ADMM-Based Image Deconvolution with Conditioned Diffusion Prior

Jacob Deutsch

Abstract—This paper presents a new approach in image deconvolution, merging the Alternating Direction Method of Multipliers (ADMM) with class-conditioned diffusion-based models. ADMM has established itself as a potent tool in image deconvolution, a key process for enhancing image clarity in various applications. Concurrently, diffusion models have gained prominence for their adaptability in computational imaging, notably excelling in generation tasks with diverse forms of conditioning. Their inherent aptitude for denoising tasks makes them particularly suitable for image deconvolution. This study explores the integration of class-conditioned diffusion models with ADMM, aiming to improve deconvolution performance by combining the unique strengths of both methodologies. The objective is to evaluate this integration and explore the impact of class conditioning on the deconvolution process.

1 Introduction

Image deconvolution is a key process in improving the quality of images in fields such as astronomy and medical imaging. This process is challenging because it involves solving an ill-posed inverse problem, where the goal is to recover clear images from blurred and noisy data. Alternating Direction Method of Multipliers (ADMM) is widely used for this purpose because of its effectiveness and flexibility. However, ADMM requires the use of denoising priors to stabilize the deconvolution process and improve results. Common denoising priors include Total Variation (TV), Non-Local Means (NLM), and Deep Neural Network-based methods like DnCNN.

In this paper, we propose combining ADMM with class-conditioned diffusion-based models. Diffusion models are adaptive and have shown excellent performance in image generation tasks. By integrating these models with ADMM, we aim to leverage their denoising ability and add context-specific guidance to the deconvolution process through class conditioning. This approach has the potential to show an improvement over traditional priors and improve the quality of deconvolved images. GitHub: https://github.com/jacobDeutsch10/admm-diffusion-ee367

2 Related Work

2.1 ADMM

Alternating Direction Method of Multipliers(ADMM) [1] is an optimization algorithm that is popular for solving a wide variety of MAP estimation and regularized inverse problems. ADMM solves problems of the form:

$$\begin{aligned} \text{minimize} \quad & f(x) + g(z) \\ \text{s.t.} \quad & Ax + Bz = c \end{aligned} \tag{1}$$

Venkatakrishnan et. al. [2] developed a framework for using ADMM to solve the problem of image deconvolution. In image deconvolution, we attempt to recover an image $x \in \mathbb{R}^N$ from a vector $b \in \mathbb{R}^M$ of noisy measurements where

$$b = Ax + \eta \tag{2}$$

where $A \in \mathbb{R}^{M \times N}$ is the linear image formation model and $\eta \in \mathbb{R}^M$ is independent additive noise. Their method proposes using ADMM to solver the regularized problem

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & \frac{1}{2}\|Ax - b\|_2^2 + \lambda \Psi(z) \\ \text{s.t.} \quad & Dz - x = 0 \end{aligned} \tag{3}$$

Where the first term in the objective is referred to as the data fidelity term and the second is the prior. The prior is used to constrain the space of possibilities. Importantly, this formulation allows users to select many different types of denoising priors in a ’plug-and-play’ setup. Denoising algorithms such as NLM, BM3D, and TV have all been shown to work well in this setup [2]. In addition, learned denoisers such as DnCNN have been shown to provide good results [3].

2.2 Denoising Diffusion Models

Denoising Diffusion Probabilistic Models (DDPMs) have recently emerged as state of the art models on generation tasks [4] [5]. DDPMs are trained by passing in a noisy image and repeatedly denoising it. At inference time DDPMs, can generate high fidelity samples by repeatedly applying denoising steps to random noise. In addition, diffusion models are able to be conditioned on text [4] or class labels [5] for more targeted image generation. It is natural then to see the potential to use these state of the art models as a denoising prior for the task of image deconvolution.

[3] applied a pretrained diffusion model as a denoising prior for ADMM image deconvolution with promising results, by running the diffusion process for a fixed number of steps depending on the variance of the noise η. However, they did not experiment with conditioning the diffusion model and only applied their method to two test images. The diffusion method achieved peak-signal-to-noise-ratios(PSNR) similar to other priors such as DnCNN, TV, and NLM. However, they conducted extensive tuning to find the best number of ADMM iterations to achieve the optimal PSNR for various priors. We will adopt the same number of iterations for each method.

