

Studio per un algoritmo di super risoluzione veloce ed adatto a grandi immagini

Sacco Domenico

Abstract

Ottenere un livello di dettaglio maggiore a partire da un'immagine iniziale è uno degli obiettivi più importanti dell'Image Processing. La tecnica di super risoluzione ha come obiettivo quello di incrementare la densità dei pixel dell'immagine di partenza. Lo stato dell'arte di questa tecnica è molto sviluppato, tuttavia raramente il problema della super risoluzione viene studiato affiancandolo al rumore di acquisizione dell'immagine.

Questo lavoro intende presentare un nuovo algoritmo di super risoluzione, cercando il modo di estendere il più possibile la sua applicabilità, processando le immagini ottenute da fotocamere comuni, in particolare dagli smartphone, le quali acquisiscono immagini di grandi dimensioni ma talvolta anche molto rumorose.

Le priorità sono quindi due: riuscire ad incrementare la qualità dell'immagine nonostante il rumore e compiere tale operazione nel più breve tempo possibile, rendendo comunque vantaggioso l'utilizzo della tecnica proposta rispetto alla tradizionale interpolazione bicubica.

Obiettivo

L'obiettivo di questo lavoro è quello di presentare un algoritmo di super risoluzione a frame singolo, idoneo per processare rapidamente le immagini acquisite da smartphone ed adatto a migliorare la qualità di brevi filmati. Si è scelto un intervallo di riferimento da 0.9 a 5 Megapixel per la dimensione delle immagini.

Per raggiungere l'obiettivo preposto, è stato necessario affrontare le seguenti problematiche:

- Le fotocamere degli smartphone sono particolarmente vulnerabili al rumore, che abbassa la qualità delle immagini acquisite.
- La dimensione delle immagini considerata è più elevata rispetto a quelle usate come riferimento per le tradizionali procedure di super risoluzione.
- Foto e filmati andrebbero processati in tempi ragionevoli, nonostante le grandi dimensioni dei frame.

Si propone un metodo che aumenti la qualità dell'immagine in misura maggiore rispetto alla tradizionale interpolazione bicubica, ma che consenta anche l'elaborazione dell'immagine in breve tempo.

I risultati ottenuti sono confrontabili, talvolta superiori, alle procedure di super risoluzione appartenenti all'attuale stato dell'arte.

Stato dell'arte delle procedure di super risoluzione

Per super risoluzione si intende una procedura di Image Processing la quale permette di ottenere da una o più immagini di partenza, chiamate "a bassa risoluzione", un'immagine di dimensioni superiori, la quale offre un maggiore livello di dettaglio data la superiore densità di pixel [1].

La super risoluzione tradizionale prevede l'impiego di più frame a bassa risoluzione per generarne uno ad alta risoluzione, è nota anche come "multi-frame super-resolution" [2].

In questo lavoro sarà invece presa in considerazione un'altra versione della super risoluzione: quella "single-framed" [3]. Talora sia infatti necessario post-processare una foto scattata in precedenza, si dispone soltanto di un frame e non è possibile acquisirne altri.

Le tecniche di super risoluzione a frame singolo possono essere divise in quattro categorie principali.

La prima è la super risoluzione basata sulle reti neurali convoluzionali [4], la quale sfrutta tale modello al fine di aumentare la qualità delle immagini [5].

La seconda tecnica di super risoluzione è quella basata sulle patch [6] [7], in questo caso l'obiettivo è principalmente quello di creare un dizionario che traduca piccole matrici, corrispondenti alle parti dell'immagine a bassa risoluzione, in grandi matrici, costituenti le parti dell'immagine ad alta risoluzione.

Come terza categoria, per ottenere un'immagine finale ad alta risoluzione, possono essere utilizzati dei metodi mirati a preservare i bordi dell'immagine [8] [9].

La quarta categoria di metodi consiste invece nell'utilizzare le trasformate Wavelet [10] per ottenere e processare più agevolmente le informazioni relative alle immagini a bassa risoluzione [11] [12].

Proposizione ed implementazione delle tecniche valutate

Le reti convoluzionali sono ispirate alla corteccia prefrontale degli animali superiori, utilizzano più strati, costituiti da un numero decrescente di neuroni.

Ciascun neurone è collegato ad un gruppo contiguo di neuroni dello strato precedente, come risultato, all'aumentare del livello, il comportamento di un neurone è influenzato da un numero sempre maggiore di neuroni appartenenti al primo strato.

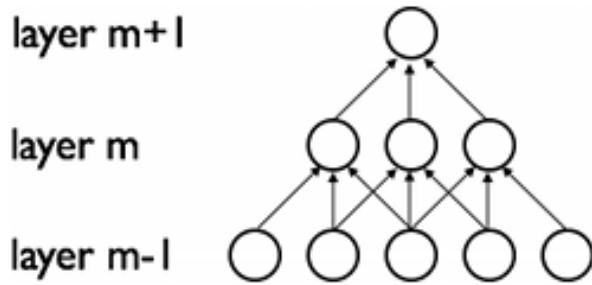


Figura 1: Semplice modello di rete neurale convoluzionale

Il sistema che implementa le patch, oltre che a creare un dizionario basato su immagini di esempio, può anche usare gli elementi ricorrenti in una stessa immagine, che si ripresentano con dimensioni differenti, determinando come ingrandirli attraverso il confronto dell'immagine con delle sue versioni sottocampionate. Le informazioni ottenute vengono integrate al dizionario già disponibile [13].

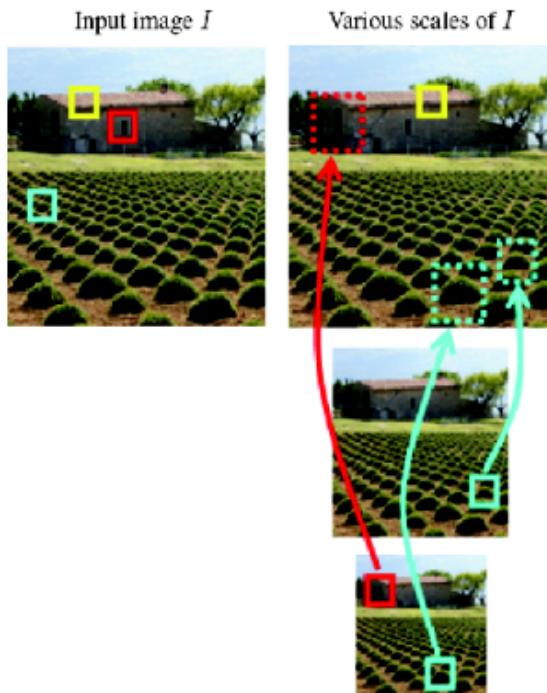


Figura 2: Utilizzo di informazioni già presenti nell'immagine nella super risoluzione patch-based

La tecnica di preservazione dei bordi può avvenire in modalità differenti, ad esempio utilizzando i campi casuali di Markov [14] per determinare le modalità di interpolazione dei pixel adiacenti [15].

Le trasformate Wavelet possono individuare le componenti ad alta frequenza dell'immagine, per

poi elaborarle mediante opportuni filtri e riconvertirle come risultato finale, migliorandone la qualità [16].

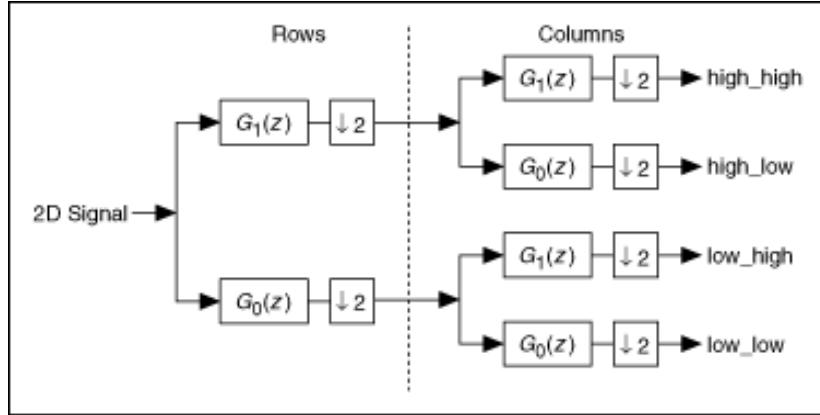


Figura 3: Trasformata DWT bidimensionale

I quattro tipi di tecniche qui presentati sono dotati di pregi e difetti, in particolare:

- L'impiego delle reti neurali è più semplice rispetto al creare degli algoritmi di elaborazione dell'immagine, tuttavia le reti neurali convoluzionali presentano un costo computazionale significativo.
- L'utilizzo delle patch è semplice da implementare, tuttavia richiede un dizionario molto completo per avere una qualità di processing superiore all' interpolazione bicubica, inoltre il tempo di ricerca nel dizionario della patch da tradurre può essere molto elevato. Talora la patch richiesta non sia disponibile è possibile considerare la più vicina a quella in questione, generalmente quella nel dizionario con il minimo Errore Quadratico Medio.
- L'utilizzo di algoritmi mirati alla preservazione dei bordi non richiede training come per le reti neurali o la creazione di dizionari come per i metodi basati sulle patch, tuttavia la capacità di ricostruire l'immagine è molto limitata, in quanto non si dispone di esempi statistici in grado di fornire informazioni aggiuntive rispetto all'immagine a bassa risoluzione di partenza.
- L'utilizzo delle trasformate Wavelet permette di estrarre le caratteristiche dell'immagine rapidamente, tuttavia presenta lo stesso problema di base del metodo precedente.

