

Relazione Computer Vision

Mappatura dei nevi automatizzata con approcci di Deep Learning

Marta Pibiri, Riccardo Balia

A.A. 2018 / 2019

1 INTRODUZIONE

Negli ultimi anni si è registrato un'incidenza maggiore di casi di melanomi. Il fenomeno ha attirato su di sé l'interesse dei ricercatori, che intendono sviluppare un sistema computerizzato che renda più veloce la diagnosi e che faciliti il lavoro dei medici. Questa metodica possiamo chiamarla *mappatura dei nevi automatizzata*, e può far uso di tecniche tipicamente usate nella Computer Vision. L'elaborato proposto vuole esplorare questo campo, in particolare nella distinzione dei nevi comuni e dei melanomi. Per poter realizzare il progetto, è stato sfruttato il dataset, reso disponibile nel paper PH2 - *A dermoscopic image database for research and benchmarking* per fini di ricerca e test. Tale dataset chiamato PH2 contiene 200 immagini dermoscopiche di nevi tra cui 80 comuni, 80 atipici e 40 melanomi. Ad ogni immagine, dei dermatologi hanno associate delle informazioni sul loro aspetto, la classe di appartenenza e le ROI che evidenziano le regioni lesionate.

2 PROBLEMI AFFRONTATI

L'obiettivo principale è stato riconoscere nevi comuni (classe: *Common Nevus*) e melanomi (classe: *Melanoma*) e effettuare una classificazione binaria utilizzando solo le immagini appartenenti a queste due classi. I nevi *Atipici* non sono stati presi in considerazione. Il dataset utilizzato si è quindi ridotto da 200 a 120 foto di cui 80 nevi comuni e 40 melanomi. Altre informazioni presenti nel dataset non sono state sfruttate, come ad esempio le ROI, il colore e altre informazioni relative alla forma del nevo come asimmetria. La realizzazione del progetto ha previsto l'utilizzo di MATLAB come strumento di sviluppo per poter utilizzare package utili, soprattutto quelli relativi alle reti neurali. Prima di poter scegliere con certezza quale fosse l'applicazione migliore, vi è stata una fase di ricerca e sperimentazione. Inizialmente, per l'estrazione delle feature si è provato l'estrazione di LBP: questo dava risultati poco accurati. La classificazione finale restituiva un'accuratezza massima del 75%. Utilizzando le reti neurali per l'estrazione delle feature, i risultati ottenuti sono stati nettamente superiori (intorno al 90%). Anche nel momento in cui sono state utilizzate le reti neurali, sono state effettuate alcune scelte relativamente ai layer: quest'ultima è stata influenzata da alcune prove e da approfondimenti

sui tutorial presenti dalla documentazione di MATLAB stesso. La maggior parte delle scelte per la classificazione sono state prese considerando la dimensione ridotta del dataset; le SVM sono state addestrate su tutto il dataset usando come tecnica di suddivisione la leave-one-out.

3 METODO

Lo svolgimento dell'elaborato si può strutturare in più fasi :

- Estrazione delle immagini dal dataset
- Creazione del augmented set
- Estrazione feature
- Classificazione
- Voting
- Stampa dei risultati

3.1 Estrazione delle immagini dal dataset

Tale fase prevede l'estrazione delle immagini e della propria label, ovvero classe di appartenenza, in un *imageDataStore*. Prima vengono copiate le immagini dei nei comuni e dei melanomi in due cartelle separate (se queste cartelle non esistono ancora), successivamente viene creato il data store.

3.2 Creazione del augmented set

Per poter incrementare il dataset, è stato scelto di creare e inserire un set di immagini augmented per ogni immagine. In particolare sono state applicate nove variazioni di rotazione; l'immagine originale e le sue varianti, popolano una riga all'interno della struttura *structIm*. La scelta di applicare solo rotazione è nata dopo una serie di esperimenti e prove: applicando un incremento della scala dell'immagine la classificazione risultava peggiore, applicando un decremento invece vi erano problemi con la fase dell'estrazione delle feature con reti neurali, dovuto al fatto che la rete richiede immagini di una dimensione specifica che l'immagine ridotta non poteva soddisfare.

