

I CUOCHI NON SANNO SCRIVERE LE RICETTE

C. Nuzzi⁽¹⁾, S. Pasinetti⁽¹⁾, F. Docchio⁽¹⁾, G. Sansoni⁽¹⁾

⁽¹⁾Dip. di Ingegneria Meccanica e Industriale, Università degli Studi di Brescia
mail autore di riferimento: c.nuzzi@unibs.it

1. INTRODUZIONE

Il processo creativo di una ricetta per un cuoco spesso è frutto di idee nate al momento, basate sulla sua esperienza e conoscenza personale. Questo comporta una certa difficoltà nel trasmettere la procedura di realizzazione di una certa ricetta a un altro individuo, che inevitabilmente la riprodurrà a suo modo. Se questo è vero già quando un cuoco junior osserva e riproduce le azioni di un cuoco esperto, diventa ancora più difficile riuscire a ricreare la ricetta originale quando la si legge descritta in modo classico, ovvero con la tipica separazione tra la lista degli ingredienti necessari e le varie fasi della preparazione. Questa problematica è ben chiara ai cuochi stessi, che faticano a trasmettere in modo intuitivo ai loro colleghi e/o sottoposti una procedura che sia il più possibile univoca e che porti alla riproduzione il più possibile simile sia del gusto che della presentazione finale del piatto.

Per ovviare a questo problema, il presente lavoro di ricerca ha come obiettivo la creazione e messa in opera di un sistema intelligente che, utilizzando diversi sensori, sia in grado di comprendere (i) le **azioni** che il cuoco compie man mano che procede con la preparazione, (ii) quali **ingredienti** usa per ogni sotto-fase e (iii) quali **strumenti** utilizza per ogni sotto-fase. In questo modo, sarà il sistema a ricondurre ogni azione analizzata ad una azione nota (ad esempio “mescolare” o “tagliare”) e, componendole in forma di flusso di lavoro, produrrà al termine una riscrittura completa della ricetta che sia di più chiaro e intuitivo utilizzo.

2. SPECIFICHE DEL PROGETTO

Come primo approccio al problema, essendo questo di notevole complessità, ci siamo concentrati sul corretto riconoscimento delle azioni. Per fare ciò, abbiamo posto i seguenti vincoli:

1. una sola ricetta complessa viene utilizzata per addestrare l’algoritmo, in modo da apprendere più azioni differenti e sotto-ricette che, in una fase finale di preparazione e impiattamento, verranno combinate nel prodotto finale. In accordo con i cuochi professionisti coinvolti nel progetto, è stata selezionata la ricetta della *Lasagna alla Bolognese*;
2. gli ingredienti previsti per la ricetta sono già presenti nelle corrette quantità e separati in appositi contenitori per alimenti pronti all’uso e posizionati su un supporto dedicato;
3. il riconoscimento degli ingredienti avviene mediante lettura dei codici QR posti al di sotto dei contenitori: quando il cuoco utilizza un ingrediente togliendolo dal supporto dedicato viene letta l’assenza del QR relativo;
4. non è previsto il riconoscimento degli strumenti, che verranno aggiunti manualmente in fase di post-processing.

3. PRIMO PROTOTIPO DI CUCINA SPERIMENTALE

Vista la natura dell’attività, è opportuno prevedere un set-up il più possibile multi-sensore allo scopo non solo di potenziare le capacità interpretative del sistema intelligente ma anche per sopperire a eventuali perdite di dati o acquisizione di dati non validi.

Un primo prototipo di cucina sperimentale è quello mostrato in Fig. 1a, nel quale sono presenti due Kinect v2 posizionate in modo da riprendere e coprire il più possibile l’area di lavoro del cuoco e una telecamera Basler RGB montata su un banchetto sperimentale (Fig. 1b) dedicato alla lettura dei codici QR. L’acquisizione dei dati avviene tramite due sistemi indipendenti connessi alla stessa rete locale ROS [1] secondo la procedura seguente:

1. Sul pc **master** (Ubuntu) sono connesse le due Kinect v2, che acquisiscono in modo indipendente immagini RGB e Depth. Per sincronizzare l’acquisizione dei due tipi di dato, ogni Kinect salva esclusivamente i frame RGB e Depth che presentano un tempo di acquisizione in un intorno di 33 Hz. I frame RGB e Depth sincronizzati vengono elaborati da una rete neurale che produce lo skeleton della persona presente nella scena, mediante l’uso del software presentato in [2] e riadattato per questo lavoro. I frame RGB e Depth e lo

skeleton ottenuto vengono inviati in un unico messaggio ROS e salvati sul pc master in *rosbag* distinte per ogni Kinect.

2. Sul pc *slave* (Windows) il software di acquisizione dei codici QR sviluppato in LabVIEW invia alla rete locale ROS un messaggio contenente tutti i codici QR letti in un certo istante di tempo. L'invio di questo messaggio avviene ogni volta che viene rilevato un cambiamento nella lettura precedente, ovvero quando viene tolto o aggiunto un ingrediente dal banchetto di prova. Il pc master salverà tutti i messaggi ricevuti in una apposita *rosbag*.

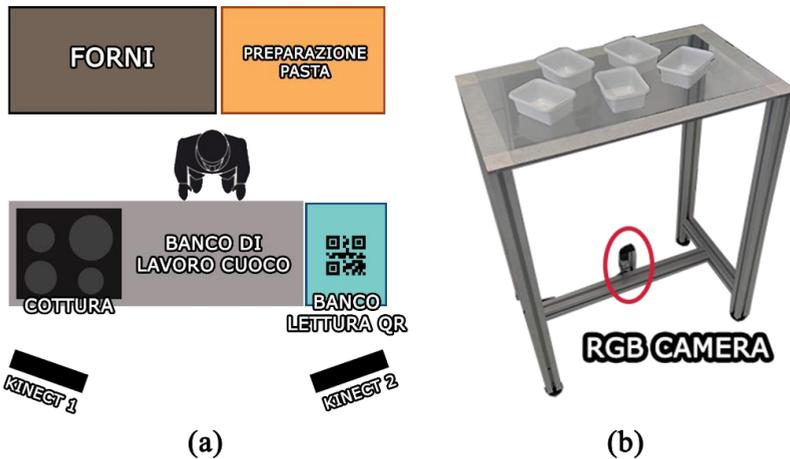


Fig. 1 (a) Schema della cucina sperimentale, (b) banchetto sperimentale adibito alla lettura dei codici QR posti sul retro dei contenitori ingrediente.

4. FUTURI SVILUPPI DEL PROGETTO

In seguito alle prove di acquisizione sul campo con lo staff di cuochi professionisti, i dati ottenuti dovranno essere accuratamente visionati e classificati, in modo da ottenere un dataset funzionale da utilizzare per l'addestramento di un algoritmo di intelligenza artificiale, che si prevede essere una variante di rete neurale ricorsiva basata su LSTM [1]. A seguito di una valutazione delle performance di quest'ultimo sui dati ottenuti, una serie di accorgimenti verranno adottati al fine di sviluppare un secondo prototipo e realizzare una seconda campagna di misure. Si prevede sin da ora l'introduzione di un metodo ad hoc per il riconoscimento automatico degli strumenti, il cui progetto è attualmente in fase iniziale.

RIFERIMENTI BIBLIOGRAFICI

- [1] M. Quigley, B. P. Gerkey, K. Conley, J. Faust, T. Foote, J. Leibs, E. Berger, R. Wheeler e A. Ng, «ROS: an open-source Robot Operating System,» in *ICRA Workshop on Open Source Software*, 2009.
- [2] C. Zimmermann, T. Welschehold, C. Dornhege, W. Burgard e T. Brox, «3D Human Pose Estimation in RGBD Images for Robotic Task Learning,» in *IEEE International Conference on Robotics and Automation (ICRA)*, 2018.
- [3] P. Wang, Li, Wanqing, P. Ogunbona, J. Wan e S. Escalera, «RGB-D-based Human Motion Recognition with Deep Learning: A Survey,» *Computer Vision and Image Understanding*, pp. 118-139, 2018.