Training Patch Analysis and Mining Skills for Image restoration Deep Neural Networks

Abstract

Problem : Network architecture development paradigm → 재현 가능성 ↓, real-world 확장 ↓

Solution : patch selection guideline 제시

Conclusion

① patch extraction guideline is below

② 실제 성능도 architecture 개선·NAS 못지 않음

I. Intro

선행연구 개관

supervised learning : small-size data → overfitting, computation ↑

semi/weakly supervised "

unsupervised learning : domain adoption method

active learning :

현재 self-supervised learning의 우려 : architecture에 집중

∵ $y = Hx + n$으로 noise image ∞개 생성 가능

핵심 question 7개

Regarding the above-stated problem, we raise several questions in this paper. We first raise obvious and intuitive simple questions:

1) Does overfitting occur in image restoration networks?
2) How good is the geometric (flip-rotation) data augmentation?

Then, we further raise some controversial questions that may be informative and eventually can provide a guideline of patch mining skill:

4) What is the effect of the number of training patches?
5) How many patches are required to bring out the capacity of the network fully?
6) Is exploiting uncertain or hard samples effective for image restoration, as evidenced in high-level vision?

Finally, we raise some questions regarding the generalization ability:

7) Are the good patches generic across different image restoration tasks?
8) Are the number of patches and the model size related?

답변은 V. Experiments 참고!

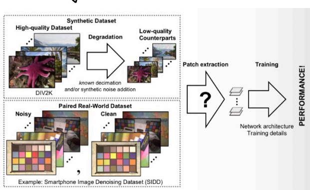
Motivation ① 재현 가능성 ↓

② 학습 효율 ↓ ∵ low-frequency data 多

Contribution ① 7개 질문에 대한 답변

② patch extraction guideline

③ verify guideline is useful



II. Related work

A. Image Restoration based on Deep Networks

- image denoising : DnCNN
  - FFDNet
  - Two-stage networks
  - denoisers for real-world images
- super-resolution : development of architecture
  - ⋮
  - NAS approach (FALSR, OISR)

B. Data-relevant Strategies for Deep Networks

- property of training images is important
  - patch-size ↑ ⇒ performance ↑
  - image quality ↑ ⇗
- hard example mining ; deep metric learning
- active learning ; efficiently train (budget ↓)

III. Backgrounds

A. Image Restoration

- image denoising $y = x + n$ (K=I)
- image super-resolution $y = (x * k)\downarrow_s + n$

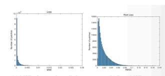
IV. Settings

A. Environments and Comparisons

- Network Architecture : EDSR baseline
- Training Dataset : DIV2K
  - 151,300 patches
  - label patch-size 96x96 stride 120
- Representative Values
  - uncertainty (=loss)
  - mean gradient magnitude
  - std of a patch
  - high frequency power (=power)

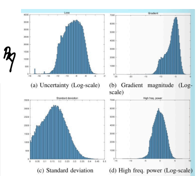
B. patch statistics

– low-informative patch 多



(a) MSE of network prediction (b) Root MSE

Fig. 2: Histogram of loss and square root of loss.



(a) Uncertainty (Log-scale) (b) Gradient magnitude (Log-scale)

(c) Standard deviation (d) High freq. power (Log-scale)

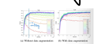
Fig. 3: Histogram of the statistics of patches.

V. Experiments

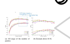
Regarding the above-stated problem, we raise several questions in this paper. We first raise obvious and intuitive simple questions:
1) Does overfitting occur in image restoration networks?
2) How good is the geometric (flip-rotation) data augmentation?
Then, we further raise some controversial questions that may be informative and eventually can provide a guideline of patch mining skill:
4) What is the effect of the number of training patches?
5) How many patches are required to bring out the capacity of the network fully?
6) Is exploiting uncertain or hard samples effective for image restoration, as evidenced in high-level vision?
Finally, we raise some questions regarding the generalization ability:
7) Are the good patches generic across different image restoration tasks?
8) Are the number of patches and the model size related?

A. Evaluations (# of patches : 10K ~ 150K)

① overfitting은 (patch-size가 < 30K / data augmentation X)면 생긴다



② geometric augmentation은 평균적으로 0.5dB 이득 PSNR



③ number of training patches는 greenish가 best (around 60K)

④ patch의 개수는 전반 정도만 있는 게 성능 best

TABLE I: PSNR results against the number of training patches. '+' denotes the results with data augmentation.

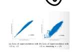
Criteria	Number of patches				
	10 K	20 K	30 K	40 K	50 K
Loss	30.76	31.17	31.32	31.45	31.54
Grad	30.90	31.29	31.44	31.56	31.64
Loss+	31.42	31.76	31.89	32.00	32.05
Grad+	31.55	31.87	31.98	32.05	32.07

⑤ hard sample과 복원 성능은 무관.

(because regression task는 outlier 반영 → 성능 ↓)

⑥ scaling factor x2, x3, σ=25는 모두 상관 ↓

(아무거나 사용해도 무관)



⑦ model size와 # of training patch는 상관 ↓

B. proposed Mining skills

① stride는 patch-size보다 더 크게 설정 (non-overlap)

② mean gradient magnitude가 큰 patch 절반만 학습시키기

③ data augmentation 적용하기 (cf. flip-rotation)

VI. Discussion

A. Comparison with Benchmarks : 성능 ≒ or better

B. Result of Image Denoising : Grad criteria is best