

# Deep learning per percezione, azione e reazione intelligente

Stefano Ghidoni, Elisa Tosello, Matteo Terreran, Giorgio Nicola

DEI – Dipartimento di Ingegneria dell’Informazione, Università degli Studi di Padova

stefano.ghidoni@unipd.it, elisa.tosello@unipd.it,  
matteo.terreran@phd.unipd.it, giorgio.nicola@phd.unipd.it

## Abstract

Questo contributo presenta tre temi di ricerca sinergici sviluppati nel gruppo di ricerca, che spaziano dalla percezione per la robotica al task e motion planning, fino ad arrivare all’interazione intelligente con l’uomo, con il comune denominatore dell’apprendimento automatico. Il contributo mostra come l’intelligenza artificiale sia cruciale per permettere ai sistemi robotici di affrontare scenari reali, che hanno un certo livello di non prevedibilità. I sistemi sviluppati sono stati applicati sia a scenari controllati in laboratorio che a sistemi industriali.

## 1 Introduzione

Il mondo della robotica e quello della percezione artificiale sono sempre più legati dal collante dell’intelligenza artificiale. Si consideri uno scenario in cui ad un robot è assegnato un compito di manipolazione in un ambiente casalingo o industriale, come preparare la tavola o assemblare pezzi in una catena di montaggio. Per poter compiere tali compiti, il robot deve segmentare la scena che lo circonda e riconoscere gli oggetti da manipolare. Note le proprietà intrinseche di tali oggetti (massa, geometria, informazioni semantiche come l’utilità dell’oggetto e la sua afferrabilità) e date le azioni che il robot sa eseguire, serve un sistema intelligente capace di generare una sequenza di azioni e movimenti che permettano di ottenere l’obiettivo desiderato. Inoltre, un robot che agisce nel mondo reale, tipicamente non deterministico, deve essere in grado di gestire situazioni impreviste, il che richiede una elevata dose di intelligenza (artificiale!).

In questo contesto, il nostro gruppo porta avanti ricerche su molteplici fronti, che presentano numerose sinergie e punti di contatto, mediante tecniche di apprendimento automatico, molto efficaci nell’affrontare scenari non deterministici. I temi di maggior interesse sono: i) visione artificiale e visione robotica; ii) task and motion planning; iii) cooperazione uomo-robot. Le ricerche portate avanti hanno non solo un forte interesse per gli aspetti di base della robotica, ma anche un forte richiamo di natura sperimentale, che abbraccia sia gli scenari della robotica di servizio che quelli della robotica industriale. In questo modo è possibile contaminare queste esperienze e questi mondi, mostrando come sia possibile

portare proficuamente nel mondo della produzione industriale la ricerca avanzata concepita in scenari più semplici e “da laboratorio”.

## 2 Segmentazione semantica

Tra i numerosi progetti di computer vision portati avanti, la segmentazione semantica ha un ruolo di primo piano. Questa consiste nella suddivisione dell’immagine in regioni che hanno un significato semantico preciso, e può essere visto come un task di classificazione per-pixel.

Il filone di ricerca sulla segmentazione semantica è partito dall’analisi dei tipici scenari della robotica di servizio. Il tema è stato dapprima affrontato con sensori attivi (telecamere 3D), che sono tipicamente usati per questo tipo di applicazioni. Sono stati inizialmente affrontati con successo [1] compiti di classificazione multi-classe (12 categorie oltre al background) in scenari dalla geometria complessa. L’approccio sviluppato è basato sulle entangled forest, un classificatore alternativo alle reti deep che offre buone prestazioni per questo task di classificazione, senza richiedere un elevato costo computazionale.

La stessa tecnologia è stata successivamente utilizzata per risolvere task di classificazione più semplici (6 categorie) in scenari industriali. Tali scenari presentano infatti una minor varietà di oggetti presenti nell’ambiente, permettendo di ridurre il numero di categorie considerate. Questo è avvenuto nel contesto del progetto europeo Coroma<sup>1</sup>, che propone un caso d’uso in cui un robot deve essere in grado di navigare autonomamente in uno stabilimento, di avvicinarsi alle barche in produzione, e di effettuare una lavorazione. La segmentazione semantica fornisce una mappa dell’ambiente di produzione, come mostrato in figura 1.

La ricerca sulla segmentazione semantica sta ora proseguendo con la sintesi di sistemi basati su deep learning. Due percorsi paralleli sono attualmente allo studio: i) la possibilità di sfruttare reti deep come estrattori di nuove feature, da fornire ai classificatori già utilizzati (entangled forest) utilizzando un approccio tipo transfer learning, e ii) lo sviluppo di reti a due rami per l’analisi separata di dati 2D e 3D, fusi poi nel contesto della stessa rete. Questi sistemi saranno applica-

<sup>1</sup>Cognitive enhanced RObot for flexible MAnufacturing of metal and composite parts, <https://www.coroma-project.eu/>

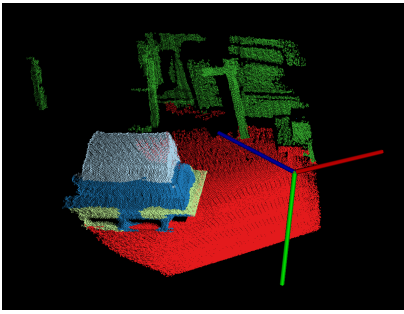


Figura 1: Esempio di segmentazione semantica per uno scenario industriale. È distinguibile l'elemento in produzione, in azzurro, e il suo supporto, in blu. Sono anche visibili il pavimento, in rosso, e il background, in verde.