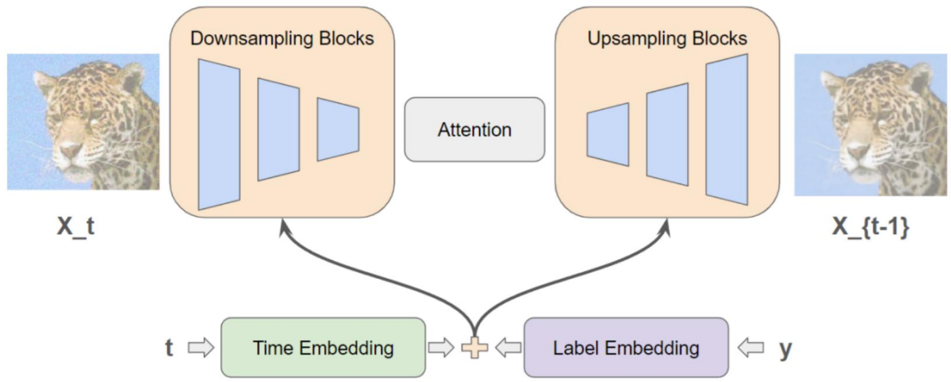

Fig. 1. Conditional Diffusion

In addition to the deconvolution task, diffusion models have been shown to perform well on a variety of inverse problems. [6] showed that diffusion could be used as a prior in a plug-and-play Half-Quadratic Splitting (HQS) framework to perform image inpainting, motion deblurring, and image super-resolution. They also do work in establishing schedules for diffusion priors as the key weakness for diffusion priors in inverse problems is needing to evaluate a large neural network many times per outer optimization iteration.

3 Proposed Method

Our proposed method involves incorporating diffusion models with and without class conditioning to perform image deconvolution. We use these diffusion models as denoising priors for ADMM-based deconvolution and compare results to NLM, TV, bilateral filtering, and DnCNN priors.

Both the conditional and unconditional diffusion models are taken from [5]. The unconditional diffusion model is a Unet architecture with ResNet downsampling blocks, attention middle blocks, and ResNet upsampling blocks. The unconditional diffusion model does take conditioning in the form of the current timestep of the diffusion process (as with all diffusion models). This timestep conditioning is applied by generating sinusoidal embeddings of the current timestep and applying two fully connected layers to the embedding before adding it to the output of each convolutional layer.

The conditional model is the same architecture as the unconditional model, but also takes as input the class label of the image. The class conditioning is performed by passing the class label to a learned embedding layer and then adding the class label embedding to the timestep embedding. The architecture is illustrated in Fig 1. We instantiate both models from pretrained checkpoints provided by [5]. These checkpoints are trained on 256x256 images from the ImageNet dataset.

We apply these models to the plug-and-play ADMM framework by applying a set number of denoising steps to the image at each outer iteration of ADMM. To set the number of denoising timesteps, we use the same heuristic as [3]. Number of timesteps is calculated by finding the timestep with noise variance closest to the noise variance of η.

