



UNIVERSITA DEGLI STUDI DI PISA

Tesi di Laurea
Scuola di Ingegneria
Dipartimento di Ingegneria dell'Informazione
A.A. 2018/2019

Sviluppo di una rete neurale per il riconoscimento
delle differenze percettive su un set di immagini alla
ricerca della minima risoluzione necessaria per la
loro individuazione

Candidato
Coccomini Davide

Relatori
Pistolesi Francesco, Lazzerini Beatrice

Indice

1	Introduzione	3
2	Colorimetria	3
2.1	Lo spettro visibile	3
2.2	Percezione del colore	4
2.3	Modello CIE XYZ	5
2.4	Modello CIE L*a*b*	6
3	Reti Neurali	7
3.1	Reti neurali e cervello umano	7
3.1.1	Dal neurone biologico a quello artificiale	7
3.2	Perceptron	8
3.2.1	Addestramento del perceptron	9
4	Obiettivo	10
5	Svolgimento	11
5.1	Creazione del dataset	11
5.1.1	Codice C++	13
5.1.2	Esempi	15
5.1.3	Conversione nello spazio colore CIE L*a*b*	15
6	Addestramento della rete	16
7	Test	16
8	Conclusioni	17

1 Introduzione

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrum exercitationem ullam corporis suscipit laboriosam, nisi ut aliquid ex ea commodi consequatur. Quis aute iure reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint obcaecat cupiditat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

2 Colorimetria

La colorimetria è la disciplina che si occupa di normalizzare la misurazione del colore attraverso lo studio dei modelli di colore.

2.1 Lo spettro visibile

Lo “spettro del visibile” si trova nella parte centrale dello spettro ottico, il quale comprende anche i raggi infrarossi e quelli ultravioletti. E’ composto da quella parte di spettro elettromagnetico che include tutte le colorazioni percepite dall’occhio umano, partendo dal rosso fino ad arrivare al viola. La colorimetria utilizza la percentuale della luce incidente che è stata riflessa (nell’intervallo del visibile (400-700 nm) per descrivere il colore dell’oggetto. Applicazioni particolari quali la misurazione della fluorescenza, del bianco ed i colori mimetici prendono in considerazione anche le radiazioni UV (350-400nm) e NIR (700-1300nm). Ciascun oggetto colorato viene pertanto definito da una curva di riflettanza, similmente alle impronte digitali nell’uomo.

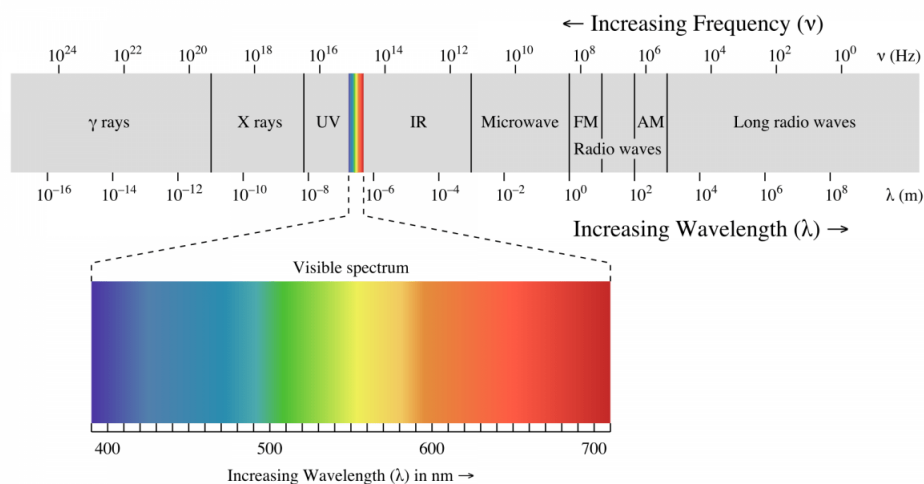


Figura 1: "Lo spettro visibile"

2.2 Percezione del colore

Il colore nasce dalla luce. La luce che colpisce un oggetto viene parzialmente assorbita a seconda del colore. La parte non assorbita viene riflessa e trasmessa ai recettori cromatici all'interno dell'occhio umano. Questi ultimi trasformano la luce assorbita in impulsi che percorrono le vie nervose fino a raggiungere il cervello, dove vengono interpretati: nasce così un'impressione cromatica. Dal punto di vista prettamente biologico il colore si genera pertanto nell'occhio dell'osservatore e costituisce un'impressione sensoriale.

