



UNIVERSITÀ DEGLI STUDI DI SALERNO

# Mediapipe per la facial expression prediction

Laboratorio di  
FONDAMENTI DI VISIONE ARTIFICIALE E  
BIOMETRIA

## Emozioni attraverso espressioni



Paul Ekman ha elaborato per primo un modello scientifico per interpretare le emozioni correlate alle espressioni facciali.

Durante i suoi studi ha rilevato che i movimenti del volto correlati ad alcune emozioni sono identici per tutti gli individui, in qualunque parte del mondo essi vivano.

Ekman ipotizza pertanto che tali atteggiamenti/movimenti siano legati a fattori biologici e che in quanto tali non sono oggetto di attività di apprendimento.

Questa biometria quindi è di tipo *comportamentale*.

## Classificare le espressioni

Le espressioni rilevabili allo stato dell'arte sono 8: Neutrale, Disgusto, Sorpresa, Tristezza, Rabbia, Paura, Gioia e Disprezzo.



La classificazione dell'emozione così esplicitata può avvenire con metodi di deep learning, machine learning o di studio di particolari features.

Più l'espressione è palese e distinta dalle altre, più semplice sarà classificarla.

Cosa succede quando l'espressione è poco palese?



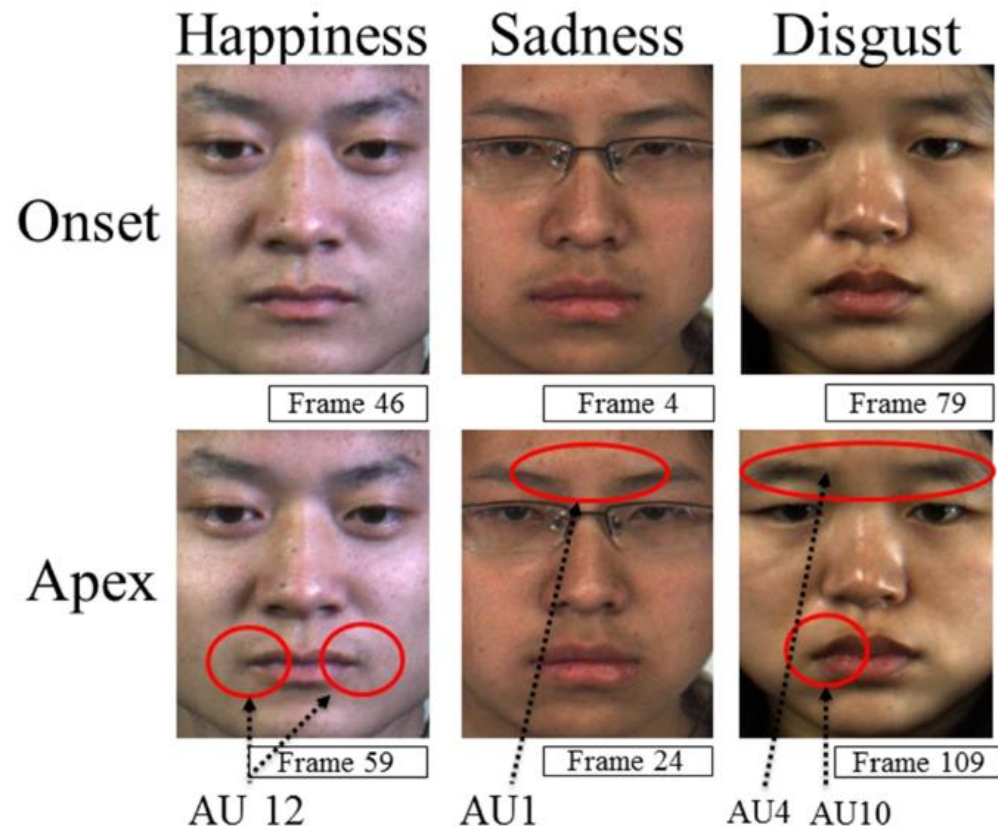
## Classificare le micro-espressioni

Le microespressioni sono caratterizzate da breve durata e bassa intensità, quindi, gli sforzi per addestrare gli esseri umani a riconoscerle hanno portato a prestazioni molto basse.

Tuttavia le micro-espressioni sono anche le più difficili da nascondere o da simulare da parte del soggetto, risultando così essere anche le più spontanee.

Il gap tra la scarsa abilità umana nel riconoscerle e l'interessante impiego delle micro-espressioni negli ambiti relativi alla sicurezza, ha portato ad un crescente impiego delle tecniche di machine learning e deep learning in questo settore.

Ma cosa c'è tra le micro e le macro espressioni?



## Predire le espressioni

Allo stato dell'arte esistono sia algoritmi in grado di classificare una macro-espressione (più semplici) che una micro-espressione (più complessi).

A metà strada tra questi due si può prevedere l'emozione che, frame dopo frame, sta andando a configurarsi.



