Opto-Electronic Technology

CN 51-1811/O4 ISSN (Print) 2097-6003 ISSN (Online) 2097-8006

Multi-scale attention residual deep convolutional dealiasing network-assisted unambiguous ultra-long baseline high-precision microwave photonic angle of arrival estimation

Xianglin Chen, Yin Li, Shiru Song, Yalin Yao, He Cui, Xuan Li, Zhe Guo, Yinlong Tan, Taolin Liu* and Tian Jiang*

Citation: Chen XL, Li Y, Song SR, Yao YL, Cui H, Li X, Guo Z, Tan YL, Liu TL, Jiang T. Multi-scale attention residual deep convolutional dealiasing network-assisted unambiguous ultra-long baseline high-precision microwave photonic angle of arrival estimation. *Opto-Electron Technol* **2**, 260007 (2026).

https://doi.org/10.29026/oet.2026.260007

Received: 27 February 2026; Accepted: 24 March 2026; Published online: 30 March 2026

OET Opto-Electronic Technology

Article

March 2026, Vol. 2, No. 1

Open Access

Multi-scale attention residual deep convolutional dealiasing network-assisted unambiguous ultra-long baseline high-precision microwave photonic angle of arrival estimation

Xianglin Chen[1], Yin Li[2], Shiru Song[3], Yalin Yao[1], He Cui[1], Xuan Li[4], Zhe Guo[4], Yinlong Tan[3], Taolin Liu[4]* and Tian Jiang[4,5]*

Abstract: Conventional interferometric angle of arrival (AOA) estimation faces a fundamental limitation: high-precision angle measurement relies on long baselines, which easily introduce phase ambiguity. This issue is particularly pronounced in ultra-wideband (UWB) systems, where traditional ambiguity resolution methods lack robustness. To overcome this challenge, this paper introduces a microwave photonic (MWP) AOA estimation algorithm enhanced by a multi-scale attention residual deep convolutional dealiasing network (MSAR-DCDN). The proposed method employs the MSAR-DCDN to directly learn the nonlinear relationship between the intermediate frequency (IF) phase and the signal's angle of arrival, thereby bypassing conventional ambiguity resolution and relaxing the traditional trade-off between baseline length and operational bandwidth. Simulations demonstrate that the algorithm maintains strong robustness across a wide signal-to-noise ratio (SNR) range from −10 dB to 25 dB, and achieves an angle estimation accuracy exceeding 93% even at a high baseline-to-wavelength ratio of 2. Outdoor experiments with an 821 mm ultra-long baseline (the baseline-to-wavelength ratio reaches 21.9) further validate the approach, yielding a root mean square error (RMSE) below 0.42°. These results demonstrate a significant performance improvement over both standard interferometric techniques and their ambiguity-resolved variants. By integrating the UWB capability of MWP with the advanced MSAR-DCDN-based deep learning mechanism, this work presents a novel and effective framework for high-precision AOA estimation in intelligent photonics sensing, which mitigates the baseline length constraint of traditional methods and realizes flexible baseline configuration.

Keywords: intelligent photonics; deep learning; microwave photonics; angle of arrival estimation; phase ambiguity; ultra-wideband; baseline constraint; baseline configuration flexibility

DOI: 10.29026/oet.2026.260007 | CSTR: 32419.14.oet.2026.260007

Citation: Chen XL, Li Y, Song SR et al. Multi-scale attention residual deep convolutional dealiasing network-assisted unambiguous ultra-long baseline high-precision microwave photonic angle of arrival estimation. *Opto-Electron Technol* **2**, 260007 (2026).

1 Introduction

Interferometer angle of arrival (AOA) estimation is a critical passive measurement technique with fast response and high precision, widely deployed in ground-based, airborne and shipborne systems to acquire electromagnetic parameters and azimuth information of non-cooperative radiation sources[1,2], serving as core support for communication positioning, autonomous driving, and integrated sensing and communication (ISAC)[3,4]. With the continuous expansion of wireless communication bandwidth and increasingly complex electromagnetic environments, ultra-wideband (UWB) interferometer AOA estimation has emerged as an effective solution, but it imposes stringent demands on system bandwidth, channel consistency calibration, antenna

Received: 27 February 2026

Accepted: 24 March 2026

Published online: 30 March 2026

[1]College of Computer Science and Technology, National University of Defense Technology, Changsha 410073, China; [2]Defense Innovation Institute, Academy of Military Sciences, Beijing 100071, China; [3]College of Advanced Interdisciplinary Studies, National University of Defense Technology, Changsha 410073, China; [4]College of Science, National University of Defense Technology, Changsha 410073, China; [5]Hunan Research Center of the Basic Discipline for Physical States, Changsha 410073, China.

*Correspondence: TL Liu, E-mail: liutaolin14@nudt.edu.cn; T Jiang, E-mail: tjiang@nudt.edu.cn

baseline design and signal processing algorithms[5]. It faces two core challenges in practical implementation: first, the bandwidth limitations of electronic components lead to transmission loss, amplitude-phase inconsistency and electromagnetic interference[6]; second, contradictions exist between long-baseline high precision and severe phase ambiguity caused by phase differences exceeding 2π[7].

At present, AOA estimation technologies have formed a pattern of parallel development of multiple technical routes. According to the core working mechanism and hardware support system, they can be divided into three categories: traditional physics-driven methods, photon-assisted methods and intelligent learning methods. Each type of technology has formed unique technical advantages, and also faces insurmountable application bottlenecks. Traditional physics-driven methods are the basic technologies for phased array AOA estimation, including four typical approaches: First, the interferometric method[7,8], the basic method studied in this paper, which estimates the AOA by measuring the phase difference of signals received by phased array elements and combining the geometric relationship between wave path difference and incident angle. It features a simple hardware structure, fast response speed and good adaptability to the characteristics of phased arrays, and can achieve high precision in ideal scenarios with short baselines and high signal-to-noise ratio (SNR). However, it has an inherent trade-off between baseline length and phase ambiguity, long baselines improve the precision but introduce severe phase ambiguity, while short baselines avoid ambiguity but limit the resolution, and its robustness is extremely poor under UWB and low SNR conditions. Second, the subspace-based methods[9−11], represented by multiple signal classification (MUSIC) and estimation of signal parameters via rotational invariance techniques (ESPRIT), achieve super-resolution AOA estimation by decomposing the covariance matrix of received signals and can distinguish close-in target signals received by phased arrays. Nevertheless, they are based on the narrowband signal assumption and are highly sensitive to the calibration precision of phased array elements and element position errors. Third, the time/frequency difference of arrival (TDOA/FDOA) methods[12,13], which estimate the AOA by measuring the time/frequency difference of signals arriving at different elements of the phased array combined with geometric positioning, without strict baseline length limitations. However, they impose stringent requirements on the time/frequency synchronization precision of each phased array element. Time delay measurement errors can be directly propagated to the AOA estimation results, leading to a significant attenuation of performance in complex electromagnetic environments with multipath propagation. Fourth, the beamforming method[14], which forms a spatial power spectrum through phased array beam scanning and determines the AOA by using the position of the spectral peak, with strong real-time performance. But its AOA estimation precision is limited by the number of phased array elements with low resolution, and the beam broadening problem in broadband scenarios further reduces the performance.

Photon-assisted AOA estimation technology is a new direction developed to solve the bandwidth bottleneck of traditional electronic-based phased arrays. Microwave photonic (MWP) technology provides core support for the reception and processing of UWB signals in phased arrays by virtue of its advantages of large bandwidth, low loss and anti-electromagnetic interference[15]. For example, MWP frequency conversion links based on Mach-Zehnder modulators (MZMs) can achieve instantaneous bandwidth of tens of GHz, significantly outperforming traditional electronic systems[5]. Through extensive research by scholars, current AOA estimation implementation schemes based on MWP technology are mainly divided into two categories: the first category combines MWP down-conversion with backend digital processing algorithms. For instance, in ref.[16], an in-phase/quadrature (I/Q) detector is used to extract the phase of intermediate frequency (IF) signals. It achieves AOA estimation with an accuracy of ±2° in the frequency range of 2~18 GHz. In ref.[17], an error of less than ±3° is realized within the angle range of 0°~81.5° and frequency range of 12~18 GHz. Building on this, recent studies have further explored modulation signal-driven photonic AOA estimation optimization schemes[18,19]. By utilizing sawtooth or triangular waves to drive electro-optic modulators, periodic phase scanning of the optical sidebands is achieved. The AOA information can be extracted by detecting the periodic phase jumps embedded in the IF signal after photodetection. This approach enables AOA estimation with an accuracy of 2.27° for signals with an instantaneous bandwidth of up to 2 GHz. Such methods eliminate the need for complex system links and backend algorithms, demonstrating unique advantages in achieving wide instantaneous bandwidth AOA estimation. Benefiting from the flexibility of backend algorithms, this scheme offers strong scalability. Previous studies have shown that combining MWP down-conversion technology with traditional algorithms can significantly improve the measurement range and accuracy. Nevertheless, AOA estimation in this scheme still relies on electrical domain algorithms (such as interferometer or MUSIC algorithms). The inherent phase ambiguity problem of these traditional algorithms when dealing with long baselines has not been resolved by the introduction of photonic technology. The second category comprises mapping-based MWP AOA estimation schemes. They directly convert the characteristics of received signals into the AOA through nonlinear mapping. This avoids explicit phase calculation[20−22]. Reference[20] achieves a maximum measurement error of ±0.4° over an angle range of 60° to 85.2° at a single frequency of 3 GHz.

Similarly, refs.[21,22] adopt the same scheme and verify that the measurement accuracy within a single frequency point and limited angle range can be within ±2°. Although such methods have a relatively simple structure and do not require ambiguity resolution, their accuracy is susceptible to power detection sensitivity and channel consistency. They are also sensitive to environmental fluctuations, resulting in insufficient engineering robustness.

Intelligent learning-based AOA estimation techniques have emerged as a research hotspot in recent years[23–25]. By establishing a nonlinear mapping between received signal features and the AOA through a data-driven approach, these methods achieve end-to-end AOA estimation. They require no explicit phase ambiguity resolution or array calibration and exhibit strong anti-interference capabilities in low SNR and UWB scenarios, offering a new technical pathway to address the phase ambiguity problem in long-baseline phased arrays. However, existing intelligent AOA methods still have notable limitations: first, most studies focus on short baselines or scenarios with small baseline-to-wavelength ratios; second, the majority of these methods have not been integrated with MWP technology and remain constrained by the bandwidth limitations of electronic devices.

An overall comparison indicates that the scheme combining MWP down-conversion with AOA estimation algorithms demonstrates greater advantages in terms of system maturity and scalability, representing a viable solution for the development of UWB interferometric AOA estimation. However, the backend signal processing based on interferometric AOA estimation algorithms still faces the inherent trade-off between baseline length and phase ambiguity. To address this challenge, traditional methods such as the combined long-short baseline approach and Chinese remainder theorem (CRT)-based ambiguity resolution algorithms have been extensively studied[8]. In long-short baseline configurations, constructing virtual baselines can further shorten the physical short baseline length, thereby extending the AOA estimation operational range of the array[26]. Additionally, using staggered baselines combined with the CRT can reduce computational complexity and expand the operational bandwidth[27–29]. Nevertheless, these methods do not fundamentally overcome the theoretical constraint between baseline length and phase ambiguity: the stepwise ambiguity resolution using long and short baselines suffers from error accumulation, and as frequency increases, the minimum baseline length reaches a practical limit. The CRT method is sensitive to noise, with significantly reduced success rates under low signal-to-noise ratio or multi-signal conditions, and it requires pairwise coprime moduli, imposing strict limitations on baseline design[30]. Moreover, as signal bandwidth expands, the performance of these methods deteriorates sharply, failing to meet the needs of flexible baseline configuration in practical applications. Therefore, there is an urgent need for a stable, high-precision UWB solution that can transcend the limitations imposed by traditional algorithms on baseline length. With the advancement of intelligent photonics, the integration of MWP systems and deep learning algorithms has emerged as a pivotal solution to break through the bottlenecks of traditional angle of arrival estimation. This fits the key development trend toward intelligence and miniaturization of radio AOA equipment. Deep learning architectures such as convolutional neural networks excel in feature extraction and nonlinear mapping, paving the way for wideband AOA estimation. Unlike physics-driven traditional methods, data-driven models can learn implicit spatial features from array-received signals and acquire AOA estimation capabilities through training. Our prior work has verified the feasibility of applying neural networks to phase-power mapping-based long-baseline MWP AOA systems, achieving high-precision and adaptive measurements[31–33].