Si è scelto di confrontare le prestazioni dell'algoritmo proposto con due tecniche di super risoluzione, una basata sulle Fast Super Resolution Convolutional Neural Networks (FSRCNN) [17] e l'altra tecnica, basata sulle patch [18].

E' stato usato come ambiente di lavoro il software MATLAB, su Windows 8, versione a 64 bit.

Le tecniche [17] e [18] non erano state precedentemente pensate per processare immagini di grandi dimensioni, pertanto sono state adattate al contesto del problema utilizzando una funzione come interfaccia, la quale permettesse di dividere l'immagine in riquadri di dimensione minore da

processare.

Il processing dell'immagine integrale infatti, talora sia di dimensioni eccessive, porta ad un blocco immediato ed irreversibile dell'elaboratore. La dimensione scelta per i riquadri riquadri è dell'ordine di 10^6 pixel, in maniera tale da non suddividere l'immagine in un numero di parti eccessivo da un lato e di non causare un overflow della memoria dell'elaboratore dall'altro.

La tecnica basata sulle patch infatti, presentava anomalie ai bordi estremi dell'immagine, i quali sono stati rimossi. Nei singoli riquadri invece non erano presenti anomalie.

Per quanto concerne invece la tecnica FSRCNN, l'adattamento ha richiesto uno studio approfondito in quanto la rete neurale scarta direttamente i bordi dei singoli riquadri, pertanto è stato necessario utilizzare un riquadro di grandezza superiore (padding su tutti e quattro i lati). Sono presentati, in funzione della grandezza del padding (4%, 8%, 12% e 16% della dimensione del riquadro), il tempo di esecuzione e la variazione del PSNR.

Le immagini di riferimento misurano 1024x1024 pixel.

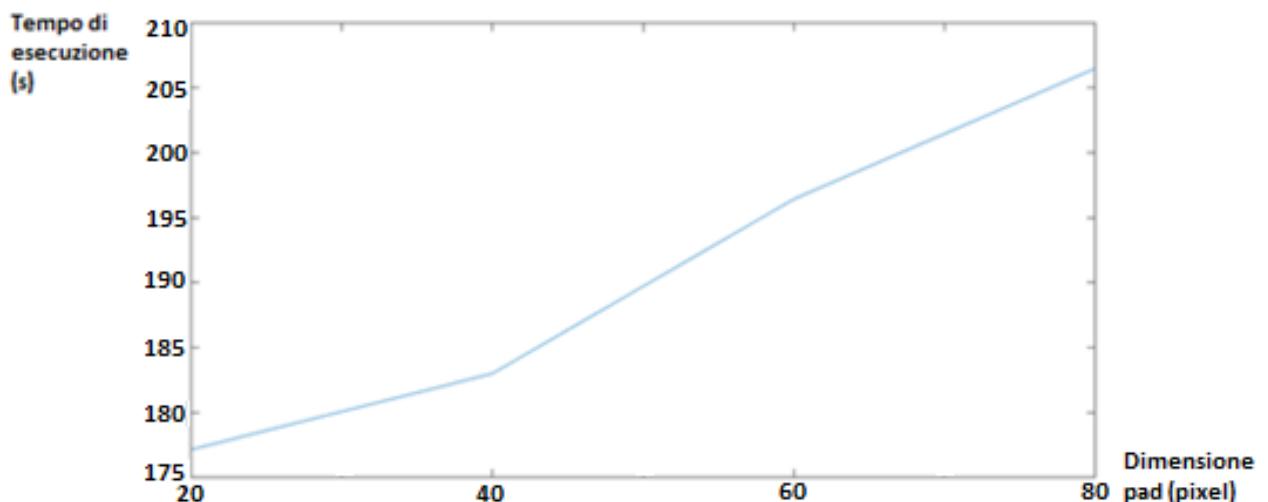


Figura 4: Tempo di esecuzione (s) in funzione del pad (pixel)

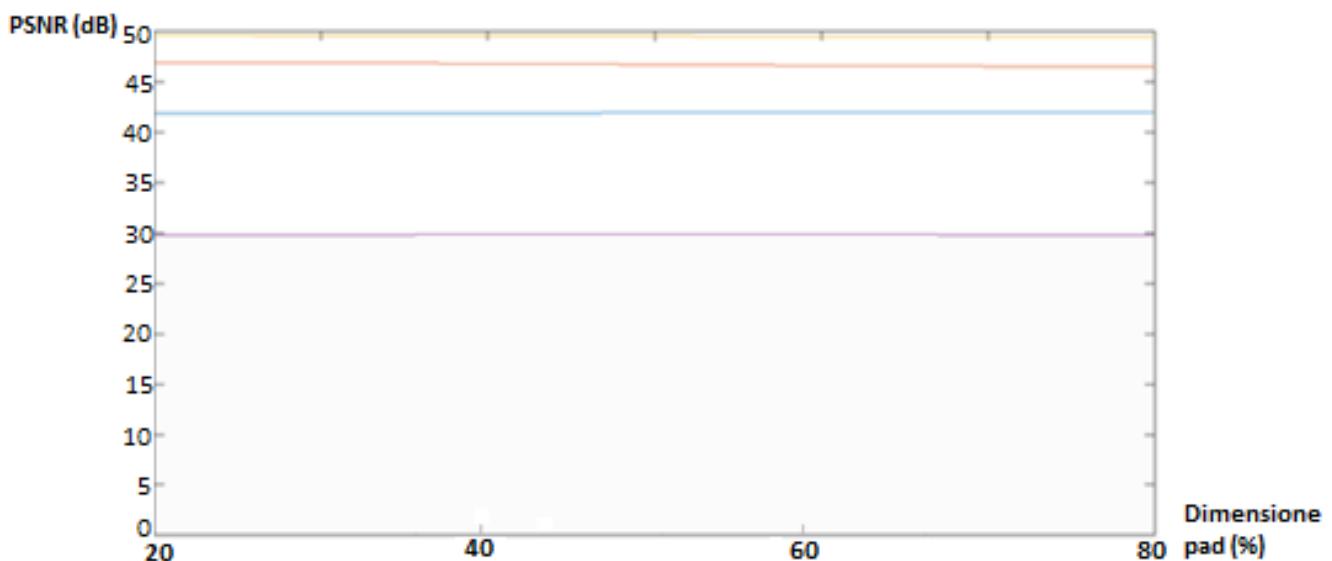


Figura 5: PSNR dell'immagine (dB) in funzione del pad (%)

Come mostrato, un decremento della dimensione del pad non fa variare in maniera significativa la qualità dell'immagine processata, pertanto si è scelto di impiegare un bordo ridotto, riducendo il tempo di esecuzione di circa il 17%.

Si evidenzia ora l'assenza di mosaicizzazione ai bordi per il padding massimo e minimo presi in considerazione, ovvero 20 ed 80 pixel rispettivamente:

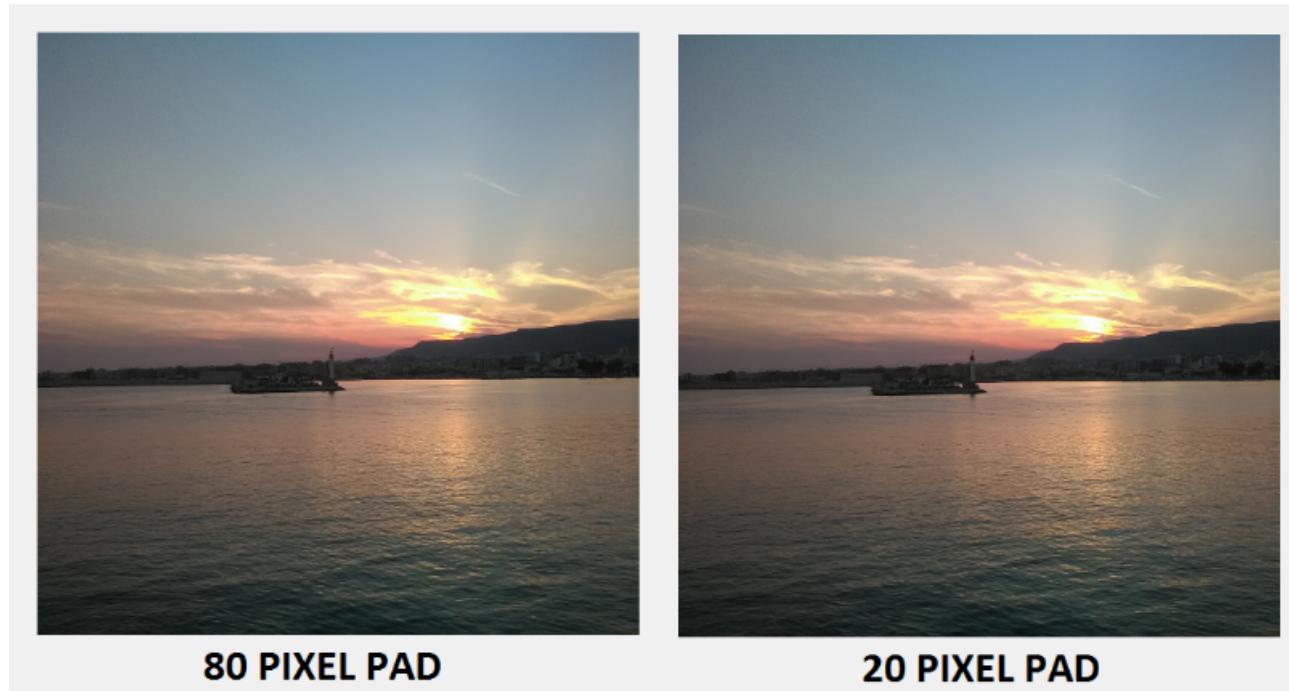


Figura 6: La stessa immagine da 1024x1024 pixel con i valori estremi di pad, nessuna mosaicizzazione presente

La tecnica proposta in questo lavoro prevede l'utilizzo della trasformata DWT bidimensionale (di tipo Haar [19]) ed implementa una tecnica di evidenziazione dei bordi attraverso l'amplificazione delle alte frequenze.