ti all'analisi di scenari industriali, quali quelli considerati nel progetto europeo Eureka<sup>2</sup>

### 3 Task and motion planning

Nel contesto della manipolazione, task planning e motion planning sono due importanti compiti della robotica, che in passato sono stati affrontati separatamente. Più di recente, la proposta di un approccio unificato ha permesso di trovare soluzioni innovative a questi problemi, sfruttando le sinergie e le dipendenze tra i due compiti. La ricerca svolta in questo ambito ha come obiettivo intermedio quello di sviluppare una serie di sistemi di apprendimento automatico capaci di risolvere problemi di task planning e di motion planning separatamente. Tale obiettivo intermedio significa, in altre parole, l'applicazione dei più recenti sistemi di intelligenza artificiale a task classici della robotica. I due componenti fondamentali, in questo ambito, sono costituiti dai sistemi di Deep Reinforcement Learning (DRL) e da sistemi avanzati di apprendimento di sequenze, le Long Short-Term Memory (LSTM).

Allo stato attuale, nel nostro gruppo sono già stati messi a punto sistemi di DRL per l'apprendimento di task e motion di complessità media e medio-alta, quali: i) l'inseguimento di un oggetto (visibile nella sequenza di figura 2); ii) la battuta di una pallina da golf utilizzando una mazza fissata al braccio; iii) il riordino di oggetti in uno spazio di lavoro in base a caratteristiche visuali [2]; iv) la collocazione di oggetti in ripiani diversi in base alle loro caratteristiche visuali; v) il pick and place.

Il passaggio intermedio appena descritto è fondamentale per raggiungere l'obiettivo più ambizioso di questa ricerca: ottenere un approccio olistico a task e motion planning fondendo assieme le reti sviluppate precedentemente. La fusione prevede, a basso livello, la connessione delle reti precedentemente sviluppate, con una successiva fase di fine tuning per ottenere un'ottimizzazione reciproca delle reti coinvolte. Questo approccio è estendibile anche alla percezione, che sarà quindi inglobata a monte del sistema di task e motion planning.

<sup>2</sup>Enhanced human-robot cooperation in cabin assembly tasks, <http://www.cleansky-eureka.eu/>

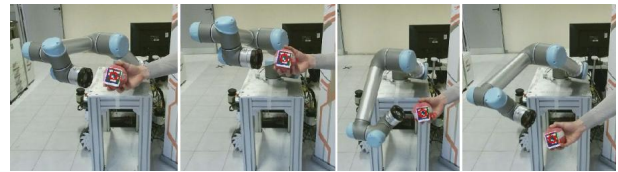


Figura 2: Esempio di inseguimento di un oggetto, ottenuto tramite Deep Reinforcement Learning.

Il percorso di ricerca illustrato fino ad ora è di forte interesse anche per le applicazioni industriali. La capacità di apprendere autonomamente compiti complessi, sia dal punto di vista delle sequenze di azioni che dei movimenti, è di fondamentale importanza per permettere l'utilizzo dei robot da parte delle persone che sono esperte di un determinato processo produttivo, ma non di robotica. L'introduzione di sistemi di intelligenza artificiale diventa quindi un elemento cruciale per poter ottenere soluzioni ottime specifiche per ogni singola applicazione, senza richiedere una fase di ottimizzazione effettuata manualmente.

### 4 Cooperazione uomo-robot

Nel campo della cooperazione uomo-robot, di per sé molto vasto, è stata condotta una ricerca molto specifica e di nicchia. Si tratta della stima delle forze applicate dal robot sugli oggetti manipolati senza l'uso di sensori ad hoc, ma basato sulle informazioni delle correnti nei motori. Un tale approccio necessita di una conoscenza profonda della struttura del robot, conoscenza difficilmente ottenibile per via diretta. Questo tema è quindi stato affrontato come un problema di stima, in collaborazione con il gruppo di controlli automatici del dipartimento. Tale problema risulta particolarmente complesso quando le velocità in gioco sono ridotte, poiché un elevato numero di effetti parassiti inficiano la stima. L'approccio seguito sfrutta un modello semiparametrico che si è dimostrato molto efficace sia in condizioni dinamiche che quasi statiche.

Gli scenari industriali in cui questo studio può essere applicato sono molteplici e di grande interesse. Un robot in grado di stimare le forze applicate senza utilizzare sensori ad hoc offre enormi vantaggi, poiché tali sensori sono molto costosi e fragili, quindi poco adatti alla manipolazione industriale.

### Riferimenti bibliografici

- [1] M. Antonello, D. Wolf, J. Prankl, S. Ghidoni, E. Mene-gatti, M. Vincze, "Multi-View 3D Entangled Forest for Semantic Segmentation and Mapping", in Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA), pp. 1855-1862, Brisbane, Australia.
- [2] F. Vendramin, E. Tosello, N. Castaman, S. Ghidoni, "Learning Robot Task Planning Primitives by means of Long Short-Term Memory Networks", in Combining Task and Motion Planning in the frame of Cloud Robotics, workshop di 2018 IEEE Simpar International Conference, Brisbane, Australia.