Research Article Vol. 34, No. 9 / 4 May 2026 / Optics Express 17148
Optics EXPRESS

# Efficient phase identification in coherent beam combination using interpretable deep learning

FEDOR CHERNIKOV,* YUNHUI XIE, JAMES A. GRANT-JACOB, YUCHEN LIU, MICHALIS N. ZERVAS, AND BEN MILLS

*Optoelectronics Research Centre, University of Southampton, UK*
*fc1e21@soton.ac.uk*

**Abstract:** Coherent beam combination (CBC) offers significant power scaling beyond the capabilities of individual fiber lasers, but its effectiveness is heavily reliant on precise phase stabilization. Recent advancements in deep learning have shown potential for phase retrieval from interference intensity patterns in a single step. However, the interpretability of deep learning models and the optimal positioning of the imaging system remain unresolved challenges. In this study, we employ a spatial light modulator to emulate a CBC system and systematically investigate the phase prediction accuracy at various axial positions of the imaging system. We demonstrate that phase-retrieval efficiency can be substantially enhanced by identifying regions of the interference pattern with higher phase sensitivity. This approach enables a significant reduction in input size, allowing for the use of a lightweight fully connected neural network to achieve a phase prediction error of approximately λ/60, with a pure fully connected neural network inference rate of approximately 35 kHz for a 7-beamlet hexagonal close-packed array.



## 1. Introduction

CBC is a promising technique for scaling the power level of fiber lasers beyond the inherent limitations that constrain the single-source performance [1–3]. By combining the outputs of multiple diffraction-limited fiber lasers such that their fields interfere constructively, CBC concentrates the combined optical power predominantly into a single spot in the far field [4,5]. This facilitates power scaling while maintaining high beam quality. However, the combined intensity distribution is highly sensitive to the relative phases of the constituent fiber lasers. Therefore, phase stabilization is essential to maintain constructive interference and ensure consistent performance over time [6,7]. Traditionally, phase control methods commonly rely on iterative feedback mechanisms which, while effective for power optimization, can limit the flexibility of the system [8–10]. This poses a challenge for applications such as adaptive manufacturing, where the additional capability of CBC of phase-only beam shaping could be exploited to dynamically tailor beam profiles to suit different processing requirements [11–14].

Recent advancements in deep learning have demonstrated significant potential for retrieving phase information directly from interference patterns in a single step [15–17]. In these approaches, Neural Networks (NNs) are trained to process camera observations of intensity distributions and predict the corresponding phase values. This circumvents the limitations of iterative phase retrieval methods and offers a simplified route to dynamic phase control in CBC systems. However, key challenges remain regarding the interpretability of these deep learning models and the optimal positioning of the imaging plane within the optical setup. It has been observed that the accuracy of the phase inference improves when the imaging system (i.e., the camera) is placed at out-of-focus positions, as these planes are not subject to the Fourier symmetry inherent to the focal plane, thereby enabling a more direct, one-to-one mapping between combined intensity

#590233 https://doi.org/10.1364/OE.590233
Journal © 2026 Received 15 Jan 2026; revised 19 Mar 2026; accepted 19 Apr 2026; published 30 Apr 2026

Optics EXPRESS

distribution and the corresponding phase values [18,19]. This raises an important question: are all out-of-focus planes equally suitable for phase inference, or is there an optimal plane where phase-relevant features are most distinguishable?

A different concern arises regarding the operational speed of the deep-learning-enabled phase information retrieval. While deep learning phase inference has demonstrated promising performance for phase inference in CBC systems and can achieve operational bandwidth of around 1 kHz [20], further increases in operational speed are desirable, particularly for applications involving dynamic beam shaping or steering. Recently, a simulation-trained lightweight CNN operating at an inference rate of approximately 14 kHz was demonstrated [21]. In this approach, however, the deep learning models was trained on full camera observations that encompass the entire interference pattern. However, do all areas of this pattern contribute equally to the phase inference task? Intuitively, peripheral areas or those with minimal spatial variation, under the effects of phase fluctuations, are likely to carry less phase-relevant information than areas, exhibiting strongly varying interference fringes. Identifying these information-rich areas could reduce input dimensionality without sacrificing phase prediction accuracy. This would not only accelerate inference but also enhance deep learning model interpretability by revealing which features the NN relies on most. These considerations give rise to an important question: is it necessary to provide NN with entire intensity distribution, or can an accurate phase inference be achieved using only selected areas of this intensity distribution?

In this study, we emulate a CBC system using an SLM to generate interference patterns corresponding to three array configurations: a two-beamlet configuration, a hexagonally close-packed array of seven beamlets, and a square array of nine beamlets. Using this setup, we first investigate how capturing intensity distributions at different positions along the optical axis affects phase prediction performance using a Convolutional Neural Network (CNN). Secondly, we examine how phase prediction changes when only specific, strategically selected areas of the interference pattern are used as inputs to the CNN for phase inference. Finally, we demonstrate that, by leveraging these identified informative areas and reducing the input size, the relatively computationally intensive CNN can be replaced with a lightweight and fast-running Fully Connected Neural Network (FCNN). We show that, at certain imaging positions and using carefully selected areas of the interference pattern, the FCNN, trained on observing the identified informative areas, can match the accuracy of a CNN trained on the full intensity distribution (with a phase prediction error of approximately $\lambda/60$ for 7-fibre hexagonal close-packed array), while significantly improving the pure FCNN inference speed (i.e., excluding data acquisition and transfer), reaching approximately 35 kHz.

## 2. Experimental setup

In this work, a Spatial Light Modulator (SLM, Thorlabs EXULUS-HD1/M, $1920 \times 1080$ pixel resolution, 6.5 µm pitch size) was used to emulate a CBC system with varying number of beamlets and to enable precise control over their relative phases. The SLM was programmed to generate three distinct patterns: a pair of circles, a hexagonally close-packed array of seven circles, and a square array of nine circles, each corresponding to a possible output configuration of fiber lasers in a CBC system. The two-beamlet array was included as it provides an intuitive and easy-to-follow example for evaluating phase inference performance. The seven-beamlet array was chosen because it represents the widely adopted hexagonal close-packed configuration, valued for its high fill factor. Lastly, the nine-beamlet array served as an example of an alternative arrangement, illustrating that the methods presented in this study are not limited to any particular fiber array geometry.

A vertically polarized light from a HeNe laser (Thorlabs, HRS015B, 632.992 nm central wavelength, 1.2 mW output power), selected for its high temporal coherence, narrow linewidth, and polarization stability, was expanded and directed onto the SLM panel at the normal angle of

Optics EXPRESS

incidence. To introduce controllable phase delays between simulated beamlets, a horizontally oriented blazed grating was applied within each circular region. The relative lateral displacement of these gratings imparted a programmable phase shift between simulated fibers [17]. Due to the diffractive nature of the phase modulation, each circular region produced multiple higher-order diffraction outputs. The zeroth-order reflection, which corresponds to specular reflection from the unmodulated background, did not carry useful interference information. In contrast, the first diffraction order contained the spatially separated and phase-modulated beamlets required for CBC emulation. To isolate this diffraction order, a 4f optical system was employed, with a spatial filter placed at the Fourier plane to block the intense but uninformative zeroth-order component (and other higher-order components except for the first-order component). This enabled clean selection of the first-order beamlets and enhanced the contrast of the resulting interference patterns, facilitating analysis of interference effects arising from the phase-controlled beamlet array.

To capture how the interference pattern evolves with propagation, the selected first-order diffraction outputs were then directed through a focusing system onto a movable camera. Specifically, after modulation on the SLM and spatial filtering of the first diffraction order, the resulting beam was directed onto a camera (Basler a2A4504-18umBAS) mounted on a motorized translation stage (ZABER X-LSQ450B-E01), as illustrated in Fig. 1(a). A convex lens with a 50 cm focal length was used to focus the beam, and the motorized stage, with a travel range of 45 cm, enabled the camera to capture intensity distributions at various positions along the optical (z) axis.