To address the inherent contradiction between baseline length and operating bandwidth in AOA estimation technology, this study proposes an multi-scale attention residual deep convolutional dealiasing network (MSAR-DCDN)-assisted MWP AOA estimation method. Its core innovations are as follows: firstly, breaking baseline length limitations: through the intelligent algorithm learning mechanism, the contradiction between baseline length and working bandwidth is fundamentally alleviated. It realizes UWB unambiguous AOA estimation under a fixed long baseline. Secondly, improving AOA estimation performance: achieving high-precision AOA estimation with large angle coverage in the UWB range. Its performance is significantly superior to traditional interferometer algorithms. The effectiveness of the proposed scheme is verified through simulations and field experiments (with a maximum baseline length of 821 mm). Results show that under the same baseline setting, the intelligent algorithm can always maintain high-precision angle prediction. In contrast, traditional algorithms completely fail due to phase ambiguity when the baseline-to-wavelength ratio (also called electrical length) is large. This research outcome opens a new path for improving baseline configuration flexibility in AOA estimation technology, providing a higher degree of freedom for baseline design. It provides new ideas for the development of future intelligent AOA estimation systems. Experimental verification shows that root mean square error (RMSE) less than 0.42° is achieved when the baseline length is 821 mm.

2 Analysis of MWP AOA estimation principles and baseline limitations

2.1 MWP AOA estimation system and AOA estimation principle

2.1.1 AOA estimation

The MWP down-conversion system, as the optical hardware foundation of intelligent optics, outputs high-quality intermediate frequency signals with retained phase information. This stable, wide-band optical signal provides an ideal data source for deep learning algorithms to learn angle features. The essence of angle of arrival (AOA) estimation is to invert the direction of incoming waves through the correlation between the path difference and phase difference of signals received by the antenna array. It requires simplifying the model based on the far-field assumption. For the convenience of subsequent analysis, a one-dimensional non-uniform antenna array model is adopted. Under the far-field assumption, it can be approximately considered that the signals arriving at each array element have the same power but differ only in phase. A schematic diagram of the antenna array receiving signals is shown in Fig. 1.

A one-dimensional antenna array with L elements is arranged horizontally. Assuming that the L elements are isotropic with spacing $d_l (d_0 = 0)$. Ignoring any interference, the incident angle of the k-th signal source $s_k(t)$ is $\theta_k \in [-\pi/2, \pi/2]$, $k = 1, 2, \cdots, K$. Then, the signal received by the l-th array element is

$$x_l(t) = \sum_{k=1}^{K} s_k(t) \mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda} \sum_0^{l-1} d_l \sin\theta_k} + n_l(t) \qquad (l = 1, 2, \cdots, L) \;, \tag{1}$$

where λ is the signal wavelength, and $n_l(t)$ denotes the additive white Gaussian noise received by the l-th array element at time t, following an $N(0, \sigma^2)$ distribution. The signal model can be written in matrix form as

$$\begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_L(t) \end{bmatrix} = \begin{bmatrix} \mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}\sum_0^0 d_l \sin\theta_1} & \mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}\sum_0^0 d_l \sin\theta_2} & \cdots & \mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}\sum_0^0 d_l \sin\theta_K} \\ \mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}\sum_0^1 d_l \sin\theta_1} & \mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}\sum_0^1 d_l \sin\theta_2} & \cdots & \mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}\sum_0^0 d_l \sin\theta_K} \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}\sum_0^{L-1} d_l \sin\theta_1} & \mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}\sum_0^{L-1} d_l \sin\theta_2} & \cdots & \mathrm{e}^{-\mathrm{j}\frac{2\pi}{\lambda}\sum_0^{L-1} d_l \sin\theta_K} \end{bmatrix} \cdot \begin{bmatrix} s_1(t) \\ s_2(t) \\ \vdots \\ s_L(t) \end{bmatrix} + \begin{bmatrix} n_1(t) \\ n_2(t) \\ \vdots \\ n_L(t) \end{bmatrix} . \tag{2}$$

In vector form:

$$\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{n}(t) \;, \tag{3}$$

where $\boldsymbol{x}(t)$ and $\boldsymbol{s}(t)$ are the array element received snapshot data vector and the signal source vector, respectively; $\boldsymbol{n}(t)$ is the additive white Gaussian noise vector; $\boldsymbol{A}$ represents the steering matrix of the signal source to the uniform antenna array. The element in the l-th row and k-th column of the steering matrix denotes the response coefficient of the l-th array element to the k-th signal:

$$a_{l,k} = e^{-j\frac{2\pi}{\lambda}\sum_0^{l-1} d_l \sin\theta_k} \;. \tag{4}$$

The steering vector of the L-element antenna array for the k-th signal is

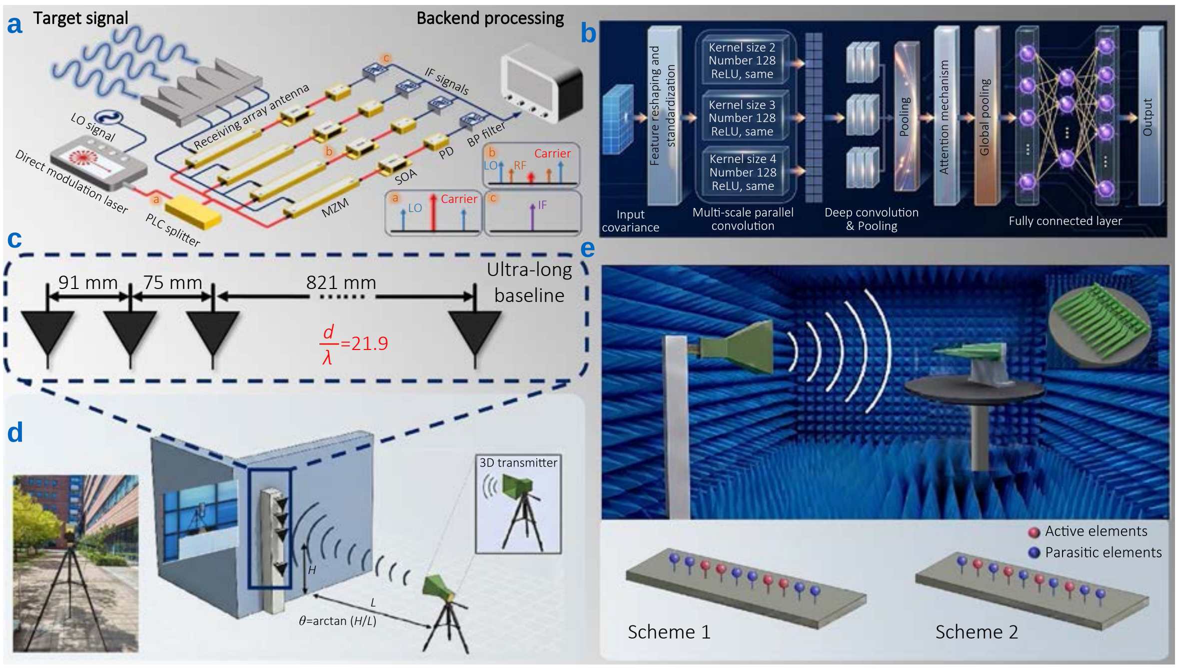

Fig. 1 | (a) Schematic diagram of the four-channel MWP down-conversion system. (b) Schematic diagram of the MSAR-DCDN network structure. (c) Schematic diagram of the ultra-long baseline setup. (d) Schematic diagram of the experimental setup for the ultra-long baseline. (e) Schematic diagram of the microwave anechoic chamber experimental scenario. LO: local oscillator; PLC: planar lightwave circuit; MZM: Mach-Zehnder modulator; SOA: semiconductor optical amplifier; PD: photodetector; BPF: bandpass filter.

$$\boldsymbol{a}_k=\boldsymbol{a}(\theta_k)=\begin{bmatrix} e^{-j\frac{2\pi}{\lambda}\sum_0^0 d_l\sin\theta_k} \\ e^{-j\frac{2\pi}{\lambda}\sum_0^1 d_l\sin\theta_k} \\ \vdots \\ e^{-j\frac{2\pi}{\lambda}\sum_0^{L-1} d_l\sin\theta_k} \end{bmatrix}. \tag{5}$$

It can be seen from the Eq. (4) that $a_{l,k}$ is only related to the array element position and the signal incident angle. This means that the AOA can be inverted as long as the phase difference between array elements (i.e., the phase information in the steering matrix) is accurately extracted. This is the core logic of all interferometer AOA estimation algorithms.

2.1.2 Down-conversion

The schematic of the four-channel MWP down-conversion system is depicted in Fig. 1(a). A directly modulated laser (DML), with an optical carrier frequency f_c and amplitude E_0, is directly modulated by a local oscillator (LO) signal at frequency $f_{\rm LO}$. The amplitude modulation coefficient and the frequency modulation coefficient of the DML at $f_{\rm LO}$ are denoted as $m_{\rm LO}$ and $M_{\rm LO}$, respectively. Consequently, the modulated optical field output from the DML is expressed as

$$E_{\rm DML}(t)=E_0\sqrt{1+m_{\rm LO}\cos(\omega_{\rm LO}t)}\,{\rm e}^{{\rm j}2\pi[f_ct+M_{\rm LO}\sin(2\pi f_{\rm LO}t)]}. \tag{6}$$

Subsequently, this optical signal is evenly distributed into four channels using a 1×4 planar lightwave circuit (PLC) splitter. Each output from the splitter is then fed into a *MZM* (*MZM$_i$*, *i*=1, 2, 3, 4), which is driven by a distinct radio frequency (RF) signal received from an antenna. The output optical fields of the four MZMs can therefore be derived as follows:

$$\begin{aligned}E_{MZM_i}(t)=&\frac{\sqrt{2t_{\rm M}}}{4}E_0\sqrt{1+m_{\rm LO}\cos(\omega_{\rm LO}t)}\\&\cdot{\rm e}^{{\rm j}2\pi[f_ct+M_{\rm LO}\sin(2\pi f_{\rm LO}t)]}\left[1+{\rm e}^{{\rm j}m_{\rm RF}\sin(2\pi f_{\rm RF}t+\varphi_i)}\right],\end{aligned} \tag{7}$$

where t_M denote the insertion loss of the MZM, and define the modulation index as $m_{\rm RF}=\pi V_{\rm RF}/V_\pi$, in which $V_{\rm RF}$ and $\omega_{\rm RF}$ represent the amplitude and frequency of the incident RF signal, respectively, and V_π is the half-wave voltage of the MZM. The phase of the RF signal applied to the MZM_i is denoted by φ_i $(i=1,2,\cdots,4)$, which carries critical information about the spatial characteristics of the incoming signal. Through this configuration, the system effectively captures and encodes the phase information of the target RF signal into the optical domain. As illustrated in Fig. 2, in the array antenna system, the spatial separation between antenna elements results in a time delay for the impinging signal across the array, and the temporal delay translates into a corresponding phase shift. For the i-th element, the time delay and phase shift are given by $\tau_i=d_i\sin\theta/c$, $\varphi_i=2\pi f_{\rm RF}\tau_{\rm i}=2\pi d_i\sin\theta/\lambda$. Here, d_i is the position of the i-th antenna element relative to the reference point (typically the first element), θ is the AOA of the signal with respect to the normal of the array plane, and c denotes the speed of light in vacuum. The RF-modulated optical signals from the four MZMs are then combined and transmitted through an optical fiber link to a centralized photodetector (PD), where they interfere and generate a photocurrent through square-law detection. The resulting photocurrent $I_i(t)$ can be expressed as

$$\begin{aligned}I_i(t)=&\frac{1}{4}\Re I_0 l_{\rm M}\{1+m_{\rm LO}\cos(2\pi f_{\rm LO}t)\\&+\sum_{n=-\infty}^{+\infty}{\rm J}_n(m_{\rm RF})\cos(2\pi nf_{\rm RF}t+\varphi_i)\\&+\sum_{n=-\infty}^{+\infty}{\rm J}_n(m_{\rm RF})\cos[2\pi(nf_{\rm RF}t\pm f_{\rm LO}t)+\varphi_i]\Big\},\end{aligned} \tag{8}$$

where $\Re$ denotes the responsivity of the PD, and ${\rm J}_n(x)$ represents the n-th order Bessel function. In a typical superheterodyne receiver architecture, the carrier frequency of the incoming signal is derived from the IF and LO signals via the relation $f_{\rm RF}=f_{\rm IF}+f_{\rm LO}$. Given that both $f_{\rm IF}$ and $f_{\rm LO}$ are known or measurable quantities, the RF frequency $f_{\rm RF}$ can be precisely determined. By accurately measuring this inter-element phase difference, the AOA can be unambiguously estimated. It is evident from the Eq. (8)that the IF signal frequency component has appeared $(f_{\rm RF}-f_{\rm LO})$ and carries the phase information of the target signal. Therefore, the phase difference between channels can be extracted by down-converting the RF signal to the IF output using a PD. In summary, the preceding derivation and analysis establish that the phase difference and the frequency of the system-generated low-frequency components suffice to determine both the target signal's AOA and frequency.

