CL-DPS: A Contrastive Learning Approach to Blind Inverse Problem Solving via Diffusion Posterior Sampling

Linfeng Ye, Pallavi Ferrao
Affiliations, University of Toronto

Motivation

Figure 1. Results of blind rotation deblurring, a challenging **non-linear** inverse problem: (a) ground truth image, (b) rotation blurred measurement, and restored images using (c) BlindDPS [11], (d) FastEM [35], (e) GibbsDDRM [41], and (f) CL-DPS (ours). Notably, all methods fail catastrophically except for CL-DPS.

Real-world applications frequently involve blind inverse problems with unknown measurements. Existing DM-based methods for blind inverse problems are limited, primarily addressing only linear measurements and thus lacking applicability to real-life scenarios that often involve non-linear operations. To overcome these limitations, we propose CL-DPS, a novel approach based on contrastive learning for solving blind inverse problems via diffusion posterior sampling.We train an auxiliary deep neural network (DNN) offline using a modified version of MoCo, a contrastive learning technique. This auxiliary DNN serves as a likelihood estimator, enabling estimation of p(y|x) without prior knowledge of the measurement operator, thereby adjusting the reverse path of the diffusion process for inverse problem solving.

Related Work

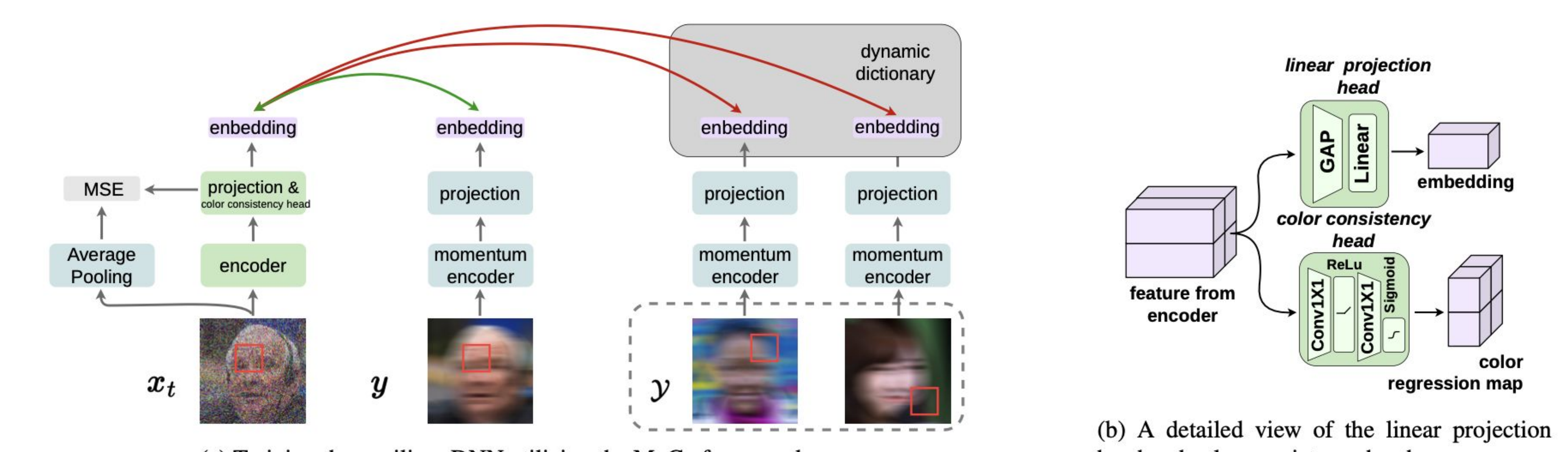
For non-blind inverse problems, methods such as diffusion posterior sampling (DPS) [1] and pseudo-guided diffusion models [48] leverage Tweedie's formula [2] to approximate the smoothed likelihood. Similarly,singular-value decomposition (SVD)-based techniques [32] are applied for related purposes. Conversely, for blind inverse problems, alongside the approaches discussed in Sec.1 [3], introduced Blind RED-Dif, an extension of the RED-diff framework [4]. This method employs variational inference to jointly estimate both the latent image and the unknown forward model parameters, addressing the challenges of unknown measurement operators.As a versatile semi-supervised learning framework, contrastive learning learns useful feature representation by clustering positive samples and dispersing negative samples. It achieves great success since instance discrimination has been proposed in [5]. Since then [6,7] advanced the field by leveraging diverse data augmentation methods and using projection head during the contrastive learning process. [6] used a momentum update mechanism to maintain a negative sample generator, rather than a physical queue of negative examples to reduce the memory consumption.