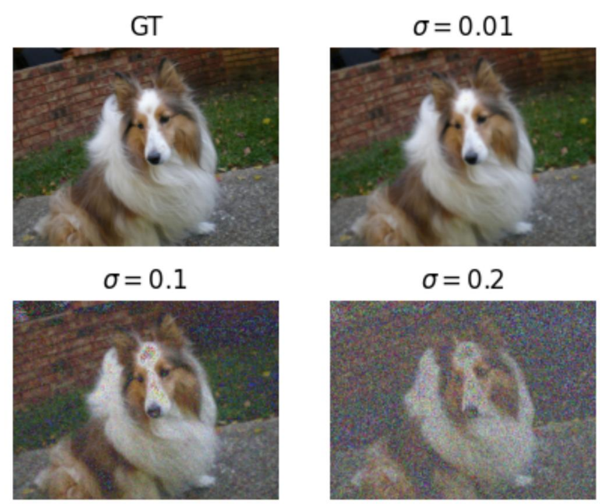

Fig. 2. Example Image at Different Noise levels

3.1 Experiment Setup

The proposed method of incorporating diffusion models as denoising priors for ADMM-based image deconvolution was evaluated on a small subset of the ImageNet validation dataset. The images were resized to 256x256 pixels and degraded by applying a blurring kernel of size 15x15 and three different levels of additive Gaussian noise with standard deviations of 0.2, 0.1, and 0.01. Fig. 2 shows an example image comparing the ground truth the the ADMM input at each noise level. The performance of both conditional and unconditional diffusion models from the pre-trained checkpoints was compared against traditional denoising priors such as Total Variation (TV), Non-Local Means (NLM), bilateral filtering, and DnCNN. Unfortunately, due to computational restrictions, we could only evaluate all methods over 20 sample images from the ImageNet validation set.

The diffusion models were integrated into the ADMM framework using the ”plug-and-play” approach, where the denoising process was run for a set number of timesteps at each outer iteration of ADMM. The number of denoising timesteps was determined by finding the timestep with noise variance closest to the noise variance of the additive noise as in [3]. For the conditional diffusion model, the class labels from the ImageNet dataset were used for class conditioning.

The ADMM algorithm was applied to solve the regularized inverse problem for image deconvolution for the number of iterations specified in [3] for each of the different priors. We evaluate the different priors qualitatively and quantitatively, by comparing the mean PSNR and Structural Similarity Index Measure (SSIM).

4 Experimental Results

4.1 Quantitative Evaluation

The quantitative results presented in the table offer valuable insights into the performance of various denoising priors for the image deconvolution task across different noise levels. At the highest noise level examined ($\sigma = 0.2$), the unconditional diffusion model prior performs well, achieving a PSNR of 23.1 dB and SSIM of 0.51, which is slightly lower than the DnCNN prior's PSNR of 23.68 dB and SSIM of 0.54. This suggests that the unconditional diffusion model prior is good at denoising and preserving image structure even in highly noisy scenarios. The conditional diffusion model prior slightly lags behind, with a PSNR of 23.08 dB and SSIM of 0.5, indicating that the class conditioning does not have a large effect in this setting. All three of the learned denoisers outperform the traditional priors at the highest noise level underscoring the effectiveness of learned priors for this task.

As the noise level decreases to $\sigma = 0.1$, the performance gap between the priors narrows, with the DnCNN prior matching the TV prior's PSNR of 24.65 dB and SSIM of 0.61. Both diffusion architectures have very similar yet slightly lower PSNRs of 24.45 dB and 24.42 dB for the unconditional and conditional models respectively. This gap is quite small and could go away if evaluating on more images. Once again, the conditioned and unconditioned diffusion models do not seem to differ much in their performance with the unconditioned diffusion model performing slightly better

At the lowest noise level ($\sigma = 0.01$), both the unconditional and conditional diffusion priors outshine all other methods, achieving the highest PSNR of 26.8 dB and SSIM of 0.75. This performance in low-noise scenarios underscores the effectiveness of the diffusion model priors in preserving image quality and structure during the deconvolution process. Although, still, adding class conditioning does not seem to have much of an effect on performance.

In contrast, the traditional priors, with the exception of TV, exhibit poor performance across all noise levels. The NLM and Bilateral priors consistently underperform, with PSNR values around 12-13 dB and SSIM values around 0.3-0.5, suggesting that they are not well-suited for the deconvolution task in this experiment. While the TV prior fares better than NLM and Bilateral, it still trails behind the learned denoisers (diffusion model priors and DnCNN) at lower noise levels, further highlighting the advantages of learned priors over traditional methods for this image deconvolution task.

These results serve as a proof of concept for the potential of diffusion model priors as effective denoising tools for image deconvolution. Their ability to outperform methods like DnCNN in low-noise scenarios (= 0.01) is noteworthy, where both the unconditional and conditional diffusion priors achieve the highest PSNR of 26.8 dB and SSIM of 0.75. While adding class conditioning does not seem to have much of an effect on performance in this specific experiment, both diffusion models performed well overall and achieved comparable results to DnCNN and TV at all noise levels.

