

Transfer learning per selezionare gli informative-frame in una laringoscopia usando learned features

Tesi di Laurea in Ingegneria Informatica

Laureando: Bortolin Simone

Relatore: Nanni Loris

A.A. 2020/2021

- 1 Introduzione
- 2 Transfer Learning
- 3 Preprocessing
- 4 Data augmentation
- 5 Risultati

Introduzione

Dataset

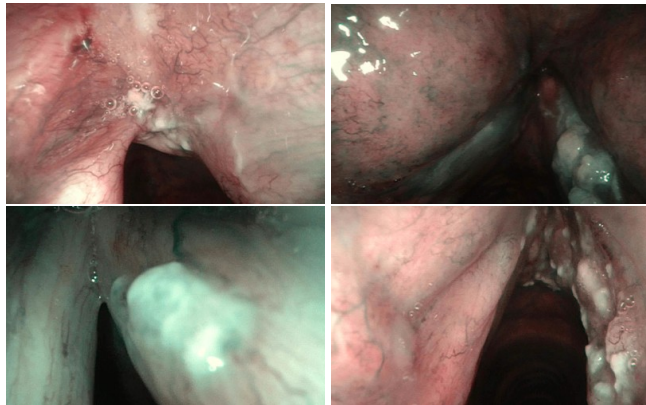
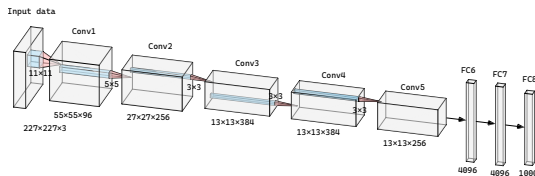


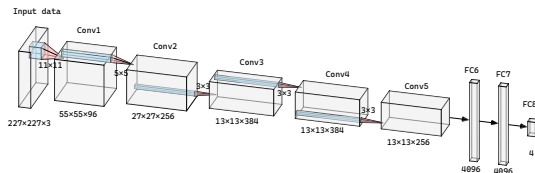
Figura: Esempio di fotogrammi endoscopici laringei

Transfer Learning

Con Transfer Learning si intende il riutilizzo di una rete preaddestrate, per altri scopi. Si parte da una rete preaddestrata, come la seguente:



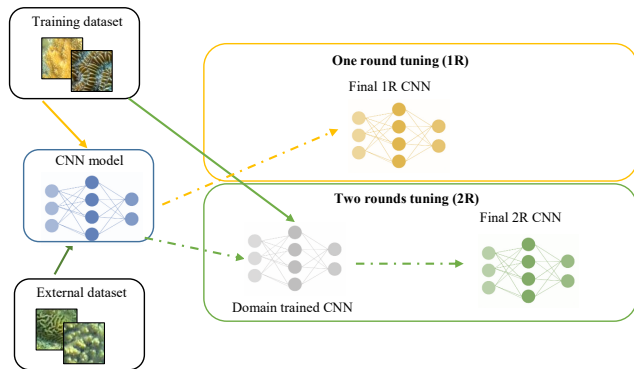
E la si riaddestra per con un altro dataset, adeguando i livelli finali:



Transfer Learning

Two Round tuning

Il Two Round tuning è una tecnica che effettua due Fine Tuning per l'allenamento della rete, di cui il primo basato su un dataset simile a quello di destinazione.



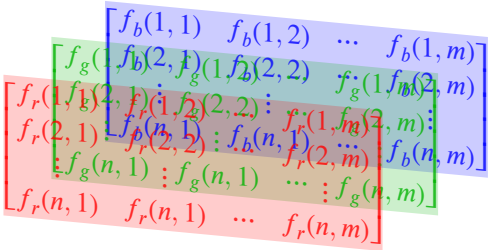
Preprocessing

Il preprocessing è l'attività che serve a migliorare le immagini, attività incentivata a migliorare i risultati nella fase di addestramento. In particolare si useranno i metodi:

- Ottimizzazione del contrasto (tramite valori hardcoded, dinamici e CLAHE)
- Correzione Gamma

Data augmentation

Partendo da una immagine raster, per esempio rappresentata nel formato tensoriale:

$$f(x, y, k) = \begin{bmatrix} f_r(1, 1) & f_r(1, 2) & \dots & f_r(1, m) \\ f_r(2, 1) & f_r(2, 2) & \dots & f_r(2, m) \\ \vdots & \vdots & \ddots & \vdots \\ f_r(n, 1) & f_r(n, 2) & \dots & f_r(n, m) \end{bmatrix}$$


A cui vengono applicate semplici trasformazioni geometriche come:

- Riflessione
- Rotazione
- Ridimensionamento
- Traslazione

Data augmentation

Discrete cosine transform

In aggiunta ai suddetti metodi, si può lavorare attraverso la DCT a metodi per modificare i valori delle frequenze dell'immagine.



Dove:

$$\text{DCTimage}(x, y) = \alpha(x)\alpha(y) \sum_{(p,q)=(1,1)}^{(n,m)} \text{Image}(p, q) \cos\left(\frac{\pi p(x-1)}{2n}\right) \cos\left(\frac{\pi q(y-1)}{2m}\right)$$

Per ogni $(x, y) \in [0..n] \times [0..m]$, dove n, m rappresentano la dimensione dell'immagine. Similmente la DCT tipo III o IDCT è definita come:

$$\text{IDCTimage}(x, y) = \sum_{(p,q)=(1,1)}^{(n,m)} \alpha(p)\alpha(q) \text{DCTimage}(p, q) \cos\left(\frac{\pi p(x-1)}{2n}\right) \cos\left(\frac{\pi q(y-1)}{2m}\right)$$

Risultati

	1R		2R	
	(1)	(2)	(1)	(2)
Semplice DA	95.4%	100%	97.5%	100%
Filtro Contrasto Semplice	96.2%	100%	96.2%	100%
Filtro Contrasto con media e STD	89.6%	91.7%	92.9%	100%
CLAHE	63.7%	45.8%		
Correzione Gamma	95.0%	100%		
Correzione Gamma e CLAHE	46.0%	0%		
Noise	96.7%	100%	95.4%	95.0%
DCT e Noise	94.6%	100%	93.8%	100%

- 1 Divisione corretta nelle 4 classi
- 2 Riconoscimento dei frame Informativi

Grazie!