

# Classificazione delle dimensioni del polipo coloretale utilizzando una rete siamese

Nomi degli autori nascosti

e-mail trattenute

Indirizzo nascosto

Redattori: in revisione per MIDL 2019

## Astratto

Il cancro del colon-retto è una delle principali cause di decessi correlati al cancro con una prevalenza in aumento. Un fattore chiave nella probabilità di differenziazione delle cellule adenomatose è il diametro del polipo. Esiste un significativo valore di cut-off di 10 mm che i medici utilizzano nella gestione della diagnosi. Proponiamo un nuovo metodo per classificare la dimensione del polipo in classi superiori o inferiori a 10 mm basate su una rete siamese. In una prima fase, una rete siamese viene addestrata per costruire un'incorporamento di caratteristiche ad alta dimensione estratte per ciascuna dimensione del polipo. Nella seconda fase, utilizziamo un approccio k-NN per classificare le dimensioni dei polipi in base alla distanza tra l'incorporamento delle caratteristiche dell'immagine di input e l'intero spazio di incorporamento appreso dalla rete siamese. Questo metodo consente una migliore classificazione binaria delle classi di dimensione del polipo sub- e sup-10 mm. I nostri dati consistono in circa 55.000 immagini di 129 film classificati in varie dimensioni di polipi che vanno da 1 a 15 mm. Abbiamo addestrato il nostro modello su 10.746 immagini e testato su 2.688 immagini equamente suddivise in ciascuna categoria binaria. Abbiamo ottenuto il 79,2% nella classificazione delle caratteristiche e il 95,7% nella classificazione delle dimensioni del polipo.

Parole chiave: Diagnosi assistita da computer, rete siamese, classificazione dei polipi

## 1. Introduzione

La dimensione del polipo coloretale è un biomarcatore critico nella diagnosi e supervisione del cancro del colon-retto, con un diametro del polipo più grande che ha una maggiore probabilità di differenziazione delle cellule adenocarcinomatose (Martínez et al., 2015; Klein et al., 2016; Summers, 2010). Ci sono due soglie significative nel processo decisionale clinico, che si verificano a 5 mm e 10 mm. Esiste un'elevata variabilità intra e interclinica nella stima della dimensione del polipo (Elwir et al., 2017), e quindi è necessario un sistema di classificazione automatica che possa aiutare i medici nelle loro stime. In questo lavoro proponiamo un modello di classificazione binaria basato su reti siamesi. Addestrando una rete neurale per apprendere caratteristiche descrittive ad alta dimensione per ciascuna categoria di polipi, ovvero inferiori o superiori a 10 mm, possiamo classificare nuove istanze invisibili di immagini di polipi nella classe più rilevante. In una prima fase, addestriamo una rete siamese per apprendere questo incorporamento e in una seconda fase utilizziamo k-Nearest Neighbors (k-NN) per calcolare il cluster di classe più vicino all'immagine della query. Per quanto ne sappiamo, presentiamo uno dei primi utilizzi delle reti siamesi per classificare le dimensioni dei polipi. Inoltre, il nostro metodo consente di costruire un'elevata comprensione dimensionale delle caratteristiche del polipo che alla fine può essere utilizzata non solo per la classificazione binaria.

Sono state condotte molte ricerche sulla classificazione dei polipi. La maggior parte di essi si concentra sul rilevamento dei polipi (Wang et al., 2018; Urban et al., 2018; Masashi et al., 2018; Zhang et al., 2017), sebbene alcuni lavori si concentrino anche sulla caratterizzazione dei polipi, come l'istopatologia (Korbar et al., 2017; Zhang et al., 2017). Tuttavia, abbiamo trovato pochi lavori che si concentrano sul polipo