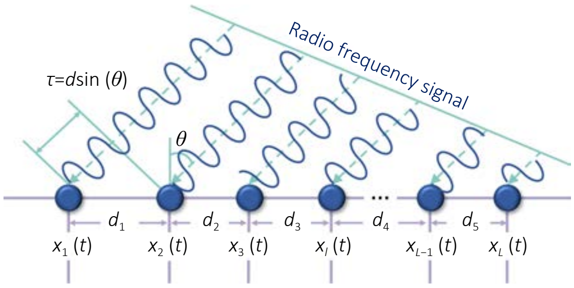

Fig. 2 | Schematic diagram of antenna array receiving microwave signal.

Based on the above analysis, MWP technology can achieve the reception and down-conversion of RF signals over an UWB range, and output intermediate frequency signals that retain the original phase information, providing a reliable data foundation for back-end algorithm processing. However, MWP links typically suffer from high transmission loss in practical applications, and back-end algorithms often impose strict requirements on signal quality. Therefore, it is particularly important to optimize the link loss. In an actual link, the input optical power to the photodetector (PD) is limited, and the optical signal output from a MZM operating at the linear point (Q point) contains a significant proportion of carrier components. This portion of carrier power does not contribute to the generation of the intermediate frequency signal, resulting in low conversion efficiency and limiting the overall system performance. Thus, improving the conversion efficiency requires optimizing the proportion of the optical carrier in the output signal. The method adopted in this work is to set the operating point of the second MZM in the link near the minimum transmission point (NULL point). By suppressing the carrier component, the proportion of useless optical power is effectively reduced, thereby enhancing the conversion efficiency of the useful signal and providing higher-quality intermediate frequency signals for back-end processing. To compensate for the associated optical power loss, a semiconductor optical amplifier (SOA) is inserted after the MZM array. The SOA boosts the optical power to above 5 dBm at the PD input. Moreover, because the LO sideband is also modulated by the RF signal in the second MZM, undesired mixing products can arise. These spurious components are suppressed by an electrical bandpass filter (BPF, centered at 1.8 GHz with a ±100 MHz passband) placed after the PD, which preserves only the desired IF component.

$$E_{\text{out}}(t)=\frac{E_{\text{in}}(t)}{2}\left[\sum_{n=-\infty}^{+\infty} j^n\mathrm{J}_n(M)\mathrm{e}^{\mathrm{j}n\omega_{\text{RF}}t+\frac{\mathrm{j}\phi_{\text{DC}}}{2}} + \sum_{n=-\infty}^{+\infty}(-1)^n j^n \mathrm{J}_n(M)\mathrm{e}^{-\mathrm{j}n\omega_{\text{RF}}t-\frac{\mathrm{j}\phi_{\text{DC}}}{2}}\right] , \tag{9}$$

here, $\phi_{\text{DC}}=\pi V_{\text{DC}}/V_\pi$ represents the phase change induced by the DC bias, and $M=\pi V_{\text{micro}}/V_\pi$ represents the phase change caused by the microwave signal. Typically, the amplitude and frequency of the modulated microwave signal can vary, whereas the DC bias V_{DC} is usually fixed at a specific value. The DC biases of the upper and lower arms are generally configured in a push–pull manner. The phase difference between the two arms can be set to 0, π/2, or π, corresponding to the modulator operating at the maximum transmission point (MAX point), the Q point, or the NULL point, respectively. By directly using the output expression E_{in} of the directly modulated laser and considering only the 0th and ±1st order terms of the Bessel function, we obtain:

$$\begin{aligned}E_{\text{out}}(t)=&\frac{E_0\sqrt{1+m_{\text{LO}}\cos(\omega_{\text{LO}}t)}\mathrm{e}^{\mathrm{j}\omega_c t}}{2}\\&\cdot[\mathrm{J}_0(M_{\text{LO}})\,(\mathrm{e}^{\mathrm{j}\omega_{\text{LO}}t}-\mathrm{e}^{-\mathrm{j}\omega_{\text{LO}}t})]\\&\cdot\left[\mathrm{J}_0(M_{\text{RF}})\cos\left(\frac{\phi_{\text{DC}}}{2}\right)\right.\\&\left.-2\sin\left(\frac{\phi_{\text{DC}}}{2}\right)\mathrm{J}_1(M_{\text{RF}})\,(\mathrm{e}^{\mathrm{j}\omega_{\text{RF}}t}+\mathrm{e}^{-\mathrm{j}\omega_{\text{RF}}t})\right] .\end{aligned} \tag{10}$$

By analyzing the above expression, we can find that the coefficient of the carrier component is $\mathrm{J}_0(M_{\text{LO}})\mathrm{J}_0(M_{\text{RF}})\cos\left(\phi_{\text{DC}}/2\right)$. The coefficient of the IF output component is $2\mathrm{J}_1(M_{\text{LO}})\mathrm{J}_1(M_{\text{RF}})\sin\left(\phi_{\text{DC}}/2\right)$. Therefore, when the DC bias voltage is equal to V_π, the carrier component coefficient is zero, achieving carrier-suppressed modulation. At this time, the IF component coefficient of the MZM output reaches the maximum value. After that, the RF signal is output through the PD. According to Eq. (8), we can derive that in the final output expression, the coefficient of the IF output component is

$$I_{\text{IF}}\propto\sin\left(\frac{\pi V_{\text{DC}}}{V_\pi}\right)\cos\left[\mathrm{J}_0(M_{\text{LO}})\mathrm{J}_0(M_{\text{RF}})\mathrm{J}_1(M_{\text{LO}})\mathrm{J}_1(M_{\text{RF}})\right]\cdot e^{j(\omega_{\text{LO}}-\omega_{\text{RF}})t} . \tag{11}$$

The Eq. (11) shows that the maximum value is achieved when $V_{\text{DC}}=V_\pi/2$. But at this time, the carrier power entering the PD also reaches the maximum value. Therefore, to concentrate as much power as possible on the IF output component, considering Eq. (9) and Eq. (10), in practical design, the DC bias voltage of the MZM needs to be adjusted to the middle value between the Q point and the NULL point. This improves the conversion efficiency of the IF signal through partial carrier suppression (reducing the proportion of carrier power) while retaining sufficient carriers to ensure the PD can normally output the IF signal for the subsequent MSAR-DCDN's phase feature learning and extraction.

2.2 Relationship between baseline length and AOA estimation performance

Modeling based on the interferometer AOA estimation principle can clearly reveal the inherent relationship and constraints among baseline length, AOA estimation accuracy, and phase ambiguity probability. According to the interferometer principle, the angle of arrival θ_{AOA}, measured phase ϕ_{M}, baseline length d_{baseline}, and signal wavelength λ satisfy:

$$\frac{2\pi d_{\text{baseline}}\sin\left(\theta_{\text{DOA}}\right)}{\lambda}=\phi_{\text{M}}+2k\pi , \tag{12}$$

where k is the ambiguity integer (when $d_{\text{baseline}}\le\lambda/2$, $k=0$, i.e., no ambiguity). Taking the differential of the Eq. (12), the expression for the AOA estimation error $\mathrm{d}\theta_{\text{AOA}}$ is obtained:

$$\mathrm{d}\theta_{\text{AOA}}=\frac{\lambda}{2\pi d_{\text{baseline}}\cos\theta_{\text{AOA}}}\cdot\left\{\mathrm{d}\phi_{\text{M}}+\left[\phi_{\text{M}}+2k\pi\left(\frac{\mathrm{d}\lambda}{\lambda}-\frac{\mathrm{d}d_{\text{baseline}}}{d_{\text{baseline}}}\right)\right]\right\} . \tag{13}$$

It can be seen from the Eq. (13) that the AOA estimation error $\mathrm{d}\theta_{\text{AOA}}$ is inversely proportional to the baseline length d_{baseline}. Ignoring wavelength and baseline errors, the larger d_{baseline} is, the smaller the angular error introduced by the phase measurement error $\mathrm{d}\phi_M$. This theoretically explains why the longest baseline is preferred for final high-precision angle calculation in engineering practice.

However, increasing the baseline length improves accuracy while significantly exacerbating the phase ambiguity problem. In the stepwise ambiguity resolution process based on multiple baselines, the phase difference measurement error of the short baseline is amplified by the ratio of adjacent baseline lengths $a_k = d_{k+1}/d_k$ and then enters the next stage of ambiguity resolution calculation, which may lead to ambiguity resolution errors[30]. Therefore, it is generally required that the phase difference measurement errors $\delta\phi_{\mathrm{M},k}$ and $\delta\phi_{\mathrm{M},k+1}$ of two adjacent baselines of the interferometer satisfy the following inequality:

$$\frac{d_{k+1}}{d_k}\left|\delta\phi_{\mathrm{M},k}\right| + \left|\delta\phi_{\mathrm{M},k+1}\right| < \pi \ . \tag{14}$$

In radar countermeasure reconnaissance, the standard deviation of the phase difference measurement error of each baseline is $\sigma_\phi = 1/\sqrt{SNR}$ (SNR is the signal-to-noise ratio of the received signal)[8]. According to the 3σ principle, the maximum phase difference measurement error $\left|\delta\phi_{\mathrm{M}}\right| = 3\sigma = 3/\sqrt{SNR}$. From the Eq. (14), the adjacent baseline ratio should satisfy:

$$a_k < \frac{\pi\sqrt{SNR}}{3} - 1 \ , \tag{15}$$

It can be seen that traditional interferometer AOA estimation systems face a fundamental trade-off in baseline design: increasing the baseline length can improve AOA estimation accuracy, but it will simultaneously reduce the robustness of ambiguity resolution. In addition, since the highest operating frequency of the system determines the maximum half-wavelength baseline required for unambiguous AOA estimation, and the lowest operating frequency affects the physical size of the antenna, this contradiction becomes particularly acute when pursuing an ultra-wide operating frequency band, ultimately manifesting as a trade-off between baseline length, AOA estimation accuracy, and system operating bandwidth. Alleviating this classic trade-off and enhancing the flexibility of baseline design is the key to realizing ultra-wideband high-precision interferometer AOA estimation.

To better evaluate the results of AOA estimation, we use estimation accuracy rate (EAR), mean absolute error (MAE), and RMSE to assess the performance of different methods:

Estimation accuracy rate:

$$Accuracy = \frac{N_{\text{within threshold}}}{N_{\text{total}}} \times 100\% \ \ , \tag{16}$$

where $N_{\text{within threshold}}$ is the number of estimates whose error falls within a predefined acceptable margin (in this paper, we set the threshold to ±1°), and N_{total} is the total number of test samples. This metric indicates the system's reliability and practical usability in real-world AOA estimation.

MAE:

$$MAE = \frac{1}{N}\sum_{i=1}^{N}\left|\hat{\theta}_i - \theta_i^{\text{true}}\right| \ \ , \tag{17}$$

where $\hat{\theta}_i$ is the estimated angle, θ_i^{true} is the true angle, and N is the number of samples. MAE measures the average magnitude of absolute errors, providing a stable view of the system's overall precision without being overly sensitive to outliers.

RMSE:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\hat{\theta}_i - \theta_i^{\text{true}}\right)^2} \ . \tag{18}$$

RMSE emphasizes larger errors due to the square term, making it especially useful for evaluating performance under challenging conditions (e.g., multipath, low SNR). It reflects both the dispersion of estimates and the impact of occasional large errors. Together, these metrics provide a comprehensive evaluation of estimation accuracy, consistency, and robustness, offering quantitative insights for overcoming the classic trade-offs in interferometer AOA estimation system design.

2.3 Limitations of traditional ambiguity resolution methods

As noted in Section 2.2, although long baselines improve AOA estimation accuracy, they also introduce phase ambiguity. Traditional de-ambiguity methods are commonly used to address this issue. However, these methods are generally designed for narrowband signals and high-SNR environments. Under UWB signals or low-SNR conditions, their performance often declines sharply, making them difficult to apply in practice. The following examines the limitations of such traditional approaches.