Neutrale



Gioia  
(micro-espressione)

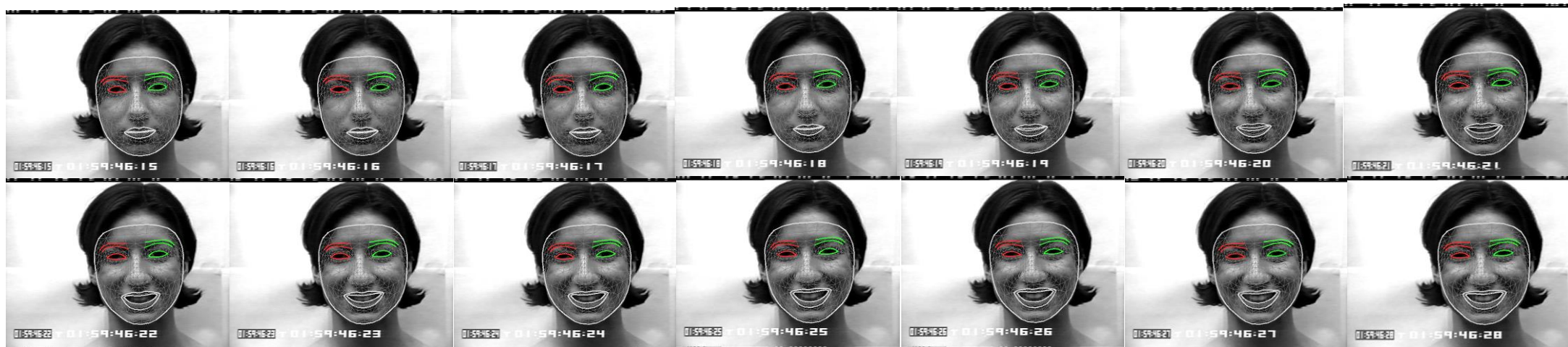


Gioia  
(macro-espressione)

**L'obiettivo di questo progetto è quello di fornire uno stato di avanzamento di una espressione sotto forma di percentuale in maniera automatizzata, per le 8 espressioni proposte, a partire da una neutrale.**

## Predire le espressioni - features

Le features da utilizzare per predire la percentuale di espressione, sono i landmark ottenuti mediante mediapipe, in particolare, le distanze tra i landmark in frame consecutivi.



Due sono gli array di distanze da tenere in considerazione:

- L'array di distanze dal frame precedente
- L'array di distanze dal frame iniziale

Il primo ci fornisce la variazione parziale, il secondo quella globale.

## Predire le espressioni - Dataset

Il Dataset che utilizzeremo è molto noto in letteratura: Cohn-Kanade Expression Dataset (CK+)

Ogni video mostra un passaggio facciale dall'espressione neutra a un'espressione di picco mirata, registrato a 30 fotogrammi al secondo (FPS) con una risoluzione di 640x490 o 640x480 pixel. Di questi video, 327 sono etichettati con una delle sette classi di espressione: rabbia, disprezzo, disgusto, paura, felicità, tristezza e sorpresa.

Ogni soggetto è in una cartella, a sua volta, ogni video è in una sottocartella di cui è disponibile l'etichettatura della emozione finale.

### ATTENZIONE:

- Non ci sono tutte le emozioni per tutti i soggetti
- Non vanno utilizzati gli stessi soggetti per fase di training e fase di testing





## Predire le espressioni - Passaggi

I passaggi del metodo da sviluppare sono i seguenti:

### ***Preparazione dataset di caratteristiche***

- Estrarre per ogni immagine del dataset i 468 landmark mediante mediapipe, collezionare nello stesso file i landmark delle stesse sequenze video.
- Costruire gli array delle distanze dei landmark: per ogni sequenza va costruito l'array di distanze locali (ogni frame dal precedente, il primo avrà distanze 0) e l'array di distanze globali (ogni frame dal primo, il primo avrà distanze 0).
- Associare ad ogni array l'etichettatura relativa all'espressione.
- A seconda di come si intende risolvere il problema, si potrebbe associare ad ogni riga una percentuale o una classe di percentuale dell'espressione.

### ***Predizione dell'espressione***

Questa fase della risoluzione può essere affrontata in due modi:

- Classificazione: costruire delle classi di percentuale (10-30; 30-50; 50-70; 70-100) per le espressioni. Il problema diventa quindi una classificazione multiclasse, dove le percentuali più basse individuano la micro-espressione e le più alte la macro-espressione.
- Regressione: il valore percentuale è ottenuto mediante regressione dei valori esistenti. Avendo un numero diverso di frame in ogni cartella del dataset, non avremo tutti i possibili valori percentuali.



## Predire le espressioni – Final remarks

- In questo progetto la decisione delle features da utilizzare e la loro analisi riveste un ruolo di fondamentale importanza. Studiare la significatività delle distanze locali e globali, come esse variano inter-espressione e intra-espressione è essa stessa un progetto (eventualmente per un gruppo più piccolo di studenti).
- La scelta di come operare in fase iniziale è a carico dello studente. Si può provare prima a cercare l'espressione e poi categorizzare la percentuale o fare tutto in un unico passaggio. La decisione ottimale va presa a valle delle considerazioni fatte nell'analisi delle features.
- Questo è un progetto **ex-novo**, l'unico materiale fornito consiste nel dataset e il codice mediapipe per estrarre i singoli landmark del volto e salvarli. Per questo motivo, l'interazione con il tutor ad ogni lezione è essenziale per discutere lo stato di avanzamento del progetto, presentare l'approccio scelto man mano e calibrare l'effort del progetto in base al numero di componenti e agli esperimenti condotti per avanzare.
- Il materiale da produrre è lo stesso già descritto per i precedenti progetti (codice commentato python o jupyter o colab, short paper) e da inviare non più tardi di 5 giorni antecedenti la data di presentazione del progetto (coincidente con la data d'esame).