classificazione delle dimensioni (Martínez et al., 2015; Itoh et al., 2019), che ottengono la classificazione utilizzando la stima della profondità. Poiché i dati sulla profondità della verità del terreno coloretale in vivo sono difficili da ottenere, il lavoro in (Martínez et al., 2015) stima la profondità utilizzando una tecnica nota come strategia di defocus, che richiede coefficienti di sfocatura, addestrati utilizzando dati fantasma. Il lavoro in (Itoh et al., 2019) utilizza reti neurali di stima della profondità monoculare preaddestrate, che addestrano utilizzando un approccio non supervisionato. Tuttavia, il modello preaddestrato è stato addestrato su funzionalità basate sulla visione artificiale e pertanto non è adatto a classificare le funzionalità di imaging medico. Proponiamo una nuova metodologia utilizzando le informazioni sull'immagine basate su reti neurali siamesi. Le reti neurali siamesi sono state utilizzate per la prima volta in (Bromley et al., 1993) per classificare le firme. Consistono nell'avere due reti neurali parallele che condividono i pesi tra di loro. Ottimizzando su una misura di somiglianza tra gli input, è possibile apprendere potenti caratteristiche discriminative. Ciò consente alla rete di generalizzare bene a dati nuovi e invisibili provenienti da distribuzioni sconosciute, e come tali negli ultimi anni sono diventati popolari nel campo della visione artificiale per le attività di riconoscimento facciale (Taigman et al., 2014; Koch et al., 2015; Varior et al., 2016), sebbene siano stati utilizzati in poche altre applicazioni (Baraldi et al., 2015; Bertinetto et al., 2016).

## 2. Metodi

### 2.1. Dati

I nostri dati sono costituiti da 129 filmati colonscopici di diversi pazienti, ciascuno della durata approssimativa di circa 15 minuti. Gli esperti hanno annotato i segmenti per ogni singolo filmato con le dimensioni del polipo e le condizioni di acquisizione pertinenti, come luce bianca, infrarosso vicino al blu e cromo. Per ogni segmento, i fotogrammi del film pertinenti sono stati estratti nella classe di dimensioni del polipo appropriata e salvati come immagini. I dati sono stati quindi separati in classi superiori o inferiori che costituiscono rispettivamente 4.478 e 51.845 immagini. Per avere un set di dati bilanciato, abbiamo selezionato casualmente 4.478 immagini appartenenti alla categoria sottostante. Abbiamo suddiviso i nostri dati in set di dati di addestramento/test utilizzando una suddivisione dell'80/20%. Per addestrare la rete siamese, abbiamo generato coppie di immagini classificate nella stessa o in una classe diversa. Pertanto, i nostri dati di addestramento consistevano in 10.746 coppie di immagini di classe: la stessa sopra, la stessa sotto e diversa. I nostri dati di test consistevano in 2.688 coppie di immagini corrispondenti alle stesse classi.

### 2.2. Struttura

Proponiamo un approccio basato sul siamese per ottenere incorporamenti di caratteristiche ad alta dimensione tali che le immagini appartenenti alla stessa classe abbiano incorporamenti simili. La struttura per il nostro metodo di classificazione è mostrata in Fig. 1. In una prima fase la rete siamese viene addestrata avendo coppie di immagini come input per una rete gemella che condivide i pesi. La rete si basa sull'architettura VGG16 (Simonyan e Zisserman, 2014). Ciascuna delle immagini è incorporata in un vettore di caratteristiche 4.096 che viene quindi fornito come input a una funzione energetica.

Questa funzione calcola una metrica che mappa la distanza tra queste rappresentazioni di feature di alto livello in modo che sia piccola se sono simili e grande se sono diverse. Usiamo la funzione di perdita contrastiva definita come:

$$L(W, Y, X_1, X_2) = (1 - Y) \frac{1}{2} (D_w(X_1, X_2))^2 + Y \frac{1}{2} (\max(0, m - D_w(X_1, X_2)))^2, \quad (1)$$

