

Article

Decoding Coherent Patterns from Arrayed Waveguides for Free-Space Optical Angle-of-Arrival Estimation

Jinwen Zhang, Haitao Zhang * and Zhuoyi Yang 

State Key Laboratory of Precision Space-Time Information Sensing Technology, Department of Precision Instrument, Tsinghua University, Beijing 100084, China; jinwen-z21@mails.tsinghua.edu.cn (J.Z.); zy-yang22@mails.tsinghua.edu.cn (Z.Y.)

* Correspondence: zhanghaitao@mail.tsinghua.edu.cn

Abstract

This paper presents a novel free-space optical Angle-of-Arrival (AOA) estimation method based on arrayed waveguide coherent mode decoding, aiming to surpass the inherent limitations of traditional AOA detection technologies, which face significant challenges in achieving miniaturization, low complexity, and high reliability. The method utilizes the AOA-related phase differences generated by the propagation and interference of incident light in an arrayed input waveguide, forming multi-beam interference fringes at the output end of the slab waveguide. This pattern is then sampled by an arrayed output waveguide to produce an intensity sequence, which is then fed into a trained CNN–Attention Regressor for AOA estimation. This study innovatively applies the method to decoding the spatial angular information of optical signals. Simulation results demonstrate the exceptional performance of our approach, achieving a Mean Absolute Error (MAE) of 0.0142° and a Root Mean Square Error (RMSE) of 0.0193° over a 40° field of view. This precision is significantly superior to conventional peak–linear calibration methods and other common neural network architectures, and exhibits remarkable robustness against simulated phase noise and manufacturing tolerances. This research demonstrates the powerful synergy between integrated photonics and deep learning, paving the way for a new class of highly integrated, robust, and high-performance on-chip optical sensors.



Academic Editor: Xin Zhu

Received: 26 October 2025

Revised: 22 November 2025

Accepted: 25 November 2025

Published: 27 November 2025

Citation: Zhang, J.; Zhang, H.; Yang, Z. Decoding Coherent Patterns from Arrayed Waveguides for Free-Space Optical Angle-of-Arrival Estimation. *Sensors* **2025**, *25*, 7231. <https://doi.org/10.3390/s25237231>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: angle-of-arrival estimation; arrayed waveguide; coherent mode decoding; convolutional neural network; attention mechanism

1. Introduction

Optical Angle-of-Arrival (AOA) estimation is a critical technique for spatial parameter sensing playing a pivotal role in fields such as Free-Space Optical (FSO) communication [1–3], LiDAR (Light Detection and Ranging) [4–6], and Laser Warning Systems (LWS) [7,8]. Its performance directly impacts the system’s capabilities in areas such as beam alignment, target localization, and threat identification [2,3,6,7]. With the increasing demand for device miniaturization, wide field of view, high precision, and high integration in modern applications [2,3,6,7], the limitations of traditional angle detection technologies have become increasingly prominent.

The landscape of optical AOA detection encompasses several distinct approaches. One category of techniques relies on discrete optical components. For example, lens-based systems determine the AOA by calculating the centroid of a light spot on a photosensitive detector, such as a Four-Quadrant Detector (QPD) [3,9]. Similarly, methods

based on detector arrays have been widely explored, inferring the AOA by measuring the differential incident optical flux across array elements positioned at different spatial locations [3,10]. While effective, these approaches generally depend on bulky and structurally complex macroscopic systems, posing significant challenges for miniaturization and seamless integration with modern photonic circuits.

To address these integration challenges, more advanced array-centric technologies have been investigated. Optical Phased Arrays (OPAs) can be utilized for optical AOA detection, where the receiver employs programmable angular selectivity to collect light from a specific angle while suppressing off-axis signals [11,12]. However, OPAs typically require active electrical control systems for beam steering, meaning they are not fully passive devices, and their on-chip integration remains complex. Many advanced concepts from the radio-frequency (RF) domain, while well-established, have not been transferred to the optical domain. For instance, a space-time modulated metasurface has been proposed for AOA estimation by measuring the amplitude imbalance between harmonic frequencies of the reflected field [13]. Similarly, leveraging time delay is a well-established principle in RF systems. This has been adapted for AOA detection, where multiple spatially distributed receivers measure the relative time delay between different signal channels to derive the angle [14,15]. On the integrated photonics front, a method utilizing a PIC-based star coupler has been demonstrated to achieve an optical Fourier transform, replacing a conventional lens for two-dimensional (2D) AOA detection [16–18]. While this approach offers a significant reduction in system Size, Weight, and Power (SWaP) [17], these advanced techniques collectively highlight a persistent challenge: they are primarily validated in the RF domain with non-trivial transitions to optical frequencies, or have not yet been tailored for passive, high-precision AOA decoding.

A review of the existing landscape reveals that current AOA detection methods are either constrained by the bulkiness of discrete components or rely on active, complex systems, with some advanced concepts yet to be demonstrated effectively and passively in the optical domain. This highlights a clear demand for a solution that is compact, passive in its optical detection stage, and amenable to chip-scale integration.

In parallel, the field of integrated photonics has revolutionized optical communication and sensing, celebrated for its advantages in miniaturization, stability, and potential for mass production [19,20]. Applying the principles of this powerful platform to the challenge of optical AOA detection, therefore, presents a compelling yet underexplored research avenue. Among the rich library of integrated photonic devices, the Arrayed Waveguide Grating (AWG) is a cornerstone component for Wavelength Division Multiplexing (WDM), widely used for multi-channel signal processing and high-precision spectral analysis due to its low loss, high channel capacity, and excellent channel isolation [21–23]. The working principle of an AWG is based on the phase accumulation from path length differences in its arrayed waveguides—light at different wavelengths accumulates different phases, enabling selective separation through interference [21]. Its high degree of tunability in geometric parameters and phase control endows the structure with significant design freedom [24]. Despite its widespread use, the potential of the arrayed waveguide for spatial information sensing, such as optical AOA detection, has remained unexplored. Given the phase-sensitive nature of the arrayed waveguides, it is theoretically plausible to convert the spatial phase front of an incident light field into a measurable intensity distribution, offering a novel pathway for highly integrated, mechanical-scan-free angle measurement. However, this direction has several research gaps, necessitating a systematic evaluation of the feasibility of arrayed waveguide devices for optical AOA detection—from theoretical modeling and device design to performance verification—to provide a methodological foundation for future implementation.

To bridge this identified research gap and harness the untapped potential of arrayed waveguides for spatial sensing, this paper proposes and numerically validates an optical AOA detection method based on arrayed waveguides. This work extends the function of an on-chip arrayed waveguide from a traditional wavelength selector/demodulator to a decoder of spatial angular information. The proposed method unfolds in four key stages. Firstly, incident free-space light is coupled into a high fill-factor arrayed waveguide, generating a phase and amplitude distribution along the propagation path that is dependent on AOA. Secondly, within the slab waveguide, these multiple beams undergo coherent superposition to form spatial interference fringes that are coupled with the incident angle. Then, an arrayed output waveguide at preset positions performs mode coupling and intensity sampling on the light field at the output plane of the slab waveguide, yielding a spatial intensity sequence. Finally, this intensity sequence is fed into a trained CNN–Attention regressor [25–29], which directly regresses the incident angle, ensuring high accuracy and robustness in the model’s predictions.

The main contributions of this paper can be summarized as follows:

- A novel application for the arrayed waveguides—optical AOA detection—is conceptually proposed and demonstrated. By fully exploiting its inherent phase sensitivity and integrated photonic platform, this approach enables miniaturized, high-precision angle sensing, extending its functionality beyond conventional wavelength division multiplexing.
- A complete hybrid detection architecture is designed that synergistically combines the arrayed waveguides with the powerful CNN–Attention regressor, creating an end-to-end AOA detection pipeline.
- Through comprehensive numerical simulations, the proposed method has been validated. The results show that for angle prediction over a 40° field of view, the method achieves a Mean Absolute Error (MAE) of 0.0142° and a Root Mean Square Error (RMSE) of 0.0193°. These findings demonstrate the method’s high accuracy and robustness in AOA estimation, providing a solid methodological basis for future experimental implementation.

2. Materials and Methods

This section details the design of the arrayed waveguide structure, the optical field simulation method, the noise model, and a convolutional neural network (CNN) combined with a self-attention mechanism (CNN–Attention Regressor) architecture used for AOA estimation.

The arrayed waveguide structure investigated in this paper is designed to operate primarily in a single plane. The core physical principle—converting the AOA into a spatial shift of the interference peak—is predominantly a one-dimensional (1D) phenomenon along the axis of the Rowland circle. To efficiently validate this fundamental concept and generate the large dataset required for training the deep learning model, we have simplified the two-dimensional (2D) optical simulation problem into a 1D model. This simplification significantly reduces computational complexity, making the extensive numerical experiments feasible.

While this 1D model captures the essential beam-steering physics and allows for a robust proof-of-concept demonstration, we acknowledge that it represents a first-order approximation of a real-world 2D device. The potential impacts and limitations arising from this simplification, such as resolution asymmetry and polarization-dependent effects, are non-trivial. A detailed discussion of these factors and their implications for practical implementation is provided in the Discussion section (Section 4). Therefore, the subsequent simulation analysis based on this 1D model serves as a foundational study to establish the

method's viability, with its conclusions providing a strong baseline for future 2D modeling and experimental verification.

2.1. Waveguide Structure and Optical Propagation

In this study, we construct an arrayed waveguide structure that converts the spatially distributed phase information corresponding to the AOA of free-space light into spatially distributed intensity information. As shown in Figure 1, the structure uses arrayed waveguides, with elements arranged equidistantly on an arc of radius $R = 2r$, to collect the phase information of the incoming free-space light. Multiple beams interfere within a slab waveguide, causing beams with the same AOA to focus at the output plane of the slab waveguide. An array of output waveguides, arranged equidistantly around a Rowland circle of radius r , samples the intensity of the output optical field, yielding a spatial optical intensity sequence that is correlated with the AOA.

The Rowland circle geometry describes the inherent focal curve of our device [21,30]. For different Angles of Arrival, the focal points trace this circle, not a straight line. Placing the output waveguides on this curve is therefore a fundamental requirement to minimize optical aberrations and ensure optimal focusing.

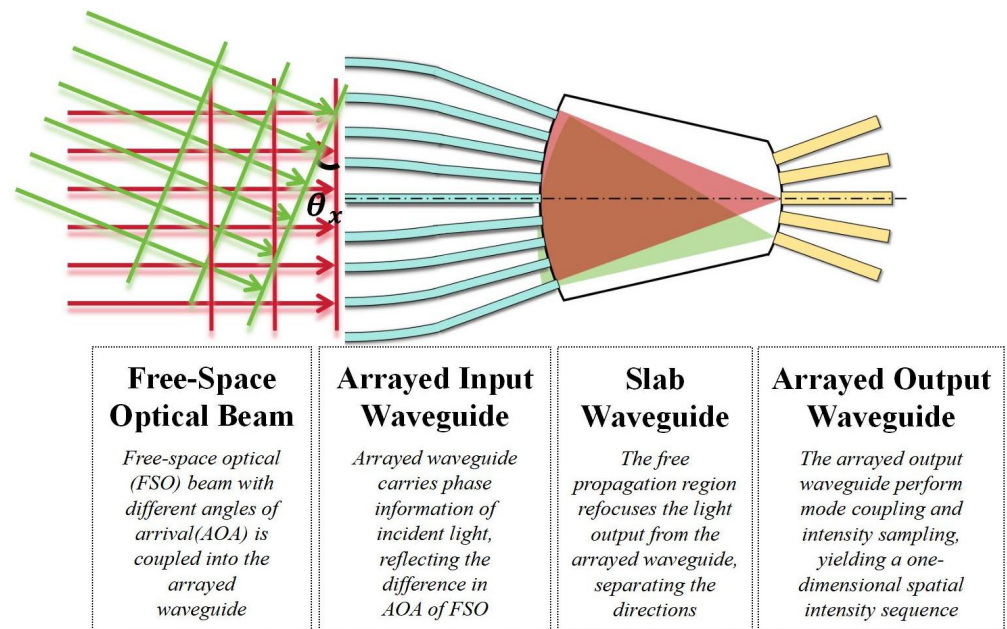


Figure 1. Basic structure of the arrayed waveguide model for Angle-of-Arrival (AOA) detection. The red lines represent normally incident beams, the green lines represent obliquely incident beams, the blue structure is the arrayed input waveguide, and the yellow structure is the arrayed output waveguide. Both are arranged equidistantly on opposite sides of the slab waveguide, following the Rowland circle geometry.

