Adaptive-basis decomposition-based low-rank network for efficient non-uniform motion deblurring*

CHEN Lei[1], XIONG Qingbo[1], ZHANG Wei[2], and LI Runde[1]**

1. School of Software, Henan University, Kaifeng 475004, China

2. School of Applied Technology, China University of Labor Relations, Beijing 100048, China

(Received 4 December 2023; Revised 18 July 2024)

In this study, we present a unified sparsity-driven framework that significantly enhances motion deblurring performance by integrating two key components: a custom-designed dataset and a low-rank module (LRM). This framework leverages the inherent sparsity of per-pixel blur kernels to bolster both deblurring accuracy and model interpretability. Firstly, we propose an adaptive-basis decomposition-based deblurring (ADD) approach, which constructs a tailored training dataset to enhance the generalization capacity of the deblurring network. The ADD framework adaptively decomposes motion blur into sparse basis elements, effectively addressing the intricacies associated with non-uniform blurs. Secondly, an LRM is proposed to improve the interpretability of deblurring models as a plug-and-play module, primarily designed to identify and harness the intrinsic sparse features in sharp images. A series of ablation studies have been conducted to substantiate the synergistic advantages of combining the proposed ADD with the LRM for overall improvement in deblurring efficacy. Subsequently, we empirically demonstrate through rigorous experimentation that incorporating the LRM into an existing Uformer network leads to substantial enhancement in reconstruction performance. This integration yields a sparsity-guided low-rank network (SGLRN). Operating under the overarching principle of sparsity, SGLRN consistently outperforms state-of-the-art methods across multiple standard deblurring benchmarks. Comprehensive experimental results, assessed through quantitative metrics and qualitative visual evaluations, provide compelling evidence of its effectiveness. The overall deblurring results are available at Google Drive.

Document code: A **Article ID:** 1673-1905(2025)01-0043-8

DOI https://doi.org/10.1007/s11801-025-3275-x

1. Introduction

Motion deblurring has been extensively investigated in recent decades, especially with the popularity of portable imaging devices such as smartphones. Thus, videos/images have become the most important way to obtain information and communicate with each other. Although a wealth of efficient methods[1-3] have been proposed to realize motion deblurring, it remains a challenging problem in computer vision because the blur kernels of motion-blurred images vary with spatial locations that are more realistic and complicated.

Conventional methods[4,5] of motion deblurring based on optimization models usually estimate the blur kernels point by point, which requires a vast amount of memory and computation. An approximate optical field is typically used to overcome these problems. However, it is difficult to extend the approximation to realistic motion-blurred scenes, which affects the accuracy of the estimated blur kernels, thus greatly reducing the performance of motion deblurring. Moreover, the effectiveness of these conventional algorithms depends heavily on the priors of sharp images, which is another difficulty. Although these methods have achieved relatively high performance with great interpretability and generalization, they often involve iterative, time-intensive, and cumbersome optimization schemes compared with convolutional neural network (CNN)-based methods.

Recently, methods based on deep learning (DL) have been proposed for motion deblurring[6-10]. They offer the advantage of enabling motion deblurring at a high efficiency and low memory by omitting the iterative optimization stage. However, these DL-based methods suffer from two bottlenecks. First, public training datasets (such as the GoPro dataset[11]) are produced by averaging consecutive frames captured with a high-speed camera. The applicability of such datasets is limited because they can only represent motion blur to a limited extent, and their corresponding pre-trained solutions may be effective only for the camera that was used, capturing transformations other than image deblurring. Consequently, DL-based methods sometimes lack generality for blurred images. Second, the design of deblurring networks relies on researchers' experiments to verify its validity. Therefore, deblurring networks typically lack interpretability

* This work has been supported by the National Natural Science Foundation of China (No.62206143), and the Key Research and Development and Promotion Special Project in Henan Province (Nos.222102210141 and 232102211015).

** ZHANG Wei is an associate professor in China Labor Relations University. He received his master's degree in 2009 from Capital Normal University. His research interests mainly revolve around computer vision and image reconstruction. E-mail: zhangwculr@163.com.

and theoretical support.

The purpose of this study was to construct an end-to-end motion deblurring network called sparsity-guide low-rank network (SLRN) with stronger generalization and better interpretability to bridge the gap between methods based on optimization and DL. To this end, we first constructed a relatively complete training dataset based on the redundancy of per-pixel kernels in non-uniform motion-blurred images. We also decomposed the per-pixel kernels as linear combinations of elements on a much smaller basis. In conducted experiments, 10 000 real-world blurry images were used to adaptively produce the base kernels, and the training dataset was then constructed by applying these base kernels to the sharp images of the ImageNet dataset[12]. Through random combination of base kernels, the completeness of the training dataset was guaranteed as much as possible, thereby endowing our SLRN with a stronger generalization. In addition, we designed a plug-and-play low-rank module (LRM) to effectively capture the sparse features of sharp images, and the LRM was then integrated into a deep convolutional neural network (DCNN). The resulting network simultaneously exhibits strong feature representation ability of the DCNN and an implicit physical constraint. Therefore, the interpretability of the SLRN model is enhanced.

