ISBN: 978-3-95450-249-3 ISSN: 2673-5350 doi: 10.18429/JACoW-IBIC2024-TUP54

ESTIMATION OF BEAM TRANSVERSE PARAMETERS THROUGH A MULTIMODE FIBER USING DEEP LEARNING

Q. Xu*, A. Hill, H. D. Zhang, C. P. Welsch
Cockcroft Institute and University of Liverpool, Warrington, UK
F. Roncarolo, G. Trad, CERN, Meyrin, Switzerland

Abstract

In response to CERN's need for alternative imaging solutions of scintillating screens due to the discontinuation of radiation-hardened VIDICON tubes, the single large-core multimode fiber has been identified as a potential medium to transmit image signals to a CMOS camera situated away from radiation-prone areas. However, significant challenges in image distortion at the fiber's output end complicate the reconstruction of the original beam distribution.

To address this, a novel machine learning-based approach was introduced that utilizes a deep convolutional encoder-regressor network. It first compresses the fiber image into a latent space. Subsequently, a fully connected regression network directly estimates the beam parameters, such as centroids and widths, from the encoder output without reconstructing the detailed image. This contribution will showcase an end-to-end system capable of estimating transverse beam parameters from the fiber output patterns and offering a safe, camera-preserving solution for beam imaging in high-radiation environments.

INTRODUCTION

At CERN, the monitoring of 2D transverse beam profiles was mostly done by rad-hard VIDICON cameras. With the cessation of this kind of camera on a global scale, the transfer to other imaging solutions is required. This has led to the exploration and adoption of CMOS cameras as a viable alternative. However one of the primary concerns using CMOS cameras is the radiation damage when operating close to the accelerator. A potential solution is to use optical fibers relaying the initial image signal so that the camera can be placed somewhere safe. Multimode fiber (MMF) has been studied as a direct imaging medium over the past decades. The large core diameter of MMF supports a high number of light propagation modes, which is proportional to the square of the fiber core radius according to the mode calculation formula [1], and this correlates with the amount of information it can encode and transmit. Since the mode is a distinct way light propagates in an MMF, finer resolution and higher bandwidth are provided for image transmission, potentially preserving most of the input information and therefore it is selected for this task.

The reconstruction challenge caused by MMF after transmission is largely due to mode coupling inside the fiber [2]. This phenomenon occurs when light energy is transferred between different propagation modes within the fiber due to changes in geometry and refractive index, making the transmission property of the fiber more dynamic. The methods to reconstruct the original image from the MMF output speckle pattern range from initial methods involving phase conjugation to wavefront shaping that uses spatial light modulators (SLM). These methods have certain limitations, such as the reconstructed image must appear at the proximal end of the fiber, which is the same side as the original light source, or requiring precise control of the image source [3], making them impractical for our scenario.

Statistical-based methods that use computers to model the optical system digitally are another popular means of MMF image reconstruction, such as a complex value transmission matrix (TM) that tries to approximate the mapping relationship between fiber input and output planes in terms of intensity and phase. However, TM is relatively susceptible to environmental perturbations and usually requires a complex equipment setup for precise measurement [4]. More recent approaches show that neural networks have very good generalization ability, capable of modeling both linear and non-linear effects within the fiber, and require only intensity measurements from relatively small datasets, making them one of the most suitable approaches for retrieving information from the distorted output.