AOA estimation is based on the phase difference of a signal received by two antennas. This phase difference $\Delta\phi$ results from the path length difference Δd between the antennas, which is given by the relation:

Since the phase difference measured by the digital phase detector ranges from (−π, π), when the spacing of the antenna array exceeds half a wavelength, the value of the phase difference is outside the measurable range of the phase detector. At this time, there will be an integer multiple of 2π difference between the measured phase difference and the true phase difference, leading to errors in the estimation of the direction of incoming waves. At this point, ambiguity resolution is required to obtain the accurate direction of the spatial signal source.

In UWB receiving scenarios, the primary challenge faced by traditional de-ambiguity algorithms stems from the significant variation of the baseline-to-wavelength ratio with frequency. The baseline length is a key parameter that determines spatial resolution and AOA estimation accuracy. However, in UWB systems, in order to cover a broad frequency range, the baseline-to-wavelength ratio changes markedly with frequency. On the other hand, in low SNR environments, noise severely disturbs the phase and amplitude characteristics of the signal, posing a serious challenge to the noise-tolerance capability of traditional de-ambiguity algorithms.

To specifically evaluate the performance of traditional de-ambiguity algorithms under low-SNR conditions, this study conducted experiments at different SNR levels, testing the de-ambiguity success rate (results are shown in Fig. 3(a)). The experiment employed a four-channel antenna array for reception. Over an SNR range from −15 dB to +15 dB, the success rate of the interferometer-based algorithm (long-short baseline de-ambiguity) was recorded for each SNR level. The experimental results show a clear positive correlation between algorithm success rate and SNR. When the SNR is below 0 dB, the success rate is below 50% with considerable result fluctuations. As the SNR rises above 10 dB, the success rate gradually stabilizes around 90%. Beyond this point, further increases in SNR improve the success rate by less than 5%, indicating that the algorithm performance has approached saturation. Thus, traditional de-ambiguity algorithms exhibit high sensitivity to low-SNR conditions: their performance degrades severely when the SNR is below 0 dB, and It only demonstrates robust performance under moderate to high signal-to-noise ratios (>5 dB) and sub-half-wavelength conditions (as shown in the bottom-right corner of Fig. 3(a)), with notable limitations in noise resistance.

To further examine the overall performance of traditional algorithms under UWB and low-SNR conditions, this study carries out Monte Carlo simulations, focusing on the de-ambiguity success rate and direction-finding error across different baseline-to-wavelength ratios and angles. As shown in Fig. 3(b), the region of high success rate for the conventional algorithm changes notably with the baseline-to-wavelength ratio, indicating that its performance varies significantly with frequency in ultra-wideband systems, and may fail entirely in certain bands. In addition, the experiments show that when the angle of arrival exceeds a certain threshold (60° in this experiment), both the success rate and the direction-finding accuracy decline clearly. The physical reason is that as the angle increases, the phase difference changes more rapidly. Under large baseline-to-wavelength ratios, the identification of the ambiguity number becomes highly sensitive to noise, making it difficult for traditional algorithms to determine the correct ambiguity number, which ultimately leads to performance breakdown. These results further confirm that traditional de-ambiguity algorithms lack adaptability and robustness in UWB receiving scenarios.

3 Multi-scale attention residual deep convolutional dealiasing network (MSAR-DCDN) for wideband AOA estimation and ambiguity resolution

3.1 Problem modeling

Deep learning technology has performed excellently in various complex tasks due to its powerful feature extraction capabilities and adaptability. In AOA estimation tasks, deep learning algorithms extract features containing the incident angle of signals from received signals and establish a nonlinear mapping relationship between these features and signal angles to complete AOA estimation. Data-driven intelligent AOA estimation methods significantly reduce the strict dependence on baseline length in traditional AOA estimation methods, enabling higher flexibility in baseline configuration. Their AOA estimation accuracy, angle resolution, real-time performance, and robustness have been greatly enhanced. To address the shortcomings of traditional algorithms in ambiguity resolution performance under UWB and low SNR scenarios, this paper proposes a MSAR-DCDN, which redefines the phase ambiguity resolution process as a data feature learning problem. This model does not rely on the traditional cascaded solution process but

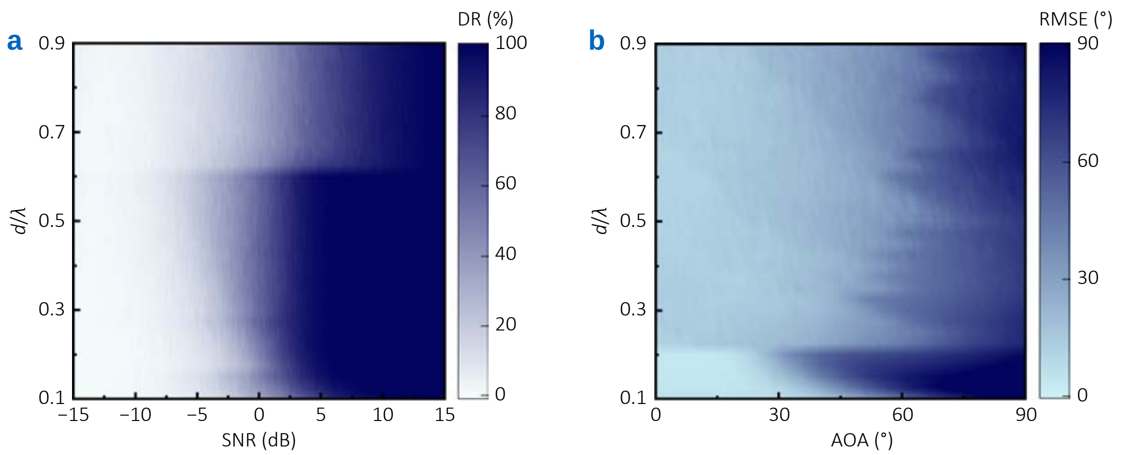

Fig. 3 | (a) de-ambiguity success rate under different baseline-to-wavelength ratios and signal-to-noise ratios. (b) RMSE under different baseline-to-wavelength ratios and angles of arrival. DR: de-ambiguity success rate; RMSE: root mean square error; SNR: signal-to-noise ratio; AOA: angle of arrival.

directly learns the nonlinear mapping between phase patterns and true angles of arrival from original received signals, effectively avoiding error transmission and accumulation. It breaks the inherent contradiction between baseline length and bandwidth in traditional AOA estimation methods, and while improving AOA estimation accuracy and resolution, it focuses on enhancing ambiguity resolution capabilities, effectively solving the problem of high-precision AOA estimation in MWP down-conversion AOA estimation links within the UWB range. AOA estimation is essentially a continuous spatial regression problem. To realize AOA estimation based on classification methods, it is necessary to uniformly discretize the angle domain to construct a discrete angle set:

$$\boldsymbol{\Theta} = \{-\theta_{\max}, \cdots, -\theta_{\text{res}}, 0, \theta_{\text{res}}, \cdots, \theta_{\max}\} , \tag{19}$$

where $\boldsymbol{\Theta}$ represents the complete set of possible arrival angles of the signal, $\pm\theta_{\max}$ respectively define the theoretical extreme boundaries of the azimuth angle, and θ_{res} is the angle resolution parameter. Thus, the number of discretized categories is

$$N = \frac{2 \cdot \theta_{\max}}{\theta_{\text{res}}} . \tag{20}$$

This study adopts a deep learning architecture to realize AOA estimation. For single-source scenarios, AOA estimation is modeled as a single-label, multi-class classification problem, whose mathematical representation is

$$\hat{\theta} = \underset{\vartheta\in\boldsymbol{\Theta}}{\operatorname{argmax}} \Pr(\theta = \vartheta|\boldsymbol{\Theta}) , \tag{21}$$

where $\Pr(\cdot)$ denotes the posterior probability distribution, and θ is the predicted AOA estimation value. For multi-signal-source scenarios, it is directly extended to a multi-label multi-target classification problem, whose expression is

$$\hat{\theta}_i = \underset{\vartheta\in\boldsymbol{\Theta}}{\operatorname{argmax}} \Pr(\theta_i = \vartheta|\boldsymbol{\Theta})(i = 1, 2, \cdots, K) , \tag{22}$$

where $\hat{\theta}_i$ represents the predicted angle of the k-th signal source.

Due to the existence of phase ambiguity, the current wideband intelligent AOA estimation algorithms have not solved the difficulty of constructing a wideband intelligent AOA estimation model, which is not because the AOA estimation frequency range has not been successfully expanded. The difficulty in constructing a wideband intelligent AOA estimation model arises from the complex relationship between the path difference and the incident angle across a wide frequency range. As can be seen from Fig. 4, when d/λ=0.35 (the ratio of d/λ is less than 0.5), the phase difference between adjacent array elements changes continuously with the increase of the incident angle, without phase ambiguity, and the azimuth features are relatively simple. At the high frequency end, phase ambiguity occurs between adjacent antenna elements (d/λ=0.7, 1.2), and this ambiguity characteristic will bring negative impacts on the input features, making the nonlinear function from input features to the incident direction extremely complex.

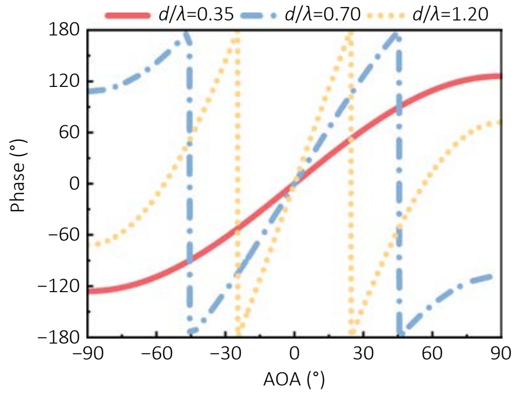

Fig. 4 | Schematic diagram of the azimuth spectrum under different baseline-to-wavelength ratios. AOA: angle of arrival.

The AOA estimation and ambiguity resolution methods based on a single convolution kernel have two core limitations: 1) a single-scale convolution kernel cannot adapt to the drastic changes of d/λ in the wideband, and is prone to lose local high-frequency details or global trends; 2) no optimization is made for "phase features being submerged by noise under low SNR", resulting in insufficient robustness of ambiguity resolution. To this end, this paper proposes a MSAR-DCDN, which breaks through the limitations of traditional methods through the integrated design of "multi-scale kernel adapting to wideband features, attention screening effective information, and residual connection stabilizing training". The MSAR-DCDN is specially designed for the wideband phase ambiguity problem, and each module of the network is targeted to solve the pain points of traditional methods, which is the fundamental reason why the network can realize ultra-long baseline and high-precision AOA estimation.

3.2 Core design logic of MSAR-DCDN

The design logic of MSAR-DCDN revolves around "unambiguous AOA estimation under the dynamic change of d/λ in wideband", and the core is to realize "adaptive virtual baseline construction, effective feature screening and deep virtual array expansion" inside the network through the innovative design of convolution modules, and at the same time be compatible with the "aperture expansion advantage of sparse array" in wideband AOA estimation. This design logic is highly consistent with the core demand of MWP AOA estimation for flexible baseline configuration, and is the key to the MSAR-DCDN breaking the baseline length constraint of traditional methods:

1) Multi-scale convolution for wideband feature adaptation: aiming at the dynamic change of d/λ, parallel convolution kernels of different sizes are adopted to perform differential transformation on the original baseline phase difference of the receiving antenna array, which equivalently generates an adaptive virtual baseline covering a wider frequency band, this module solves the problem of traditional single-scale convolution kernel being unable to adapt to wideband d/λ changes, and lays a feature foundation for the MSAR-DCDN to realize wideband unambiguous AOA estimation.

2) Attention mechanism for effective information screening: aiming at the fuzzy virtual baseline dominated by noise under low SNR, the channel attention subnetwork dynamically strengthens the feature channels that meet the unambiguous condition ($\phi < \pi$) and suppresses redundant noise, this module significantly improves the anti-noise performance of the MSAR-DCDN, making the network still able to accurately extract phase features under low SNR conditions and solve the problem of traditional algorithms being sensitive to noise.

3) Residual connection for training guarantee and feature reuse: aiming at the problem of effective virtual baseline loss in deep convolution, the residual unit retains the underlying unambiguous features and generates new short baseline combinations at the same time to expand the aperture of the virtual sparse array, this module ensures the stable training of the deep MSAR-DCDN network and the full reuse of effective phase features, and further expands the baseline adaptability of the network.