Fig.1 shows the deblurring performance of the SLRN trained on the GoPro dataset and the proposed dataset respectively. It was empirically verified that the proposed SLRN model trained on the proposed dataset can achieve a better deblurring performance on several datasets compared with the same model trained on the GoPro dataset.

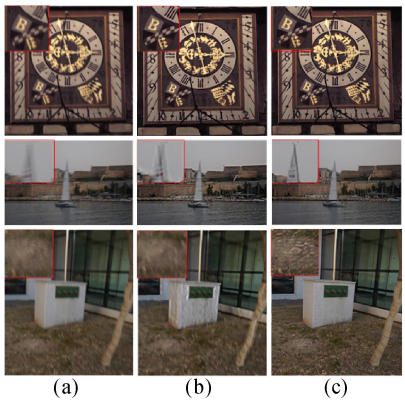

Fig.1 Cross-datasets deblurring performance of the SLRN: (a) Blur images taken from KÖHLER[13], LAI[14] and REALBLUR_J[15], respectively from top to bottom; (b) Deblurring results achieved by the SLRN trained on the GoPro dataset; (c) Deblurring results achieved by the SLRN trained on the proposed dataset

2. Related work

A wealth of sparse-based priors in optimization-based methods have been proposed to simultaneously recover latent sharp images and blur kernels, such as a mixture of Gaussian models[16], total variation regularization[17], and a student-t prior[18]. In addition, XU et al[19] employed an unnatural sparse representation in a gradient space to recover the latent sharp image. ZHANG et al[20] presented a collective latent image recovery algorithm that uses sparse representation to solve the challenging problem of facial image deblurring and recognition. TOFIGHI et al[21] used structured sparse representations and to automatically obtain an estimate of the kernel support and image support for blind motion deblurring. The dark channel prior that originated in the single-image haze removal problem was adopted by PAN et al[22] for blind image deblurring. Although these sparse priors have been successfully applied in natural image deblurring, they fail in some specific scenarios because of the non-applicability of image priors, such as text, face, and low-illumination images.

Recently, DL-based methods have been used for image restoration, wherein the sparsity property of clean images has been widely utilized for designing networks and achieving superior performances. To improve the deblurring performance, LIU et al[23] combined deep image priors with traditional TV regularization in a Unet architecture. These authors demonstrated that the addition of TV regularization results in significant performance gain when tested on image denoising and deblurring. CAI et al[24] proposed a dark and bright channel priors embedded network (DBCPeNet) to plug channel priors into a neural network for effective dynamic scene deblurring, where sparse regularization was introduced to regularize the DBCPeNet model learning. WU et al[25] proposed a structured analysis sparse coding method to learn external and internal image priors for image restoration. This method was then unfolded into a deep neural network. To maintain the sparsity of image edges, LI et al[26] proposed an updated scale-recurrent network (SRN) architecture for image deblurring that embeds the norm into the network architecture. ZHA et al[27] proposed a low-rankness guided group sparse representation (LGSR) model for image restoration, demonstrating its effectiveness in various applications. MANDRACCHIA et al[28] also highlighted the importance of optimal sparsity in achieving reliable system-aware restoration of fluorescence microscopy images. This approach was found to yield remarkable restoration of the fluorescence signal across different microscopy systems. According to the aforementioned studies, the networks constrained by the sparsity property can characterize the effective features of sharp images, which is helpful for the interpretability of the network architecture and the performance gain of image reconstruction. However, for image deblurring, the effective representation of sparse features is not sufficient to obtain a high deblurring performance, and the incompleteness of the training dataset leads to lack of generalization.

3. Method

An overview of the proposed method is shown in Fig.2. It contains two main parts: the generated training dataset and the designed network architecture. From a mathematical point of view, we first present theoretical support for training dataset generation and the network architecture.

