Chapter 05. 이미지 복원 (Image Reconstruction)

정량 지표

MSE(L-2 Loss) Mean Squared Error

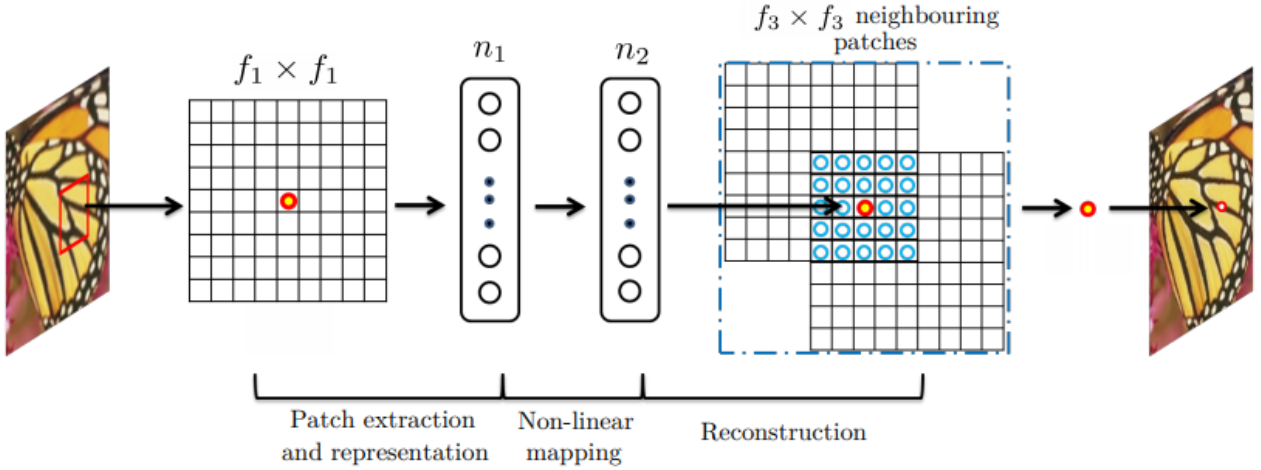


Fig. 3. An illustration of sparse-coding-based methods in the view of a convolutional neural network.

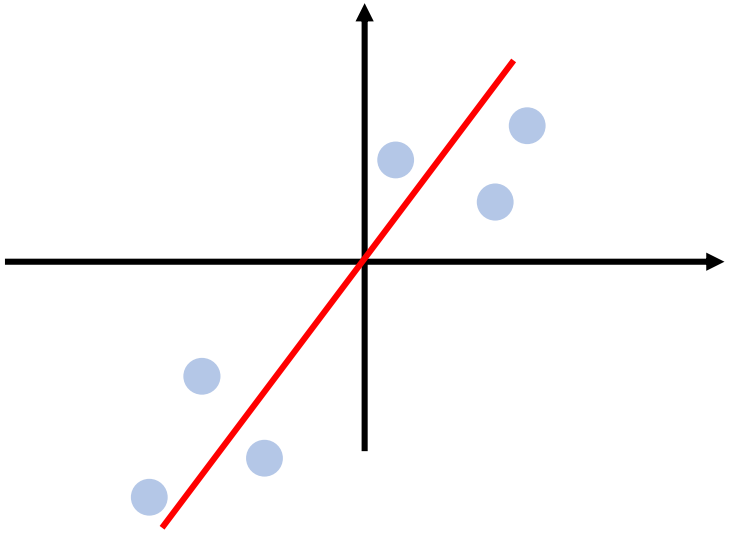
$$L(\Theta) = \frac{1}{n}\sum_{i=1}^{n} ||F(\mathbf{Y}_i;\Theta) - \mathbf{X}_i||^2, \qquad (4)$$

Squared(제곱)된 Error(잔차)의 Mean(평균)을 구하는 것이 MSE이다.

SRCNN에서 관습적으로 사용되었다.

