

TurboMPLE: Joint Infrared Turbulence Mitigation and Physical Fields Estimation via Mutual Progressive Layered Extraction

Yitong An¹, Yubo Jiang¹, and Xiangzhi Bai^{1,2,3*}

¹ Image Processing Center, Beihang University, Beijing 102206, China

² State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing 100191, China

³ Key Laboratory of Spacecraft Design Optimization and Dynamic Simulation Technology, Ministry of Education, Beihang University, Beijing 102206, China
{20374038, jbond0409, jackybxz}@buaa.edu.cn

Abstract. Atmospheric turbulence severely degrades thermal infrared imaging quality, yet most existing mitigation methods focus solely on image restoration while ignoring underlying physical characteristics. In this study, we construct a large-scale infrared turbulence imaging dataset and proposed the Turbulence-oriented Mutual Progressive Layered Extraction (TurboMPLE), a joint network for infrared turbulence mitigation and physical fields estimation. Through Mutual PLE Blocks, TurboMPLE achieves progressive feature-level collaboration, introducing turbulence priors into the mitigation process while incorporating richer visual information into physical fields estimation. In addition, embedding the Lambert–Beer law and turbulent physics constraints enables more interpretable results. With a lightweight and computationally efficient design, TurboMPLE achieves state-of-the-art performance on both tasks and demonstrates strong generalization capability on real-world data. The code is available at <https://github.com/Ayt777/TurboMPLE>.

Keywords: Atmospheric turbulence · Turbulence mitigation · Physical fields estimation · Multi-task learning

1 Introduction

Imaging systems suffer severe quality degradation due to atmospheric turbulence, which primarily induces spatiotemporal blur and geometric distortion through random refractive index fluctuations. Effective turbulence mitigation (TM) remains a fundamental challenge in image restoration. Conventional TM methods [18, 46, 47, 57] are mostly based on frame alignment and temporal fusion, and thus struggle to handle dynamic scenes effectively. In recent years, deep learning-based methods [14, 15, 29, 35, 42, 45, 50–52] have achieved remarkable progress. However, most of them neglect the turbulence-related physical characteristics

* Corresponding author

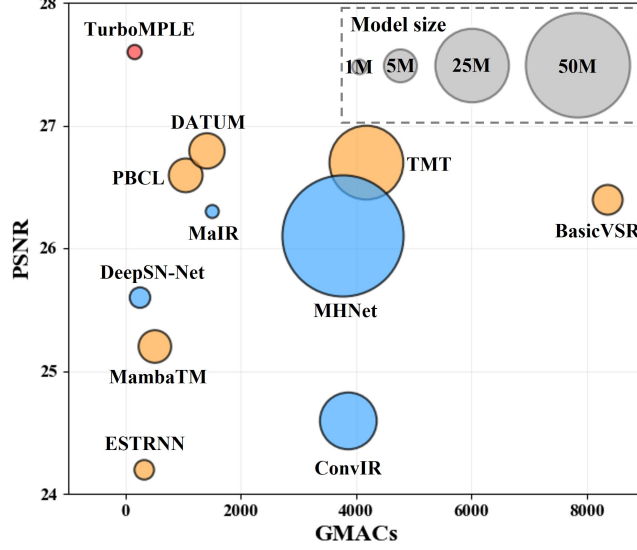


Fig. 1: Benchmarking models for turbulence mitigation. The proposed TurboMPLE (red) sets the state-of-the-art, outperforming video- (orange) and image-based (blue) networks in both performance and efficiency. All computational costs (GMACs) are measured on a 30-frame sequence of 256×256 images.

directly associated with image degradation, thereby rendering their generalization capability entirely dependent on synthetic training data. To address this issue, [51] introduced a latent phase distortion (LPD) module to estimate the tilt and blur representations, and further proposed a phase-distortion-guided Mamba block to enable degradation-aware turbulence mitigation. However, the LPD representations, learned through a variational autoencoder, remain latent variables lacking direct physical interpretability. [43] introduced a physically boosted cooperative learning (PBCL) framework that adopts a three-stage strategy: independent estimation of the turbulence strength fields, image restoration conditioned on the estimated fields, and subsequent refinement of the estimated fields based on restored images. Such sequential designs suffer from error accumulation and require multi-stage training, thereby hindering end-to-end joint optimization and constraining the ability to learn complementary representations. In addition, [43] assumes static turbulence fields and neglects the essential temporal dynamics, limiting its capability in video turbulence mitigation.

To address these issues, we propose the Turbulence-oriented Mutual Progressive Layered Extraction (TurboMPLE), a joint framework for infrared turbulence mitigation and physical fields estimation. The estimated fields include the refractive index structure constant (C_n^2) and temperature structure constant (C_T^2), which are commonly used to quantify the strength of atmospheric turbu-

lence, the energy dissipation rate (ε), which characterizes the decay of turbulent kinetic energy, and the Reynolds stress (RS), which quantifies momentum transfer resulting from velocity fluctuations. Leveraging feature-level mutual design, TurboMPLE achieves improved performance on both tasks. Specifically, we make the following key contributions:

1. We constructed a large-scale infrared turbulence imaging dataset comprising both synthetic and real-world data, totaling 181,290 frames, with dynamic physical fields available for synthetic data.
2. TurboMPLE is the first method to jointly perform turbulence mitigation and physical fields estimation within a multi-task learning framework, enabling end-to-end optimization and incorporating feature-level mutual guidance. Also, embedding Lambert–Beer law and turbulent physics constraints promote interpretable results.
3. Extensive experiments on synthetic and real-world data demonstrate TurboMPLE achieves state-of-the-art performance on both tasks, while maintaining a lightweight architecture and computational efficiency.

2 Related works

2.1 Turbulence mitigation

Conventional turbulence mitigation methods typically address the problem as a many-to-one image restoration task. The standard pipeline involves two main stages: frame alignment and temporal fusion [18, 46, 47, 57]. Alignment corrects turbulence-induced pixel displacement and tilt, typically by creating a tilt-free reference (e.g., temporal average [37, 57]) and registering all frames to it. Fusion exploits the “lucky effect” [10] through methods like lucky-region fusion [1] to aggregate the clearest regions across frames. For dynamic scenes, conventional methods usually assume rigid motion in moving areas, restoring only the static background via the standard pipeline [28, 31].

Recently, deep learning methods have significantly advanced the field of turbulence mitigation, and can be broadly categorized based on their input domain: single-frame and multi-frame methods. Single-frame TM models [14, 29, 42, 45] process individual images relying on complex network architectures, such as generative models [30, 33], to restore the image quality. Multi-frame TM algorithms [15, 35, 50–52] leverage the temporal dimension, aggregating information across a sequence of images to address the distortions and blur caused by turbulence. However, most existing methods primarily target the singular objective of image restoration, thereby neglecting valuable physical information, such as the underlying turbulent parameter fields directly inherent to the degradation process.

