

IA per la diagnosi di patologie neurocognitive e della pelle

Marco Leo, Pierluigi Carcagnì, Cosimo Distante

Consiglio Nazionale delle Ricerche
Istituto di Scienze Applicate e Sistemi intelligenti Campus Ecotekne, Via Monteboni - 73100 Lecce
{marco.leo,pierluigi.carcagni,cosimo.distante}@cnr.it

Abstract

Il documento descrive due tecniche evolute, basate su computer vision ed intelligenza artificiale, per la valutazione e diagnosi di deficit neuro-cognitivi e patologie della pelle di tipo canceroso. La prima tecnica consiste in un approccio computazionale per valutare la capacità di produrre correttamente espressioni facciali. La seconda tecnica prevede l'impiego di reti a convoluzione per classificare 7 differenti tipologie di patologie di natura cancerosa della pelle. Entrambe le tecniche sono state valutate su dati reali e hanno dimostrato un'ottima capacità di fornire evidenza numerica a supporto della diagnosi e della valutazione clinica.

1 Introduzione

La diagnosi medica automatica è un centro emergente di interesse per la visione delle macchine e per l'intelligenza artificiale. L'obiettivo è fornire informazioni obiettive non invasive sulle condizioni di un paziente a partire da immagini digitali acquisite anche attraverso strumenti non professionali (quali ad esempio lo smartphone). Questo introduce la possibilità di monitorare il paziente da remoto, migliorando così la qualità di vita oltre che di intervenire in modo precoce in caso di problemi. In questo documento sono descritte due tecniche evolute, basate su computer vision ed intelligenza artificiale, per la valutazione e diagnosi di deficit neuro-cognitivi e patologie cutanee. La prima tecnica consiste in un approccio computazionale per valutare la capacità di produrre correttamente espressioni facciali, che sono alla base delle relazioni sociali e la cui abilità di produzione rappresenta uno degli indicatori più rilevanti per la diagnosi e valutazione di malattie neurocognitive, ed in particolari dei disordini dello spettro autistico. La seconda tecnica prevede l'impiego di reti a convoluzione per classificare 7 differenti patologie di natura cancerosa della pelle.

2 Tecniche di CV e IA per la diagnosi e valutazione clinica

Per la valutazione dei disordini neurocognitivi è stato sviluppato un sistema basato su sequenze di immagini acquisite da fotocamere commerciali. La pipeline algoritmica è composta



Figura 1: struttura algoritmica del sistema di analisi dei volti per la valutazione di disordini neurocognitivi

da cinque componenti principali: un rilevatore di volti, un rilevamento facciale e un blocco di localizzazione, una stima della posizione della testa e un blocco di osservazione dello sguardo, uno stimatore dell'intensità dell'unità di azione facciale e un modulo di alto livello per l'analisi semantica dei dati.

I principali elementi di innovazione sono:

1. La pipeline di elaborazione basata su metodi di visione artificiale è in grado di analizzare le espressioni facciali in condizioni non controllate;
2. Le intensità delle unità di azione sono stimate analizzando l'aspetto locale e quindi entrambe le relazioni temporali e geometriche, apprese dalle reti neurali convoluzionali, sono sfruttate per regolarizzare le stime raccolte;
3. Viene formulata una modellizzazione statistica delle configurazioni del volto non emotivo, in grado di incorporare movimenti stereotipati e di evidenziare anche i movimenti volontari sottili dei muscoli facciali.

Per il riconoscimento automatico di sette classi di patologie cancerose della pelle, a partire da immagini dermoscopiche, è stata implementata una rete neurale convoluzionale (CNN), basata sull'architettura DenseNet [Huang *et al.*, 2017]. L'intera architettura è riportata in Tabella 1.

Partendo dai parametri ottenuti in seguito al pre-addestramento su un dataset generico di oggetti¹, è stato poi

¹ www.image-net.org

Layers	Output Size	Densenet-62
Convolution	112×112	7×7 conv, stride 2
Pooling	56×56	3×3 max pool, stride 2
Dense Block (1)	56×56	$\begin{matrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{matrix} \times 6$
Transition Layer (1)	56×56 28×28	1×1 conv 2×2 average pool, stride 2
Dense Block (2)	28×28	$\begin{matrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{matrix} \times 12$
Transition Layer (2)	28×28 14×14	1×1 conv 2×2 average pool, stride 2
Dense Block (3)	14×14	$\begin{matrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{matrix} \times 11$
Classification Layer	1×1	7×7 global average pool 7D fully-connected, softmax