This logic deeply binds the "virtual expansion of physical baseline" with the "accurate extraction of intelligent features", and finally realizes end-to-end wideband ambiguity resolution and AOA estimation, solving the performance bottleneck of traditional methods in wideband and low SNR scenarios. The MSAR-DCDN's design logic is highly targeted to the core challenges of MWP ultra-long baseline AOA estimation, which is the reason why the network can realize the organic combination of MWP and deep learning, and break the traditional trade-off between baseline length and operating bandwidth.

The three groups of kernels (2×2, 3×3, 4×4) scan the input original phase difference matrix in parallel, and the output features are spliced and then the dimension is compressed by 1×1 convolution. A single convolution can generate differentiated virtual baselines, which greatly expands the baseline adaptability of the MSAR-DCDN and makes the network able to cover the wideband d/λ dynamic change range. The multi-scale virtual baseline is equivalent to a "dynamic sparse array", which not only retains the aperture advantage of the original array, but also avoids the phase ambiguity caused by the change of d/λ in the physical array, this coordination makes the MSAR-DCDN fully compatible with the MWP system's antenna array design, and realizes the organic combination of optical hardware and intelligent algorithm.

This design solves the problem that "a single convolution kernel cannot cover the wideband phase features", and provides a multi-granularity feature foundation for unambiguous AOA estimation of the MSAR-DCDN, which is the first key step for the network to realize ultra-long baseline high-precision AOA estimation.

The virtual baseline generation of the MSAR-DCDN can be described by the following formula:

$$Virtualbaseline = F_{\text{conv}}^{(L)} \circ \cdots \circ F_{\text{conv}}^{(2)} \circ F_{\text{conv}}^{(1)}(R) , \tag{23}$$

$$R = \frac{1}{T} \sum_{t=1}^{T} x(t) x^{H}(t) . \tag{24}$$

Among them, $F_{\text{conv}}^{(l)}$ is the convolution operation of the l-th layer of the MSAR-DCDN, and R is the covariance matrix of the received signal. The MSAR-DCDN realizes the deep expansion of the virtual baseline through the stacked convolution operation, and the generated large-aperture virtual sparse array is the key to the network realizing ultra-long baseline high-precision AOA estimation, which makes the MSAR-DCDN no longer limited by the physical baseline length and fundamentally breaks the traditional baseline constraint. Finally, the core structure and parameters of the MSAR-DCDN model are shown in Table 1.

Table 1 | Core structure and quantization parameters of the MSAR-DCDN network.

Layer type	Layer parameters	Output shape
Input layer	(4, 4, 2) (real and imaginary dual channels)	4×4×2
Conv2D (first Conv block)	32 2×2 convolution kernels, ReLU, same padding	4×4×32
BatchNormalization	for 32 channels	4×4×32
Conv2D (second Conv block)	64 3×3 convolution kernels, ReLU, same padding	4×4×64
BatchNormalization	for 64 channels	4×4×64
MaxPooling2D	2×2 pooling kernel	2×2×64
Conv2D (third Conv block)	64 4×4 convolution kernels, ReLU, same padding	4×4×64
BatchNormalization	for 64 channels	2×2×64
Conv2D	128 2×2 convolution kernels, ReLU, same padding	2×2×128
BatchNormalization	for 128 channels	2×2×128
GlobalAveragePooling2D	Global average pooling	128
Dense	128 neurons, ReLU	128
Dropout (0.4)	dropout	128
Dense	64 neurons, ReLU	64
Dropout (0.3)	Dropout	64
Output dense	181 neurons, softmax	181

4 Simulation and experimental results analysis

4.1 Simulation analysis

4.1.1 Simulation setup

The simulation adopts a four-channel received signal model, strictly simulating the actual workflow of the MWP down-conversion system, with a RF of 10 GHz and an IF of 750 MHz. In the process of constructing the training dataset, the angles of incoming waves are uniformly sampled within the full field of view of ±90°, 100 samples are generated for each angle, and additive white Gaussian noise of [−10, 30] dB is randomly added to simulate noise interference in complex electromagnetic environments; the test dataset adopts a fixed-step SNR setting to avoid the impact of data distribution deviation on evaluation results.

Unified training and implementation details of deep learning methods: the deep learning methods share the same training/test split (8:2). The incident angles are uniformly sampled across the full field of view of ±90°, with 100 samples generated per angle. Additive white Gaussian noise ranging from −10 dB to 30 dB is randomly added to the data. For hyperparameter settings, both models uniformly adopt the Adam optimizer with an initial learning rate of 0.001, an exponential decay strategy (decay factor of 0.96 every 1000 steps), and early stopping with a patience of 30 epochs. The loss function is focal loss ($\gamma = 2.0$, $\alpha = 0.25$), with a batch size of 32, 100 training epochs, and a dropout rate of 0.3~0.4. All models are trained and inferred on the same workstation equipped with an NVIDIA RTX 4090 GPU to eliminate hardware-related discrepancies.

Unified Implementation Details of Classical Signal Processing Algorithms: the test dataset for all classical algorithms is consistent with that used for the deep learning methods. For system calibration, IM (interferometer method), IM-ARM (interferometer algorithm with the ambiguity resolution method), and MUSIC all use calibration data from the normal direction (0°) to eliminate inherent channel phase differences and array position errors. In terms of parameter configuration, the MUSIC algorithm employs the same covariance matrix input as MSAR-DCDN. IM-ARM consistently uses a unified long-short baseline configuration and the Chinese remainder theorem (CRT)-based ambiguity resolution strategy across all experiments. All classical algorithms and deep learning methods are implemented on the same hardware platform, and inference latency is statistically measured under the same snapshot number of 512.

To comprehensively evaluate the AOA estimation performance of the algorithm in terms of accuracy, stability and engineering usability, the performance evaluation indicators adopt the average RMSE, average MAE, and average EAR of all angle samples. The accuracy sets an allowable threshold of 1°, and a prediction is determined as effective if the difference between the predicted angle and the true angle is less than the threshold, which is consistent with the practical application requirements of high-precision AOA estimation.

4.1.2 Model performance verification

The MSAR-DCDN model is specifically designed for 4×4 array complex covariance matrix input (4×4×2 dimensions, real and imaginary parts as dual channels). The core architectural parameters and computational metrics are summarized in Table 2.

Table 2 | Key parameters of the MSAR-DCDN network.

Metric	Value
Input dimension	4×4×2 (real + imaginary parts)
Total trainable params	144,000 (144k)
Total FLOPs per inference	About 231 kFLOPs
Model size	0.55 MB
Dropout rate	0.3–0.4 (to avoid overfitting)
Optimization strategy	Adam (lr=0.0005), ReduceLROnPlateau (patience=15)

To verify the superiority of the proposed MSAR-DCDN model structure, two sets of comparative experiments were conducted in this section.

Verification of structural advantages: Under the premise of keeping the training dataset, test dataset and hardware platform unchanged, comparative tests were carried out on the MSAR-DCDN model and a simple CNN model (a basic convolutional network with multi-scale convolution, attention mechanism and residual connections removed) under three typical baseline-to-wavelength ratios (d/λ=0.35, 0.7, 1.2), corresponding to the unambiguous short baseline, slightly ambiguous medium baseline and severely ambiguous ultra-long baseline scenarios, respectively. The experimental results are presented in Table 3 and Fig. 5.

Table 3 | Estimation accuracy rate comparison of MSAR-DCDN and simple CNN.

d/λ	Estimation accuracy rate (%)	
	MSAR-DCDN	Simple CNN
0.35	100	93
0.7	97	81
1.2	94	66

It can be clearly seen from the results that the proposed MSAR-DCDN achieves significantly superior performance over the simple CNN model under all baseline configurations, and the performance gap expands further with the increase of d/λ. Especially in the ultra-long baseline scenario with severe phase ambiguity at d/λ=1.2, the estimation accuracy rate of MSAR-DCDN is 28 percentage points higher than that of the simple CNN model (94% vs 66%).

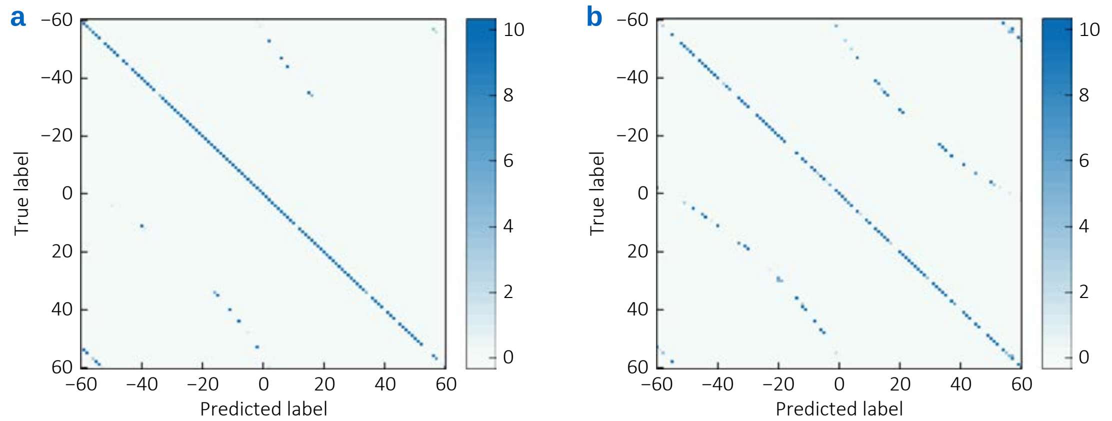

Fig. 5 | Comparison of model prediction confusion matrix results at electrical length 1.2. (a) Result of MSAR-DCD. (b) Result of simple CNN.

This fully verifies that the integrated design of MSAR-DCDN can effectively address the phase dealiasing problem in AOA estimation with long baselines, which is the core key for the model to achieve performance improvement.

Inference latency comparison: To quantitatively verify the real-time inference performance of the proposed MSAR-DCDN algorithm, a comparative experiment on the algorithm inference time under different snapshot numbers was conducted. Taking four typical snapshot numbers (64, 128, 256, 512) as variables, the AOA estimation was completed by the classic MUSIC algorithm, the traditional interferometer algorithm and the MSAR-DCDN algorithm proposed in this paper, respectively, and the total inference time of each algorithm was counted. All experiments were implemented on the same workstation to ensure the fairness of the comparison, and the statistical results of the inference time are shown in Table 4.

Table 4 | Inference time comparison of different algorithms under different snapshot numbers.

Snapshot	Time (ms)		
	MUSIC	Interferometer	MSAR-DCDN
64	7.13	0.05	0.184
128	7.13	0.06	0.176
256	9.61	0.16	0.182
512	11.01	0.27	0.171

It can be clearly seen from the experimental data that the inference time of the traditional MUSIC algorithm and interferometer algorithm increases significantly with the increase of the snapshot number, while the inference time of the MSAR-DCDN algorithm remains basically stable and is not affected by the change of data complexity (snapshot number). Specifically, in the scenario of the maximum data volume with 512 snapshots, the total inference latency of the MSAR-DCDN algorithm is only 0.171 ms, which consists of two parts: 0.014 ms for covariance matrix (CM) computation and 0.157 ms for network inference. The core reason for this performance advantage is that the input of the MSAR-DCDN model is the signal CM, whose dimension is only related to the number of array elements of the antenna array and has no correlation with the snapshot number. Therefore, the increase of the snapshot number will not improve the computational complexity of the model, and the inference time remains essentially unchanged. In contrast, the MUSIC algorithm needs to complete complex operations such as system calibration and covariance matrix decomposition, resulting in the longest inference time (11.01 ms at 512 snapshots) and the worst real-time performance. Although the traditional interferometer algorithm does not require covariance matrix computation, it still needs to execute the system calibration step, and its inference time is 0.27 ms at 512 snapshots, which is slightly higher than that of the proposed algorithm.