A proposito di impressione sensoriale: ciascun individuo "percepisce" il colore in modo differente. Tale fenomeno non è riconducibile solamente al fatto che non esistono mai due occhi uguali tra loro. Anche l'interpretazione del colore varia infatti da individuo ad individuo. Perfino la stessa persona può percepire differentemente il colore in momenti diversi ed in base allo stato d'animo. Il colore stesso può pertanto generare sensazioni differenti.

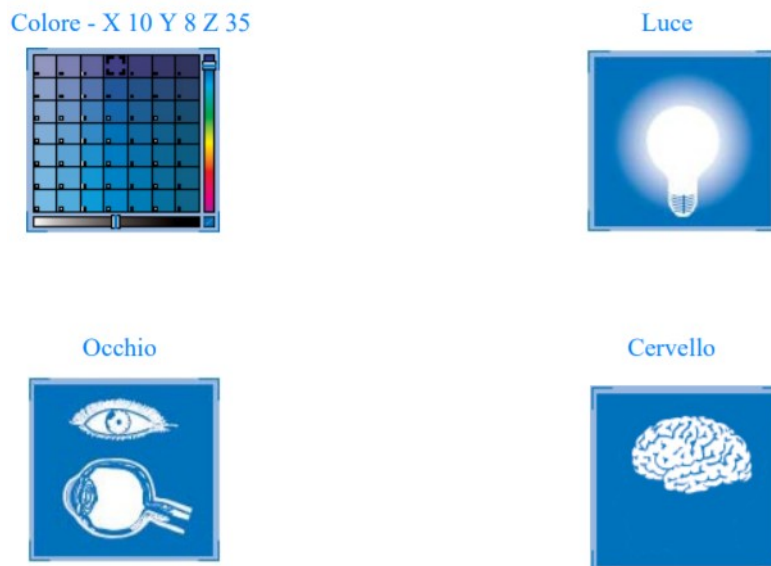


Figura 2: "Percezione del colore"

2.3 Modello CIE XYZ

Il modello CIE XYZ rappresenta tutti i colori caratterizzati da tre parametri: luminosità, tinta e purezza. Tinta e purezza rappresentano la cromaticità del colore. Del solido viene solitamente rappresentata soltanto la sezione secondo il piano XY, in cui X ed Y indicano la cromaticità e la luminanza non viene rappresentata.

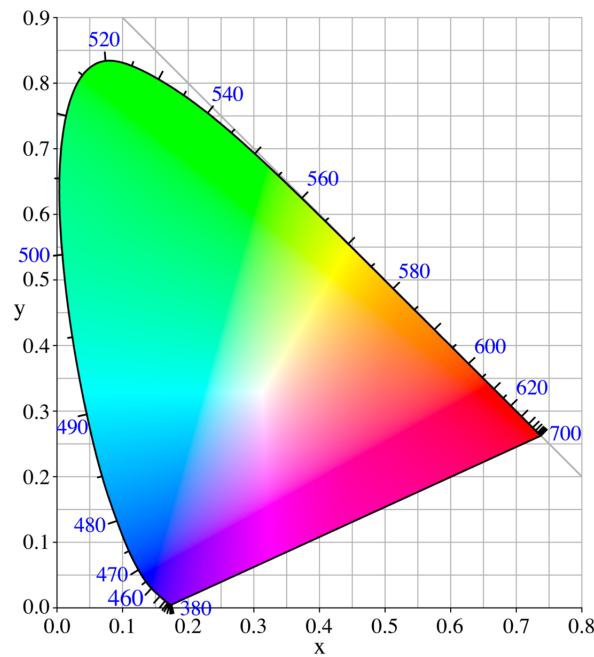


Figura 3: "Spazio colore CIE XYZ"