3.3 Estrazione feature

In questo passo, viene sfruttata la struttura *structIm* appena creata e le reti neurali *ResNet18*, *AlexNet*, *GoogLeNet*. Si tratta di reti neurali pre-addestrate per l'estrazione di feature, e sono state scaricate e installate come estensioni di MATLAB. Sono stati scelti i seguenti layers per poter estrarre le feature:

- Per ResNet18, 'pool5'
- Per AlexNet, 'fc7'
- Per GoogLeNet, 'loss3-classifier'.

La struttura *structIm* contiene attualmente l'immagine originale preso dal dataset e una serie di nove trasformazioni. Ciascuna di queste viene nuovamente trasformata per creare un *augmentedImageDatastore* che ha lo scopo principale di ridimensionare le immagini con misure richieste da ogni specifica rete. Il ri-dimensionamento viene effettuato con un crop sulla parte centrale dell'immagine. Nonostante questo rischi di escludere parti utili dall'immagine da cui estrarre le features, dei test hanno mostrato che questo approccio consente di estrarre features più rappresentative rispetto a una riduzione della scala dell'immagine. Le immagini a questo punto vengono passate alla rete neurale in modo da ottenere un vettore di feature, ogni riga

corrisponde ad una immagine: nella fase successiva bisogna considerare ogni dieci righe si fa riferimento a una diversa immagine del dataset e alla trasformazione associata.

3.4 Classificazione

Per la classificazione si è scelto un classificatore *SVM lineare*, che è stato addestrato sulle feature estratte dalle reti neurali. A causa delle dimensioni ridotte del dataset a nostra disposizione, si è optato per una sua suddivisione con la metodologia di *cross validation leave-one-out*. Questo prevede che il test set contenga un solo elemento, il training set invece contiene tutti gli elementi del dataset fatta eccezione per quello già incluso nel test set e tutte le trasformazioni associate. Il processo è iterato per ogni elemento del dataset. Per quanto oneroso, ha consentito un addestramento su tutti i dati a disposizione, e una classificazione ogni elemento su SVM quanto più addestrate possibili.

3.5 Voting

Prese in considerazione le classificazione ottenute con le tre diverse reti neurali, esse sono state utilizzate per il *voting*: una immagine veniva classificata come *CommonNevus* se la maggior parte delle reti la definiva tale, o viceversa per *Melanoma*. Il voting restituisce una nuova classificazione delle immagini.