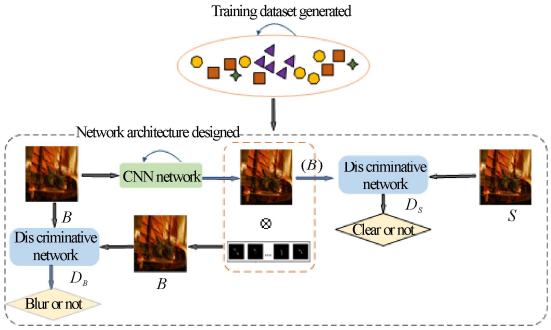

Fig.2 Overview of the proposed method

3.1 Sparsity-based motion blur degradation

For non-uniform motion-blurred images u, the blur kernel at each pixel is spatially varying. Given a sharp image with the size of $H\times W$, assuming that it is blurred at each pixel by the different blur kernel K_i with the size of $K\times K$, the observed blurry image can be formulated as

$$v_i = u_{nn(i)} \otimes k_i + n_i, \tag{1}$$

where $u_{nn(i)}$ is a window of size $K\times K$ around pixel i in image u, $\otimes$ is the two-dimensional (2D) convolution operator and n_i is additive noise. We assume that kernels are non-negative (no negative light) and of area one (conservation of energy).

Predicting all parameters of per-pixel kernels k_i would introduce an estimation problem in a very high-dimensional space K^2HW, being intractable for large images and kernels. We deal with this problem by assuming that there exists significant redundancy between the per-pixel kernels. To this end, we apply the low-rank constraint on the parameters of the kernels to realize the assumption. Specifically, the per-pixel blur kernels are decomposed as linear combinations of elements on a much smaller basis. If the blur kernel basis has B elements, then, only B coefficients m_i^b are required per pixel instead of the original $K\times K$. Additionally, the B basis elements need to be estimated resulting in an estimation problem of dimension $B(K^2+HW)$. Thus, the per-pixel kernel k_i results from the convex combination of the basis kernel, and the degradation model becomes

$$v_i = \langle u_{nn(i)}, \sum k^b m_i^b \rangle + n_i. \tag{2}$$

The process of sparsity-based motion blur degradation is shown in Fig.3. The coefficients of the weight mask are normalized so that they sum to one at each pixel location.

Based on the above analysis, given a pair of sharp/blur images, a group of basis kernels can be obtained by using convex optimization adaptively, and the optimization formula is as follows

$$\min_{k^1,k^2,\cdots k^B} \sum_i \left(\| v_i - u_{nn(i)} \otimes \left(\sum_{b=1}^{B} k^b m_i^b \right) \|^2 + \sum_{b=1}^{B} R(k^b) \right), \tag{3}$$

where $R(k^b)$ represents the prior of the basis kernel k^b, and $\|\cdot\|_1$ or $\|\cdot\|_2$ is often used to characterize the sparse characteristics.

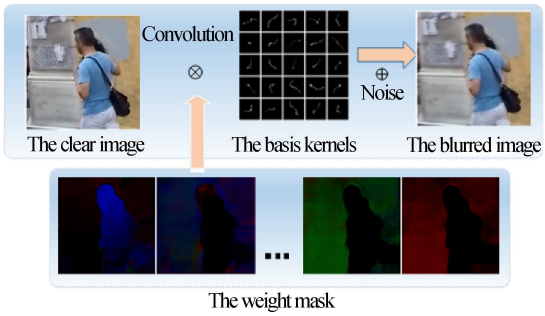

Fig.3 Process of sparsity-based motion blur degradation

3.2 Sparsity feature extraction and LRM module

Un-uniform motion blur images are intrinsically sparse (such as the coefficients in the frequency domain), we intend to design an appropriate network module to capture such sparse features. From a mathematical point of view, the richness of image information can be expressed by the rank of image matrix. In this case, sparse feature motion-blurred images are transformed into the 2D matrix, where each column represents a specific feature vector, and the rank of this 2D matrix can be regarded as the correlation of the columns. If the correlation is strong, it indicates that the matrix can be projected to the linear subspace with a lower dimension. Therefore, the sparse features of un-uniform motion-blurred images can be formulated as

$$\boldsymbol{\Psi}_{\mathrm{LR}}(U) = \boldsymbol{PQ}, \tag{4}$$

where $\boldsymbol{\Psi}_{\mathrm{LR}}(U)$ represents the sparse features of a motion-blurred image U, $\boldsymbol{P} \in R^{M\times K}$ and $\boldsymbol{Q} \in R^{K\times N}$ are the low-rank matrix with $K \ll M, N$. Thus low-rank matrices $\boldsymbol{\Psi}_{\mathrm{LR}}(U)$ can be reconstructed using the matrix product of $\boldsymbol{P}$ and $\boldsymbol{Q}$. Based on the above analysis, the upper limit of the rank for the reconstructed $\hat{\boldsymbol{\Psi}}_{\mathrm{LR}}(U)$ is K, which means $\hat{\boldsymbol{\Psi}}_{\mathrm{LR}}(U)$ is constrained to be a low-rank matrix. This matrix decomposition method will be applied in designing the deblurring network module to extract sparse features.

