

Low-rank Adaptation-based All-Weather Removal for Autonomous Navigation

Sudarshan Rajagopalan and Vishal M. Patel

Abstract—All-weather image restoration (AWIR) is crucial for reliable autonomous navigation under adverse weather conditions. AWIR models are trained to address a specific set of weather conditions such as fog, rain, and snow. But this causes them to often struggle with out-of-distribution (OoD) samples or unseen degradations which limits their effectiveness for real-world autonomous navigation. To overcome this issue, existing models must either be retrained or fine-tuned, both of which are inefficient and impractical, with retraining needing access to large datasets, and fine-tuning involving many parameters. In this paper, we propose using Low-Rank Adaptation (LoRA) to efficiently adapt a pre-trained all-weather model to novel weather restoration tasks. Furthermore, we observe that LoRA lowers the performance of the adapted model on the pre-trained restoration tasks. To address this issue, we introduce a LoRA-based fine-tuning method called LoRA-Align (LoRA-A) which seeks to align the singular vectors of the fine-tuned and pre-trained weight matrices using Singular Value Decomposition (SVD). This alignment helps preserve the model’s knowledge of its original tasks while adapting it to unseen tasks. We show that images restored with LoRA and LoRA-A can be effectively used for computer vision tasks in autonomous navigation, such as semantic segmentation and depth estimation.

I. INTRODUCTION

Image restoration under adverse weather conditions is a widely studied problem which is particularly important for autonomous navigation applications. The advent of deep learning has prompted the development of CNN and transformer-based restoration approaches including, [1], [2], [3], [4], [5], [6], [7], [8], Restormer [9], MPRNet [10] and SwinIR [11]. These methods were designed to handle a single degradation at a time, making them impractical for real-world scenarios as they require storing multiple sets of weights for different degradations. To address these challenges, all-weather or all-in-one models such as All-in-one [12], TransWeather [13], Airnet [14], PromptIR [15] and [16] have been proposed. These models employ various techniques including multiple encoders, contrastive learning, prompt learning and domain translation, to simultaneously tackle multiple degradations. Despite their proficiency in handling multiple degradations, these networks are fundamentally limited to work for a specific set of training degradations, such as fog, rain and snow. This limitation significantly hinders their deployability for autonomous navigation, as it often necessitates retraining or fine-tuning for any new weather task. As previously mentioned, retraining is not a feasible solution due to its time-consuming nature and the need for access to the original

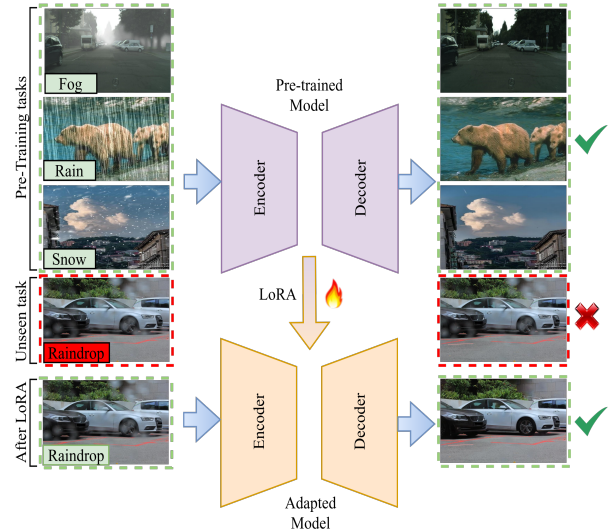


Fig. 1. The AWIR model shown above is pre-trained for fog, rain and snow, for which it works well. However, it fails for the novel or unseen task of raindrop removal. The performance for raindrop removal improves significantly after parameter efficient adaptation using LoRA.

training dataset, which is often large for AWIR tasks. A more practical alternative is to fine-tune the model on a dataset consisting of samples from the new task. However, this is computationally expensive due to the large number of parameters involved.

Recently, Parameter Efficient Fine-Tuning (PEFT) has emerged as an efficient alternative to fine-tuning deep networks. PEFT fine-tunes a pre-trained model for a new task by using only a few learnable parameters. A pioneering PEFT work, LoRA [17], showed that the fine-tuning weight updates are often low-rank, enabling a substantial reduction in the number of learnable parameters. They also showed that PEFT methods such as adapters [18], [19] and prompt tuning [20], [21] under-perform compared to LoRA while having drawbacks such as additional inference latency and difficult optimization. While LoRA has been studied extensively for LLMs [22], [23], [24] and large vision models [25], [26], [27], there has been little to almost no focus on its application to the low-level vision task of AWIR.

In this paper, we propose to employ LoRA to efficiently fine-tune AWIR models for novel restoration tasks. Fig. 1 presents an illustration of its working. Images affected by the novel degradation can then be restored by the adapted model and used for autonomous navigation tasks such as depth estimation and semantic segmentation (see Sec. IV-E). We also conduct a comprehensive analysis to provide valuable

insights on the effective use of LoRA for AWIR tasks. Additionally, we observe that LoRA can lower performance of the adapted model on the original pre-trained tasks which is a disadvantage, especially for autonomous navigation applications. To address this issue, we propose LoRA-Align (LoRA-A) which combines the parameter efficiency of LoRA while preserving the original task performance by seeking alignment of the singular vectors of the pre-trained and fine-tuned weight matrices. This alignment enables the adapted model to be more effective for the original tasks as well, improving its real-world deployability.

In summary our contributions are as follows:

- 1) We propose LoRA as an efficient fine-tuning technique to adapt pre-trained all-weather restoration models for unseen tasks. We provide valuable insights into the incorporation of LoRA for AWIR models.
- 2) We propose a novel SVD-based alignment method for LoRA called LoRA-Align (LoRA-A) to preserve pre-trained task performance while adapting to novel all-weather restoration tasks.
- 3) Additionally, we show that the images restored by LoRA and LoRA-A can be successfully used for downstream tasks such as semantic segmentation and depth estimation, thereby aiding autonomous navigation under adverse weather conditions.