4.1.3 Performance comparison under different SNR conditions

The baseline length is fixed at half-wavelength, and the performance of each algorithm is investigated over a SNR range from −10 dB to 25 dB with a 5 dB step. At each SNR level, the EAR and MAE of the interferometer algorithm, the interferometer algorithm with ambiguity resolution method, MUSIC algorithm and the proposed MSAR-DCDN intelligent algorithm are examined. The results shown in Fig. 6 indicate that the intelligent algorithm exhibits significantly superior anti-noise capability compared to traditional algorithms across the entire SNR range. Under low SNR conditions (−10 dB to 0 dB), the traditional interferometer algorithm is significantly affected by phase measurement errors, with a MAE exceeding 15°. Although an attempt is made to correct it using an ambiguity resolution method, the MAE of interferometer algorithm further increases. This is because under extremely low SNR environments, the phase information has been severely contaminated, leading to complete failure of the ambiguity resolution algorithm and introducing additional errors. In contrast, the intelligent algorithm can still maintain a high precision with an MAE of less than 2° when the SNR is −5 dB. As the SNR increases, the performance of each algorithm improves: when the SNR exceeds 0 dB, the EAR of the intelligent algorithm tends to stabilize and is less affected by noise. The performance of the MUSIC algorithm improves rapidly with increasing SNR, and its MAE can be stably controlled below 2° when the SNR is higher than 5 dB. However, the estimation accuracy of the traditional interferometer algorithm is still highly dependent on the SNR level, and its MAE can only reach within 2° when the SNR rises above 20 dB. This experiment demonstrates that through a data-driven approach, the intelligent algorithm learns the essential features of signal phases from a large number of samples, reducing the dependence on accurate instantaneous phase difference measurement, thereby exhibiting stronger robustness and stability in complex noise environments.

4.1.4 Performance analysis under different electrical length conditions

Electrical length (baseline-to-wavelength ratio) is the core indicator reflecting the baseline configuration flexibility and phase ambiguity degree, and to verify the ability of the MSAR-DCDN algorithm to break through the traditional baseline length constraint and realize wideband unambiguous AOA estimation, the SNR is fixed at 20 dB, and the performance of each algorithm is investigated as the electrical length varies from 0.2 to 2. The average EAR, MAE, and RMSE within the ±90° range are examined for the three methods under different electrical lengths. From Fig. 7(a–c), it can be clearly observed that the performance of the conventional interferometer and MUSIC algorithm deteriorates significantly as the electrical length increases. This is because a longer electrical length results in a smaller unambiguous angle range. Although the ambiguity resolution method alleviates the phase ambiguity problem to some extent, its accuracy remains around 60% as the electrical length increases. In contrast, the proposed MSAR-DCDN network does not exhibit this issue, maintaining an accuracy above 93% and keeping both MAE and RMSE within 2° even at a high electrical length of 2, which fully reflects the algorithm's insensitivity to electrical length changes.

Subsequently, we provide a detailed comparison of the AOA estimation performance of the three methods under different electrical lengths and angles, as shown in Fig. 8(a–h). The light-colored region in the figures represent the angular ranges where correct AOA estimation is achieved. The correctly estimated range of the interferometer and MUSIC algorithm gradually narrows as the electrical length increases. Even when the ambiguity resolution algorithm is incorporated, certain ambiguous regions remain unresolved. In comparison, the intelligent algorithm maintains correct

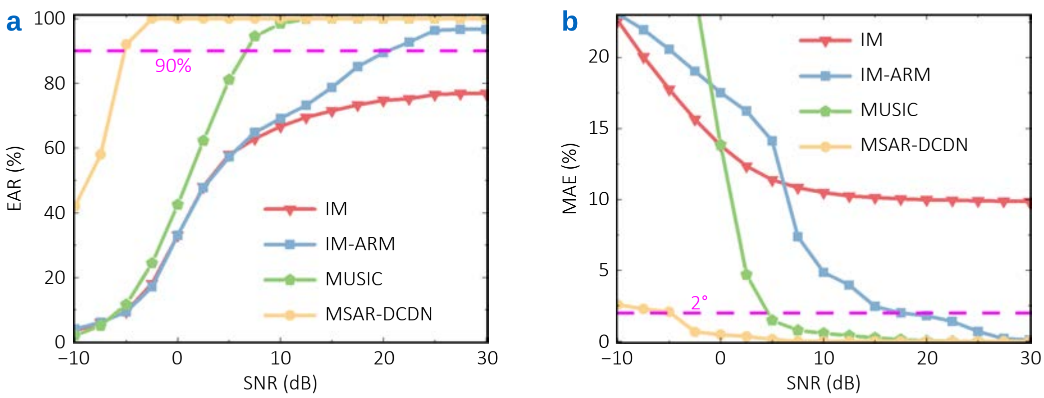

Fig. 6 | Comparison of algorithm performance under different SNR levels. (a) Comparison of estimation accuracy results. (b) Comparison of MAE results.

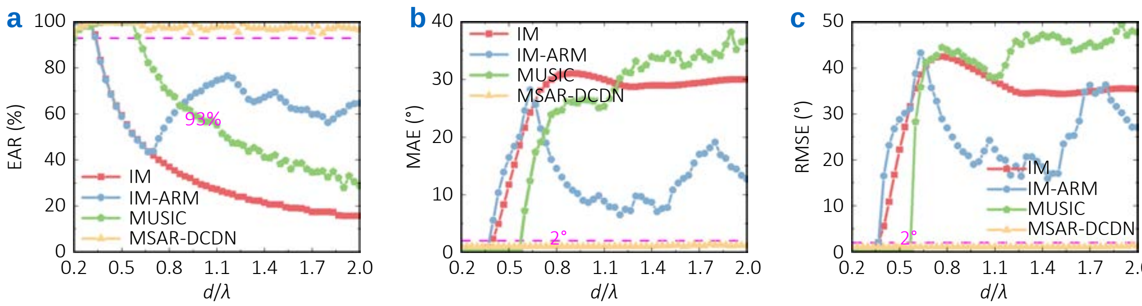

Fig. 7 | Comparison of algorithm performance under different electrical lengths. (a) EAR comparison of three methods. (b) MAE comparison of three methods. (c) RMSE comparison of three methods.

AOA estimation across the full angular range despite increasing electrical length. These results further demonstrate that, through its data-driven learning mechanism, the intelligent algorithm effectively mitigates the performance degradation caused by phase ambiguity in conventional methods, exhibiting high reliability and adaptability even under ultra-long baseline conditions.

4.2 Experimental verification

4.2.1 Experimental setup

This chapter designs and implements microwave anechoic chamber experiments and outdoor complex environment experiments to verify the effectiveness and robustness of the proposed MSAR-DCDN intelligent AOA estimation algorithm in actual systems from multiple dimensions. The MWP down-conversion system used in the experiment is consistent with that shown in Fig. 1(a). The MWP down-conversion module utilizes a DML (EM440, 1550 nm) and a MZM (Fujitsu FTM7920FBA) to generate an IF signal in the optical domain. A SOA (JSA-BT515G25-PM) is inserted before the photodetector, boosting the optical power to approximately 5 dBm. The IF signal is then extracted by a 20 GHz PD (Discovery DSC-R401HG) and filtered by a 1.8 GHz (±100 MHz) BPF to remove spurious mixing products. A data acquisition module digitizes the output at 300 Mega Samples per second (MSaps), incorporating built-in digital down-conversion. A dual-channel broadband signal generator (SinoLink SLVS06D) supplies the excitation, which is then amplified (MWPA-020180G10) and transmitted via a dual-polarized horn antenna (2–18 GHz). The experiments are performed in a standard anechoic chamber with hybrid ferrite/polyurethane absorbers to suppress interference. Reception is handled by a 10-element Vivaldi antenna array (2–18 GHz, element spacing d= 17 mm).

4.2.2 Microwave anechoic chamber experimental results

The microwave anechoic chamber experiment provides a controlled ideal electromagnetic environment for accurately evaluating the baseline performance of the algorithm under known conditions. Its experimental configuration is shown in Fig. 1(e). The receiving antenna array is installed on a turntable, which can rotate precisely within the range of ±90° to simulate different directions of incoming waves. The transmitting end is driven by a signal source to transmit radio frequency signals with a fixed frequency of 8 GHz. The experiment sets the following two baseline schemes for comparative verification:

Scheme 1 (long and short baselines): the spacing between adjacent array elements is 17 mm, 51 mm, and 17 mm respectively, corresponding to electrical lengths of approximately 0.45 and 0.91, aiming to use the short baseline for ambiguity resolution reference.

Scheme 2 (all baselines exceeding half-wavelength): the spacing between all adjacent array elements is 34 mm (electrical length > 0.5), and there is theoretically no ambiguity-free phase interval under this condition, which constitutes a challenging test scenario that traditional algorithms cannot handle.

For the traditional interferometer algorithm, first collect calibration data in the normal direction (0°) to eliminate the

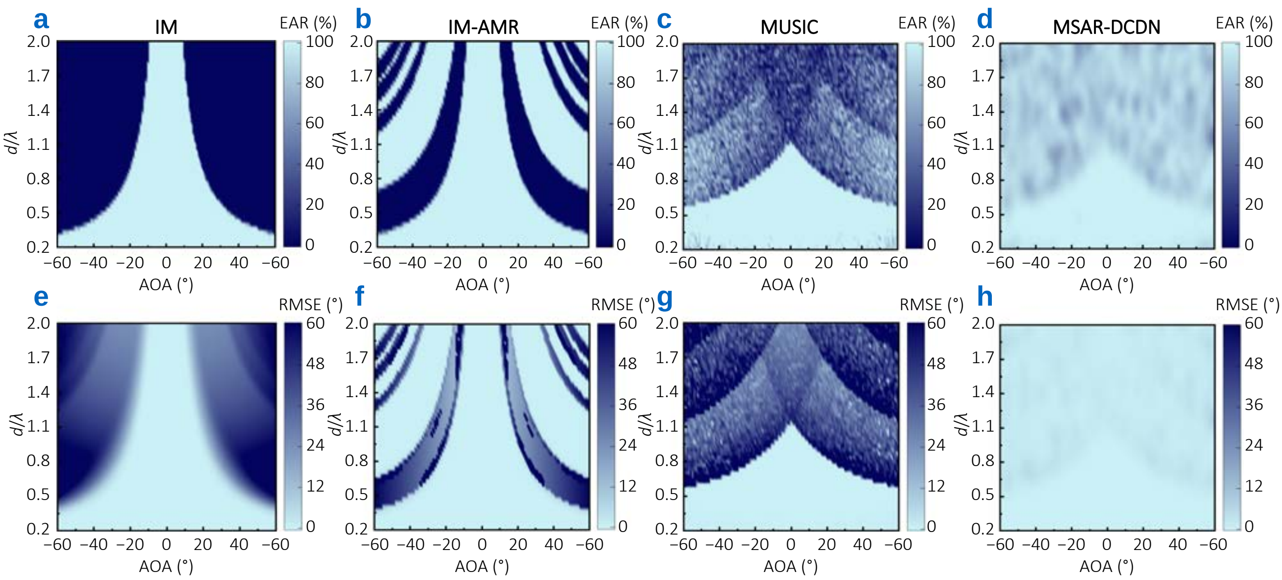

Fig. 8 | Performance comparison of algorithms for different electrical lengths and angle of arrivals. (a) EAR of IM. (b) EAR of IM-ARM. (c) EAR of MUSIC. (d) EAR of MSAR-DCDN. (e) RMSE of IM. (f) RMSE of IM-ARM. (g) RMSE of MUSIC. (h) RMSE of MSAR-DCDN. IM: interferometer algorithm; IM-ARM: interferometer algorithm with the ambiguity resolution method; MUSIC: multiple signal classification; AOA: angle of arrival.

inherent phase difference between system channels and array position errors. Subsequently, the basic interferometer algorithm (without ambiguity resolution method) and the interferometer algorithm combined with the ambiguity resolution method are used for AOA estimation respectively. In contrast, the MSAR-DCDN intelligent algorithm adopts an end-to-end data-driven mode: 80% of the collected original data is used as the training set, which is directly input into the network for training. The trained model can directly estimate the azimuth angle of the remaining 20% of the test data without any pre-calibration or post-processing steps. The results are shown in Fig. 9.

For **Scheme 1** (long-short baseline combination), the basic interferometer algorithm can only achieve ambiguity-free estimation within ±20° due to phase ambiguity, and the ambiguity resolution method still fails at extreme angles (±50° and beyond) because the short baseline phase reference is unreliable. In contrast, the MSAR-DCDN algorithm achieves high-precision estimation across the full ±90° range with RMSE below 0.3°, and does not require the cooperation of short baselines for ambiguity resolution, verifying the core advantage of the MSAR-DCDN in decoupling AOA estimation accuracy from baseline length. For **Scheme 2** (all baselines exceed half-wavelength), traditional algorithms completely fail due to the lack of ambiguity-free phase interval, while the MSAR-DCDN algorithm still maintains stable and accurate estimation. This fully verifies that the MSAR-DCDN can break through the physical baseline constraints of traditional methods.