According to Eq.(4), we design the LRM to obtain sparse features of motion-blurred images. As shown in Fig.4, the input feature $\boldsymbol{F} \in R^{W\times H\times C}$ is convolved to be $\boldsymbol{F}_{rs} \in R^{WH\times C}$, then $\boldsymbol{Q} \in R^{WH\times K}$ is obtained by downsampling and other related operations, where $K=C/r$, r represents the ratio of downsampling. Subsequently, $\boldsymbol{P} = \boldsymbol{QF}^{\mathrm{T}} \in R^{K\times C}$ is achieved by matrix multiplication.

Finally, feature concatenation is completed by using skip-connection, so as to realize the LRM module according to Eq.(4) and mine low-rank prior of feature maps sufficiently.

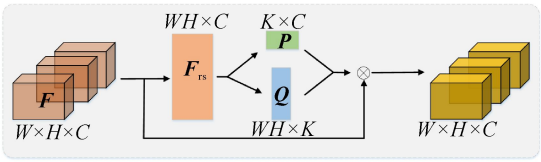

Fig.4 Process of sparsity-based motion blur degradation

The low-rank decomposition inherently captures the underlying structure and patterns within the data, which is particularly relevant in image deblurring tasks where the blur process can often be modeled as a linear transformation with a structured kernel. By decomposing the feature maps or the learned representations into lower-dimensional components, we enable the extraction of meaningful and sparse factors that contribute to the final restoration result. The rationale behind this approach lies in the fact that natural images typically exhibit high degrees of correlation and redundancy across their features, which can be effectively represented in a low-rank space. The module forces the network to learn such representations, thereby revealing the essential components involved in the deblurring process. Furthermore, the sparsity constraint applied on these decomposed components ensures that only the most influential and discriminative elements are retained, further enhancing the interpretability. Each component could potentially correspond to specific aspects of the blur (e.g., motion direction, blur intensity), making it easier to trace how the model interprets and mitigates the blur effect. In the following experimental section, we observed improved performance metrics, which indirectly attest to the model's capability to better understand and address the degradation phenomenon.

3.3 Implementation details

3.3.1 Training dataset generation

In the process of generating our training dataset, we randomly collected 10 000 motion blurred images from the Internet with the aim to utilize Eq.(3) for obtaining a diverse set of basis kernels (corresponding to different degradation factors). Subsequently, employing Eq.(2), we procedurally synthesized 50 000 pairs of sharp and blurred training images by operating on this kernel set and an independent collection of 10 000 clear images sourced from the ImageNet dataset. Specifically, for each clear image u in ImageNet, a group of blur basis kernels $\{k\}$ of blur kernels was chosen randomly and convolution was performed between u and $\{k\}$, as per the prescribed equation, incorporating Gaussian noise in the process. As illustrated in Fig.3, a random weight mask m was multiplied with the varied convolution results to ensure energy conservation within the degradation process. Consequently, this methodology yielded tuples for training, which consisted of the original clear image u^{GT}, the synthetically blurred image v^{GT}, the corresponding kernel set $\{k\}^{\mathrm{GT}}$, and the applied weight masks $\{m\}^{\mathrm{GT}}$.

3.3.2 Deblur network

Recently, excellent results have been achieved in image reconstruction using transformer networks. Inspired by this, we introduced the LRM into a transformer network to obtain better deblurring performance; the complete network model is shown in Fig.5. The Uformer network (proposed in Ref.[29] for image restoration) was adopted for our deblurring experiments, maintaining the encoder-decoder structure and the hierarchical structure of the skip connection. As shown in Fig.5(a), the proposed LRM was inserted in the middle of this deblurring network to extract sparse low-rank features. Specifically, given a degraded image $\boldsymbol{I} \in R^{3\times H\times W}$, Uformer first applied a 3×3 convolutional layer with a Leaky ReLU to extract low-level features $\boldsymbol{X}_0 \in R^{C\times H\times W}$. Next, the feature maps $\boldsymbol{X}_0$ are passed through K-encoder stages. Each stage contains a stack of locally enhanced window (LeWin) transformer blocks and one downsampling layer. A bottleneck stage with a stack of LeWin transformer blocks was added at the end of the encoder. For feature reconstruction, the decoder also contains K stages that consist of an upsampling layer and a stack of LeWin transformer blocks, similar to the encoder. A LeWin transformer block, shown in Fig.5(b), benefits from self-attention in the transformer to capture long-range dependencies and involves the convolution operator in the transformer to capture a useful local context. A multiscale restoration modulator applies multiple modulators to the Uformer