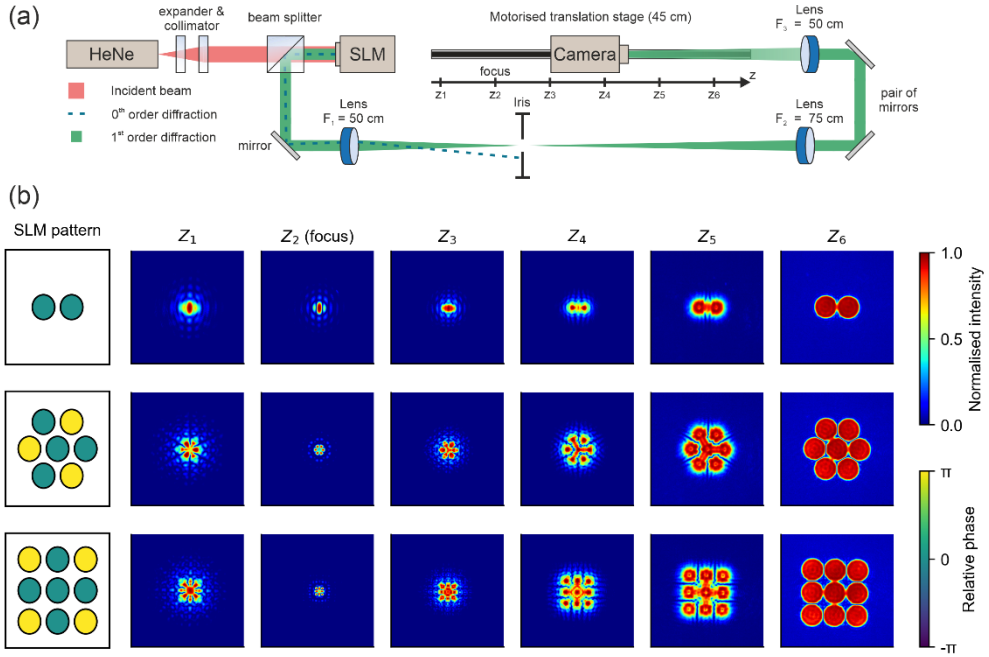


**Fig. 1.** (a) Schematic of the experimental setup (b) Representative camera observations of intensity distributions for 2-, 7-, and 9-beamlet CBC array at various positions along optical axis. Here, $z_6$ corresponds to a plane closer to the SLM (i.e., simulated CBC head), $z_2$ is the focal plane, and $z_1$ lies beyond the focus (green and yellow colors in the CBC arrays correspond to 0 and π phase).

Figure 1(b) presents examples of camera observations of the intensity distributions produced by 2-, 7-, and 9-beamlet CBC arrays, recorded at six equally spaced positions along the optical

Optics EXPRESS

axis ($z_1$-$z_6$), where $z_6$ is closest to the focusing lens $F_3$, $z_2$ corresponds to the focal plane, and $z_1$ is a plane located beyond focus. Only one observation point beyond the focal plane was included, as the primary focus of this study is on the near-field to far-field region typically employed in CBC phase control systems. At $z_6$, the individual beamlets are clearly distinguishable with interference predominantly occurring in the areas between them. As the beamlets propagate towards the focus, they increasingly overlap, resulting in more complex intensity patterns. Intuitively, the optical axis can be divided into two distinct regions. In the first region, closer to the focusing lens (i.e., $F_3$), interference occurs predominantly between adjacent beamlets. The resulting intensity distribution can be interpreted as a superposition of multiple two-adjacent-beamlet interference patterns. In the second region, as the beamlets propagate toward the focal plane, their spatial overlap increases, and more than two beamlets simultaneously contribute significantly to the interference pattern. This raises a key question: does the complexity of the interference pattern, whether dominated by a simple two-adjacent-beamlet contributions or multi-beamlet interactions, influence the accuracy of phase inference?

## 3. Data collection and neural network training

To evaluate how phase inference accuracy depends on the position of the imaging plane along the optical axis, the training data were collected using the following procedure. For each of the three CBC configurations, one beamlet was designated as the reference and its phase was fixed at $0\pi$ rad. The phases of the remaining beamlets were independently randomized, drawn from a uniform distribution over the interval $[-\pi, \pi)$. This resulted in 1, 6, and 8 varying phase values for the 2-, 7- and 9-beamlet arrays, respectively. After applying these phase values to the SLM, the resulting interference pattern was recorded at a specific imaging plane along the optical axis.

Each captured camera observation was cropped to include the entire inference pattern and subsequently resized to a $256 \times 256$ 8-bit grayscale image. Although the resize factor varied depending on the axial position of the camera (since the interference pattern is more spatially compact near the focal plane), the goal was to maintain a consistent image resolution and ensure that the interference pattern occupied a comparable area across all cropped camera observations. Each recorded intensity distribution, along with its corresponding set of relative phase values, constituted one training pair.

In total, 5,000, 20,000, and 25,000 training pairs were collected for 2-, 7-, and 9-beamlet configurations, respectively, at each imaging position. While the dataset sizes differ across CBC configurations, this does not impact the analysis, as the primary objective is to compare phase inference accuracy between imaging planes within a given CBC configuration, rather than across different CBC fiber arrays. Thus, absolute error values may vary by configuration, but performance trends as a function of imaging position remain directly comparable.

To investigate not only how phase inference accuracy depends on the position of the imaging plane, but also which areas of the interference pattern are most informative for phase inference, the following method was applied. Specifically, the objective is to evaluate whether accurate phase inference is achievable using only specific areas of the interference pattern.

To isolate these areas, binary masks based on the spatial distribution of intensity variation were first constructed. For a given axial position and CBC configuration, the pixel-wise standard deviation was computed using all collected camera observations. This standard deviation map reflects how much the intensity of each pixel fluctuates as a function of random phase changes, as visualized in Fig. 2(a).

The rationale behind this approach is as follows: assuming the only source of intensity variation is the changing phase values, pixels with higher standard deviations are more responsive to these phase changes and therefore are likely to contain more phase-relevant information. For example, at position $z_6$, as shown in Fig. 2(b), the standard deviation maps for 2-, 7-, and 9-beamlet CBC arrays show higher variation in areas between the neighboring beamlets. This is expected, as at

Optics EXPRESS

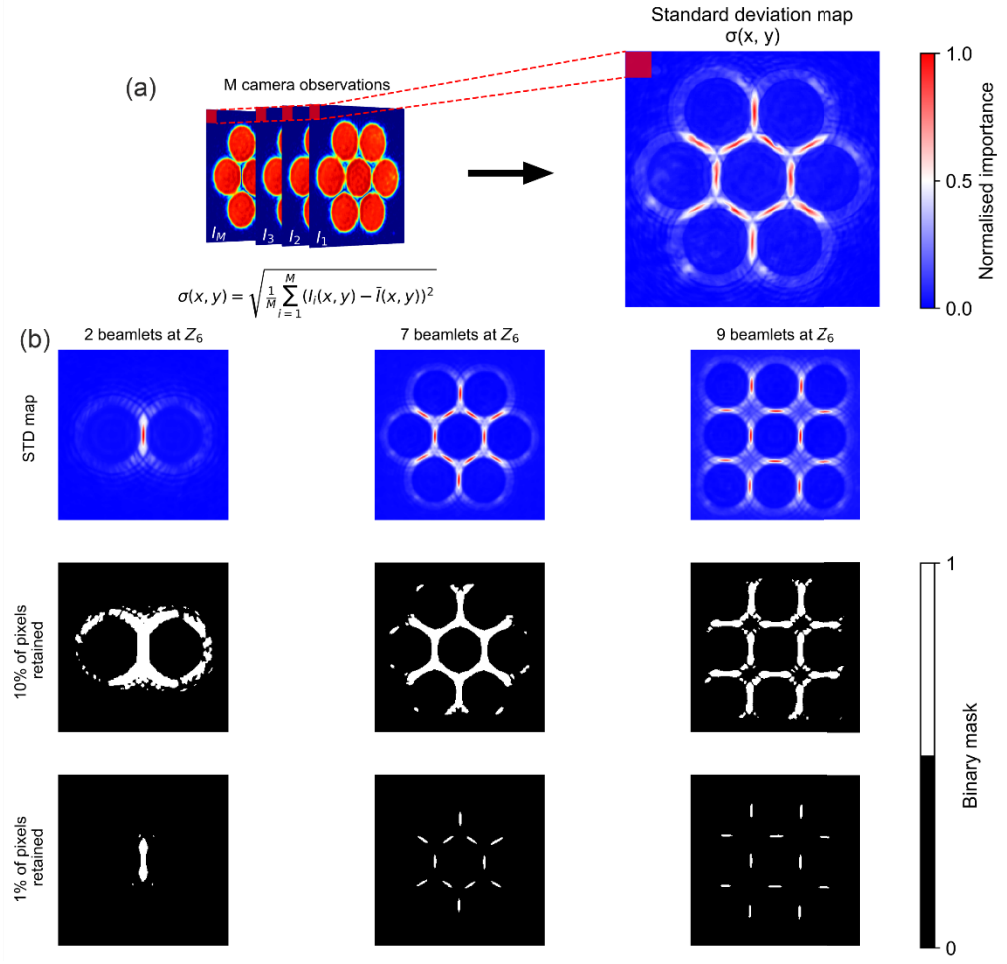


**Fig. 2.** (a) Schematic illustration of the construction of a pixel-wise standard deviation map, using the 7-beamlet CBC array at position $z_6$ as an example (b) Standard deviation maps for 2-, 7-, and 9-beamlet CBC arrays at positions $z_6$, along with corresponding binary masks retaining 10% and 1% of the total number of pixels.

this propagation distance the beamlets are just beginning to overlap, making the inter-beamlet areas the primary zones of interference.

Once the standard deviation map was computed, binary masks were generated by applying a threshold to retain only the pixels considered most informative. Specifically, pixels with a standard deviation value above the threshold were retained (i.e., assigned value of 1), while those below the threshold were excluded (i.e., assigned value of 0). Figure 2(b) shows examples of such binary masks, where the thresholds were chosen to retain approximately 10% (i.e., ~6,500 pixels) and 1% (i.e., ~650 pixels) of the total number of pixels of camera observation.