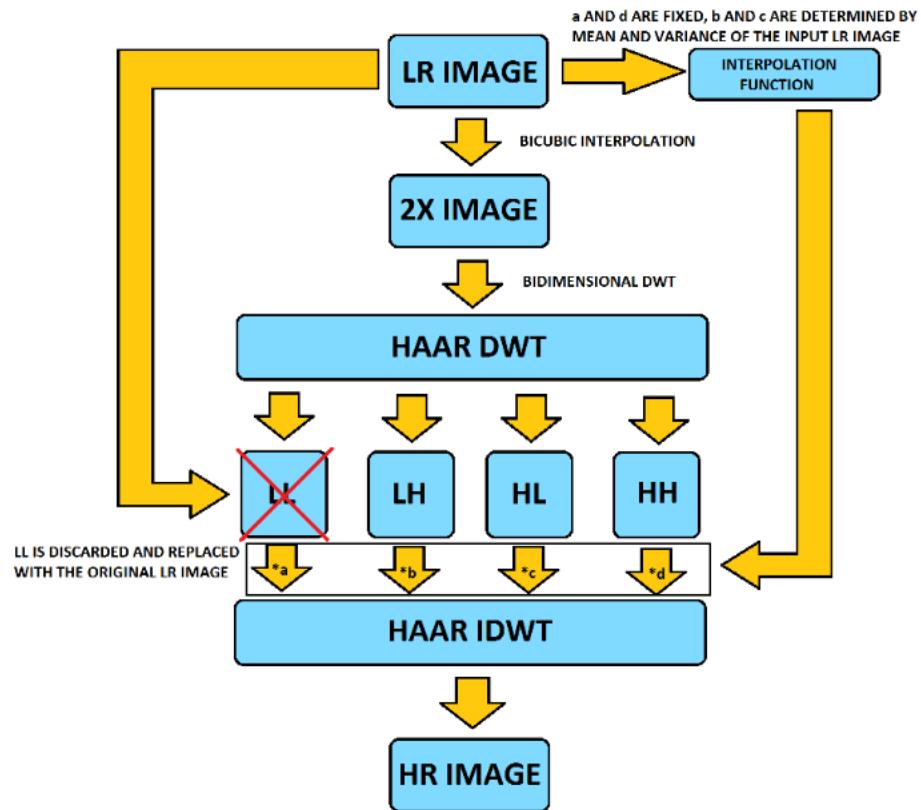


Figura 7: La tecnica proposta

I coefficienti a e d sono fissi e rispettivamente pari a 2 e 1,9. I coefficienti c e d sono stati ottenuti studiando il campione di immagini a disposizione, creando una funzione che permettesse di ottenerli conoscendo la varianza e la media dell'immagine in ingresso.

Il sotto campionamento prevede l'utilizzo di un filtro di averaging e successivamente la decimazione della matrice così ottenuta.

Di seguito sono confrontate le trasformate logaritmiche della risposta all'impulso:

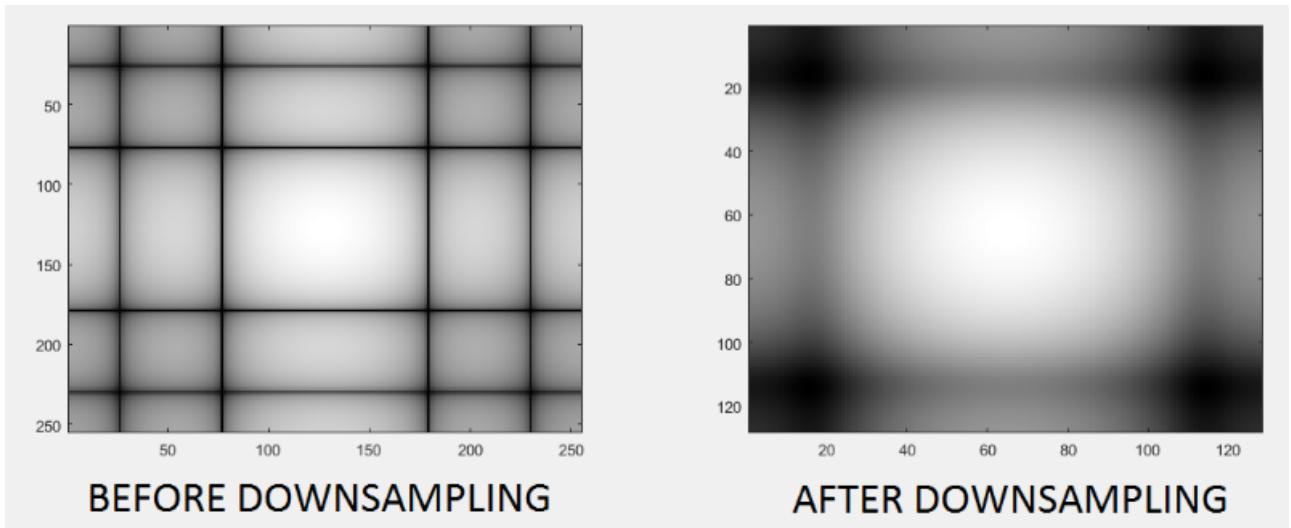


Figura 8 Trasformata logaritmica prima e dopo il sotto campionamento

Le risposte in frequenza relative alle matrici LL e LR IMAGE (ovvero l'immagine di partenza a bassa risoluzione) su Lenna sono le seguenti:

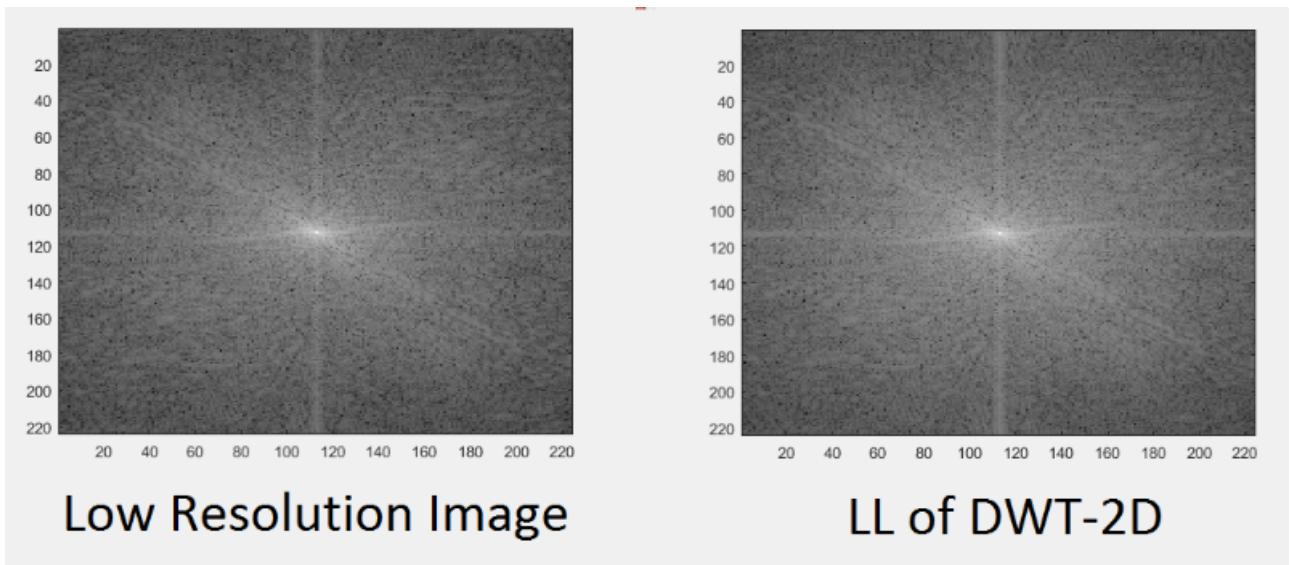


Figura 9 Le due risposte in frequenza, con trasformata logaritmica

Pur sembrando molto simili, le due immagini presentano dei punti in cui sono differenti:



Figura 10 Differenza fra l'immagine di partenza (LR IMAGE) e LL

Tali punti rappresentano le alte frequenze le quali, come ci si aspetta, sono presenti nell'immagine a bassa risoluzione di partenza ma non in quella LL, dato il filtraggio passa-basso che quest'ultima ha subito.

Alla procedura descritta, si è scelto di affiancare un filtro passa basso (averaging) per eliminare il rumore alle alte frequenze ed uno passa alto per evidenziare i bordi, si osservi il confronto:



Figura 11: A sinistra la tradizionale interpolazione bicubica, a destra la tecnica proposta

Si è deciso, per ridurre l'utilizzo della memoria in caso di matrici di grandi dimensioni, di suddividere in blocchi anche questa tecnica.