2.2 Turbulence estimation

Atmospheric turbulence has been quantified through both physical measurements and data-driven estimations. Traditional methods include radar sens-

ing, temperature-based observation, and optical beam transmission. Radar techniques [6, 13, 17] infer parameters such as C_n^2 and ε from very-high-frequency or Doppler radar variables, including wind-velocity variance, spectral width, and echo power. Temperature fluctuation methods [44] employ thermosondes to measure local thermal variations, from which C_T^2 is derived and converted to C_n^2 via:

$$C_T^2 = \langle [T(x) - T(x+r)]^2 \rangle r^{-2/3}, \quad l_0 \ll r \ll L_0, \quad (1)$$

$$C_n^2 = (79 \times 10^{-6} P/T^2)^2 C_T^2, \quad (2)$$

where $\langle \cdot \rangle$ is the ensemble average, T is the temperature, P is the local atmospheric pressure, l_0 and L_0 represent the inner and outer scales of turbulence, respectively. Optical-beam transmission methods [9, 27, 56] analyze intensity or phase fluctuations of propagating light to estimate integrated or path-resolved turbulence strength.

Recent advances leverage visual observations for turbulence estimation. Phase-based image analysis has been used to derive path-integrated C_n^2 and atmospheric coherence length from temporal motion patterns [3]. A physics-informed convolutional network further couples spatiotemporal features with differentiable gradient modeling for C_n^2 estimation from video sequences [36]. To extend beyond one-dimensional inference, PBCL quantifies two-dimensional turbulence strength (C_n^2 , C_T^2) from infrared imagery [43]. However, it still assumes static fields despite the inherently dynamic nature of atmospheric turbulence.

2.3 Multi-task learning

Multi-task learning (MTL) addresses multiple tasks within a single model and is widely applied in image processing [2, 20, 49]. Without building separate models, MTL reduces computational costs while leveraging shared information across related tasks. Early MTL approaches employ hard parameter sharing [4], where tasks share a common bottom network before diverging into task-specific towers. However, Shared-Bottom models suffer from gradient conflicts and sensitivity to task correlations. To address this, Multi-gate Mixture-of-Experts (MMoE) structure [24] introduces multiple shared experts and task-specific gates that dynamically weight expert contributions. Building on MMoE, Progressive Layered Extraction (PLE) model [39] decouples shared and task-specific experts via a multi-layer structure, where task-specific gates fuse experts at each layer to prevent interference while preserving collaborative learning.

3 Dataset Construction

For the multi-task learning framework, a dataset comprising both real-world and synthetic turbulence-distorted infrared sequences was constructed. The real-world data were acquired using a FLIR A615 thermal infrared camera in Beijing,

Hubei, and Chongqing, China, amounting to 80,580 frames at 880×640 resolution. Specifically, the infrared sequences were collected in Beijing (Changping District, $116^\circ 10'E-116^\circ 18'E$, $40^\circ 8'N-40^\circ 10'N$; altitude 40–50 m), Chongqing (Yuzhong District and Liangjiang New District, $106^\circ 31'E-106^\circ 34'E$, $29^\circ 33'N-29^\circ 35'N$; altitude 200–300 m), and Hubei (Taichi Lake and Wudang Mountains, $110^\circ 56'E-111^\circ 15'E$, $32^\circ 22'N-32^\circ 35'N$; altitude 400–700 m), with a corresponding atmospheric pressure range of 980–1015 hPa. The collected scenes feature predominantly static backgrounds, including buildings, streets, and vegetation, while also containing dynamic targets such as humans and vehicles, all affected by atmospheric turbulence. Recordings were conducted in high-temperature conditions ($34-40^\circ C$), within the long-wave infrared range ($8-13 \mu m$), and at distances of several tens of meters.

The synthetic dataset contains 3,357 sequences, totaling 100,710 frames at 256×256 resolution. Each degraded frame is paired with its corresponding restoration label and the turbulence physical fields with spatial resolution 16×16 . The synthetic data were organized into training, validation, and test sets, with 2,664, 255, and 438 sequences, respectively. The dataset is further characterized by the statistics of its turbulence fields: C_n^2 ($m^{-2/3}$), C_T^2 ($K^2 m^{-2/3}$), ε (m^2/s^3), and RS (m^2/s^2) span ranges of $[1.39 \times 10^{-19}, 8.32 \times 10^{-12}]$, $(0, 13.59]$, $(0, 8152.42]$, and $(0, 87.27]$, with mean/std values of $1.28 \times 10^{-12}/7.78 \times 10^{-13}$, $1.87/1.21$, $660.73/532.56$, and $12.70/7.95$, respectively. For the test set, sequences were categorized into weak, medium, and strong turbulence levels based on the simulated magnitudes of C_n^2 and temperature. Details of the turbulence simulation pipeline are provided in the supplementary material.

4 Method

4.1 Overview

Existing turbulence mitigation methods often neglect the turbulence-related physical characteristics. [43] introduced a physically boosted cooperative learning (PBCL) framework to address this. As shown in Fig. 2(a), PBCL employs a multi-stage training design that separates turbulence estimation and mitigation, which leads to error accumulation and prevents end-to-end optimization. Moreover, it assumes static turbulence fields, ignoring temporal dynamics crucial for video restoration. In contrast, our proposed TurboMPLE (Fig. 2(b)) adopts a multi-task learning framework that enables feature-level mutual learning among experts, while explicitly modeling the temporal dynamics of turbulence to achieve more coherent video restoration. Fig. 2(c) shows the architecture of the proposed TurboMPLE. Specifically, given an image sequence $I \in \mathbb{R}^{T \times 1 \times H \times W}$, the encoder employs a Bouguer–Lambert–Beer law-based embedding, stacked convolutions, and temporal attention to capture spatio-temporal feature $F_{st} \in \mathbb{R}^{T \times C \times H/4 \times W/4}$ with the feature channel dimension C set to 64. The feature is then processed with N MutualPLE blocks, each comprising task-specific experts, shared experts, a Cross-Task Mutual Block (CTMB),

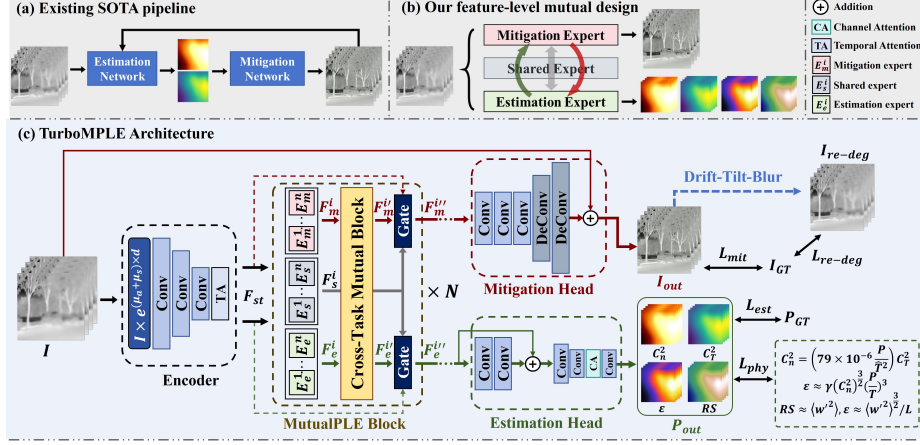


Fig. 2: The proposed TurboMPLE network. (a) and (b) illustrate the workflow of the existing state-of-the-art method [43] and TurboMPLE, respectively. (c) shows the architecture of TurboMPLE. The detailed architectures of experts and CTMB are provided in the supplementary material.

