

Visual recognition nell'analisi di immagini di risonanza magnetica in pazienti affetti da neoplasia cerebrale: confronto tra algoritmi di cognitive computing e performance umana.

Camilla Russo¹, Paolo Maresca¹ and Alfredo Marinelli^{2,3}

- 1) *Dipartimento di Ingegneria Elettrica e delle Tecnologie dell'Informazione, Università degli Studi di Napoli "Federico II", Italia*
- 2) *Dipartimento di Medicina Clinica e Chirurgia, Università degli Studi di Napoli "Federico II", Italia*
- 3) *IRCCS Neuromed Istituto Neurologico Mediterraneo Pozzilli (INM), Pozzilli, Italia*

Abstract

Questo contributo è finalizzato a confrontare le performance classificative di un sistema innovativo di visual recognition e cognitive computing, basato sul framework di deep learning open-source per l'analisi di dati di imaging medicale denominato FuseMedML, rispetto alle performance classificative umane ottenute coinvolgendo nella valutazione degli stessi dati radiologi con diverso grado di esperienza nella valutazione in Risonanza Magnetica delle neoplasie cerebrali. In tale studio pilota su un modello di setting neuro-oncologico, gli strumenti di intelligenza computazionale si sono dimostrati globalmente più rapidi da addestrare ed utilizzare grazie alla moderna potenza di calcolo, con performances classificative superiori a quelle umane (soprattutto se confrontate con quelle di osservatori di minor esperienza), e con risultati che sottolineano come i tempi e le risorse per l'addestramento necessari ad ottenere prestazioni diagnostiche soddisfacenti siano notevolmente implementati con tale supporto tecnologico. La potenzialità finale di strumenti di intelligenza computazionale è quella di fornire un supporto informatico tale da rendere più efficiente l'analisi di immagini biomedicali ed agevolare gli orientamenti diagnostico-terapeutici da essa derivanti.

Keywords

Intelligenza artificiale, cognitive computing, neuroncologia, neuroimaging.

1. Introduzione

Nel campo della radiologia, l'accurato riconoscimento e la caratterizzazione di sospette lesioni tumorali sono fondamentali per garantire ai pazienti il più corretto orientamento diagnostico; infatti, le informazioni raccolte dal radiologo al momento della diagnosi di sospetto tramite la valutazione e l'interpretazione dei dati di imaging forniscono agli oncologi e ai chirurghi gli strumenti necessari per avviare il paziente verso la più efficace strada terapeutica [1]. Nel campo della neuroncologia, la valutazione di un sospetto tumore cerebrale mediante imaging di risonanza magnetica (MRI) è generalmente un

passaggio essenziale nell'iter diagnostico; tuttavia, come per molte metodiche diagnostiche operatore-dipendenti, anche l'interpretazione di questo tipo di immagini è fortemente influenzata dalla specifica esperienza del neuroradiologo [2]. Questo risulta ancor più vero se si considera la variabilità clinico-radiologica della patologia tumorale del sistema nervoso centrale (SNC) ed il possibile affaticamento degli specialisti radiologi e più in generale di tutte le figure professionali coinvolte in tale processo di valutazione. In quest'ottica, un sempre crescente numero di tecnologie computer-assisted e di tecniche orientate all'intelligenza computazionale stanno nascendo a supporto del medico nella sua attività

quotidiana, con sempre nuovi scenari di impiego anche nel campo dell'oncologia e della neuroncologia [3-6].

2. Background

In neuroncologia le due più comuni lesioni neoplastiche del SNC dell'adulto sono rappresentate dalle metastasi cerebrali e dal glioblastoma [7-8]. Le metastasi, definite come localizzazioni secondarie al SNC di lesioni neoplastiche primarie che insorgono in un altro organo o sistema, rappresentano in termini di frequenza il tumore cerebrale maligno più comune dell'età adulta; generalmente multiple e più frequentemente diagnosticate in stadi avanzati di malattia in pazienti con anamnesi positiva per tumore, le metastasi raramente possono essere singole (SBM) e rappresentare la prima manifestazione clinica d'esordio in pazienti con precedente anamnesi oncologica silente [9]. Al secondo posto in termini di incidenza figurano i tumori cerebrali primitivi, definiti come lesioni neoplastiche originanti dalle cellule del tessuto nervoso stesso; tra queste lesioni, il sottotipo più frequente nell'età adulta è rappresentato dal glioblastoma (GBM), neoplasia primitiva derivante dalle cellule gliali del SNC, con scarsa tendenza a metastatizzare in altri organi o tessuti ma con elevatissima aggressività biologica locale e gravata da una prognosi generalmente molto infausta [10].

