

Riconoscere un volto

Capacità di individuare volti da qualunque angolatura, persino quando sono parzialmente occultati, è stata una capacità unicamente umana, almeno fino a poco tempo fa.

The Physics arXiv Blog

Nel 2001, due scienziati informatici, Paul Viola e Michael Jones, hanno dato il via a una rivoluzione nel campo del riconoscimento facciale computerizzato. Dopo anni di ristagno, questi ricercatori hanno sviluppato un algoritmo in grado di individuare i volti nelle immagini in tempo reale. Di fatto, il cosiddetto algoritmo Viola-Jones era talmente veloce e semplice che subito è diventato parte degli standard per le fotocamere compatte.

Il trucco consisteva nell'ignorare il problema assai più complesso del riconoscimento facciale, concentrandosi unicamente sul rilevamento dei volti fotografati frontalmente, ignorando qualunque altro volto preso da angolature differenti. Dati questi limiti, ci si rese conto che il ponte del naso forma solitamente una linea verticale più luminosa rispetto alle orbite oculari ai lati e che gli occhi restano spesso in ombra, formando quindi una fascia orizzontale più scura.

Viola e Jones hanno quindi creato un algoritmo in grado di ricercare nelle immagini le bande verticali luminose che potrebbero corrispondere con un naso, le bande scure orizzontali che potrebbero corrispondere agli occhi, oltre ad altre caratteristiche generiche associate ai volti.

Presa singolarmente, nessuna di queste caratteristiche allude particolarmente a un volto. Quando però questi elementi vengono rilevati a cascata uno dopo l'altro, il risultato è una buona indicazione della presenza di un volto in un'immagine, da cui il nome del processo: "rilevatore a cascata". Tuttavia, l'algoritmo di Viola-Jones consentiva di rilevare e individuare i volti visti di fronte, mentre la stessa accuratezza non si poteva ottenere da qualunque altro angolo, il che limitava l'utilizzo del sistema per i motori di ricerca facciale.

È per questo motivo che Yahoo si è interessata al problema. Oggi, Sachin Farfede e Mohammad Saberian, degli Yahoo Labs in California, e Li-Jia, della vicina Università di Stanford, hanno reso noto un nuovo approccio che permette di rilevare volti ripresi d'an-

golo, anche quando sono parzialmente occultati. Secondo questi ricercatori, il nuovo approccio sarebbe più semplice rispetto ad altri, ma consentirebbe ugualmente prestazioni allo stato dell'arte.

Per realizzare il proprio modello Farfede e colleghi hanno utilizzato un approccio fondamentalmente differente, concentrandosi sui progressi conseguiti negli ultimi anni da una forma di apprendimento automatico conosciuta come rete neurale convoluzionale profonda. L'idea consiste nell'addestrare una rete neurale formata da molteplici strati utilizzando un vasto database di esempi, costituiti in questo caso da immagini di volti ripresi da angoli differenti.

Al fine di addestrare la loro rete neurale, Farfede ha creato un database di 200mila immagini, che include volti ripresi da diverse angolature e di profilo, oltre a 20 milioni di immagini senza volti. Il gruppo ha quindi addestrato la rete neurale con gruppi di 128 immagini per oltre 50mila cicli. Il risultato è un singolo algoritmo in grado di individuare con una precisione sorprendente volti ripresi da vari angoli, persino quando sono parzialmente occultati.

Il gruppo ha battezzato il metodo "Deep Dense Face Detector" e sostiene che possa

reggere il confronto con altri algoritmi: «Abbiamo confrontato il nostro metodo con altri metodi basati sull'apprendimento approfondito, dimostrando che il nostro porta a risultati più rapidi e precisi».

Oltretutto, l'algoritmo è particolarmente efficace nell'individuare volti capovolti, un aspetto che altri metodi non hanno ancora perfezionato. Secondo il gruppo, l'algoritmo potrebbe venire migliorato ulteriormente utilizzando database che includano più volti a testa in giù. «In futuro utilizzeremo strategie di campionamento migliori e tecniche di potenziamento dei dati più sofisticate per migliorare le prestazioni del nostro metodo nel rilevare volti coperti o ruotati».

Si tratta di un lavoro interessante, che dimostra la rapidità con cui si sta evolvendo il riconoscimento facciale computerizzato. La tecnica della rete neurale convoluzionale profonda ha appena qualche anno di età, ma ha già portato a importanti progressi nel riconoscimento di oggetti e volti.

La grande promessa di questo genere di algoritmo sta nella ricerca delle immagini. Al momento, si tratta di una ricerca diretta alle immagini scattate in un luogo o in un momento specifico. È però difficile riuscire a trovare immagini di persone specifiche, ma non è difficile immaginare che presto una funzione del genere diventerà una comune. Quel giorno, il mondo sarà un luogo molto più piccolo. Non solo le fotografie, ma anche le innumerevoli immagini, fotografiche e video, che sono state raccolte nel corso della storia, potranno venire ritrovate. In un modo o nell'altro, questa tecnologia acquisirà un potere enorme. ■