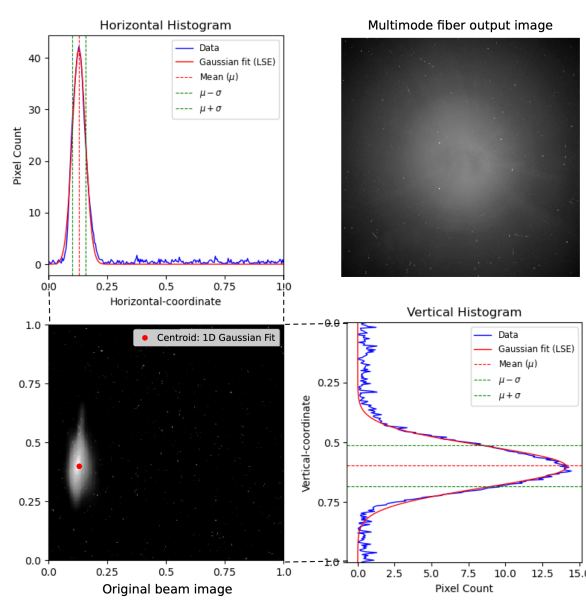

Figure 1: A representative example of transverse parameter calculation from a scintillating screen and the corresponding MMF output pattern.

* qiyuan.xu@liverpool.ac.uk

Content from this work may be used under the terms of the CC BY 4.0 licence (© 2024). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI.

ISBN: 978-3-95450-249-3 ISSN: 2673-5350 doi: 10.18429/JACoW-IBIC2024-TUP54

This study focuses on estimating four basic transverse beam parameters from fiber output: two that determine beam position and two that describe beam size. The calculation of these parameters begins with capturing the ground truth beam image on a scintillating screen, followed by normalizing the image dimensions and reducing background noise. Histograms based on pixel intensity are then created for both horizontal and vertical directions. Two Gaussian distributions are subsequently fitted to each histogram using a non-linear least squares algorithm as Fig. 1 demonstrates. The mean of these Gaussians represents the beam centroids, and their standard deviations (std) define the beam widths. Together, these describe the beam position and size.

DATASET AND MODEL

In this experiment, real beam data were used for both training and evaluation; the dataset was collected at CERN's CLEAR facility, where the light produced by an electron beam of 150 MeV interacting with a Chromox scintillating screen ($Al_2O_3 : CrO_2$) is captured. A Quadrupole scan is applied for obtaining variations of the beam transverse distributions [5]. As Fig. 2 shows, the first CMOS camera is placed near the beam pipe viewport to capture the ground truth beam image, which will later be used for model training. A beam splitter is used to divide this initial beam image into a secondary optical path, which is shrunk and coupled via a lens system (L1, doublet of 100 mm focal length, OBJ1, microscope objective). After that, the signal is transmitted by a ⌀1500 µm step-index MMF. At the distal end of this fiber, another set of lenses is used to decouple the light signal from the MMF (OBJ2, microscope objective, L2, doublet of 100 mm focal length). This output light is eventually projected onto the second CMOS camera, where the fiber output speckle pattern is recorded and used as the input of the model.

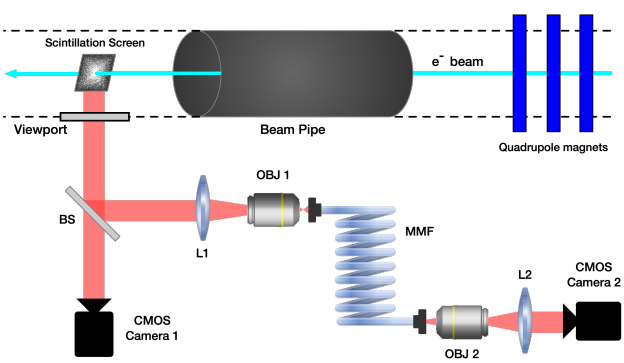

Figure 2: Schematic diagram of experimental data collection of ground truth images and MMF speckle patterns.

Approximately 6000 pairs of single-channel image samples were gathered and preprocessed (resized and normalized). we used the previously described method to calculate the four normalized beam parameters as the labels, and subsequently remove abnormal samples. The dataset was then randomly sampled into training, validation, and test sets with a ratio of 8:1:1. The model was trained exclusively using the training set, with performance evaluation conducted through inference on the test set. Due to the specific nature of this experiment, data augmentation was not utilized, as the shift or rescale of labels does not linearly correspond to the shift and rescale of fiber output patterns.