II. RELATED WORKS

We now discuss relevant research on adverse weather removal, parameter efficient fine-tuning and transfer learning. **Adverse weather removal.** Early works such as [28], [29] and [30] incorporated degradation-physics to restore images. Subsequently, deep learning based approaches such as SPANet [4], [31], [6], [32] and [5] for deraining, [1], [2], [3] and [33] for dehazing, and [34], [35] and [7] for desnowing were proposed. More recent methods such as [10], [11] and [9] have been proposed to handle multiple degradations. These methods require storing multiple-sets of weights for each degradation, making them impractical. To address this limitation, AWIR methods have been explored.

All-in-one [12] used neural architecture search with multiple encoders while TransWeather [13] employed a unified network with a single encoder for all-weather restoration. Airnet [14] and [36] used contrastive regularization to learn enhanced degradation representations. [37] fused generic and specific weather features for restoration. [38] utilized a patch-based denoising diffusion model while [15] introduced learnable prompt embeddings for AWIR. Despite these advancements, the deployability of AWIR methods for autonomous navigation is limited by the pre-defined set of training degradations. Adapting to new degradations involves retraining or fine-tuning, both of which are practically inefficient due to the need for learning a large number of parameters.

Parameter efficient fine-tuning and transfer learning. The enormous size of LLMs makes full-finetuning for downstream tasks challenging. To address this, parameter efficient fine-tuning (PEFT) and transfer learning methods have been proposed for LLMs. Prompt engineering [20] involves

designing task-specific instructions or prompts to adapt a pre-trained model for a new task. Its variants include few-shot prompting [39] and chain of thought prompting [21]. Visual prompt tuning (VPT) [40] introduced learnable input tokens to achieve a similar functionality for computer vision tasks. SAM [41] utilized prompting for various downstream segmentation tasks. Another prominent approach is the use of adapters [18], [19], which are learnable layers added within transformer blocks for novel task adaptation. T2I-adapter [42] leveraged adapters for controllable text to image generation. However, adapters introduce additional latency during inference [17]. A significant advancement in PEFT is LoRA [17], which proposed a low-rank decomposition of the weight updates of transformer layers. Since its introduction, LoRA has inspired numerous works such as LoRA+ [22], VeRA [43], AdaLoRA [23], DoRA [44]. Despite the growing popularity of LoRA, there has been little focus on its application to the low-level vision task of AWIR. In this work, we aim to demonstrate the potential of LoRA for efficiently adapting pre-trained AWIR models to new tasks, thereby aiding autonomous navigation.

III. PROPOSED METHODOLOGY

In this section, we present our approach for LoRA-based adaptation of pre-trained AWIR models.

A. Low-rank adaptation

We briefly discuss the working of LoRA [17] for parameter efficient fine-tuning on AWIR models. Fig. 2(a) provides an illustration of how LoRA operates. Consider a dense layer in a pre-trained all-weather model with weights represented by $W \in \mathbb{R}^{m \times n}$. During fine-tuning, the weights are updated to $W + \Delta W$ where $\Delta W \in \mathbb{R}^{m \times n}$ and W is kept frozen. LoRA constrains the learnable parameters in ΔW to a low-rank decomposition $\Delta W = AB$, where $A \in \mathbb{R}^{m \times r}$ and $B \in \mathbb{R}^{r \times n}$. Here, r represents the intrinsic rank of ΔW and is typically quite small, as demonstrated by [17]. With LoRA, the output x_{out} of a dense layer for an input x_{in} is given by

$$x_{\text{out}} = W \cdot x_{\text{in}} + \Delta W \cdot x_{\text{in}}. \quad (1)$$

Our experiments (see Sec. IV) reveal that low-rank adaptation of pre-trained AWIR models achieves competitive performance on novel restoration tasks, comparable to that of full fine-tuning, while having to learn only a fraction of the parameters. The adapted model can then be used for autonomous navigation tasks, as discussed in Sec. IV-E. However, we also observed a fall in performance of the adapted model for the original tasks. To address this issue, we propose an SVD-based alignment method called LoRA-Align (LoRA-A), which can preserve the model's performance on pre-trained tasks while efficiently adapting to new tasks, further enhancing its applicability for autonomous navigation.

B. LoRA-Align

We now introduce LoRA-Align (illustrated in Fig. 2(b)), designed to preserve original task performance while adapting a pre-trained AWIR model to novel tasks using LoRA.

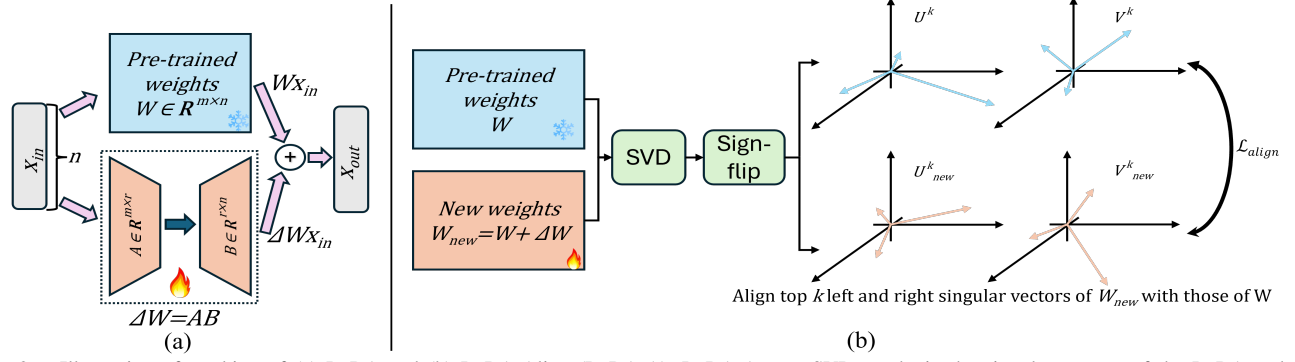


Fig. 2. Illustration of working of (a) LoRA and (b) LoRA-Align (LoRA-A). LoRA-A uses SVD to obtain the singular vectors of the LoRA-updated weights W_{new} and the pre-trained weights W . Subsequently, we resolve the sign ambiguity in SVD and obtain the top k sign-corrected left and right singular vectors (U_{new}^k and V_{new}^k) of W_{new} and those (U^k and V^k) of W . The alignment loss $\mathcal{L}_{\text{align}}$ is calculated between U_{new}^k , U^k and V_{new}^k , V^k in order to align the singular vectors. This results in LoRA-A retaining performance on the pre-trained tasks. *Graphs are drawn for representation purposes only.*