These binary masks were then applied to the respective datasets. For instance, the 9-beamlet mask at position $z_6$ was applied via pixel-wise multiplication to each camera observation in the 9-beamlet dataset acquired at the same position $z_6$. This operation produced a new dataset in which each camera observation retained only the selected by mask high-variation pixels (i.e., pixel values remained as they were), while the value of the remaining pixels was set to zero. This process was repeated for each CBC configuration at every axial position. As a result, for each

Optics EXPRESS

z-position and CBC configuration, three datasets were obtained: one containing the full camera observations, one retaining only the top 10% of pixels based on standard deviation, and one retaining only the top 1% of pixels.

Thus far, the training datasets consist of either full or masked camera observations (i.e., $256 \times 256$ 8-bit images) and their corresponding phase values (represented as vectors). The NN used for phase inference is a CNN known as MobileNetV3-small [22], selected for its optimized balance between accuracy and computational efficiency. Several modifications were made to the original MobileNetV3-small architecture. First, the input image size was adjusted from the original $224 \times 224$ to $256 \times 256$ to match the dimensions of the input data. Second, the output layer was modified from the original shape $N \times 1000$ to $N \times 1$, $N \times 6$, or $N \times 8$, corresponding to the number of phases to be inferred in the 2-, 7-, and 9-beamlet CBC systems, respectively. Finally, due to the periodic nature of phase values, the loss function was adapted accordingly. The modified loss function takes the form [23]:

$$\text{Loss} = (\sin(\phi) - \sin(\varphi))^2 + (\cos(\phi) - \cos(\varphi))^2$$

where $\phi$ is the predicted phase and $\varphi$ is the ground-truth phase. This formulation ensures that the loss correctly accounts for the phase wrapping at $2\pi$. Each of the datasets described above was a split into training and validation sets using a 90/10 ratio. A separate CNN, with randomly initialized parameters, was trained on each training set without any pre-training. All models were trained for 200 epochs using the Adam optimizer [24] with a learning rate of $6 \times 10^{-5}$ and a batch size of 32. For clarity in subsequent discussion, CNN models trained on full camera observations are denoted as $\text{CNN}^{100}$, those trained on inputs retaining 10% of pixels as $\text{CNN}^{10}$, and those retaining only 1% as $\text{CNN}^{1}$.

## 4. Neural network evaluation

Despite being trained on datasets acquired at different axial positions and with different fractions of the image retained, all models converged at approximately the same stage of training, reaching a stable plateau after a similar number of epochs, For consistency, after the final epoch, the trained model weights were saved, and each CNN was evaluated on the respective validation set, which contained previously unseen data. The performance was quantified using the mean phase prediction error, defined here as the mean absolute error between the predicted relative phases and the ground truth values, averaged across all the beamlet channels. Figure 3(a), Fig. 3(b), and Fig. 3(c) present the validation results of the CNN validation for 2-, 7-, and 9-beamlet CBC arrays, respectively. Above each graph, standard deviation maps are shown for all six axial positions considered. These maps were used to generate binary masks: the top inset displays the masks retaining 10% of the total number of pixels, while bottom inset shows masks retaining only 1%. Each graph contains three curves corresponding to the phase prediction error of the CNN trained on full camera observation (i.e., $\text{CNN}^{100}$, shown in blue), as well as those trained on masked observations retaining 10% (i.e., $\text{CNN}^{10}$, shown in orange) and 1% (i.e., $\text{CNN}^{1}$, shown in green) of the pixels.

For 2-beamlet CBC array, as shown in Fig. 3(a), $\text{CNN}^{100}$ exhibited a nearly flat phase prediction error across all axial positions, with mean phase error values ranging from approximately 0.03 to 0.06 rad for positions $z_1$ to $z_6$. The performance of $\text{CNN}^{10}$ and $\text{CNN}^{1}$ was comparable, maintaining prediction errors below 0.08 rad. Notably, no increase in error was observed at the focus (i.e., at $z_2$), which can be explained by the inherently unambiguous mapping between the interference pattern and the phase difference in the two-beamlet case (i.e., when the phase spans the full $[-\pi, \pi]$ range). Overall, this simple two-beamlet case demonstrates that phase prediction accuracy shows no strong dependence on the axial position of the imaging system, and that accurate phase inference can be achieved even when a substantial portion of the input data is masked.

Optics EXPRESS

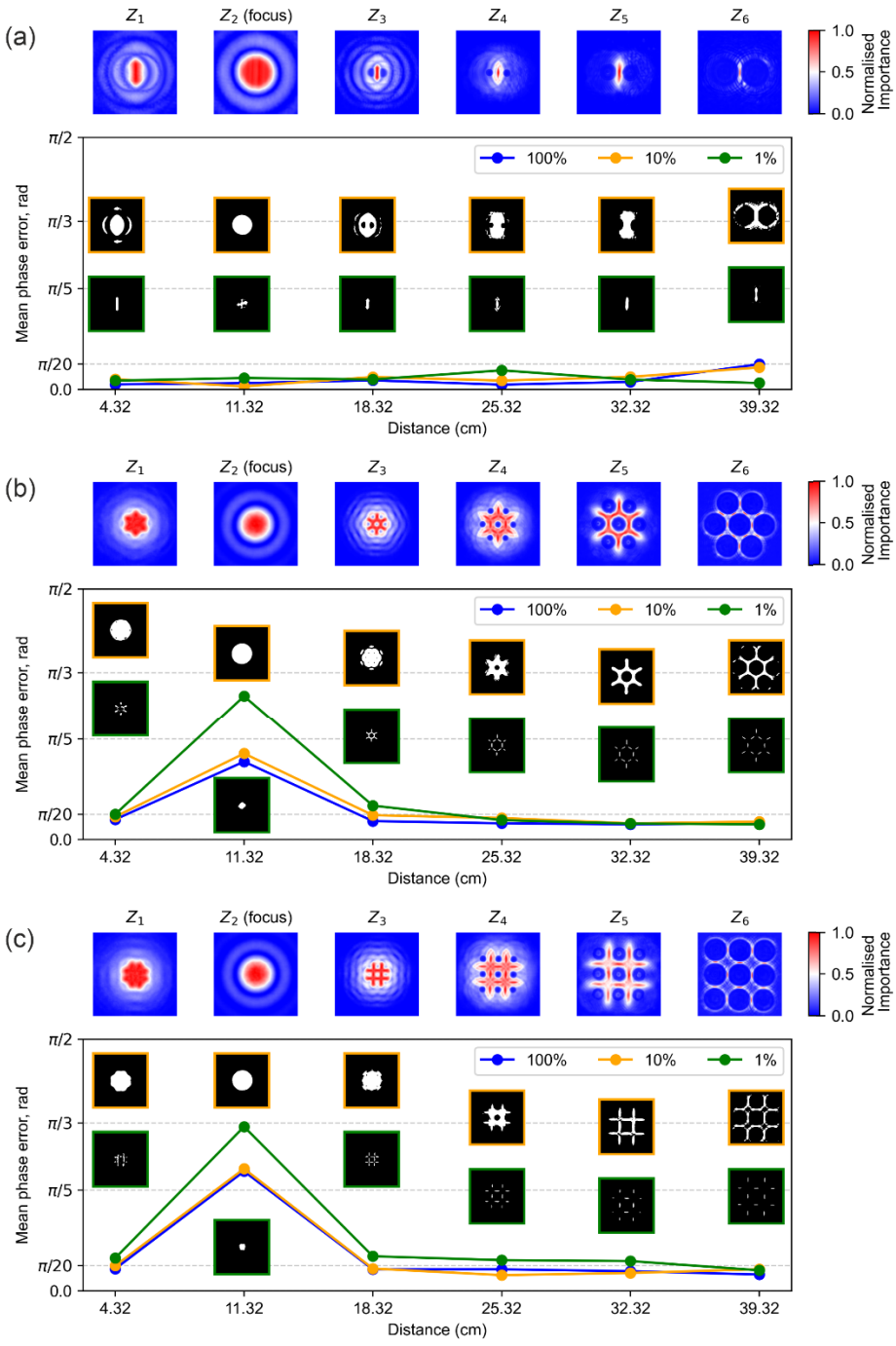


**Fig. 3.** Phase prediction error as a function of imaging system position when the CNN was trained on full (blue), 10%-masked (orange), and 1%-masked (green) camera observations for (a) 2-, (b) 7-, and (c) 9-beamlet CBC arrays. The top portion of each plot shows the corresponding standard deviation map, while the insets illustrate the binary masks used: the upper displays 10% mask and the lower inset the 1% mask.

Optics EXPRESS

For 7-beamlet CBC array, as shown in Fig. 3(b), the $CNN^{100}$ maintained consistent phase prediction performance across most axial positions, with phase prediction errors around 0.1 rad at $z_1$, $z_3$, $z_4$, $z_5$, and $z_6$. However, a pronounced peak in phase prediction error emerged at the focal position $z_2$, where the prediction error rose sharply to 0.48 rad. This behavior is a result of the Fourier symmetry inherent in the interference pattern at the focal plane, which leads to ambiguity in phase-to-pattern mapping. The focal-plane error spike was a consistent feature across both $CNN^{10}$ and $CNN^{1}$, reflecting the intrinsic phase ambiguity at this position. For non-focal positions, $CNN^{10}$ and $CNN^{1}$ generally performed comparably to $CNN^{100}$. Notably, position $z_5$ exhibited the smallest difference across the three models, with prediction errors of 0.094 rad ($CNN^{100}$), 0.101 rad ($CNN^{10}$), and 0.099 rad ($CNN^{1}$), suggesting that this plane provides a robust signal for phase inference even with substantially reduced input information.