References

[1] Hyungjin Chung, Jeongsol Kim, Michael Thompson Mc-cann, Marc Louis Klasky, and Jong Chul Ye. Diffusion posterior sampling for general noisy inverse problems. In The Eleventh International Conference on Learning Representations, 2023.
[2]Bradley Efron. Tweedie's formula and selection bias. Journal of the American Statistical Association, 106(496):1602– 663 1614, 2011.
[3]Hyungjin Chung, Jeongsol Kim, Sehui Kim, and Jong Chul Ye. Parallel diffusion models of operator and image for blind inverse problems. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 6059–6069, 2023.
[4] Morteza Mardani, Jiaming Song, Jan Kautz, and Arash Vahdat. A variational perspective on solving inverse problems with diffusion models. In The Twelfth International Conference on Learning Representations
[5] Zhirong Wu, Yuanjun Xiong, Stella X Yu, and Dahua Lin. Unsupervised feature learning via non-parametric instance discrimination. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 3733–3742, 2018.
[6]Kaiming He, Haoqi Fan, Yuxin Wu, Saining Xie, and Ross 686 Girshick. Momentum contrast for unsupervised visual rep- 687 resentation learning. In Proceedings of the IEEE/CVF con- 688 ference on computer vision and pattern recognition, pages 689 9729–9738, 2020.
[7]Ting Chen, Simon Kornblith, Mohammad Norouzi, and Ge- offrey Hinton. A simple framework for contrastive learning of visual representations. In International conference on ma- chine learning, pages 1597–1607. PMLR, 2020.

New Technique

We propose CL-DPS, an inverse problem solver using diffusion models for the blind setting. CL-DPS incorporates an auxiliary DNN, trained using MoCo, to serve as a likelihood estimator. Unlike previous blind solvers, which are limited to recovering images only under linear measurements, CL- DPS is capable of recovering images for both linear and non-linear measurements.
Estimating the posterior pt(x(t)|y) requires an estimation of the likelihood pt(y|x(t)). To achieve this, we aim to train an auxiliary DNN offline (prior to applying diffusion models for inverse problem-solving) which is able to estimate the likelihood pt(y|x(t)). Note that at this the time the measurement parameters ψ are unknown. This auxiliary DNN will then be employed during the diffusion- based inverse problem-solving process to adjust the reverse diffusion path accordingly.



(a) Training the auxiliary DNN utilizing the MoCo framework.

(b) A detailed view of the linear projection head and color consistency head.

$$\mathcal{L}_{p(\boldsymbol{y}|\boldsymbol{x}_t)} = -\log \frac{\exp(\langle f(\boldsymbol{x}_t), f(\boldsymbol{y})\rangle/\tau)}{\sum_{\tilde{\boldsymbol{y}}\in\mathcal{Y}} \exp(\langle f(\boldsymbol{x}_t), f(\tilde{\boldsymbol{y}})\rangle/\tau)}.$$