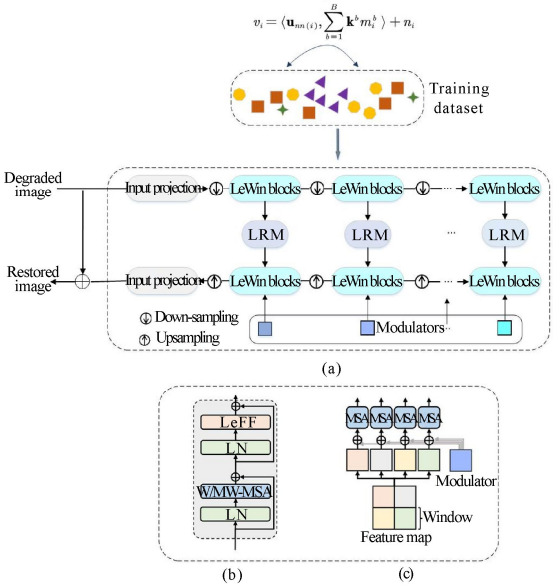

Fig.5 (a) Uformer network plugged by the LRM for the conducted experiments; (b) LeWin Transformer block; (c) Illustration of how the modulators modulate the W-MSAs in each LeWin Transformer block, which is named as MW-MSA in (b)

decoder, as shown in Fig.5(a) and (c). Note that the focus of this study was not to propose an optimal deblurring network but to verify the effectiveness of the proposed LRM and training dataset generation, which can be arbitrarily applied to existing deblurring networks.

4. Experiments

The focus of this study was to verify the effectiveness of the proposed LRM and generalized training dataset. To provide a fair comparison, we used standard deblurring benchmarks: GoPro[11], KÖHLER[13], LAI[14], and REALBLUR_J[15] to evaluate their effectiveness. The experiments were implemented on a personal computer (PC) equipped with an NVIDIA RTX 4090 graph processing unit (GPU). Unless otherwise specifically stated, all the experiments followed the same settings, i.e., the maximum number of iterations was set to 10 000 for training, the Adam optimizer was used on patches of 256×256, the batch size was set to 4, and the learning rate was set to 0.000 1. To ensure fairness in comparisons, all comparative models employ the mean squared error (*MSE*) as the loss function. This consistency in the choice of optimization metric guarantees equitable benchmarking across the board.

4.1 Ablation experiments

In this study, we validate the effectiveness of our proposed LRM through empirical experimentation. To substantiate its performance enhancement, a series of ablation studies were conducted on the widely recognized GoPro dataset, comparing the deblurring capabilities of existing models such as multi-scale convolutional neural network (MSCNN)[11], robust and accurate deblurring network (RADN)[8], and Uformer[29]. The experimental results for different deblurring methods with and without the inclusion of LRM are presented in Tab.1.

Observably, the integration of LRM into these deblurring methodologies leads to a notable improvement in deblurring performance, evidenced by an approximate increase of ~0.20 dB in peak signal-to-noise ratio (*PSNR*) compared to their respective baselines. This underscores our objective to integrate the LRM within a deblurring network to extract more refined features. Tab.1 demonstrates that the models augmented with the LRM consistently achieve superior deblurring performance in terms of both *PSNR* and structural similarity index measure (*SSIM*), thereby providing robust empirical evidence for the efficacy of the LRM. Bolded values indicate better deblurring performance. These findings not only validate the utility of our proposed LRM but also highlight its potential to enhance the state-of-the-art in image deblurring tasks.

To verify the effectiveness of the generalized training dataset, the deblurring methods of MSCNN, RADN, and Uformer were trained on the GoPro and proposed generalized training datasets, and then the pre-trained models were tested directly on three datasets, namely GoPro[11], KÖHLER[13], and LAI[14], to evaluate the generalization. The GoPro dataset was constructed using 240 fps videos captured using the GoPro camera; these videos contain diverse three-dimensional scenes, where 2 103 pairs were used for training and 1 111 pairs for evaluation. The KÖHLER and LAI datasets are standard deblurring benchmarks that contain 48 and 100 pairs of synthetic motion blur/sharp images, respectively.

Tab.1 Comparison of different deblurring methods with/without LRM on the GoPro dataset

Description	*PSNR*	*SSIM*
MSCNN[11]	29.23	0.916 2
MSCNN+LRM	**29.35**	**0.920 1**
RADN[8]	30.11	0.928 7
RADN+LRM	**30.43**	**0.931 0**
Uformer[29]	32.97	0.967 1
Uformer+LRM	**33.14**	**0.969 7**

The experimental results are listed in Tab.2. When possible, we reproduced the results using the available code; otherwise, the values were obtained from their own paper. ^{*G} represents the deblurring models trained on the GoPro dataset and ^{*P} represents the deblurring models trained on the proposed training dataset. The bold values represent better performance. It can be observed that the models trained on the GoPro dataset present the best deblurring performance on the test images of GoPro. However, a lower performance is obtained on the KÖHLER and LAI datasets, which indicates that the GoPro dataset limits the generalization of existing deblurring models. In contrast, models trained on the proposed training dataset can achieve better deblurring performance on the KÖHLER and LAI datasets. Even on the test images of GoPro, the deblurring performance of the pre-trained models trained on the proposed dataset was only slightly lower than that trained on the GoPro dataset. Therefore, the proposed training dataset generation is helpful for enhancing the generalization of the deblurring models.