For 9-beamlet CBC array, shown in Fig. 3(c), the $CNN^{100}$ demonstrated stable phase prediction across most imaging planes, with errors generally below 0.14 rad. However, a substantial spike was observed at the focal plane $z_2$, where the prediction error increased to 0.75 rad. This behavior is consistent with previous observations and is attributed to the Fourier symmetry at focus, which introduces ambiguity in mapping intensity to phase. $CNN^{10}$ performed comparably to $CNN^{100}$ across most positions. $CNN^{1}$, while still following the overall trend, showed slightly higher prediction errors than $CNN^{10}$ at nearly all positions. This suggests that while significant input masking (i.e., down to 1% of pixels) still preserves much of the relevant phase information, some degradation in accuracy becomes noticeable in more complex interference configurations such as the 9-beamlet array.

Overall, these results show that for the CBC arrays with symmetrical fiber structure the phase prediction error does not depend strongly on the position of the imaging system as long as it is outside the focal plane. The results also demonstrate that $CNN^{1}$ and $CNN^{10}$ can produce nearly equivalent to $CNN^{100}$ inference quality outside the focal plane, and that even with as little as 1% of the input preserved, meaningful phase retrieval remains achievable, albeit with a modest drop in accuracy. Notably, in several cases the masked-input models exhibit marginally lower prediction error than $CNN^{100}$ most prominently at positions closer to the focusing lens; this behavior is likely attributed to the elimination of redundant or weakly informative pixels, which effectively improves the signal-to-noise ratio of the input representation and can enhance generalisation performance. In principle, prediction accuracy could be further improved through dataset-specific hyperparameter tuning. However, since the primary objective of this study is to evaluate relative trends in phase prediction performance across imaging positions and masking levels, maintaining consistent settings ensures a fair basis for such comparisons. To provide additional insight into the spatial regions of the diffraction pattern that inform the CNN's predictions, an interpretability analysis based on four attribution methods is provided in Supplement 1.

## 5. Simulated detectors

While the results presented thus far demonstrate that accurate phase inference can be achieved using only specific areas of the interference pattern, a practical challenge remains. Although fewer non-zero (i.e., informative) pixels are retained, the input dimensionality in the CNN architecture remains fixed at $256 \times 256$. Consequently, the reduction of non-zero pixels does not translate into any improvement in inference speed. This raises an important question: how can the reduced pixel count be leveraged to improve computational efficiency?

To move beyond the fixed-size inputs (i.e., $256 \times 256$ image) and take advantage of the reduced pixel count, the concept of a "simulated detector" was introduced. A "simulated detector" (hereafter referred to as a detector) is defined as a square region of $N \times N$ pixels within the camera observation, where the pixel intensities are summed to yield a single value (i.e., detector reading). While this does not replicate the physical characteristics of a true photodetector, it mimics its core function: spatial integration over a finite area. The motivation behind this abstraction is to

Optics EXPRESS

simulate the effect of spatially aggregated measurements, allowing us to investigate whether a sparse set of such detectors, strategically placed in informative areas of the interference pattern, can yield sufficient information for accurate phase inference. The overall effectiveness of this approach is therefore shaped not only by how detectors are defined, but also, critically, by their spatial arrangement, the number of detectors deployed, and their size.

To establish a systematic approach, we begin by examining the interference patterns at position $z_6$. At this axial position, the beamlets begin to interfere, and the standard deviation map indicates that the most sensitive pixels to phase changes lie in the area between the neighboring beamlets. At $z_6$, where interference primarily occurs between two neighboring beamlets, variations in relative phase lead to a shift of the interference fringes along the line connecting the centers of the beamlets. This suggests that, for effective phase inference, monitoring intensity variations along this axis is most appropriate. Although there is no strictly defined "center" of a beamlet as it propagates along the optical axis, the standard deviation map effectively captures the spatial structure of the initial CBC array, as shown in Fig. 3. This allows us to identify points that are functionally equivalent to the centers of the beamlets, and, in turn, to define the axes connecting them. Having identified these axes, the next step is to determine the optimal placement of detectors along them to extract the most phase-informative signal.

Figure 4(a) presents an example of a camera observation of the intensity distribution of a 7-beamlet array at axial position $z_6$, with a black square indicating the region shown in more detail in Fig. *4*(b). The black line in Fig. 4(b) lies along the axis connecting the centers of the central and leftmost beamlets, with three black squares marking the positions of simulated detectors placed along this line. Figure 4(c) shows the corresponding intensity profile along the line from Fig. *4*(b), with dashed lines indicating the boundaries of the simulated detectors. At this axial position and the particular phase relation between the corresponding beamlets, the intensity distribution exhibits a single trough, indicating a local decrease in intensity between two beamlets. To monitor such variations in intensity and thereby infer the phase of one beamlet relative to another, a group of three detectors was manually positioned. The central detector (i.e., 2) was positioned near the trough of the intensity profile, while adjacent (i.e., 1 and 3) are placed symmetrically on either side. The size and spacing of the detectors were selected such that they capture the regions of steepest intensity gradient, allowing such arrangement of detectors to effectively track fringe displacement. For axial position $z_6$, the detectors were defined as $3 \times 3$ pixel regions, with approximate 4-pixel separation between their centers (depending slightly on their orientation). This arrangement offers a practical compromise, capturing sufficient spatial detail for effective phase tracking while preserving computational efficiency.

The optimal placement of each group of detectors can be conveniently visualized using previously defined 10% binary mask. Figure *5*$a_1$ (position $z_6$) shows this binary mask for 7-beamlet CBC array, with red squares indicating six regions where each detector group was positioned. Figure *5*$a_2$ ("Detectors") schematically illustrates the internal arrangement of detectors within a red square, with their alignment following the lines connecting the centers of the interfering beamlets. This visualization highlights that detectors are placed primarily within the high-variance regions defined by mask, ensuring that each group lies within areas exhibiting strong spatial changes in intensity. As a result of this detector approach, each camera observation for the 7-beamlet CBC array at $z_6$ can be represented by 18 readings, each corresponding to the sum of pixel intensities within a $3 \times 3$ pixel region.

The detector placement strategy developed for position $z_6$ in the 7-beamlet CBC array was subsequently extended to other axial positions (i.e., $z_2$ through $z_5$) as well as to the 9-beamlet CBC array. At positions $z_4$ and $z_5$, the intensity profiles observed along the lines connecting the centers of beamlets predominantly exhibited single peak/trough structures similar to those at $z_6$. Conveniently, at $z_4$ and $z_5$, the structure of the binary masks closely resembles that observed at $z_6$, exhibiting a comparable grid-like pattern, as illustrated in Fig. *5*$a_1$ and Fig. *5*$b_1$. Therefore,

Optics EXPRESS

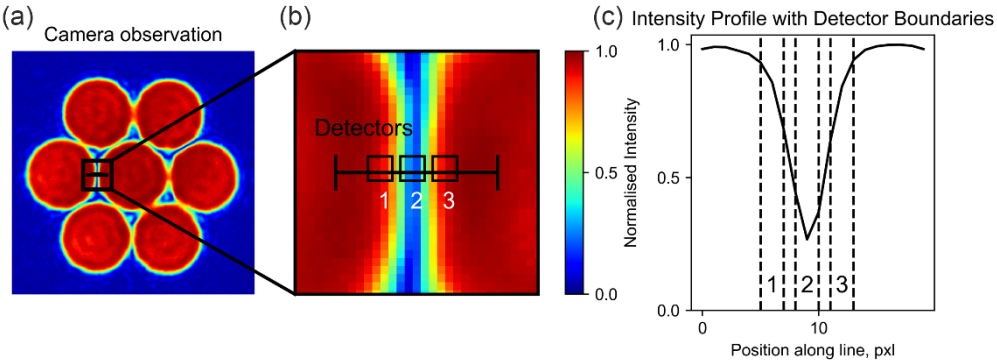


**Fig. 4.** Schematic of a detector placement choice (a) Example of camera observation of intensity distribution at axial position $z_5$, with a black square indicating the zoomed-in region shown in (b). The black line denotes the slice along which intensity is extracted, and the black squares represent the locations of simulated detectors. (b) Magnified view of the selected regions from (a), highlighting the slice and detector positions. (c) Intensity profile along the slice indicated in (b), showing spatial variation of intensity with dashed line illustrating the boundaries of each of three detectors.

detector group placement at these positions followed the same procedure as that established for $z_6$, with each group positioned within a distinct high-variance region of the binary mask. For $z_5$, both detector size and spacing remained identical to $z_6$ (i.e., $3\times3$ with ~4-pixel separation), while at $z_4$ the detector size was increased to $6\times6$ pixels with a center-to-center spacing of approximately 9 pixels to match the spatial scale of the pattern.

