Development of radiation-tolerant beam imaging via multimode fiber and synthetic data-driven machine learning

Q. Xu,[1,2,3,*] H. D. Zhang,[1,2,†] G. Trad,[3] A. Hill,[1,2] F. Roncarolo,[3] and C. P. Welsch[1,2]

[1]*Cockcroft Institute, Warrington, WA4 4AD, United Kingdom*
[2]*University of Liverpool, Liverpool, L69 7ZE, United Kingdom*
[3]*CERN, 1211 Geneva 23, Switzerland*

(Received 22 August 2025; accepted 11 February 2026; published 9 March 2026)

Transverse beam profile monitoring in high-radiation areas is challenging due to camera degradation. A proposed solution employs a multimode fiber (MMF) to relay optical signals from the radiation zone to a shielded area, where a standard CMOS camera can operate safely. However, MMF transmission introduces significant distortions, producing complex speckle patterns at the output. We present a machine-learning–based method to reconstruct the transverse beam distributions from these patterns. The model was trained solely on synthetic data generated via stochastic Gaussian mixture simulations, in which the samples were displayed on a laser-illuminated digital micromirror device (DMD) and relayed through a 5-m MMF. The same setup was used for testing, with beam images on a scintillating screen from CERN's CLEAR facility replayed on the DMD. The model achieved a 1.49% mean relative root mean square error across four key beam parameters—77% lower than the 6.59% baseline, demonstrating its potential for diagnostics in radiation-constrained environments.

DOI: 10.1103/wddb-m37l

I. INTRODUCTION

Beam imaging systems are essential in particle accelerators for measuring transverse profiles, enabling beam alignment, emittance evaluation, and ensuring overall machine protection [1]. These systems typically use cameras or other optical detectors to capture visible light emitted when the beam interacts with interceptive screens via scintillation or optical transition radiation effects, which is observed through a viewport on the beam pipe [2]. Some of these diagnostics operate in radiation-exposed areas, such as beam dumps or fixed-target experiments, where radiation exposure significantly affects the durability and performance of cameras and their communication systems [3]. This challenge has become more critical following the discontinuation of radiation-hardened VIDICON tubes used at CERN for such applications. Consequently, there is a growing need for alternative imaging technologies capable of operating reliably in high-radiation environments.

At CERN, several approaches have been proposed and tested to address beam instrumentation challenges in these environments, including the use of single-mode fiber bundles to relay light signals to low-radiation areas, optical relay lines using mirrors to guide light away from high-radiation zones, and the development of radiation-tolerant cameras [4–6]. While these methods have been validated in principle, practical implementation still faces challenges [7]. Additionally, to quantify radiation effects on cameras, an irradiation test was conducted at CHARM using a radiation-tolerant camera, where a sharp drop in visibility was observed when the total integrated dose (TID) exceeded 250 Gy. However, instrumentation in some areas is expected to experience TID levels exceeding 400 Gy [8], indicating the need for more radiation-hardened solutions.

In this work, we propose an approach that uses a single large-core multimode fiber (MMF) to relay the screen light signal to a low-radiation area, allowing safe operation of standard cameras. MMF was selected for its simplicity, modularity, ease of installation and replacement, high mode capacity for preserving spatial information, and its documented radiation response in prior irradiation studies [9–11]. However, effects such as scattering, mode coupling, and intermodal dispersion cause the fiber output to become a complex superposition of the input, which is commonly referred to as the "speckle pattern" [12]. Machine learning models are well suited to reconstruct such patterns back to the original input image, but a key challenge lies in obtaining sufficient high-quality training data, as real beam

[*]Contact author: qiyuan.xu@liverpool.ac.uk
[†]Contact author: haozhang@liverpool.ac.uk

data are limited. To address this, we developed a synthetic training dataset using stochastic Gaussian mixture (SGM) simulation.

The proposed approach was validated in the lab using an optical setup comprising a laser-illuminated digital micromirror device (DMD) as light source and a 5-m-long MMF for transmission. A modified convolutional autoencoder (CAE) was used for reconstruction, trained exclusively on synthetic data, and evaluated on real beam images acquired from a previous CLEAR experiment. We outline the challenges of using real beam data and motivate the need for synthetic alternatives in Sec. III. The simulation-based generation of training data using an SGM model is detailed in Sec. IV. The experimental validation setup is presented in Sec. V, and the reconstruction performance using real beam data is discussed in Sec. VI.

II. MULTIMODE FIBER IMAGING

The standard expression for estimating the number of guided modes in an MMF is given in Eq. (1) [13], where M denotes the estimated mode count, n is the spatially varying refractive index within the fiber core, n_{cl} is the cladding refractive index, and the integral is taken over the fiber cross-sectional area. This relationship implies that light in an MMF can propagate along tens of thousands of distinct paths or modes. The effect is particularly pronounced in large-core MMFs, where the number of supported modes scales with the square of the core radius. Such a high mode capacity permits the transmission of rich spatial information, enabling direct, analogy-based image transport [14].

$$M \approx \frac{\pi}{\lambda^2} \int (n^2 - n_{\mathrm{cl}}^2)\, \mathrm{d}A. \tag{1}$$

As the signal propagates through the MMF, interactions with the medium act to scramble and distort the original image. These include dispersion, scattering, and mode coupling. Light disperses both temporally and spectrally, with energy mixing across various modes. Environmental perturbations, including temperature fluctuations and mechanical stress, further complicate transmission by altering the fiber's refractive index [15,16], leading to a complex spatial interference pattern at the output.

Despite these challenges, image reconstruction from speckle patterns has been extensively studied, with recent advances driven by statistical and computational methods. A classic technique is the transmission matrix approach [17], which models the fiber as a linear system: $I = |Wx|^2$, where x is the input image in vector form, W is the transmission matrix, and I is the output speckle intensity. Extensions of this model incorporate phase and polarization information [18,19]. Neural network-based approaches have been introduced to better capture nonlinear effects. For instance, a single-layer dense neural network can effectively reconstruct images transmitted through MMF [20]. More broadly, convolutional neural networks, particularly U-Net architectures, have shown strong performance in MMF reconstruction tasks, including under minor perturbations [21,22]. We use a U-Net like 2D CAE with no skip connections and a compact bottleneck, so the reconstruction is driven by a compressed latent space, which helps reduce overfitting and improve generalization. The trade-off is mild smoothing of very fine details, which is acceptable for the beam distribution reconstruction task considered here.

III. REAL BEAM DATA AND LIMITATIONS

MMF imaging-based reconstruction tasks typically use datasets composed of paired images: the transverse beam distribution before transmission and the corresponding speckle pattern recorded at the MMF output. The speckle pattern serves as the model input, and the original distribution is used as the label. To ensure generalization, the dataset should span a diverse range of beam distributions covering the relevant input space. To evaluate dataset quality and define performance metrics, we extract four transverse beam parameters from each sample: horizontal and vertical centroids and widths. These parameters are used to assess reconstruction accuracy. Each raw beam image is preprocessed by forming 1D horizontal and vertical projections; for each projection, we subtract its minimum value (removing an additive pedestal) and then normalize the resulting profiles to the [0, 1] range. Gaussian functions are fitted to the histograms using nonlinear least squares; the mean μ and standard deviation σ of each fit represent the centroid and width, respectively. This baseline correction removes only an additive offset in the extracted profiles; structured or time-varying backgrounds would require dedicated dark/background acquisition or more advanced estimation.

A previously collected dataset from the CLEAR facility was used for evaluation [23], as summarized in Table I. The dataset comprises 6257 real beam samples, acquired using a scintillating screen with an emission peak at 693 nm. Transverse beam variations were introduced via a quadrupole triplet scan to modulate beam size and dipole magnets to adjust beam position. Each data sample includes a pair of single-channel images—the MMF output and the corresponding beam distribution—recorded at a resolution of

TABLE I. Transverse beam image dataset collection at CERN.