4.2 Qualitative Results

In addition to our quantitative analysis, we qualitatively evaluate a number of sample images. Figure 3 shows a comparison of the various methods at the highest noise level ($\sigma = 0.2$) as well as the Ground Truth and input noisy images. We see that both NLM and bilateral both fail to recover a high fidelity image from the noisy input. Both methods produce a noisy image with noticeable darkening and blurring in the case of NLM. While TV achieves relatively high PNSR and SSIM statistics compared to the other traditional denoisers, the TV images are still noticeably noisy.

The learned denoisers far outperform the traditional denoisers. Upon inspection of the example images the case for the diffusion model becomes stronger than simply looking at PSNR and SSIM. Both diffusion models produce crisp images with very little noticeable noise remaining in the image. This is especially apparent in the image of the snake in the top row. The unconditioned and conditional diffusion models both produce crisp images while the image produced with the DnCNN prior still has some noticeable noise. This difference in the blownup version of the same images in Fig 4. From these it is clear that the DnCNN prior produces images that are visually more noisy. However, the DnCNN images have higher PSNRs and SSIMs. This may mean that the diffusion models may have better performance in denoising while sacrificing the structural fidelity and intricate details of the image. This is reinforced by the second image. The diffusion models seem to muddle the faces of the two skiers and blur the details of the mountain range in the backdrop. The third image of the

	TV		**NLM**		**Bilateral**		**DnCNN**		**Diffusion**		**Diffusion Conditioned**	
σ	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
0.2	20.94	0.36	12.67	0.32	12.87	0.3	**23.68**	**0.54**	23.1	0.51	23.08	0.5
0.1	24.65	0.61	12.68	0.32	12.95	0.4	**24.65**	**0.61**	24.45	0.58	24.42	0.57
0.01	25.72	0.67	12.73	0.33	12.98	0.49	25.03	0.64	**26.80**	**0.75**	**26.8**	**0.75**