Tab.2 Cross-datasets quantitative comparison for image deblurring

Method	*PSNR*			*SSIM*		
	GoPro	KÖHLER	LAI	GoPro	KÖHLER	LAI
MSCNN^{*G}	**29.23**	25.13	16.51	**0.916 2**	0.710 7	0.810 2
MSCNN^{*P}	28.95	**25.47**	**16.91**	0.913 7	**0.732 3**	**0.821 5**
RADN^{*G}	**30.11**	27.25	19.34	**0.928 7**	0.763 7	0.842 7
RADN^{*P}	30.03	**27.94**	**19.61**	0.921 9	**0.792 7**	**0.851 9**
Uformer^{*G}	**32.97**	28.07	21.57	**0.973 1**	0.801 5	0.861 4
Uformer^{*P}	32.68	**28.39**	**22.04**	0.961 5	**0.823 4**	**0.871 3**

Fig.6 and Fig.7 show some visual comparisons on the

LAI and KÖHLER datasets between the pre-trained models trained on the GoPro dataset and the proposed-training dataset. The first row shows the deblurring results of the models trained on the GoPro dataset and the second row shows the deblurring results of the models trained on the proposed training dataset. As can be observed, the deblurring results in the second row have sharper edges, resulting in better visual effects (for example, the person in Fig.6 and the moon in Fig.7). These observations further verified the generalization of the proposed training dataset.

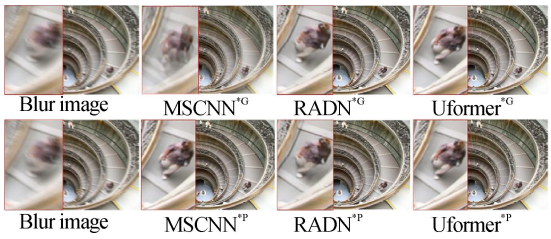

Fig.6 Visual comparison of the pre-trained model on the LAI dataset[14]

Fig.7 Visual comparison of the pre-trained model on the KÖHLER dataset[13]

4.2 Performance on Uformer plugged by the LRM module

Uformer is a general U-shaped transformer for image restoration in Ref.[29], which we empirically verified that integrating the LRM with Uformer trained on the proposed training dataset can outperform state-of-the-art methods on the KÖHLER, LAI, and REALBLUR_J datasets for non-uniform motion deblurring, indicating the generalization of the proposed method. First, we incorporated the LRM into Uformer (as shown in Fig.5), and then conducted quantitative and qualitative comparisons with recently proposed DL-based methods, including MSCNN[11], DeblurGAN[30], spatially variant recurrent neural network (SVRNN)[7], SRN[31], parameter selective sharing (PSS)-SRN[32], RADN[8] and non-autoregressive fusion network (NAFNet)[33]. The above-mentioned six methods were trained on the GoPro dataset, and our method was trained on the proposed generalized dataset. *PSNR* and *SSIM* were also used as quantitative metrics.

The experimental results are presented in Tab.3. It can be observed that our model ranks the first among all these end-to-end methods in terms of *PSNR* and *SSIM*, except for REALBLUR_J. This may be because the images in REALBLUR_J are low-light, which is significantly different from the proposed training dataset generated from ImageNet[12], resulting in insufficient priors for image restoration. Even so, our method ranks the second in deblurring performance in terms of *SSIM*, which demonstrates the generalization of the proposed method on different datasets.

Fig.8 shows some visual comparisons between the proposed model and state-of-the-art deblurring algorithms on real-world images from the LAI dataset, for which no corresponding ground truth exists for the real blurry images, allowing only visual comparison. It can be observed that Uformer plugged by the LRM can achieve the best visual effect of reconstruction, outperforming the rest of methods, which indicates that the proposed model can process real scenes. Owing to space limitations, only four deblurring results are presented in Fig.8. However, the overall deblurring results on the four datasets, i.e., GoPro, KÖHLER, LAI, and REALBLUR_J, are available at Google Drive as reference.