We accomplish this with the help of SVD. For a dense layer, the SVD of its pre-trained weight matrix $W \in \mathbb{R}^{m \times n}$ is given by $W = U\Sigma V^T$, where $U \in \mathbb{R}^{m \times m}$, $\Sigma \in \mathbb{R}^{m \times n}$, and $V \in \mathbb{R}^{n \times n}$. The columns of U contain the left-singular vectors of W and the columns of V contain its right-singular vectors. Σ is a diagonal matrix consisting of the singular values of W arranged in descending order. After one iteration of LoRA, the new weight matrix of the dense layer becomes $W_{\text{new}} = W + \Delta W$. We decompose the updated weight matrix W_{new} using SVD to obtain $W_{\text{new}} = U_{\text{new}}\Sigma_{\text{new}}V_{\text{new}}^T$ where the dimensions of each of the resulting matrices remain the same as that of W . To preserve performance on previous tasks, we align the left and right singular vectors of W with those of W_{new} , respectively (see Fig. 2), after each iteration of LoRA. This is achieved by selecting the top k singular vectors from U, U_{new} and V, V_{new} , and computing an alignment score between them. However, an important issue arises due to the inherent sign ambiguity in the singular vectors obtained from SVD. This problem is further compounded as we aim to resolve the sign ambiguity across two different weight matrices W and W_{new} , where W_{new} changes iteratively. To resolve this issue, we propose the following steps:

- 1) Obtain top k singular vectors from U, V, U_{new} and V_{new} as $U^k, V^k, U_{\text{new}}^k$ and V_{new}^k , respectively.
- 2) Compute the ratio $R^i = \frac{\|\mathbf{u}^i - \mathbf{u}_{\text{new}}^i\|_2}{\|\mathbf{u}^i + \mathbf{u}_{\text{new}}^i\|_2}$ for every left singular vector, $\mathbf{u}^i, i = 1, 2, \dots, k$. The rationale behind the definition of R^i is as follows. If the singular vectors \mathbf{u}^i and $\mathbf{u}_{\text{new}}^i$ are closely related but sign-flipped versions of each other, then $\|\mathbf{u}^i + \mathbf{u}_{\text{new}}^i\|_2$ will be very small but $\|\mathbf{u}^i - \mathbf{u}_{\text{new}}^i\|_2$ will be a much larger value. Thus, the ratio R^i will become very large in such cases. Conversely, if \mathbf{u}^i and $\mathbf{u}_{\text{new}}^i$ are closely related but not sign flipped, then the ratio R^i will be quite low.
- 3) If $R^i > T$ (a suitably chosen value), we flip the sign of $\mathbf{u}_{\text{new}}^i$ and its corresponding right singular vector $\mathbf{v}_{\text{new}}^i$.

Since our sign-flipping method would not work correctly if W_{new} is significantly different from W , we perform the alignment at every LoRA iteration, starting from the first iteration. Subsequently, we calculate the dot product with the sign-corrected vectors and obtain S_{align}^U and S_{align}^V as the

alignment scores for left and right singular vectors of W and W_{new} , respectively.

$$S_{\text{align}}^U = U_{\text{new}}^{kT} \cdot U^k, S_{\text{align}}^U \in \mathbb{R}^{k \times k} \quad (2a)$$

$$S_{\text{align}}^V = V_{\text{new}}^{kT} \cdot V^k, S_{\text{align}}^V \in \mathbb{R}^{k \times k}. \quad (2b)$$

These scores are then used to compute an alignment loss, $\mathcal{L}_{\text{align}}$, for aligning the singular vectors of W and W_{new} .

$$\mathcal{L}_{\text{align}} = 0.5 \cdot (\text{Mean}(\text{Diag}(I_{k \times k} - S_{\text{align}}^U)^2) + \text{Mean}(\text{Diag}(I_{k \times k} - S_{\text{align}}^V)^2)), \quad (3)$$

where $\text{Mean}(\cdot)$ represents averaging and $\text{Diag}(\cdot)$ selects the diagonal elements from a matrix. $\mathcal{L}_{\text{align}}$ is computed for each LoRA layer in the model and averaged to obtain the mean alignment loss $\mathcal{L}_{\text{align}}^{\text{avg}}$. The final training loss $\mathcal{L}_{\text{train}}$ is

$$\mathcal{L}_{\text{train}} = \mathcal{L}_1(y, \hat{y}) + w_{\text{align}} \cdot \mathcal{L}_{\text{align}}^{\text{avg}}, \quad (4)$$

where y is the ground truth, \hat{y} is the restored image, $\mathcal{L}_1(\cdot)$ is the L1 loss and w_{align} is a weighting factor for the alignment loss. Directly computing the alignment loss between W and ΔW results in unstable gradients due to ΔW not being full-rank. This is a documented issue with the SVD function in PyTorch [48].

IV. EXPERIMENTS

We now demonstrate the effectiveness of LoRA for efficient fine-tuning, and show that LoRA-Align (LoRA-A) helps preserve original task performance during adaptation. We evaluate both methods on two restoration networks: PromptIR [15] and TransWeather [13]. Additionally, we conduct comprehensive analysis and ablation studies to provide valuable insights into the application of LoRA and LoRA-A for AWIR tasks. Finally, we present results on autonomous navigation tasks such as semantic segmentation and depth estimation using images restored by the adapted models.