In contrast, at positions $z_2$ and $z_3$, the interference patterns no longer exhibit well-defined single peak/trough structures along the axis connecting centers. The standard deviation maps become more uniform, and the binary masks begin to lose their clear grid-like structure, appearing more uniform overall. Despite this, to maintain a consistent detector configuration and to preserve the one-detector-group-per-beamlet approach, detector regions were defined based on the binary masks. Detector groups were preferentially placed toward outer areas of the binary mask region, thereby avoiding excessive overlap in the central region and ensuring spatial diversity. While this approach may not represent the optimal detector placement for $z_2$ and $z_3$, it ensures structural consistency across all axial positions, facilitating deep learning models to learn from similarly structured inputs regardless of axial distance from the focus.

As a result, a total of 6 and 8 regions were defined at each z-position for 7- and 9-beamlet arrays, respectively (i.e., one per beamlet, excluding the reference beamlet). For the 9-beamlet array, 12 interference regions can in principle be identified between adjacent beamlets in the defocused diffraction pattern. However, only a minimal set of regions sufficient to uniquely recover the relative phases was used in this work. In an N-channel CBC system, the phase state can be fully described by N-1 independent phase differences with respect to a reference beamlet. Therefore selecting 8 regions for the 9-beamlet array provides a complete but non-redundant set of measurements for phase reconstruction. While additional detector regions could potentially improve the prediction accuracy, preliminary tests indicated that increasing the number of regions did not significantly improve the phase prediction accuracy or training convergence, while increasing the dimensionality of the input vector. With three detectors per region, each camera observation was thus represented by 18 and 24 values for the 7- and 9-beamlet configurations, respectively. This enabled the construction of new datasets in which each $256\times256$ camera observation was replaced with a compact vector representation. From a practical standpoint,

Optics EXPRESS

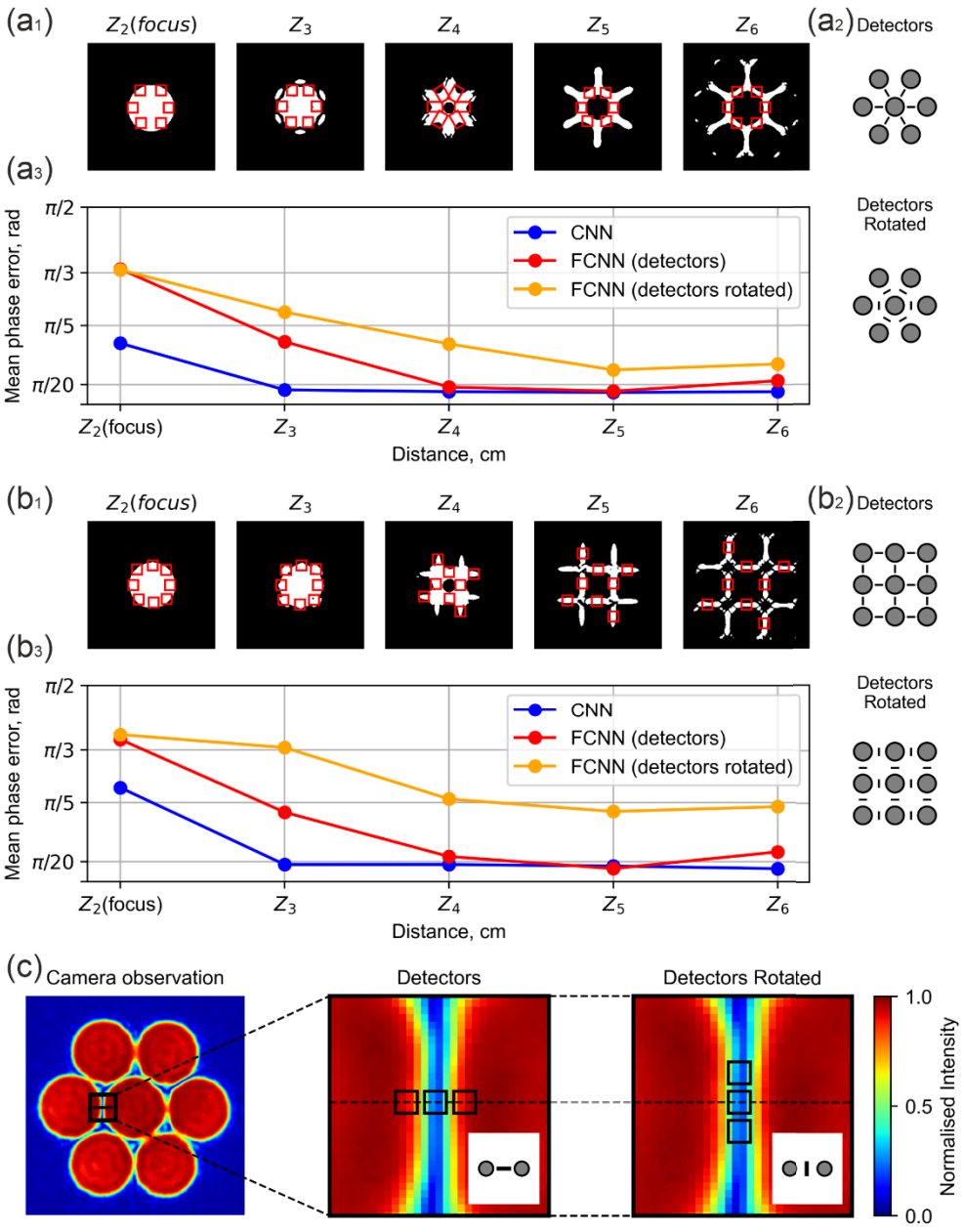


**Fig. 5.** Mean prediction error as a function of axial position for CNN100 (blue) and two FCNN models: one trained on detector dataset aligned along imaginary line connecting centers of the simulated beamlets (red), and another one trained on a dataset from detectors with rotated placement (orange). The top of each plot shows the 10% binary mask used to guide detector placement, with selected areas for detector positions highlighted as red squares. Results are shown for (a) 7-beamlet, and (b) 9-beamlet CBC array. (c) Schematic representation of the detector orientations; the two-beamlet insets illustrate the geometric relationship between the beamlets and the detectors for both the "Detectors" (aligned) and "Detectors Rotated" (orthogonal) configurations used in (a2) and (b2).

Optics EXPRESS

this enables replacing a CNN, which takes images as input, with a FCNN trained on 18- or 24-dimensional inputs and corresponding 6- or 8-dimensional phase outputs, respectively.

To assess the significance of the detector orientation, an alternative relative placement was evaluated, referred to as "Detectors Rotated" in Fig. 5$a_2$ and Fig. 5$b_2$. In this placement, three detectors were positioned orthogonal to the line connecting the centers of the beamlets, as demonstrated in Fig. 5c ("Detectors Rotated"). This contrasts with the original layout, where detectors were aligned along the line connecting beamlet centers, as shown in Fig. 5c ("Detectors"). To clarify the schematic representations used in Fig. 5$a_2$ and Fig. 5$b_2$, the insets in Fig. 5c provide simplified two-beamlet examples. These insets show the relative position between the group of detectors and the beamlets for both the "Detectors" and "Detectors Rotated" cases. Using this new relative position, additional datasets were constructed for positions $z_2$ through $z_6$ for both the 7- and 9-beamlet CBC arrays.

For the 7-beamlet array, the FCNN was implemented with an 18-32-64-128-256-6 architecture, utilizing an 18-dimensional input vector to represent individual detector readings and a 6-node output layer for the relative phases of the outer beamlets with respect to the central reference. For the 9-beamlet CBC array, the FCNN operated on a 24-dimensional input vector and produced 8 outputs. For each detector orientation (i.e., Detector and Detector Rotated), a separate FCNN was trained using the corresponding datasets across positions $z_2$ to $z_6$. Each dataset was split into training and validation subsets using a 90/10 ratio. Each FCNN model was trained using the Adam optimizer with a learning rate of $2 \times 10^{-4}$ for 60 epochs and a batch size of 8. After training, each model was evaluated on its respective validation set, and the mean phase prediction error was recorded to facilitate performance comparison between detector approach at different z-positions and detector orientations.

Figure 5($a_3$) (7-beamlet) and Fig. 5($b_3$) (9-beamlet) present a comparative analysis between the $CNN^{100}$ (i.e., trained on camera observations incorporating full interference pattern) and two FCNN, each trained on datasets constructed using detector-based approach. The FCNN differ in how their detectors are oriented: either along the axis connecting beamlets centers (Detectors, shown in red) or rotated orthogonally (Detector Rotated shown in orange). For both 7- and 9-beamlet CBC array, the FCNN trained on detectors aligned along the centers followed a consistent trend across axial positions: the highest phase prediction mean error occurred at $z_2$, followed by a minimum at $z_5$, and then a slight increase at $z_6$. In the 7-beamlet case, the mean errors were 1.07 rad at $z_2$, 0.10 rad at $z_5$, and 0.18 rad at $z_6$. For the 9-beamlet case, the corresponding errors were 1.13 rad, 0.10 rad, and 0.23 rad, respectively.