2.1.1. Basic Assumptions

In this research, we build an arrayed waveguide AOA estimation architecture on a simulation platform for optical field simulation. The simulation environment is based on rigorous electromagnetic field theory to ensure the physical accuracy of the calculation results.

It is assumed that the input is a tilted Gaussian beam with a finite beam waist [31], incident from a certain position in space onto the center of the arrayed input waveguide structure. Its complex amplitude distribution can be expressed as

$$E_{\text{tilt}}^2(x, y, z) = \exp \left[-\frac{(\eta^2 + \xi^2)}{\omega^2(\chi)} \right] \times \exp \left\{ -i \times k \left[\frac{(\eta^2 + \xi^2)}{2\Delta R} + \chi \right] - \Delta\psi \right\}, \quad (1)$$

with

$$\eta = \cos \theta_x x + \sin \theta_y \sin \theta_x y - \sin \theta_x \cos \theta_y z - x_0, \quad (2)$$

$$\zeta = \cos \theta_y y + \sin \theta_y z - y_0, \quad (3)$$

$$\chi = \sin \theta_x x - \sin \theta_y \cos \theta_x y + \cos \theta_x \cos \theta_y z, \quad (4)$$

$$\Delta R = Z_0(\chi/Z_0 + Z_0/\chi), \quad (5)$$

$$\Delta\psi = \arctan(\chi/Z_0), \quad (6)$$

$$w(\chi) = w_0 \sqrt{1 + (\chi/Z_0)^2}, \quad (7)$$

$$z_0 = \frac{\pi \omega_0^2}{\lambda}, \quad (8)$$

Assuming the waveguide operates under the weakly guiding approximation, the Effective Index Method (EIM) is used to calculate the effective refractive index of the modes [32]. This method decomposes the three-dimensional waveguide structure into two-dimensional slab waveguide problems. The effective refractive index of the guided modes is obtained by solving the eigenvalue equation:

For TE modes, the eigenvalue equation is

$$\kappa h = m\pi + \arctan\left(\frac{p_0}{\kappa}\right) + \arctan\left(\frac{p_2}{\kappa}\right), \quad (9)$$

For TM modes, the eigenvalue equation is

$$\kappa h = m\pi + \arctan\left(\frac{n_1^2 p_0}{n_0^2 \kappa}\right) + \arctan\left(\frac{n_1^2 p_2}{n_2^2 \kappa}\right), \quad (10)$$

with

$$\kappa = (k_0^2 n_1^2 - \beta^2)^{1/2}, \quad (11)$$

$$p_0 = (\beta^2 - k_0^2 n_0^2)^{1/2}, \quad (12)$$

$$p_2 = (\beta^2 - k_0^2 n_2^2)^{1/2}, \quad (13)$$

where n_1 , n_2 , and n_0 are the refractive indices of the core, upper cladding, and substrate, respectively.

The simulation employs a dispersion model based on the Sellmeier equation to describe the refractive index of materials at specific wavelengths [33]. For the waveguide material, the following third-order Sellmeier equation is used.

$$n^2(\lambda) = 1 + \frac{s_1 \lambda^2}{\lambda^2 - \lambda_1^2} + \frac{s_2 \lambda^2}{\lambda^2 - \lambda_2^2} + \frac{s_3 \lambda^2}{\lambda^2 - \lambda_3^2}, \quad (14)$$

where the Sellmeier coefficients s_1 , s_2 , s_3 and the resonant wavelengths λ_1 , λ_2 , λ_3 are either constants or polynomial functions of temperature.

2.1.2. Optical Field Propagation

The arrayed input waveguide structure is shown in Figure 2a, where the light-yellow structure represents the upper cladding and substrate, and the light-blue structure is the waveguide core. The energy of the incident light beam is coupled into and received by the cores of the arrayed input waveguides. The meanings of the relevant structural parameters are given in Table 1. It is important to note that N_i in the figure is any positive integer, and $L_i = L_0 \pm N_i * \frac{\lambda}{n_c}$ represents the waveguide length of the i -th channel of the arrayed input waveguide.

For a plane wavefront incident at a certain oblique AOA, there exists an optical path difference $\Delta L = d_g \sin \theta$ between adjacent arrayed waveguides, as shown in Figure 2b. This path difference is determined by the geometric distance d_g between the centers of two waveguides and the sine of the incident angle θ , resulting in a phase difference $\Delta \varphi = 2\pi \frac{d_g}{\lambda} \sin \theta$. The waveguide length of the central channel of the arrayed input waveguide is L_0 . For a given wavelength, to ensure that the arrayed input waveguide structure itself does not introduce any additional phase difference, the lengths of the arrayed input waveguides should be $L_i = L_0 \pm N_i * \frac{\lambda}{n_c}$. In this way, the optical path difference between different channels of the arrayed input waveguide structure is related only to the incident angle and the spacing between waveguide centers, thus achieving the collection of phase information related to the AOA.

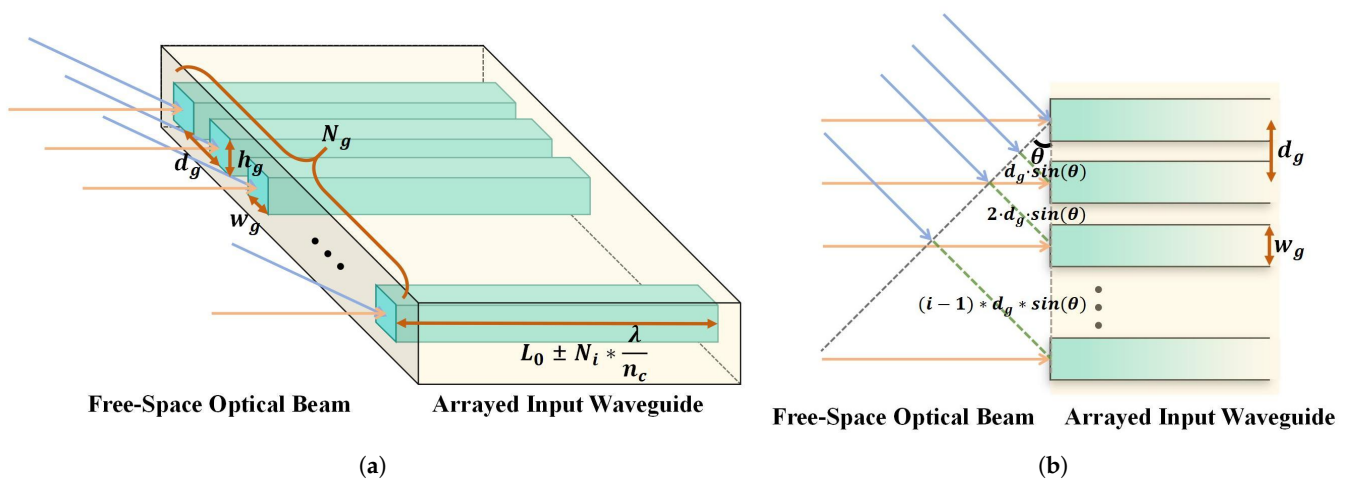


Figure 2. Structure and principle of the arrayed input waveguide. (a) Schematic of the arrayed input waveguide structure and its parameters. The light-yellow structure is the upper cladding and substrate, and the light-blue is the waveguide core. Here, N_i is a positive integer and $L_i = L_0 \pm N_i * \frac{\lambda}{n_c}$ indicates that the optical path difference caused by the length difference between different channels of the arrayed input waveguide should be an integer multiple of the wavelength. (b) Mechanism of phase difference generation for beams incident at different AOAs. The light-yellow structure is the upper cladding and substrate, and the light-blue is the waveguide core.

Table 1. Main structural parameters and their meanings for the arrayed waveguide AOA detection model.

Symbol	Parameter Explanation
λ	Wavelength of the optical signal
θ	AOA of the incident beam
w_g	Core width of the arrayed input waveguide
h_g	Core thickness of the arrayed input waveguide
d_g	Center-to-center spacing of adjacent arrayed input waveguides
N_g	Number of arrayed input waveguides
L_0	Length of the central channel arrayed input waveguide
R	Focusing length of the slab waveguide (Rowland circle diameter)
w_o	Core width of the arrayed output waveguide
h_o	Core thickness of the arrayed output waveguide
d_o	Center-to-center spacing of adjacent arrayed output waveguides
N_o	Number of arrayed output waveguides

After the optical field is collected by the arrayed input waveguide, it continues to propagate through the slab waveguide. The propagation process of the optical field in the waveguide is primarily simulated using the Rayleigh-Sommerfeld diffraction integral [34]. This method, based on scalar diffraction theory, simulates the propagation from the near-field to the output plane. The integral formula is as follows:

$$V(x, y, z) = \iint V(x', y', 0) \left[\frac{1}{2\pi} \frac{z}{R} (1 - ikR) \frac{\exp(ikR)}{R^2} \right] dx' dy', \quad (15)$$

where $V(x, y, z)$ and $V(x', y', 0)$ represent the complex amplitude distributions at the output and input planes, respectively, $R = \sqrt{(x - x')^2 + (y - y')^2 + z^2}$ is the propagation distance, and $k = 2\pi/\lambda$ is the wavenumber.

This method provides good accuracy in both near-field and far-field regions, making it suitable for diffraction calculations in arrayed waveguide structures.

The arrayed waveguide structure is shown in Figure 3, which illustrates the internal layout of the model and the relationships between its components [21]. It can be seen that the arrayed input waveguide and the arrayed output waveguide are distributed on opposite sides of the slab waveguide, arranged equidistantly on two arcs of the slab waveguide. The central axes of all channels of the arrayed input waveguide point towards the center of the output waveguide structure, and the propagation length is the Rowland circle diameter R .

2.1.3. Structural Parameter Design

For constructive interference to occur at a specific position on the output facet of the slab waveguide, the phase difference for beams from different channels of the arrayed input waveguide arriving at the same focal point must be an integer multiple of 2π . This satisfies the grating equation:

$$\beta_c L_0 + \beta_s \left(R - \frac{d_g}{2} \sin \theta_0 \right) = \beta_0 d_g \sin \theta + \beta_c L_0 + \beta_s \left(R + \frac{d_g}{2} \sin \theta_0 \right) - 2m\pi, \quad (16)$$

where β_0 , β_s and β_c are the propagation constants of the beam in air, in the slab waveguide, and in the arrayed waveguides, respectively. The parameter θ represents the angle between the input beam and the central axis of the arrayed waveguide structure, i.e., the AOA. The parameter θ_0 represents the angle between the output waveguide and the axis of the slab waveguide's output facet. Other parameters and their meanings are given in Table 1.

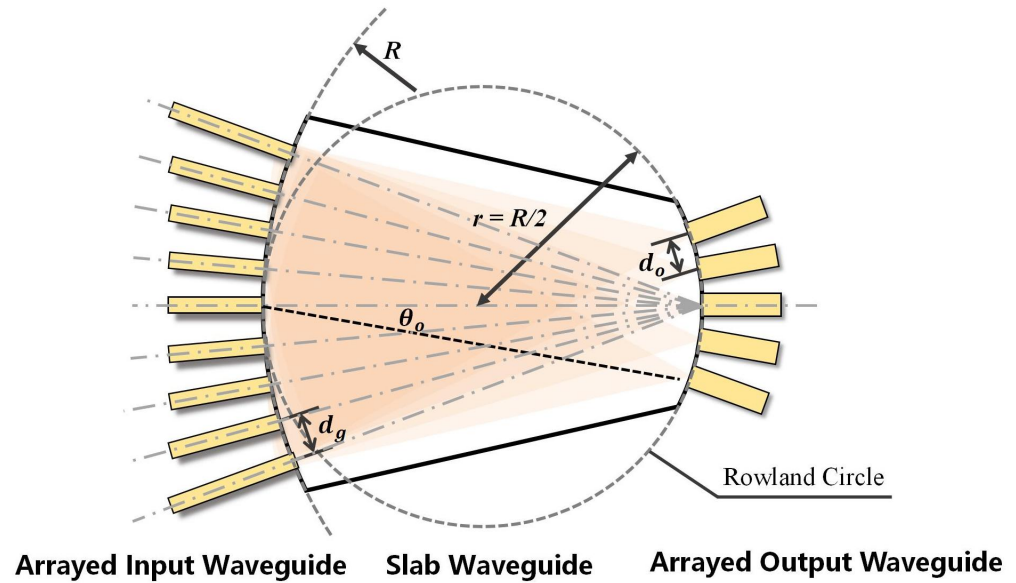


Figure 3. Structural details of the slab waveguide in the arrayed waveguide AOA detection model.