Tuttavia, GBM e SBM possono presentare caratteristiche simili alla MRI convenzionale, e pertanto sollevare importanti dubbi in termini di identificazione precoce e classificazione soprattutto negli osservatori meno esperti; questo a causa di alcuni elementi di imaging comuni ad entrambe le condizioni patologiche (enhancement circinato, necrosi e/emorragia, edema periferico, etc) [11-12]. In questi casi, ulteriori e più invasive indagini strumentali, che generalmente richiedono invasività, tempi e costi maggiori, si rendono spesso necessarie.

3. Scopo

Scopo di questo lavoro è confrontare le performance classificative di un sistema innovativo basato su tecnologie di visual recognition e cognitive computing (framework di deep learning open-source basato su PyTorch per dati di imaging medico denominato FuseMedML) [13] rispetto alle performance

classificative umane ottenute coinvolgendo nella valutazione dei dati MRI radiologi con diverso grado di esperienza nell'interpretazione delle immagini di neoplasie cerebrali.

4. Materiali e Metodi

In questa sezione sono illustrati le caratteristiche del dataset adoperato per lo scopo prefissato, le tecnologie utilizzate e il tipo di analisi statistica applicata ai dati ottenuti.

4.1. Dataset: caratteristiche ed organizzazione

Sono state retrospettivamente raccolte immagini MRI T1W volumetriche post-contrastografiche univoche ottenute a scopo preoperatorio per neuronavigazione, provenienti da 947 esami MRI differenti di altrettanti pazienti affetti da neoplasia cerebrale sospetta per GBM o SBM. Per ogni paziente, l'immagine MRI più rappresentativa della neoplasia cerebrale è stata selezionata e poi esportata dal formato DICOM al formato JPEG; per evitare possibili oversampling, lesioni con diametro massimo <5cm sono state campionate una volta nel livello più rappresentativo, mentre lesioni più grandi con diametro massimo >5cm sono state campionate due volte a livello delle due fette più rappresentative (distanza minima tra le immagini selezionate: 4cm). Pazienti con fattori di confondimento, quali severi artefatti (da movimento o da presenza di devices) o significative comorbidità del SNC, sono stati esclusi dalla successiva analisi; alla fine, è stato ottenuto un numero complessivo di 878 soggetti e 1412 immagini. La valutazione preliminare delle immagini e l'inclusione/esclusione dei pazienti è stata operata da un neuroradiologo con 10 anni di esperienza nel campo del neuroimaging oncologico.

Le immagini così ottenute sono poi state suddivise in due classi sulla base della diagnosi definitiva istopatologica: il 59% dei pazienti ha ricevuto diagnosi finale di GBM e il 41% di SBM da un tumore primario sconosciuto (come successivamente confermato tramite ulteriori indagini strumentali volte ad accertare la presenza di un tumore primitivo ignoto al momento dell'individuazione della lesione espansiva cerebrale); è da notare che, nonostante il rapporto finale del numero di soggetti mostri un lieve

squilibrio (1.4:1) in favore del GBM, tale differenza riflette l'effettiva diversa distribuzione epidemiologica di queste due condizioni patologiche [7-8].

4.2 FuseMedML

Come già illustrato in un precedente conference paper oggetto di pubblicazione [14], la prima fase di applicazione di modelli pre-addestrati basati sul transfer-learning nel riconoscimento di diversi tipi di tumore cerebrale a partire da immagini MRI è stata effettuata utilizzando il framework di deep learning per dati medici noto come FuseMedML [13]. Come anticipato nella sezione introduttiva, FuseMedML è un framework open-source basato su Python specificamente progettato per accelerare le scoperte in campo biomedico attraverso tecnologie avanzate di machine learning; la sua versione iniziale è basata su PyTorch.