Questo spazio colore si basa sui valori di tristimolo, cioè dei parametri che definiscono il modo in cui l'essere umano percepisce i colori. I valori tristimolo di un colore con una distribuzione di potenza spettrale $I(\lambda)$ sono date in termini di un osservatore standard da:

$$X = \int_0^\infty I(\lambda) \bar{x}(\lambda) d\lambda$$

$$Y = \int_0^\infty I(\lambda) \bar{y}(\lambda) d\lambda$$

$$Z = \int_0^\infty I(\lambda) \bar{z}(\lambda) d\lambda$$

2.4 Modello CIE $L^*a^*b^*$

Il Modello CIE $L^*a^*b^*$, è un'evoluzione del modello CIE XYZ. E' un modello tridimensionale che si sviluppa lungo tre assi ortogonali. Due assi sul piano orizzontale riguardano la cromaticità: l'asse a si estende nel verde (-a) al rosso (+a) e l'asse b dal blu (-b) al giallo (+b); un asse verticale riguarda la luminosità o luminanza (L) che diminuisce dal basso verso l'alto. Questo particolare spazio colore include tutti i colori percepibili e quindi anche tutto il gamut degli spazi RGB e CMYK ed è indipendente dal dispositivo che li rappresenta. Essendo quindi molto più simile al modo in cui l'essere umano riesce a percepire i colori, questo spazio colore si presta particolarmente bene quando si devono individuare le differenze percettive all'interno di più immagini.

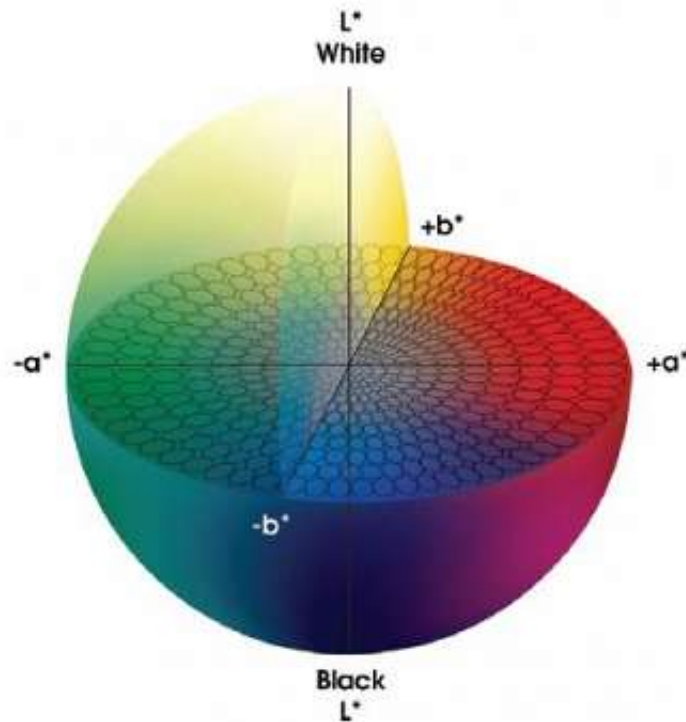


Figura 4: "Spazio colore CIE $L^*a^*b^*$ "

3 Reti Neurali

3.1 Reti neurali e cervello umano

Le reti neurali artificiali sono nate per riprodurre attività tipiche del cervello umano come la percezione di immagini, il riconoscimento di forme, la comprensione del linguaggio, il coordinamento senso-motorio e molto altro. A tale scopo si sono studiate le caratteristiche del cervello umano. Nel sistema nervoso esistono miliardi di neuroni (cellule nervose). Un neurone è formato da un corpo cellulare e da molti prolungamenti ramificati, detti dendriti, attraverso i quali il neurone riceve segnali elettrici da altri neuroni. Ogni neurone ha anche un prolungamento filamentoso chiamato assone. All'estremità l'assone si ramifica formando terminali attraverso i quali i segnali elettrici vengono trasmessi ad altre cellule (ad esempio ai dendriti di altri neuroni). Tra un terminale di un assone e la cellula ricevente esiste uno spazio. I segnali superano questo spazio per mezzo di sostanze chimiche dette neurotrasmettitori. Il punto di connessione tra terminale e dendrite è detto sinapsi. Un neurone si "attiva", cioè trasmette un impulso elettrico lungo il suo assone quando si verifica una differenza di potenziale elettrico tra l'interno e l'esterno della cellula. L'impulso elettrico provoca la liberazione di un neurotrasmettitore dai terminali dell'assone, che a loro volta possono, ad esempio, influenzare altri neuroni.