and adaptive gating mechanisms, yielding task-specific representations for turbulence mitigation ($F_m^{N''} \in \mathbb{R}^{T \times C \times H/4 \times W/4}$) and parameter estimation ($F_e^{N''} \in \mathbb{R}^{T \times C \times H/4 \times W/4}$), respectively. These task-specific features are subsequently fed into the mitigation head and estimation head to produce the restored images $I_{out} \in \mathbb{R}^{T \times 1 \times H \times W}$ and the predicted physical fields $P_{out} \in \mathbb{R}^{T \times 4 \times H/16 \times W/16}$. Details are provided in the following subsections.

4.2 TurboMPLE architecture

Bouguer–Lambert–Beer Law–Embedded Encoder. In thermal infrared imaging, atmospheric turbulence leads to both optical distortions (blur and tilt) and grayscale drift induced by heat dissipation. To better capture the optical distortions, the Bouguer–Lambert–Beer Law (BLB) is embedded at the beginning of encoder, aiming to suppress grayscale drift and alleviate thermal radiation artifacts. Specifically, the attenuation of infrared radiation in the atmosphere is described by BLB as [40]:

$$I_1 = I \times e^{(\mu_a + \mu_s) \times d} \quad (3)$$

where I_1 is the compensated input, μ_a , μ_s denote attenuation coefficient of absorption and scattering, and d is the transmission distance. The learnable parameters μ_a , μ_s , and d are initialized to 0, 0, and 1, respectively, and are constrained via a sigmoid mapping to $[0, 1]$ for μ_a and μ_s , and $[0, 10]$ for d . Subsequently, I_1 is processed through multiple convolutional layers and temporal attention unit [38] to produce the spatio-temporal feature F_{st} .

Mitigation, Estimation, and Share Experts. Each MutualPLE block contains task-specific experts and shared experts. Since turbulent degradation exhibits both spatial distortion patterns and temporal dynamics, we adopt a dual-branch architecture following [48] to decompose the feature learning in shared experts. The convolutional branch integrates a turbulence-aware modulation mechanism that dynamically adjusts features. Specifically, a lightweight estimator predicts a turbulence map $T = \sigma(f_{est}(F_{st}))$, which guides an adaptive modulator to produce spatially varying weights $W = \sigma(f_{mod}([F_{st}, T]))$. Enhancing turbulence-sensitive regions while suppressing redundant responses, the modulated feature is obtained as $F_{conv} = F_{st} \odot W$, where $\odot$ represents element-wise multiplication. Subsequent convolutions further refine the modulated features to obtain rich spatial degradation features. Meanwhile, the SSM branch exploits the strength of state-space models in efficiently capturing long-range temporal dependencies by employing the 2D-Selective-Scan (SS2D) [21] to model the temporal dynamics of turbulent sequences, thereby complementing the spatial representations extracted by the turbulence-aware convolution branch. Finally, the outputs from two branches are concatenated and refined through a channel-shuffle operation to form the shared feature F_s^i . For the mitigation and estimation experts, a convolutional dual-branch architecture is employed to efficiently extract global-local information, resulting in F_m^i and F_e^i .

Cross-Task Mutual Block (CTMB). To enable feature-level mutual learning between the mitigation and estimation tasks, the CTMB establishes a bidirectional information exchange pathway. Specifically, two cross-attention modules are deployed in opposite directions, where one guides the mitigation experts to integrate turbulence priors from the estimation branch, and the other enables the estimation experts to utilize visual features for richer turbulence strength information:

$$\begin{aligned}\tilde{F}_m^i &= \text{CrossAttn}_{m \leftarrow e}(F_m^i, F_e^i), \\ \tilde{F}_e^i &= \text{CrossAttn}_{e \leftarrow m}(F_e^i, F_m^i),\end{aligned}\tag{4}$$

where $\tilde{F}_m^i$ and $\tilde{F}_e^i$ are enhanced mitigation and estimation features. The features are then adaptively fused with the original features using the progressive mutual weight:

$$W_m^i = \omega^i \Phi_m([F_m^i, \tilde{F}_m^i]), \quad W_e^i = \omega^i \Phi_e([F_e^i, \tilde{F}_e^i])\tag{5}$$

$$\begin{aligned}F_m^{i'} &= F_m^i \odot (1 - W_m^i) + \tilde{F}_m^i \odot W_m^i, \\ F_e^{i'} &= F_e^i \odot (1 - W_e^i) + \tilde{F}_e^i \odot W_e^i,\end{aligned}\tag{6}$$

where $[\cdot, \cdot]$ denotes channel concatenation, Φ_m and Φ_e are pointwise convolutions followed by a Sigmoid activation, and ω^i corresponds to the progressive mutual weight, defined as $\omega^i = 0.2 + i \times 0.4$ for the i -th MutualPLE block, gradually increasing to enhance cross-task cooperation in deeper blocks. Finally, the gating networks are employed to adaptively fuse task-specific and shared features:

$$\begin{aligned}F_m^{i''} &= \alpha_m^i[F_s^i, F_m^{i'}], \quad \alpha_m^i = \mathcal{G}(F_m^{i-1}), \\ F_e^{i''} &= \alpha_e^i[F_s^i, F_e^{i'}], \quad \alpha_e^i = \mathcal{G}(F_e^{i-1}),\end{aligned}\tag{7}$$

where $\mathcal{G}$ denotes the gating function implemented by global pooling, linear projection and Softmax activation, α_m^i and α_e^i are adaptive gating weights, $F_m^{i''}$ and $F_e^{i''}$ are the outputs of the i -th MutualPLE block. F_e^{i-1} and F_m^{i-1} are input features of the i -th MutualPLE block. When $i = 1$, F_e^{i-1} corresponds to F_{st} .