## PSNet

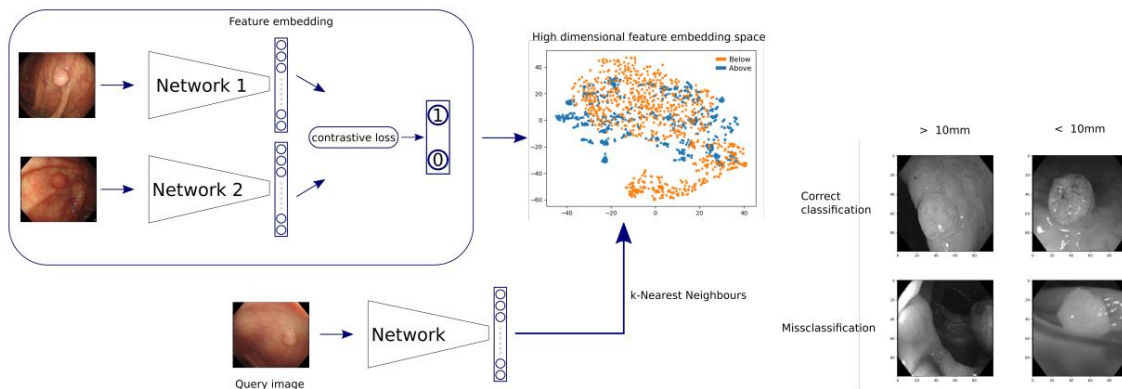


Figura 1: A sinistra: quadro di classificazione dei polipi. La rete siamese è mostrata in alto a sinistra. Le reti 1 e 2 rappresentano reti gemelle con pesi condivisi, basate sull'architettura VGG16. Le nuove immagini di query possono essere passate attraverso la rete addestrata e una funzione di distanza kNN viene utilizzata per calcolare l'incorporamento dell'output al resto degli incorporamenti delle funzionalità. Una rappresentazione T-SNE 2D dello spazio di incorporamento può essere vista in alto a destra. A destra: esempio di risultati di classificazione.

dove  $W$  rappresenta i pesi della rete,  $X_1$  e  $X_2$  rappresentano le caratteristiche incorporate delle coppie di immagini, con  $Y = 1$  se sono simili e  $Y = 0$  se sono dissimili.  $D_w$  rappresenta la norma  $l_2$ , tale che  $D_w(X_1, X_2) = \|X_1 - X_2\|_2$ , e  $m$  è il margine. Addestrando la rete, è possibile ottenere uno spazio di incorporamento in cui le caratteristiche sono ben separate. Fig. 1.

In un secondo passaggio usiamo la rete addestrata per calcolare il vettore di caratteristiche di una nuova immagine di query e usiamo l'algoritmo k-NN per calcolare la distanza dal cluster di classe più vicino.

### 3. Risultati e conclusione

Abbiamo valutato la rete siamese su 2.688 coppie di immagini appartenenti alle tre classi menzionate nella sezione 2.1. Questo ha valutato il potere discriminante dello spazio di incorporamento di alto livello appreso dalla nostra rete. La nostra rete è stata in grado di classificare se coppie di immagini erano simili o dissimili con una precisione del 70,2%. Abbiamo quindi utilizzato le 1792 immagini di prova appartenenti alla classe above e below per classificarle utilizzando un approccio k-NN. Siamo stati in grado di classificarli con una precisione del 95,7%, utilizzando  $k = 100$ . Nei nostri esperimenti, ogni fotogramma è stato separato in modo casuale in dati di addestramento e test. Sebbene vi sia una netta separazione tra i dati di addestramento e di test per quanto riguarda ciascun fotogramma, dobbiamo considerare diversi fotogrammi ottenuti dallo stesso video clip. Poiché anche le immagini di prova appartengono ai 129 film, riteniamo che possa essersi verificato un errore di classificazione. Miriamo a ottenere più film per eseguire una convalida più appropriata.

Abbiamo presentato un approccio preliminare alla classificazione delle dimensioni del polipo basato sul lavoro di rete siamese. Sebbene questo lavoro classifichi i polipi in classi binarie, la nostra metodologia può essere facilmente adattata per funzionare per la versione multiclasse di questo problema.

trattenuto

## Ringraziamenti

Riconoscimenti trattenuti.

## Riferimenti

Lorenzo Baraldi, Costantino Grana, Rita Cucchiara. Una rete siamese profonda per il rilevamento di scene nei video trasmessi. In ACM Multimedia, 2015.

Luca Bertinetto, Jack Valmadre, Jo˜ao F. Henriques, Andrea Vedaldi, and Philip HS Torre. Reti siamesi completamente convoluzionali per il tracciamento di oggetti. Nella conferenza europea sulla visione artificiale, 2016.

Jane Bromley, Isabelle Guyon, Yann LeCun, Eduard Sˆackinger e Roopak Shah. Verifica della natura del sigillo utilizzando una rete neurale a ritardo di tempo "siamese". In Proceedings of the 6th International Conference on Neural Information Processing Systems, NIPS'93, pagine 737–744, 1993.

S Elwir, A Shaukat, M Shaw, J Hughes e J Colton. Variabilit  e fattori associati al dimensionamento dei polipi da parte degli endoscopisti in una grande pratica comunitaria. Open internazionale di endoscopia, 5(8), 2017.