TABLE 1
PSNR and SSIM of all methods at each noise level

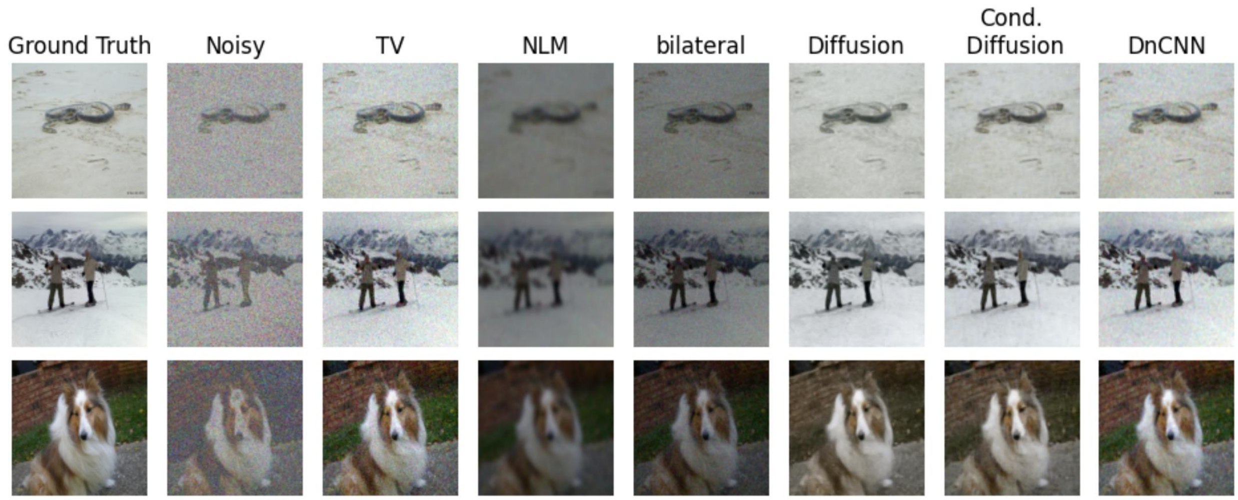

Fig. 3. Example Images at $\sigma = 0.2$

dog reinforces this and makes another trend more apparent. The diffusion models noticeably reduce the intensity of the colors particularly the green of the grass becomes more gray and the light hair on the dog's face becomes less bright compared to the rest of the dog's hair. DnCNN does a much better job of maintaining the vibrancy of the colors when compared to the ground truth. In addition, the diffusion models seem to due worse at recovering the high-frequency detail of the dog's fur when compared to DnCNN. Notably, the conditional diffusion model seems to do a slightly better job at recovering the dogs hair, and perhaps this is the result of the additional contextual information that the image contains a dog. Although it is difficult to draw any concrete conclusions from this one example, and the conditional model did not show any improvement over the conditioned model in our qualitative analysis.

5 Discussion

The experimental results clearly demonstrate the superiority of learned denoisers, particularly diffusion models and DnCNN, over traditional denoising methods like TV, NLM, and bilateral filtering for the image deconvolution task. This underscores the importance of leveraging powerful data-driven approaches to address challenging inverse problems in image processing. However, it is noteworthy that at higher noise levels (= 0.2), the DnCNN prior outperformed both the unconditional and conditional diffusion model priors in terms of quantitative metrics like PSNR and SSIM. This highlights the potential trade-off between denoising performance and computational complexity, as DnCNN is a much lighter and faster model compared to the large diffusion models.

While diffusion models showed impressive denoising capabilities, especially at lower noise levels, their practical deployment for image deconvolution tasks may be hindered by their substantial computational requirements. Running multiple iterations of the diffusion process at each outer ADMM iteration can be prohibitively expensive, particularly for high-resolution images or real-time applications. In contrast, DnCNN only needs to be evaluated once per ADMM iteration, making it more computationally efficient. Future work could explore strategies to accelerate diffusion models or develop more efficient architectures that retain their denoising prowess while reducing computational overhead.

Interestingly, the experiment did not yield significant improvements with class conditioning for the diffusion models. This could be attributed to the fact that images within the same ImageNet class may not necessarily share strong structural similarities, which are crucial for effective denoising and deconvolution. Future research could investigate more robust conditioning strategies, such as leveraging natural language descriptions or leveraging multi-modal information (e.g., text and image) to better capture the contextual information needed for effective class-conditional denoising.

Furthermore, the qualitative analysis revealed that while diffusion models excel at removing noise, they may sometimes compromise fine details, color vibrancy, and high-frequency information. This trade-off between denoising and detail preservation warrants further investigation to strike an optimal balance. Potential avenues could include exploring hybrid approaches that combine the strengths of diffusion models and more detail-preserving methods like DnCNN, or developing new architectures that can better retain image fidelity while effectively denoising.

6 Conclusion

In this study, we evaluated the use of diffusion models and other learned denoisers like DnCNN as priors for ADMM-based image deconvolution. The results showed that these learned denoisers clearly outperformed traditional denoising methods such as TV, NLM, and bilateral filtering across different noise levels. At high noise levels