Le immagini MRI selezionate sono state preliminarmente sottoposte ad un pre-processing con anonimizzazione, seguito da cropping, deskulling e resizing per eliminare aree indesiderate e rendere omogenee le dimensioni delle immagini [15-16].

Tra le reti neurali pre-addestrate, i migliori risultati sono stati ottenuti con VGG16, sviluppato dal Visual Graphics Group dell'Università di Oxford [17]. Questo modello raggiunge una precisione del 92,7% nei 5 migliori test di ImageNet e si compone di 13 livelli convoluzionali, 5 pooling layers e 3 dense layers. La versione di VGG16 che è stata utilizzata in questo caso è rappresentata dal sottomodulo Keras di TensorFlow. Un'illustrazione schematica del modello adoperato è mostrata in figura 1.

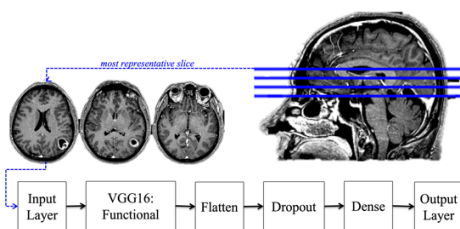


Figura 1. Architettura proposta per la rete neurale pre-addestrata.

Una volta stabilita l'architettura del modello, si è proceduto alla fase di apprendimento. Per la corretta validazione dei modelli, si è suddiviso il

dataset in sub-set: un training set (80% del campione) utilizzato in apprendimento per l'aggiornamento dei pesi della rete neurale, un test set (10% del campione) per la valutazione della capacità della rete di generalizzare su dati/immagini, ed un validation set (10% del campione) per la validazione su dati/immagini mai incontrati in apprendimento. Per l'addestramento del modello è stato necessario un tempo complessivo di 5'5''.

4.3 Performance umana

Le immagini MRI pre-processate come precedentemente descritto sono state adoperate per addestramento e successiva valutazione delle performance classificative umane. A tale scopo, sono stati coinvolti tre osservatori con diversa esperienza nel campo del neuroimaging oncologico:

- un neuroradiologo con 10 anni di esperienza nel neuroimaging e nel neuroimaging oncologico (osservatore 1);
- un radiologo generalista con 3 anni di esperienza di neuroimaging e con rudimenti di neuroimaging oncologico convenzionale/avanzato (osservatore 2);
- un radiologo generalista senza esperienza specifica nel campo del neuroimaging oncologico (osservatore 3).

Ai 3 osservatori è stato chiesto di classificare le immagini MRI nelle due categorie precedentemente descritte, rispettivamente GBM e SBM. Analogamente a quanto effettuato per l'addestramento della rete neurale, i 3 osservatori hanno proceduto autonomamente ed indipendentemente ad una fase di addestramento sul training set (80% del campione), disponendo liberamente del dataset per circa 1 mese; hanno poi eseguito una prima fase di testing sul test set (10% del campione), per affinare le competenze acquisite sul training set. Infine, è stata eseguita un'ultima fase di valutazione indipendente dei dati in cieco sul validation set (10% del campione, N=150 immagini).

4.4 Analisi statistica

I risultati sono stati analizzati in termini di sensibilità, specificità, valore predittivo positivo, valore predittivo negativo, ed accuratezza diagnostica. Per ciascun gruppo sono state poi ricavate le curve ROC, che misurano in termini di

area under the curve (AUC) e rappresentano graficamente l'efficienza del classificatore binario. Per l'analisi statistica è stato adoperato il software statistico aggiuntivo XLSTAT disponibile in Excel (XLstat package, 2019.7).

5. Risultati

Si è quindi proceduto con la raccolta dei risultati sia della rete neurale pre-addestrata che dei tre osservatori umani indipendenti di diversa esperienza. L'accuratezza diagnostica della rete neurale pre-addestrata ottenuta adoperando il framework di deep learning FuseMedML è stata pari al 96.45%. L'accuratezza diagnostica ottenuta interpolando le valutazioni dei 3 osservatori umani è stata pari al 84.33%, con valori nettamente superiori e quasi comparabili con la rete neurale pre-addestrata per l'osservatore 1 esperto (90.15%) rispetto all'osservatore 3 inesperto (75.39%). Le curve ROC che rappresentano graficamente la performance del modello di classificazione binaria proposto nei due setting (rete neurale pre-addestrata vs valutazione umana) sono mostrate in figura 2; la differenza statistica tra le due AUC (rispettivamente 0.99 e 0.81) risulta statisticamente significativa ($p < 0.000001$, $z = 5.0642$).

