione è disponibile per video inserzioni, film e serie tv, oppure sulla piattaforma di Youtube con la generazione automatica dei sottotitoli. Ma è fruibile in qualsiasi piattaforma? Quante persone sfruttano le storie o le dirette di Instagram per esprimere i propri pensieri ed esternare la propria identità sociale? Inoltre durante il periodo del lockdown c'è stato un incremento nell'utilizzo utilizzo dei social come strumento di propaganda ed intrattenimento. Tuttavia, non vi è la possibilità di inserire una sottotitolazione automatica (benché in molti video il volto dell'interlocutore è ripreso, ed è visibile il labiale).

[24] La tecnologia di Dandelion API permette di analizzare in real time la semantica di un testo. Indicando se l'analisi del sentimento esprime un'opinione positiva, negativa o neutra.

a destra  
immagine mockup di Beyond the mask in un contesto d'uso



Ma cosa succede quando il labiale non è inquadrato? Inoltre, se non si ha la possibilità di attivare l'audio perché ci si trova in un luogo pubblico, come si può seguire una diretta? A seguito di questa riflessione si potrebbe pensare a uno sviluppo futuro plausibile. Implementare un filtro instagram che integri in automatico la sottotitolazione nei video sfruttando librerie e/o framework che svolgono la funzione che in fase di prototipazione ha svolto p5.speech. Aumentando la probabilità di raggiungere un pubblico più ampio, includendo anche persone con disabilità uditive.

Cosa succederebbe se si ribaltassero i ruoli, ovvero se si potesse tradurre in parole la lingua dei segni? L'applicazione potrebbe aiutare entrambe le parti a comunicare ed interagire nella maniera più completa. Si dovrà addestrare un algoritmo che riconosca il linguaggio dei segni, creare dei modelli su cui addestrare la PoseNet<sup>[25]</sup>.