A. Implementation details

For all-weather pre-training, we followed the specifications mentioned in the papers of the restoration networks. For fine-tuning, we used a learning rate of 5×10^{-5} . For LoRA and LoRA-A we used a learning rate of 5×10^{-4} . We employed the AdamW optimizer with $\beta_1 = 0.9, \beta_2 = 0.99$, an exponential learning rate scheduler with $\gamma = 0.95$ and a batch size of 32 for all experiments. Our ablations (see

TABLE I
QUANTITATIVE COMPARISONS OF LoRA AND LoRA-A WITH THE PRE-TRAINED AND FINE-TUNED VERSIONS OF PROMPTIR [15] AND TRANSWEATHER [13] FOR THE NOVEL TASK OF RAINDROP REMOVAL.

Method	Trainable parameters (M)	Pre-trained Tasks				Novel Task
		Fog [45]	Rain100L [46]	Rain100H [46]	Snow100k [34]	Raindrop [47]
PromptIR [15] pre-trained	35.4	28.58/0.972	34.49/0.956	27.25/0.834	33.77/0.936	23.79/0.838
PromptIR Fine-tuned	35.4	26.66/0.945	25.85/0.822	24.71/0.786	28.52/0.892	29.80/0.904
PromptIR LoRA	0.55	24.62/0.931	25.51/0.828	24.52/0.781	27.70/0.882	29.63/0.900
PromptIR LoRA-A	0.55	26.74/0.949	26.88/0.848	24.98/0.790	29.30/0.905	29.35/0.897
TransWeather [13] pre-trained	38.31	28.19/0.958	31.58/0.935	26.03/0.805	31.69/0.912	24.01/0.841
TransWeather Fine-tuned	38.31	18.18/0.864	26.72/0.860	24.10/0.754	26.71/0.855	28.89/0.888
TransWeather LoRA	0.26	20.29/0.890	27.06/0.870	24.04/0.767	27.44/0.867	28.30/0.883
TransWeather LoRA-A	0.26	23.47/0.930	27.70/0.890	25.07/0.793	28.00/0.883	27.84/0.878

TABLE II
QUANTITATIVE COMPARISONS OF FINE-TUNING AND LoRA FOR VARIOUS PRE-TRAINING AND UNSEEN TASKS. THE RESULTS FOR THE NOVEL TASK OF RAIN (ROW 3) ARE ON THE RAIN100H [46] DATASET.

Pre-trained tasks	Novel task	Pre-trained	Fine-tuned	LoRA
Rain, Snow, Raindrop	Fog	15.59/0.827	23.72/0.944	25.41/0.947
Rain, Raindrop, Fog	Snow	23.17/0.771	29.85/0.892	29.60/0.889
Snow, Fog, Raindrop	Rain	12.13/0.344	25.03/0.760	24.78/0.750

Sec. IV-D) reveal that LoRA works well while adapting both Attention and MLP layers of the transformer blocks with a rank of 4. For LoRA-A, we used $k = 16$, $w_{\text{align}} = 100$ and $T = 7$ along with the above specifications for LoRA. We used the above parameters for Secs. IV-C and IV-E.

B. Datasets

We considered four tasks for all our experiments: defogging, deraining, desnowing and raindrop removal. The datasets used were as follows:

- 1) Defogging: City Fog dataset [45] with 8925 paired images for training and 4575 paired images for testing.
- 2) Deraining: SRD dataset [10] with 13,711 paired images for training. For testing, Rain100H and Rain100L datasets [46] with 100 paired images each.
- 3) Desnowing: Snow100k dataset [34] with 50,000 paired images for both training and testing.
- 4) Raindrop removal: Raindrop [47] dataset consisting of 861 images for training and 58 images for testing.

The specific pre-training and adaptation tasks used in different experiments will be described as needed. For adaptation to a new task, we use 1000 randomly selected images from its training set (on average) rather than the entire dataset to enable fast adaptation.

C. Results and analysis

Performance of LoRA on unseen tasks. In this experiment, we pre-train PromptIR and TransWeather for defogging, deraining, and desnowing. Then we adapt the models to the novel task of raindrop removal using naive fine-tuning, LoRA

and LoRA-A. Table I shows that while the pre-trained models perform well for their original tasks, their performance on the novel task is poor. After fine-tuning and LoRA, the performance of both PromptIR and TransWeather on the novel task improves significantly. Notice that LoRA is on par with the full fine-tuning performance for the novel task of raindrop removal while it needs to learn only a fraction of the parameters ($\sim 1.55\%$ for PromptIR and $\sim 0.68\%$ for TransWeather). The qualitative results for this experiment are shown in Fig. 3.

Dependence of LoRA on pre-trained tasks. Our approach of leveraging LoRA for adapting AWIR models to novel restoration tasks does not rely on specific pre-training tasks. To demonstrate this, we pre-train PromptIR on sets of 3 different tasks chosen from defogging, deraining, desnowing and raindrop removal and use the 4th unseen task for adaptation. The results of this experiment are presented in Table II. Once again, we observe that LoRA achieves competitive performance to full fine-tuning while using only a fraction of the parameters. For the novel task of defogging (see row 1 of Table II), LoRA even outperforms fine-tuning, thus, demonstrating its potential for adapting AWIR models.

Comparison of LoRA with re-training. For this experiment, we first pre-train PromptIR [15] for defogging, deraining and desnowing, and adapt it for raindrop removal using LoRA. Next, we re-train a PromptIR model for all four tasks, i.e. defogging, deraining, desnowing and raindrop removal. We then compare the performance of both models for the task of raindrop removal. Table III shows that LoRA with a rank of 4 achieves nearly the same performance as re-training while using only a fraction of the learnable parameters and trained solely for the novel task of raindrop removal. This experiment demonstrates that LoRA-based adaptation can achieve performance close to that of full retraining.