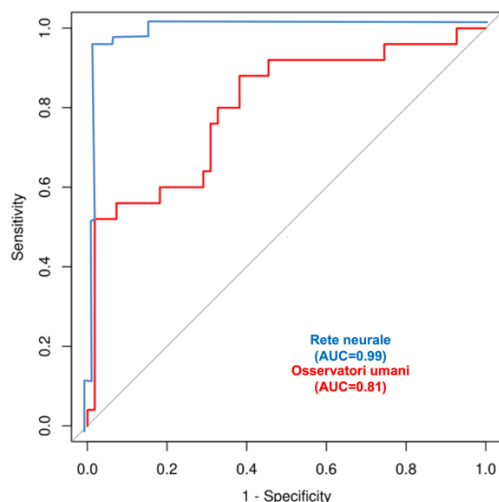


Figura 2. Curve ROC della rete neurale pre-addestrata (blu) e degli osservatori umani (rossa).

6. Discussione e Conclusioni

Le tecnologie di visual recognition e cognitive computing (come quella basata sul framework

open-source FuseMedML adoperato nel presente studio) nell'interpretazione dei dati di imaging medicale si sono dimostrate strumenti di grande affidabilità diagnostica, mostrando performance classificative spesso superiori rispetto ai classificatori umani in generale. In particolare, gli strumenti di intelligenza computazionale risultano globalmente più efficienti da addestrare e utilizzare grazie alla moderna potenza di calcolo. Infatti, il divario nelle performance risulta tanto maggiore quanto più inesperto è l'osservatore umano, sottolineando come i tempi e le risorse per la formazione necessaria ad ottenere prestazioni diagnostiche soddisfacenti possano essere rapidamente implementati con un tale supporto tecnologico.

In conclusione, non solo le tecnologie computer-assisted e le tecniche orientate all'intelligenza computazionale hanno in sé il potenziale di superare le performances umane in termini di tempo, accuratezza e affidabilità su dati da classificare [18-19], ma adeguano rapidamente le proprie performances all'aumentare del pool di dati da analizzare e possono potenzialmente essere applicati in tempi relativamente brevi a corpora di ricerca molto più ampi di quelli comunemente accessibili agli osservatori umani, per quanto esperti [20-22].

7. Ringraziamenti

Grazie ai Proff. A. Brunetti e E. M. Covelli per il prezioso contributo operativo nella raccolta e trasmissione dei dati oggetto di analisi.

Grazie ai Dott.ri F. Giordano, R. Ronza e A. Tucci per la collaborazione nella valutazione ed interpretazione dei dati oggetto di analisi.

8. Bibliografia

- [1] Faehndrich J, Weidauer S, Pilatus U, et al. Neuroradiological viewpoint on the diagnostics of space-occupying brain lesions. *Clinical Neuroradiology*. 2011 Sep;21(3):123-139.
- [2] Due-Tønnessen, P., Pinho, M.C., Emblem, K.E. et al. The Impact of MRI Features and Observer Confidence on the Treatment Decision-Making for Patients with Untreated Glioma. *Sci Rep* 2019; 9, 19898.
- [3] Meng W, Sun Y, Qian H, Chen X, Yu Q, Abiyasi N, Yan S, Peng H, Zhang H and Zhang X. Computer-Aided Diagnosis

