Supplemental Material for MetaVIn: Meteorological and Visual Integration for Atmospheric Turbulence Strength Estimation

Ripon Kumar Saha
Arizona State University
rsaha8@asu.edu

Scott McCloskey
Kitware, Inc.
scott.mccloskey@kitware.com

Suren Jayasuriya
Arizona State University
sjayasuri@asu.edu

1. Blind IQA Correlation Plots

The main paper presents a table for the quantitative comparison of Blind Image Quality Analysis (BIQA) based C_n^2 estimation, evaluated on the BTS3 and BTS4 datasets. We further illustrate these BIQA methods' correlations through scatter plots in Figure 1. PaQ-2-PiQ [4] achieves the highest performance with a Spearman correlation coefficient (ρ) of 0.14, followed by MUSIQ [2] with ρ = 0.09. Several methods show negative correlations, including ILNIQE (ρ = -0.15), DBCNN (ρ = -0.13), and CLIP-IQA [3] (ρ = -0.09). These results demonstrate that existing BIQA methods generally fail to establish strong correlations with atmospheric turbulence strength.

It's worth noting that most of these methods are designed to perform overall image quality analysis, including image aesthetics, and some consider human preferences as well. Therefore, they might not particularly show strength in correlation with physics-based atmospheric turbulence strength or atmospheric image degradation estimation, especially in the presence of camera motion.

2. Meteorological Details

In the BRIAR dataset [1], meteorological data was captured alongside long-range imagery for biometrics tasks. The collected meteorological parameters include:

- **Temperature** (°C): Ambient temperature, measured in degrees Celsius, capturing the overall thermal conditions of the environment.

- **Wind Chill** (°C): The perceived temperature accounting for wind speed, indicating how cold it feels due to wind.

- **Heat Index** (°C): The perceived temperature considering both humidity and temperature, indicating how hot it feels.

- **Dew Point** (°C): The temperature at which air becomes saturated with moisture, leading to dew formation.
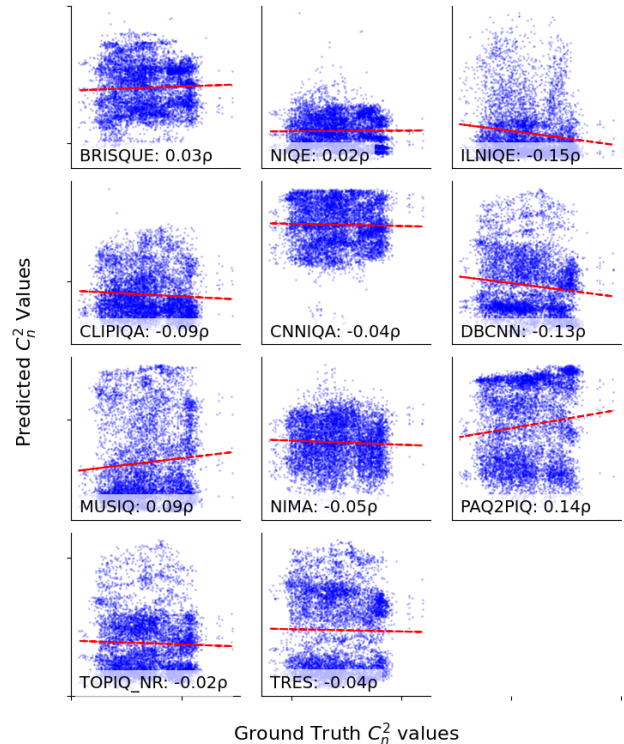

Figure 1. Correlation between various image quality assessment (IQA) scores and C_n^2 values (log scale) associated with of spearman's rank correlation coefficient (ρ) is displayed. Traditional (top 1 row) and deep learning-based BIQA methods (2nd, 3rd and 4th row) show weak correlation with turbulence strength, indicating their ineffectiveness in this application.

- **Relative Humidity (%)**: The percentage of moisture in the air relative to its maximum capacity at a given temperature.

- **Wind Speed (m/s)**: The instantaneous wind speed, measured in meters per second.

- **Wind Direction** (°): The instantaneous wind direction in degrees from true north.

- **Barometric Pressure (hPa)**: The atmospheric pressure at the surface, measured in hectopascals.

- **Solar Loading (W/m²)**: The amount of solar radiation energy received per unit area, measured in watts per square meter.