Performance of LoRA-Align. For analysing the performance of LoRA-A, we use the same PromptIR and TransWeather training configurations as for LoRA. LoRA-A has the same parameter efficiency as LoRA while maintaining better performance on the pre-trained tasks (see Table I). Moreover, the table shows that LoRA and LoRA-A achieve near identical performance for the novel task. In the case of PromptIR, for a trade-off of just 0.28 dB with respect

TABLE III

QUANTITATIVE COMPARISON OF LoRA AND RE-TRAINING FOR THE NOVEL TASK OF RAINDROP REMOVAL.

Method	Rank	#Parameters (M)	PSNR/SSIM
Re-trained	-	35.4M	29.93/0.905
LoRA	4	0.55	29.63/0.900
LoRA	32	4.4	29.93/0.907
LoRA	64	8.9	29.91/0.907

TABLE IV

ABLATIONS FOR LoRA ON PROMPTIR TO DETERMINE THE BEST TRANSFORMER LAYERS TO TUNE FOR RAINDROP REMOVAL.

Encoder	Decoder	Attention	MLP	PSNR/SSIM
✓	-	✓	✓	28.92/0.885
-	✓	✓	✓	29.28/0.897
✓	✓	✓	-	29.14/0.896
✓	✓	✓	✓	29.63/0.900

to LoRA on the novel raindrop removal task, LoRA-A remarkably recovers over 1.38 dB of performance (on the average) on the pre-trained tasks. For TransWeather, LoRA-A preserves 1.35 dB more performance (on the average) on the pre-trained tasks than LoRA for the raindrop removal task. Fig. 3 shows qualitative results for one of the pre-training tasks (defogging) and the novel task. Observe that LoRA-A delivers identical performance to LoRA on the novel task while maintaining pre-training task performance. Incidentally, when LoRA is trained with rank 2 so as to reduce its fine-tuning for the novel task (see Table V), it yields very similar performance as LoRA-A on the novel task. However, LoRA achieves a PSNR of only 25.70 dB on the pre-trained task of Rain100L compared to 26.88 dB of LoRA-A.

D. Ablations

We conduct ablations on LoRA and LoRA-Align to find the optimal restoration-specific settings for both methods. For all ablations, we use PromptIR pre-trained for defogging, deraining and desnowing, and adapt it to raindrop removal.

First, we identify the best rank and transformer block layers to apply LoRA. Table IV contains the results obtained by using LoRA on different transformer block layers of the PromptIR model. Specifically, we compare adapting the attention layers, MLP layers, encoder and decoder of PromptIR with a rank of 4. The results from Table IV indicate that adapting both the attention and MLP layers in the encoder and decoder yields best performance for the novel task of raindrop removal. Table V presents the novel task performance for different ranks and the number of learnable parameters involved for each rank. We also provide the performance of full fine-tuning for comparison. It is evident that LoRA achieves competitive performance with full fine-tuning across almost all ranks. In fact, ranks of 32 and 64 surpass the performance of full fine-tuning while using considerably fewer parameters. We used a rank of 4 for all our experiments as it provides the best trade-off between performance and parameter efficiency.

TABLE V

EFFECT OF RANK OF LoRA ON THE NUMBER OF LEARNABLE PARAMETERS (IN MILLIONS (M)) IN PROMPTIR AND ITS PERFORMANCE ON THE NOVEL TASK OF RAINDROP REMOVAL.

Method	Rank	#Parameters (M)	PSNR/SSIM
Fine-tuning	-	35.4	29.80/0.904
LoRA	2	0.277	29.39/0.899
LoRA	4	0.554	29.63/0.900
LoRA	8	1.10	29.67/0.903
LoRA	16	2.2	29.71/0.905
LoRA	32	4.4	29.93/0.907
LoRA	64	8.9	29.91/0.907

TABLE VI

EFFECT OF VARYING k IN THE LoRA-A FRAMEWORK FOR THE PRE-TRAINED TASK OF DERAINING ON RAIN100L [46] AND THE NOVEL TASK OF RAINDROP REMOVAL ON RAINDROP [47].

Method	k	Rain100L [46]	Raindrop [47]
LoRA	-	25.51/0.828	29.63/0.900
LoRA-Align	2	25.96/0.828	29.59/0.901
LoRA-Align	4	26.15/0.832	29.47/0.899
LoRA-Align	8	26.15/0.832	29.52/0.900
LoRA-Align	16	26.88/0.848	29.35/0.897
LoRA-Align	32	27.05/0.854	29.14/0.894
LoRA-Align	64	27.91/0.869	28.14/0.880

Next, we perform ablations on the value of k in LoRA-A where k specifies the number of singular vectors to align for retaining pre-training task performance. We vary k to examine its effect on both the pre-training task and novel task performance. During this experiment, LoRA specifications are used as determined from the previous ablations. Table VI shows that increasing the value of k leads to better performance on the pre-training task of deraining on Rain100L. However, the novel task (raindrop removal) performance gradually decreases and the drop becomes quite significant beyond $k = 32$. This behaviour is to be expected as we are constraining more singular vectors of the adapted weight matrix (W_{new}) to align with those of the pre-trained weight matrix W . We selected value of $k = 16$ as it produces the best trade-off between performance on novel task and pre-trained tasks.

E. Autonomous navigation applications

In this section, we utilize the images restored by PromptIR adapted using LoRA and LoRA-A for two important autonomous navigation tasks: semantic segmentation and depth estimation. For both tasks, we adapt PromptIR pre-trained on defogging, deraining and desnowing for the task of raindrop removal. We use the optimal parameters obtained in Sec. IV-D for LoRA and LoRA-A. Results are presented for the novel task and one of the pre-training tasks (defogging).