Importantly, at $z_5$, where interference is primarily between adjacent beamlets, the performance of FCNN approached that of the $CNN^{100}$. For example, at $z_5$, the $CNN^{100}$ achieved 0.09 rad (7-beamlet) and, at $z_6$, 0.10 rad (9-beamlet), nearly identical to the best-case performance of FCNN. This suggests that, under suitable conditions, the FCNN performs similarly compared to the CNN, despite having access only to sparse, integrated intensity readings.

This behavior can be attributed to the design principle behind the detector approach, which was optimized based on two-beamlet interference. At $z_5$, where beamlet interference predominantly occurs between pairs of adjacent beamlets, this assumption holds well, and the sparse detector configuration captures sufficient phase-relevant information. In contrast, at $z_2$, where beamlets largely overlap and the interference pattern becomes indistinct, the two-beamlet approximation breaks down. At $z_6$, although the interference is pairwise in nature, the fringe visibility is diminished due to weaker overlap and a lower signal-to-noise ratio, both of which reduce phase sensitivity.

To confirm the significance of detector orientation, FCNNs trained using rotated detector configurations consistently yielded higher phase prediction errors. In the 7-beamlet case, the minimum error achieved was 0.23 rad at $z_5$, and for the 9-beamlet case, 0.55 rad at $z_5$. These results underscore the importance of aligning detectors along physically meaningful axes, particularly

Optics EXPRESS

those corresponding to the direction of maximum intensity variation, for effective central "fringe" location and unambiguous phase inference.

A comparative analysis of inference speed was conducted between the $CNN^{100}$ and FCNN architectures for phase inference in a 7-beamlet array. Benchmarking was performed on a system running Windows 10 with an Intel Core i7-6700 CPU. Prior to benchmarking, the trained models were converted from PyTorch (.pth) [25] to ONNX (.onnx) [26] to improve inference efficiency. The $CNN^{100}$ achieved an average inference rate of approximately 440 Hz for pure model inference, while the operational bandwidth decreased to approximately 230 Hz when data transfer was included. The FCNN demonstrated significantly faster performance, achieving around 35 kHz for the model inference alone, and approximately 3700 Hz when data transfer overhead was included. While these results depend on the underlying hardware, the observed trend is consistent with theoretical expectations: due to their simpler structure, FCNNs typically requires less computational resources and are consequently more efficient during phase inference. In this case, the dominant limitation arises from data transfer rather than the FCNN inference itself.

In summary, these findings demonstrate that, by identifying "information rich" areas in intensity patterns and carefully selecting the information provided to the deep learning model, it is possible to extract compact yet highly informative representations. This approach enables substantial reductions in input dimensionality and significantly faster inference, all while preserving accuracy in phase prediction.

## 6. Conclusions

This work presents a novel strategy for efficient phase inference in CBC systems using deep learning, replacing high-dimensional camera observations with sparse, spatially integrated intensity measurements. We demonstrate that, at specific axial positions of the imaging system along the optical axis, input dimensionality can be drastically reduced without compromising prediction accuracy. FCNNs trained on these compact representations achieved phase prediction errors comparable to CNNs operating on full-resolution camera images (approximately $\lambda/60$ for a 7-beamlet hexagonal CBC array). Notably, the FCNN approach yielded a substantial gain in inference speed (excluding data acquisition and transfer), achieving an effective inference rate of approximately 35 kHz compared to 440 Hz for the CNN, representing more than 80-fold increase. These results underscore the critical importance of informed input selection and demonstrate that deep learning models can operate effectively on physically motivated, low-dimensional representations. Overall, this approach offers a scalable and computationally efficient pathway for phase retrieval in high-power laser systems.

## Appendix A: uniqueness of single defocused intensity pattern

Although it has been demonstrated that defocused interference patterns enable convergence of deep learning algorithms and allow phase identification with high accuracy, the question remains whether the defocused plane mitigates or merely hides the non-uniqueness that is well known at the focal plane. To address this question, we provide an analysis of near-field intensity patterns, demonstrating the uniqueness of the phase-intensity mapping at early axial propagation positions.

## Appendix B: analytical results

Consider a standard 7-channel hexagonal closely packed CBC array. The central and rightmost beamlets are denoted 0 and 1, respectively. The rest are denoted as 2 to 6. For a fixed observation plane $z_0$ within the Fresnel domain (i.e., $z_0$ small), in the region between beamlet 0 and 1, the interference pattern is dominated by the relative phase ($\varphi$) between beamlets 0 and 1. The contributions of the other beamlets are much smaller and their relative phases play a secondary

Optics EXPRESS

role. In this case, we consider all the relative phases between beamlet 0 and 2-6 fixed and vary only the relative phase ($\varphi$) between beamlets 0 and 1. At the fixed observation plane $z_0$, in the space between beamlets 0 and 1, the total field distribution can be expressed as

$$E(x, y, z_0; \varphi) = E_0(x, y, z_0) + E_1(x, y, z_0)e^{i\varphi} \tag{1}$$

where $E_0$ is the sum of all the contributions of beamlets 0 plus 2-6, and $E_1$ is the contribution of beamlet 1. The total intensity distribution is given by:

$$I(\varphi) = |E(\varphi)|^2 = |E_0|^2 + |E_1|^2 + E_0E_1^*e^{-i\varphi} + E_0^*E_1e^{i\varphi} \tag{2}$$

which becomes

$$I(\varphi) = \alpha + \beta\cos(\varphi) + \gamma\sin(\varphi) \tag{3}$$

where

$$\alpha = |E_0|^2 + |E_1|^2; \beta = 2Re\{E_0^*E_1\};\ \gamma = -2Im\{E_0^*E_1\} \tag{4}$$

In Eqs. (2)–(4) we have dropped the $(x,\ y,\ z_0)$ dependence for simplicity. The intensity is sampled at three points $x_1, x_2, x_3$ along the $y = 0$ line connecting the centers of beamlets 0 and 1. The corresponding intensities are given by:

$$I_k(\varphi) = \alpha_k + \beta_k\cos(\varphi) + \gamma_k\sin(\varphi); k = 1,\ 2,\ 3 \tag{5}$$

For observation planes $z_0$ close to the beamlet exit plane, the interference pattern between beamlets 0 and 1 is dominated by an intensity notch, which can be fitted by a parabola. In this case, we apply a centered quadratic fit, expressed as

$$f(x) = A(\varphi)(x - x_0)^2 + B(\varphi)(x - x_0) + C(\varphi) \tag{6}$$

with

$$A(\varphi) = \frac{I_1 - 2I_2 + I_3}{2h^2}; B(\varphi) = \frac{I_3 - I_1}{2h}; C(\varphi) = I_2 \tag{7}$$

where $x_0 \equiv x_2$ and $|x_1 - x_2| = |x_3 - x_2| = h$. The coefficient $A$, $B$ and $C$ give the curvature, linear term and offset of the fitted parabola.

Since $A$, $B$, $C$ are linear combinations of $I_1$, $I_2$, $I_3$, they can also take the form

$$\begin{aligned} A(\varphi) &= A_0 + A_c\cos(\varphi) + A_s\sin(\varphi) \\ B(\varphi) &= B_0 + B_c\cos(\varphi) + B_s\sin(\varphi) \\ C(\varphi) &= C_0 + C_c\cos(\varphi) + C_s\sin(\varphi) \end{aligned} \tag{8}$$

where

$$\begin{gathered} A_0 = \tfrac{\alpha_1 - 2\alpha_2 + \alpha_3}{2h^2};\ A_c = \tfrac{\beta_1 - 2\beta_2 + \beta_3}{2h^2};\ A_s = \tfrac{\gamma_1 - 2\gamma_2 + \gamma_3}{2h^2} \\ B_0 = \tfrac{\alpha_3 - \alpha_1}{2h};\ B_c = \tfrac{\beta_3 - \beta_1}{2h};\ B_s = \tfrac{\gamma_3 - \gamma_1}{2h} \\ C_0 = \alpha_2;\ C_c = \beta_2;\ C_s = \gamma_2 \end{gathered} \tag{9}$$

In the coefficient $(A,\ B,\ C)$ space, $A(\varphi)$, $B(\varphi)$ and $C(\varphi)$ form a coefficient vector $\bar{F}(\varphi)$ whose tip moves in space as the relative phase ($\varphi$) is varied. The coefficient vector $\bar{F}(\varphi)$ can be expressed

Optics EXPRESS

as:

$$\bar{F}(\varphi) = \begin{bmatrix} A(\varphi) \\ B(\varphi) \\ C(\varphi) \end{bmatrix} = \bar{F}_0 + \bar{u}\cos(\varphi) + \bar{v}\sin(\varphi) \tag{10}$$

with

$$\bar{F}_0 = \begin{bmatrix} A_0 \\ B_0 \\ C_0 \end{bmatrix};\ \bar{u} = \begin{bmatrix} A_c \\ B_c \\ C_c \end{bmatrix};\ \bar{v} = \begin{bmatrix} A_s \\ B_s \\ C_s \end{bmatrix} \tag{11}$$

Equation (10) shows that the tip of the coefficient vector follows an ellipse in a 2D plane defined by the directions $\bar{u}$ and $\bar{v}$ with offset $\bar{F}_0$ from the origin, imbedded in the 3D space $(A,\ B, C)$. Because the phase curve lies in a single plane, it cannot fold arbitrarily in 3-D space. Instead, the curve is a simple ellipse, and it has no self-intersection. This shows that in the case of three-detector fringe measurement $(I_1(\varphi), I_2(\varphi), I_3(\varphi))$ and simple quadratic fitting, unique phase $(\varphi)$ recovery is possible from $(A(\varphi), B(\varphi), C(\varphi))$ (except for $2\pi$ periodicity).