Hayato Itoh, Holger R. Roth, Yuichi Mori, Masashi Misawa, Masahiro Oda, Shin-ei Kudo e Mori Kensaku. Classificazione delle dimensioni del polipo con caratteristiche rgb-d per la colonscopia. Proc. SPIE 10950, Medical Imaging 2019: Diagnosi assistita da computer, 1095015, 2019.

JL Klein, M Okcu, Karl Preisegger e Heinz Hammer. Distribuzione, dimensione e forma degli adenomi coloretali determinati da un colonscopista con un alto tasso di rilevamento delle lesioni: influenza dell'et , del sesso e dell'indicazione della colonscopia. Giornale europeo unito di gastroenterologia, 4(3):438–448, 2016.

G Koch, R Zemel e R Salakhutdinov. Reti neurali siamesi per immagini one-shot riconoscimento. In ICML 2015 Deep Learning Workshop, 2015.

B Korbar, AM Olofson, AP Mirafior, CM Nicka, MA Suriawinata, L Torresani, AA Suriawinata, and S Hassanpour. Apprendimento approfondito per la classificazione dei polipi coloretali su immagini di diapositive intere. J Pathol Inform., 8:30, 2017.

Fabio Mart ınez, Josu e Ruano, Mart ın G omez e Eduardo Romero. Stima della dimensione dei polipi durante le procedure endoscopiche effettive utilizzando una caratterizzazione spazio-temporale. Imaging e grafica medica computerizzata, 43: 130–136, 2015.

Misawa Masashi, Kudo Shin-ei, Mori Yuichi, Cho Tomonari, Kataoka Shinichi, Yamauchi Akihiro, Ogawa Yushi, Maeda Yasuharu, Takeda Kenichi, Ichimasa Katsuro, Nakamura Hiroki, Yagawa Yusuke, Toyoshima Naoya, Ogata Noriyuki, Kudo Toyoki, Hisayuki Tomokazu, Hayashi Takemasa, Wakamura Kunihiko, Baba Toshiyuki, Ishida Fumio, Itoh Hayato, Roth Holger, Oda Masahiro e Mori Kensaku. Rilevazione di polipi assistita da intelligenza artificiale per colonscopia: esperienza iniziale. Gastroenterologia, 154:2027 – 2029, 2018.

K. Simonyan e A. Zisserman. Reti convoluzionali molto profonde per il riconoscimento di immagini su larga scala. CoRR, abs/1409.1556, 2014.

RM estati. Misurazione della dimensione del polipo alla colonografia ct: cosa sappiamo e cosa abbiamo bisogno di sapere? Radiologia, 255:707–720, 2010.

Yaniv Taigman, Ming Yang, Marc'Aurelio Ranzato e Lior Wolf. Deepface: colmare il divario rispetto alle prestazioni a livello umano nella verifica facciale. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), giugno 2014.

Gregor Urban, Priyam Tripathi, Talal Alkayali, Mohit Mittal, Farid Jalali, William Karnes e Pierre Baldi. Il deep learning localizza e identifica i polipi in tempo reale con una precisione del 96% nello screening della colonscopia. Gastroenterologia, 155(4):1069 – 1078, 2018.

Rahul Rama Varior, Mrinal Haloi e Gang Wang. Architettura di rete neurale convoluzionale siamese recitata per la reidentificazione umana. Nella conferenza europea sulla visione artificiale, 2016.

Pu Wang, Glissen Brown Jeremy R. Xiao, Xiao, Tyler M. Berzin, Mengtian Tu, Fei Xiong, Xiao Hu, Peixi Liu, Yan Song, Di Zhang, Xue Yang, Liangping Li, Jiong He, Xin Yi, Jingjia Liu, e Xiaogang Liu. Sviluppo e validazione di un algoritmo di deep learning per il rilevamento di polipi durante la colonscopia. Ingegneria biomedica della natura, 2:741–748, 2018.

R. Zhang, Y. Zheng, TWC Mak, R. Yu, SH Wong, JYW Lau e CCY Poon.  
Rilevamento automatico e classificazione dei polipi coloretali trasferendo le caratteristiche cnn di basso livello dal dominio non medico. IEEE Journal of Biomedical and Health Informatics, 21(1):41–47, 2017.