5. Conclusion

We presented a method for generalized training dataset generation and an LRM to enhance the generalization and interpretability of existing deblurring networks. The LRM was designed to refine the sparse features of motion-blurred images. The LRM can be integrated into existing deblurring networks as a plug-and-play module to improve the interpretability of the deblurring models. To verify their effectiveness, a comparison of the deblurring performance of some existing models plugged by

Tab.3 Quantitative comparisons of Uformer integrated by LRM module on different datasets

Method	*PSNR*			*SSIM*		
	REALBLUR_J	KÖHLER	LAI	REALBLUR_J	KÖHLER	LAI
MSCNN	27.87	25.13	16.51	0.827 1	0.710 7	0.810 2
DeblurGAN	28.03	26.15	17.23	0.834 3	0.753 1	0.831 5
SVRNN	28.15	27.03	17.84	0.842 6	0.812 3	0.845 1
SRN	27.96	26.37	17.51	0.837 5	0.801 4	0.832 4
PSS-SRN	28.47	27.06	18.76	0.849 7	0.814 5	0.840 3
RADN	28.61	27.25	19.34	0.861 8	0.763 7	0.842 7
NAFNet	28.84	28.19	20.95	0.860 1	0.815 2	0.865 3
Proposed	28.95	28.39	21.04	0.857 3	0.823 4	0.871 3

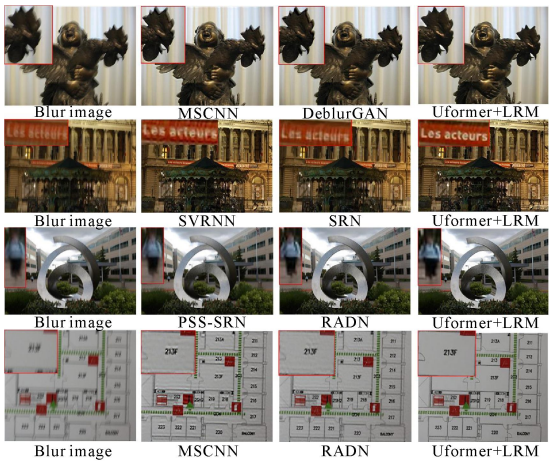

Fig.8 Visual comparisons of deblurring results generated by the proposed model and the state-of-the-art deblurring models on the real-world images of the LAI dataset[14]

the LRM was presented. Moreover, we reported the deburring performance on different test datasets for methods trained on the GoPro dataset and the proposed training dataset. It was empirically verified that the proposed SLRN model (Uformer plugged by the LRM) trained on the proposed training dataset can outperform state-of-the-art methods on three standard deblurring benchmarks, which further indicates the effectiveness of the proposed training dataset and LRM.

Ethics declarations

Conflicts of interest

The authors declare no conflict of interest.

References

[1] WANG M, ZHU F, BAI Y. An improved image blind deblurring based on dark channel prior[J]. Optoelectronics letters, 2021, 17(1): 40-46.

[2] LU Y C, LIU T P, LIN C H. Two-stage single image deblurring network based on deblur kernel estimation[J]. Multimedia tools and applications, 2023, 82(11): 17055-17074.

[3] SUN Y, ZHI X, JIANG S, et al. Image fusion for the novelty rotating synthetic aperture system based on vision transformer[J]. Information fusion, 2024, 104: 102163.

[4] ZHANG S, TANG G, LIU X, et al. Retinex based low-light image enhancement using guided filtering and variational framework[J]. Optoelectronics letters, 2018, 14(2): 156-160.

[5] DELBRACIO M, GARCIA-DORADO I, CHOI S, et al. Polyblur: removing mild blur by polynomial reblurring[J]. IEEE transactions on computational imaging, 2021, 7: 837-848.

[6] GONG D, YANG J, LIU L, et al. From motion blur to motion flow: a deep learning solution for removing heterogeneous motion blur[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, Hawaii, USA. New York: IEEE, 2017: 2319-2328.

[7] ZHANG J, PAN J, REN J, et al. Dynamic scene deblurring using spatially variant recurrent neural networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June 18-22, 2018, Salt Lake City, Utah, USA. New York: IEEE, 2018: 2521-2529.

[8] PUROHIT K, RAJAGOPALAN A N. Region-adaptive dense network for efficient motion deblurring[C]//Proceedings of the AAAI Conference on Artificial Intelligence, February 7-12, 2020, New York, USA. Washington: AAAI Press, 2020: 11882-11889.

[9] WEN Y, CHEN J, SHENG B, et al. Structure-aware motion deblurring using multi-adversarial optimized cyclegan[J]. IEEE transactions on image processing, 2021, 30: 6142-6155.

[10] LI Y, REN D, SHU X, et al. Learning single image defocus deblurring with misaligned training pairs[C]//Proceedings of the AAAI Conference on Artificial Intelligence, February 7-10, 2023, Washington DC, USA. Washington: AAAI Press, 2023, 37(2): 1495-1503.

[11] NAH S, HYUN KIM T, MU LEE K. Deep multi-scale convolutional neural network for dynamic scene deblurring[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, Hawaii, USA. New York: IEEE, 2017: 3883-3891.