We note that the terms $\beta_c L_0$ and $\beta_s R$ cancel on both sides. Rearranging the remaining terms and using the relations $\beta_0 = \frac{2\pi}{\lambda}$ and $\beta_s = \frac{2\pi n_s}{\lambda}$, where n_s is the effective refractive index of the slab waveguide, we obtain the basic grating equation for the arrayed waveguide AOA detection model:

$$d_g \sin \theta + n_s d_g \sin \theta_o = m\lambda, \quad (17)$$

- Angular and Positional Sensitivity to AOA

The grating equation shows that a change in the incident angle will lead to a difference in the beam's output angle, ultimately causing it to focus at a different position. Therefore, an angular-AOA sensitivity equation can be formulated to describe the relationship between the angle of the diffraction maximum and the AOA. Similarly, a positional-AOA sensitivity equation can be used to describe the correspondence between the focal position of the output beam and the AOA.

By differentiating the grating equation with respect to the incident angle θ , we get

$$d_g \cos \theta + n_s d_g \cos \theta_o \frac{d\theta_o}{d\theta} = 0, \quad (18)$$

Thus, we can obtain the angular-AOA sensitivity equation:

$$\frac{d\theta_o}{d\theta} = -\frac{\cos \theta}{n_s \cos \theta_o} \quad (19)$$

Similarly, the positional-AOA sensitivity equation can be derived:

$$\frac{dx}{d\theta} = R \frac{d\theta_o}{d\theta} = -\frac{R \cos \theta}{n_s \cos \theta_o} \quad (20)$$

The angular resolution of the device is influenced by structural parameters such as the spacing and refractive index of the arrayed waveguides. For instance, a larger waveguide spacing and a smaller refractive index can effectively increase the separation between the maxima of interference patterns formed by beams with different incident angles, thereby enhancing angular resolution. Concurrently, a larger clear aperture can produce sharper and finer interference fringes, further improving the accuracy and robustness of angle estimation.

- Angular Detection Range

We define $\Delta\theta$ as the difference between the angles of two incident beams of different directions from adjacent diffraction orders that focus at the same final position. Starting from the grating equation, when the difference in the direction angle of the input beam is $\Delta\theta$ the m -th diffraction order of the central wavelength coincides with the $(m + 1)$ -th order:

$$d_g \sin(\theta + \Delta\theta) + n_s d_g \cos\theta_0 = (m + 1)\lambda, \quad (21)$$

$$d_g \sin\theta + n_s d_g \cos\theta_0 = m\lambda, \quad (22)$$

Subtracting the two equations gives

$$\Delta\theta = \sin^{-1}\left(\frac{\lambda_c}{d_g} + \sin\theta\right) - \theta, \quad (23)$$

From the analysis of the above formula, it is known that the angular detection range is related to the incident angle, wavelength, and the spacing of the arrayed input waveguides. Furthermore, when $\theta = \sin^{-1}\left(-\frac{\lambda_c}{2d_g}\right)$ the angular detection range has an extremum. At this point, it can be calculated that

$$\Delta\theta = 2\sin^{-1}\left(\frac{\lambda_c}{2d_g}\right), \quad (24)$$

This formula can be used to approximate the angular detection range.

2.2. Angle Estimation Methods

For the fringe interference light field at the output of the planar waveguide, this study investigates the use of arrayed output waveguides for mode coupling. Ultimately, through intensity sampling, the light field coupling efficiency on different waveguide channels is calculated, and the light intensity received by different channels is outputted to obtain a light intensity sequence of length N_o . In response to the intensity sequence signal generated by simulation, this study proposes an end-to-end regression method based on a CNN–Attention model for angle estimation and compares its performance with the conventional peak–linear calibration method as well as other deep learning architectures.

2.2.1. CNN–Attention Regressor Method

This paper adopts a hybrid neural network architecture for angle regression, designed to accurately predict the AOA of incident free-space light from fringe data. This network combines the feature extraction capabilities of a convolutional neural network with the sequence modeling advantages of a multi-head self-attention mechanism. The overall architecture of the model and the process of data dimensionality change within the network are detailed in Figure 4. The model is primarily composed of the following three core modules: a CNN feature extractor, a Transformer-based encoder (including multi-head attention and a feed-forward network), and an MLP regression head. Through the synergistic action of these modules, the model can effectively extract light intensity sequence features from the input data and learn the complex relationships between these features, ultimately achieving precise angle prediction.

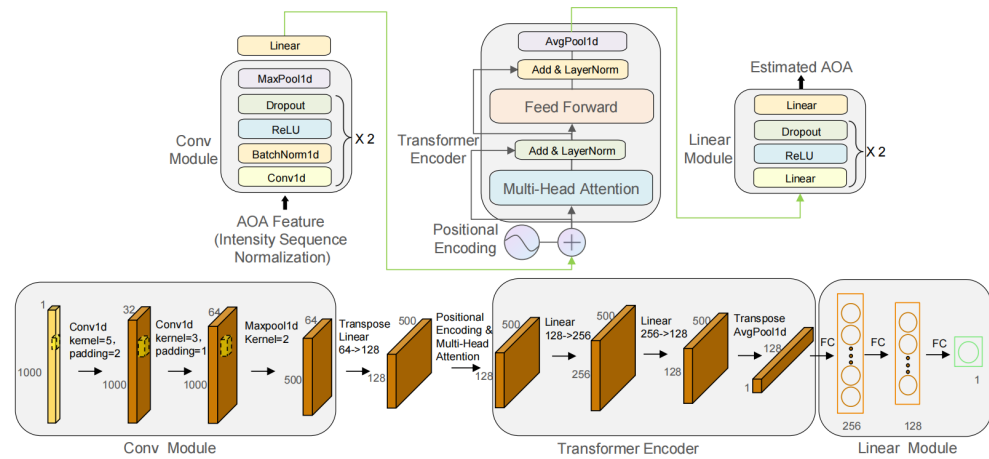


Figure 4. The architecture and data flow diagram of the proposed CNN-Attention regressor.

- CNN Feature Extractor [25–28]

The CNN feature extraction module is formed by stacking two convolutional layers and max-pooling layers. Each convolutional layer is followed by a batch normalization layer, a ReLU activation function, and a Dropout layer to extract local interference pattern features from the input data and increase the model's robustness. The max-pooling layer is used to reduce the spatial resolution of the feature maps, decrease the computational load, and extract more representative features. Through multi-layer convolution and pooling operations, the model can effectively capture the dependencies within the interference pattern in the input data. To adapt the feature maps output by the CNN to the Transformer decoder architecture, the model uses a linear layer to transform the dimensionality of the convolutional features into the dimension required by the attention mechanism, thereby better capturing the relationships between features.

- Transformer Encoder [29]

This module is the core of the model, used for modeling the global dependencies within the feature sequence. Since the Transformer's self-attention mechanism itself lacks the ability to perceive sequence order, this study employs positional encoding using sine and cosine functions. This embeds the absolute or relative position information of each position in the sequence into the feature vectors, enabling the model to perceive positional information. The encoding formulas are as follows:

$$PE(pos, 2i) = \sin(pos / (10000^{2i/d_{model}})), \quad (25)$$

$$PE(pos, 2i + 1) = \cos(pos / (10000^{2i/d_{model}})), \quad (26)$$

where PE is the final output positional encoding data, pos represents the position in the sequence, i represents the dimension index, and $d_{model} = 128$ represents the feature dimension.

Subsequently, by computing multiple attention heads in parallel, the model can capture complex relationships between different positions in the input sequence. Each attention head independently calculates attention weights and transforms the input sequence into a weighted representation. This paper uses Scaled Dot-Product Attention with the following calculation formula:

$$\text{Attention}(Q, K, V) = \text{softmax}(QKT / \sqrt{dk})V, \quad (27)$$

where Q , K , and V are the query, key, and value matrices, respectively, and dk is the dimension of the key vectors.

The final part of the encoder is a feed-forward network. The feed-forward neural network consists of two linear layers and a ReLU activation function, used for non-linear transformation and feature reshaping of the attention mechanism's output.

- MLP Regression Head

After the Transformer encoding, we obtain a feature sequence containing global information. To perform regression prediction, we first apply global average pooling to all position features in the sequence to get a global feature vector of length 128. Finally, the model uses an Multi-Layer Perceptron (MLP) regression head to map the extracted features to the target angle value. This regression head consists of multiple linear layers, ReLU activation functions, and Dropout layers to achieve precise angle prediction.

2.2.2. Benchmark Models for Comparison

To comprehensively evaluate the proposed CNN–Attention architecture and to understand the contribution of its components, two additional deep learning models were implemented as benchmarks: a standalone Convolutional Neural Network (CNN) and a Long Short-Term Memory (LSTM) network. Both models take the 1000-element intensity sequence as input and output a single angle value.

- CNN Model

This model is designed to assess the effectiveness of convolutional feature extraction alone for AOA estimation. It begins with a feature extractor consisting of two stacked 1D convolutional layers, each followed by ReLU activation, batch normalization, and dropout. A max-pooling layer is applied after the second convolutional block, effectively halving the sequence length. The output features are then transformed by a linear layer to a dimension of 128. To replace the attention mechanism, these features undergo a ReLU activation and are subsequently aggregated across the sequence dimension using global average pooling. Finally, an MLP regression head, comprising three linear layers, ReLU activations, and dropout, predicts the final angle. This architecture closely mirrors the CNN feature extraction and MLP regression head components of our proposed CNN–Attention model, but without the Transformer Encoder module.

- LSTM Model

This model evaluates the capability of recurrent neural networks to capture sequential dependencies in the intensity data. The 1000-element input intensity sequence is first projected to a hidden dimension of 128 via a linear layer. This projected sequence is then fed into a stack of two LSTM layers, each with 128 hidden units. The hidden state from the last time step of the final LSTM layer is extracted and passed to an MLP regression head. This head consists of two linear layers, ReLU activation, and dropout, ultimately outputting the predicted angle.

2.2.3. Peak–Linear Calibration Method

To highlight the advantages of the deep learning method adopted in this paper, we chose a conventional method based on a physical optics model and signal processing techniques as a comparative baseline. The core idea of this method is to utilize the linear correspondence between the position of the principal maximum (i.e., the peak) in the light intensity sequence and the AOA of the spatial light for fitting and prediction [35].

This method mainly includes two stages:

- Peak Position Detection

By setting a specific angle range, it can be ensured that only one principal maximum exists in the light intensity sequence within this range. This allows for the establishment of a one-to-one mapping relationship between the magnitude of AOA and the peak position. By searching and locating the position of the principal maximum peak in the signal, a basis for the model's subsequent fitting is provided.

- Mapping Relationship and Angle Estimation

Through validation with simulated data, we found that the peak position x and the AOA y exhibit an approximately linear relationship throughout the measurement range. Therefore, we use a linear function combined with the least squares method to perform a global fit on a set of data to establish the functional relationship between the AOA and the peak position $y = f(x)$. This linear model is represented as $y = f(x) = kx + b$.

2.3. Error Models

In the simulation framework for AOA detection in arrayed waveguides, this study explicitly modeled phase perturbations and manufacturing tolerances to evaluate the sensitivity of the end-to-end output light field intensity sequence to multi-source errors.

2.3.1. Phase Noise

Phase perturbations within the arrayed input waveguide, arising from factors such as temperature drift, mechanical stress, and refractive index fluctuations, are modeled as an additive phase screen. In this simulation, this is primarily implemented by constructing a pixel-wise independent Gaussian white noise phase screen $\phi_{phase}(x, y) \sim N(0, \sigma_\phi^2)$.

2.3.2. Manufacturing Tolerances

In this work, manufacturing tolerances are broken down into two categories: global or local channel phase drift caused by arrayed input waveguide length errors and the addition of a linear phase gradient within the channel due to arrayed input waveguide pointing deviations.

The length error term is implemented by adding fixed, uniform random noise $\phi_{length}(N) \sim U(0, 1)$ to different channels of the arrayed input waveguide.

