

Deep Learning efficiente e interazione gestuale per l'industria

Andrea Bragagnolo^{1,2}, Attilio Fiandrotti¹, Cristina Gena¹, Marco Grangetto¹, Fabiana Vernerio¹, and Andrea Basso²

¹Dipartimento di Informatica, Università degli Studi di Torino

²Synesthesia s.r.l., Corso Dante 118, 10126, Torino

[nome.cognome]@unito.it

Abstract

Questo documento presenta due temi, relativi all'applicazione dell'Intelligenza Artificiale in ambito industriale, che sono oggetto di ricerca presso il Dipartimento di Informatica dell'Università degli Studi di Torino. Il primo tema riguarda la compressione delle reti neurali, utili ad esempio per l'analisi automatica di prodotti a scaffale, al fine di ridurre i requisiti hardware. Il secondo tema riguarda l'adozione di forme di interazione gestuale senza contatto nel contesto di una linea di produzione.

1 Introduzione

Questo contributo, avente come oggetto l'applicazione dell'intelligenza artificiale nell'ambito dell'industria, è presentato dal Dipartimento di Informatica dell'Università degli Studi di Torino. Nel nodo sono attivi due gruppi che esplorano questo tema.

Il gruppo EIDOS¹ si occupa di analisi efficiente delle immagini tramite reti neurali profonde (deep learning). Le reti neurali stanno sostituendo gli algoritmi di analisi dell'immagine tradizionali in varie applicazioni industriali in ragione delle loro superiori prestazioni. Tali prestazioni sono ottenute al prezzo di una complessità che ne ostacola l'utilizzo in ambienti dove le risorse hardware siano scarse, es.: macchine utensili, dispositivi operativi mobili. Il gruppo ha sviluppato metodi per comprimere le reti neurali riducendone i requisiti hardware. Si porterà come caso d'esempio lo sviluppo di un sistema per l'analisi automatica di prodotti a scaffale sviluppato con un partner industriale.

Il secondo gruppo, composto da Cristina Gena e Fabiana Vernerio, si specializza nell'Interazione Uomo-Macchina. In questo contesto, studia come l'interazione gestuale, abilitata dal riconoscimento di movimenti e contrazioni muscolari tramite tecniche di intelligenza artificiale [Di Girolamo *et al.*, 2017], possa rappresentare una forma di interazione conveniente in contesti come quello di una linea di montaggio, dove vincoli di sicurezza impongono a chi vi lavora di indossare dispositivi di protezione individuale che renderebbero poco efficiente l'utilizzo di tecnologie di input standard

come tastiere, mouse, touchscreen. Il gruppo ha collaborato, nel contesto del progetto regionale HOME² (Hierarchical Open Manufacturing Europe), con IIT³ (Istituto Italiano di Tecnologia) e Synarea⁴.

2 Tema 1: Deep Learning efficiente per l'industria

Da diversi anni ormai, le reti neurali artificiali profonde [LeCun *et al.*, 2015] raggiungono prestazioni allo stato dell'arte in diversi settori industriali come ad esempio il rilevamento ottico di difetti di produzione, l'assistenza di operatori addetti all'utilizzo di macchine utensili, etc. Tali prestazioni sono frutto dell'elevato numero di neuroni artificiali e relative interconnessioni che è possibile modellare ed allenare tramite le recenti tecniche di apprendimento automatico e sistemi di calcolo parallelo. L'elevato numero di neuroni e relative interconnessioni si traducono tuttavia in requisiti di risorse di calcolo e soprattutto di memoria che non sempre sono disponibili in ambito industriale, quando tali reti devono essere impiegate ad esempio su macchine utensili e dispositivi operativi mobili.

Una possibile soluzione a tale problema è ricorrere ad un server di calcolo collocato in cloud con il compito di attendere le richieste della macchina utensile, elaborare tale richieste tramite rete neurale e fornire la risposta. Tale sistema però presenta diversi svantaggi. Innanzitutto, la connessione dati fra macchina e cloud potrebbe non essere sempre disponibile o non con costanza. Secondo, la banda disponibile potrebbe essere limitata, con una ripercussione sulla latenza della risposta che non permetterebbe alla macchina di operare in tempo reale. Terzo, l'Invio di dati industriale verso il cloud potrebbe porre problemi di riservatezza di processi industriali.

Per le ragioni sopra descritte, è fondamentale ridurre i requisiti di memoria e di calcolo delle reti neurali affinché queste possano essere applicate localmente all'impianto industriale. La soluzione proposta è semplificare la topologia della rete neurale, riducendo così le risorse necessarie al suo utilizzo e permettendone praticamente l'applicazione locale in ambito industriale. Lo strumento individuato è quello del *pruning*,

²<https://www.home-opensystem.org/>

³<https://www.iit.it/>

⁴<https://www.synarea.com/>

¹<http://eidos.di.unito.it/>

ovvero la rimozione selettiva di neuroni e/o connessioni fra essi con conseguente riduzione della complessità della rete e relativi requisiti computazionali. In tal senso, all'interno del gruppo sono state sviluppate diverse tecniche di pruning con la peculiarità che la selezione dei neuroni e delle interconnessioni da rimuovere avviene già in fase di allenamento e rimuovendo quegli elementi della rete che influiscono meno sulle prestazioni della rete stessa.

2.1 Progetti e applicazioni

I metodi di pruning di cui sopra sono stati prototipati all'interno del progetto *AI For Retail*, sviluppato con il partner industriale Synesthesia⁵. Il progetto ha vinto una serie di bandi regionali e nazionali tra cui il progetto GAP-AI bandito dal centro di competenza nazionale sul manufacturing (CIM4.0)⁶ il progetto GADAI relativo al bando PASS⁷ indetto dalla regione Piemonte e il bando VIR⁸. L'obiettivo del progetto è sfruttare il deep learning per l'analisi automatica del piano di vendita (per esempio supermercati) e in particolare di scaffali in contesti retail, ovvero vendita al dettaglio. Per un brand, è fondamentale pianificare correttamente l'esposizione del merchandise dato che la maniera in cui i prodotti sono presentati sugli scaffali è in grado di incrementare sostanzialmente le vendite. Per questo motivo l'analisi dello scaffale è un processo molto importante che permette sia di analizzare la disposizione corrente, e apportare eventuali modifiche, sia verificare che gli accordi presi vengano rispettati. Ad oggi queste verifiche vengono compiute prevalentemente a mano da operatori che analizzano visivamente la disposizione dei prodotti sullo scaffale e la annotano su dispositivo mobile con un processo lungo e pronò all'errore. Il progetto mira a sviluppare una app per dispositivi mobili capace di estrarre dalla foto di uno scaffale tutte le informazioni di interesse per il brand (es: presenza e posizione di un determinato prodotto) grazie ad una rete neurale. Come requisito di progetto, l'app deve poter funzionare anche senza accesso alla rete dati, che in ambienti chiusi come i supermercati è spesso assente. La sfida tecnologica consiste quindi nel far funzionare la rete neurale sul dispositivo mobile, in modo che questo possa analizzare le foto scaffale e inviare un report alla centrale operativa quando la rete torna disponibile. Le tecniche di pruning di cui sopra sono state sfruttate per generare delle reti neurali capaci di riconoscere prodotti a scaffali ed al contempo di rispettare i vincoli di capacità di calcolo e memoria disponibili sul dispositivo mobile utilizzato dall'operatore. L'applicazione che incorpora tali reti ha dimostrato di poter essere utilizzata tramite normali tablet Android o iOS già in uso dagli store account managers e consente di individuare la posizione di alcuni prodotti di interesse a partire da una foto di uno scaffale.

⁵<https://www.synesthesia.it/>

⁶<https://cim40.com/projects/gap-ai/>

⁷<https://bandi.regione.piemonte.it/contributi-finanziamenti/bando-pass>

⁸<https://www.finpiemonte.it/bandi/dettaglio-bando/>

[v-ir-voucher-acquisizione-servizi-per-ricerca-innovazione/](https://www.finpiemonte.it/bandi/dettaglio-bando/v-ir-voucher-acquisizione-servizi-per-ricerca-innovazione/)

2.2 Prospettive e sfide

I risultati della nostra sperimentazione hanno verificato che il pruning delle reti neurali è sufficiente a permetterne il loro utilizzo su dispositivi come smartphone e tablet. Tuttavia, i dispositivi wireless che costituiranno lo scheletro dell'Internet of Things (IoT), tecnologia abilitante di Industria 4.0 e Smart Factory, dispongono di capacità di calcolo ordini di grandezza inferiori. al fine di abilitare l'utilizzo delle reti neurali anche in ambito IoT, è necessario ridurre ulteriormente i relativi requisiti computazionali. I risultati preliminari della nostra ricerca suggeriscono che la rappresentazione della rete neurale su un numero di bit ridotto tramite ovvero alla *quantizzazione* sia la chiave per un'applicazione industriale pratica delle reti neurali anche in ambito IoT.

2.3 Pubblicazioni e altre risorse

Il metodo di pruning delle reti neurali di cui sopra sono discussi nel dettaglio in diverse pubblicazioni internazionali a conferenza e su rivista con peer review [Tartaglione *et al.*, 2020], [Tartaglione *et al.*, 2021], [Tartaglione *et al.*, 2022], [Bragagnolo *et al.*, 2021], [Bragagnolo e Barbano, 2022] ed è stata depositata domanda di brevetto di utilità industriale [Tartaglione *et al.*, 2019].

3 Tema 2: Interazione gestuale nella smart factory

Le fabbriche intelligenti si possono intendere come sistemi ciberfisici orientati alla produzione [Riedl *et al.*, 2014], in grado di sfruttare le tecnologie dell'informazione per garantire una gestione dei processi flessibile, efficiente, ecologica, economica e di alta qualità [Wang *et al.*, 2016]. In questo contesto, una sfida per chi si occupa di Interazione Uomo-Macchina è progettare opportune modalità per comunicare, visualizzare, interagire con i dati e le informazioni disponibili, specie se si prendono in considerazione i molteplici vincoli ambientali e di sicurezza. Ad esempio, tastiere e dispositivi touch sarebbero scomodi gli operatori che lavorano alle linee di produzione e devono indossare guanti protettivi, mentre tecnologie basate sull'interazione vocale rischierebbero di essere poco efficaci a causa del rumore di fondo. L'obiettivo del nostro gruppo è dunque quello di esplorare le applicazioni dell'interazione gestuale senza contatto, in combinazione con l'utilizzo di grandi display.

3.1 Progetti e applicazioni

Il tema in oggetto ha trovato spazio all'interno del progetto regionale HOME sull'Industria 4.0, in cui il gruppo ha collaborato con IIT, Synarea e con il partner industriale Galeasso s.r.l.⁹, un'azienda specializzata nella lavorazione della lamiera, che ha fornito un *case study* riguardante l'assemblaggio e la saldatura di componenti in metallo.

Il riconoscimento delle *gesture* è affidato ad un bracciale intelligente, sviluppato *ad hoc* dal partner IIT sulla base di lavori precedenti nel campo della riabilitazione neuromotoria

⁹<https://www.galeasso.com/>

[Di Girolamo *et al.*, 2017]. Tale bracciale permette di rilevare movimenti e contrazioni muscolari e utilizza una libreria di machine learning per permettere la calibratura e l'apprendimento di gesti specifici, definiti dall'utente. Rispetto a tecnologie commerciali basate su *computer vision*, come ad esempio Kinect¹⁰, la soluzione proposta ha il vantaggio di sollevare meno timori legati alla violazione della privacy.

Nel corso del progetto, sono stati condotti vari *user studies*. Un primo *guessability study* (19 partecipanti) aveva l'obiettivo di aiutarci a definire dei gesti appropriati per una serie di azioni (*scorrere una lista*, *selezionare un'opzione*, e *ruotare*, *traslare*, *ridimensionare* un modello 3D) che gli operatori eseguono abitualmente quando usano software di modellazione 3D per visualizzare i prodotti da assemblare [Andolina *et al.*, 2019].

Successivamente, abbiamo svolto due valutazioni con saldatori esperti per valutare l'efficacia e la facilità d'uso dell'interazione gestuale nell'effettivo contesto di lavoro, attraverso la simulazione di due task quotidiani, quali la selezione della prossima lavorazione da svolgere e la manipolazione di un modello 3D del prodotto finito. Indossando il bracciale intelligente, gli operatori coinvolti potevano interagire attraverso delle *gesture* con una versione semplificata del software di gestione delle lavorazioni, proiettata su un grande display collocato vicino alla loro postazione di lavoro. A seguito della prima valutazione, è stata messa a punto una nuova versione del bracciale con l'obiettivo di migliorare l'accuratezza nel riconoscimento delle *gesture*, evitando in particolare di rilevare piccoli movimenti involontari.

Per quanto gli studi sul campo siano stati condotti con un numero molto limitato di partecipanti (4 in totale), i nostri risultati sono incoraggianti. Gli operatori coinvolti, infatti, hanno apprezzato la nostra soluzione, hanno valutato positivamente la loro *user experience* e sarebbero interessati ad utilizzare l'interazione gestuale nelle loro attività quotidiane. Resta da capire se l'esperienza possa rimanere positiva anche nel lungo periodo, a seguito di un possibile, naturale calo nel livello di motivazione e dell'eventuale insorgere della fatica fisica legata all'esecuzione delle *gesture*.

3.2 Prospettive e sfide

I nostri risultati hanno mostrato che il modo in cui viene calibrato il bracciale ha un ruolo fondamentale nel determinare la *user experience* e, quindi, l'effettiva adottabilità dell'interazione gestuale: se un'eccessiva sensibilità del bracciale può causare errori nell'interazione, il suo opposto determina lentezza e scarsa responsività dell'interfaccia. La situazione è complicata dal fatto che i parametri corretti variano da un individuo all'altro. Se, da una parte, si potrebbe pensare di risolvere questo problema sfruttando la capacità del bracciale di apprendere ed effettuando una calibratura *ad hoc* per ogni utente, dall'altra è necessario considerare che, in un contesto reale, in cui è probabile che un numero limitato di dispositivi sia condiviso tra più operatori, le frequenti operazioni di ricalibratura rappresentano un costo poco sostenibile. Parallelamente, è necessario considerare che le *gesture* più facilmente e universalmente rilevabili non necessariamente sono

intuitive e naturali per l'utente, il quale potrebbe quindi vedersi costretto a memorizzare un insieme di movimenti convenzionali che potrebbero anche assumere significati diversi a seconda del contesto, con un conseguente sovraccarico cognitivo e peggioramento dell'esperienza d'uso complessiva. La sfida che intravediamo sta quindi nella soddisfazione di tre esigenze in conflitto ma ugualmente importanti: la naturalezza delle *gesture*, la loro riconoscibilità e la necessità di una calibratura "leggera".

3.3 Pubblicazioni e altre risorse

Il tema di ricerca discusso è stato oggetto di due pubblicazioni, presentate in conferenze internazionali: IEEE SMC 2019¹¹, [Andolina *et al.*, 2019] e INTERACT 2021¹², [Andolina *et al.*, 2021].