Semantic segmentation. We use Mask2Former [49] to compare the predictions obtained on the degraded images and restored images produced by the adapted models (using LoRA and LoRA-A). The quantitative and qualitative results of this experiment are presented in Table VII and Fig. 4, respectively. We use the mean Intersection over Union (mIoU) to evaluate segmentation performance. For the

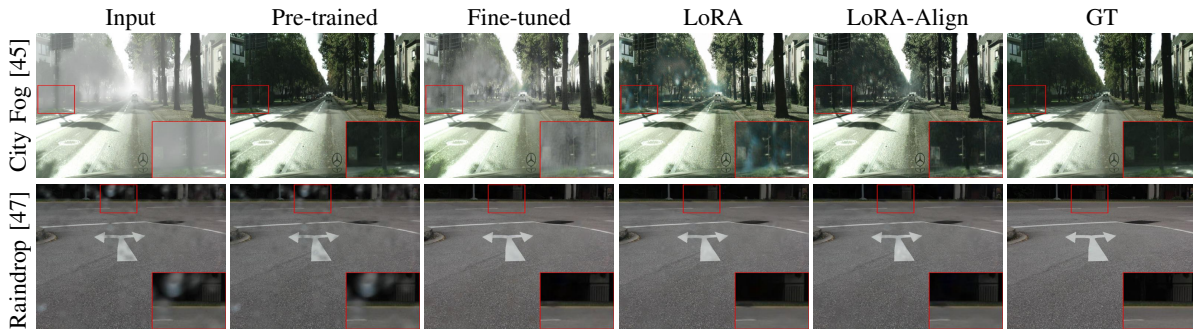


Fig. 3. Qualitative comparisons of fine-tuning, LoRA and LoRA-A for the novel task of raindrop removal using PromptIR pre-trained for defogging, deraining and desnowing. Results for one of the pre-training tasks (defogging) is also shown.

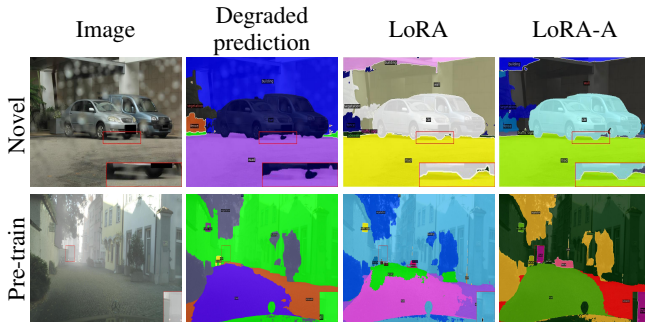


Fig. 4. Predictions of Mask2Former [49] on novel task of raindrop removal and pre-training task of defogging. After restoration using LoRA and LoRA-A, the results improve significantly. Furthermore, LoRA-A yields better segmentation results for the pre-training task of defogging.

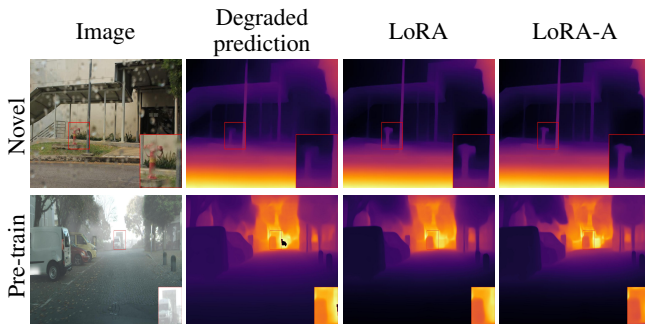


Fig. 5. Predictions of Depth-Anything [50] on the novel task of raindrop removal and pre-training task of defogging. The predictions improve significantly after restoration using LoRA and LoRA-A. LoRA-A yields better depth results for the pre-training task of defogging.

novel task of raindrop removal, performance significantly improves after restoration using LoRA and LoRA-A. Note that the predictions of Mask2Former on the clean images were used as ground truth for raindrop removal as Raindrop [47] does not contain segmentation annotations. For the pre-training task of defogging, LoRA-A preserves restoration performance thereby leading to better predictions of Mask2Former. Thus, the adapted models can greatly aid in improving the segmentation performance.

Depth estimation. For the task of depth estimation, we use a foundation model called Depth Anything [50]. Table VIII and Fig. 5, respectively provide the quantitative and qualitative comparisons of Depth Anything on the degraded image and

TABLE VII
MIOU SCORES OF MASK2FORMER ON IMAGES RESTORED BY LoRA AND LoRA-A ADAPTED MODELS FOR THE PRE-TRAINING TASK OF DEFOGGING AND NOVEL TASK OF RAINDROP REMOVAL.

Task	Degradation	Degraded	LoRA	LoRA-A
Novel	Raindrop	39.19	43.82	42.25
Pre-train	Fog	57.00	58.16	59.14

TABLE VIII
QUANTITATIVE EVALUATION OF DEPTH ANYTHING [50] ON IMAGES RESTORED BY LoRA AND LoRA-A FOR DEFOGGING (PRE-TRAINING TASK) AND RAINDROP REMOVAL (NOVEL TASK).

Task, Degradation	$\delta_1 \uparrow$	$\delta_2 \uparrow$	$\delta_3 \uparrow$	SILOG \downarrow
Novel, Raindrop (Degraded)	0.950	0.986	0.995	0.137
Novel, Raindrop (LoRA)	0.977	0.995	0.998	0.106
Novel, Raindrop (LoRA-A)	0.974	0.995	0.998	0.112
Pre-train, Fog (Degraded)	0.978	0.989	0.991	0.199
Pre-train, Fog (LoRA)	0.980	0.991	0.992	0.190
Pre-train, Fog (LoRA-A)	0.982	0.991	0.993	0.186

the images restored using LoRA and LoRA-A. For quantitative evaluation we use $\delta_1, \delta_2, \delta_3$ thresholds and SILOG metrics (see [51]). To calculate metrics for both raindrop removal and defogging, the predictions of Depth Anything on the clean images were used as ground truth. There is a significant improvement in depth estimation on the novel task after restoration using LoRA and LoRA-A. Furthermore, LoRA-A performs better on the pre-training tasks leading to superior depth predictions for defogging.