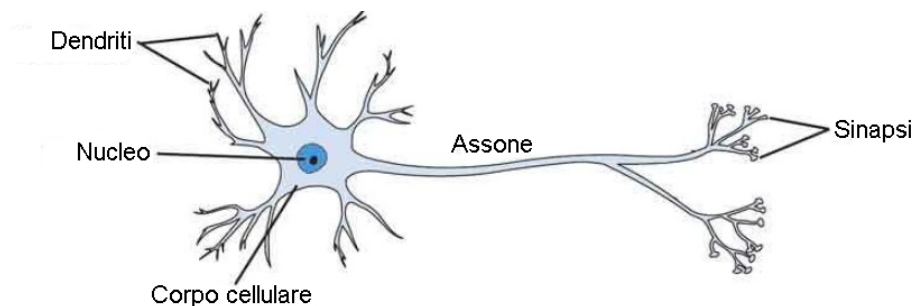


Figura 5: "Neurone biologico"

3.1.1 Dal neurone biologico a quello artificiale

Per riprodurre artificialmente il cervello umano occorre realizzare una rete di elementi molto semplici che sia una struttura distribuita, massicciamente parallela, capace di apprendere e quindi di generalizzare. Tipicamente, il neurone artificiale ha molti ingressi ed una sola uscita. Ogni ingresso ha associato un peso, che determina la conducibilità del canale di ingresso. L'attivazione del neurone è una funzione della somma pesata degli ingressi. Il metodo più usato per addestrare una rete neurale consiste nel presentare in ingresso alla rete un insieme di esempi (training set). La risposta fornita

dalla rete per ogni esempio viene confrontata con la risposta desiderata, si valuta la differenza (errore) fra le due e, in base a tale differenza, si aggiustano i pesi.

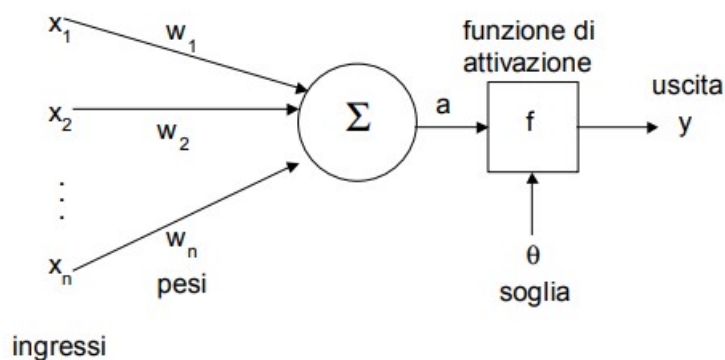


Figura 6: "Modello di neurone"

Abbiamo n canali di ingresso x_1, \dots, x_n , a ciascuno dei quali è associato un peso. I pesi w_i sono numeri reali che riproducono le sinapsi. Se $w_i > 0$, il canale è detto eccitatorio, se $w_i < 0$, il canale è inibitorio. Il valore assoluto di un peso rappresenta la forza della connessione. L'uscita, cioè il segnale con cui il neurone trasmette la sua attività all'esterno, è calcolata applicando la funzione di attivazione alla somma pesata degli ingressi.

$$y = f\left(\sum_{i=1}^n w_i x_i\right)$$

3.2 Perceptron

In molte applicazioni si incontrano problemi di classificazione di un insieme di oggetti, cioè occorre associare ogni oggetto alla classe corretta. Supponiamo di voler classificare in due classi distinte oggetti rappresentati mediante punti nel piano. Se le due classi sono linearmente separabili, possiamo usare una rete neurale che approssima una retta di separazione tra le due classi. Un oggetto sarà quindi classificato rappresentandolo come punto nel piano ed assegnandolo a quella delle due classi individuata dal semipiano in cui cade il punto. Tale classificazione può essere facilmente ottenuta addestrando una rete neurale con due ingressi ed un solo neurone di uscita con funzione di attivazione a soglia: Tale rete è un esempio di perceptron, costituito da più ingressi confluenti in un neurone di uscita.