4.3 Losses

For turbulence mitigation, the loss function L_m is defined as:

$$L_m = L_{mit} + \beta_1 L_{re-deg}, \quad (8)$$

where L_{mit} measures the L1 distance between the restored images I_{out} and the ground truth I_{GT} , and L_{re-deg} denotes the re-degradation loss computed by applying a differentiable pipeline that sequentially simulates grayscale drift, blur, and tilt on I_{out} , followed by L1 comparison with the original input I . The weighting coefficient β_1 is set to 0 during the initial training stages and switched to 1 at the last several epochs. For physical fields estimation, the loss function L_e is defined as:

$$L_e = L_{est} + \beta_2 L_{phys}, \quad (9)$$

where β_2 is set to 0.5, L_{est} measures the L1 distance between the predicted physical fields P_{out} and the ground truth P_{GT} , and L_{phys} incorporates prior physical knowledge of turbulence to guide the network training. Specifically, Eq. (2) reveals the intrinsic relationship between C_n^2 and C_T^2 in atmospheric turbulence, highlighting how temperature fluctuations govern the refractive index structure. Beyond this, the kinetic energy dissipation rate ε can be inferred through C_n^2 [12] via:

$$\varepsilon = \gamma (C_n^2)^{3/2} (T/P)^3, \quad (10)$$

where $\gamma = 1.08 \times 10^{22}$. Similarly, the Reynolds stress, characterizing velocity fluctuation intensity, is formulated as the mean square of turbulent velocity and can be inferred from ε [23]:

$$RS \approx \langle w'^2 \rangle, \varepsilon \approx \langle w'^2 \rangle^{3/2} / L, \quad (11)$$

with L is the typical scale of atmospheric turbulence ($l_0 \ll L \ll L_0$). By constraining the relationships among these physical fields in Eqs. (2), (10) and (11), L_{turb} is formulated as:

$$\begin{aligned} L_{phys} = \mathbb{E}(\| \hat{C}_n^2 - \left(79 \times 10^{-6} P / \hat{T}^2 \right)^2 \hat{C}_T^2 \| \\ + \| \hat{\varepsilon} - \gamma (\hat{C}_n^2)^{\frac{3}{2}} (\hat{T}/P)^3 \| + \| \widehat{RS} - (\hat{\varepsilon} L)^{\frac{2}{3}} \|). \end{aligned} \quad (12)$$

where $\hat{C}_n^2$, $\hat{C}_T^2$, $\hat{\varepsilon}$, $\widehat{RS}$ are estimated fields. Intuitively, L_{phys} enforces consistency between turbulence fields, promoting physically plausible estimations.

Table 1: Quantitative evaluation of turbulence mitigation. Bold and underlined values indicate the best and second-best performance, respectively. All methods are evaluated with identical inputs of 256×256 spatial resolution and 30-frame sequences.

Turbulence Level	Weak		Medium		Strong		Overall	
Method	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
BasicVSR	29.2718	0.8695	25.8040	0.7474	24.1317	0.6714	26.4142	0.7632
ESTRNN	25.1907	0.7898	24.1390	0.7178	23.1995	0.6609	24.1810	0.7231
PBCL	28.3768	0.8502	<u>26.2520</u>	<u>0.7693</u>	<u>25.0672</u>	<u>0.7265</u>	26.5729	<u>0.7823</u>
ConvIR	26.9256	0.8007	23.9879	0.6641	22.8554	0.6016	24.5989	0.6893
TMT	29.3165	0.8697	26.0301	0.7498	24.7026	0.6825	26.6936	0.7678
DATUM	29.1793	0.8683	26.2379	0.7613	24.9530	0.6964	<u>26.7997</u>	0.7757
MambaTM	26.8179	0.7848	24.9059	0.6923	23.9590	0.6445	25.2341	0.7075
MaIR	<u>29.3921</u>	<u>0.8705</u>	25.5369	0.7443	24.0329	0.6675	26.3329	0.7612
DeepSN-Net	28.3611	0.8512	24.9958	0.7216	23.3838	0.6485	25.5916	0.7409
MHNet	28.9540	0.8587	25.5120	0.7251	23.9925	0.6503	26.1641	0.7452
TurboMPLE	29.8887	0.8790	27.1542	0.7918	25.6436	0.7274	27.5718	0.7998

5 Experiments

5.1 Training setting

For turbulence mitigation, in addition to representative turbulence mitigation networks [43, 50–52], we evaluated several restoration [7, 8, 11, 19], deblurring [55] and super-resolution networks [5] to provide a comprehensive benchmark. To train the proposed model, we employed the Adam optimizer [16] with an initial learning rate of $1e^{-4}$ and a batch size of 2. Inputs, consisting of 30 frames, were randomly cropped to 128×128 , and the training was conducted for 100 epochs. We trained MambaTM [51], DATUM [50], TMT [52], ESTRNN [55], and BasicVSR [5] with the same configuration as TurboMPLE. For MaIR [19], MHNet [11], ConvIR [7] and DeepSN-Net [8], which take single-frame inputs, the training procedure was otherwise identical.

For physical fields estimation, we compared TurboMPLE with FADE [22], FoVFormer [25], STA-DRN [32], MSU [54], UwU-Net [26], GRCNN [41], and UNet [34], all of which were trained using the same strategy as the proposed model. PBCL [43] was trained according to the original paper, including both separate and cooperative training.

5.2 Results on turbulence mitigation

Fig. 3 compares different methods on the synthetic dataset. Across both challenging cases ((a1) and (a2)), TurboMPLE achieves the most faithful restoration of



Fig. 3: Qualitative comparison of turbulence mitigation on synthetic data.

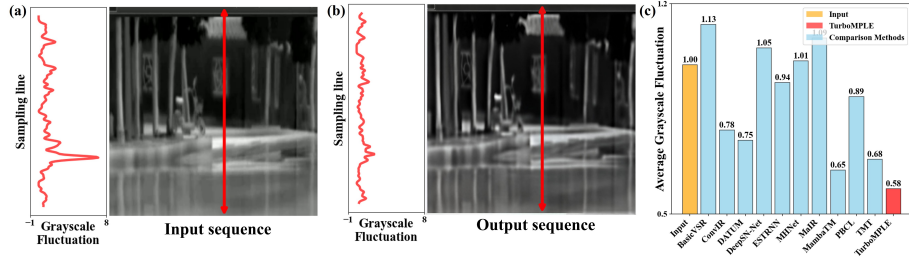


Fig. 4: Turbulence mitigation results on real-world data. (a) Degraded input sequence. (b) TurboMPLE output. (c) Comparison showing TurboMPLE achieves the most effective mitigation relative to competing methods.

fine details, especially in the magnified regions, effectively suppressing blur and geometric distortions induced by atmospheric turbulence, with PSNR exceeding the second-best method by over 1 dB. Quantitatively, as shown in Tab. 1, TurboMPLE attains the highest PSNR and SSIM, exceeding the second-best method by 0.77 dB and 0.0175, respectively, and maintaining superiority across all turbulence levels, demonstrating its robustness against severe atmospheric distortions. Moreover, with only 145 GMACs—over 50× more efficient than Ba-