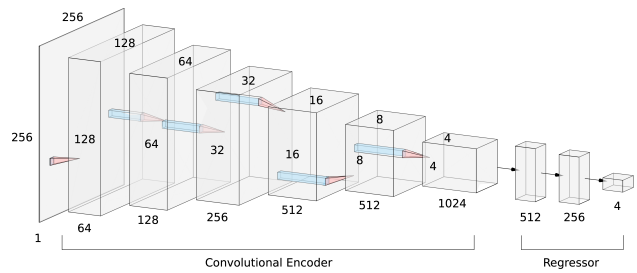

Figure 3: Convolutional encoder-regressor network.

On the model side, the neural network used for this experiment features a convolutional encoder for input processing, as Fig. 3 illustrates. Convolutional neural networks (CNN) are generally considered the best feature extractor for image data [6]. They reduce the dimension of the input image while increasing the number of abstract representations layer by layer, with each kernel searching for specific patterns. The encoder branch comprises a total of 6 layers of convolution blocks, each halving the dimension of the single channel 256 × 256 input image down to 1024 total 4 × 4 representation maps. Pooling is not used; instead, reduction is achieved by a stride-two Conv2d layer which could preserve more information. Each layer includes batch normalization to prevent gradient vanishing. To avoid over-fitting caused by providing the network with excessive pixel-level details, there are no skip connections in the network. Compressing the image data into a relatively abstract level of representation is beneficial, the encoder branch extracts the invariant from the image and is more resistant to noise and environmental changes. The second part of this network is a regressor, consisting of only three layers of fully connected neurons to extract the final beam parameters from the encoder's output latent space. The network is trained using an appropriate learning rate, Adam optimizer, and 50 training epochs, during which the validation loss stabilizes. Early stopping is also applied. The training employs a Mean Squared Error (MSE) loss function, defined as:

$$\text{Loss}_{\text{MSE}} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{m}\sum_{j=1}^{m}(y_{ij} - \hat{y}_{ij})^2 \quad (1)$$

Where n represents the number of samples in a batch, and m denotes the length of the output vector; in this case, $m = 4$ since four beam parameters are being predicted. y stands for the actual values, and $\hat{y}$ represents the predicted values. Using this MSE, errors for four normalized beam parameters are summed and averaged to determine the value that will backpropagate and update the network weights.

Content from this work may be used under the terms of the CC BY 4.0 licence (© 2024). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI.

ISBN: 978-3-95450-249-3 ISSN: 2673-5350 doi: 10.18429/JACoW-IBIC2024-TUP54

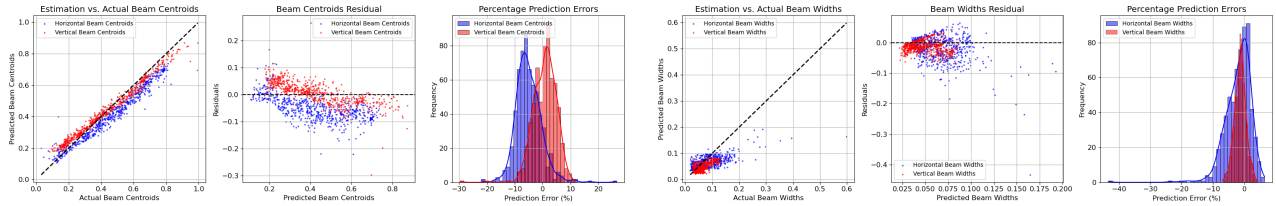

Figure 4: Test set prediction statistics on beam centroids and beam widths.