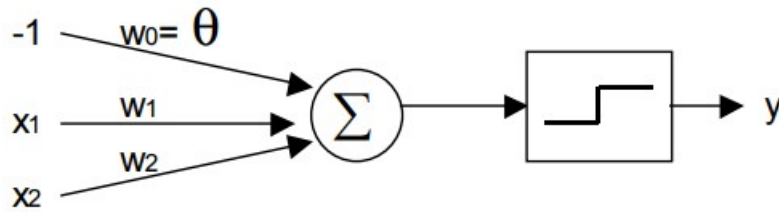


Figura 7: "Perceptron"

3.2.1 Addestramento del perceptron

Per addestrare un perceptron bisogna realizzare un training set adeguato per il problema che si vuole risolvere. Dopo aver predisposto il dataset è necessario seguire i seguenti passi:

1. si inizializzano i pesi w_i con valori casuali;
2. si presenta alla rete un ingresso x_k insieme al valore t_k desiderato in uscita;
3. si calcola la risposta y_k della rete e si aggiornano i pesi mediante la delta rule;
4. si ripete il ciclo dal passo 2, finché la risposta della rete non risulti soddisfacente.

In particolare la delta rule è una regola usata per aggiustare i valori dei pesi di un neurone andando alla ricerca di quelli più adeguati. Dati t ed y , rispettivamente, l'uscita desiderata e l'uscita neurale, l'errore δ è dato da:

$$\delta = t - y$$

Fissato un numero reale η compreso tra 0 e 1 detto learning rate, la delta rule stabilisce che la variazione del generico peso Δw_i è:

$$\Delta w_i = \eta \delta x_i$$

Dopo l'addestramento la rete viene testata controllandone il comportamento su un insieme di dati, detto test set, non utilizzati durante la fase di training. La fase di test ha quindi lo scopo di valutare la capacità di generalizzazione della rete neurale.

4 Obiettivo

L'individuazione delle differenze percettive all'interno di un dataset di immagini è uno dei grandi problemi dell'industria della colorimetria moderna che ha la necessità di emulare il comportamento umano nella percezione dei colori. E' infatti spesso necessario riprodurre un materiale o un tessuto partendo da un originale, nel tentativo di emularlo il più fedelmente possibile. Per fare ciò, sono stati realizzati macchinari in grado di acquisire immagini ad alta definizione dei soggetti originali e di quelli riprodotti così da poterli confrontare attraverso indici matematici più o meno precisi (ad esempio SSIM) nel tentativo di identificare eventuali differenze percettive. Ovviamente, sia per l'acquisizione che per il confronto tra immagini ad alta definizione, sono necessari tempi e costi non indifferenti che sarebbero abbattuti se questi processi potessero avvenire a risoluzioni più basse. Lo scopo di questa ricerca è quindi quello di realizzare una rete neurale che sia in grado di identificare le differenze tra le immagini e sfruttarla per capire qual'è la minima risoluzione e qualità che le immagini devono possedere affinché queste differenze siano individuabili.

5 Svolgimento

5.1 Creazione del dataset

Per poter realizzare la rete è necessario creare un dataset significativo affinché questa possa essere addestrata e testata. Per fare ciò sono state prese in considerazione 40 immagini ad una risoluzione di 4K sulle quali sono state successivamente fatte delle elaborazioni.

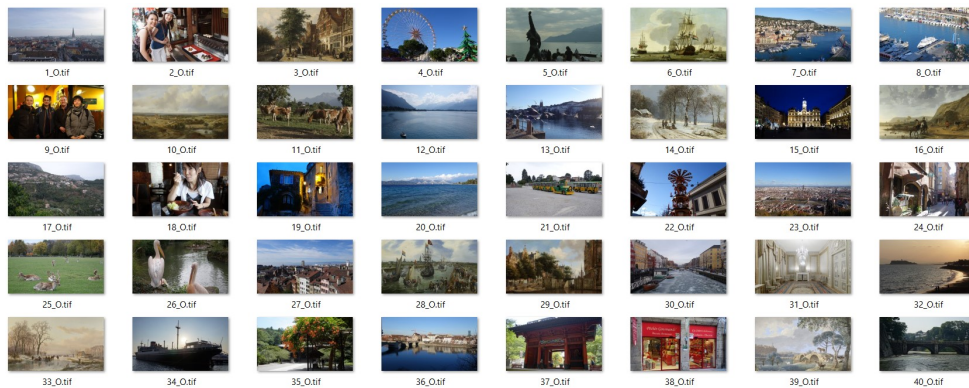


Figura 8: "Dataset iniziale"

In particolare le elaborazioni sono state:

1. Per ogni immagine nel dataset iniziale è stata generata un'immagine in una risoluzione minore o uguale (4K, 3K, 2K, 1K, 720p, 480p) ed un'immagine compressa in jpg con un livello di compressione variabile (0%, 20%, 40%, 60%, 80%, 100%);
2. Per ogni immagine generata al punto 1 è stato applicato un filtro blur su 8 intensità diverse.