Riferimenti bibliografici

- [Andolina *et al.*, 2019] Salvatore Andolina, Paolo Ariano, Davide Brunetti, Nicolo Celadon, Guido Coppo, Alain Favetto, Cristina Gena, Sebastiano Giordano, e Fabiana Venero. Experimenting with large displays and gestural interaction in the smart factory. In *2019 IEEE International Conference on Systems, Man and Cybernetics, SMC 2019, Bari, Italy, October 6-9, 2019*, pages 2864–2869. IEEE, 2019.
- [Andolina *et al.*, 2021] Salvatore Andolina, Paolo Ariano, Davide Brunetti, Nicolò Celadon, Guido Coppo, Alain Favetto, Cristina Gena, Sebastiano Giordano, e Fabiana Venero. Introducing gestural interaction on the shop floor: Empirical evaluations. In Carmelo Ardito, Rosa Lanziloti, Alessio Malizia, Helen Petrie, Antonio Piccinno, Giuseppe Desolda, e Kori Inkpen, editors, *Human-Computer Interaction - INTERACT 2021 - 18th IFIP TC 13 International Conference, Bari, Italy, August 30 - September 3, 2021, Proceedings, Part V*, volume 12936 of *Lecture Notes in Computer Science*, pages 451–455. Springer, 2021.
- [Bragagnolo *et al.*, 2021] Andrea Bragagnolo, Enzo Tartaglione, Attilio Fiandrotti, e Marco Grangetto. On the role of structured pruning for neural network compression. In *2021 IEEE International Conference on Image Processing (ICIP)*, pages 3527–3531, 2021.
- [Bragagnolo e Barbano, 2022] Andrea Bragagnolo e Carlo Alberto Barbano. Simplify: A python library for optimizing pruned neural networks. *SoftwareX*, 17:100907, 2022.
- [Di Girolamo *et al.*, 2017] M. Di Girolamo, N. Celadon, S. Appendino, A. Turolla, e P. Ariano. Emg-based bio-feedback system for motor rehabilitation: A pilot study. In *2017 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, pages 1–4, Oct 2017.
- [LeCun *et al.*, 2015] Yann LeCun, Yoshua Bengio, e Geoffrey Hinton. Deep learning. *nature*, 521(7553):436–444, 2015.

¹⁰<https://azure.microsoft.com/en-us/services/kinect-dk/>

¹¹<https://smc2019.org/>

¹²<https://www.interact2021.org/>

- [Riedl *et al.*, 2014] Matthias Riedl, Holger Zipper, Marco Meier, e Christian Diedrich. Cyber-physical systems alter automation architectures. *Annual Reviews in Control*, 38:123–133, 2014.
- [Tartaglione *et al.*, 2019] Enzo Tartaglione, Marco Grangetto, Francesco Odierna, Andrea Bragagnolo, e Attilio Fiandrotti. Method and apparatus for pruning neural networks, October 13 2019. IT Patent App. IT201900018821A1.
- [Tartaglione *et al.*, 2020] Enzo Tartaglione, Andrea Bragagnolo, e Marco Grangetto. Pruning artificial neural networks: A way to find well-generalizing, high-entropy sharp minima. In Igor Farkaš, Paolo Masulli, e Stefan Wermter, editors, *Artificial Neural Networks and Machine Learning – ICANN 2020*, pages 67–78, Cham, 2020. Springer International Publishing.
- [Tartaglione *et al.*, 2021] Enzo Tartaglione, Andrea Bragagnolo, Francesco Odierna, Attilio Fiandrotti, e Marco Grangetto. Serene: Sensitivity-based regularization of neurons for structured sparsity in neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, pages 1–14, 2021.
- [Tartaglione *et al.*, 2022] Enzo Tartaglione, Andrea Bragagnolo, Attilio Fiandrotti, e Marco Grangetto. Loss-based sensitivity regularization: Towards deep sparse neural networks. *Neural Networks*, 146:230–237, 2022.
- [Wang *et al.*, 2016] Shiyong Wang, Jiafu Wan, Di Li, e Chunhua Zhang. Implementing smart factory of industrie 4.0: An outlook. *International Journal of Distributed Sensor Networks*, 12(1):3159805, 2016.