The pointing error term is implemented by adding a linear phase gradient to different channels of the arrayed input waveguide, expressed by the formula:

$$\phi_{orientation}(x) = k_i \sin(\theta_{dev})x, \quad (28)$$

where k_i is a proportionality coefficient following a uniform distribution $k_i \sim U(0, 1)$, and θ_{dev} is the maximum deviation angle of a single rectangular waveguide in the array. The term $k_i \theta_{dev}$ approximates the angle by which the actual rectangular waveguide deviates from its ideal position.

3. Results

This section details the setup of the simulation experiments, including the construction of the dataset and parameter selection, and presents the simulation results. We propose an end-to-end regression method based on a CNN–Attention regressor. This method utilizes a convolutional neural network to automatically extract effective features from the data and employs an attention mechanism to focus on important sequence information, thereby achieving more accurate predictions. The performance differences between this

method and the peak–linear calibration method will be comprehensively evaluated through comparative experiments.

3.1. Simulation Structure Parameters

The simulation parameters are configured as shown in Table 2 below. First, regarding the channel count design, the arrayed input waveguide is set to 5000 to ensure a reasonable optical aperture size. Meanwhile, the number of arrayed output waveguides is set to 1000 to increase the length of the optical intensity sequence, thereby improving the prediction accuracy of the model. In terms of arrayed waveguide dimensions, the spacing between adjacent waveguides is controlled to keep the crosstalk level below 30 dB. The aperture of the arrayed input waveguide is designed to be 1.23 μm to ensure that the waveguide operates in a single-mode transmission state. The aperture of the arrayed output waveguide is designed to be 3 μm to ensure that the receiving waveguide has a high duty cycle.

To simulate the impact of phase noise and manufacturing tolerances, we introduced several error models as described in Section 2.3. For the experiments detailed below, the additive phase noise was modeled with a standard deviation of $\sigma_\phi = \frac{\pi}{4}$. The length error term was set to a range corresponding to a phase shift of $[-\frac{\pi}{2}, \frac{\pi}{2}]$, and the maximum pointing deviation angle was set to 5° . A detailed analysis of the model's robustness to phase noise and other parameter variations is provided in Section 3.4. Based on the above design, the coupling efficiency of the arrayed input waveguide for free-space beams with different angles of arrival fluctuates within the range of 61.4% to 81.2%. The arrayed input waveguide coupling efficiency reaches a maximum value of 81.2% when the AOA is 0° . Based on these simulation parameters, the proposed model can achieve high-precision detection of AOAs within a 40° field of view.

Table 2. Simulation structure parameter values for arrayed waveguide AOA estimation model.

Symbol	Parameter Name	Parameter Value
λ	Optical Signal Wavelength	1550 nm
θ	Incident Beam AOA	$(-20^\circ, 20^\circ)$
w_g	Arrayed Input Waveguide Core Width	1.23 μm
h_g	Arrayed Input Waveguide Core Thickness	1 μm
d_g	Arrayed Input Waveguide Center Spacing	1.91 μm
N_g	Number of Arrayed Input Waveguides	5000
L_0	Center Channel Arrayed Input Waveguide Length	500 μm
R	Rowland Circle Diameter	10 mm
w_o	Arrayed Output Waveguide Core Width	3 μm
h_o	Arrayed Output Waveguide Core Thickness	1 μm
d_o	Arrayed Output Waveguide Center Spacing	3.68 μm
N_o	Number of Arrayed Output Waveguides	1000
z	Gaussian Beam Propagation Distance	1 m
b_w	Gaussian Beam Waist Radius	2 mm
-	Upper Cladding Material	SiO_2
-	Core Material	Si_3N_4
-	Substrate Material	SiO_2

3.2. Simulation Results

To illustrate the fundamental working principle of the device, we first simulate the light field propagation at a 0° AOA, while including phase perturbation noise and manufacturing tolerances. Figure 5 demonstrates the complete journey of the light field through the core components of the architecture.

The process begins with a Gaussian beam serving as the input light field (Figure 5b), incident on the arrayed input waveguide. After propagating through this array, the continuous light field is discretized into 5000 dense, individual beamlets (Figure 5c). While microscopically discrete, the macroscopic intensity profile of these beamlets retains a Gaussian envelope.

These discrete beamlets then enter the free-propagation region (the slab waveguide implementing the Rowland circle). Finally, upon reaching the focal plane of the Rowland circle, the light converges into a single, focused spot. This spot is then sampled by the arrayed output waveguide, generating a light intensity sequence (Figure 5d). The spatial position of this sequence's peak is subsequently used to determine the AoA.

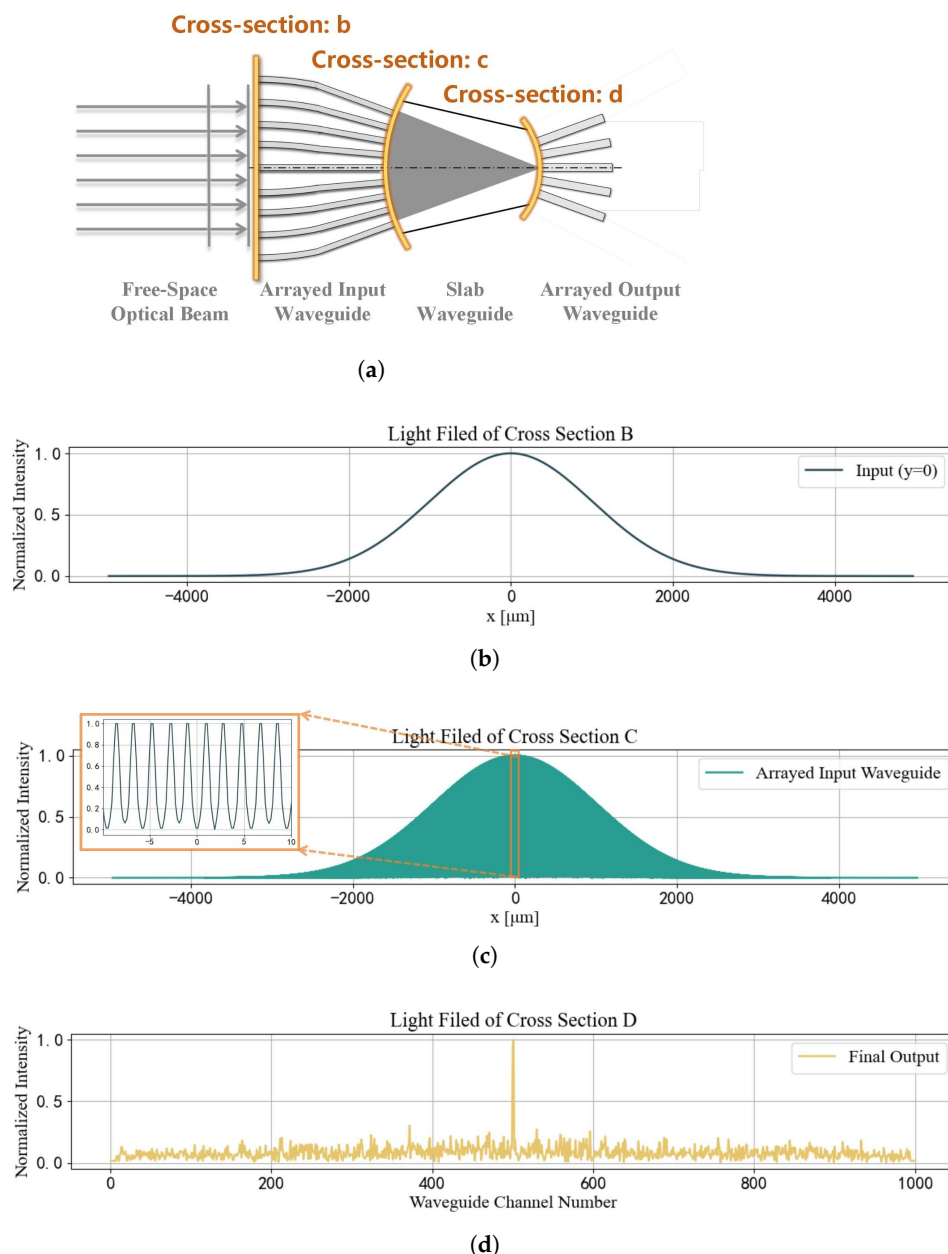


Figure 5. Input and output light field intensity distribution of the arrayed waveguide AOA detection architecture. (a) The cross-sectional position of the following light fields in the arrayed waveguides. (b) Input light field intensity distribution of the arrayed waveguide structure. (c) Output light field intensity distribution of the arrayed input waveguide. (d) Output light field intensity distribution of the arrayed output waveguide.

Figure 6 visualizes the light field evolution within the slab waveguide to illustrate the physical mechanism of beam steering, comparing a 0° on-axis with a 15° off-axis incidence. For the 0° AoA (Figure 6a), the light propagates symmetrically to form a focus at the center of the output arc. In contrast, the input wavefront's phase tilt at 15° AoA (Figure 6b) steers the entire propagation path, resulting in a correspondingly shifted focus spot. This comparison vividly demonstrates the device's core function: deterministically mapping the input AOA to a unique spatial position.

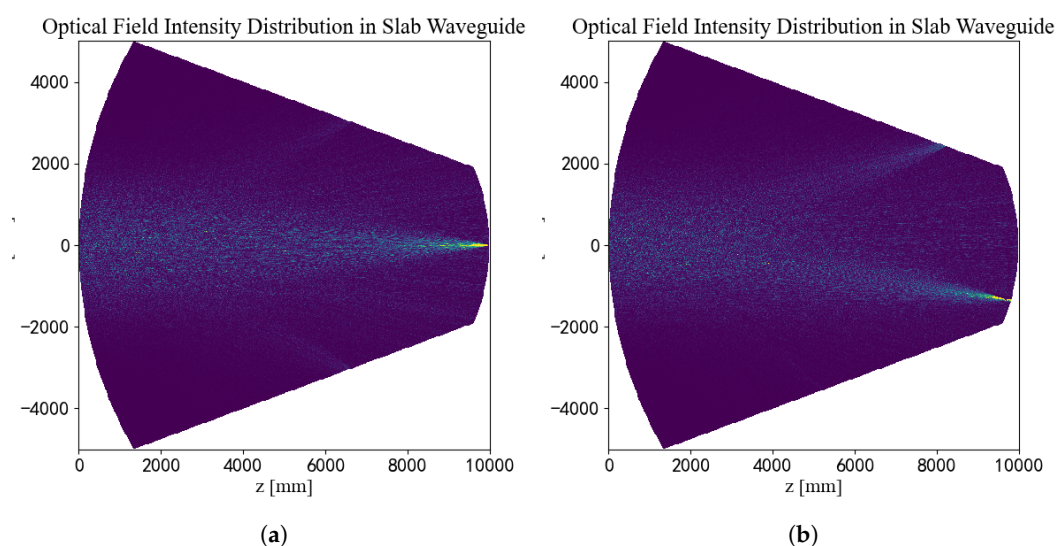


Figure 6. Visualization of the light field intensity evolution within the slab waveguide for different input angles. (a) For a 0° AoA, the beam focuses symmetrically onto the center. (b) For a 15° AoA, the entire propagation path is tilted, shifting the final focus spot to a corresponding off-center location.

Building on this principle, we selected different AOA values to show the corresponding final output light intensity sequence captured by the arrayed output waveguide, as shown in Figure 7. It can be seen that the position and intensity distribution of the focused spot vary with the angle. The greater the difference in the AOA, the more obvious the difference in the spatial position of the main intensity peak, which is the direct consequence of the beam steering effect illustrated in Figure 6.

3.3. AoA Estimation

To evaluate AoA estimation performance, this study compares a traditional signal processing method with a deep learning approach. For this purpose, a total of 2000 datasets were established for training and testing the two methods.

3.3.1. AoA Estimation by Peak–Linear Calibration

The traditional approach investigated is the peak–linear calibration method. This method estimates the AoA by first identifying the waveguide channel with the highest output intensity in the arrayed output waveguide. A linear mapping is then established between the position of this peak intensity channel and the corresponding angle. For this method, which does not require a validation set, the dataset was divided into 1700 samples for the training set and 300 samples for the test set.

The results of this method are summarized in Figure 8. The linear regression performed on the training data, shown in Figure 8a, establishes the mapping function. It exhibits a strong linear correlation, suggesting a good initial fit to the data.