Prendendo come riferimento il fattore di magnificazione pari a 2, i tempi medi di esecuzione su un'immagine di 2,5 megapixel sono stati i seguenti:

TECNICA	BICUBICA	FSRCNN	PATCH-BASED	PROPOSTA (RGB)	PROPOSTA (SOLO Y)
TEMPO	0.9s	550s	620s	4.3s	2.7s

Come si nota, la velocità della tecnica proposta è di gran lunga superiore alle altre tecniche di super risoluzione, tanto da poter permettere, a differenza della rete neurale FSRCNN, l'applicazione a tutte e tre le componenti RGB in tempi rapidi.

E' possibile velocizzare ulteriormente l'algoritmo proposto considerando la trasformazione YCbCr e processando solo il canale di luminanza. Talora si abbia a disposizione l'immagine (o l'insieme dei frame da processare) direttamente in formato YCbCr, le prestazioni temporali dell'algoritmo migliorano notevolmente, dovendo evitare la conversione da RGB a YCbCr e quella opposta per fornire come uscita il risultato nuovamente in formato RGB. Tuttavia anche in presenza di immagini in RGB la tecnica che processa solo la luminanza resta notevolmente più veloce.

Per il confronto delle tecniche si sono scelti gli indici PSNR (Peak Signal-to-Noise Ratio) [20] e SSIM (Structure Similarity Index) [21].

Nel set di test, composto da 10 immagini da 5 Megapixel scattate da cellulare, la versione YCbCr dell'algoritmo si è rivelata non soltanto più veloce, ma anche leggermente più efficiente.

Analisi del PSNR (dB):

IMMAGINE	BICUBICA	FSRCNN	PATCH-BASED	PROPOSTA (LUMINANZA)	PROPOSTA (RGB)
1	38.0425	39.1333	38.9725	38.6721	38.7922
2	42.7707	43.6545	43.5284	43.2307	43.3094
3	49.5249	49.6119	50.3824	50.1307	49.8394
4	51.5006	50.4767	51.4904	51.8282	50.8319
5	46.1310	46.5338	46.4091	46.5437	46.5233
6	30.9941	31.6246	31.5498	31.3494	31.3872
7	32.9584	33.6739	33.3793	33.3859	33.4317
8	50.2422	49.8126	50.2813	50.6253	49.6873
9	36.4377	37.6192	37.4648	37.1373	37.2612
10	35.5587	37.4385	37.3477	36.3153	36.5154

Analisi del SSIM:

IMMAGINE	BICUBICA	FSRCNN	PATCH-BASED	PROPOSTA (LUMINANZA)	PROPOSTA (RGB)
1	0.9594	0.9661	0.9664	0.9646	0.9652
2	0.9880	0.9880	0.9892	0.9892	0.9893
3	0.9962	0.9905	0.9964	0.9964	0.9961
4	0.9951	0.9895	0.9950	0.9954	0.9948
5	0.9941	0.9923	0.9946	0.9947	0.9945
6	0.9738	0.9771	0.9770	0.9760	0.9762
7	0.8751	0.8970	0.8981	0.8898	0.8911
8	0.9947	0.9902	0.9947	0.9952	0.9944
9	0.9569	0.9646	0.9654	0.9628	0.9636
10	0.9601	0.9705	0.9708	0.9662	0.9671

Viene di seguito presentata la media dei risultati ottenuti:

METODO DI SUPER RISOLUZIONE	PSNR MEDIO (db)	SSIM MEDIA
INTERPOLAZIONE BICUBICA	41.4161	0.9693
FSRCNN	41.9579	0.9726
PATCH-BASED	42.0806	0.9748
ALGORITMO PROPOSTO (RGB)	41.7579	0.9732
ALGORITMO PROPOSTO (SOLO Y)	41.9219	0.9730

Dalle tabelle si evince che, nonostante le tecniche presentate siano un centinaio di volte più rapide, ottengono prestazioni paragonabili, se non in certi casi superiori, alla tecnica di super risoluzione attraverso la rete neurale convoluzionale e all'algoritmo basato sulle patch.

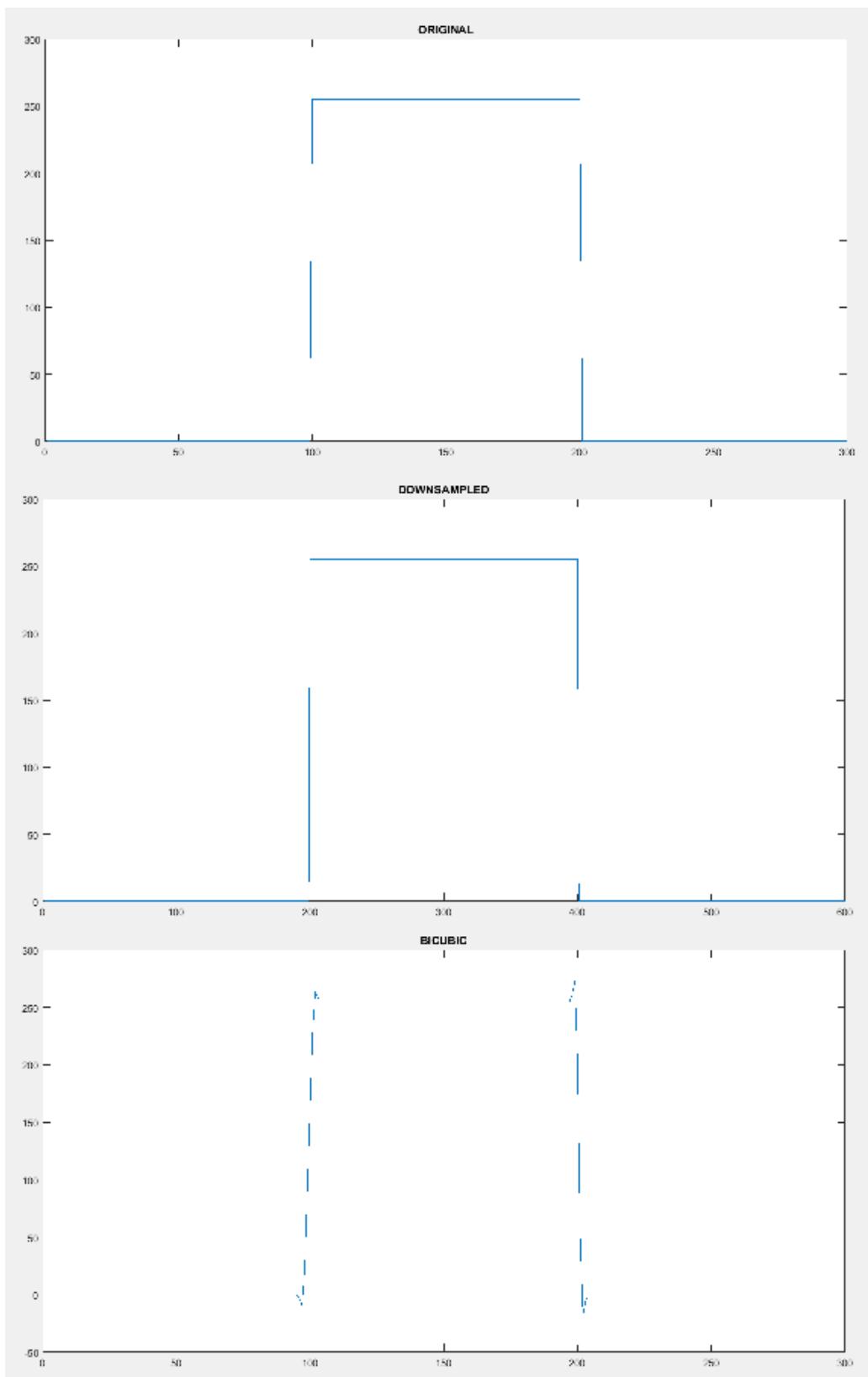
Si presenta ora un confronto fra le tecniche nell'immagine di Lenna, ovvero in un contesto diverso dalle foto scattate attraverso uno smartphone:

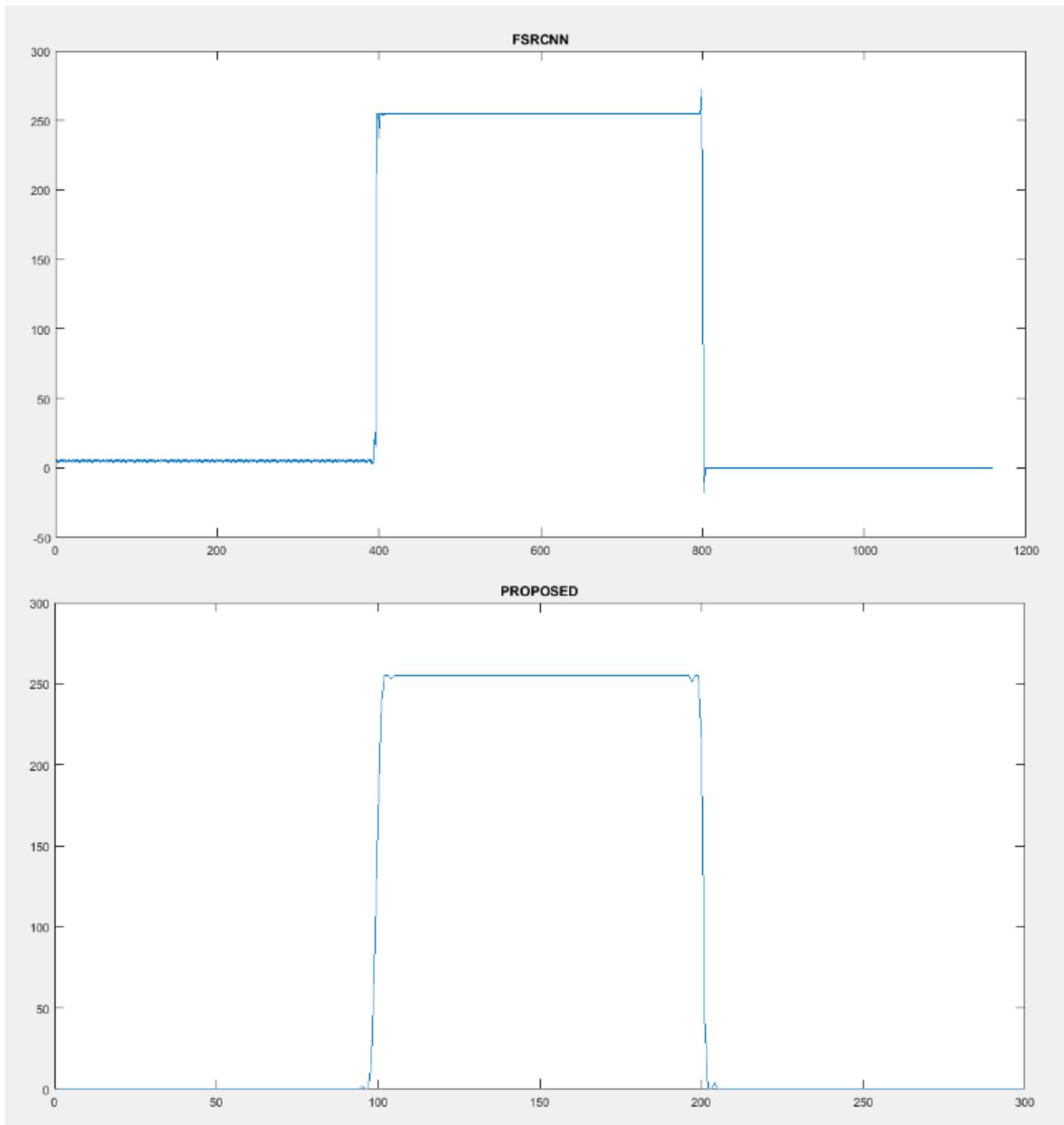


TECNICA	PSNR	SSIM
BICUBICA	0.9478	30.6258
FSRCNN	0.9649	32.8937
PATCH-BASED	0.9630	32.5303
PROPOSTA (RGB)	0.9543	31.1593
PROPOSTA (Y)	0.9536	31.0709

Risposta al profilo delle tecniche presentate

Si è usata un'immagine a scacchiera di riferimento (nero-bianco-nero), sfocata e decimata, per ottenere la risposta al profilo delle tecniche presentate:





Come ci si attendeva, a risposta al profilo dell'immagine sfocata e sottocampionata ha un livello intermedio fra massimo e minimo e la bicubica presenta subito prima della salita un valore negativo ed una sovra elongazione.

La risposta della rete neurale è molto particolare, presenta un valore minore di nero dopo la sequenza bianca.

La risposta relativa alla tecnica proposta è atipica: prima della salita presenta un lieve picco e subito dopo questa una lieve diminuzione, il comportamento per la discesa è simmetrico.

Comportamento delle tecniche in presenza di rumore Gaussiano

Si è scelto di introdurre del rumore additivo gaussiano a media nulla (Additive White Gaussian Noise) nelle immagini da elaborare, in modo da determinare quale degli algoritmi presentati fosse il migliore per aumentare la risoluzione nonostante il rumore.

VARIANZA

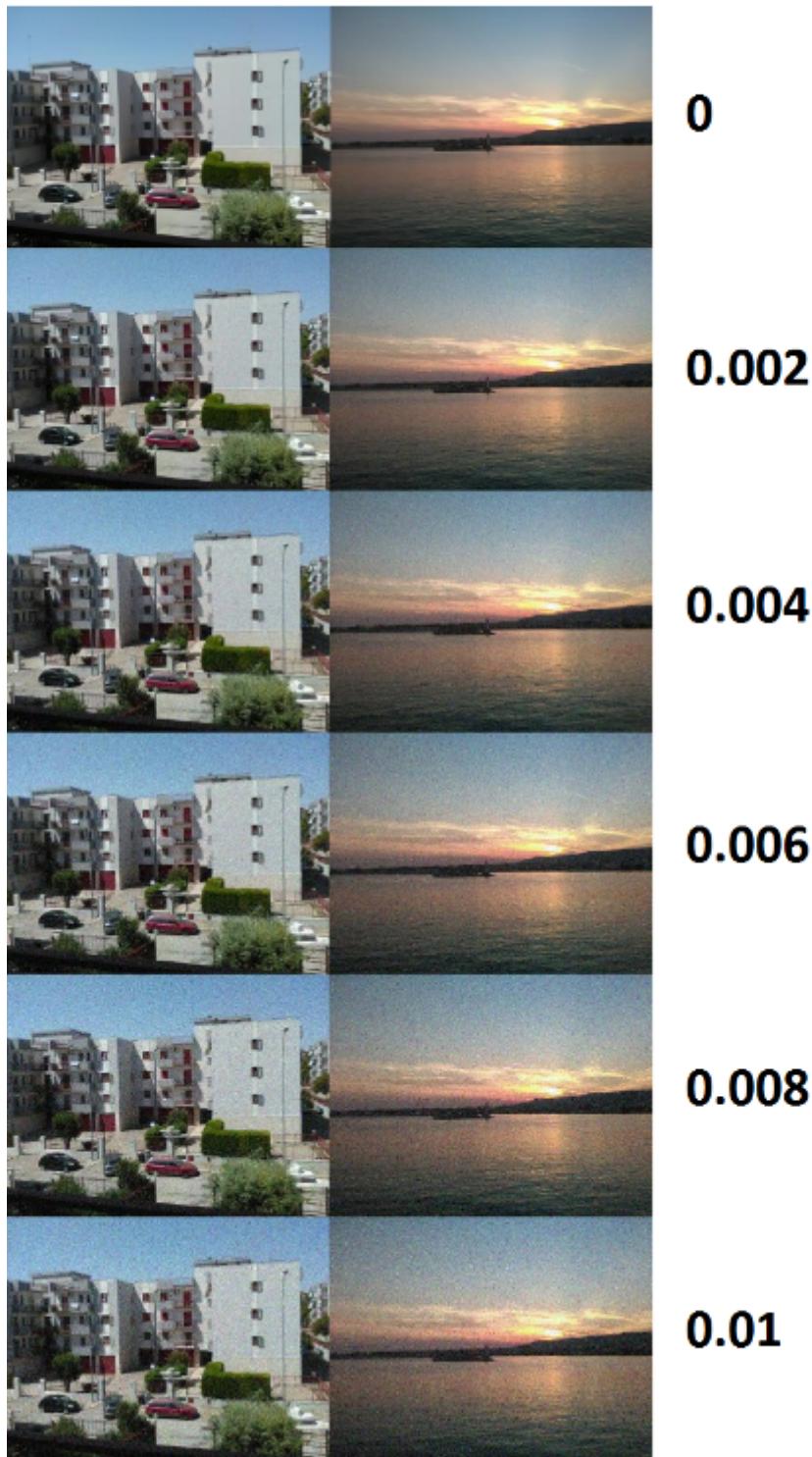


Figura 14: Rumore gaussiano ad intensità crescente in due delle immagini del set

Si è scelto di sperimentare sulle tecniche da confrontare i filtri di averaging 2×2 , 3×3 e 4×4 al fine di contrastare il rumore AWGN. All'aumentare del rumore si è riscontrato più performante applicare filtri di dimensioni superiori, il filtro 4×4 è stato quello rivelatosi più efficiente, ma filtri di dimensioni superiori comporterebbero la perdita di una consistente parte dell'informazione

presente nell'immagine, costituita dalle alte frequenze .