In summary, the proposed MSAR-DCDN AOA estimation algorithm maintains high accuracy and stability under different baseline settings. It overcomes the strict dependence of traditional methods on baseline length, demonstrating significant technical advantages and application potential.

4.2.3 Results of the ultra-long baseline experiment

To further evaluate the performance of the MSAR-DCDN in a real and complex electromagnetic environment, we designed an outdoor ultra-long baseline experiment. The geometric relationship of the experimental scene is shown in Fig. 2(d). The receiving end is deployed at a fixed position with a height H = 26.5 m, and the transmitting end is placed on the ground. The true angle of arrival θ ($\theta = \arctan(H/L)$) can be determined by accurately measuring the horizontal distance L. The transmitting end adopts a dual-polarized ridged horn antenna to transmit 8 GHz signals, which are amplified to compensate for path loss. The receiving array adopts an ultra-long baseline design, with the spacing between adjacent units being 821 mm, 75 mm, and 91 mm respectively (the 821 mm baseline corresponds to a baseline-to-wavelength ratio of 21.9, which is the longest baseline setting in the existing MWP AOA estimation related research), aiming to maximize AOA estimation accuracy. The system generates a LO frequency of 9.8 GHz, and after mixing with a RF frequency of 8 GHz, a 1.8 GHz IF signal is generated for acquisition. The experiment selects four typical angles of θ = 10°, 13°, 17°, and 26° for fixed-point testing to cover different spatial scenarios.

In this experiment, AOA estimation was performed using three distinct methods on the same dataset: a traditional interferometer algorithm, an enhanced interferometer algorithm incorporating the CRT for phase ambiguity resolution method[28,29,34], and MSAR-DCDN intelligent algorithm

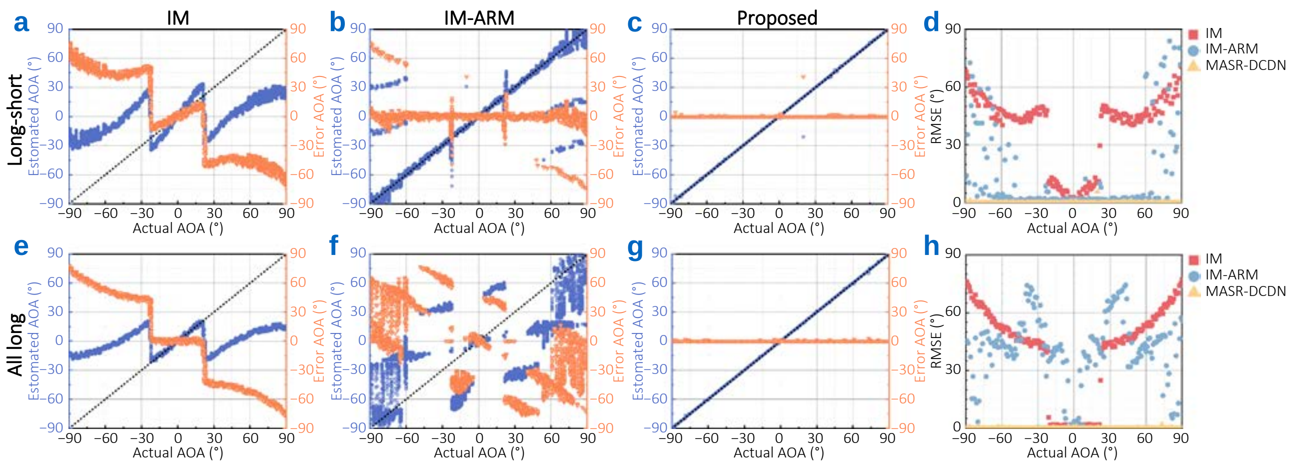

Fig. 9 | Microwave anechoic chamber experimental results. (a) IM prediction results under the long-short baselines configuration. (b) IM-ARM prediction results under the long-short baselines configuration. (c) MSAR-DCDN prediction results under the long-short baselines configuration. (d) RMSE comparison of the three methods under the long-short baselines configuration. (e) IM prediction results under the all-long baseline configuration. (f) IM-ARM prediction results under the all-long baseline configuration. (g) MSAR-DCDN prediction results under the all-long baseline configuration. (h) RMSE comparison of the three methods under the all-long baseline configuration. IM: interferometer algorithm; IM-ARM: interferometer algorithm with the ambiguity resolution method; AOA: angle of arrival.

proposed in this paper. As shown in Table 5, without phase ambiguity resolution, the traditional interferometer algorithm maintains high accuracy only at 10°, while it completely fails at other angles due to severe phase ambiguity. After incorporating ambiguity resolution, its performance deteriorates sharply as the angle increases, with accuracy dropping to 54% and RMSE reaching 16.65° at 26°. In stark contrast, the MSAR-DCDN achieves 100% EAR across all tested angles, with both MAE and RMSE consistently below 0.42° under the ultra-long baseline with a baseline-to-wavelength ratio of 21.9. Meanwhile, it is clearly demonstrated in Fig. 10(a,b) that the proposed algorithm exhibits superior stability and accuracy compared to conventional algorithms, realizing high-precision AOA estimation under the longest baseline condition to date, and verifying the core performance advantage of the algorithm proposed in this paper. This result fully demonstrates the excellent stability and environmental adaptability of the MSAR-DCDN algorithm in complex outdoor electromagnetic environments, and its engineering application potential in high-precision AOA estimation with flexible baseline configuration. Two experiments (microwave anechoic chamber and outdoor ultra-long baseline) analyzed the performance of the MSAR-DCDN under ideal and real-world conditions. The results show that the proposed intelligent AOA estimation algorithm not only overcomes the inherent phase ambiguity of traditional methods and eliminates dependence on specific baseline designs, but also demonstrates strong robustness and high accuracy in complex outdoor electromagnetic environments. Additionally, the data-driven paradigm of the MSAR-DCDN enables it to bypass explicit calibration: the training dataset contains both target phase information related to AOA and inevitable inter-channel phase deviations (e.g., from inconsistencies in microwave photonic links or antenna array imperfections), which enables the network to automatically learn to distinguish and compensate for these errors while extracting angle-dependent features. This makes it capable of achieving AOA estimation without the need for any calibration steps. This work provides a solid technical foundation for engineering next-generation, high-precision, and highly adaptive AOA estimation systems with flexible baseline configuration.

We compare the MSAR-DCDN with the existing MWP AOA estimation related research (as shown in Fig. 10(c)). The comparison results show that the MSAR-DCDN algorithm proposed in this paper achieves AOA estimation with an error within 0.42° under the condition of the longest baseline-to-wavelength ratio (21.9) in the existing research, breaking the performance bottleneck that the traditional MWP AOA estimation algorithm can only achieve high precision under the short baseline or long-short baseline matching condition. Previous works have realized high-precision AOA estimation in MWP systems, but they all limit the baseline length to a small range (baseline-to-wavelength ratio <5) to avoid phase ambiguity, and their estimation accuracy will drop sharply once the baseline is extended[35–39]. In contrast, the MSAR-DCDN algorithm in this paper decouples the AOA estimation accuracy from the baseline length through the multi-scale attention residual deep learning structure, realizing the dual goals of ultra-long baseline and high precision that cannot be achieved by previous works, and providing a new technical path for the development of next-generation MWP AOA estimation systems with flexible baseline configuration and high precision.

5 Conclusions and prospects

This paper proposes a MSAR-DCDN-assisted microwave

Table 5 | Experimental results of the ultra-long baseline.

Methods	IM				IM-ARM				MSAR-DCDN			
AOA	10	13	17	26	10	13	17	26	10	13	17	26
MAE (°)	0.47	13.47	13.67	24.5	0.47	1.63	4	10.7	0.39	0.11	0.06	0.37
RMSE (°)	0.56	13.48	13.67	24.5	0.56	3.98	7.35	16.65	0.39	0.11	0.07	0.42
EAR (%)	99.8	0	0	0	99.8	80.6	65.8	54	100	100	100	100

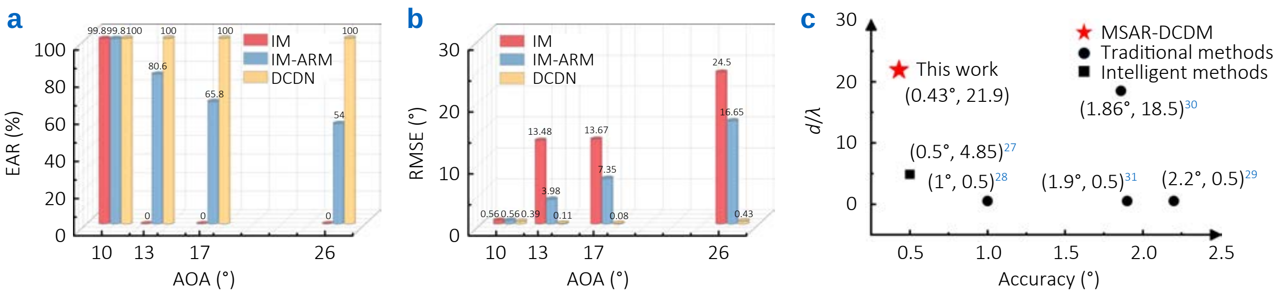

Fig. 10 | (a) EAR results. (b) RMSE results. (c) The comparison results of the MSAR-DCDN with the existing MWP AOA estimation related research. IM: interferometer algorithm; IM-ARM: interferometer algorithm with the ambiguity resolution method; EAR: estimation accuracy rate.

photonic AOA estimation method for the core demand of "optical hardware + intelligent algorithm" synergy in the field of intelligent optics, and achieves the following key contributions: 1) breaking baseline constraints: the contradiction between baseline length and operating bandwidth is fundamentally alleviated through data-driven nonlinear mapping, realizing UWB unambiguous AOA estimation under fixed long baselines. 2) Improving system performance: the method achieves high-precision (RMSE < 0.42° with 821 mm baseline), wide-field-of-view AOA estimation in UWB scenarios, and its performance is significantly superior to traditional algorithms. 3) Providing a new paradigm: the integration of MWP and deep learning provides a flexible baseline design solution for intelligent optical sensing, laying a technical foundation for the development of next-generation AOA estimation systems.

6 Discussion

The proposed deep learning-assisted AOA estimation method exhibits a noteworthy advantage: it maintains an estimation error within 2° even under negative SNR conditions (−5 dB), while demonstrating stable performance. This stands in sharp contrast to the conventional interferometer-based approach, whose performance deteriorates sharply or fails completely under the same conditions. We attribute this robustness to the inherent feature learning and noise suppression capabilities of deep learning models, which have been validated and theoretically supported in related literature[40,41]. Traditional methods rely on explicit phase difference calculations and covariance matrix estimation steps that are highly sensitive to noise and prone to significant distortion when noise power exceeds signal power. In contrast, deep learning establishes an end-to-end nonlinear mapping from noisy received signals to the true AOA, enabling implicit extraction of noise-invariant signal representations. Specifically, the proposed MSAR-DCDN model employs a multi-scale parallel convolutional structure along with a channel attention mechanism, allowing it to adaptively focus on weak yet informative signal components submerged in strong noise. Furthermore, through hierarchical feature fusion and residual connections, the model effectively suppresses irrelevant noise interference, thereby preserving phase and amplitude information critical for AOA estimation. This internal mechanism ensures that the model maintains high accuracy and strong stability even in challenging low-SNR environments. Beyond noise robustness, system miniaturization and SWaP (size, weight, power) optimization are also key considerations for practical deployment, which highlights inherent challenges in MWP systems. Notably, the integration of MWP systems is not a simplistic "photonic integration" process but a coordinated integration of multiple heterogeneous material platforms, which brings intrinsic difficulties to achieving significant SWaP reduction[42]. Unlike traditional RF electronic devices, which benefit from mature, highly miniaturized, and standard-packaged discrete components, microwave photonic systems require the integration of multiple specialized material platforms to realize their core functions. These include indium phosphide (InP)-based DMLs and PDs, lithium niobate (LN)-based MZMs, and silica-based PLCs. The physical and material incompatibilities among these platforms necessitate complex packaging and coupling structures, which impede further miniaturization and increase overall system weight. These practical challenges highlight that significant progress in SWaP reduction for microwave photonic systems still lies ahead. And it should be clarified that the ultra-wideband (UWB) capability claimed in this work refers to the ultra-wide operating bandwidth of the system rather than ultra-wide instantaneous bandwidth, which is realized by the synergy of the inherent large-bandwidth advantage of the MWP link and the "offline training by typical fixed frequencies + online model calling by real-time detected frequency" workflow of the MSAR-DCDN algorithm. Specifically, during the initial deployment of the system, we generate dedicated training datasets for multiple typical fixed frequency points within the preset UWB range, train independent MSAR-DCDN models for each frequency point, and store the model parameters locally. In routine online operation, the system directly loads the pre-trained model corresponding to the real-time detected signal frequency for AOA estimation without iterative training or additional data collection, which effectively avoids the feature confusion and precision attenuation caused by single-model cross-frequency training.