Questo procedimento ha così generato circa 3960 immagini, utilizzabili per addestrare e testare la rete.

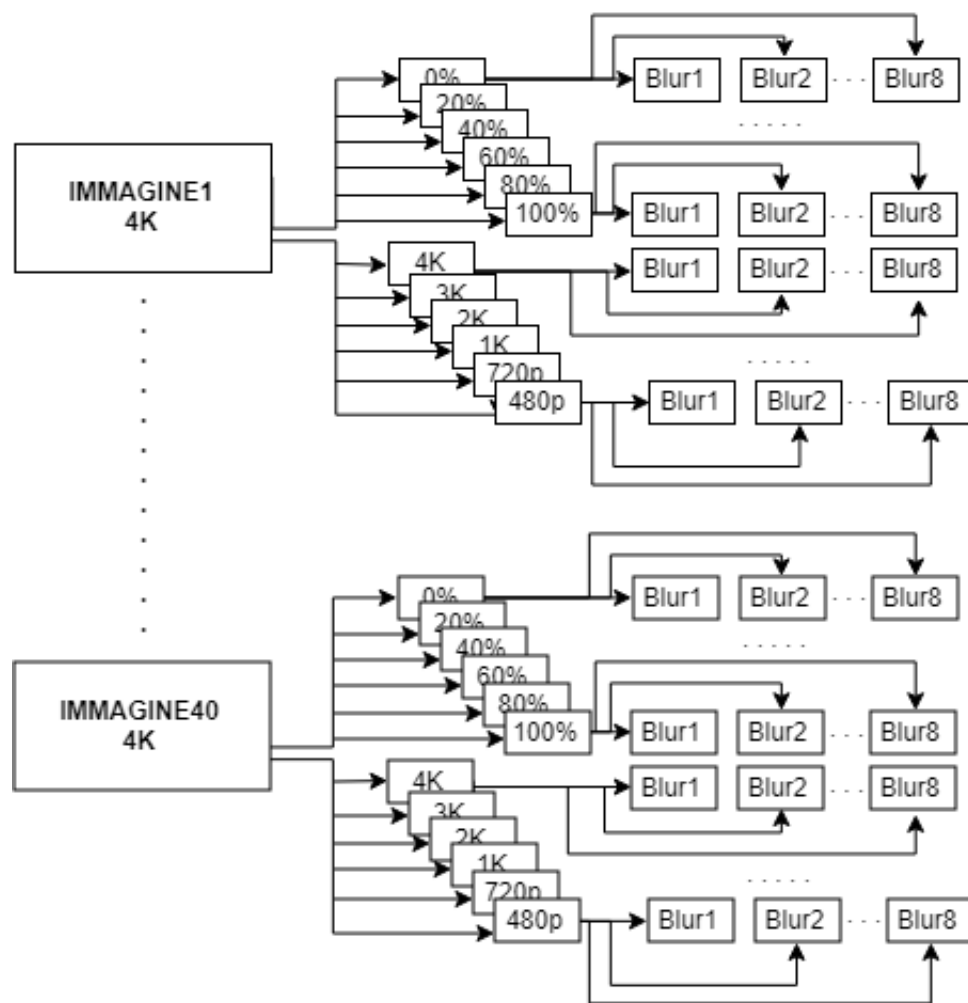


Figura 9: "Generazione del dataset"