Sono di seguito presentati i risultati, confrontati con le varie versioni dell'algoritmo proposto:

- VER 1.1: Nessun filtro, tutte le componenti RGB
- VER 1.2: Solo averaging, tutte le componenti RGB
- VER 1.3: Averaging e sharpening, tutte le componenti RGB
- VER 2.1: Nessun filtro, solo luminanza
- VER 2.2: Solo averaging, solo luminanza
- VER 2.3: Averaging e sharpening, solo luminanza

Analisi del PSNR (dB) all'aumentare della varianza del l'AWGN:

TECNICA	VAR=0	VAR=0.002	VAR=0.004	VAR=0.006	VAR=0.008	VAR=0.01
RIFERIMENTO	Inf	32.0670	29.1696	27.4805	26.2897	25.3711
BICUBICA	41.416 1	32.4143	30.0853	28.6196	27.5507	26.7042
BICUBICA e 4X4	36.618 5	33.8481	32.4302	31.4064	30.6075	29.9390
FSRCNN	41.957 9	28.7247	26.0861	24.5131	23.3892	22.5175
FSRCNN e 4X4	36.822 3	33.5156	31.9332	30.8153	29.9516	29.2361
PATCH-BASED	42.080 6	29.2790	26.6346	25.0482	23.9079	23.0234
PATCH-BASED e 4X4	36.907 7	33.5218	31.9202	30.8072	29.9296	29.2188
VER 1.1	41.757 9	30.7906	28.2791	26.7474	25.6399	24.7767
VER 1.2	34.697 2	33.2626	32.2779	31.5142	30.8776	30.3317
VER 1.3	37.145 2	32.9858	31.2152	30.0056	29.0786	28.3243
VER 2.1	41.921 9	30.8719	28.3630	26.8374	25.7322	24.8674

VER 2.2	34.690 3	33.2647	32.2753	31.4986	30.8555	30.3026
VER 2.3	37.131 8	33.0849	31.3498	30.1613	29.2487	28.5047

Analisi del SSIM:

TECNICA	VAR=0	VAR=0.002	VAR=0.004	VAR=0.006	VAR=0.008	VAR=0.01
RIFERIMENTO	1.000 0	0.6610	0.5401	0.4683	0.4188	0.3813
BICUBICA	0.969 3	0.7110	0.6007	0.5312	0.4817	0.4435
BICUBICA e 4X4	0.944 8	0.8444	0.7782	0.7278	0.6873	0.6533
FSRCNN	0.972 6	0.6287	0.5117	0.4429	0.3954	0.3596
FSRCNN e 4X4	0.947 7	0.8404	0.7707	0.7182	0.6766	0.6417
PATCH-BASED	0.974 8	0.6476	0.5301	0.4600	0.4111	0.3744
PATCH-BASED e 4X4	0.947 9	0.8400	0.7702	0.7180	0.6762	0.6417
VER 1.1	0.973 2	0.6287	0.5114	0.4426	0.3950	0.3595
VER 1.2	0.929 6	0.8627	0.8136	0.7743	0.7410	0.7125
VER 1.3	0.951 0	0.7927	0.7053	0.6440	0.5975	0.5602
VER 2.1	0.973 0	0.6848	0.5708	0.5010	0.4518	0.4142
VER 2.2	0.929 5	0.8666	0.8201	0.7821	0.7502	0.7226
VER 2.3	0.950	0.8116	0.7306	0.6722	0.6272	0.5907

Nonostante i lunghi tempi di esecuzione degli algoritmi FSRCNN e Patch-Based, l'interpolazione bicubica, molto più rapida, supera entrambe in termini di PSNR e SSIM. Tale considerazione vale sia in presenza che in assenza del filtro di averaging 4x4.

Le versioni 1.2 e 2.2 delle tecniche proposte si sono rivelate le più efficaci.

I valori inferiori del PSNR e del SSIM delle tecniche 1.2 e 2.2 rispetto alla bicubica quando la varianza è nulla o bassa possono essere attribuiti al fatto che l'immagine di riferimento, ovvero quella priva di rumore aggiunto, sia in realtà intrinsecamente rumorosa a causa del rumore termico dei sensori della fotocamera.

Pertanto le immagini ottenute dalle tecniche 1.2 e 2.2, anche se meno rumorose di quelle ottenute dalle altre tecniche, non hanno PSNR e SSIM più alti dato che vengono confrontate con un'immagine di riferimento pur sempre rumorosa.

Quando la varianza del rumore AWGN aggiunto artificialmente agli scopi dell'analisi diventa preponderante, risalta invece dalle metriche utilizzate la maggiore efficienza delle tecniche proposte (si vedano i risultati delle ultime 3 colonne sulla destra).

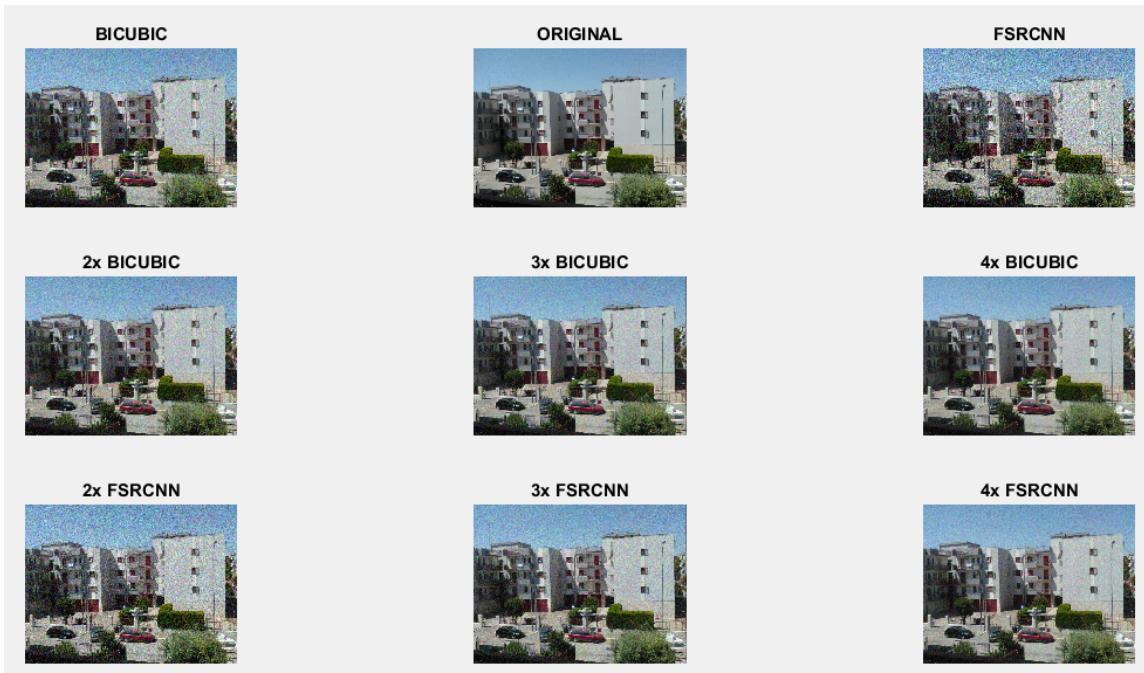


Figura 15: Reazione al rumore massimo (var=0.01) dell'algoritmo FSRCNN e della interpolazione bicubica, con e senza averaging.

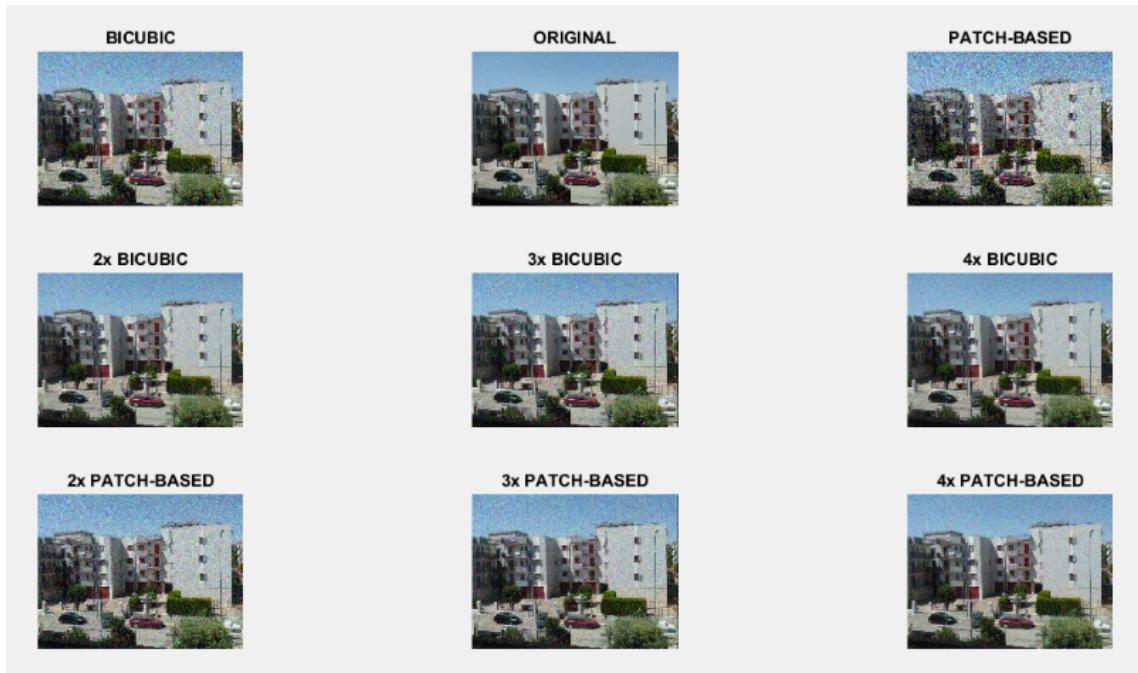


Figura 16: Reazione al rumore massimo ($\text{var}=0.01$) dell'algoritmo Patch-based e della interpolazione bicubica, con e senza averaging.