http://mmlab.ie.cuhk.edu.hk/projects/SRCNN.html

MSE(L-2 Loss) Mean Squared Error

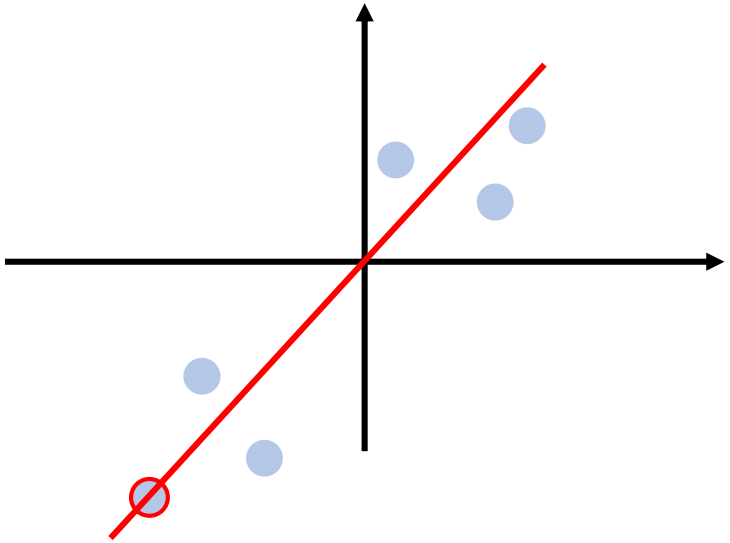


$$\arg\min_{w} \sum \|y_i - wx_i\|_2^2$$

- 에러가 클 수록 더욱 큰 페널티
- 데이터의 평균 – 존재하지 않는 값
- 데이터를 Smoothing하는 효과
- Outlier에 취약한 단점이 있음

보통 별 의심 없이 써 오던 MSE Loss는 '평균'을 나타내는 특성이 있다.

MAE(L-1 Loss) Mean Absolute Error



$$\arg\min_{w} \sum |y_i - wx_i|$$

- 에러가 커져도 동일한 페널티
- 데이터의 중간값 – 존재하는 정확한 값
- 존재하는 값을 사용하여 샤프한 특성
- 적게 존재하는 값을 무시하는 특성

MAE Loss는 '중간값'의 특성이 있으며, Outlier에 강건한 특성이 있다.

Loss Function Trends

TABLE 1
Parameters comparison of CNN-based SR algorithms. GRL stands for Global residual learning, LRL means Local residual learning, MST is abbreviation of Multi-scale training.

Method	Input	Output	Blocks	Depth	Filters	Parameters	GRL	LRL	MST	Framework	Loss
SRCNN	bicubic	Direct		3	64	57k				Caffe	ℓ_2
FSRCNN	LR	Direct		8	56	12k				Caffe	ℓ_2
ESPCN	LR	Direct		3	64	20k				Theano	ℓ_2
SCN	bicubic	Prog.	✓	10	128	42k				Cuda-CovNet	ℓ_2
REDNet	bicubic	Direct		30	128	4,131k	✓	✓		Caffe	ℓ_2
VDSR	bicubic	Direct		20	64	665k	✓	✓		Caffe	ℓ_2
DRCN	bicubic	Direct		20	256	1,775k	✓			Caffe	ℓ_2
LapSRN	LR	Prog.	✓	24	64	812k	✓			MatConvNet	ℓ_1
DRRN	bicubic	Direct	✓	52	128	297k	✓	✓	✓	Caffe	ℓ_2
SRGAN	LR	Direct	✓	33	64	1500k				Theano/Lasagne	ℓ_2
DnCNN	bicubic	Direct		17	64	566k			✓	MatConvNet	ℓ_2
IRCNN	bicubic	Direct		7	64	188k			✓	MatConvNet	ℓ_2
FormResNet	bicubic	Direct	✓	20	64	671k	✓		✓	MatConvNet	ℓ_2, ℓ_{TV}
EDSR	LR	Direct	✓	65	256	43000k	✓	✓		Torch	ℓ_1
MDSR	LR	Direct	✓	162	64	8,000k	✓	✓	✓	Torch	ℓ_1
ZSSR	LR	Direct		8	64	225k	✓			Tensorflow	ℓ_1
MemNet	bicubic	Direct	✓	80	64	677k	✓	✓	✓	Caffe	ℓ_2
MS-LapSRN	LR	Prog.	✓	84	64	222k	✓	✓	✓	MatConvNet	ℓ_1
CMSC	bicubic	Direct	✓	35	64	1220k	✓	✓	✓	PyTorch	ℓ_2
CNF	bicubic	Direct		15	64	337K				Caffe	ℓ_2
IDN	LR	Direct	✓	31	64	796k	✓	✓		Caffe	ℓ_2, ℓ_1
BTSRN	LR	Direct	✓	22	64	410K	✓	✓		Tensorflow	ℓ_2
SelNet	LR	Direct		22	64	974K	✓	✓		MatConvNet	ℓ_2
CARN	LR	Direct	✓	32	64	1,592K	✓	✓	✓	PyTorch	ℓ_1
SRMD	LR	Direct		12	128	1482k				MatConvNet	ℓ_2
SRDenseNet	LR	Direct	✓	64	16-128	-	✓	✓		TensorFlow	ℓ_2
EnhanceNet	LR	Direct	✓	24	64	-		✓		TensorFlow	ℓ_2, ℓ_t, GAN
SRFeat	LR	Direct	✓	54	128	-	✓	✓		TensorFlow	ℓ_2, ℓ_p, GAN
SRRAM	LR	Direct	✓	64	64	1,090K	✓	✓	✓	Tensorflow	ℓ_1
D-DBPN	LR	Direct	✓	46	64	10000K	✓	✓		Caffe	ℓ_2
RDN	LR	Direct	✓	149	64	21900k	✓	✓		Torch	ℓ_1
ESRGAN	LR	Direct	✓	115	64	-	✓	✓		Pytorch	ℓ_1
RCAN	LR	Direct	✓	500	64	16,000k	✓	✓	✓	Pytorch	ℓ_1