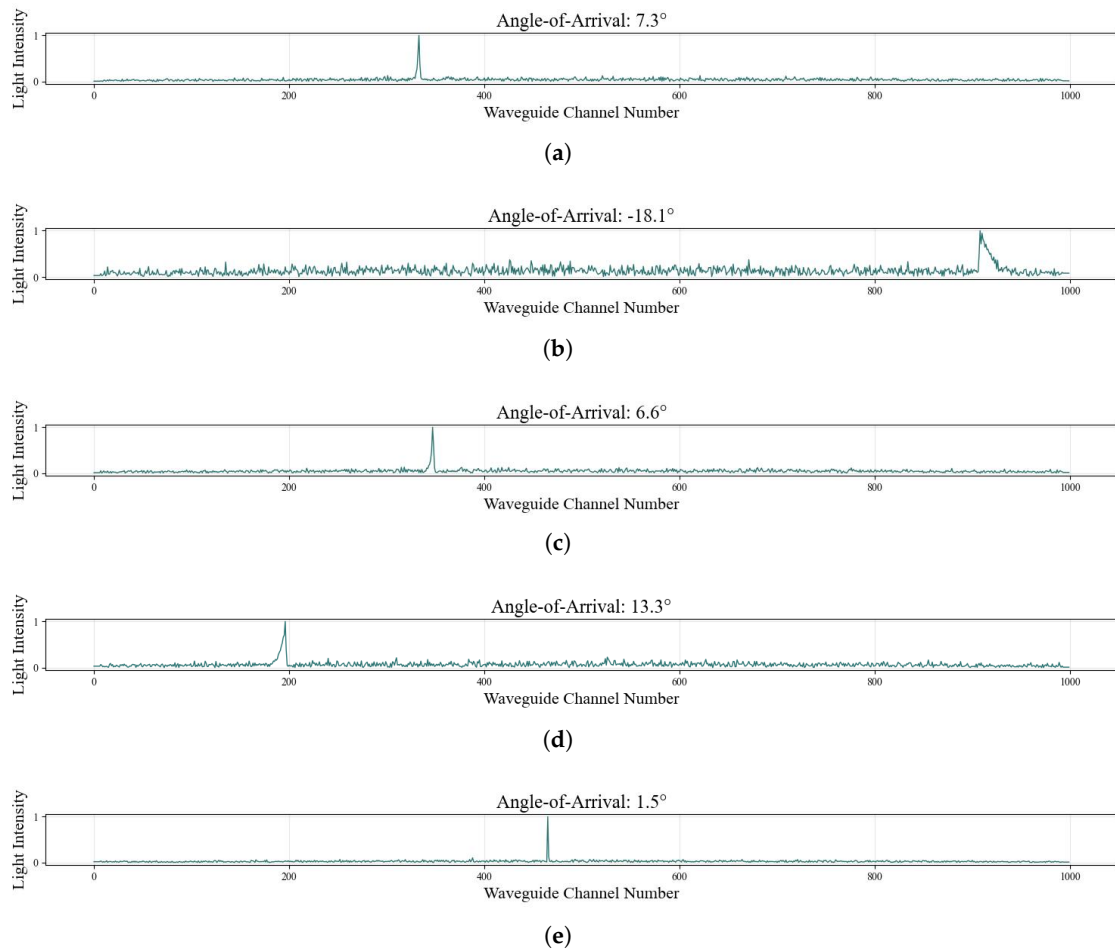


Figure 7. Examples of AOA and arrayed output light field intensity sequences in the arrayed waveguide AOA detection architecture. (a) Light intensity sequence at 7.3° AOA. (b) −18.1°. (c) 6.6°. (d) 13.3°. (e) 1.5°.

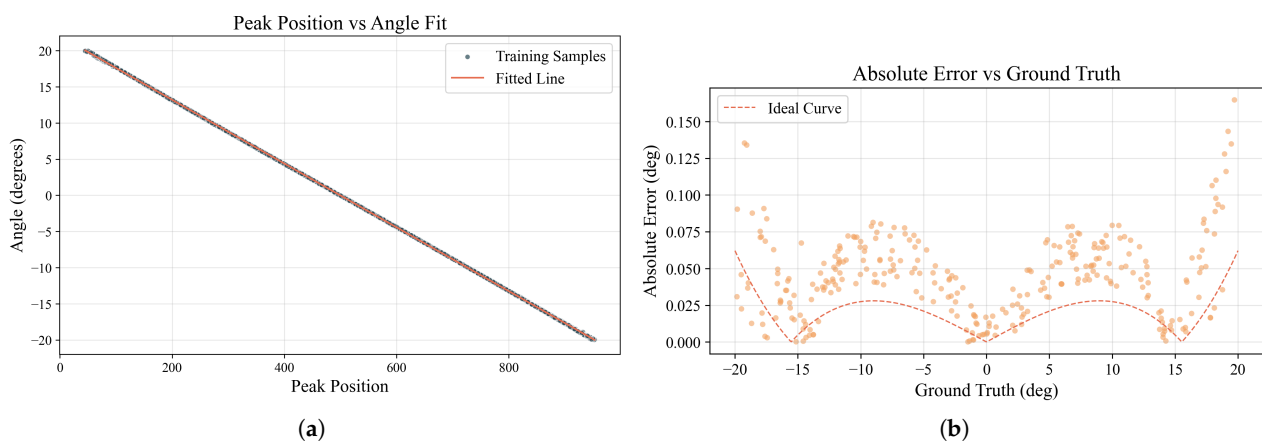


Figure 8. Performance evaluation of the peak–linear calibration method. (a) The linear regression fit between the peak channel index and the AoA for the training data. (b) The absolute error distribution as a function of the true AoA for the test data, revealing systematic errors at larger angles. The theoretical absolute error curve (dashed line) is also presented, which is derived by applying a linear approximation to the device’s inherently nonlinear grating equation.

A more detailed analysis of the model’s performance on the test set, illustrated in Figure 8b, reveals inherent operational boundaries. The plot of absolute error versus ground truth AoA exhibits a distinct, symmetrical “W-shaped” dependency. The model

achieves its highest accuracy in the low-error troughs of this “W” shape, located at AoAs of approximately 0° and $\pm 15^\circ$. In stark contrast, the most significant performance degradation occurs at the extreme edges of the field of view (as AoA approaches $\pm 20^\circ$). This nonuniform, angle-dependent accuracy underscores the need for a more sophisticated modeling approach to ensure consistently high performance across the entire operational range.

3.3.2. AOA Estimation by CNN–Attention Regressor

The conventional peak–linear calibration method serves as an effective benchmark, yet to push the boundaries of angle prediction accuracy, we developed a sophisticated deep learning model. This model is a regression-focused Convolutional Neural Network featuring an attention mechanism (CNN–Attention). The dataset for this endeavor was allocated as follows: 1400 samples for training, 300 for validation, and 300 for testing.

During the model training process, the mean squared error (MSE) was selected as the loss function. To further optimize the training effect of the model, the Adam optimization algorithm was adopted, which can effectively handle non-convex optimization problems. A cosine annealing scheduler was used for the learning rate to smoothly adjust it. By setting the batch size to 32 and the number of training epochs to 1000, the model is ensured to undergo sufficient training, completing the extraction of diffraction pattern features. As shown in Figure 9, the training process was stable, with both training and validation losses steadily converging to a low value. The small gap between the two curves demonstrates the model’s effective generalization and absence of overfitting.

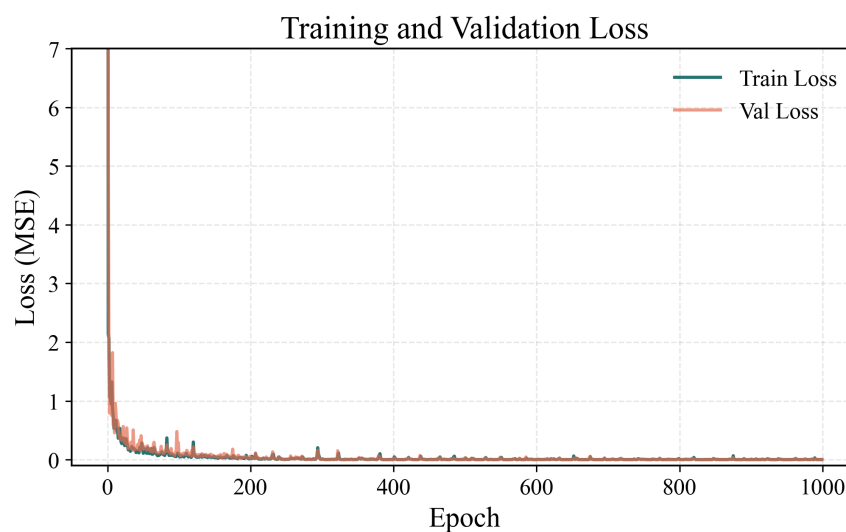


Figure 9. Loss curves for the training and validation sets during model training.

For practical benchmarking, the computational time for training was recorded. The model was trained on a workstation equipped with an Intel Core i9-10900KF CPU (Intel Corporation, Santa Clara, CA, USA), 32 GB of RAM, and an NVIDIA GeForce RTX 3060 (12 GB) GPU (NVIDIA Corporation, Santa Clara, CA, USA). The entire training process for 1000 epochs completed in just 21.73 min. This efficient training time underscores the model’s practicality for rapid development and deployment.

The prediction performance of the model is visualized in Figure 10. We obtained high-precision AOA prediction results, with a MAE of 0.0142° , a RMSE of 0.0193° , and a coefficient of determination R^2 of 0.999997.

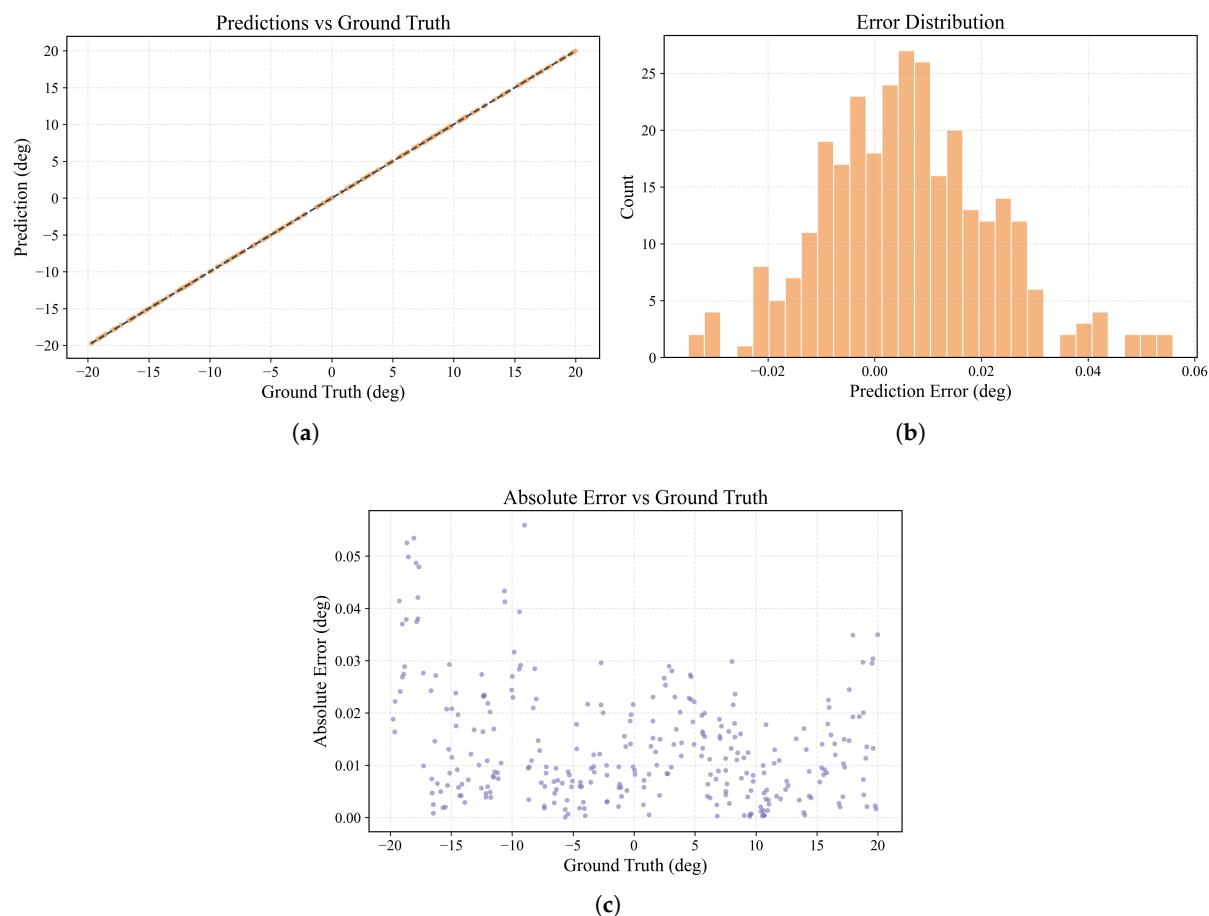


Figure 10. The CNN–Attention regression method angle prediction effect display. (a) Predicted vs. Ground truth angles. (b) Error distribution histogram. (c) Error magnitude distribution across angular range.