Experimental Results

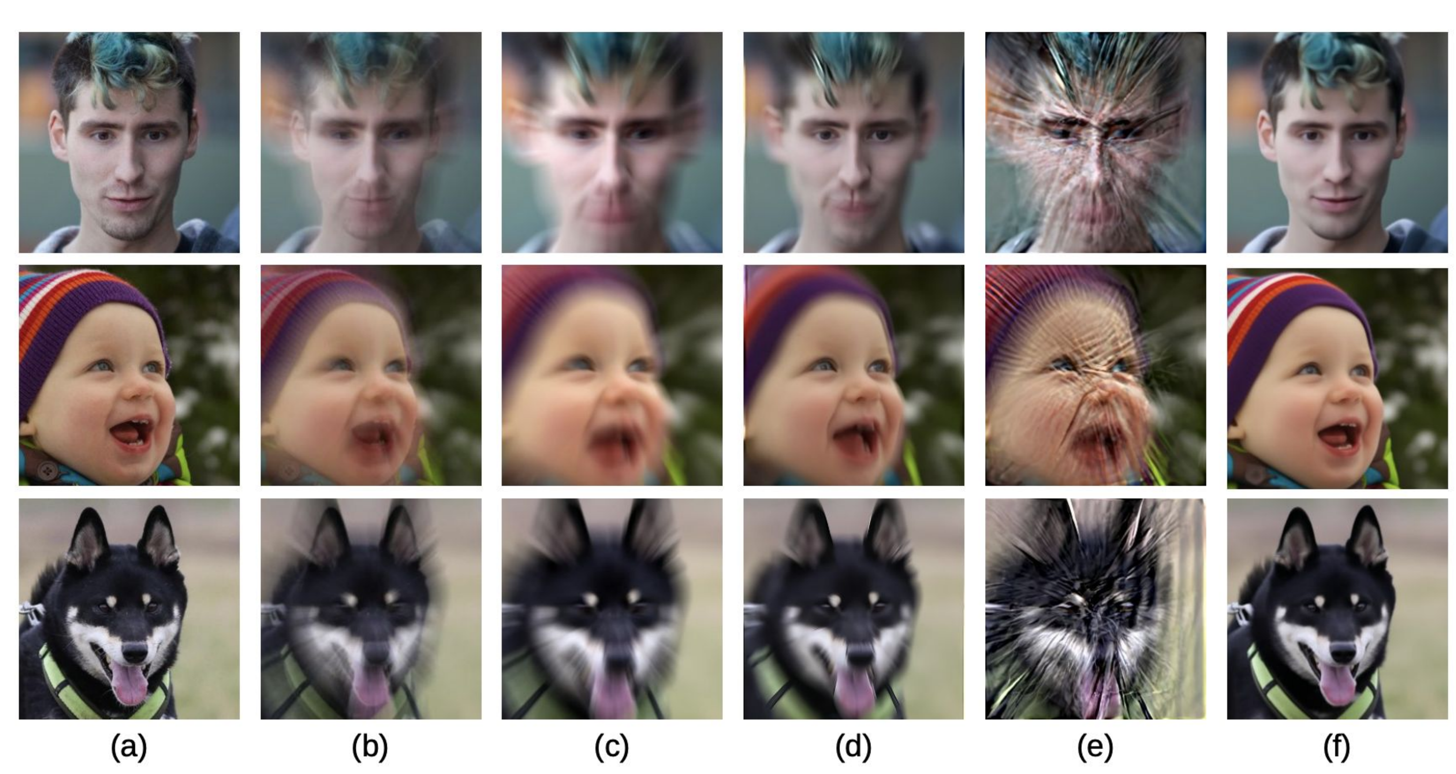


Figure 7. Results of blind zoom deblurring, a challenging **non-linear** inverse problem: (a) ground truth image, (b) zoom blurred measurement, and restored images using (c) BlindDPS [11], (d) FastEM [35], (e) GibbsDDRM [41], and (f) CL-DPS (ours). Notably, all methods fail catastrophically except for CL-DPS.

Method	FFHQ (256 × 256)						AFHQ (256 × 256)					
	Rotation			Zoom			Rotation			Zoom		
	PSNR ↑	FID ↓	LPIPS ↓	PSNR ↑	FID ↓	LPIPS ↓	PSNR ↑	FID ↓	LPIPS ↓	PSNR ↑	FID ↓	LPIPS ↓
CL-DPS (Ours)	**22.74**	**33.66**	**0.302**	**20.68**	**42.61**	**0.435**	**21.46**	**36.96**	**0.319**	**19.63**	**57.54**	**0.468**
BlindDPS [11]	16.87	343.76	0.552	16.39	292.91	0.780	13.25	200.46	0.674	11.75	279.57	0.607
FastEM [35]	15.96	268.43	0.597	18.68	303.25	0.623	11.57	289.19	0.680	15.60	310.06	0.797
GibbsDDRM [41]	18.43	236.55	0.565	15.45	327.42	0.802	15.24	263.49	0.628	14.57	280.54	0.549

Table 1. **Non-linear** blind inverse problems: Blind rotation and zoom deblurring results on the FFHQ and AFHQ datasets for CL-DPS and benchmark methods. CL-DPS successfully restores the input images with high quality, whereas all other methods fail. **Bold** and underlined values denote the best and second-best results, respectively.

Method	FFHQ (256 × 256)						AFHQ (256 × 256)					
	Motion			Gaussian			Motion			Gaussian		
	PSNR ↑	FID ↓	LPIPS ↓	PSNR ↑	FID ↓	LPIPS ↓	PSNR ↑	FID ↓	LPIPS ↓	PSNR ↑	FID ↓	LPIPS ↓
CL-DPS (Ours)	22.93	32.44	0.157	**24.82**	**26.64**	0.348	**22.06**	42.25	0.280	**23.76**	20.56	**0.225**
SelfDeblur [45]	10.83	270.0	0.717	11.36	235.4	0.686	9.081	300.5	0.768	11.53	172.2	0.662
DeblurGANv2 [34]	17.75	220.7	0.571	19.69	185.5	0.529	17.64	186.2	0.597	20.29	86.87	0.523
Pan_l0 [44]	15.53	242.6	0.542	19.94	92.70	0.415	15.34	235.0	0.627	21.41	62.76	0.395
BlindDPS [11]	22.24	**29.49**	0.281	24.77	27.36	**0.233**	20.92	**23.89**	0.338	23.63	**20.54**	0.287
FastEM [35]	24.68	-	0.34	-	-	-	-	-	-	-	-	-
LatentDEM [56]	22.65	-	0.167	-	-	-	-	-	-	-	-	-
GibbsDDRM [41]	**25.80**	38.71	**0.115**	-	-	-	22.01	48.00	**0.197**	-	-	-

Table 2. **Linear** blind inverse problems: Blind motion and Gaussian deblurring results on the FFHQ and AFHQ datasets for CL-DPS and benchmark methods. CL-DPS achieves competitive results compared to other benchmark methods.