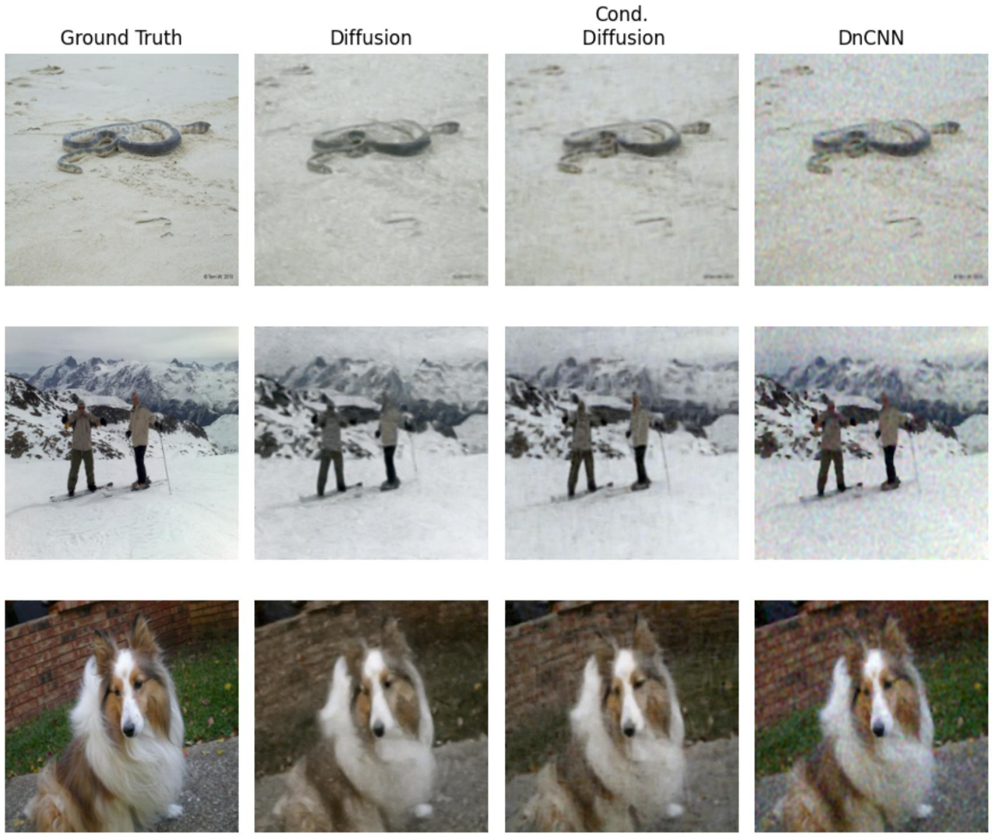

Fig. 4. Example Images at $\sigma = 0.2$

($\sigma = 0.2$), DnCNN achieved the best quantitative performance in terms of PSNR and SSIM metrics. However, at lower noise levels ($\sigma = 0.01$), both the unconditional and conditional diffusion model priors outperformed all other methods.

While the diffusion models demonstrated impressive denoising capabilities, especially at low noise, they come with a significant computational cost due to having to run multiple iterations of the diffusion process per ADMM iteration. DnCNN, being a much lighter model, is more computationally efficient as it only needs to be evaluated once per iteration.

The experiment did not show a clear benefit of using class conditioning for the diffusion models in this deconvolution task. One potential reason is that images within the same ImageNet class may not share strong structural similarities required for effective class-conditional denoising.

Qualitative analysis revealed some trade-offs with the diffusion model outputs. While excelling at noise removal, they sometimes compromised fine details, color vibrancy, and high-frequency information compared to DnCNN and the ground truth images.

Future work could explore strategies to improve the computational efficiency of diffusion models, develop better conditioning approaches leveraging multimodal data, and strike a balance between denoising performance and detail preservation. Overall, this work demonstrates the potential of learned denoisers like diffusion models for image deconvolution, while highlighting some current limitations to be addressed.

References

[1] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends in Machine Learning*, vol. 3, pp. 1–122, 01 2011.

[2] S. V. Venkatakrishnan, C. A. Bouman, and B. Wohlberg, "Plug-and-play priors for model based reconstruction," in *2013 IEEE Global Conference on Signal and Information Processing*, 2013, pp. 945–948.

[3] Z. Patel, A. Haritas, and K. Kumar, "Deconvolution using admm with diffusion denoising prior," 2022. [Online]. Available: https://www.cs.toronto.edu/~lindell/teaching/2529/past_projects/2022/report/akash_haridas-kartikaeya_kumar-zakaria_patel.pdf

[4] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-resolution image synthesis with latent diffusion models," *CoRR*, vol. abs/2112.10752, 2021. [Online]. Available: https://arxiv.org/abs/2112.10752

[5] P. Dhariwal and A. Nichol, "Diffusion models beat gans on image synthesis," *CoRR*, vol. abs/2105.05233, 2021. [Online]. Available: https://arxiv.org/abs/2105.05233

[6] Y. Zhu, K. Zhang, J. Liang, J. Cao, B. Wen, R. Timofte, and L. V. Gool, "Denoising diffusion models for plug-and-play image restoration," 2023.