5.1.1 Codice C++

Per ottenere il risultato precedentemente descritto è stato sviluppato un programma in C++ sfruttando la libreria OpenCV. Per ogni file nella directory delle immagini vengono prima generate le immagini compresse e le rispettive immagini filtrate:

```
1      for (int i = 0; i <= 100; i += 20) {
2          cout << "Compressing image " << fileName << " of " << i << "%"
3              << endl;
4          std::vector<int> params;
5          params.push_back(CV_IMWRITE_JPEG_QUALITY);
6          params.push_back(100 - i + i%100);
7          string format = (i==0) ? ".tif" : ".jpg";
8          string newPath = basePath + "compressed/" + to_string(i) + "/" +
9              fileName + format;
10
11         if (i == 0)
12             copyFile(filePath, newPath);
13         else
14             imwrite(newPath, imageOriginal, params);
15
16         for (int intensity = 10; intensity <= 80; intensity += 10) {
17             cout << "Blurring image " << fileName << " with intensity " <<
18                 intensity << endl;
19             Mat imageFiltered = applyFilter(newPath, "blur", intensity);
20             string newPathFiltered = basePath + "compressed/" + to_string(i) + "
21                 /" + explode(fileName, '_')[0] + "_B" + to_string(intensity/10) +
22                 format;
23             imwrite(newPathFiltered, imageFiltered);
24         }
25     }
```

Successivamente viene effettuato un procedimento simile per tutte le risoluzioni necessarie:

```
1     for (dimension resolution : resolutions){
2         cout << "Resizing image " << fileName << " to " << resolution.name
           << endl;
3         Size size(resolution.height, resolution.width);
4         Mat resizedImage;
5         resize(imageOriginal, resizedImage, size);
6         string newPath = basePath + "resized/" + resolution.name + "/" +
           fileName + ".tif";
7
8         if (resolution.name.compare("4K") == 0)
9             copyFile(filePath, newPath);
10        else
11            imwrite(newPath, resizedImage);
12
13        for (int intensity = 10; intensity <= 80; intensity += 10) {
14            cout << "Blurring image " << fileName << " with intensity "
               << intensity << endl;
15            Mat imageFiltered = applyFilter(newPath, "blur", intensity);
16            string newPathFiltered = basePath + "resized/" + resolution.name
               + "/" + explode(fileName, '_')[0] + "_B" + to_string(intensity /
               10) + ".tif";
17            imwrite(newPathFiltered, imageFiltered);
18        }
19    }
```

5.1.2 Esempi

Di seguito vengono riportati alcuni esempi di immagini generate con il precedente procedimento le cui differenze sono facilmente percepibili ad occhio nudo.

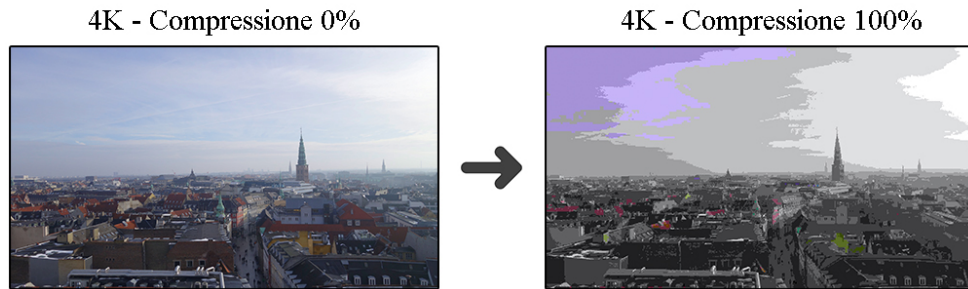


Figura 10: "Esempio di compressione"



Figura 11: "Esempio di applicazione del filtro"

5.1.3 Conversione nello spazio colore CIE L*a*b*

Le immagini del dataset sono inizialmente nello spazio RGB ma poiché si vuole cercare di simulare la percezione dell'occhio umano, si effettua la conversione nello spazio CIE L*a*b* utilizzando la libreria ImageMagick e convertendo tutte le immagini compresse e trasformate in formato JPG in TIF, un formato capace di memorizzare immagini in questo spazio colore.

6 Addestramento della rete

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrum exercitationem ullam corporis suscipit laboriosam, nisi ut aliquid ex ea commodi consequatur. Quis aute iure reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint obcaecat cupiditat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

7 Test

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrum exercitationem ullam corporis suscipit laboriosam, nisi ut aliquid ex ea commodi consequatur. Quis aute iure reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint obcaecat cupiditat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

8 Conclusioni

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrum exercitationem ullam corporis suscipit laboriosam, nisi ut aliquid ex ea commodi consequatur. Quis aute iure reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint obcaecat cupiditat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.