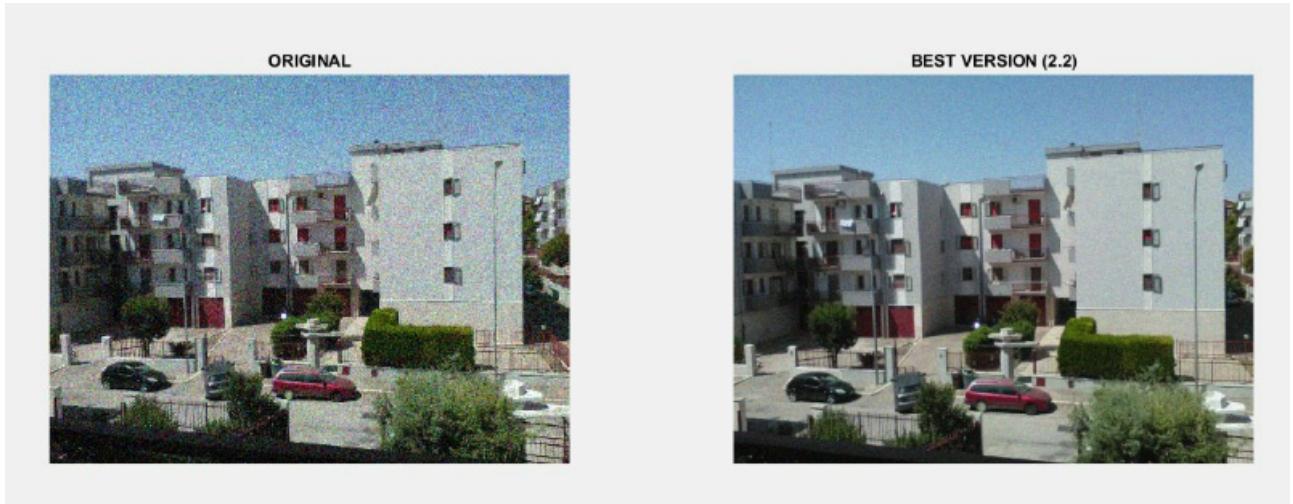


Figura 17: Reazione al rumore massimo ($\text{var}=0.01$) della migliore tecnica presentata.

Per cercare una motivazione in ragione della quale l'interpolazione bicubica riesca a ricostruire l'immagine meglio dei metodi FSRCNN e Patch-Based in presenza del rumore AWGN, si può dimostrare l'affermazione intuitiva che l'interpolazione bicubica sia intrinsecamente passa - basso e pertanto elmini il rumore in maniera abbastanza efficace. A supporto di tale affermazione si è valutata la risposta all'impulso bidimensionale dell'interpolazione bicubica e dei filtri di averaging di diverse dimensioni:

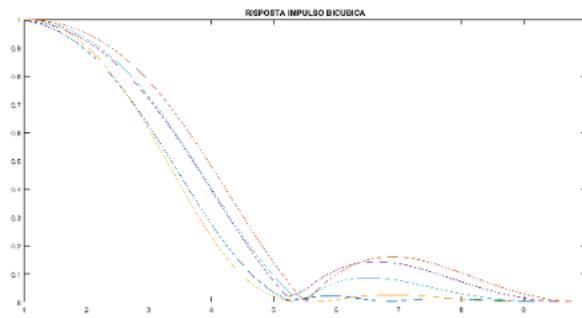


Figura 18: Risposta all'impulso dell'interpolazione bicubica

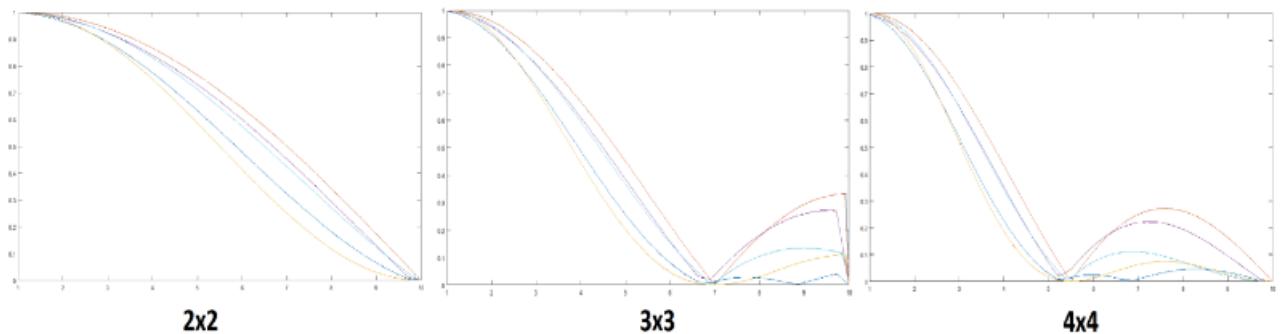


Figura 19: Risposta all'impulso dei filtri di averaging impiegati per contrastare il rumore

Avendo appreso dalle tecniche precedentemente viste la maggiore efficienza dell'interpolazione bicubica a causa del suo effetto passa - basso, si è scelto di usare proprio l'interpolazione bicubica come base per la tecnica proposta, in maniera tale da avere come punto di partenza un metodo già di per sé efficace, sia come tempo di esecuzione che come contrasto del rumore.

Applicazioni real time

La tecnica proposta può essere impiegata al fine di trasmettere in streaming contenuti video, aumentandone la risoluzione in misura migliore rispetto all'interpolazione bicubica, mantenendo un frame rate accettabile dato il ridotto costo computazionale. L'algoritmo mantiene parte della sua efficienza anche con fotogrammi contenuti in video compressi in formato mp4.



BICUBIC

PROPOSED TECNHIQUE

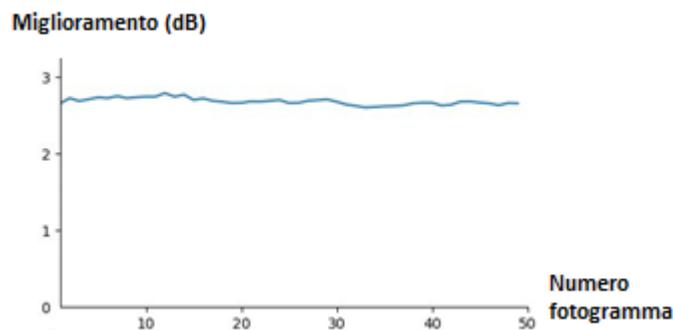


Figura 20: Aumento del PSNR nei primi 50 fotogrammi del video di prova impiegato
(<https://www.youtube.com/watch?v=shnV2H6E0Q8>)

Un'altra applicazione riguarda la comunicazione attraverso videoconferenza, in questo caso il collo di bottiglia che riduce il frame rate è situato nella trasmissione dei pacchetti attraverso la rete e non nel tempo di elaborazione delle immagini.

Si è verificato che, anche sulla stessa macchina (e quindi con tempo di attraversamento della rete nullo) il frame rate si rivela sufficientemente alto da trasmettere in maniera fluida la sequenza di immagini prelevate dalla webcam.

Si può pensare quindi di trasmettere le immagini a metà della risoluzione, applicando la tecnica proposta sui due (o più) computer che comunicano, risparmiando banda e nel contempo recuperando buona parte della qualità perduta attraverso la tecnica proposta nella sua versione più rapida (VER 2.1).

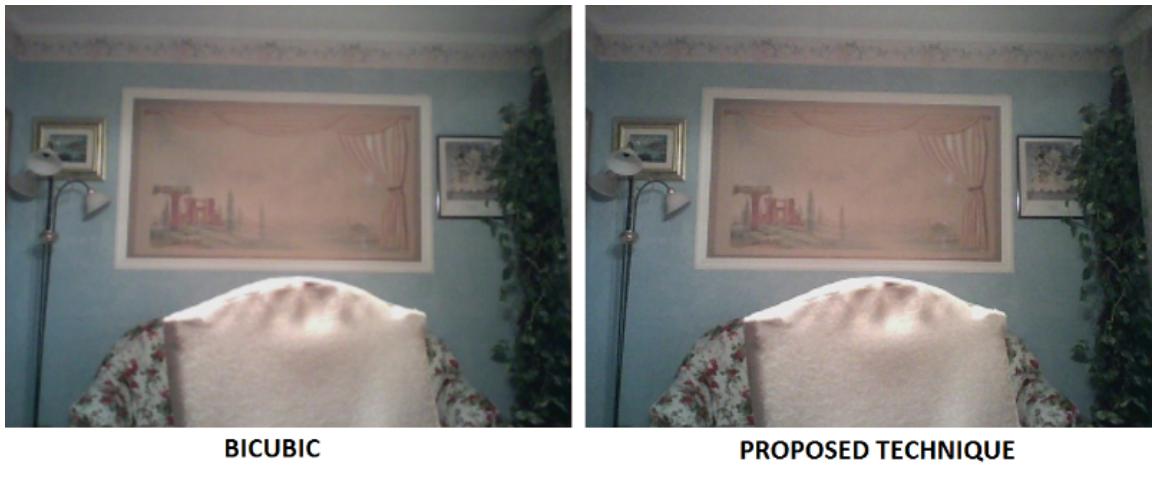


Figura 21: Confronto fra i due video acquisiti da webcam

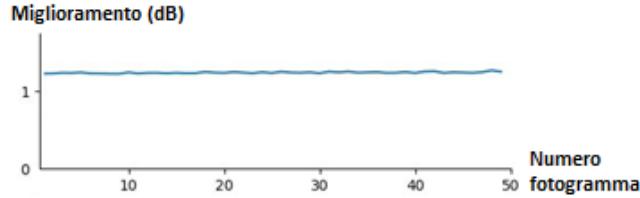


Figura 22: Aumento del PSNR in 50 fotogrammi acquisiti da una webcam ad immagine fissa, da notare come il rumore termico influenzi solo marginalmente il miglioramento, il quale resta pressoché costante

Applicazione a Telegram

Si è deciso di creare un bot per Telegram [22] collegato ad un sorgente Python, il quale potesse applicare le tecniche di super risoluzione proposte in questo lavoro alle foto che gli vengono inviate. Oltre a tale funzionalità, il bot è stato dotato della capacità di migliorare la qualità dei video in formato mp4.