V. CONCLUSIONS

We proposed leveraging LoRA to efficiently adapt pre-trained AWIR models to novel restoration tasks, achieving near-identical performance to full fine-tuning but at a fraction of the learnable parameters. Additionally, we introduced LoRA-Align, which significantly recovers pre-training task performance with a minor trade-off on novel tasks while maintaining LoRA’s parameter efficiency. Finally, we demonstrated that AWIR models adapted using LoRA and LoRA-A can enhance the performance of autonomous navigation tasks such as semantic segmentation and depth estimation for adverse weather conditions. We believe that our work will aid the real-world deployability of AWIR models.

VI. ACKNOWLEDGMENTS

This work is supported by the Intelligence Advanced Research Projects Activity (IARPA) via Department of Interior/ Interior Business Center (DOI/IBC) contract number 140D0423C0076. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. Disclaimer: The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DOI/IBC, or the U.S. Government.

REFERENCES

- [1] H. Zhang, V. Sindagi, and V. M. Patel, "Joint transmission map estimation and dehazing using deep networks," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 7, pp. 1975–1986, 2020.
- [2] H. Zhang and V. M. Patel, "Densely connected pyramid dehazing network," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 3194–3203.
- [3] H. Wu, Y. Qu, S. Lin, J. Zhou, R. Qiao, Z. Zhang, Y. Xie, and L. Ma, "Contrastive learning for compact single image dehazing," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021, pp. 10546–10555.
- [4] T. Wang, X. Yang, K. Xu, S. Chen, Q. Zhang, and R. W. Lau, "Spatial attentive single-image deraining with a high quality real rain dataset," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 12262–12271.
- [5] W. Wei, D. Meng, Q. Zhao, Z. Xu, and Y. Wu, "Semi-supervised transfer learning for image rain removal," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 3872–3881.
- [6] R. Quan, X. Yu, Y. Liang, and Y. Yang, "Removing raindrops and rain streaks in one go," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021, pp. 9143–9152.
- [7] K. Zhang, R. Li, Y. Yu, W. Luo, and C. Li, "Deep dense multi-scale network for snow removal using semantic and depth priors," *IEEE Transactions on Image Processing*, vol. 30, pp. 7419–7431, 2021.
- [8] W.-T. Chen, H.-Y. Fang, J.-J. Ding, C.-C. Tsai, and S.-Y. Kuo, "Jstasr: Joint size and transparency-aware snow removal algorithm based on modified partial convolution and veiling effect removal." Berlin, Heidelberg: Springer-Verlag, 2020. [Online]. Available: https://doi.org/10.1007/978-3-030-58589-1_45
- [9] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, and M.-H. Yang, "Restormer: Efficient transformer for high-resolution image restoration," in *CVPR*, 2022.
- [10] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, M.-H. Yang, and L. Shao, "Multi-stage progressive image restoration," in *CVPR*, 2021.
- [11] J. Liang, J. Cao, G. Sun, K. Zhang, L. Van Gool, and R. Timofte, "Swinir: Image restoration using swin transformer," in *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, 2021, pp. 1833–1844.
- [12] R. Li, R. T. Tan, and L.-F. Cheong, "All in one bad weather removal using architectural search," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 3172–3182.
- [13] J. J. Valanarasu, R. Yasarla, and V. M. Patel, "Transweather: Transformer-based restoration of images degraded by adverse weather conditions," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 2343–2353.
- [14] B. Li, X. Liu, P. Hu, Z. Wu, J. Lv, and X. Peng, "All-in-one image restoration for unknown corruption," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022, pp. 17431–17441.
- [15] V. Potlapalli, S. W. Zamir, S. Khan, and F. Khan, "Promptir: Prompting for all-in-one image restoration," in *Thirty-seventh Conference on Neural Information Processing Systems*, 2023.
- [16] P. W. Patil, S. Gupta, S. Rana, S. Venkatesh, and S. Murala, "Multi-weather image restoration via domain translation," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, October 2023, pp. 21 696–21 705.
- [17] E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, and W. Chen, "Lora: Low-rank adaptation of large language models," *arXiv preprint arXiv:2106.09685*, 2021.
- [18] N. Houlsby, A. Giurgiu, S. Jastrzebski, B. Morrone, Q. De Laroussilhe, A. Gesmundo, M. Attariyan, and S. Gelly, "Parameter-efficient transfer learning for nlp," in *International conference on machine learning*. PMLR, 2019, pp. 2790–2799.
- [19] J. Pfeiffer, A. Kamath, A. Rücklé, K. Cho, and I. Gurevych, "Adapterfusion: Non-destructive task composition for transfer learning," *arXiv preprint arXiv:2005.00247*, 2020.
- [20] P. Liu, W. Yuan, J. Fu, Z. Jiang, H. Hayashi, and G. Neubig, "Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing," *ACM Comput. Surv.*, vol. 55, no. 9, jan 2023. [Online]. Available: <https://doi.org/10.1145/3560815>
- [21] J. Wei, X. Wang, D. Schuurmans, M. Bosma, F. Xia, E. Chi, Q. V. Le, D. Zhou, *et al.*, "Chain-of-thought prompting elicits reasoning in large language models," *Advances in neural information processing systems*, vol. 35, pp. 24 824–24 837, 2022.
- [22] S. Hayou, N. Ghosh, and B. Yu, "Lora+: Efficient low rank adaptation of large models," *arXiv preprint arXiv:2402.12354*, 2024.
- [23] Q. Zhang, M. Chen, A. Bukharin, N. Karampatziakis, P. He, Y. Cheng, W. Chen, and T. Zhao, "Adalora: Adaptive budget allocation for parameter-efficient fine-tuning," *arXiv preprint arXiv:2303.10512*, 2023.
- [24] Y. Mao, Y. Ge, Y. Fan, W. Xu, Y. Mi, Z. Hu, and Y. Gao, "A survey on lora of large language models," 2024. [Online]. Available: <https://arxiv.org/abs/2407.11046>
- [25] S. Doveh, A. Arbelle, S. Harary, E. Schwartz, R. Herzig, R. Giryas, R. Feris, R. Panda, S. Ullman, and L. Karlinsky, "Teaching structured vision & language concepts to vision & language models," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 2657–2668.
- [26] M. Zanella and I. Ben Ayed, "Low-rank few-shot adaptation of vision-language models," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 1593–1603.
- [27] S. Khanna, M. Irgau, D. B. Lobell, and S. Ermon, "Explora: Parameter-efficient extended pre-training to adapt vision transformers under domain shifts," 2024. [Online]. Available: <https://arxiv.org/abs/2406.10973>
- [28] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 1956–1963.
- [29] S. Roth and M. Black, "Fields of experts: a framework for learning image priors," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 2, 2005, pp. 860–867 vol. 2.
- [30] L.-W. Kang, C.-W. Lin, and Y.-H. Fu, "Automatic single-image-based rain streaks removal via image decomposition," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 1742–1755, 2012.
- [31] X. Li, J. Wu, Z. Lin, H. Liu, and H. Zha, "Recurrent squeeze-and-excitation context aggregation net for single image deraining," in *European Conference on Computer Vision*. Springer, 2018, pp. 262–277.
- [32] H. Porav, T. Bruls, and P. Newman, "I can see clearly now: Image restoration via de-raining," in *2019 International Conference on Robotics and Automation (ICRA)*, 2019, pp. 7087–7093.
- [33] Y. Qu, Y. Chen, J. Huang, and Y. Xie, "Enhanced pix2pix dehazing network," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 8152–8160.
- [34] Y.-F. Liu, D.-W. Jaw, S.-C. Huang, and J.-N. Hwang, "Desnownet: Context-aware deep network for snow removal," *IEEE Transactions on Image Processing*, vol. 27, no. 6, pp. 3064–3073, 2018.
- [35] W.-T. Chen, H.-Y. Fang, C.-L. Hsieh, C.-C. Tsai, I.-H. Chen, J.-J. Ding, and S.-Y. Kuo, "All snow removed: Single image desnowing algorithm using hierarchical dual-tree complex wavelet representation and contradict channel loss," in *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2021, pp. 4176–4185.
- [36] W.-T. Chen, Z.-K. Huang, C.-C. Tsai, H.-H. Yang, J.-J. Ding, and S.-Y. Kuo, "Learning multiple adverse weather removal via two-stage knowledge learning and multi-contrastive regularization: Toward a unified model," 2022.