The trained model exhibits significant limitations in scenario adaptation; its effective applicability is strictly constrained by the data collection scenario and core parameter characteristics of the training data, and its generalization capability heavily depends on the distribution characteristics of the training data. Specifically, the model can only stably process input data that completely matches the training data in terms of collection conditions. When key conditions during data acquisition change (such as adjustment of signal acquisition frequency or reconfiguration of sensor array layout), significant data distribution shift occurs. This shift leads to a mismatch between the feature representations learned by the model and the data from new scenarios, resulting in noticeable degradation in prediction accuracy or processing stability. In such cases, the model must be retrained or fine tuned to adapt to the new application scenario[32]. This limitation essentially reflects the generalization bottleneck of traditional machine learning models under distribution shift conditions and even minor adjustments to local parameters can hardly avoid the need for retraining through existing adaptation strategies.

Beyond the machine learning and traditional microwave photonic AOA estimation methods elaborated above, AOA sensing techniques based on novel optical modulation have become a prominent research direction. By virtue of unique light-matter interaction mechanisms, these methods open up alternative technical avenues for electromagnetic wave angle detection. For example, 2D metasurface-based optical modulation schemes[43] and astigmatic metalens (AML)-assisted LiDAR architectures[44] have been proposed for wide-angle sensing; such approaches feature ultra-compact planar structures, high resolution and wide field of view, providing valuable insights for the miniaturization and integration of AOA estimation systems. Nevertheless, these metasurface-based techniques are currently limited by high design and fabrication costs, and are only suitable for fixed-baseline, narrow-bandwidth angle sensing in high-frequency bands (e.g., visible light, near-infrared). Their micro-nano structural sensitivity also leads to performance degradation under environmental perturbations such as temperature drift and mechanical vibration. Notably, microwave photonic AOA estimation technology is now in an era of multi-technical integration and innovation. The combination of micro-nano optical regulation and data-driven intelligent signal processing is anticipated to become a new research hotspot. This integration can synergistically exploit the ultra-compact structure and wide-angle sensing capability of metasurfaces with the strong anti-interference performance and flexible baseline adaptability of deep learning algorithms, thus paving a promising way for the further advancement and expansion of AOA estimation research.

References

1. Wei ZX, Liu F, Masouros C et al. Toward multi-functional 6G wireless networks: integrating sensing, communication, and security. *IEEE Commun Mag* **60**, 65–71 (2022).
2. Liu F, Cui YH, Masouros C et al. Integrated sensing and communications: toward dual-functional wireless networks for 6G and beyond. *IEEE J Sel Areas Commun* **40**, 1728–1767 (2022).
3. Luo HL, Gao FF, Yuan WM et al. Beam squint assisted user localization in near-field integrated sensing and communications systems. *IEEE Trans Wirel Commun* **23**, 4504–4517 (2024).
4. Pin Tan DK, He J, Li YC et al. Integrated sensing and communication in 6G: motivations, use cases, requirements, challenges and future directions. In *2021 1st IEEE International Online Symposium on Joint Communications & Sensing (JC&S)* 1–6 (IEEE, 2021). http://doi.org/10.1109/JCS52304.2021.9376324.
5. Tang ZZ, Li YF, Yao JP et al. Photonics-based microwave frequency mixing: methodology and applications. *Laser Photonics Rev* **14**, 1800350 (2020).
6. Spezio AE. Electronic warfare systems. *IEEE Trans Microw Theory Tech* **50**, 633–644 (2002).
7. Lipsky SE. *Microwave Passive Direction Finding* (SciTech Publishing, 2004).
8. Shi R. *Principle, Method and Application for Direction Finding by Interferometer* (Publishing House of Electronics Industry, Beijing, 2023).
9. Schmidt R. Multiple emitter location and signal parameter estimation. *IEEE Trans Antennas Propag* **34**, 276–280 (1986).
10. Gupta P, Kar SP. MUSIC and improved MUSIC algorithm to estimate direction of arrival. In *2015 International Conference on Communications and Signal Processing (ICCSP)* 0757–0761 (IEEE, 2015). http://doi.org/10.1109/ICCSP.2015.7322593.
11. Roy R, Paulraj A, Kailath T. Direction-of-arrival estimation by subspace rotation methods - ESPRIT. In *ICASSP '86. IEEE International Conference on Acoustics, Speech, and Signal Processing* 2495–2498 (IEEE, 1986). http://doi.org/10.1109/ICASSP.1986.1168673.
12. Cao YL, Peng L, Li JZ et al. A new iterative algorithm for geolocating a known altitude target using TDOA and FDOA measurements in the presence of satellite location uncertainty. *Chin J Aeronaut* **28**, 1510–1518 (2015).
13. Kim J. Source localization based on FDOA and TDOA with bounded sensor location errors and changing carrier frequency. *IEEE Internet Things J* **12**, 22478–22488 (2025).
14. Jiang SY, Liu S, Yan FG et al. Over-relaxed ADMM-based unitary adaptive beamforming scheme with lobe-level constraints. *IEEE Trans Signal Process* **73**, 3656–3670 (2025).
15. Capmany J, Novak D. Microwave photonics combines two worlds. *Nat Photonics* **1**, 319–330 (2007).
16. Biernacki PD, Nichols LT, Enders DG et al. A two-channel optical downconverter for phase detection. *IEEE Trans Microw Theory Tech* **46**, 1784–1787 (1998).
17. Huang CJ, Chan EHW. Photonic frequency down conversion approach to measure microwave signal angle of arrival. *IEEE Photonics J* **14**, 7226608 (2022).
18. Zhang X, Chi H, Gao YR et al. Photonic angle-of-arrival measurement of microwave signals using a triangular wave. *Opt Lett* **48**, 5013–5016 (2023).
19. Li P, Yan LS, Ye J et al. Angle-of-arrival estimation of microwave signals based on optical phase scanning. *J Light Technol* **37**, 6048–6053 (2019).
20. Ni Y, Kong X, Wang RX et al. Photonic angle-of-arrival and time-difference-of-arrival measurement based on dual drive 1×2 MZM. *Chin Opt Lett* **11**, 030605 (2013).
21. Chen H, Chan EHW. Simple approach to measure angle of arrival of a microwave signal. *IEEE Photonics Technol Lett* **31**, 1795–1798 (2019).
22. Li GY, Shi DF, Wang L et al. Unambiguous measurement of AOA using a DDMZM. *Opt Commun* **514**, 128132 (2022).
23. Cheng XC, Yuan XJ, Jiang WJ et al. Deep active learning for mmWave array-based multi-source AoA tracking. *IEEE Trans Wirel Commun* **23**, 12368–12381 (2024).
24. Huang Y, Zhang YJ, Tao J et al. Off-grid DOA estimation via a deep learning framework. *Sci China Inf Sci* **66**, 222305 (2023).
25. Zheng SL, Yang Z, Shen WG et al. Deep learning-based DOA estimation. *IEEE Trans Cogn Commun Netw* **10**, 819–835 (2024).
26. Bai J, Ge ZQ, Ma L. Phase center calibration for UWB phase interferometer direction finding by virtual baseline. *IEICE Trans Fundam Electron Commun Comput Sci* **E102–A**, 1383–1386 (2019).
27. Li XP, Cao YH, Yao BB et al. Robust generalized Chinese-remainder- theorem-based DOA estimation for a coprime array. *IEEE Access* **6**, 60361–60368 (2018).
28. Cao CH, Zhao YB, Pang XJ et al. A method based on Chinese remainder theorem with all phase DFT for DOA estimation in sparse array. *J Syst Eng Electron* **31**, 1–11 (2020).
29. Yang L, Han YB. A joint spatio-temporal sub-Nyquist sampling structure for wideband receivers. *IEEE Trans Instrum Meas* **73**, 1–15 (2024).
30. Lin KY, Krishna B, Krishna H. Rings, fields, the Chinese remainder theorem and an extension-part I: theory. *IEEE Trans Circuits Syst II Analog Digit Signal Process* **41**, 641–655 (1994).
31. Li Y, Cai QS, Yang J et al. A novel deep learning-assisted microwave photonic direction finding system based on long-baseline array. *J Light Technol* **41**, 4971–4978 (2023).
32. Zhang LL, Li Y, Liao X et al. A low-latency and high-performance microwave photonic AOA and IFM system based on deep learning and FPGA. *IEEE Sens J* **25**, 9934–9945 (2025).
33. Li Y, Cai QS, Yang J et al. Adaptive microwave photonic angle-of-arrival estimation based on BiGRU-CNN [Invited]. *Chin Opt Lett* **21**,

090001 (2023).

34. Xia XG, Liu KJ. A generalized Chinese remainder theorem for residue sets with errors and its application in frequency determination from multiple sensors with low sampling rates. *IEEE Signal Process Lett* **12**, 768–771 (2005).
35. Humeur R. A new high-resolution direction finding architecture using photonics and neural network signal processing for miniature air vehicle applications (Naval Postgraduate School, Monterey, 2015).
36. Li GY, Meng X, Wang L et al. Simultaneous and unambiguous identification of DFS and AOA without dependence on echo signal power. *Opt Lett* **48**, 1028–1031 (2023).
37. Li GY, Shi DF, Wang L et al. Photonic system for simultaneous and unambiguous measurement of angle-of-arrival and Doppler-frequency-shift. *J Light Technol* **40**, 2321–2328 (2022).
38. Doan VS, Huynh ‐ The T, Hoang VP et al. Phase ‐ difference measurement ‐ based angle of arrival estimation using long ‐ baseline interferometer. *IET Radar Sonar Navig* **17**, 449–465 (2023).
39. Chen H, Chan EHW. Angle-of-arrival measurement system using double RF modulation technique. *IEEE Photonics J* **11**, 7200110 (2019).
40. Papageorgiou GK, Sellathurai M, Eldar YC. Deep networks for direction-of-arrival estimation in low SNR. *IEEE Trans Signal Process* **69**, 3714–3729 (2021).
41. Zeng WM, Zhang JL, Zhang TT. A robust machine learning based UWB AOA estimation method. In *2024 IEEE 99th Vehicular Technology Conference (VTC2024-Spring)* 1–5 (IEEE, 2024). http://doi.org/10.1109/VTC2024-Spring62846.2024.10683435.
42. Marpaung D, Yao JP, Capmany J. Integrated microwave photonics. *Nat Photonics* **13**, 80–90 (2019).
43. Pu MB, Li X, Guo YH et al. Nanoapertures with ordered rotations: Symmetry transformation and wide-angle flat lensing. *Opt Express* **25**, 31471–31477 (2017).
44. Gong SJ, Guo YH, Li XY et al. Spectral-acoustic-coordinated astigmatic metalens for wide field-of-view and high spatiotemporal resolution 3D imaging. *Light Sci Appl* **15**, 85 (2026).

Acknowledgements

We are grateful for financial supports from the National Natural Science Foundation of China (No. 62471478).

Author contributions

Xianglin Chen: Writing – original draft, Methodology, Investigation, Formal analysis, Data curation. Yin Li: Writing - review & editing, Supervision, Methodology. Shiru Song: Methodology, Investigation, Formal analysis, Data curation. Yalin Yao: Investigation, Formal analysis. He Cui: Methodology, Formal analysis. Xuan Li: Methodology, Investigation. Zhe Guo: Visualization, Software. Yinlong Tan: Writing – review & editing, Formal analysis. Taolin Liu: Writing – review & editing, Supervision, Formal analysis, Conceptualization. Tian Jiang: Writing – review & editing, Supervision, Funding acquisition, Conceptualization.

Competing interests

The authors declare no competing financial interests.

Published by Editorial Office of *Opto-Electronic Technology*, Institute of Optics and Electronics, Chinese Academy of Sciences.

Opto-Electronic Journals Group
www.oejournal.org

Scan for Article PDF