L'utilizzo di Telegram è stato ritenuto ottimale in quanto permette di evitare problemi nella comunicazione tra due dispositivi a causa dei sistemi di protezione, per esempio in caso di impiego del NAT [23].

In questo caso Telegram agisce come un vero e proprio relay server, allo scopo di permettere la condivisione di testo e file multimediali fra due dispositivi connessi ad Internet.

L'architettura del programma che viene eseguito sul server al fine di garantire l'attività del bot è divisa in due parti fondamentali:

- Il codice sorgente in Python, il quale si occupa dell'invio, della ricezione e dell'interpretazione dei messaggi.
- I moduli compilati con MATLAB, ciascuno dei quali svolge l'omonimo ruolo del comando

assegnatogli.

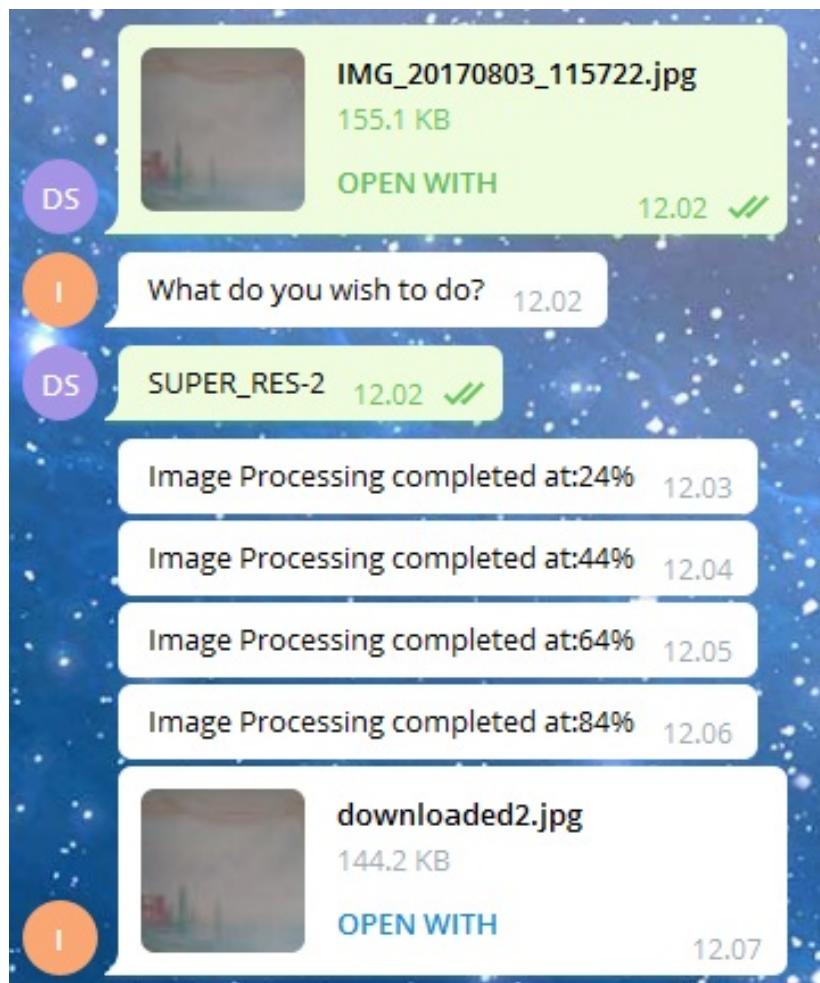


Figura 23: Esempio di sessione, applicazione della super risoluzione FSRCNN con fattore 2

Per usufruire del bot, basta trovarlo sul motore di ricerca di Telegram, in cui risponde al nome di "Image_processing_super_resolution", avviare la conversazione con il comando "/start" ed inviare come file (quindi senza compressione) una qualsiasi immagine con estensione jpg o un breve video in formato mp4. Il bot risponde ai seguenti comandi:

- FAST_RES : applica la tecnica di super risoluzione proposta.
- PATCH_RES-[2,3,4] : applica l'algoritmo di super risoluzione patch-based, in cui i numeri rappresentano il fattore di ingrandimento desiderato.
- SUPER_RES-[2,3,4] : applica l'algoritmo basato sulla rete neurale FSRCNN, anche in questo caso i numeri rappresentano il fattore di ingrandimento desiderato.

- VIDEO_PROCESS : applica la tecnica di super risoluzione proposta, un filtro passa basso per ridurre il rumore ed uno passa-alto per migliorare la qualità estetica dell'immagine.
- HELP : lista dei comandi disponibili.

Conclusioni

In questo lavoro è stato presentato un metodo di super risoluzione adatto alle immagini di grandi dimensioni, contraddistinto dalla capacità di essere veloce e resistente al rumore, le cui prestazioni sono confrontabili ai metodi appartenenti allo stato dell'arte.

Tale algoritmo è caratterizzato da un tempo di esecuzione molto ridotto e può quindi essere utilizzato per la comunicazione in tempo reale, processare in tempi ragionevoli dei brevi filmati oppure immagini di dimensione significativa.

Bibliografia e sitografia

[1] Pradeep Gaidhani, SUPER RESOLUTION,

http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/AV1011/Super_Resolution_CVonline.pdf

[2] Xuelong Li, Yanting Hu, Xinbo Gao and Beijia Ning, A MULTI-FRAME IMAGE SUPER-RESOLUTION METHOD

https://www.researchgate.net/publication/223740682_A_multi-frame_image_super-resolution_method

[3] Hardik Raval and Ashish Kothari, Analysis of Single Frame Super Resolution Methods,

<https://www.ijecs.in/issue/v3-i11/23%20ijecs.pdf>

[4] DeepLearning.net , Convolutional Neural Networks (LeNet),

<http://deeplearning.net/tutorial/lenet.html>

[5] Chao Dong, Chen Change Loy, Kaiming He and Xiaoou Tang, Image Super-Resolution Using Deep Convolutional Networks,

<https://arxiv.org/pdf/1501.00092.pdf>

[6] Kwang In Kim and Younghee Kwon, Example-based Learning for Single-Image Super-resolution,

<https://people.mpi-inf.mpg.de/~kkim/supres/LearningImageSuperresol.pdf>

[7] Wei-long Chen, Li Guo, Wu He, Wei Wu and Xiao-min Yang, Super-resolution Reconstruction Algorithm Based on Patch Similarity and Back-projection Modification,

<http://www.jsoftware.us/vol9/jsw0907-22.pdf>

[8] Qiang Zhou, Shifeng Chen, Jianzhuang Liu and Xiaoou Tang, Edge-Preserving Single Image Super-Resolution,

http://mmlab.ie.cuhk.edu.hk/archive/2011/mm11_edge_preserving_SR.pdf

[9] Seema S Bhat and Dr.M.C.Hanumantharaju, Edge Preserving Single Image Super Resolution Techniques– A Comprehensive Study,

<http://www.ijlret.com/Papers/NC3PS2016/29.%20Seema%20Bhat.pdf>

[10] Meyer, Yves (1992). Wavelets and Operators. Cambridge: Cambridge University Press. ISBN 0-521-42000-8.

[11] Hui Ji and Cornelia Fermuller, Robust wavelet-based super-resolution reconstruction: Theory and Algorithm,

http://www.math.nus.edu.sg/~matjh/depository/PAMI_2009_superresolution.pdf

[12] Mathew .K and Dr. S. Shibu, Wavelet based Technique for Super Resolution Image Reconstruction,

<http://research.ijcaonline.org/volume33/number7/pxc3875754.pdf>

[13] Jason Pacheco, Final Project: Super-Resolution From a Single Image (Writeup),

<http://cs.brown.edu/courses/csci1950-g/results/final/pachecoj/>

[14] Andrew Blake and Pushmeet Kohli, Introduction to Markov Random Fields,

https://mitpress.mit.edu/sites/default/files/titles/content/9780262015776_sch_0001.pdf

[15] Kishor P. Upla, Prakash P. Gajjar and Manjunath V. Joshi, A fast approach for edge preserving super-resolution,

<http://ieeexplore.ieee.org/document/6011945/>

[16] Anumolu Lasmika and K. Raveendra , Improving Quality of Satellite Image by Wavelet Transforming & Morphological Filtering,

https://www.ijirset.com/upload/2014/july/31_Improving.pdf

[17] Chao Dong, Chen Change Loyand Xiaoou Tang, Accelerating the Super-Resolution Convolutional Neural Network,

<http://mmlab.ie.cuhk.edu.hk/projects/FSRCNN.html>

[18] Kwang In Kim and Younghee Kwon, Example-based Learning for Single-Image Super-resolution,

<https://people.mpi-inf.mpg.de/~kkim/supres/supres.htm>

[19] Emil Mikulic, Haar Wavelet Transform,

<https://unix4lyfe.org/haar/>

[20] NATIONAL INSTRUMENTS, Peak Signal-to-Noise Ratio as an Image Quality Metric,

<http://www.ni.com/white-paper/13306/en/>

[21] Mathworks, Structural Similarity Index (SSIM) for measuring image quality,

<https://it.mathworks.com/help/images/ref/ssim.html>

DATA DI ULTIMA CONSULTAZIONE DEGLI URL: 30/08/2017

[22] Telegram Messenger, Telegram a new era of messaging,

<https://telegram.org>

[23] Lixia Zhang, A Retrospective View of Network Address Translation

<http://web.cs.ucla.edu/~lixia/papers/NAT-in-Retrospect.pdf>