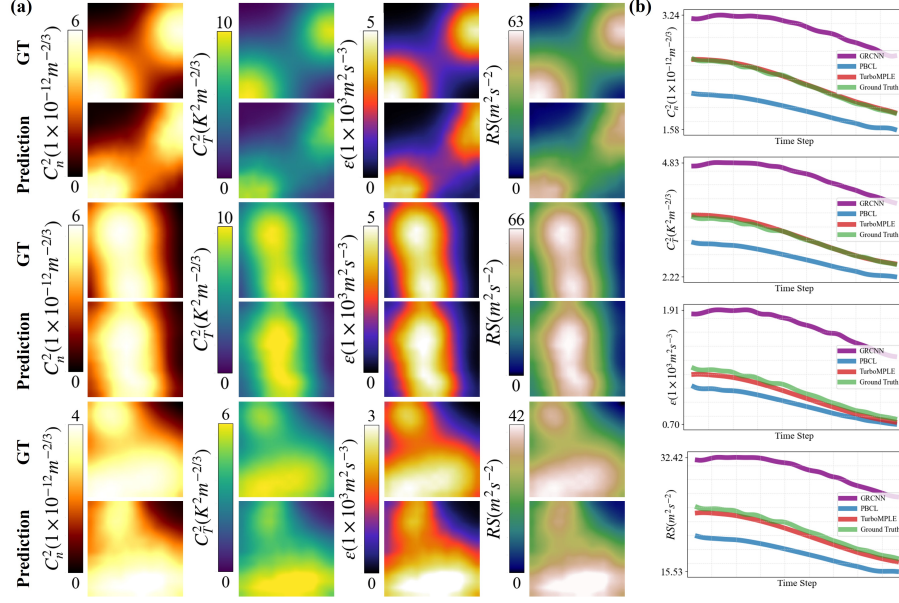


Fig. 5: (a) Turbulence estimation results on synthetic data. (b) Comparison of temporal dynamics of turbulence fields predicted by TurboMPLE, PBCL and GRCNN.

sicVSR (8355 GMACs)—TurboMPLE delivers both state-of-the-art quality and exceptional efficiency.

Fig. 4 demonstrates the effectiveness of TurboMPLE in mitigating spatio-temporal turbulence on real-world video data. The degraded input sequence, captured in a high-temperature environment, exhibits pronounced intensity fluctuations along the sampling line near the ground. The output of TurboMPLE substantially suppresses these fluctuations, restoring temporal and spatial consistency, and achieves the most effective turbulence mitigation, outperforming the second-best method by 20% in grayscale stabilization. Additional results are provided in the supplementary material. These results highlight the superior capability of TurboMPLE in preserving fine details and stabilizing intensity variations under challenging real-world turbulence.

5.3 Results on turbulence estimation

Fig. 5 presents three representative results of TurboMPLE on synthetic data. Rows 1–2 show two spatially separated turbulence regions, whereas Rows 3–4 depict spatially coupled ones. In both cases, the two regions exhibit comparable intensity levels, and TurboMPLE accurately reconstructs their spatial positions and magnitudes. Rows 5–6 illustrate two turbulence regions that differ markedly in both scale and intensity, yet TurboMPLE still delivers precise estimations. The average R^2 scores for these configurations are 0.917, 0.944, and 0.902, respec-

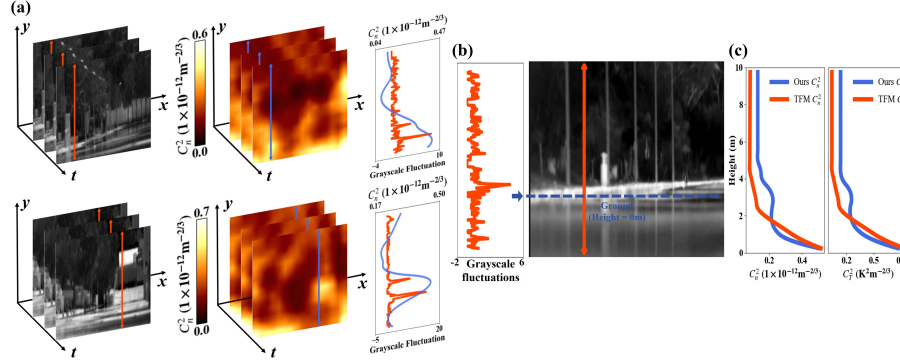


Fig. 6: Turbulence estimation results on real-world data.

tively, demonstrating the strong ability of TurboMPLE to estimate complex, spatially heterogeneous turbulence fields. Visualizations including comparison methods can be found in the supplementary material. Quantitative results in Tab. 2(a) further confirm the superiority of TurboMPLE, which achieves the lowest spatial MAE, 32.3% lower than the second-best method (PBCL), and an average spatial R^2 improvement of 0.054 across four physical fields.

Beyond spatial accuracy, we further evaluate TurboMPLE’s ability to capture temporal dynamics. As shown in Tab. 2(b), TurboMPLE achieves the best performance across all physical fields, reducing MAE by 49.2% and improving R^2 by 0.031 over PBCL. Fig. 5(b) illustrates the temporal dynamics of estimated fields, where TurboMPLE attains an average R^2 above 0.96 and an MAE that is only 16.2% and 6.2% of those of PBCL and GRCNN, respectively, highlighting its superior temporal fidelity and stability.

Fig. 6(a) presents the evaluation results obtained from real-world turbulence data. The first column displays input image sequences severely distorted by atmospheric turbulence while the second column shows the estimated dynamic C_n^2 fields. The third column compares the grayscale fluctuation profiles with the corresponding average C_n^2 distributions extracted along vertical sampling lines (orange for inputs and blue for estimations), revealing a pronounced spatial correspondence between them. Since grayscale fluctuations originate from light-angle deviations caused by refractive index inhomogeneities, the consistent alignment between these quantities provides strong physical evidence that the estimated physical fields faithfully capture the real turbulence characteristics present in the imaging process.

To further substantiate the effectiveness of our method on real-world turbulence data, we adopt the temperature fluctuation method (TFM) to obtain physically grounded, quantitative reference estimates of turbulence strength. Specifically, temperature measurements are acquired using thermocouple thermometers equipped with four probes, from which C_T^2 and C_n^2 are derived and treated as reference ground-truth. As shown in Fig. 6(b), the input image sequence ex-

Table 2: Quantitative evaluation of turbulence field estimation in both spatial (a) and temporal (b) metrics. Bold and underlined values indicate the best and second-best performance, respectively. All methods are evaluated with identical inputs of 256×256 spatial resolution and 30-frame sequences. The definition of evaluation metrics are detailed in the supplementary material.