Tabella 1: Architettura della CNN impiegata

effettuato un *fine-tuning* mediante il dataset *HAM10000* [Tschandl *et al.*, 2018] costituito da 10015 immagini dermo-scopiche suddivise in sette patologie cancerose della pelle: melanoma, nevo melanocitico, carcinoma basocellulare, cheratosi attinica, cheratosi benigna, dermatofibroma, vascolare. Al fine di aumentare la capacità discriminante della rete, nella fase di apprendimento è stato impiegato l’approccio basato su *center-loss*, proposto in [Wen *et al.*, 2016], per la funzione di costo. La procedura di *fine-tuning* è stata eseguita mediante un innovativo approccio multi livello sugli ultimi 10 layers della rete che ha permesso di ottenere 6 modelli ognuno dei quali maggiormente discriminante per una particolare classe di patologie. I modelli ottenuti sono stati quindi utilizzati per l’addestramento di un classificatore di tipo SVM. Il principale elemento di innovazione consiste nella creazione di modelli neurali generati dalla stessa architettura di rete a ridotta complessità, che consente un’implementazione in sistemi embedded quali dispositivi portatili per l’utilizzo sia in ambito clinico che per la diagnosi domestica.

3 Risultati Sperimentali

Per dimostrare la correlazione tra i risultati numerici della pipeline proposta e le capacità individuali di eseguire espressioni facciali, è stata eseguita una valutazione preliminare su un set di dati di sequenza contenente persone con sviluppo psicologico e fisico normotipico, durante l’esecuzione o meno delle espressioni facciali [Lucey *et al.*, 2010]. La pipeline proposta è stata poi testata su un gruppo di 17 bambini con diagnosi ASD. La prima prova dell’efficacia del metodo computazionale proposto per la stima delle capacità di produzione delle espressioni facciali è stata l’evidenza di poter quantificare le abilità dei singoli bambini e anche nelle loro diverse parti del viso. Inoltre è stato possibile osservare che i risultati dell’approccio proposto possono essere confrontati con le annotazioni degli psicologi (vedere figura 2)

La validità del sistema per il riconoscimento delle patologie cutanee è stata fatta mediante il confronto con i risultati riportati nel Challenge ISIC 2018² per quanto riguarda il Task 3 relativo alla classificazione delle sette categorie di patologie relative al dataset HAM10000. Il Task 3 è stato affrontato da 141 gruppi di ricerca e, escludendo dataset esterni in aggiunta al solo HAM10000 per l’addestramento, i migliori ri-

²<https://challenge2018.isic-archive.com>

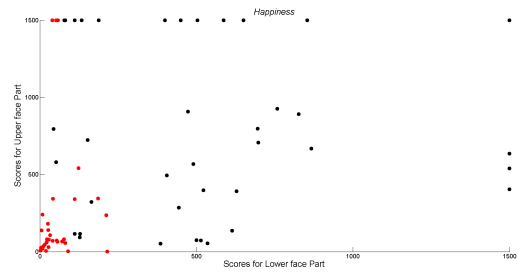


Figura 2: Confronto tra gli output del sistema (coordinate nel grafico) e le annotazioni degli psicologi (rosso=espressione non fatta; nero=espressione fatta).

sultati sono stati ottenuti con l’approccio basato su *ensemble* di modelli (molto pesante computazionalmente e complesso nella definizione) descritto in [Jiaxin Zhuangy, 2018] con una precisione media di $82,6\% \pm 2.0$ su 7 classi. L’approccio proposto in questo documento ha invece ottenuto una precisione dell’ 88% sulle stesse 7 classi con in aggiunta una notevole semplificazione dei modelli e della loro definizione.

4 Ringraziamenti

Per le attività di ricerca relative alla classificazione automatica delle lesioni cutanee si ringraziano Dedalus SPA ed Infotel srl. Per le attività di ricerca relative alla valutazione e diagnosi dei disordini dello spettro autistico si ringraziano Amici di Nico Onlus, Adelpia Onlus ed il Dipartimento di Storia Società e Studi sull’uomo dell’Università del Salento.

Riferimenti bibliografici

- [Huang *et al.*, 2017] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, e Kilian Q Weinberger. Densely connected convolutional networks. In *CVPR*, volume 1, page 3, 2017.
- [Jiaxin Zhuangy, 2018] Siyamalan Manivannanz Jiaxin Zhuangy, Weipeng Liy. Skin lesion analysis towards melanoma detection using deep neural network ensemble. 2018.
- [Lucey *et al.*, 2010] Patrick Lucey, Jeffrey F Cohn, Takeo Kanade, Jason Saragih, Zara Ambadar, e Iain Matthews. The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. In *Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2010 *IEEE Computer Society Conference on*, pages 94–101. IEEE, 2010.
- [Tschandl *et al.*, 2018] Philipp Tschandl, Cliff Rosendahl, e Harald Kittler. The ham10000 dataset: A large collection of multi-source dermatoscopic images of common pigmented skin lesions. *arXiv preprint arXiv:1803.10417*, 2018.
- [Wen *et al.*, 2016] Yandong Wen, Kaipeng Zhang, Zhifeng Li, e Yu Qiao. A discriminative feature learning approach for deep face recognition. In *European Conference on Computer Vision*, pages 499–515. Springer, 2016.