Data collection configuration	Specifications
Facility	CLEAR
Particle source	e^-
Energy	150 MeV
Repetition rate	10 Hz
Bunch charge	0.12–0.26 nC/bunch
Screen type	Chromox ($Al_2O_3 : Cr_2O_3$)
Central emission wavelength	693 nm

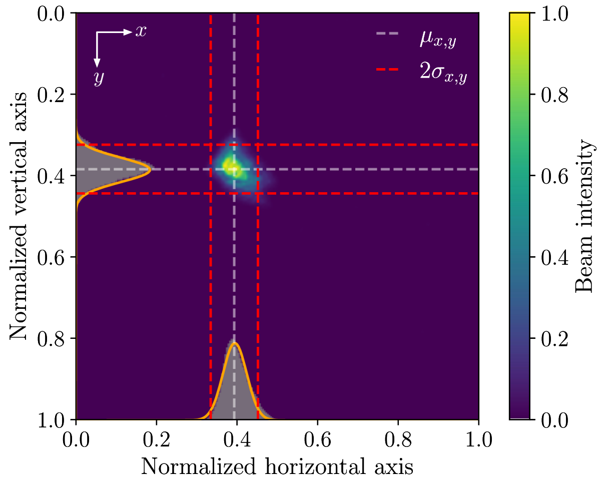

FIG. 1. Transverse beam sample from the CLEAR facility, captured on a Chromox scintillating screen. The original image size is 256×256 pixels. The transverse parameters μ_x, μ_y and σ_x, σ_y denote the beam centroid and beam widths in the horizontal and vertical axis, respectively.

256×256. One typical sample from a real beam on a scintillating screen is shown in Fig. 1, where the transverse beam distribution is represented by the intensity of the captured image. For visualization, the Gaussian fitting process is illustrated, and a beam width of 2σ is marked with red dashed lines. For evaluation, a beam width of σ is used to ensure a consistent scale of prediction errors.

Beam parameters were computed over the dataset to analyze their statistical distributions, as shown in Fig. 2. The results reveal distributional bias in the dataset. For example, the skewness (γ_1) of σ_x is 1.77, indicating a strongly right-skewed distribution, where a small number of large values extend the right tail. In contrast, the skewness of μ_x is -0.12, suggesting a slight left skew. Additionally, the parameter distributions are uneven across dimensions: the beam centroid varies more in the vertical direction, while the beam width shows greater variation horizontally. In our previous work [24], an end-to-end Encoder-Regressor model was trained to predict the four beam parameters from speckle images using this dataset, and model bias was observed due to these distributional imbalances. Together with limited beam time, radiation-induced degradation of the ground-truth camera, and practical constraints in scanning the full beam parameter space, these challenges motivated the development of a simulation framework capable of generating large, diverse, and unbiased samples for model training.

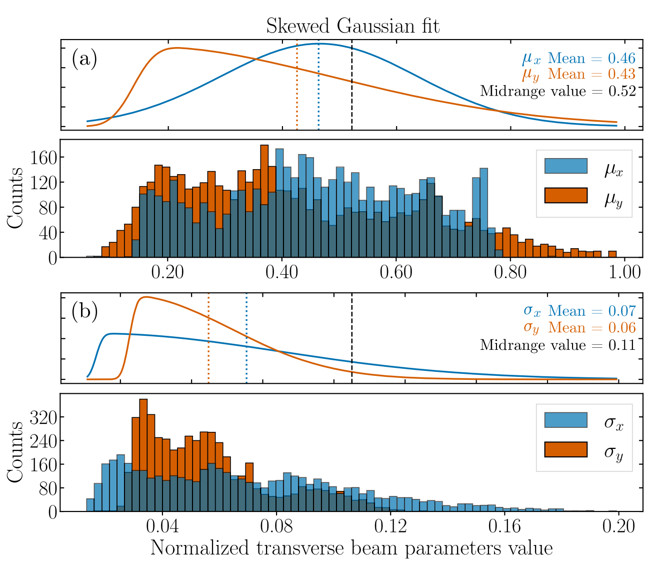

FIG. 2. Statistical distributions of the four transverse beam parameters and their skewed Gaussian fits: (a) Horizontal and vertical beam centroids. (b) Horizontal and vertical beam widths.

IV. SYNTHETIC DATA GENERATION

This work uses the SGM model as the basis for generating simulated training datasets. The approach constructs complex two-dimensional distributions by superposing multiple Gaussian components. Each component is defined by independently controlled parameters: position, amplitude, standard deviations, and rotation angle. To systematically generate these parameters, corresponding probability distributions are introduced, as described in Eqs. (2)–(5).

The final intensity distribution, denoted as $I(x, y)$ in Eq. (2), is modeled as a superposition of K two-dimensional Gaussian components, each represented by $\mathcal{N}$. Each component is characterized by its own mean vector μ_i, standard deviation vector $\boldsymbol{\sigma}_i$, and amplitude A_i. To account for orientation variability, each Gaussian is rotated by a random angle θ_i, applied via the rotation matrix R_{θ_i} to the input coordinates (x, y). Parameters $\mu_i = (\mu_x, \mu_y)$ and $\sigma_i = (\sigma_x, \sigma_y)$ are sampled uniformly from ranges informed by the measured operating regime and extended slightly beyond it: we sample $\sigma_x, \sigma_y \in [0.015, 0.25]$ and μ_x, μ_y uniformly over [0.05, 0.95] of the normalized field of view.

$$I(x, y) = \sum_{i=1}^{K} A_i \, \mathcal{N}\left(R_{\theta_i}\begin{bmatrix} x \\ y \end{bmatrix}\middle|\mu_i, \boldsymbol{\sigma}_i\right). \tag{2}$$

The amplitude term A is heuristically motivated by charge-density scaling and is sampled to keep generated intensities close to typical operating levels while still allowing controlled excursions. We split A into a baseline term and an area-dependent term, $A = U_1 + U_2$, with $U_1 \sim \mathcal{U}(I_{\min}, 100)$ and $U_2 \sim \mathcal{U}[0, 0.25/(\sigma_x\sigma_y)]$ [i.e., $c_1 = I_{\min}$, $c_2 = 100$, and $c_3 = 0.25$ in Eq. (3)]. Here $I_{\min}$ denotes the generator-defined lower bound for sampling U_1 and is set per sample as a decreasing function of the number of active Gaussian components. To model saturation, pixel values are clipped at the 8-bit full-scale value (255) before applying the same normalization used for real images.

$$A = U_1 + U_2 \tag{3}$$

$$U_1 \sim \mathcal{U}(c_1, c_2), \qquad U_2 \sim \mathcal{U}\left(0, \frac{c_3}{\sigma_x \cdot \sigma_y}\right). \tag{4}$$

The number of active Gaussian components is modeled using a binomial distribution, as described in Eq. (5). This probabilistic approach reflects our bottom-up simulation design, in which activation is determined independently at the component level (with $N = 100$ and an expected number of active components $K = 4$), rather than by sampling a total count globally.

$$\mathrm{P}(K) = \binom{N}{K} p^K (1-p)^{N-K}, \qquad K \in \{0, \cdots, N\}. \tag{5}$$

To illustrate the modeling capability of the SGM simulation, a fitting example using only three Gaussian components is shown in Fig. 3. A complex transverse beam distribution is approximated using nonlinear least squares fitting. This example is not intended as a signal reconstruction method, but to demonstrate the potential of Gaussian components as flexible building blocks.