3.6 Stampa dei risultati

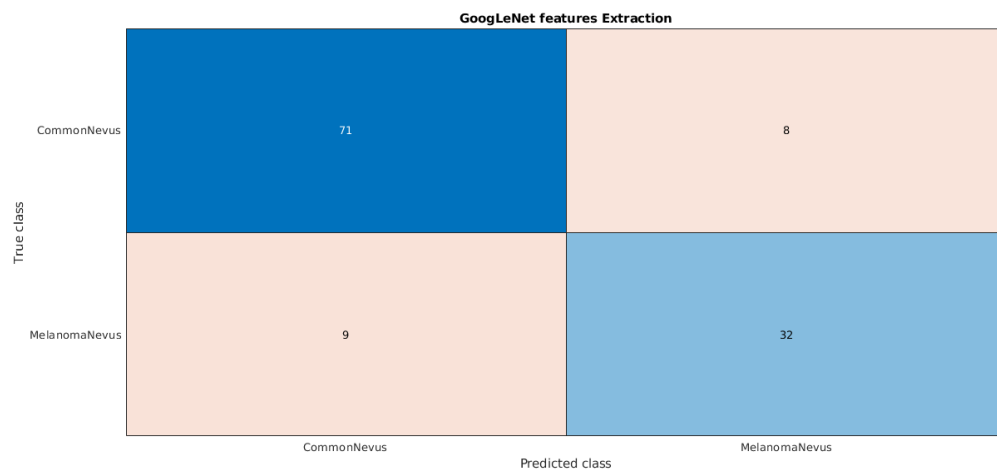
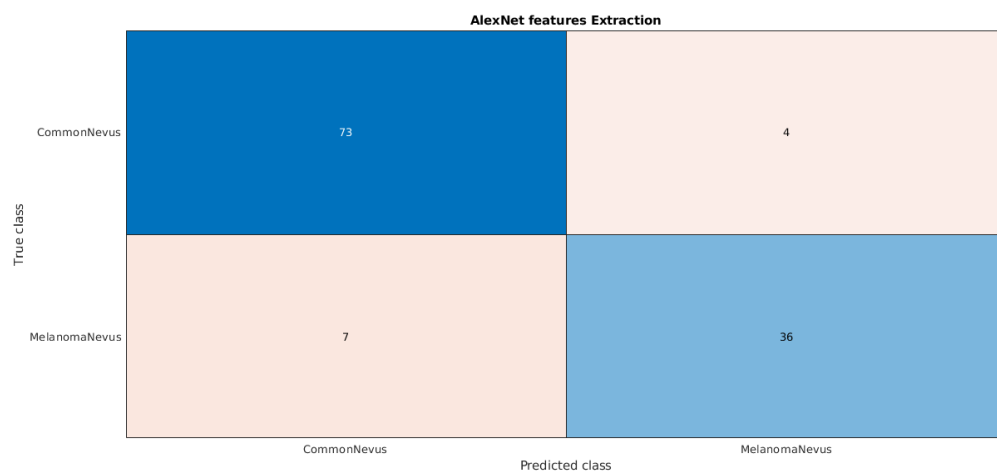
La visualizzazione dei risultati è fatta tramite matrice di confusione e tabelle. Inoltre sono state stampate le immagini classificate in maniera scorretta nella fase di voting.

4 RISULTATI OTTENUTI

4.1 Matrici di confusione

	Predicted	
	Positive	Negative
Actual True	TP	FN
Actual False	FP	TN

Per analizzare i risultati ottenuti, sono state realizzate le matrici di confusione corrispondenti ad ogni metodica utilizzata. La matrice di confusione rappresenta l'accuratezza del metodo: ogni colonna della matrice rappresenta i valori predetti, mentre ogni riga rappresenta i valori reali.



		Voting	
True class	CommonNevus	74	4
	MelanomaNevus	6	36
		CommonNevus	MelanomaNevus
		Predicted class	

Dopo aver ottenuto la matrice di confusione, è stata calcolata l'*accuratezza* per ogni metodica. Essa indica in percentuali quanti record sono stati classificati in modo corretto ed è ottenibile come somma degli elementi sulla diagonale principale principale rapportata al totale di tutti i valori della matrice.

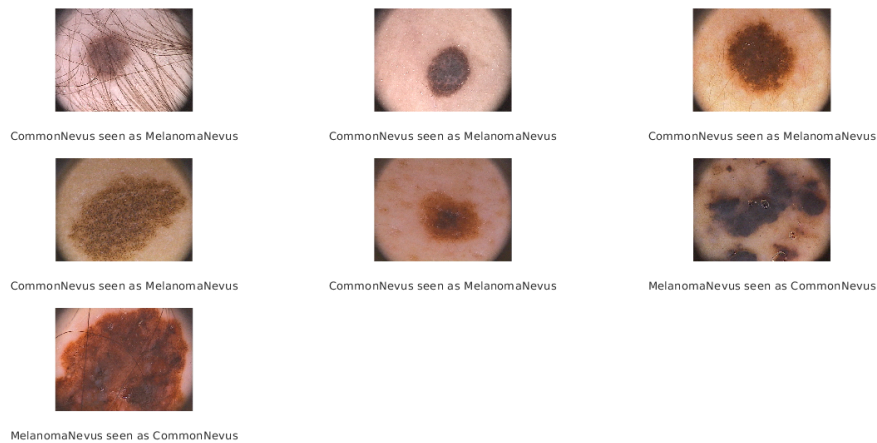
$$A = \frac{TN + TP}{TN + TP + FP + FN} \quad (4.1)$$