[12] DENG J, DONG W, SOCHER R, et al. Imagenet: a large-scale hierarchical image database[C]//Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, June, 2009, Miami, Florida, USA. New York: IEEE, 2009: 248-255.

[13] HIRSCH M, SRA S, SCHÖLKOPF B, et al. Efficient filter flow for space-variant multiframe blind deconvolution[C]//Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June, 2010, San Francisco, California, USA. New York: IEEE, 2010: 607-614.

[14] LIU Y L, LAI W S, CHEN Y S, et al. Single-image HDR reconstruction by learning to reverse the camera pipeline[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, June, 2020, Seattle, Washington, USA. New York: IEEE, 2020: 1651-1660.

[15] RIM J, LEE H, WON J, et al. Real-world blur dataset for learning and benchmarking deblurring algorithms[C]//16th European Conference on Computer Vision, August 23-28, 2020, Glasgow, UK. Heidelberg: Springer International Publishing, 2020: 184-201.

[16] LJUBENOVIĆ M, FIGUEIREDO M A T. Blind image deblurring using class-adapted image priors[C]//

Proceedings of the 2017 IEEE International Conference on Image Processing, September, 2017, Beijing, China. New York: IEEE, 2017: 490-494.

[17] XIE J, HOU G, WANG G, et al. A variational framework for underwater image dehazing and deblurring[J]. IEEE transactions on circuits and systems for video technology, 2021, 32(6): 3514-3526.

[18] MOHAMMAD-DJAFARI A, DUMITRU M. Bayesian sparse solutions to linear inverse problems with non-stationary noise with Student-t priors[J]. Digital signal processing, 2015, 47: 128-156.

[19] HU Z, HUANG J B, YANG M H. Single image deblurring with adaptive dictionary learning[C]//Proceedings of the 2010 IEEE International Conference on Image Processing, September, 2010, Hong Kong, China. New York: IEEE, 2010: 1169-1172.

[20] ZHANG H, YANG J, ZHANG Y, et al. Close the loop: joint blind image restoration and recognition with sparse representation prior[C]//Proceedings of the 2011 International Conference on Computer Vision, November, 2011, Barcelona, Spain. New York: IEEE, 2011: 770-777.

[21] TOFIGHI M, LI Y, MONGA V. Blind image deblurring using row-column sparse representations[J]. IEEE signal processing letters, 2017, 25(2): 273-277.

[22] PAN J, SUN D, PFISTER H, et al. Deblurring images via dark channel prior[J]. IEEE transactions on pattern analysis and machine intelligence, 2017, 40(10): 2315-2328.

[23] LIU J, SUN Y, XU X, et al. Image restoration using total variation regularized deep image prior[C]//Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, May, 2019, Brighton, UK. New York: IEEE, 2019: 7715-7719.

[24] CAI J, ZUO W, ZHANG L. Dark and bright channel prior embedded network for dynamic scene deblurring[J]. IEEE transactions on image processing, 2020, 29: 6885-6897.

[25] WU F, DONG W, HUANG T, et al. Hybrid sparsity learning for image restoration: an iterative and trainable approach[J]. Signal processing, 2021, 178: 107751.

[26] LI M, GAO S, ZHANG C, et al. Blind motion deblurring via L0 sparse representation[J]. Computers & graphics, 2021, 97: 248-257.

[27] ZHA Z, WEN B, YUAN X, et al. Low-rankness guided group sparse representation for image restoration[J]. IEEE transactions on neural networks and learning systems, 2022.

[28] MANDRACCHIA B, LIU W, HUA X, et al. Optimal sparsity allows reliable system-aware restoration of fluorescence microscopy images[J]. Science advances, 2023, 9(35): 9245.

[29] WANG Z, CUN X, BAO J, et al. Uformer: a general U-shaped transformer for image restoration[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, June, 2022, New Orleans, Louisiana, USA. New York: IEEE, 2022: 17683-17693.

[30] KUPYN O, BUDZAN V, MYKHAILYCH M, et al. Deblurgan: blind motion deblurring using conditional adversarial networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June, 2018, Salt Lake City, Utah, USA. New York: IEEE, 2018: 8183-8192.

[31] TAO X, GAO H, SHEN X, et al. Scale-recurrent network for deep image deblurring[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, June, 2018, Salt Lake City, Utah, USA. New York: IEEE, 2018: 8174-8182.

[32] GAO H, TAO X, SHEN X, et al. Dynamic scene deblurring with parameter selective sharing and nested skip connections[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, June, 2019, Long Beach, California, USA. New York: IEEE, 2019: 3848-3856.

[33] CHEN L, CHU X, ZHANG X, et al. Simple baselines for image restoration[C]//18th European Conference on Computer Vision, August, 2022, Copenhagen, Denmark. Heidelberg: Springer International Publishing, 2022: 17-33.