- **Turbulence Strength** (C_n^2) **(m$^{-2/3}$)**: The refractive index structure parameter, representing atmospheric turbulence strength (shown in log10 scale).

3. Effect of Distance on Image degradation:

In this section, we investigate the impact of the distance from the camera to the scene on image degradation. We analyze how distance affects various image sharpness measures. Specifically, we compute the sharpness, tenengrad, and variance of gradient metrics from the first frame of all videos. These metrics are crucial for understanding the impact of image quality on atmospheric turbulence strength prediction. The BRIAR dataset [1] provides the distances from the camera to the person, facilitating this analysis.

Table 1. Sharpness Metrics for Different Distances

Distance	Sharpness	Tenengrad	Var. of Grad.
100	2628 ± 3114	89 ± 40	226 ± 193
180	638 ± 621	46 ± 30	72 ± 68
200	263 ± 275	36 ± 23	46 ± 47
300	617 ± 381	63 ± 35	94 ± 78
330	289 ± 272	34 ± 22	52 ± 49
370	523 ± 325	63 ± 29	99 ± 80
400	157 ± 195	24 ± 16	25 ± 29
490	138 ± 150	26 ± 14	23 ± 23
500	329 ± 513	34 ± 32	54 ± 83
600	217 ± 337	32 ± 25	43 ± 60
720	93 ± 130	18 ± 13	11 ± 14
800	8 ± 4	9 ± 3	1 ± 1
1000	29 ± 8	12 ± 5	6 ± 3

Table 1 presents the calculated sharpness, tenengrad, and variance of gradient metrics for images taken at various distances. As the distance increases, sharpness values decrease significantly, with notable drops at farther distances. For instance, the sharpness value at 100m (2628 ± 3114) is substantially higher than at 800m (8 ± 4) and 1000m (29 ± 8). The tenengrad metric shows a similar trend, with the mean value dropping as distance increases, indicating reduced edge sharpness. The variance of gradient also decreases with distance, suggesting less variation in image gradients at greater distances. These metrics highlight the impact of distance on image quality, with greater distances leading to lower sharpness, reduced edge detail, and decreased gradient variance.
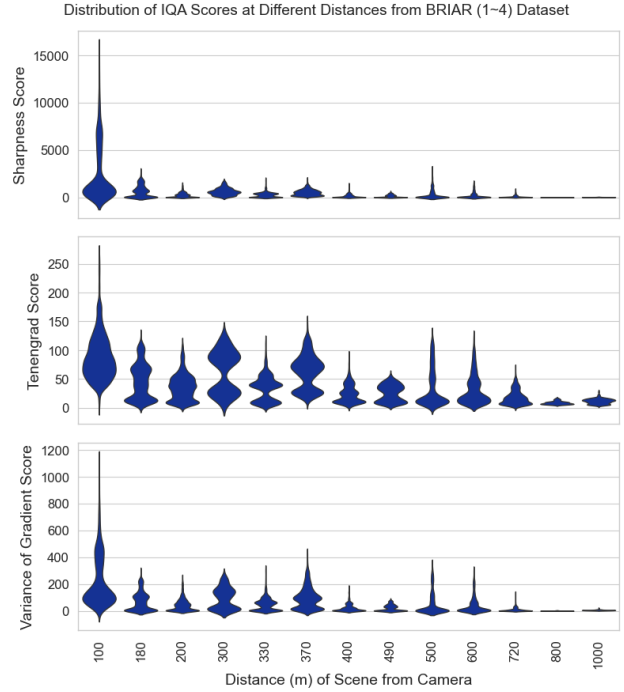

Figure 2. The impact of distance on image quality is evident in the figure, where increasing distance correlates with a decline in sharpness, tenengrad, and variance of gradient metrics. This downward trend signifies a deterioration in image quality, with farther distances resulting in reduced sharpness, less defined edges, and decreased gradient variation.

Figure 2 visualizes these values at different distances on a logarithmic scale. The results show that as the distance increases, sharpness, tenengrad, and variance of gradient values decrease, indicating increased image degradation. This observation aligns with the common understanding that objects appear clearer and sharper when imaged from a closer distance compared to a zoomed view from further away. Furthermore, distance must be factored into almost all systems, including large aperture scintillometer-based C_n^2 estimation and image gradient-based C_n^2 estimation. Since C_n^2 provides a path-averaged turbulence strength, incorporating distance is essential for accurate C_n^2 estimation.

This demonstrates that distance is a significant factor and is effectively utilized in MetaVIn. However, computing distance manually or using a range finder adds complexity and device dependency. Therefore, we used a state-of-the-art monocular depth estimation model to calculate the pseudo-depth of the scene, which was then used in our approach.

4. Ablation study on KAN Architecture

This study investigates the impact of varying architecture choices and hyperparameters on the performance of a 3-layer Knowledge-Aware Network (KAN) model. We

conducted a series of ablation studies to examine the effects of different KAN architectures, layer sizes, and grid sizes on the model's performance. Our results show that the choice of KAN architecture significantly affects performance, with the (13, 8, 1) architecture achieving the highest Spearman correlation and lowest errors. Additionally, we found that increasing the layer size improves performance up to a certain point, and that a grid size of 5 strikes the best balance between spatial resolution and computational efficiency. These findings provide valuable insights into the optimal design of KAN models for achieving good results. For this ablation study we have used learning rate of 0.9, and train them for 50 epochs using LBFGS optimizer and L1 loss function.

Details of the experiments and results are presented in the following sections. Our findings indicate that the model's performance remains stable and high-performing, with minimal effects resulting from modifications to the network configuration.

4.1. Varying Architecture Choices

Table 2 presents a comparison of different KAN architectures with their corresponding performance metrics. The architecture (13, 8, 1) achieves the highest Spearman correlation (ρ) of 0.942, along with a low MAE of 0.179 and relative error of 0.006. Other architectures, such as (13, 4, 2, 1) and (13, 8, 2, 4, 1), also demonstrate promising results with ρ above 0.9 and relatively low errors. The architecture (13, 8, 2, 2, 1) notably underperforms, exhibiting a negative Spearman correlation and substantially higher error values. These results indicate that KAN architecture choice significantly impacts performance, making optimal architecture selection crucial for achieving superior results.

Table 2. Architecture Comparison

Architecture	Sp. Corr.	MAE	Relative Error
(13, 8, 1)	**0.942**	**0.179**	**0.006**
(13, 4, 2, 1)	0.929	0.204	0.007
(13, 4, 4, 1)	0.917	0.215	0.007
(13, 8, 8, 1)	0.940	0.187	0.006
(13, 8, 2, 2, 1)	-0.054	13.769	1.000
(13, 8, 2, 4, 1)	0.930	0.197	0.007
(13, 8, 4, 2, 1)	0.932	0.194	0.007
(13, 8, 4, 4, 1)	0.924	0.206	0.007
(13, 8, 8, 8, 1)	0.938	0.182	0.006

4.2. Layer Size Optimization

We investigated the impact of varying layer sizes on performance in the [13, Layer, 1] Kan architecture while maintaining a grid size of 5. The results, presented in Table 3, show that increasing the layer size improves performance,

with the highest Spearman correlation (0.943) and lowest Mean Absolute Error (0.179) and Relative Error (0.006) achieved at a layer size of 8. However, increasing the layer size beyond 8 leads to a slight decrease in performance. For a layer size of 3, the network exhibited instability and failed to converge, resulting in significantly elevated MAE and poor correlation values.

Table 3. KAN Layer Size Ablation Study Results.

Layer	Spearman.Corr	MAE	Relative Error
2	0.916	0.214	0.008
3	0.386	13.770	1.000
4	0.918	0.218	0.008
5	0.929	0.196	0.007
6	0.940	0.185	0.007
7	0.934	0.196	0.007
8	**0.943**	**0.179**	**0.006**
9	0.936	0.195	0.007
10	0.939	0.185	0.007
11	0.935	0.197	0.007
12	0.931	0.199	0.007

4.3. Grid Size Optimization

We conducted an ablation study to investigate the effect of varying grid sizes on the performance of our model. While keeping all other hyperparameters constant, including the layer architecture ([13, 8, 1]) and learning rate (0.9), we systematically varied the grid size from 2 to 8. The results, presented in Table 4, reveal that the optimal grid size is 5, achieving the highest Spearman correlation (0.943) and lowest Mean Absolute Error (0.179) and Relative Error (0.006). These findings suggest that a grid size of 5 strikes the best balance between spatial resolution and computational efficiency, leading to improved accuracy and generalization capabilities.

Table 4. Grid Size Ablation Study Results

Grid	Spearman.Corr	MAE	Relative Error
2	0.941	0.188	0.007
3	0.933	0.198	0.007
4	0.927	0.202	0.007
5	**0.943**	**0.179**	**0.006**
6	0.933	0.195	0.007
7	0.937	0.184	0.007
8	0.935	0.194	0.007

Scene	Snowfall	Low Contrast	Motion Blur
GT C_n^2	2.51×10^{-13}	7.58×10^{-14}	4.67×10^{-14}
Ours.	3.02×10^{-13}	7.58×10^{-14}	4.07×10^{-14}

Scene	Overexposed (top)	Haze	Perfect illumination
GT C_n^2	2.23×10^{-13}	1.81×10^{-13}	4.07×10^{-14}
Ours.	2.34×10^{-13}	1.54×10^{-13}	5.12×10^{-14}

Figure 3. MetaVIn Prediction under different situations including snowfall, haze, poor illumination, and overexposure.

5. MetaVIn in Different Environmental Situations

The BRIAR dataset [1] includes images captured under a wide range of environmental conditions, such as snowfall, haze, motion blur, overexposure, and varying directional lighting. These diverse conditions pose significant challenges for purely image-based models to accurately predict the degradation caused by atmospheric turbulence. However, our MetaVIn approach, which integrates visual and meteorological data, demonstrates robustness and provides close predictions that generalize well across various environmental conditions.

Figure 3 illustrates MetaVIn's performance under different situations. For instance, under snowfall, the ground truth C_n^2 value is 2.51×10^{-13}, while MetaVIn predicts 3.02×10^{-13}. In low contrast scenarios, both the ground truth and MetaVIn predictions align at 7.58×10^{-14}. Despite the challenges posed by motion blur, overexposure, and haze, MetaVIn maintains a high degree of accuracy, showcasing its effectiveness in diverse lighting and environmental conditions. This capability ensures reliable C_n^2 estimation across a broad spectrum of real-world scenarios, underscoring the practical utility of our approach.

Figure 4 provides a comprehensive comparison of Ground Truth and MetaVIn predictions of C_n^2 across a spectrum of values from very low to high, observed from a particular viewpoint. The rightmost image in the third row illustrates the impact of extreme turbulence on image quality. These samples, sourced from the BRIAR BGC4 dataset, depict various atmospheric turbulence conditions recorded on

(Ground Truth C_n^2, Predicted C_n^2 from MetaVIn)

Figure 4. First three rows show a comparison of Ground Truth and MetaVIn Predictions of C_n^2 from low to high C_n^2 values captured on different days and using different camera systems. Values are in the format (GT C_n^2, Our C_n^2). The bottom row shows some failure cases with higher error.

different days using different camera systems. The values are displayed in the format (GT C_n^2, Our C_n^2). The first three rows demonstrate the robustness of our approach under various conditions, including poor lighting, different camera systems, snowy scenes, and motion blur, with a close alignment between predicted and actual C_n^2 values. However, the bottom row highlights instances where our method struggled, particularly in extreme low-light conditions and images severely affected by blurring and complex motion. These examples underscore the challenges and potential areas for improvement in enhancing the model's robustness.

6. Data Availability and Reproducibility

The BRIAR dataset used in this paper was made available by the authors of [1]. As noted in their paper for those looking to obtain the dataset, "Interested parties should con-

tact the authors, who will forward the request to the appropriate US government representatives". Therefore, we encourage readers who wish to reproduce the results on the BRIAR data to follow these instructions. We plan to make the code for the KAN publicly available after peer review to be used in conjunction with the dataset.

References

[1] David Cornett, Joel Brogan, Nell Barber, Deniz Aykac, Seth Baird, Nicholas Burchfield, Carl Dukes, Andrew Duncan, Regina Ferrell, Jim Goddard, et al. Expanding accurate person recognition to new altitudes and ranges: The briar dataset. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pages 593–602, 2023. 1, 2, 4

[2] Junjie Ke, Qifei Wang, Yilin Wang, Peyman Milanfar, and Feng Yang. Musiq: Multi-scale image quality transformer. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 5148–5157, 2021. 1

[3] Jianyi Wang, Kelvin CK Chan, and Chen Change Loy. Exploring clip for assessing the look and feel of images. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 37, pages 2555–2563, 2023. 1

[4] Zhenqiang Ying, Haoran Niu, Praful Gupta, Dhruv Mahajan, Deepti Ghadiyaram, and Alan Bovik. From patches to pictures (paq-2-piq): Mapping the perceptual space of picture quality. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 3575–3585, 2020. 1