Metodo	Accuratezza
Features by ResNet	0.94166
Features by AlexNet	0.90833
Features bu GoogLeNet	0.85833
Voting	0.91666

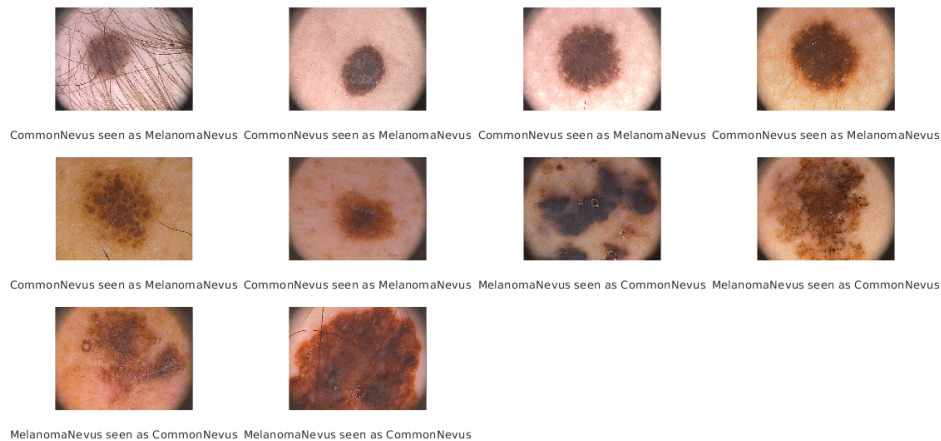
Table 4.1: tabella rappresentante l'accuratezza associata al metodo usato

Osservando tale tabella, possiamo individuare come metodologia migliore, tra quelle presentate, l'estrazione di feature con ResNet e come seconda la metodologia Voting.

4.2 Classificazione errata - ResNet



4.3 Classificazione errata - Voting



Osservando le immagini classificate scorrettamente, si possono supporre le motivazioni e si può denotare che più o meno sono sempre le stesse immagini a non essere classificate bene. In particolare, sono le immagini che presentano elementi estranei (si osservi la prima immagine in entrambi gli insiemi) o che rappresentano nevus o melanomi in maniera differente, inteso come inquadratura e inclinazione della foto.

5 CONCLUSIONI

La soluzione adottata ha reso buoni risultati nonostante le varie problematiche presenti o nate durante la realizzazione dell'elaborato. Nonostante esse, si è potuto sottolineare come l'estrazione delle features tramite reti neurali ha reso migliore la classificazione finale rispetto a estrazione di feature tramite LBP. Anche la scelta dell'addestramento con suddivisione di leave-one-out risulta vincente in questa situazione per via della dimensione del dataset. Probabilmente anche la qualità di alcune foto non è eccellente come abbiamo potuto constatare dalle foto non classificate bene ma anche trovare dataset ricchi di fotografie "ottimali" non è semplice; anche il paper PH2 - A dermoscopic image database for research and benchmarking riporta lo stesso problema. Nella ricerca non è semplice trovare o utilizzare dataset di foto mediche perché spesso non sono accessibili. In ogni caso, tale tipologie di ricerca sono importanti sia nel campo della computer vision sia nel campo medico perché le metodiche trovate potrebbero facilitare per davvero l'individuazione di patologie e problemi, nel nostro caso di melanomi e per questo motivo è giusto andare avanti con la ricerca e la sperimentazione.