3.3.3. Performance Comparison

To comparatively analyze the performance of the two methods, we conducted three independent trials. The comparison of the angle prediction performance of the two schemes in different trials is presented in Table 3, and a histogram is used to visually display the angle prediction effect of the two methods, as shown in Figure 11.

From the charts, we can observe that both methods have high R^2 values, which means that this arrayed waveguide AOA detection architecture has excellent AOA estimation accuracy and robustness, and can accurately recover the initial AOA information of the model from the influence of noise such as phase interference and manufacturing tolerances. Furthermore, we observe that in the three trials, the angle prediction performance of the CNN–Attention regression method is significantly higher than that of the conventional peak–linear calibration method, and it has better effects under different performance evaluation indicators. Finally, in terms of the MAE value, the AOA prediction effect of the deep learning method is 3.25 times that of the conventional method; for the RMSE value, the improvement is 2.81 times.

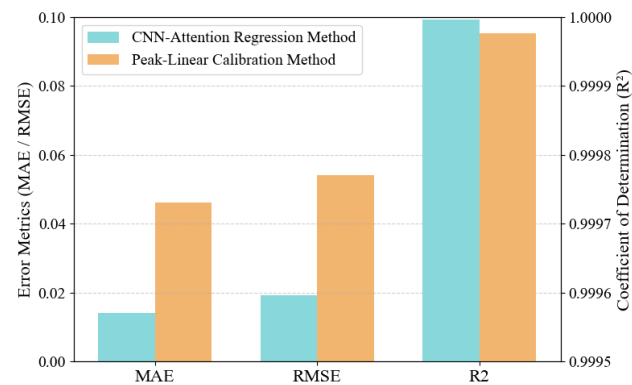


Figure 11. Comparison histogram of key performance indicators of two AOA estimation methods (MAE: Mean Absolute Error; RMSE: Root Mean Square Error; R^2 : Coefficient of determination).

Table 3. Comparison of evaluation metrics across different methods. MAE, RMSE, and R^2 stand for Mean Absolute Error, Root Mean Square Error, and the coefficient of determination, respectively.

Evaluation Metrics		MAE		RMSE		R²	
Method	Peak-Linear Calibration Method	CNN-Attention regression method	Peak-Linear Calibration Method	CNN-Attention regression method	Peak-Linear Calibration Method	CNN-Attention regression method	
Trial 1	0.0461	0.0134	0.0542	0.0173	0.999977	0.999998	
Trial 2	0.0461	0.0153	0.0542	0.0214	0.999977	0.999996	
Trial 3	0.0461	0.0140	0.0542	0.0193	0.999977	0.999997	
Mean	0.0461	0.0142	0.0542	0.0193	0.999977	0.999997	

The selection of the hybrid CNN-Attention architecture is further justified by a detailed comparison against other common deep learning models: a standalone CNN and an LSTM. This analysis, summarized in Table 4, provides a comprehensive evaluation across multiple metrics, including prediction accuracy, model complexity, and computational efficiency.

A clear performance hierarchy emerges from this comparison. The standalone CNN model performed the worst, with a MAE of 0.4122° . While CNNs are excellent at extracting local features from the interference fringes, they struggle to comprehend the global context—that is, the absolute position of the entire fringe pattern, which is the primary indicator of the AOA. This fundamental limitation leads to significantly larger prediction errors.

The LSTM model, which processes the data as a sequence, achieved a much better MAE of 0.0515° . This performance is comparable to that of the direct physics-based peak detection algorithm, suggesting that a standard sequence model can effectively capture the overall shift of the intensity peak.

In stark contrast, the proposed CNN-Attention model proves highly effective by synergizing the strengths of both approaches. The CNN component first extracts rich local pattern features, and the attention mechanism subsequently weighs the importance of these features across the entire sequence. This allows the model to overcome the CNN's limitation by effectively identifying the global position of the most salient features, resulting in a superior performance that validates our architectural choice.

Notably, our proposed model achieves this state-of-the-art accuracy with a more efficient parameter count than the LSTM. While its inference time of 1.03 ms is marginally higher than the others, it remains well within the threshold for real-time applications, thus

offering the best overall balance of accuracy, model efficiency, and practical usability. This result strongly validates our architectural choice.

Table 4. Detailed comparison of different deep learning architectures.

Model	MAE (°)	RMSE (°)	Parameters (k)	Training Time (s/Epoch)	Inference Time (ms/Sample)
CNN	0.4122	0.6040	80.96	0.24	0.98
LSTM	0.0515	0.0696	272.77	1.18	0.69
CNN-Attention (Proposed)	0.0142	0.0193	213.44	1.30	1.03

3.4. Robustness Analysis

To validate the model's suitability for real-world implementation, we performed a comprehensive robustness analysis. This evaluation investigates the model's resilience against deviations in key operational parameters, structural imperfections arising from fabrication, and noise inherent in signal measurement, directly addressing key practical benchmarks.

3.4.1. Robustness to Phase Noise and Fabrication Imperfections

First, we assessed the model's resilience to signal noise and common manufacturing tolerances.

- Phase Noise

Our physical simulations indicate that the sampled optical signal maintains a distinguishable structure even when the standard deviation of the phase screen (σ_ϕ) is increased to a tested maximum of 0.6π . Under this severe noise condition, the model's prediction performance degrades, yielding a MAE of 0.704° . While this error is higher than the excellent results achieved at $\sigma_\phi = \frac{\pi}{4}$ (reported in Section 3.3.2), it defines the operational limit of the model under extreme phase perturbations.

- Manufacturing Tolerances

We evaluated the model's performance under the waveguide length errors and pointing deviations defined in our simulation setup. The results confirm that the model maintains high prediction accuracy even in the presence of random length errors (causing phase shifts in the range of $[-\frac{\pi}{2}, \frac{\pi}{2}]$) and pointing deviations up to 5° . This demonstrates that the proposed method effectively compensates for these standard fabrication and alignment imperfections without performance degradation.

3.4.2. Robustness to Operational and Structural Parameter Variations

In addition to the above analysis, we further investigated the model's generalization capability when faced with parameters outside its training distribution. Specifically, we analyzed its performance under variations in operating wavelength and waveguide spacing (d_g).

- Wavelength Variation

A test set was generated using a wavelength that deviated by 10% from the central wavelength used for training. Under this condition, the model achieved a MAE of 0.0775° . As illustrated in Figure 12a, the prediction accuracy is notably higher within a smaller

angular range (e.g., $\pm 10^\circ$), where the MAE is only 0.0329° . This indicates a strong robustness to moderate wavelength fluctuations.

- Structural Variation

We simulated a 10% deviation in the waveguide spacing to simulate significant structural variations or fabrication errors. The test yielded a remarkably low MAE of 0.0134° . This extremely low error value indicates that our trained model possesses outstanding robustness to changes in waveguide spacing.

It is noteworthy that when parameters like wavelength or waveguide spacing undergo further modification, the effective angular detection range of the device, as defined by Equation (24), also changes. For instance, when we tested the model at a wavelength of 1064 nm, which is far from the training data, we found that angle prediction remained accurate within a $\pm 15^\circ$ range. However, for angles between 15° and 20° , the model failed to produce valid predictions, as shown in Figure 12b. This confirms that the model's prediction range is limited by the device's physical laws, rather than a failure of the model itself.

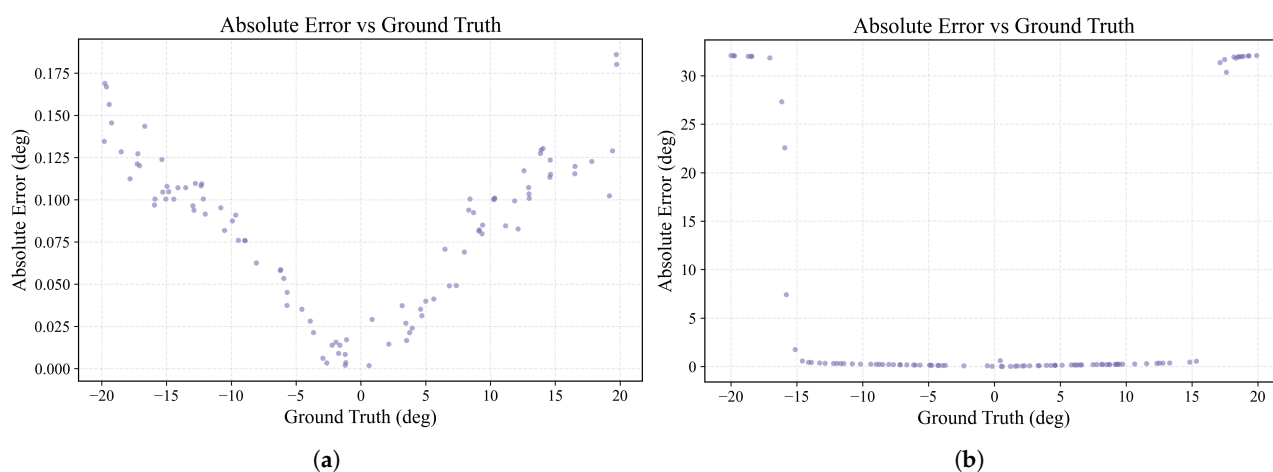


Figure 12. Absolute prediction error under wavelength variations. (a) Error distribution with 10% wavelength deviation. (b) Error distribution at 1064 nm wavelength, far outside the training distribution. Predictions remain accurate within the $\pm 15^\circ$.

3.4.3. Summary of Robustness

In summary, these comprehensive tests confirm that our method is not only highly accurate under ideal conditions but also robust against a range of real-world challenges. These include operational parameter shifts (wavelength), structural variations (waveguide spacing), signal noise (phase noise), and other fabrication imperfections (length and pointing errors). This multi-faceted robustness reinforces the model's strong potential for practical deployment.

4. Discussion

This study successfully demonstrates a novel AOA detection method based on arrayed waveguide coherent mode decoding. This method transforms the spatially distributed phase information of a FSO AOA into spatially distributed intensity information and employs an end-to-end regression method based on a CNN–Attention model to recover the beam's AOA from the intensity sequence. Simulations have verified its excellent performance in high-precision angle prediction, particularly demonstrating strong robustness against various sources of interference.

4.1. Underlying Principles of High-Precision Detection

The high-precision AOA prediction performance of this method can be attributed to the synergistic effect of physical optics principles and a deep learning model. From a physical perspective, the core mechanism lies in the multi-beam interference phenomenon guided by the arrayed waveguide structure. When beams from different waveguides are aligned to the same detection position within the slab waveguide, coherent combination occurs, concentrating energy at the center of the principal maximum and forming a distinct principal maximum interference fringe.

On the computational level, the CNN–Attention regression method surpasses traditional methods that rely solely on peak detection. It is capable of learning and extracting complex nonlinear features from the complete signal output by the arrayed output waveguide. This capability allows the model to accurately estimate the AOA even when the signal is distorted by noise. Furthermore, through its powerful feature extraction capabilities, the CNN–Attention regressor can learn to compensate for signal variations caused by specific manufacturing tolerances, thus exhibiting stronger robustness against device fabrication errors.

The systematic “W-shaped” error from the peak–linear method (Figure 8b) is physically explained by the device’s governing grating equation, $\theta_0 = -\arcsin(\sin \theta/n_s)$ (Equation (19)). This inherent nonlinearity means that forcing a linear fit between the input AOA and the output peak position introduces a predictable, systematic error, whose theoretical shape is shown in Figure 8b. While the linear approximation is adequate near the center of the field-of-view, the pronounced nonlinearity at the angular extremes causes large deviations, resulting in the observed “W-shaped” error pattern. The proposed CNN–Attention regressor, however, is not bound by this linear constraint. It directly learns the complex nonlinear mapping from the intensity sequence, thus avoiding such systematic biases and achieving uniformly high accuracy across the entire detection range.