The strength of the SGM lies in its ability to generate a wide variety of distributions by controlling the parameters of multiple components. Although real beam profiles are often unimodal, the SGM purposefully generates multi-component patterns. This design leverages the approximate linear superposition property of the MMF system, where multiple spatial features in the input field contribute quasi-independently to the output speckle pattern [25]. As a result, complex multi-Gaussian inputs can encode greater diversity while keeping the dataset compact. From a system perspective, the MMF can be viewed as a transmission channel with a complex transfer function linking input distributions to speckled output patterns. The SGM thus offers an efficient means of sampling this high-dimensional input-response space for building a comprehensive training dataset.

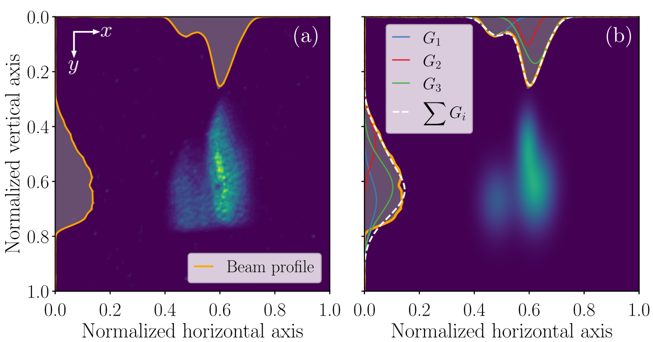

FIG. 3. (a) A transverse beam distribution from CLEAR with its corresponding horizontal and vertical beam profiles. (b) Simulation fitting using an SGM model (white dashed lines) comprised of three Gaussian components. This approach highlights the capability of SGMs to represent even complex transverse beam profiles (orange lines).

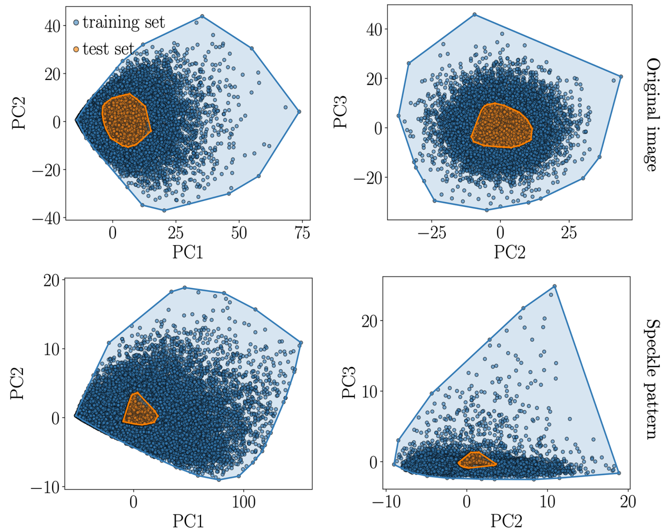

FIG. 4. Principal component projections of the training (blue) and test (orange) datasets. Top row: PCA of the original images, with each panel showing the projection onto two of the three leading components along with its explained variance. Bottom row: PCA of the corresponding MMF speckle patterns, similarly annotated.

Using the SGM simulation, 25,000 synthetic samples were generated. To assess and compare the diversity of the training and testing datasets, principal component analysis (PCA) was applied. As shown in Fig. 4, each image sample is flattened into a vector and projected onto the first three principal components. This analysis shows that the synthetic training set spans a broader region of the PCA space and covers the variance observed in the real test set along all major directions, indicating higher diversity.

V. EXPERIMENT AND MODEL TRAINING

The overall design of the experiment is shown in Fig. 5. The objective is to evaluate whether it is possible to use the synthetic dataset as the training set for a machine learning model and then use this model to reconstruct unseen transverse beam distributions transmitted through the same MMF-based optical system. To this end, two distinct datasets are prepared: one synthetic, generated using the SGM simulation, and one real, derived from beam images acquired at the CLEAR beamline. Both datasets are spatially encoded and transmitted through a laser-illuminated optical system under identical conditions.

A. Optical setup and data acquisition

As depicted in Fig. 6, a 532-nm laser (CPS532-C2) emits light that is expanded and collimated approximately tenfold

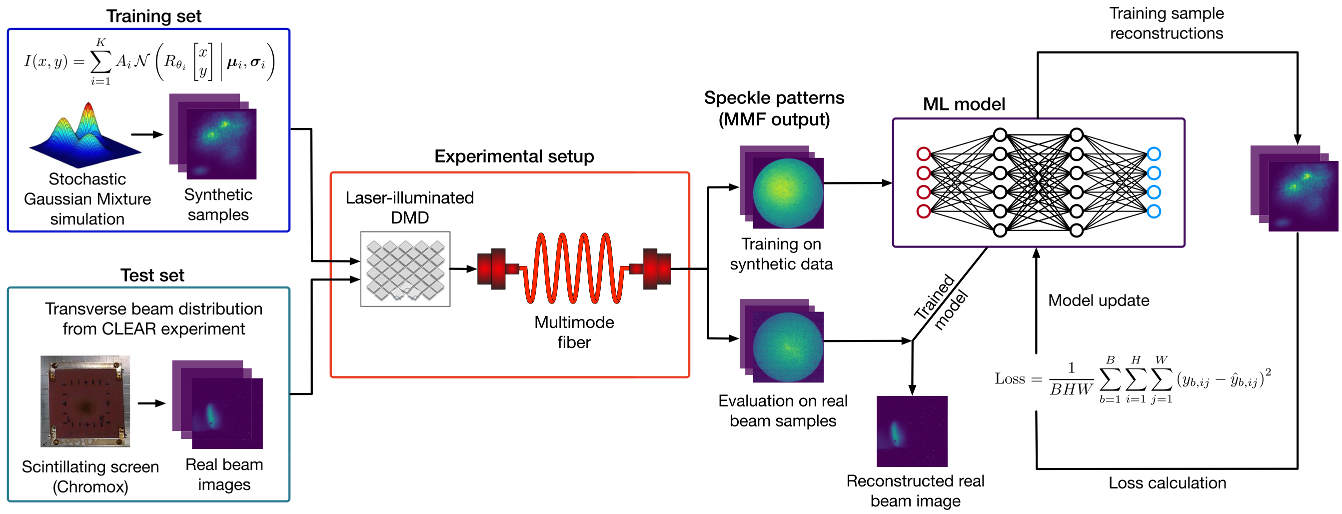

FIG. 5. Overview of the full simulation-to-reconstruction framework. Synthetic complex distributions are generated using the SGM model and used exclusively for training. Both simulated and real beam images are loaded onto a laser-illuminated DMD, and the resulting optical patterns are transmitted through an MMF. A neural network is trained to reconstruct the original transverse distribution from the MMF output.

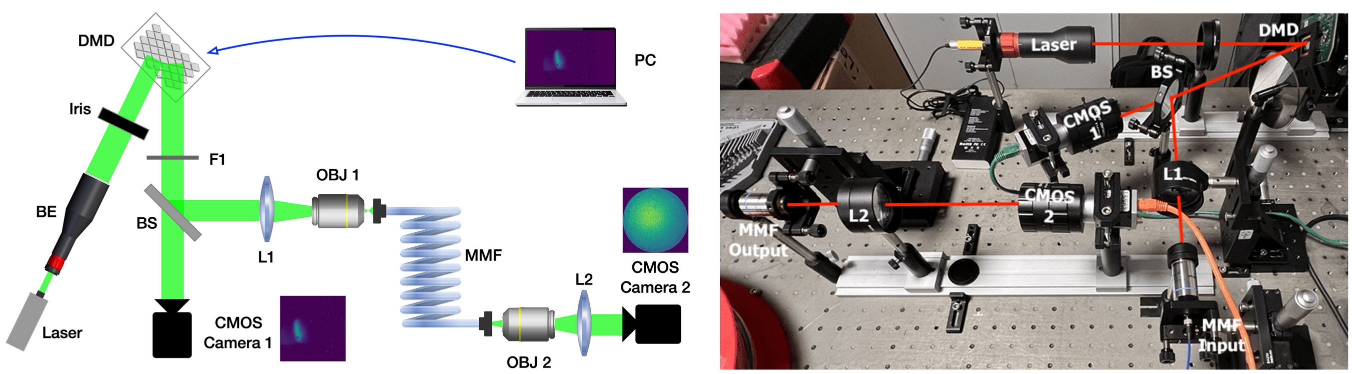

FIG. 6. Left: schematic diagram of the DMD-laser setup for collecting original image-speckle pattern pairs. Right: experimental setup at the DITALab in the Cockcroft Institute. The 5-m MMF, coiled around a cylinder, and a PoE switch for camera power and synchronization are not visible in the photo.