(a) Spatial Evaluation												
Method	C_n^2			C_T^2			ε			RS		
	MAE	RMSE	R^2	r	MAE	RMSE	R^2	r	MAE	RMSE	R^2	r
UNet	0.076	0.102	0.604	0.893	0.071	0.097	0.564	0.890	0.053	0.076	0.523	0.870
GRCNN	0.060	0.074	0.771	0.888	0.054	0.068	0.763	0.883	0.039	0.053	0.712	0.853
UwUNet	0.118	0.142	0.124	0.386	0.105	0.129	0.109	0.387	0.072	0.095	<0.1	0.368
PBCL	0.036	0.049	<u>0.891</u>	<u>0.948</u>	0.034	<u>0.047</u>	<u>0.882</u>	<u>0.944</u>	0.024	0.037	<u>0.860</u>	<u>0.929</u>
MSU	0.137	0.158	<0.1	0.200	0.123	0.144	<0.1	0.196	0.091	0.111	<0.1	0.164
STA-DRN	0.057	0.079	0.748	0.869	0.053	0.075	0.732	0.860	0.039	0.058	0.687	0.832
FoVFormer	0.058	0.077	0.718	0.888	0.053	0.072	0.694	0.879	0.040	0.057	0.603	0.839
FADE	0.048	0.068	0.754	0.932	0.044	0.064	0.734	0.927	0.037	0.056	0.621	0.909
TurboMPLE	0.024	0.042	0.943	0.973	0.023	0.042	0.935	0.972	0.017	0.037	0.919	0.963

(b) Temporal Evaluation												
Method	C_n^2			C_T^2			ε			RS		
	MAE	RMSE	R^2	r	MAE	RMSE	R^2	r	MAE	RMSE	R^2	r
UNet	0.586	0.613	0.962	0.984	0.939	0.559	0.962	0.984	0.399	0.535	0.954	0.972
GRCNN	0.257	0.875	0.945	0.969	0.406	0.866	0.940	0.965	0.192	0.826	0.918	0.944
UwUNet	0.806	0.123	0.761	0.390	1.223	<0.1	0.753	0.386	0.521	<0.1	0.705	0.378
PBCL	0.162	<u>0.957</u>	<u>0.985</u>	<u>0.988</u>	0.259	<u>0.952</u>	<u>0.984</u>	<u>0.985</u>	0.105	<u>0.953</u>	<u>0.978</u>	<u>0.982</u>
MSU	0.919	<0.1	<0.1	0.130	1.400	<0.1	<0.1	0.124	0.607	<0.1	<0.1	0.145
STA-DRN	0.348	0.806	0.920	0.513	0.528	0.810	0.914	0.501	0.227	0.789	0.901	0.556
FoVFormer	0.342	0.790	0.984	0.983	0.548	0.765	0.982	0.981	0.259	0.668	0.974	0.969
FADE	0.324	0.778	0.968	0.987	0.498	0.759	0.967	<u>0.985</u>	0.254	0.646	0.962	0.977
TurboMPLE	0.072	0.990	0.997	0.997	0.139	0.983	0.997	0.997	0.059	0.981	0.995	0.995

hibits pronounced grayscale fluctuations near the ground (height = 0 m) along the sampling line, indicating strong turbulence. Fig. 6(c) presents a quantitative comparison between the TurboMPLE predictions and the TFM-derived reference profiles along the sampling line, yielding $r = 0.944$ for C_n^2 , and $r = 0.968$ for C_T^2 , respectively. Additional TFM validations under varying-altitude conditions are provided in the supplementary material. This quantitative evaluation provides direct evidence that the turbulence parameters predicted by TurboMPLE are physically meaningful and well aligned with real-world turbulence conditions.

Table 3: Ablation study. R_s^2 and R_t^2 denote R^2 in spatial and temporal evaluation of physical fields estimation, respectively.

Components	PSNR	SSIM	R_s^2	R_t^2
Baseline	26.470	0.781	0.890	0.967
+ Dual-branch Experts	26.726	0.786	0.900	0.974
+ CTMB	27.113	0.787	0.909	0.983
+ BLB embedding	27.452	0.797	0.916	0.972
+ L_{phys}	27.572	0.800	0.934	0.985

5.4 Ablation study

Influence of key components in TurboMPLE design and training. For the baseline in Tab. 3, the Dual-branch Experts are substituted with residual dense blocks [53] to isolate the contributions, whereas the other components are excluded directly. Adding each module consistently enhances both tasks, with the complete model achieving improvements of 1.102 dB in PSNR and 0.019 in SSIM for turbulence mitigation, and R_s^2 and R_t^2 gains of 0.044 and 0.018 respectively for physical fields prediction, over the baseline.

Influence of input frame number. We trained TurboMPLE with 10, 15, 20, 25, and 30 input frames. As shown in Fig. 7, mitigation performance improves with more input frames, reflecting the zero-mean stochastic nature of turbulent

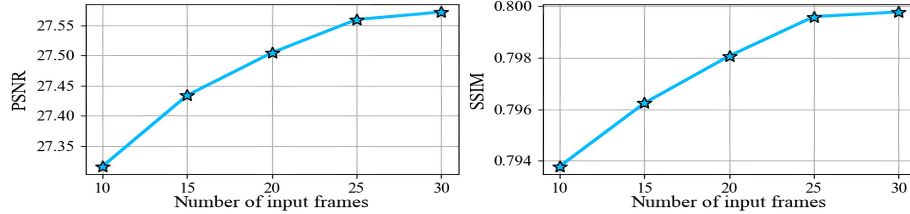


Fig. 7: Influence of input frame number on turbulence mitigation.

fluctuations where additional frames enable better estimation of the underlying clear state. Performance gain saturates beyond 25 frames, therefore a 30-frame configuration is adopted in this study. Also, this demonstrates that TurboMPLE maintains stable performance across varying frame counts, yielding only a marginal PSNR gain of 0.256 dB when increasing the input from 10 to 30 frames.

Influence of MutualPLE block number. We evaluated TurboMPLE with various numbers of MutualPLE blocks ($N = 1, 2, 3$). As shown in Tab. 4, increasing from 1 to 2 substantially improves performance, with PSNR and SSIM gains of 1.09 dB and 0.017, and corresponding R_s^2 and R_t^2 improvements of 0.038 and 0.019. Adding a third block yields only marginal PSNR gains while increasing computational cost, motivating the adoption of $N = 2$ in this study.

Table 4: Influence of MutualPLE block number.

N	PSNR	SSIM	R_s^2	R_t^2	GMACs	Params
1	26.482	0.783	0.896	0.966	96.3	0.645
2	27.572	0.800	0.934	0.985	145.1	0.942
3	27.477	0.795	0.936	0.981	193.8	1.239

6 Conclusion

In this study, we constructed a large-scale infrared turbulence imaging dataset and introduced TurboMPLE, a joint network for turbulence mitigation and physical fields estimation. By feature-level mutual design, TurboMPLE introduces turbulence priors into mitigation process while incorporating richer visual information into physical fields estimation. Physical embedding and constraints further enable interpretable results. Extensive experiments show that TurboMPLE delivers state-of-the-art performance in both tasks with a lightweight architecture and computationally efficient design, providing a robust solution for real-world restoration and turbulence-aware physical fields modeling.

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China under grant 62271016 and Beijing Natural Science Foundation under grant L258070.