4.2. Advantages and Application Prospects

Compared to traditional AOA detection technologies, the method proposed in this study demonstrates significant advantages and broad application prospects in several aspects. A core contribution is its high potential for miniaturization. Based on mature Planar Lightwave Circuit (PLC) technology [36,37], the entire arrayed waveguide structure can be compactly integrated on a chip, providing a highly integrated AOA estimation solution for portable and embedded applications.

A second key advantage is the significant reduction in system complexity. The system’s optical front-end is entirely passive, obviating the need for the active components (e.g., phase shifters, dynamic beam steering) or complex beamforming networks common in traditional phased arrays. This architecture strategically transfers the primary complexity from the difficult-to-integrate optical domain to the mature and robust domain of digital electronic processing, where an active processing unit executes the decoding algorithms. As a result, this approach not only leads to a substantially simplified optical architecture but also requires only a single, non-recurring calibration, significantly reducing both deployment and operational complexity.

Regarding computational efficiency for real-time systems, our model’s inference time of 1.03 ms is sufficient for many applications like robotic navigation and beam tracking, representing a favorable trade-off between the substantial accuracy gain and modest computational cost. Both its input and output ends are arrayed waveguide structures, giving it great potential for integration with existing waveguide devices like waveguide combiners. This integration requires no additional optical interfaces and holds the promise of using multiple arrayed waveguide AOA detection structures for corresponding channel

combinations to enhance optical signal energy for long-range detection. It also offers the possibility of simultaneously realizing multiple functions on a single chip, such as WDM and spatial direction sensing, thus opening new pathways for the construction of next-generation multi-functional photonic integrated systems.

4.3. Limitations and Future Work

Despite the promising simulation results, it is important to recognize the applicability boundaries and limitations of the current model, which in turn point to key directions for future research.

4.3.1. Device Specificity and Wavelength Dependency

Our analysis, grounded in the grating equation, reveals that the model's performance exhibits strong robustness to minor manufacturing variations but requires retraining for systematic changes to core parameters like wavelength. This distinction in sensitivity to geometric parameters versus wavelength has direct implications for device design and generalization. Consequently, future research will prioritize overcoming this wavelength dependency by developing multi-wavelength systems, either through advanced multi-task learning models or innovative broadband waveguide structures.

- Wavelength Dependency and Material Dispersion

The model's wavelength dependency primarily stems from material dispersion. Although our design utilizes the zeroth-order diffraction ($m = 0$) where the grating equation $d_g \sin \theta + n_s d_g \sin \theta_o = m\lambda$ appears wavelength-independent, the effective refractive index varies with wavelength due to material dispersion. This variation alters the output optical field's phase profile, consequently affecting prediction accuracy. This explains the observed increase in prediction error when the wavelength deviates from the training value, as detailed in Section 3.4.2.

- Robustness to Waveguide Spacing Variations

In contrast, the model demonstrates exceptional robustness to variations in waveguide spacing d_g . Under the zeroth-order diffraction condition, the influence of d_g is significantly diminished. The neural network has effectively learned a phase-to-angle mapping that is inherently insensitive to d_g variations, maintaining high prediction accuracy even with a 10% deviation in d_g .

- Physical Limits on Model Generalization

However, the model's generalization capability remains fundamentally constrained by physical principles. The detectable angular range $\Delta\theta$ is ultimately limited by the diffraction limit $\Delta\theta = 2 \arcsin(\lambda/2d_g)$. Experimental verification at 1064 nm confirmed accurate predictions within the theoretical range of approximately $\pm 15^\circ$, with failure beyond this boundary (Figure 12b). This confirms that the model's operational limits are physically determined rather than being solely defined by the training data range.

4.3.2. Limitations of the Simplified Simulation Model

The current study relies on a simplified simulation model. A primary limitation, as noted in Section 2, is the use of a 1D simulation model as a first-order approximation. While this approach was instrumental for demonstrating the core principle and training the deep learning regressor, transitioning to a full 2D model and a physical device will introduce several factors that must be considered. Our future work will involve detailed simulations and experimental validation to address these factors. The key aspects are outlined below.

- Asymmetry in Angular Resolution

The current study assumes a 1D variation in AOA. In a practical 2D scenario with both azimuth and elevation angles, the rectangular cross-section of the waveguides ($1.23\ \mu\text{m} \times 1\ \mu\text{m}$) will lead to different effective mode field diameters in the horizontal and vertical directions. This asymmetry will cause the diffraction-limited spot size at the output to be different for the two axes, likely resulting in non-uniform angular resolution. For instance, the smaller vertical dimension could lead to a larger diffraction angle and thus a lower resolution in the elevation direction compared to the azimuth.

- Increased Crosstalk and Scattering

Transitioning from a 1D linear array to a 2D planar array fundamentally alters the crosstalk environment. In a 2D array, each waveguide is surrounded by a larger number of neighbors compared to the two adjacent ones in a 1D model. While the strongest coupling still occurs between orthogonally adjacent waveguides, the cumulative crosstalk power received by a single waveguide increases due to the contribution from all surrounding neighbors. The overall increase in integrated crosstalk can lead to a higher optical noise floor, potentially broadening the main interference peak at the output and reducing the signal-to-noise ratio. Furthermore, out-of-plane scattering losses, an inherent 3D effect primarily caused by sidewall roughness [38], are not captured by the 1D model and could reduce the overall device efficiency in a physical implementation.

4.3.3. Unmodeled Physical Effects and Environment Factors

While our proposed model demonstrates robustness against fabrication tolerances, a comprehensive analysis requires the consideration of additional physical phenomena and environmental factors. In this section, we address these potential deviations and discuss strategies to mitigate them in practical experiments.

- Power Fluctuations and Propagation Loss

To evaluate the model's robustness against material absorption and scattering losses, we performed a supplementary simulation incorporating a typical propagation loss of 1 dB/cm. Additionally, to account for optical power instability, we introduced multiplicative noise drawn from a Gaussian distribution ($\mu = 1$, $\sigma = 0.05$) to the power of each output channel. The resulting MAE for AOA estimation was 0.0168° , which is only marginally higher than the ideal case of 0.0142° . This result demonstrates the resilience of the proposed method, indicating that the model effectively learns the relative intensity distribution rather than relying on absolute power levels.

- Incoherent Background Light

In practical environments, incoherent background light may degrade the signal-to-noise ratio. This influence can be suppressed using standard signal processing techniques. For instance, by modulating the input optical signal, lock-in amplification can be employed to effectively distinguish the coherent signal from static background noise [39]. This strategy will be integral to the design of subsequent experiments.

- Nonlinear Effects

Nonlinear optical effects are currently not included in the model, as the sensing application is designed to operate at low optical power levels where such phenomena are typically insignificant. However, ensuring the device operates within its linear regime is crucial; therefore, verifying the linear dynamic range will be a priority in future experimental characterization.

- Polarization State

The polarization state of the light beam can impact performance due to fabrication-induced birefringence, a well-known challenge in integrated photonics. To ensure

polarization-independent operation in future implementations, we plan to adopt established solutions [21], such as designing waveguides with a square cross-section to minimize birefringence, applying dispersion compensation, or utilizing polarization diversity schemes.

4.3.4. Practical Implementation and Fabrication Challenges

Finally, the physical implementation of the device on a silicon-based or silicon nitride platform presents further practical challenges. The following discussion outlines the proposed implementation route and an analysis of the expected challenges, which will form the basis of our future experimental research.

- Fabrication Process and Tolerance Effects

We propose fabricating the device on a silicon nitride (Si_3N_4 -on- SiO_2) platform, which is well-suited for this application due to its low propagation loss and broad transparency window. The fabrication can be accomplished using standard silicon photonics manufacturing techniques [40–44]: (1) Starting with a silicon wafer with a thick thermal oxide (SiO_2) bottom cladding. (2) Depositing a Si_3N_4 core layer via LPCVD or PECVD. (3) Defining the waveguide pattern using EBL for prototyping or DUV lithography for mass production. (4) Transferring the pattern to the Si_3N_4 layer using RIE. (5) Encapsulating the device with a SiO_2 upper cladding.

However, real-world fabrication is subject to tolerances that will affect device performance. The manufacturing tolerances modeled in our simulation (phase perturbations, length and pointing errors) are direct consequences of physical process variations. For instance, non-uniformity in lithography and etching can lead to variations in waveguide width and height. These dimensional deviations alter the effective refractive index, which in turn affects the phase accumulation and modifies the interference condition. This can cause a shift in the focal spot position and a degradation of the interference fringe contrast. Similarly, sidewall roughness from etching is a primary source of propagation loss and random phase errors [38]. A key objective of our future experimental work will be to characterize the impact of these fabrication-induced non-uniformities and validate the robustness of our CNN–Attention regressor, which is designed to learn and compensate for such systematic deviations.

- Expected Optical Loss

The total on-chip optical loss is another critical parameter to be determined experimentally. We anticipate several major sources of loss [40,41,44,45]: (1) Propagation loss, standard LPCVD processes for stoichiometric Si_3N_4 consistently yield propagation losses in the range of 0.1 to 1.0 dB/cm at the 1550 nm wavelength. Moreover, values below 0.1 dB/cm are routinely achievable through process optimization such as high-temperature annealing. (2) Coupling loss, our simulations in Section 3.1 quantify this input coupling efficiency, which is inherently dependent on the AOA. The efficiency ranges from a maximum of 81.2% (a loss of 0.9 dB) at the center of the field of view to 61.4% (a loss of 2.1 dB) at the edges. In addition to this primary input loss, there are also smaller internal transition losses at the interfaces where light couples from the input waveguides into the slab waveguide, and subsequently from the slab into the output waveguides. (3) Bending loss, which is expected to be negligible given the large radius of curvature in our design. The characterization and minimization of these losses will be a central part of the device's experimental validation.

5. Conclusions

This research aims to address the growing demand for miniaturized, high-precision AOA detection in modern optical detection systems. We innovatively propose and theoretically validate a novel AOA detection scheme based on coherent mode decoding with arrayed waveguides. The study confirms a precise, one-to-one correspondence between the AOA of the incident light beam and the main peak position of the decoded output light intensity, which establishes a solid physical foundation for high-precision angle retrieval using data-driven methods. To achieve high-precision AOA detection, this study designs and employs an end-to-end regression model that combines a convolutional neural network with an attention mechanism. Simulation results strongly demonstrate the effectiveness of this method, achieving MAE and RMSE as low as 0.0142° and 0.0193° , respectively. This level of precision is significantly superior to conventional methods based on peak detection and linear calibration, highlighting the powerful capability of deep learning models in accurately resolving complex optical patterns. Furthermore, the method exhibits excellent robustness under simulated phase noise and fabrication tolerance-induced perturbations, demonstrating its great potential for practical applications in the future.

The contribution of this research lies in providing a viable path for on-chip integrated AOA detection systems that combine high precision, low complexity, and high robustness, thereby advancing the interdisciplinary field of photonics and artificial intelligence. Looking ahead, our top priority is to finalize the device fabrication and experimental validation. Building upon this foundation, we will focus on optimizing both the decoding algorithm and the system architecture for multi-beam AoA detection. These efforts are aimed at unlocking its full application potential in fields such as LiDAR, LWS and optical communications.