## Appendix C: symmetric two-beam special case

If we consider the influence of beamlets 2-6 on beamlet 0 is negligible or the special case of two identical symmetrically placed beamlets about $x_0 \equiv x_2$ $(y = 0)$, as shown in Fig. 6, the corresponding electric fields at the detector positions $x_1,\ \ x_2,\ \ x_3$ can be expressed as:

$$E_0(x_1, 0, z_0) = E_1(x_3, 0, z_0);\ E_0(x_3, 0, z_0) = E_1(x_1, 0, z_0);\ E_0(x_2, 0, z_0) = E_1(x_2, 0, z_0) \tag{12}$$

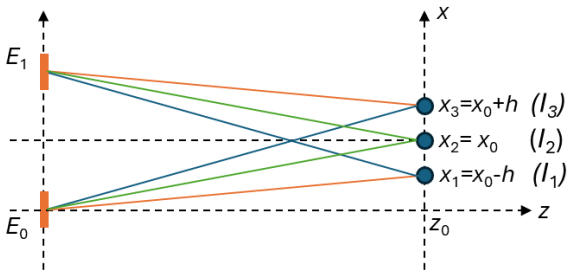


**Fig. 6.** Schematic of two identical symmetrically placed beamlets about $x_0 \equiv x_2$.

In this case, it is easily shown that intensities $I_1$, $I_2$ and $I_3$ fulfill the following relations

$$I_1(\varphi) = I_3(-\varphi);\ I_2(\varphi) = I_2(-\varphi) \tag{13}$$

where $I_k = |E(x_k, 0, z_0; \varphi)|^2; k = 1,\ 2,\ 3$. Substituting Eqs. (13) into (7), we obtain the following generic relations for the coefficients of the centered quadratic fit:

$$A(\varphi) = A(-\varphi);\ B(\varphi) = -B(-\varphi);\ C(\varphi) = C(-\varphi) \tag{14}$$

This means that $A$ and $C$ are even in phase, while $B$ is odd in phase. So, in the ideal two-beam symmetric case, from Eqs. (8) we obtain

$$\begin{aligned} A(\varphi) &= A_0 + A_c\cos(\varphi) \\ B(\varphi) &= B_0 + B_s\sin(\varphi) \\ C(\varphi) &= C_0 + C_c\cos(\varphi) \end{aligned} \tag{15}$$

Optics EXPRESS

From Eqs. (10) and (11), the phase curve takes the simplified form:

$$\bar{F}(\varphi) = \begin{bmatrix} A_0 \\ 0 \\ C_0 \end{bmatrix} + \begin{bmatrix} A_c \\ 0 \\ C_c \end{bmatrix} \cos(\varphi) + \begin{bmatrix} 0 \\ B_s \\ 0 \end{bmatrix} \sin(\varphi) \tag{16}$$

In this case, the ellipse cosine direction is in the $A$-$C$ plane and the sine direction along $B$.

## Appendix D: experimental results

We begin the experimental analysis by considering a single beamlet in a 7-channel CBC array, specifically the rightmost beamlet. Figure 7(a) shows three example intensity patterns. In the leftmost image, the phase of the selected beamlet is -π (relative to the central reference fiber). In the central image, the phase is 0π, and in the right image the phase is π. In these examples, the phases of the remaining beamlets are random and their exact values are not important for the present discussion.

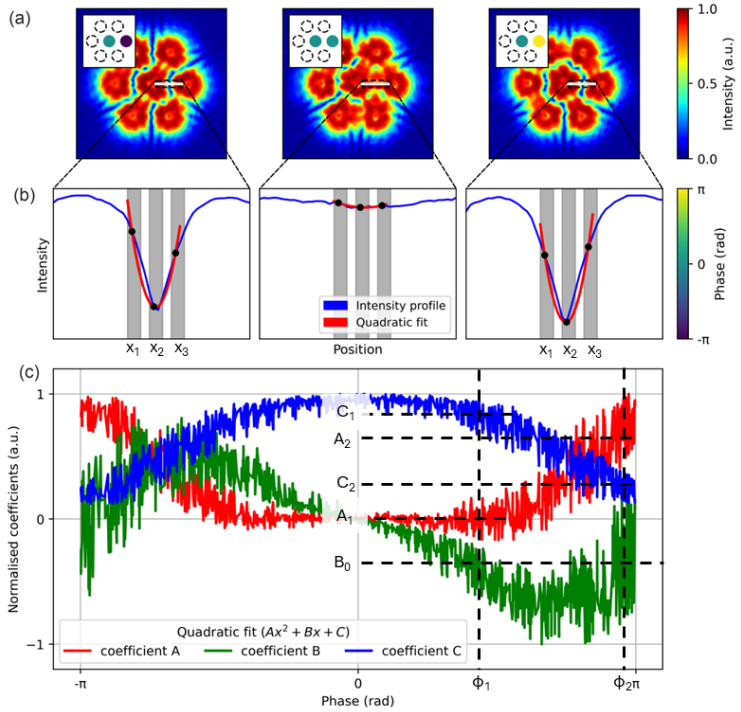


**Fig. 7.** (a) Example intensity patterns of a 7-channel CBC system recorded at the axial position $z_5$ for different relative phases of the selected rightmost beamlet. (b) One-dimensional intensity distribution along the line indicated in (a). The grey regions mark the detector positions used for intensity sampling. (c) Coefficients $A$, $B$, and $C$ of the quadratic fit $y = Ax^2 + Bx + C$ plotted as functions of the phase of the selected rightmost beamlet.

Figure 7(b) illustrates a one-dimensional slice taken along the white line overlaid on Fig. 7(a), showing the intensity as a function of position. At this axial position (i.e., $z_5$), the intensity profile exhibits a parabolic shape when the phase of the beamlet under consideration is ±π, while the parabola becomes shallower when the beamlet phase is 0π. The grey regions indicate the

Optics EXPRESS

positions of the three detectors defined in the main text of the manuscript. Each detector averages the intensity over its area, resulting in three measured values. The placement of the detectors effectively samples three points along the parabola. These values are sufficient to perform a quadratic fit, shown as the red curve in Fig. 7(b).

Next, we consider 1000 intensity patterns, each generated by applying random phases to the outer beamlets within the range -π to π. For each intensity pattern, a quadratic function of the form $y = Ax^2 + Bx + C$ is fitted to the detector measurements obtained from the fixed detector positions. The resulting coefficients $A$, $B$, and $C$ are then plotted as functions of the beamlet phase, as shown in Fig. 7(c). Although some noise is present, the coefficients $A$ and $C$ exhibit clear symmetric (even) dependence around 0π with respect to the phase of beamlet, whereas the coefficient $B$ shows a distinctly asymmetric (odd) dependence. This dependence is in very good agreement with the analytical relations expressed by Eqs. (14). The observed noise primarily comes from the contributions of the remaining beamlets and is significantly smaller in the simple 2-beamlet case (where no additional beamlets are present) at the same axial position $z_5$.

Due to the even symmetries around 0π, observed in Fig. 7(c), coefficients $A$ and $C$ or their combination $A + C$ cannot be used to uniquely distinguish a given phase from its negative counterpart. Coefficient $B$, on the other hand, although showing odd symmetry around 0π, it cannot be used to uniquely distinguish the relative phase of the rightmost beamlet. For example, a value $B_0$ on its own can provide two phases ($\phi_1$ and $\phi_2$ shown in Fig. 7(c)), resulting in a phase ambiguity. However, if $B_0$ is combined with $A_1(A_2)$, their combination can determine $\phi_1(\phi_2)$ unambiguously. Similarly, if $B_0$ is combined with $C_1(C_2)$ or $A_1 + C_1(A_2 + C_2)$, these combinations again can determine the phase $\phi_1(\phi_2)$ unambiguously.

To quantify the predictions outlined above, we carried out an additional analysis in which four additional FCNNs were trained and validated on fitted coefficient-beamlet phase pairs. Specifically, the detector readings from the main manuscript were used to fit a quadratic function to each intensity pattern, yielding coefficients $A$, $B$, and $C$. One FCNN was trained using all three coefficients ($A$, $B$, $C$) as input to predict the phase of the rightmost beamlet (i.e., input dimensionality of 3 m output dimensionality of 1). The remaining three FCNNs were trained using all possible pairs of coefficients (i.e., $A + B$, $B + C$, $A + C$, input dimensionality of 2, output dimensionality of 1). The prediction accuracy of these FCNNs was subsequently compared qualitatively to that of the original FCNN trained on the raw detector readings for the rightmost beamlet. The results are plotted in Fig. 8. In Fig. 8(a), training and validation were carried out by using the raw intensities recorded directly by the three detectors and used as a reference. In Fig. 8(b-e) training and validation were carried out by using the fitted parabola coefficients $A + B + C$, $A + B$, $A + C$ and $B + C$, respectively. In Fig. 8(b) it is shown that using all three coefficients $A$, $B$, $C$ results in a validation map almost indistinguishable from the three-detector reference (see Fig. 8(a)). The same applies when coefficient B is combined with either $A$ ($B + A$ - Fig. 8(c)) or $C$ ($B + C$ – Fig. 8(d)). On the contrary, when only coefficients A and C are taken into consideration, the corresponding validation map, shown in Fig. 8(e), demonstrates that no reliable phase extraction can be achieved. All these findings are in close agreement with the conclusions, drawn above from the symmetries of coefficients $A$, $B$, $C$, shown in Fig. 7(c). Although the $A$, $B$, $C$ coefficient curves in Fig. 7(c) are effected by noise, the experimental results of Fig. 8 clearly demonstrate that use of the NN results in quite accurate and reliable prediction-ground truth validation maps.