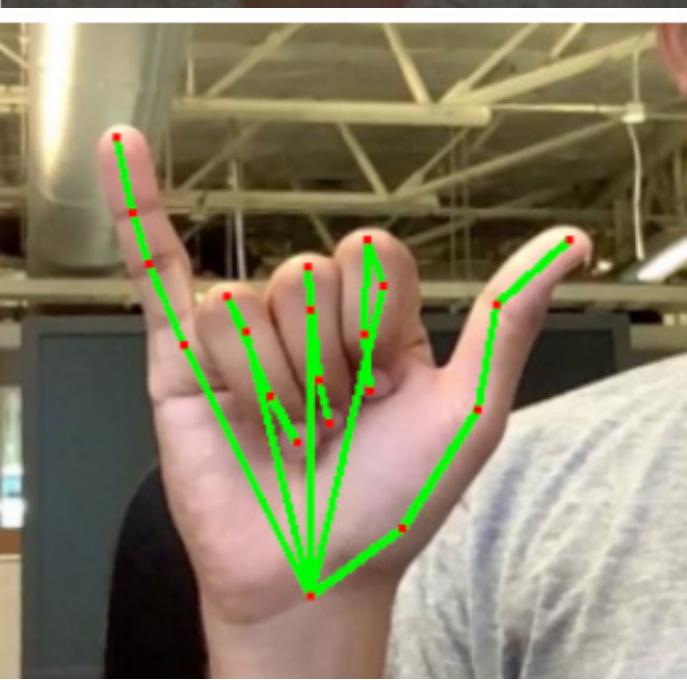
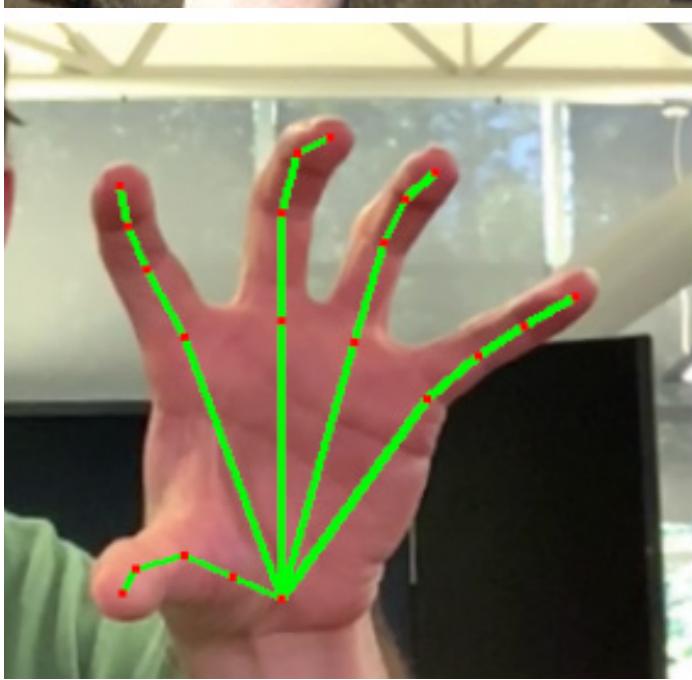
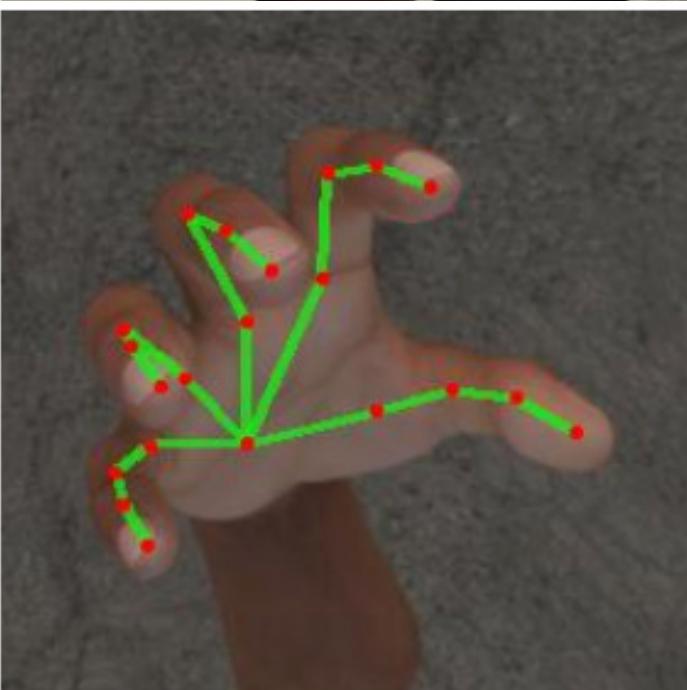
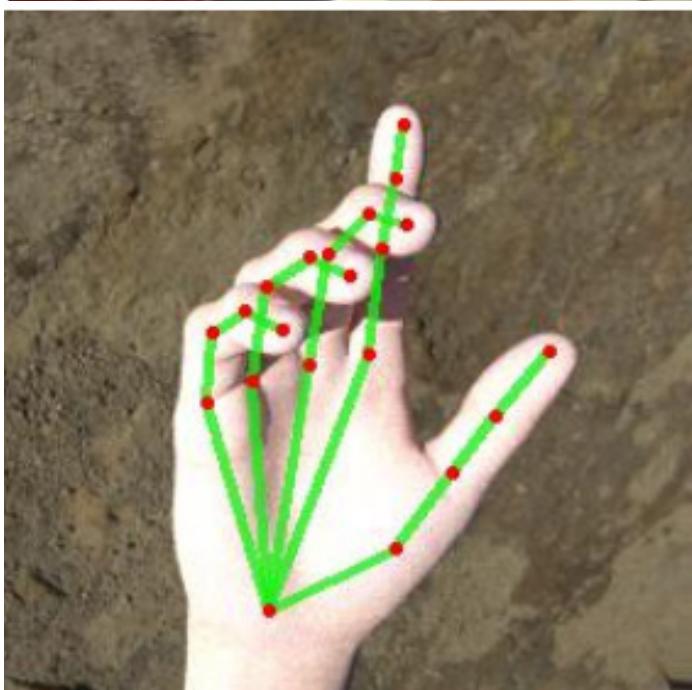
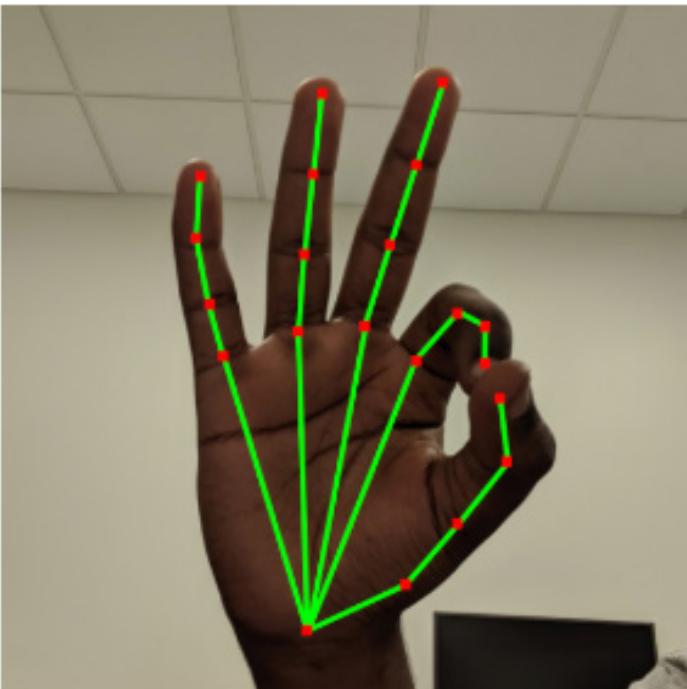
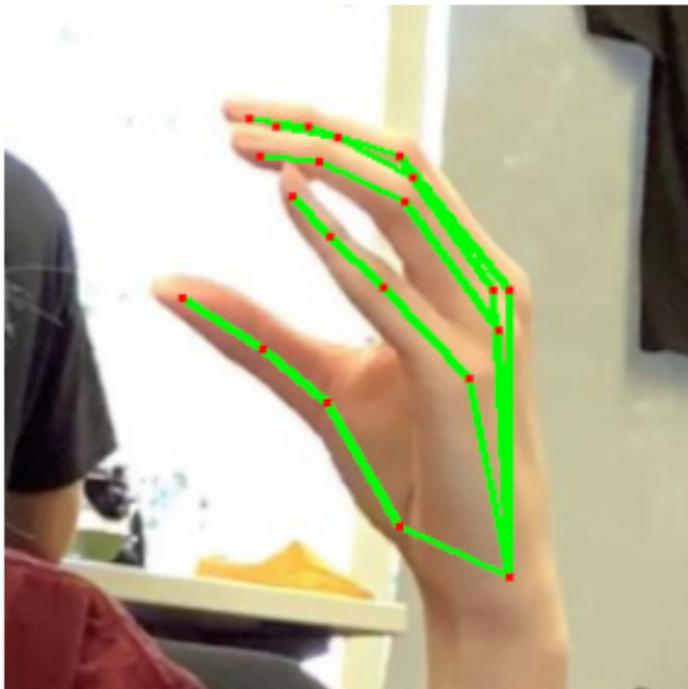
Il laboratorio di intelligenza artificiale di Google ha creato un algoritmo di hand-tracking che potrebbe portare al riconoscimento del linguaggio dei segni<sup>[26]</sup>. Si tratta di sistemi di apprendimento automatico per produrre, in tempo reale, una mappa altamente accurata della mano e di tutte le sue dita, utilizzando una telecamera. I ricercatori hanno ridotto le quantità di dati da elaborare<sup>[27]</sup> per ottenere un algoritmo rapido e preciso.

[25] PoseNet è un modello di apprendimento automatico che consente la stima della posa umana in tempo reale.

[26] <https://techcrunch.com/2019/08/19/this-hand-tracking-algorithm-could-lead-to-sign-language-recognition/>

[27] Hanno inizialmente creato il sistema che trova il palmo, che è quadrato, una forma semplice che l'algoritmo può gestire. Poi un algoritmo separato guarda l'immagine e assegna 21 coordinate ad essa, coordinandosi approssimativamente a nocche e punte delle dita.

**a destra**  
immagini dell'algoritmo di hand-tracking di Google, 2019



## **Bibliografia.**

Ekman, P. (1972). *Universals and Cultural Differences in Facial Expressions of Emotions*. In Cole, J. (Ed.), Nebraska Symposium on Motivation (pp. 207-282).

Goethe, J. W. (1810). *La teoria dei colori*. Tübingen: Cotta.

Mehrabian, A. (1972). *Nonverbal Communication*. New Jersey: Transaction Publishers.

Sacks, O. (2014). *Vedere voci: Un viaggio nel mondo dei sordi*. Milano: Adelphi.

Russo, C. T., & Volterra, V. (2007), *Le lingue dei segni. Storia e semiotica*. Roma: Carocci.

## **Sitografia.**

Abiliaproteggere. (2020). *Le mascherine trasparenti per una comunicazione inclusiva*.  
<https://www.abiliaproteggere.net>

Chuck, A. M. (2019). *Can Google's Translatotron Actually Help the Accuracy of AI Translation?*  
<https://www.business2community.com>

Coldewey, D. (2019). *TechCrunch is now a part of Verizon Media*. <https://techcrunch.com>

Contributori di Wikipedia. (2020). *Comunità sorda*.  
<https://it.wikipedia.org>

Cutecircuit. (2020). *The SoundShirt*.  
<https://cutecircuit.com>

Enciclopedia Treccani, n.d. *Sordità nell'Enciclopedia Treccani*. <http://www.treccani.it/enciclopedia/sordita>

Giglioli, D. (2008). *Una panoramica della situazione dei sordi italiani in generale e della lingua dei segni italiana in particolare*. <http://www.unapeda.asso.fr>

Glaiel, T. (2020). *How to make a voice activated LED facemask (DIY Guide)*. <https://medium.com>

Glaiel, T. (2020). *TylerGlaiel/voicemask*.  
<https://github.com>

Jia, Y., & Weiss, R. (2019). *Introducing Translatotron: An End-to-End Speech-to-Speech Translation Model*. <https://ai.googleblog.com>

McMullan, T. (2019). *This Outfit Can Help Deaf People Feel Music*. <https://onezero.medium.com>

Montinar, A. S. (2020). *Final Project – Slime Language*. <https://courses.ideate.cmu.edu>

Sbarbati, S. (2015). *Toyota Sensitive Concert: quella volta che ho “toccato” la musica*. <https://www.frizzifrizzi.it>

SyncDev. (2016). *Minimum Viable Product | SyncDev*.  
<http://www.syncdev.com>

<https://www.yourtypeface.com/>

<https://www.facebook.com/business/help/1675722002698686?id=603833089963720>

<https://webempath.net/lp-eng/>