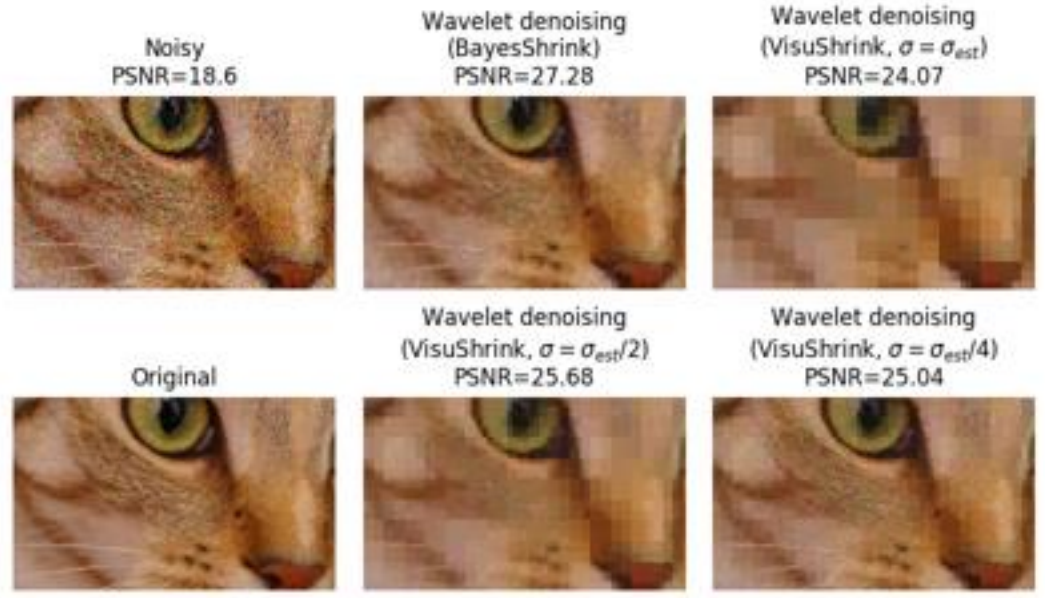
L-2 loss (MSE)보다 L-1 loss의 채용이 더 많아지는 추세이다. GAN Loss와 TV Regularization도 보인다.

PSNR Peak Signal-to-Noise Ratio

255 or 1.0

$$PSNR = 10\log_{10}\left(\frac{MAX_I^2}{MSE}\right)$$

$$= 20\log_{10}\left(\frac{MAX_I}{\sqrt{MSE}}\right)$$



PSNR은 잡음(에러)가 얼마나 되는지 측정하는 성능 지표 중 하나로, 영상 복원을 평가하는 데에 많이 사용한다.

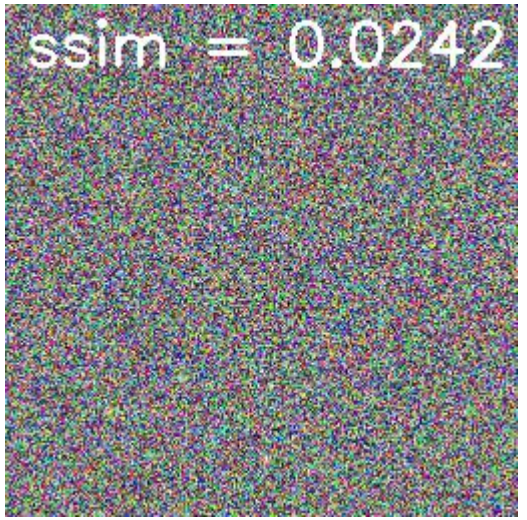
Image Reconstruction/Restoration 외에도 Image Encoding/Decoding 평가에 많이 쓰인다.

SSIM Structural Similarity Index Measure

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

with:

- μ_x the average of x;
- μ_y the average of y;
- σ_x^2 the variance of x;
- σ_y^2 the variance of y;
- σ_{xy} the covariance of x and y;
- $c_1=(k_1 L)^2$, $c_2=(k_2 L)^2$ two variables to stabilize the division with weak denominator;
- L the dynamic range of the pixel-values (typically this is $2^{\#bits\ per\ pixel}-1$);
- $k_1=0.01$ and $k_2=0.03$ by default.



구조적인 특징이 얼마나 일치하는지 측정하는 SSIM. 미분 가능하기 때문에 Loss function으로도 사용된다.

https://en.wikipedia.org/wiki/Structural_similarity