- Evaluation of the Correlation Between Magnetic Resonance Imaging With Molecular Subtypes in Breast Cancer. *Front. Oncol.* 2021; 11:693339.
- [4] Wang S, Burt K, Turkbey B, Choyke P, Summers RM. Computer aided-diagnosis of prostate cancer on multiparametric MRI: a technical review of current research. *Biomed Res Int.* 2014; 2014:789561.
- [5] Kluk J, Ogiela MR. AI Approaches in Computer-Aided Diagnosis and Recognition of Neoplastic Changes in MRI Brain Images. *Applied Sciences.* 2022; 12(23):11880.
- [6] Hsieh KLC, Tsai RJ, Teng YC, Lo CM (2017) Effect of a computer-aided diagnosis system on radiologists' performance in grading gliomas with MRI. *PLOS ONE* 12(2): e0171342.
- [7] Pizzini, F.B., Thust, S., Jäger, H.R. (2019). Clinical Presentations, Differential Diagnosis, and Imaging Work-Up of Cerebral Mass Lesions. In: Barkhof, F., Jäger, H., Thurnher, M., Rovira, À. (eds) *Clinical Neuroradiology*. Springer, Cham.
- [8] Yamamoto, A.K. (2019). Imaging Techniques for Neurosurgical Planning of Tumor Resection. In: Barkhof, F., Jäger, H., Thurnher, M., Rovira, À. (eds) *Clinical Neuroradiology*. Springer, Cham.
- [9] Pope WB. Brain metastases: neuroimaging. *Handb Clin Neurol.* 2018;149:89-112.
- [10] Wirsching, HG., Weller, M. (2017). Glioblastoma. In: Moliterno Gunel, J., Piepmeyer, J., Baehring, J. (eds) *Malignant Brain Tumors*. Springer, Cham.
- [11] Fordham AJ, Hacherl CC, Patel N, Jones K, Myers B, Abraham M, Gendreau J. Differentiating Glioblastomas from Solitary Brain Metastases: An Update on the Current Literature of Advanced Imaging Modalities. *Cancers (Basel).* 2021 Jun 13;13(12):2960.
- [12] Elefante A, Russo C. (2022) Ruolo dell'imaging nei tumori cerebrali intrassiali dell'adulto In: *Neuroradiologia, per Studenti e Medici di Medicina Generale.* (eds) Edizioni Idelson Gnocchi
- [13] IBM Research, Haifa: FuseMedML. [https://github.com/IBM/fuse-med-ml\(2021\)](https://github.com/IBM/fuse-med-ml(2021))
- [14] Russo C, Maresca P, Marinelli A. Cognitive computing tools for identification and classification of brain tumors starting from Magnetic Resonance Imaging: preliminary results. 2022 IEEE 21st Mediterranean Electrotechnical Conference (MELECON) | 978-1-6654-4280-0/22.
- [15] Kang, J.; Ullah, Z.; Gwak, J. MRI-Based Brain Tumor Classification Using Ensemble of Deep Features and Machine Learning Classifiers. *Sensors* 2021, 21, 2222.
- [16] Zhang et al. Finding Extreme Points in Contours with OpenCV. In *PyImageSearch*. Available online :<https://www.pyimagesearch.com/2016/04/11/finding-extremepoints-in-contours-with-opencv>
- [17] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems* 2017, 60 (6), 84-90. Previously published in the Proceedings of the 25th International Conference on Neural Information Processing Systems, Lake Tahoe, NV Dec. 2012, 1097-1105
- [18] Badža MM, Barjaktarović MČ. Classification of Brain Tumors from MRI Images Using a Convolutional Neural Network. *Applied Sciences.* 2020; 10(6):1999.
- [19] Xu J, Meng Y, Qiu K, Topatana W, Li S, Wei C, Chen T, Chen M, Ding Z and Niu G (2022) Applications of Artificial Intelligence Based on Medical Imaging in Glioma: Current State and Future Challenges. *Front. Oncol.* 12:892056.
- [20] Iqbal S, Khan MUG, Saba T, Rehman A. Computer-assisted brain tumor type discrimination using magnetic resonance imaging features. *Biomed Eng Lett.* 2017 Oct 4;8(1):5-28.
- [21] De Man R, Gang GJ, Li X, Wang G. Comparison of deep learning and human observer performance for detection and characterization of simulated lesions. *J Med Imaging (Bellingham).* 2019 Apr;6(2):025503.
- [22] Sunwoo L, Kim YJ, Choi SH, Kim KG, Kang JH, et al. (2017) Computer-aided detection of brain metastasis on 3D MR imaging: Observer performance study. *PLOS ONE* 12(6): e0178265.