In this Appendix, we have shown analytically and experimentally that monitoring the characteristic intensity dip (“dark” fringe), formed by interference between two adjacent beamlets (0 and 1) in the near field, using three closely spaced detectors, breaks the one-to-many inverse phase recovery problem, pertinent to CBC systems, and results in an unambiguous extraction of the beamlet relative phase. We should add, however, that as we move the observation plane outside the near field region, approaching the focal plane, the propagated fields from beams 2–6

Optics EXPRESS

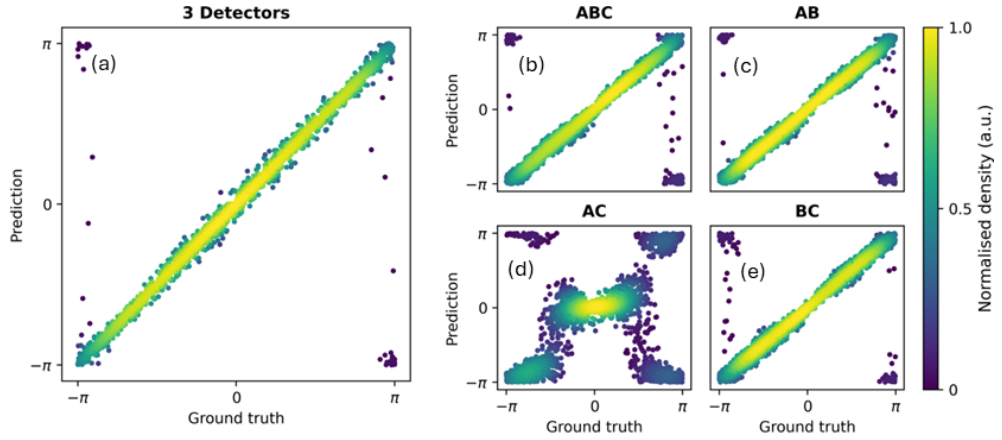


**Fig. 8.** Prediction-ground truth validation maps of a 7-channel CBC system recorded at the axial position $z_5$ for different relative phases of the selected rightmost beamlet. Training and validation were carried out by using (a) intensities recorded by the three detectors and the fitted parabola coefficients (b) A + B + C, (c) A + B, (d) A + C and (e) B + C.

spread and interfere more strongly inside the local fitting window between beamlets 0 and 1. The quadratic fit coefficients $A$, $B$, $C$ change in magnitude and relative phase, and the chosen local window may stop tracking a single intensity dip. That alone does not yet create self-intersections, but it makes the simple ellipse, discussed in section 1, less physically meaningful. As a result, the phase extraction ambiguity increases and the FCNN mean phase error increases (see Fig. 5($a_3$) and Fig. 5($b_3$)).

**Funding.** Engineering and Physical Sciences Research Council (EP/W028786/1, EP/T026197/1, EP/Z002567/1).

**Disclosures.** The authors declare no conflicts of interest.

**Data availability.** Data underlying the results presented in this paper are available in Ref. [27].

**Supplemental document.** See Supplement 1 for supporting content.

## References

1. C. L. Linslal, P. Ayyaswamy, S. Maji, *et al.*, "Challenges in coherent beam combining of high power fiber amplifiers: a review," ISSS J. Micro and Smart Sys. **11**(1), 277–293 (2022).
2. T. Y. Fan, "Laser beam combining for high-power, high-radiance sources," IEEE J. Sel. Top. Quantum Electron. **11**(3), 567–577 (2005).
3. A. Klenke, M. Muller, H. Stark, *et al.*, "Coherent Beam Combination of Ultrafast Fiber Lasers," IEEE J. Sel. Top. Quantum Electron. **24**(5), 1–9 (2018).
4. K. Jin, H. Chang, R. Su, *et al.*, "Coherent beam combining of 7 fiber amplifiers based on all-fiber internal phase-locking technique," Opt. Laser Technol. **171**, 110456 (2024).
5. H. Chang, Q. Chang, J. Xi, *et al.*, "First experimental demonstration of coherent beam combining of more than 100 beams," Photonics Res **8**(12), 1943–1948 (2020).
6. T. M. Shay, V. Benham, J. T. Baker, *et al.*, "First experimental demonstration of self-synchronous phase locking of an optical array," Opt. Express **14**(25), 12015–12021 (2006).
7. P. Ma, H. Chang, Y. Ma, *et al.*, "7.1 kW coherent beam combining system based on a seven-channel fiber amplifier array," Opt. Laser Technol. **140**, 107016 (2021).
8. T. Rousseaux, J. Primot, J.-C. Chanteloup, *et al.*, "HIBISCUS: Simulator for the development of control strategies for coherent beam combining lasers with PISTIL technique," EPJ Web Conf. **287**, 08019 (2023).
9. H. Zhou, X. Feng, L. Xie, *et al.*, "Comprehensive investigation of LOCSET and SPGD algorithms in coherent beam combining applications," Opt. Laser Technol. **181**, 111568 (2025).
10. X. Zhang, P. Li, Y. Zhu, *et al.*, "Coherent beam combination based on Q-learning algorithm," Opt. Commun. **490**, 126930 (2021).
11. H. Pang and T. Haecker, "Laser Cutting with annular intensity distribution," Procedia CIRP **94**, 481–486 (2020).
12. H. Le, P. Penchev, A. Henrottin, *et al.*, "Effects of Top-hat Laser Beam Processing and Scanning Strategies in Laser Micro-Structuring," in *Micromachines*, (2020).

Optics EXPRESS

13. C. Prieto, E. Vaamonde, D. Diego-Vallejo, *et al.*, "Dynamic laser beam shaping for laser aluminium welding in e-mobility applications," Procedia CIRP **94**, 596–600 (2020).
14. A. Nissenbaum, N. Armon, E. Shekel, *et al.*, "Dynamic beam lasers based on coherent beam combining," in *Proc. SPIE*, 2022), 119810B.
15. B. Mills, J. A. Grant-Jacob, M. Praeger, *et al.*, "Single step phase optimisation for coherent beam combination using deep learning," Sci. Rep. **12**(1), 5188 (2022).
16. B. M. James, A. Grant-Jacob, and M. N. Zervas, "Phase identification despite amplitude variation in coherent beam combination using deep learning," Optics Continuum (2023).
17. F. Chernikov, Y. Xie, J. A. Grant-Jacob, *et al.*, "Deep learning for simultaneous phase and amplitude identification in coherent beam combination," Sci. Rep. **15**(1), 11757 (2025).
18. Y. Xie, F. Chernikov, B. Mills, *et al.*, "Single-step phase identification and phase locking for coherent beam combination using deep learning," Sci. Rep. **14**(1), 7501 (2024).
19. T. Hou, Y. An, Q. Chang, *et al.*, "Deep-learning-based phase control method for tiled aperture coherent beam combining systems," High Power Laser Sci. Eng. **7**, e59 (2019).
20. W. Kerridge-Johns, F. Chernikov, W. Yu, *et al.*, "Neural Network Coherent Beam Combination with Low Differential Phase Noise Ytterbium Fiber Amplifiers," in *Laser Congress 2024 (ASSL, LAC, LS&C)*, Technical Digest Series (Optica Publishing Group, 2024), ATh1A.6.
21. N. Hulard, B. Rouzé, L. Lombard, *et al.*, "Experimental demonstration of coherent beam combination by a simulation-trained deep neural network," Opt. Lett. **51**(2), 381–384 (2026).
22. A. Howard, M. Sandler, B. Chen, *et al.*, "Searching for MobileNetV3," in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 2019), 1314–1324.
23. H. Zhou, R. Tao, X. Feng, *et al.*, "Machine learning phase control of filled-aperture coherent beam combining: principle and numerical demonstration," High Power Laser Sci. Eng. **13**, e39 (2025).
24. D. P. Kingma, "Adam: A method for stochastic optimization," (2014).
25. A. Paszke, S. Gross, F. Massa, *et al.*, "Pytorch: An imperative style, high-performance deep learning library," Advances in neural information processing systems **32** (2019).
26. O. R. Developers, "ONNX Runtime" (2021, 2021), retrieved https://onnxruntime.ai/.
27. F. Chernikov, et al., "Dataset to support Efficient phase identification in coherent beam combination using interpretable deep learning," (2026).