using a beam expander (Thorlabs BE10-532). A portion of the expanded beam is selected by a threaded iris diaphragm, adjusted to match the working area (1024×1024 micromirror region, approximately 8×8 mm) of a DMD (ViALUX V-6501 with DLP6500, 1920×1080), simulating the behavior of a scintillating screen. DMD-based setups, widely explored in accelerator diagnostics and beam instrumentation [26–29], enable programmable spatial modulation of light. The laser beam then illuminates the DMD, onto which either a simulated pattern or a measured beam distribution is loaded. The DMD modulates the spatial light pattern by tilting its micromirrors to direct light into the desired reflection path. The modulated beam then passes through a filter (NDC-100C-2) to reduce intensity and prevent saturation of both cameras. The beam is equally split by a beamsplitter (BSW16). One part is directed to the first CMOS camera (Basler ace acA1920-40gm), which records the original pattern and serves as the ground-truth image in the dataset. The other part passes through a lens assembly consisting of a doublet lens (L1, AC508-080-AB-ML) and a microscope objective (OBJ1, RMS40X), which reduce the image size for efficient coupling into a step-index multimode fiber (FP1500URT, 5 m; commercially available large-core patch cable used for convenient laboratory integration). After transmission through the MMF, the output is decoupled and magnified by a second lens assembly comprising another doublet (L2, AC508-080-AB-ML) and objective lens (OBJ2, RMS40X). The resulting image is captured by a second CMOS camera and serves as the input to the reconstruction model. The MMF is securely mounted on the optical table to maintain a fixed geometry and prevent fiber configuration change. The entire data acquisition process, including both training and testing datasets, is completed within 6 h and consists of 25,000 SGM-simulated images and 6257 real beam images, each paired with its corresponding speckle pattern.

B. Synchronization and alignment

To ensure that both cameras capture images simultaneously, we employ the precision time protocol (PTP, IEEE 1588) along with scheduled action commands supported by the Basler cameras. This approach achieves highly synchronized imaging and reduces timing mismatches to the microsecond level [30]. The exposure time is set to 40 ms, which is well within this synchronization tolerance. In addition, a fiber coupling profile is established by sequentially displaying a single moving square block on the active area of the DMD, which is divided into eight rows and eight columns, forming 64 blocks in total, as Fig. 7 shows. For each block, the corresponding fiber output image is recorded, and a coupling ratio is calculated as the total output intensity divided by the total input intensity. Due to the higher gain of the camera recording the speckle pattern, the relative values in the MMF coupling rate profile can exceed 1. In practice, values above 2.5 are generally considered sufficient. To ensure effective alignment, a procedure is established to verify that all four corner blocks exhibit sufficient coupling rates before data acquisition, as low transmission typically occurs near the fiber input edges. To optimize MMF mode excitation, the input angle should approach but not exceed the fiber's acceptance angle. This maximizes the number of supported modes that are excited without incurring excessive losses.

C. Model architecture and training

After dataset collection, the next step is training the reconstruction model. The designed model follows a convolutional autoencoder architecture, as shown in Fig. 8. The encoder processes speckle patterns from the MMF output using a series of Conv2D layers. Each layer includes batch normalization and Leaky ReLU activation to improve gradient stability during training. Instead of max pooling, stride-2 convolutions are employed to achieve spatial downsampling. This choice replaces the nonparametric and heuristic nature of pooling with a learnable operation. After six layers, the latent representation is compressed to 1024 feature maps of size 4×4.

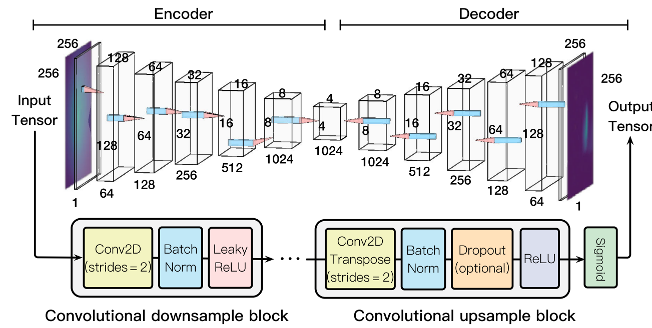

FIG. 8. Architecture of the CAE for beam image reconstruction from MMF speckle patterns, featuring downsampling blocks in the encoder branch and upsampling blocks in the decoder branch.

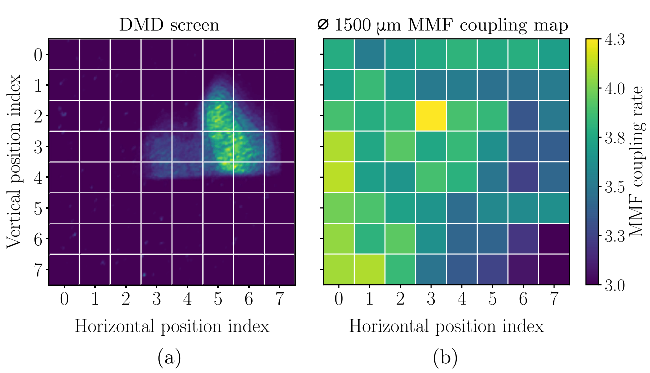

FIG. 7. Spatial-based input coupling efficiency in an MMF. (a) Example input distribution at the fiber's proximal end. (b) Heat map showing the transmission efficiency for each input location.

The decoder reconstructs the speckle patterns through six transposed convolutional layers from those abstract representation, restoring the original resolution of 256×256 pixels. Batch normalization and ReLU activations are used in the decoder, with dropout applied to the first three layers to mitigate overfitting. Since high-frequency details in the speckle image are not essential for extracting core beam parameters, skip connections were omitted to encourage reliance on the latent representation. This bottleneck-driven design acts as a regularizer, helping suppress noise and irrelevant distortions while improving generalization.

The neural network model was developed using TensorFlow and trained on a Linux-based high-performance computing (HPC) cluster equipped with GPU acceleration. An initial learning rate of 1.0×10^{-4} was employed, and the model was optimized using the Adam algorithm, with training proceeding for a maximum of 100 epochs. Early stopping was implemented based on validation loss, resulting in convergence at epoch 59. For model evaluation, the dataset was partitioned such that 1200 samples were randomly selected from the test set to form a validation set, leaving 5000 samples for final testing. The loss function utilized was the mean squared error (MSE), calculated as the average of the squared differences between the reconstructed and ground-truth images on a per-pixel basis, as defined in Eq. (6).

$$\mathrm{Loss}_{\mathrm{MSE}} = \frac{1}{BHW} \sum_{b=1}^{B} \sum_{i=1}^{H} \sum_{j=1}^{W} (y_{b,ij} - \hat{y}_{b,ij})^2, \qquad (6)$$

where B denotes the batch size, while H and W represent the image height and width in pixels, respectively. An alternative approach considered was an end-to-end loss function based on beam parameters. However, this was deemed unsuitable for two primary reasons: (i) the training data comprised SGM patterns lacking explicit beam parameters, and (ii) different transverse beam distributions can correspond to identical sets of beam parameters, potentially leading to information loss and reduced sample efficiency during training. Consequently, the pixelwise MSE loss was adopted to provide detailed reconstruction spatial error information.

VI. TRANSVERSE BEAM RECONSTRUCTION

By applying the CAE model trained on synthetic data to the test set, we evaluate both the reconstruction accuracy and the generalization performance. Representative reconstruction results for selected real beam samples are shown in Fig. 9, featuring beams with varying sizes and positions. The red dashed line indicates the contour of the original beam distribution at the 90th percentile intensity level, while the yellow dashed line shows the corresponding contour of the reconstructed output. Beam centroids are also marked for reference. These results demonstrate that regardless of beam size or position, the model is able to recover the main structure of the beam—even for samples with complex, non-Gaussian distributions.

However, some high-frequency variations, particularly fine details at the beam edges, are not fully captured, resulting in a smoother predicted profile. This is primarily attributed to the model's bottleneck architecture, which restricts the level of detail passed to the decoder, and to the nature of the generated training dataset, in which a lower bound on the Gaussian widths limits the smallest feature scale. These design choices intentionally prioritize generalization ability over precise reconstruction of fine shape details.

After the beam distribution is reconstructed, the four transverse beam parameters are extracted from each sample for quantitative comparison. The statistical results are presented in Fig. 10. Figures 10(a)–10(c) illustrate the beam position reconstruction on test set, with blue and yellow representing the horizontal and vertical parameters, respectively. In Fig. 10(a), the scatter points closely follow the 45° line, indicating strong agreement between predicted and actual values. The residuals, computed as the difference between actual and predicted values, are shown in Fig. 10(b). The residual plot shows that most samples have minimal error, with only a few exhibiting larger deviations. Since the image dimension is normalized to 1, errors for all four parameters are consistently scaled, allowing direct percentage computation by multiplying by 100. Figure 10(c) displays the residual distribution, clipped within $\pm 10\%$ for better visualization, with 99.3% of predictions concentrated within $\pm 5\%$. Similarly, beam width reconstruction results are shown in Figs. 10(d)–10(f), where errors remain within a smaller range. As seen in Fig. 10(f), 98.7% of residuals also lie within $\pm 5\%$, and more importantly, the centered distribution indicates no systematic prediction bias.

Table II presents the final benchmark results for three different models trained on either the CLEAR or the SGM dataset. The scores are given as RMSE values in the form of mean $\pm$ standard deviation in percentage, based on three independent runs to reflect stochastic variations from model initialization and data sampling. For models trained on the CLEAR dataset, evaluation is performed on a fixed test set obtained by sampling one tenth of the data (training split: 8:1:1). For models trained on the SGM dataset,

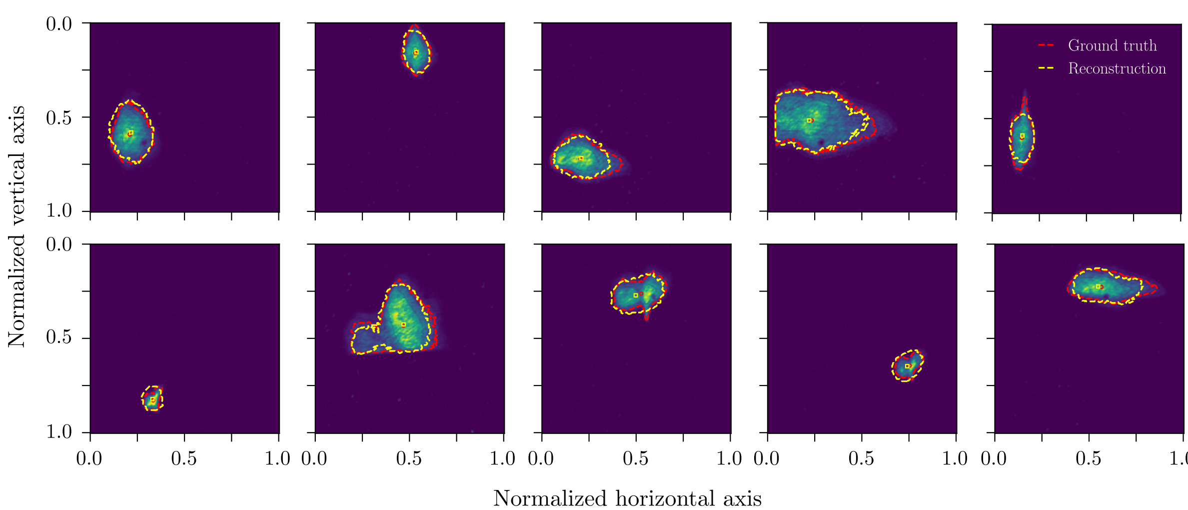

FIG. 9. Comparative visualization of reconstructed transverse beam distributions using real beam data at the 90th percentile intensity level. The samples shown are representative cases selected to reflect variation in beam size, shape, and position. Ground-truth values are depicted with red dashed lines, and model predictions with yellow dashed lines; beam centroids follow the same color scheme. For clarity, the background beam intensity distributions are rescaled to their respective min-max ranges.

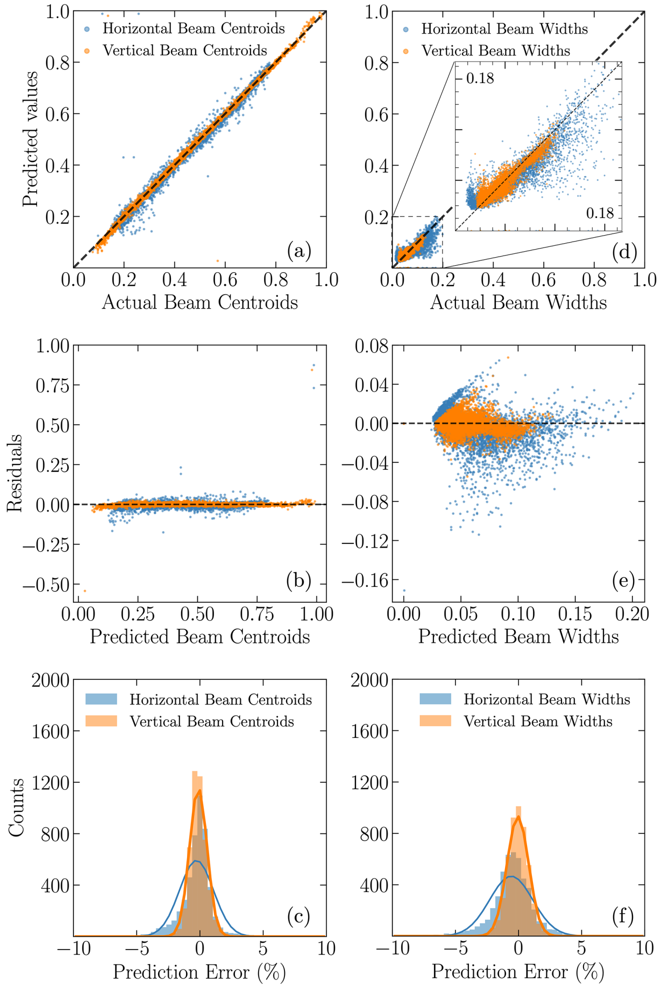

FIG. 10. Statistics of the reconstruction errors for four transverse beam parameters on the test set. The first column shows beam centroids and the second column shows beam widths, with blue for the horizontal parameter and yellow for the vertical. The top row plots actual vs predicted values, the middle row shows residuals (prediction minus real), and the bottom row displays the residual distributions.

evaluation is carried out on almost the entire CLEAR dataset (excluding the validation set), providing an even more comprehensive measure of performance. The mean indicates the model reconstruction error, and the standard deviation represents the model stability.

On the CLEAR dataset, three models were trained: Pix2Pix, a conditional GAN-based model [31]; the Encoder-Regressor Network (ERN) [24]; and the proposed CAE, which reconstructs the full transverse beam distribution. The ERN retains only the encoder branch of an autoencoder and appends a fully connected regressor to predict four scalar targets $(\mu_x, \mu_y, \sigma_x, \sigma_y)$ from the encoder output. As a supervised regression model, ERN requires ground-truth parameter labels for every training sample; therefore, it can only be trained on datasets where these targets are well defined, making the SGM dataset unsuitable for ERN. Although ERN obtained the lowest RMSEs on the CLEAR dataset, its design likely limits its performance ceiling, since different transverse beam distributions can yield identical beam centroid and width, and the loss function could be insensitive to such variations. This suggests that even with a dataset tailored to its requirements (such as a single-Gaussian set), the achievable performance is still inherently constrained. The CAE, on the other hand, by reconstructing the beam distribution end to end, forces the model to learn the optical relay system's response, resulting in higher modeling capability. Another observation is that Pix2Pix exhibited the largest variability, consistent with the known instability of GAN training [32].

On the SGM dataset, the CAE achieved both the highest accuracy and the highest stability, demonstrating the effectiveness of the SGM dataset and the model optimization. The Pix2Pix was first tuned for training on the CLEAR dataset, and for consistency, the same configuration was applied to the SGM dataset. However, the model complexity may have been insufficient for this relatively larger dataset, which likely explains why the performance improvement from dataset change was not as large as that observed for the CAE.

It can be seen from Table II that many models exhibit larger errors in the horizontal (x) beam parameters compared to the vertical ones. This discrepancy can be attributed to the characteristics of the specific test set used in this experiment. As illustrated in Fig. 2, the range of σ_x extends up to 0.20, nearly twice the range observed in the σ_y (around 0.12). This broader variation introduces higher distributional variance, making horizontal beam reconstruction more challenging. As a result, errors in horizontal beam width are often larger.

In addition, the results show a correlation between beam width and position errors: larger errors in width tend to coincide with larger errors in centroid position, and vice versa. This correlation arises because both quantities are derived from the same reconstructed transverse profile. The centroid is obtained as the mean of a Gaussian fit to the predicted profile and is therefore sensitive to the overall reconstructed shape. Asymmetric reconstruction errors (e.g., overestimating intensity on one side) can bias the fitted mean toward the elongated tail, reducing position accuracy. Likewise, distortions in the reconstructed profile can affect the fitted width, so the two errors can be correlated. Consequently, errors in beam width reconstruction can indirectly degrade the precision of position measurement. Mild smoothing applied to the reconstructed profiles could potentially mitigate small asymmetric artifacts, though at the cost of some loss of edge detail.

The correlation between prediction error and beam parameter values was also examined. A general trend is observed: prediction errors tend to increase for more extreme or boundary values of the parameters. For beam

TABLE II. RMSEs for four transverse beam parameters. CLEAR-trained models are evaluated on a test set comprising one tenth of CLEAR; SGM-trained models on the full CLEAR dataset excluding validation samples. Values are mean ± standard deviation over three runs with different random seeds. The last column is the mean RMSE across all parameters.

Model	$\delta\mu_x$	$\delta\mu_y$	$\delta\sigma_x$	$\delta\sigma_y$	δ_{avg}
			Trained on CLEAR dataset		
Pix2Pix	9.21 ± 2.63	7.39 ± 3.53	5.41 ± 1.18	4.34 ± 0.51	6.59 ± 2.96
Encoder–regressor	4.41 ± 0.46	3.61 ± 0.29	4.15 ± 0.59	3.20 ± 0.70	3.90 ± 0.29
Conv–autoencoder	6.78 ± 0.76	7.66 ± 0.14	4.52 ± 0.06	3.44 ± 0.07	5.60 ± 0.19
			Trained on SGM dataset		
Pix2Pix	7.50 ± 0.96	7.39 ± 1.46	2.83 ± 0.16	3.18 ± 0.28	5.23 ± 0.64
Conv–autoencoder	1.94 ± 0.27	1.24 ± 0.39	1.88 ± 0.16	0.90 ± 0.03	1.49 ± 0.19

centroids, higher errors are more common at the edges of the distribution range, whereas samples near the center typically yield lower errors. Similarly, for beam widths, extreme cases—particularly the largest anisotropic shapes—are associated with higher prediction errors, while more typical, midrange values are reconstructed with better accuracy. This behavior is likely due to the relatively underrepresentation of boundary cases in the training set, leading to reduced generalization performance in those regions of the parameter space.

VII. DISCUSSION

The scaling effect of the training set is investigated by keeping the CAE architecture and training hyperparameters fixed while varying only the size of the training dataset. The dataset size starts at 5000 randomly sampled training points and increases in steps of 5000, up to a maximum of 25,000. For each dataset size, an independent model is trained and evaluated on the same test dataset. The results are presented in Fig. 11. Initially, as the dataset size increases, the beam parameter reconstruction errors decrease significantly. With further increases in the dataset size, the errors continue to decrease, albeit at a slower rate. For the largest training set, the validation loss stabilizes after 59 epochs, indicating that the model requires sufficient training time to learn from the larger dataset. In contrast, for the smallest dataset (5000 samples), the model converges also slowly, likely due to insufficient training data, resulting in underfitting and instability. Overall, these results suggest that there is still room for further improvement with additional training data, as the model has not yet reached saturation.

Another important factor examined is the relationship between the average intensity of the MMF speckle pattern and the reconstruction error of beam parameters. To analyze this effect, test set samples were grouped by the average intensity of their speckle patterns. For each bin, the mean absolute error (MAE) was calculated as the average, over all samples in the bin, of the absolute errors across their four transverse beam parameters. This correlation is shown in Fig. 12. A linear fit (yellow line) reveals a clear trend: as input intensity increases, the MAE decreases from approximately 1.0 to 0.4. At lower intensity levels, the MAE is higher—likely due to a reduced signal-to-background noise ratio, which degrades the quality of the extracted features. This suggests that, in future data

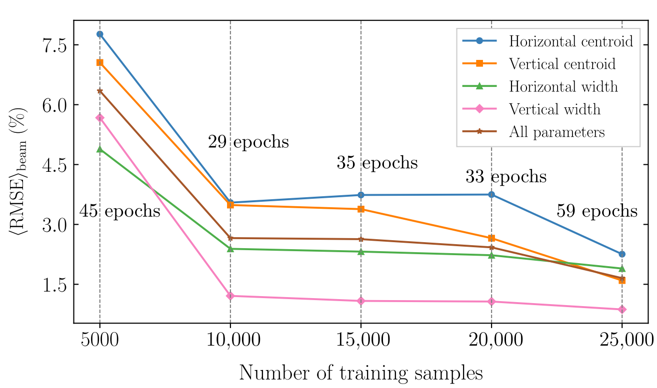

FIG. 11. Effect of training dataset size (SGM) on the total RMSE % for the four reconstructed beam parameters, evaluated on the same test set. The dataset size is incremented by 5000 at each step, and the training convergence speed (in epochs) is also shown.

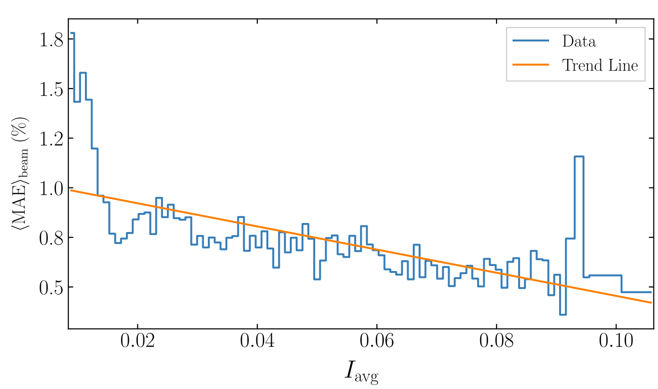

FIG. 12. Observed correlation between the total beam-parameter reconstruction error and the normalized (0–1) mean intensity of the MMF speckle input in the test set. The yellow line shows the first-order fit.

collection experiments, maintaining a minimum input intensity threshold may help improve reconstruction accuracy by ensuring sufficient signal quality.

