



UNIVERSITÀ
DEGLI STUDI
FIRENZE

Scuola di Scienze Matematiche, Fisiche e Naturali
Corso di Laurea Magistrale in Informatica

Tesi di Laurea

RICONOSCIMENTO DI AZIONI UMANE
USANDO TECNICHE DI APPRENDIMENTO
PROFONDO PER LA STIMA DELLA POSA

HUMAN ACTION RECOGNITION USING
DEEP LEARNING TECHNIQUES FOR POSE
ESTIMATION

ANDREA MOSCATELLI

Relatore: *Marco Bertini*

Anno Accademico 2018-2019

Andrea Moscatelli: *Riconoscimento di azioni umane usando tecniche di apprendimento profondo per la stima della posa*, Corso di Laurea Magistrale in Informatica, © Anno Accademico 2018-2019

ABSTRACT

I recenti progressi nel campo della visione artificiale hanno permesso alla comunità scientifica di spostarsi verso problemi ancora più articolati rispetto a quelli classici ed il riconoscimento di azioni corporee tramite l'analisi della posa umana sta attraendo recentemente una notevole attenzione. Il suo successo è senzaltro dovuto non solo agli ottimi risultati ottenuti, ma anche alla sua efficiente semplificazione della struttura umana riducendo di fatto i costi computazionali e le risorse necessarie allo stoccaggio dati.

In questo lavoro di tesi ci siamo dedicati al riconoscimento e alla classificazione di azioni umane tramite tecniche di apprendimento profondo per la stima della posa. A tale scopo abbiamo deciso di ideare un algoritmo che non avesse bisogno di informazioni iniziali complesse, come ad esempio la posizione dei giunti dei soggetti inquadrati, ma che attraverso l'uso dei soli video RGB fosse in grado di estrapolare tutte le informazioni necessarie.

Al fine di ottenere il miglior algoritmo abbiamo eseguito un serie di esperimenti strutturati secondo un procedimento ben preciso, che permettesse l'esplorazione rapida di ogni tecnica ideata affinando progressivamente i risultati per quelle più promettenti.

Quello che abbiamo ottenuto alla fine del processo sono due algoritmi con una discreta capacità di classificazione delle azioni umane e per la loro semplicità anche un'elevata portabilità.

Questi due algoritmi, pur facendo uso dei soli video RGB, hanno prestazioni comparabili a molti altri lavori scientifici facenti uso dalla posizione dei giunti fornita da dataset stesso, inoltre per la loro semplicità sono facilmente migliorabili.