Author Contributions: Conceptualization, Z.Y.; Methodology, J.Z. and Z.Y.; Software, J.Z.; Validation, J.Z.; Formal Analysis, J.Z.; Investigation, J.Z. and Z.Y.; Resources, H.Z.; Data Curation, J.Z.; Writing—Original Draft Preparation, J.Z.; Writing—Review & Editing, Z.Y. and H.Z.; Visualization, J.Z.; Supervision, Z.Y. and H.Z.; Funding Acquisition, H.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This study was financially supported by Sichuan Science and Technology Program (No. 2024YFHZ0002) and the Open Foundation of the State Key Laboratory of Precision Space-time Information Sensing Technology (No. STSL2025-B-04-01(L)).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The simulation data supporting the findings of this study are available within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Guiomar, F.P.; Fernandes, M.A.; Nascimento, J.L.; Rodrigues, V.; Monteiro, P.P. Coherent Free-Space Optical Communications: Opportunities and Challenges. *J. Light. Technol.* **2022**, *40*, 3173–3186. [\[CrossRef\]](#)
2. Tsai, M.C.; Bashir, M.S.; Alouini, M.S. Improved Angle-of-Arrival Estimation of Narrow Gaussian Beams for Mobile FSO Platforms. *IEEE Trans. Aerosp. Electron. Syst.* **2024**, *60*, 2778–2790. [\[CrossRef\]](#)
3. Zhao, S.; Zhu, L.; Shen, S.Y.; Du, H.; Wang, X.Y.; Chen, L.; Wang, X.D. Angle of Arrival for the Beam Detection Method of Spatially Distributed Sensor Array. *Sensors* **2025**, *25*, 1625. [\[CrossRef\]](#)
4. Premachandra, C.; Athukorala, A.; Tan, U.X. All-UWB SLAM Using UWB Radar and UWB AOA. *IEEE Robot. Autom. Lett.* **2025**, *10*, 8171–8178. [\[CrossRef\]](#)

5. Zhang, Y.; Li, J.; Zhao, S. Fusion Positioning of Ultra-wideband Single Base Station and Visual Inertial Odometry Based on Arrival of Angle and Time of Flight. In Proceedings of the 2023 2nd International Conference on Robotics, Artificial Intelligence and Intelligent Control (RAIIC), Mianyang, China, 11–13 August 2023; pp. 365–371. [\[CrossRef\]](#)
6. Zhang, K.H.; Zhang, Z.C.; Zhu, B.C. Beacon LED Coordinates Estimator for Easy Deployment of Visible Light Positioning Systems. *IEEE Trans. Wirel. Commun.* **2022**, *21*, 10208–10223. [\[CrossRef\]](#)
7. Mohammadnejad, S.; Aasi, M. Analysis of structures and technologies of various types of photodetectors used in laser warning systems: A review. *Opt. Eng.* **2023**, *62*, 090901. [\[CrossRef\]](#)
8. Shi, J.; Zhang, R.; Niu, J.Q.; Yang, X.M.; Li, M.W.; Wang, Z.B. Research on two-dimensional laser warning integrated detection technology. *Optik* **2022**, *270*, 170034. [\[CrossRef\]](#)
9. Safi, H.; Dargahi, A.; Cheng, J.L. Beam Tracking for UAV-Assisted FSO Links with a Four-Quadrant Detector. *IEEE Commun. Lett.* **2021**, *25*, 3908–3912. [\[CrossRef\]](#)
10. Jakubaszek, M.; Wojtanowski, J.; Zygmunt, M.; Mlodzianko, A. Sensor of light incidence angle. In Proceedings of the 13th Symposium on Laser Technology, Jastarnia, Poland, 24–28 September 2018; pp. 135–143. [\[CrossRef\]](#)
11. Guo, Y.J.; Guo, Y.H.; Li, C.S.; Zhang, H.; Zhou, X.Y.; Zhang, L. Integrated Optical Phased Arrays for Beam Forming and Steering. *Appl. Sci.* **2021**, *11*, 4017. [\[CrossRef\]](#)
12. Abiri, B.; Fatemi, R.; Hajimiri, A. A 1-D Heterodyne Lens-Free Optical Phased Array Camera with Reference Phase Shifting. *IEEE Photonics J.* **2018**, *10*, 1–12. [\[CrossRef\]](#)
13. Fang, X.Y.; Li, M.M.; Han, J.Z.; Ramaccia, D.; Toscano, A.; Bilotti, F.; Ding, D.Z. Accurate Direction-of-Arrival Estimation Method Based on Space-Time Modulated Metasurface. *IEEE Trans. Antennas Propag.* **2022**, *70*, 10951–10964. [\[CrossRef\]](#)
14. Wu, M.M.; Ma, J.X.; Zhang, Q. Photonic-assisted angle-of-arrival measurement system for both broadband and single-frequency radar signals. *IEEE Trans. Antennas Propag.* **2025**, *577*, 131392. [\[CrossRef\]](#)
15. Li, X.Y.; Wen, A.J.; Li, X.R.; Wang, Z.D. Photonic-Assisted Approach to Simultaneous Measurement of Frequency and Angle-of-Arrival. *J. Light. Technol.* **2023**, *41*, 7379–7389. [\[CrossRef\]](#)
16. Sinigaglio, H.; Murakowski, J.; Creazzo, T.; Prather, D. Angle-of-Arrival Detection Using Photonic Integrated Circuit and Array-Beam Space Mapping. In Proceedings of the 2024 IEEE International Symposium on Phased Array Systems and Technology (ARRAY), Boston, MA, USA, 15–18 October 2024; pp. 1–7. [\[CrossRef\]](#)
17. Sinigaglio, H.I.; Murakowski, J.; Creazzo, T.; Schneider, G.J.; Prather, D.W. Passive Angle of Arrival Detection Using Antenna Array and Photonic Integrated Circuit. *IEEE Trans. Microw. Theory Tech.* **2025**, *73*, 6971–6980. [\[CrossRef\]](#)
18. Creazzo, T.; Stine, C.; Creavin, C.; Harrity, C.; Shreve, K.; Wang, F.; Yao, P.; Murakowski, J.; Schneider, G.; Shi, S.; et al. Silicon Photonic Integrated Circuit Beamformer for RF Photonic Applications. In Proceedings of the 2024 IEEE/MTT-S International Microwave Symposium—IMS 2024, Washington, DC, USA, 16–21 June 2024; pp. 70–73. [\[CrossRef\]](#)
19. Pelucchi, E.; Fagas, G.; Aharonovich, I.; Englund, D.; Figueroa, E.; Gong, Q.H.; Hannes, H.; Liu, J.; Lu, C.Y.; Matsuda, N.; et al. The potential and global outlook of integrated photonics for quantum technologies. *Nat. Rev. Phys.* **2022**, *4*, 194–208. [\[CrossRef\]](#)
20. Siew, S.Y.; Li, B.; Gao, F.; Zheng, H.Y.; Zhang, W.; Guo, P.; Xie, S.W.; Song, A.; Dong, B.; Luo, L.W.; et al. Review of Silicon Photonics Technology and Platform Development. *J. Light. Technol.* **2021**, *39*, 4374–4389. [\[CrossRef\]](#)
21. Smit, M.K.; vanDam, C. PHASAR-based WDM-devices: Principles, design and applications. *IEEE J. Sel. Top. Quantum Electron.* **1996**, *2*, 236–250. [\[CrossRef\]](#)
22. Pan, X.F.; Yi, J.J.; Wang, S.; Ruan, Z.L.; Huang, F.; Yu, Z.J.; Dai, D.X.; Liu, L. Electrically and Arbitrarily Reconfigurable Arrayed Waveguide Grating on Thin Film Lithium Niobate. *Laser Photonics Rev.* **2025**, *19*, e00507. [\[CrossRef\]](#)
23. Yi, J.J.; Guo, C.J.; Ruan, Z.L.; Chen, G.X.; Wei, H.Q.; Lu, L.W.; Gong, S.Q.; Pan, X.F.; Shen, X.W.; Guan, X.W.; et al. Anisotropy-free arrayed waveguide gratings on X-cut thin film lithium niobate platform of in-plane anisotropy. *Light Sci. Appl.* **2024**, *13*, 147. [\[CrossRef\]](#)
24. van Wijk, A.C.; Doerr, C.R.; Akca, I.B. Custom Arrayed Waveguide Gratings with Improved Performance. *Adv. Photon. Res.* **2023**, *4*, 2300198. [\[CrossRef\]](#)
25. Lecun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **1998**, *86*, 2278–2324. [\[CrossRef\]](#)
26. Kiranyaz, S.; Avci, O.; Abdeljaber, O.; Ince, T.; Gabbouj, M.; Inman, D.J. 1D convolutional neural networks and applications: A survey. *Mech. Syst. Signal Process.* **2021**, *151*, 107398. [\[CrossRef\]](#)
27. Liu, W. Super resolution DOA estimation based on deep neural network. *Sci. Rep.* **2020**, *10*, 19859. [\[CrossRef\]](#)
28. Wu, L.; Liu, Z.M.; Huang, Z.T. Deep Convolution Network for Direction of Arrival Estimation With Sparse Prior. *IEEE Signal Process. Lett.* **2019**, *26*, 1688–1692. [\[CrossRef\]](#)
29. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention Is All You Need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 1–13.
30. Marz, R. *Integrated Optics. Design and Modeling*; Artech House Inc.: Norwood, MA, USA, 1994.

31. Zhou, P.; Chen, Z.; Wang, X.; Li, X.; Liu, Z.; Xu, X. Tolerance on tilt error for coherent combining of fiber lasers. *Chin. Opt. Lett.* **2009**, *7*, 39–42. [[CrossRef](#)]
32. Hocker, G.B.; Burns, W.K. Mode dispersion in diffused channel waveguides by the effective index method. *Appl. Opt.* **1977**, *16*, 113–118. [[CrossRef](#)] [[PubMed](#)]
33. Frey, B.J.; Leviton, D.B.; Madison, T.J. Temperature-dependent refractive index of silicon and germanium. In Proceedings of the Conference on Optomechanical Technologies for Astronomy, Orlando, FL, USA, 6 July 2006; pp. 790–799. [[CrossRef](#)]
34. Marathay, A.S.; McCalmont, J.F. On the usual approximation used in the Rayleigh-Sommerfeld diffraction theory. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **2004**, *21*, 510–516. [[CrossRef](#)]
35. Engel, T. 3D optical measurement techniques. *Meas. Sci. Technol.* **2023**, *34*, 032002. [[CrossRef](#)]
36. Ma, H.; Jen, A.K.Y.; Dalton, L.R. Polymer-based optical waveguides: Materials, processing, and devices. *Adv. Mater.* **2002**, *14*, 1339–1365. [[CrossRef](#)]
37. Tsuda, H. Silicon photonics platforms for optical communication systems, outlook on future developments. *IEICE Electron. Express* **2020**, *17*, 20202002. [[CrossRef](#)]
38. Roberts, S.; Ji, X.; Cardenas, J.; Corato-Zanarella, M.; Lipson, M. Measurements and Modeling of Atomic-Scale Sidewall Roughness and Losses in Integrated Photonic Devices. *Adv. Opt. Mater.* **2022**, *10*, 2102073. [[CrossRef](#)]
39. Zhang, Q.; Jeong, W.; Kang, D.J. Lock-in amplifiers as a platform for weak signal measurements: Development and applications. *Curr. Appl. Phys.* **2024**, *66*, 95–109. [[CrossRef](#)]
40. Blumenthal, D.J.; Heideman, R.; Geuzebroek, D.; Leinse, A.; Roeloffzen, C. Silicon Nitride in Silicon Photonics. *Proc. IEEE* **2018**, *106*, 2209–2231. [[CrossRef](#)]
41. Pfeiffer, M.H.P.; Herkommer, C.; Liu, J.; Morais, T.; Zervas, M.; Geiselmann, M.; Kippenberg, T.J. Photonic Damascene Process for Low-Loss, High-Confinement Silicon Nitride Waveguides. *IEEE J. Sel. Top. Quantum Electron.* **2018**, *24*, 6101411. [[CrossRef](#)]
42. Bogaerts, W.; Selvaraja, S.K.; Dumon, P.; Brouckaert, J.; De Vos, K.; Van Thourhout, D.; Baets, R. Silicon-on-Insulator Spectral Filters Fabricated with CMOS Technology. *IEEE J. Sel. Top. Quantum Electron.* **2010**, *16*, 33–44. [[CrossRef](#)]
43. Moss, D.J.; Morandotti, R.; Gaeta, A.L.; Lipson, M. New CMOS-compatible platforms based on silicon nitride and Hydex for nonlinear optics. *Nat. Photon.* **2013**, *7*, 597–607. [[CrossRef](#)]
44. Bauters, J.F.; Heck, M.J.R.; John, D.; Dai, D.; Tien, M.C.; Barton, J.S.; Leinse, A.; Heideman, R.G.; Blumenthal, D.J.; Bowers, J.E. Ultra-low-loss high-aspect-ratio Si₃N₄ waveguides. *Opt. Express* **2011**, *19*, 3163–3174. [[CrossRef](#)] [[PubMed](#)]
45. Sharma, T.; Wang, J.; Kaushik, B.K.; Cheng, Z.; Kumar, R.; Wei, Z.; Li, X. Review of Recent Progress on Silicon Nitride-Based Photonic Integrated Circuits. *IEEE Access* **2020**, *8*, 195436–195446. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.