The four beam parameters are expected to be relatively independent of absolute beam intensity, since scaling the amplitude of a distribution does not affect the calculation of its μ or σ, provided that the signal remains above the background noise level. This implies that the task has stronger tolerance to errors in absolute beam intensity, provided that the relative intensity profile is preserved. However, ensuring this in practice still requires a training dataset that spans a broad range of intensities, to reconstruct the relative intensity across varying signal strengths. This assumption also holds for non-Gaussian distributions as shape-based statistics generally remain invariant under uniform amplitude scaling.

One of the key assumptions underlying the SGM dataset is the superposition property of the MMF imaging system. To evaluate this, we numerically stacked multiple speckle patterns from the test set by summing them pixelwise to form composite input images. Each composite represents the superposition of speckle patterns from two to nine individual beam samples. These images were then fed into the trained model to assess its response.

The results show that the model reconstructed the underlying beam structures across all stacking levels, with only minor degradation, demonstrating its ability to interpret superposed speckle inputs. Minor pixel saturation effects were observed only at the highest stacking levels (e.g., eight or nine overlaid samples), but they did not significantly degrade reconstruction quality. This experiment was designed to assess both the physical superposition property of the MMF system and the model's ability to adapt to it. The outcome validates the superposition assumption and explains why a training dataset composed of multiple-distribution beam profiles can still generalize effectively to single-distribution test samples without introducing artifacts. Moreover, this property enhances dataset efficiency, an important consideration for practical deployment, where minimizing data collection and training time is essential for on-site usability.

While the laser-based proof of concept achieves high accuracy, translation to beamline operation will rely on scintillation, where illumination characteristics and long-term MMF stability are key considerations. The MMF transmission function is wavelength dependent [33], and broadband, lower coherence scintillation can reduce speckle contrast through spectral averaging and shift the output statistics [34]. Polarization differences can, in principle, also affect the MMF output via polarization mode coupling. In practice, this can be mitigated by engineering the illumination to better reflect the deployed scintillation spectrum (e.g., an LED or broadband with a bandpass filter chosen to approximate the screen emission band around the peak). In addition, synthetic-to-experiment transfer has been demonstrated in related optical fiber inverse problems [35], and realistic noise and mismatch augmentation has been shown to improve simulation-to-experiment generalization [36].

Here the training and evaluation data were collected in one continuous measurement campaign with the MMF rigidly mounted; longer term (multiday) drift was not assessed and will be investigated in future work. Over longer timescales, temperature drift and small mechanical perturbations may lead to gradual changes in mode coupling and intermodal phases; reported mitigations include auxiliary measurements to identify and track fiber state changes (e.g., reference-assisted bending compensation [37] and active fiber configuration sensing [38]), as well as reflection mode transmission matrix recovery [39], which avoids distal access. Learning-based strategies designed to generalize across dynamically perturbed fiber conformations have also been reported [40].

VIII. CONCLUSION

In this work, we explored an MMF-based image transmission system that potentially enables a standard camera, positioned in a low-dose area, to perform beam imaging tasks. To address the challenge of acquiring large-scale, high-quality real beam data for training, we proposed an SGM simulation to generate synthetic training samples. A modified CAE was employed as the reconstruction model. A complete optical setup was established, in which a laser-illuminated DMD served as a light source to emulate emission from a scintillating screen. Both synthetic and real beam distributions were spatially encoded and projected through the same optical system; however, only synthetic data were used for model training, while real beam images were reserved for evaluation. The proposed approach achieved an overall RMSE of 1.49% across beam parameters, outperforming all previously tested models and datasets. The synthetic dataset mitigated distributional gaps and class imbalance present in real beam data. These results indicate that the SGM simulation provides a robust strategy for reconstructing diverse transverse beam distributions without requiring prior experimental data.

To further improve reconstruction accuracy and advance toward robust transverse beam diagnostics on real beamlines, several enhancements are under consideration. First, we aim to improve the similarity between the DMD-based light source and actual scintillation-screen emission, particularly in terms of wavelength, coherence, and polarization, to better replicate beamline conditions. Second, qualification and selection of candidate large-core, radiation-tolerant MMFs will be part of ongoing follow-up work, including quantitative irradiation tests (e.g., radiation-induced attenuation in the relevant wavelength band), as a next step beyond the current proof-of-concept fiber. Third, the dataset can be refined by increasing the variance of the Gaussian distributions in both position and size to better cover boundary cases. On the model side, recent work has demonstrated the potential of transformer-based architectures [41], which leverage self-attention mechanisms to capture both local and global structure in the input and offer improved representational capacity for complex beam profiles. Finally, the impact of environmental perturbations on the fiber configuration, such as mechanical shifts or temperature fluctuations, will be investigated to ensure long-term system robustness.

ACKNOWLEDGMENTS

This work was supported by the Science and Technology Facilities Council (STFC) through the LIV.INNO Centre for Doctoral Training under grant agreement ST/W006766/1. Support from CERN as a Project Associate within the Beam Instrumentation Group is also gratefully acknowledged. The authors thank J. Wolfenden and D. Ghosal for providing necessary equipment, and N. Kumar, M. Patel, and R. Ghagi for helpful discussions. Model training was performed on Barkla, part of the High Performance Computing facilities at the University of Liverpool, United Kingdom.

DATA AVAILABILITY

The data that support the findings of this article are not publicly available upon publication because it is not technically feasible and/or the cost of preparing, depositing, and hosting the data would be prohibitive within the terms of this research project. The data are available from the authors upon reasonable request.

[1] E. Bravin, Transverse beam profiles, arXiv:2005.07400.

[2] P. Forck, *Lecture Notes on Beam Instrumentation and Diagnostics* (Joint University Accelerator School, Archamps, 2010), https://www-bd.gsi.de/conf/juas/juas_script.pdf.

[3] M. Brugger, Radiation damage to electronics at the LHC, in *Proceedings of IPAC2012, New Orleans, Louisiana, USA* (JACoW, Geneva, Switzerland, 2012), Vol. 1205201, THPPP006.

[4] D. Celeste, E. Bravin, S. Burger, F. Roncarolo *et al.*, Development and test of a beam imaging system based on radiation tolerant optical fiber bundles, in *Proceedings of the 10th International Particle Accelerator Conference, IPAC-2019, Melbourne, Australia* (JACoW, Geneva, Switzerland, 2019), pp. 2658–2660.

[5] S. Burger, E. Bravin, A. Topaloudis, E. Veyrunes, F. Velotti, and F. Roncarolo, New CERN SPS beam dump imaging system, in *Proceedings of the 10th International Beam Instrumentation Conference, IBIC-2021, Pohang, Republic of Korea* (JACoW, Geneva, Switzerland, 2021), p. 254.

[6] Institut Supérieur de l'Aéronautique et de l'Espace, ISAE-SUPAERO, https://www.isae-supaero.fr/en/ (2023).

[7] F. Roncarolo, S. Burger, G. Trad, F. Velotti, A. Topaloudis, T. Levens, A. Guerrero, and T. Lefevre, Review of CERN beam instrumentation for fixed target experiments, in *Proceedings of the 14th International Particle Accelerator Conference, Venice, Italy* (JACoW, Geneva, Switzerland, 2023), THPL080.