TRANSVERSE BEAM PARAMETERS REGRESSION

Previously, Generative Adversarial Network (GAN) had been used to reconstruct the 1:1 detailed beam image [7] and subsequently use this reconstructed image to calculate the beam parameters. This led to the idea of developing an end-to-end system that both extracts information from the fiber output and estimates beam parameters in a single neural network. The advantage of this approach is that without the image decoder branch, the model size is nearly halved, and the training and inference time is correspondingly reduced. Figure 4 shows the evaluation results for beam parameters estimation. About 600 test samples were input into the trained model. The left three graphs display the beam centroids estimation. The first one indicates the correlation between predictions and the ground truth value. The second and third illustrate the distributions of residual and prediction errors. Similarly, the right three images are beam width prediction results. The normalized Root Mean Square Error (RMSE) is less than 0.069 for all four parameters and has an average prediction RMSE of 0.043, indicating a good prediction performance. A slightly biased prediction toward smaller values can be observed. This is likely due to the unbalanced distribution of the dataset. As Table 1 shows, γ_1 represents the skewness of data distribution, where positive skewness suggests more data are concentrated toward the lower end (left), and negative skewness indicates more data are concentrated toward the higher end (right). Although the horizontal centroid has a negative γ_1, the mean value μ for each beam centroid is lower than 0.5, also indicating a slight shift toward smaller values.

Table 1: Dataset Distribution Statistics (Normalized Beam Parameter Values)

Beam Parameters	μ	σ	γ_1
Centroid (Horizontal)	0.471	0.193	-0.118
Centroid (Vertical)	0.421	0.203	0.486
Width (Horizontal)	0.087	0.054	1.776
Width (Vertical)	0.060	0.024	0.866

DISCUSSION AND FUTURE PLANS

This paper presents a deep convolutional encoder-regressor network for the direct estimation of the transverse beam parameters from MMF output patterns. A maximum test set prediction RMSE of 0.069 is achieved. The future plan is to solve the prediction shift and further increase the accuracy by simulating a high-variance simulation dataset using advanced imaging devices, such as the Digital Micromirror Device (DMD), to load the pattern and pretrain the network.

Furthermore, the perturbation of fiber could also be systematically studied. It is widely known that environmental changes such as temperature, fiber stress, and fiber geometry deformation can alter the refractive index, thus changing the fiber transmission properties. This makes the neural network trained on data from a specific configuration less effective on other fiber configurations. Potential methods such as transfer learning or more complex models with better generalization capabilities like Vision Transformer need to be designed and evaluated.

Another interesting observation during training is that with the same model, dataset and training settings, the model produces slightly different inference distribution each time. This variability is likely due to factors such as random initialization and the stochastic nature of the training process. Ensemble methods like Bootstrap Aggregating can be applied in the future to improve the stability of the model.

ACKNOWLEDGEMENTS

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

REFERENCES

[1] *Multimode Fiber Specifications*, Thorlabs, 2024. `https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id=10417`

[2] 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*, vol. 15, no. 2, pp. 524-612, Jun. 2023. `doi:10.1364/aop.484298`

[3] D. Psaltis and C. Moser, "Imaging with multimode fibers," *Opt. Photonics News*, vol. 27, no. 1, pp. 24-31, Jan. 2016. `doi:10.1364/opn.27.1.000024`

[4] 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.*, vol. 104, no. 10, p. 100601, Mar. 2010. `doi:10.1103/physrevlett.104.100601`

Content from this work may be used under the terms of the CC BY 4.0 licence (© 2024). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI.

ISBN: 978-3-95450-249-3 ISSN: 2673-5350 doi: 10.18429/JACoW-IBIC2024-TUP54

[5] G. Trad and S. Burger, "Artificial Intelligence-Assisted Beam Distribution Imaging Using a Single Multimode Fiber at CERN", in *Proc. IPAC'22*, Bangkok, Thailand, Jun. 2022, pp. 339–342. doi:10.18429/JACoW-IPAC2022-MOPOPT041

[6] Y. Zhang, "A better autoencoder for image: Convolutional autoencoder," in *Proc. ICONIP17-DCEC*, 2018. http://users.cecs.anu.edu.au/Tom.Gedeon/conf/ABCs2018/paper/ABCs2018_paper_58.pdf

[7] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul. 2017, pp. 1125-1134. doi:10.1109/cvpr.2017.632

Content from this work may be used under the terms of the CC BY 4.0 licence (© 2024). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI.