

UNIVERSITÀ DI PISA

Scuola di Ingegneria Corso di laurea in Ingegneria Informatica A.A. 2018/2019

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

Candidato
Davide Coccomini

Relatori Beatrice Lazzerini Francesco Pistolesi

Indice

1	Intr	roduzione	3
2	Colorimetria		
	2.1	Lo spettro visibile	4
		2.1.1 Percezione del colore	5
	2.2	Spazio colore	6
		2.2.1 Modello CIE XYZ	6
		2.2.2 Modello CIE L*a*b*	7
	2.3	Risoluzione e compressione	8
3	Reti Neurali		
	3.1	Reti neurali e cervello umano	9
		3.1.1 Dal neurone biologico a quello artificiale	9
	3.2	Perceptron	10
		3.2.1 Addestramento del perceptron	11
4	Svolgimento 1		
	4.1	Creazione del dataset	12
		4.1.1 Codice C++	14
		4.1.2 Esempi	16
		4.1.3 Conversione nello spazio colore CIE L*a*b*	16
	4.2	Addestramento della rete	17
5	Tes	t	18
6	Cor	nclusioni	19
7	Rin	graziamenti	20

1 Introduzione

L'individuazione delle differenze percettive tra due o più 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.

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

Tutte le colorazioni percepite dall'occhio umano compongono lo "spettro del visibile" che si trova nella parte centrale dello spettro ottico, il quale comprende anche i raggi infrarossi e quelli ultravioletti. La colorimetria utilizza la percentuale della luce incidente che è stata riflessa compresa nell'intervallo del visibile (400-700 nm) per descrivere il colore dell'oggetto. Ciascun oggetto colorato viene pertanto definito da una curva di riflettanza, similmente alle impronte digitali nell'uomo.

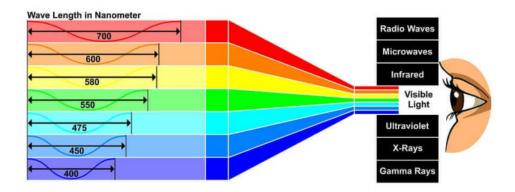


Figura 1: "Lo spettro visibile"

2.1.1 Percezione del colore

Il colore nasce dalla luce. La luce che colpisce un oggetto viene parzialmente assorbita a seconda del materiale che lo compone. 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. Proprio perché nella percezione del colore vengono coinvolte componenti biologiche, ciascun individuo "percepisce" il colore in modo differente. Tale fenomeno può essere ricondotto sia al fatto che non esistono mai due occhi uguali tra loro sia al fatto che l'interpretazione del colore varia da individuo ad individuo. Perfino la stessa persona può percepire differentemente il colore in momenti diversi ed in base allo stato d'animo. È quindi molto complesso definire in maniera oggettiva un colore.

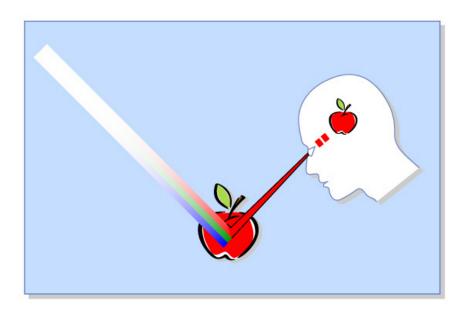


Figura 2: "Percezione del colore"

2.2 Spazio colore

Gli spazi colore sono schemi di rappresentazione dei colori. Ognuno di essi si basa su alcuni parametri e facendoli variare riesce a rappresentare un numero più o meno grande di colori. Esistono numerosi spazi colore che, in base alle loro peculiarità, si prestano meglio ad uno specifico utilizzo piuttosto che ad un altro.

2.2.1 Modello CIE XYZ

Il modello CIE XYZ rappresenta tutti i colori caratterizzati da tre parametri: luminosità, tinta e purezza, rappresentati attraverso un solido. Del solido viene solitamente rappresentata soltanto la sezione secondo il piano XY, in cui X ed Y indicano la cromaticità (tinta e purezza) mentre la luminosità non viene rappresentata.

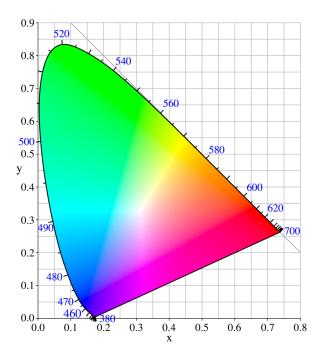


Figura 3: "Spazio colore CIE XYZ"

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:

$$\mathbf{X} = \int_0^\infty I(\lambda) \, \overline{x}(\lambda) \, d\lambda; \, \mathbf{Y} = \int_0^\infty I(\lambda) \, \overline{y}(\lambda) \, d\lambda; \, \mathbf{Z} = \int_0^\infty I(\lambda) \, \overline{z}(\lambda) \, d\lambda;$$

2.2.2 Modello CIE L*a*b*

Il modello CIE L*a*b*, è un'evoluzione del modello CIE XYZ. È 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.

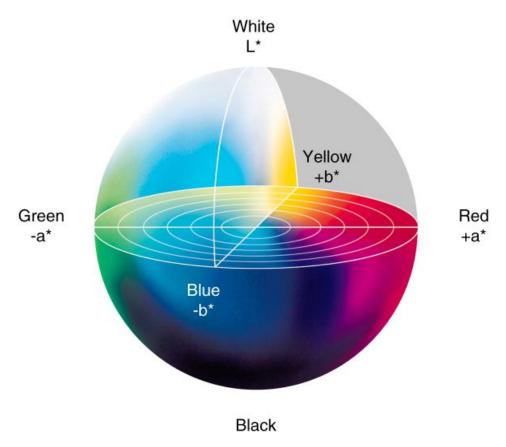


Figura 4: "Spazio colore CIE L*a*b*"

2.3 Risoluzione e compressione

La risoluzione di un'immagine corrisponde al numero di pixel per pollice che contiene, indicata con il termine in DPI (punti per pollice). Un maggior numero di dpi si traduce in una maggiore quantità di informazioni e quindi in una migliore qualità dell'immagine. Ovviamente, maggiore è il numero di informazioni che un'immagine contiene, maggiore sarà il peso del file. Per questa ragione sono state sviluppate alcune tecniche di compressione tra cui quella JPEG. Questa tecnica stabilisce due metodi di compressione di base, uno basato sull'uso della trasformata discreta del coseno, con compressione di tipo "lossy" e cioè con perdita di informazione, e l'altro sull'uso di un metodo predittivo con compressione di tipo lossless, cioè senza perdita di informazione. La perdita di informazioni durante la compressione può compromettere drasticamente la qualità dell'immagine.



Figura 5: "Esempio di compressione JPEG"

3 Reti Neurali

3.1 Reti neurali e cervello umano

Le reti neurali artificiali sono nate per emulare attività tipiche del cervello umano. Risultano quindi molto utili per risolvere problemi di riconoscimento delle immagini o del linguaggio umano. Per riuscire ad emulare questi comportamenti ci si è quindi ispirati al funzionamento del cervello umano. Nel sistema nervoso esistono miliardi di neuroni. 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. 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 influenzare altri neuroni.

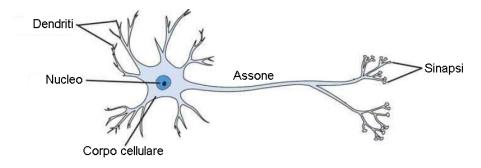


Figura 6: "Neurone biologico"

3.1.1 Dal neurone biologico a quello artificiale

Per riprodurre artificialmente il cervello umano occorre realizzare una rete di elementi molto semplici in grado di imparare e generalizzare. Tipicamente, il neurone artificiale ha molti ingressi ed una sola uscita. Per determinate la conducibilità e quindi l'importanza del canale di ingresso, ogni neurone ha associato un peso. 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 cercando di ottenere il risultato desiderato.

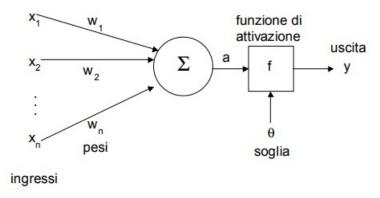


Figura 7: "Modello di neurone"

Abbiamo n canali di ingresso x_1, \ldots, 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:

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

3.2 Perceptron

Un perceptron è una rete neurale ad un singolo livello che si dimostra molto utile nei problemi di classificazione. In generale è composto da una serie $x_1, ..., x_n$ di inputs a cui sono associati dei pesi $w_1, ..., w_n$. Vi è poi una funzione che si occupa di sommare i prodotti dei valori di input e dei rispettivi pesi. Il risultato passa infine per una funzione di attivazione che genera l'output.

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 Svolgimento

4.1 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.

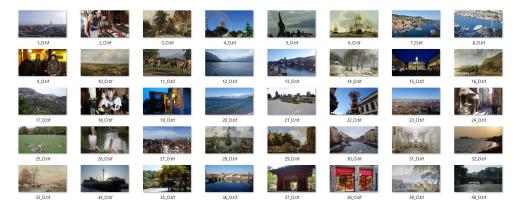


Figura 8: "Dataset iniziale"

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 JPEG 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 (sfocatura) su 8 intensità diverse.

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

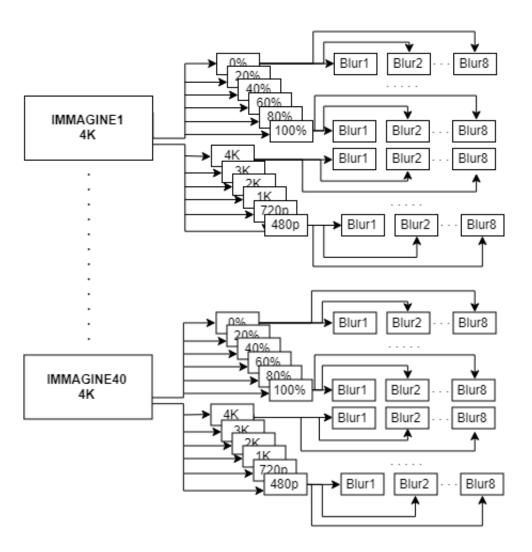


Figura 9: "Generazione del dataset"

4.1.1 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 << "%" <<
                                                           endl;
   3
                                           std::vector<int> params;
                                           params.push_back(CV_IMWRITE_JPEG_QUALITY);
   5
                                           params.push_back(100 - i + i\%100);
                                           string format = (i==0) ? ".tif": ".jpg";
   6
                                           string newPath = basePath + "compressed/"+to_string(i)+"/" + fileName +
   7
   8
   9
                                           if (i == 0)
10
                                                     copyFile(filePath , newPath);
11
                                           else
12
                                                     imwrite(newPath, imageOriginal, params);
13
14
                                           for (int intensity = 10; intensity \leq 80; intensity += 10) {
15
                                                     cout << "Blurring image " << fileName << " with intensity " <<
                                                                      intensity << endl;
                                                     Mat imageFiltered = applyFilter(newPath, "blur", intensity);
16
                                                     string \ newPathFiltered = basePath + \verb"compressed/" + to\_string(i) + \verb"/" + to\_string(i) + to\_string(i) + \verb"/" + to\_string(i) + to\_string(i
17
                                                                      + explode(fileName, '_')[0] + "_B" + to_string(intensity/10) + format
18
                                                     imwrite(newPathFiltered, imageFiltered);
19
20
                       }
```

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

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

4.1.2 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"

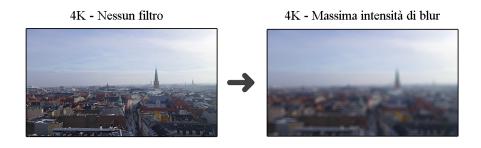


Figura 11: "Esempio di applicazione del filtro"

4.1.3 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 con metodo JPEG in TIF, un formato capace di memorizzare immagini in questo spazio colore.

4.2 Addestramento della rete

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

5 Test

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

6 Conclusioni

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

7 Ringraziamenti

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.