- [37] Y. Zhu, T. Wang, X. Fu, X. Yang, X. Guo, J. Dai, Y. Qiao, and X. Hu, "Learning weather-general and weather-specific features for image restoration under multiple adverse weather conditions," in *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023, pp. 21 747–21 758.
- [38] O. Özdenizci and R. Legenstein, "Restoring vision in adverse weather conditions with patch-based denoising diffusion models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1–12, 2023.
- [39] H. Touvron, T. Lavril, G. Izacard, X. Martinet, M.-A. Lachaux, T. Lacroix, B. Rozière, N. Goyal, E. Hambro, F. Azhar, *et al.*, "Llama: Open and efficient foundation language models," *arXiv preprint arXiv:2302.13971*, 2023.
- [40] M. Jia, L. Tang, B.-C. Chen, C. Cardie, S. Belongie, B. Hariharan, and S.-N. Lim, "Visual prompt tuning," in *European Conference on Computer Vision*. Springer, 2022, pp. 709–727.
- [41] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo, *et al.*, "Segment anything," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 4015–4026.
- [42] C. Mou, X. Wang, L. Xie, Y. Wu, J. Zhang, Z. Qi, and Y. Shan, "T2i-adapter: Learning adapters to dig out more controllable ability for text-to-image diffusion models," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 38, no. 5, 2024, pp. 4296–4304.
- [43] D. J. Kopiczko, T. Blankevoort, and Y. M. Asano, "Vera: Vector-based random matrix adaptation," *arXiv preprint arXiv:2310.11454*, 2023.
- [44] S.-Y. Liu, C.-Y. Wang, H. Yin, P. Molchanov, Y.-C. F. Wang, K.-T. Cheng, and M.-H. Chen, "Dora: Weight-decomposed low-rank adaptation," *arXiv preprint arXiv:2402.09353*, 2024.
- [45] M. Hahner, D. Dai, C. Sakaridis, J.-N. Zaech, and L. V. Gool, "Semantic understanding of foggy scenes with purely synthetic data," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, 2019, pp. 3675–3681.
- [46] W. Yang, R. T. Tan, J. Feng, J. Liu, Z. Guo, and S. Yan, "Deep joint rain detection and removal from a single image," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 1685–1694.
- [47] R. Qian, R. T. Tan, W. Yang, J. Su, and J. Liu, "Attentive generative adversarial network for raindrop removal from a single image," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.
- [48] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in pytorch," 2017.
- [49] B. Cheng, I. Misra, A. G. Schwing, A. Kirillov, and R. Girdhar, "Masked-attention mask transformer for universal image segmentation," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 1290–1299.
- [50] L. Yang, B. Kang, Z. Huang, X. Xu, J. Feng, and H. Zhao, "Depth anything: Unleashing the power of large-scale unlabeled data," in *CVPR*, 2024.
- [51] D. Eigen, C. Puhrsch, and R. Fergus, "Depth map prediction from a single image using a multi-scale deep network," in *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2*, ser. NIPS'14. Cambridge, MA, USA: MIT Press, 2014, p. 2366–2374.