References

1. Aubailly, M., Vorontsov, M.A., Carhart, G.W., Valley, M.T.: Automated video enhancement from a stream of atmospherically-distorted images: the lucky-region

- fusion approach. In: *Atmospheric Optics: Models, Measurements, and Target-in-the-Loop Propagation III*. vol. 7463, pp. 104–113. SPIE (2009)
2. Bhattacharjee, D., Zhang, T., Süssstrunk, S., Salzmann, M.: Mult: An end-to-end multitask learning transformer. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 12031–12041 (2022)
 3. Bose-Pillai, S.R., McCrae, J.E., Rice, C.A., Wood, R.A., Murphy, C.E., Fiorino, S.T.: Estimation of atmospheric turbulence using differential motion of extended features in time-lapse imagery. *Optical Engineering* **57**(10), 104108–104108 (2018)
 4. Caruana, R.: Multitask learning. *Machine learning* **28**(1), 41–75 (1997)
 5. Chan, K.C., Wang, X., Yu, K., Dong, C., Loy, C.C.: Basicvsr: The search for essential components in video super-resolution and beyond. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 4947–4956 (2021)
 6. Chen, Z., Tian, Y., Lü, D.: Turbulence parameters in the troposphere—lower stratosphere observed by beijing mst radar. *Remote Sensing* **14**(4) (2022). <https://doi.org/10.3390/rs14040947>, <https://www.mdpi.com/2072-4292/14/4/947>
 7. Cui, Y., Ren, W., Cao, X., Knoll, A.: Revitalizing convolutional network for image restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **46**(12), 9423–9438 (2024)
 8. Deng, X., Zhang, C., Jiang, L., Xia, J., Xu, M.: Deepsn-net: Deep semi-smooth newton driven network for blind image restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2025)
 9. Diskin, Y., Whiteley, M., Grose, M., Jackovitz, K., Drye, R., Hampshire, B., Owens, M., Smith, E., Magee, E., Kalensky, M., Jumper, E., Gordeyev, S., Archibald, A.: Aircraft to ground profiling: Turbulence measurements and optical system performance modeling. *AIAA Journal* **59**(11), 4610–4625 (2021). <https://doi.org/10.2514/1.J060580>, <https://doi.org/10.2514/1.J060580>
 10. Fried, D.L.: Probability of getting a lucky short-exposure image through turbulence. *Journal of the Optical Society of America* **68**(12), 1651–1658 (1978)
 11. Gao, H., Zhang, Y., Yang, J., Dang, D.: Mixed hierarchy network for image restoration. *Pattern Recognition* **161**, 111313 (2025)
 12. Hocking, W.K.: Measurement of turbulent energy dissipation rates in the middle atmosphere by radar techniques: A review. *Radio Science* **20**(6), 1403–1422 (1985)
 13. Jaiswal, A., Phanikumar, D.V., Bhattacharjee, S., Naja, M.: Estimation of turbulence parameters using ariess radar and gps radiosonde measurements: First results from the central himalayan region. *Radio Science* **55**(8), e2019RS006979 (2020). <https://doi.org/https://doi.org/10.1029/2019RS006979>, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2019RS006979>, e2019RS006979 2019RS006979
 14. Jaiswal, A., Zhang, X., Chan, S.H., Wang, Z.: Physics-driven turbulence image restoration with stochastic refinement. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. pp. 12170–12181 (2023)
 15. Jin, D., Chen, Y., Lu, Y., Chen, J., Wang, P., Liu, Z., Guo, S., Bai, X.: Neutralizing the impact of atmospheric turbulence on complex scene imaging via deep learning. *Nature Machine Intelligence* **3**(10), 876–884 (2021)
 16. Kinga, D., Adam, J.B., et al.: A method for stochastic optimization. In: *International Conference on Learning Representations (ICLR)*. vol. 5. California; (2015)
 17. Kohma, M., Sato, K., Tomikawa, Y., Nishimura, K., Sato, T.: Estimate of turbulent energy dissipation rate from the vhf radar and radiosonde observations

- in the antarctic. *Journal of Geophysical Research: Atmospheres* **124**(6), 2976–2993 (2019). <https://doi.org/https://doi.org/10.1029/2018JD029521>, <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018JD029521>
18. Lao, D., Wang, C., Wong, A., Soatto, S.: Diffeomorphic template registration for atmospheric turbulence mitigation. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 25107–25116 (2024)
 19. Li, B., Zhao, H., Wang, W., Hu, P., Gou, Y., Peng, X.: Mair: A locality-and continuity-preserving mamba for image restoration. In: *Proceedings of the Computer Vision and Pattern Recognition Conference*. pp. 7491–7501 (2025)
 20. Liu, Y., Lu, Y., Liu, H., An, Y., Xu, Z., Yao, Z., Zhang, B., Xiong, Z., Gui, C.: Hierarchical prompt learning for multi-task learning. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 10888–10898 (2023)
 21. Liu, Y., Tian, Y., Zhao, Y., Yu, H., Xie, L., Wang, Y., Ye, Q., Jiao, J., Liu, Y.: Vmamba: Visual state space model. *Advances in Neural Information Processing Systems* **37**, 103031–103063 (2024)
 22. Lu, H., Liu, W., Fu, H., Cao, Z.: Fade: A task-agnostic upsampling operator for encoder-decoder architectures. *International Journal of Computer Vision* **133**(1), 151–172 (2025). <https://doi.org/10.1007/s11263-024-02191-8>
 23. Luce, H., Kantha, L., Hashiguchi, H., Lawrence, D., Doddi, A.: Turbulence kinetic energy dissipation rates estimated from concurrent uav and mu radar measurements. *Earth, Planets and Space* **70**(1), 207 (2018)
 24. Ma, J., Zhao, Z., Yi, X., Chen, J., Hong, L., Chi, E.H.: Modeling task relationships in multi-task learning with multi-gate mixture-of-experts. In: *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. pp. 1930–1939 (2018)
 25. Ma, X., Wang, L., Wang, X., Song, W., Sun, Q., Zhu, L., Huang, H.: Learning refractive-diffractive optics with unidirectional transformer for large field-of-view imaging. *International Journal of Computer Vision* pp. 1–21 (2025). <https://doi.org/10.1007/s11263-025-02546-9>
 26. Manifold, B., Men, S., Hu, R., Fu, D.: A versatile deep learning architecture for classification and label-free prediction of hyperspectral images. *Nature Machine Intelligence* **3**(4), 306–315 (2021). <https://doi.org/10.1038/s42256-021-00309-y>
 27. Mao, J., Zhang, Y., Li, J., Gong, X., Zhao, H., Rao, Z.: Novel detection of atmospheric turbulence profile using mie-scattering lidar based on non-kolmogorov turbulence theory. *Entropy* **25**(3), 477 (2023)
 28. Mao, Z., Chimitt, N., Chan, S.H.: Image reconstruction of static and dynamic scenes through anisoplanatic turbulence. *IEEE Transactions on Computational Imaging* **6**, 1415–1428 (2020)
 29. Mao, Z., Jaiswal, A., Wang, Z., Chan, S.H.: Single frame atmospheric turbulence mitigation: A benchmark study and a new physics-inspired transformer model. In: *European Conference on Computer Vision*. pp. 430–446. Springer (2022)
 30. Mei, K., Patel, V.M.: Ltt-gan: Looking through turbulence by inverting gans. *IEEE Journal of Selected Topics in Signal Processing* **17**(3), 587–598 (2023)
 31. Oreifej, O., Li, X., Shah, M.: Simultaneous video stabilization and moving object detection in turbulence. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **35**(2), 450–462 (2012)
 32. Pan, Y., Shang, Y., Liu, T., Shao, Z., Guo, G., Ding, H., Hu, Q.: Spatial-temporal attention network for depression recognition from facial videos. *Expert Systems*

- with Applications **237**, 121410 (2024). <https://doi.org/10.1016/j.eswa.2023.121410>
33. Rai, S.N., Jawahar, C.: Removing atmospheric turbulence via deep adversarial learning. *IEEE Transactions on Image Processing* **31**, 2633–2646 (2022)
 34. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: *Medical image computing and computer-assisted intervention—MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, proceedings, part III* 18. pp. 234–241. Springer (2015)
 35. Saha, R.K., Qin, D., Li, N., Ye, J., Jayasuriya, S.: Turb-seg-res: A segment-then-restore pipeline for dynamic videos with atmospheric turbulence. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 25286–25296 (2024)
 36. Saha, R.K., Salcin, E., Kim, J., Smith, J., Jayasuriya, S.: Turbulence strength c n 2 estimation from video using physics-based deep learning. *Optics Express* **30**(22), 40854–40870 (2022)
 37. Shimizu, M., Yoshimura, S., Tanaka, M., Okutomi, M.: Super-resolution from image sequence under influence of hot-air optical turbulence. In: *2008 IEEE Conference on Computer Vision and Pattern Recognition*. pp. 1–8. IEEE (2008)
 38. Tan, C., Gao, Z., Wu, L., Xu, Y., Xia, J., Li, S., Li, S.Z.: Temporal attention unit: Towards efficient spatiotemporal predictive learning. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 18770–18782 (2023)
 39. Tang, H., Liu, J., Zhao, M., Gong, X.: Progressive layered extraction (ple): A novel multi-task learning (mtl) model for personalized recommendations. In: *Proceedings of the 14th ACM Conference on Recommender Systems*. pp. 269–278 (2020)
 40. Vollmer, M., Möllmann, K.P.: *Infrared thermal imaging: fundamentals, research and applications*. John Wiley & Sons, Weinheim (2018)
 41. Wang, J., Hu, X.: Convolutional neural networks with gated recurrent connections. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **44**(7), 3421–3435 (2021). <https://doi.org/10.1109/TPAMI.2021.3054614>
 42. Wang, X., López-Tapia, S., Katsaggelos, A.K.: Real-world atmospheric turbulence correction via domain adaptation. In: *2024 IEEE International Conference on Image Processing (ICIP)*. pp. 1466–1472. IEEE (2024)
 43. Wang, Y., Jin, D., Chen, J., Bai, X.: Revelation of hidden 2d atmospheric turbulence strength fields from turbulence effects in infrared imaging. *Nature Computational Science* **3**(8), 687–699 (2023)
 44. Wu, S., Hu, X., Han, Y., Wu, X., Su, C., Luo, T., Li, X.: Measurement and analysis of atmospheric optical turbulence in lhasa based on thermosonde. *Journal of Atmospheric and Solar-Terrestrial Physics* **201**, 105241 (2020). <https://doi.org/https://doi.org/10.1016/j.jastp.2020.105241>, <https://www.sciencedirect.com/science/article/pii/S1364682620300584>
 45. Xia, Y., Zhou, C., Zhu, C., Teng, M., Xu, C., Shi, B.: Nb-gtr: narrow-band guided turbulence removal. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 24934–24943 (2024)
 46. Xie, Y., Zhang, W., Tao, D., Hu, W., Qu, Y., Wang, H.: Removing turbulence effect via hybrid total variation and deformation-guided kernel regression. *IEEE Transactions on Image Processing* **25**(10), 4943–4958 (2016)
 47. Xu, S., Sun, R., Chang, Y., Cao, S., Xiao, X., Yan, L.: Long-range turbulence mitigation: a large-scale dataset and a coarse-to-fine framework. In: *European Conference on Computer Vision*. pp. 311–329. Springer (2024)

48. Yue, Y., Li, Z.: Medmamba: Vision mamba for medical image classification. arXiv preprint arXiv:2403.03849 (2024)
49. Zhang, W., Li, X., Shi, G., Chen, X., Qiao, Y., Zhang, X., Wu, X.M., Dong, C.: Real-world image super-resolution as multi-task learning. *Advances in Neural Information Processing Systems* **36**, 21003–21022 (2023)
50. Zhang, X., Chimitt, N., Chi, Y., Mao, Z., Chan, S.H.: Spatio-temporal turbulence mitigation: A translational perspective. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 2889–2899 (2024)
51. Zhang, X., Chimitt, N., Wang, X., Yuan, Y., Chan, S.H.: Learning phase distortion with selective state space models for video turbulence mitigation. In: *Proceedings of the Computer Vision and Pattern Recognition Conference*. pp. 2127–2138 (2025)
52. Zhang, X., Mao, Z., Chimitt, N., Chan, S.H.: Imaging through the atmosphere using turbulence mitigation transformer. *IEEE Transactions on Computational Imaging* **10**, 115–128 (2024)
53. Zhang, Y., Tian, Y., Kong, Y., Zhong, B., Fu, Y.: Residual dense network for image super-resolution. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. pp. 2472–2481 (2018)
54. Zhang, Z., Gao, S., Zhou, M., Yan, M., Cao, S.: Mapping network-coordinated stacked gated recurrent units for turbulence prediction. *IEEE/CAA Journal of Automatica Sinica* **11**(6), 1331–1341 (2024). <https://doi.org/10.1109/JAS.2024.124335>
55. Zhong, Z., Gao, Y., Zheng, Y., Zheng, B., Sato, I.: Real-world video deblurring: A benchmark dataset and an efficient recurrent neural network. *International Journal of Computer Vision* **131**(1), 284–301 (2023)
56. Zhou, H., Su, X., Duan, Y., Song, H., Zou, K., Zhang, R., Song, H., Hu, N., Tur, M., Willner, A.E.: Atmospheric turbulence strength distribution along a propagation path probed by longitudinally structured optical beams. *Nature Communications* **14**(1), 4701 (2023)
57. Zhu, X., Milanfar, P.: Removing atmospheric turbulence via space-invariant deconvolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **35**(1), 157–170 (2012)