[8] CERN Radiation Working Group, Beam profile measurements using beam camera at the LHC, Available at https://indico.cern.ch/event/1207408/contributions/5078453/attachments/2544392/4381289/RadWG_BeamCam_2022_11_09.pdf (2022) [accessed April 25, 2023].

[9] J. Blanc, F. Achten, A. Alessi, A. Amezcua, J. Kuhnhenn, A. Pastouret, D. Ricci, and I. Toccafondo, Characterization of radiation-resistant multimode optical fibers for large-scale procurement, IEEE Trans. Nucl. Sci. **68**, 1407 (2021).

[10] S. Girard, J. Kuhnhenn, A. Gusarov, B. Brichard, M. Van Uffelen, Y. Ouerdane, A. Boukenter, and C. Marcandella, Radiation effects on silica-based optical fibers: Recent advances and future challenges, IEEE Trans. Nucl. Sci. **60**, 2015 (2013).

[11] C. Campanella, V. De Michele, A. Morana, G. Mélin, T. Robin, E. Marin, Y. Ouerdane, A. Boukenter, and S. Girard, Radiation effects on pure-silica multimode optical fibers in the visible and near-infrared domains: Influence of OH groups, Appl. Sci. **11**, 2991 (2021).

[12] D. L. Smith, L. V. Nguyen, M. I. Reja, E. P. Schartner, H. Ebendorff-Heidepriem, D. J. Ottaway, and S. C. Warren-Smith, Harnessing the power of complex light propagation in multimode fibers for spatially resolved sensing, Photonics Res. **12**, 411 (2024).

[13] R. Paschotta, V number, RP photonics encyclopedia, available online at https://www.rp-photonics.com/v_number.html (2024).

[14] D. Psaltis and C. Moser, Imaging with multimode fibers, Opt. Photonics News **27**, 24 (2016).

[15] W. Xiong, P. Ambichl, Y. Bromberg, B. Redding, S. Rotter, and H. Cao, Principal modes in multimode fibers: Exploring the crossover from weak to strong mode coupling, Opt. Express **25**, 2709 (2017).

[16] H. Cao, T. Čižmár, S. Turtaev, T. Tyc, and S. Rotter, Controlling light propagation in multimode fibers for imaging, spectroscopy, and beyond, Adv. Opt. Photonics **15**, 524 (2023).

[17] S. M. Popoff, G. Lerosey, R. Carminati, M. Fink, A. C. Boccara, and S. Gigan, Measuring the transmission matrix in optics: An approach to the study and control of light propagation in disordered media, Phys. Rev. Lett. **104**, 100601 (2010).

[18] P. Caramazza, O. Moran, R. Murray-Smith, and D. Faccio, Transmission of natural scene images through a multimode fibre, Nat. Commun. **10**, 2029 (2019).

[19] W. Fan, Z. Chen, V. V. Yakovlev, and J. Pu, High-fidelity image reconstruction through multimode fiber via polarization-enhanced parametric speckle imaging, Laser Photonics Rev. **15**, 2000376 (2021).

[20] C. Zhu, E. A. Chan, Y. Wang, W. Peng, R. Guo, B. Zhang, C. Soci, and Y. Chong, Image reconstruction through a multimode fiber with a simple neural network architecture, Sci. Rep. **11**, 896 (2021).

[21] N. Borhani, E. Kakkava, C. Moser, and D. Psaltis, Learning to see through multimode fibers, Optica **5**, 960 (2018).

[22] A. Abdulaziz, S. P. Mekhail, Y. Altmann, M. J. Padgett, and S. McLaughlin, Robust real-time imaging through flexible multimode fibers, Sci. Rep. **13**, 11371 (2023).

[23] G. Trad and S. Burger, Artificial intelligence-assisted beam distribution imaging using a single multimode fiber at CERN, in *Proceedings of the 13th International Particle Accelerator Conference, IPAC-2022, Bangkok, Thailand* (JACoW, Geneva, Switzerland, 2022), p. 339.

[24] Q. Xu, A. Hill, H. Zhang, C. Welsch, F. Roncarolo, and G. Trad, Estimation of beam transverse parameters through a multimode fiber using deep learning, in *Proceedings of the 13th International Beam Instrumentation Conference, IBIC-2024, Beijing* (JACoW, Geneva, Switzerland, 2024), p. 170.

[25] J. Pauwels, G. Van der Sande, and G. Verschaffelt, Space division multiplexing in standard multi-mode optical fibers based on speckle pattern classification, Sci. Rep. **9**, 17597 (2019).

[26] H. Zhang, R. Fiorito, A. Shkvarunets, R. Kishek, and C. Welsch, Beam halo imaging with a digital optical mask, Phys. Rev. ST Accel. Beams **15**, 072803 (2012).

[27] B. Riddick, E. Montgomery, R. Fiorito, H. Zhang, A. Shkvarunets, Z. Pan, and S. Khan, Photocathode quantum efficiency mapping at high resolution using a digital micromirror device, Phys. Rev. ST Accel. Beams **16**, 062802 (2013).

[28] S. Li, S. Alverson, D. Bohler, A. Egger, A. Fry, S. Gilevich, Z. Huang, A. Miahnahri, D. Ratner, J. Robinson *et al.*, Ultraviolet laser transverse profile shaping for improving x-ray free electron laser performance, Phys. Rev. Accel. Beams **20**, 080704 (2017).

[29] S. Li, F. Cropp, K. Kabra, T. Lane, G. Wetzstein, P. Musumeci, and D. Ratner, Electron ghost imaging, Phys. Rev. Lett. **121**, 114801 (2018).

[30] F. Karstens, GigE multi-camera synchronization, White Paper, Basler AG, 2016.

[31] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, Image-to-image translation with conditional adversarial networks, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, New York, 2017), pp. 1125–1134.

[32] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. A. Bharath, Generative adversarial networks: An overview, IEEE Signal Process. Mag. **35**, 53 (2018).

[33] K. Okamoto, *Fundamentals of Optical Waveguides*, 3rd ed. (Academic Press, New York, 2022).

[34] A. Efimov, Coherence and speckle contrast at the output of a stationary multimode optical fiber, Opt. Lett. **43**, 4767 (2018).

[35] F. Ren, Y. Xie, X. Pi, and X. Wang, Bridge the gap between simulated and real-world data in optical fiber mode decomposition for accuracy improvement: A deep learning-based co-learning framework with visual similarity-based matching, Expert Syst. Appl. **256**, 124937 (2024).

[36] B. Yan, J. Zhang, S. Mi, M. Wang, C. Wang, G. Fan, and P. Zhang, MT-SCUNet: A hybrid neural network for enhanced mode decomposition in optical fibers, Opt. Fiber Technol. **93**, 104196 (2025).

[37] S. Farahi, D. Ziegler, I. N. Papadopoulos, D. Psaltis, and C. Moser, Dynamic bending compensation while focusing through a multimode fiber, Opt. Express **21**, 22504 (2013).

[38] R. Zhu, J. Luo, X. Zhou, H. Feng, and F. Xu, Anti-perturbation multimode fiber imaging based on the active measurement of the fiber configuration, ACS Photonics **10**, 3476 (2023).

[39] G. S. D. Gordon, M. Gataric, A. G. C. P. Ramos, R. Mouthaan, C. Williams, J. Yoon, T. D. Wilkinson, and S. E. Bohndiek, Characterizing optical fiber transmission matrices using metasurface reflector stacks for lensless imaging without distal access, Phys. Rev. X **9**, 041050 (2019).

[40] S. Resisi, S. M. Popoff, and Y. Bromberg, Image transmission through a dynamically perturbed multimode fiber by deep learning, Laser Photonics Rev. **15**, 2000553 (2021).

[41] A. Dosovitskiy *et al.*, An image is worth 16x16 words: Transformers for image recognition at scale, arXiv:2010.11929.