

# Dynamic polarization fusion network (DPFN) for imaging in different scattering systems

BING LIN, XUEQIANG FAN, PENG PENG, AND ZHONGYI GUO\* 

*School of Computer and Information, Hefei University of Technology, Hefei 230009, China*

\*[guozhongyi@hfut.edu.cn](mailto:guozhongyi@hfut.edu.cn)

**Abstract:** Deep learning has broad applications in imaging through scattering media. Polarization, as a distinctive characteristic of light, exhibits superior stability compared to light intensity within scattering media. Consequently, the de-scattering network trained using polarization is expected to achieve enhanced performance and generalization. For getting optimal outcomes in diverse scattering conditions, it makes sense to train expert networks tailored for each corresponding condition. Nonetheless, it is often unfeasible to acquire the corresponding data for every possible condition. And, due to the uniqueness of polarization, different polarization information representation methods have different sensitivity to different environments. As another of the most direct approaches, a generalist network can be trained with a range of polarization data from various scattering situations, however, it requires a larger network to capture the diversity of the data and a larger training set to prevent overfitting. Here, in order to achieve flexible adaptation to diverse environmental conditions and facilitate the selection of optimal polarization characteristics, we introduce a dynamic learning framework. This framework dynamically adjusts the weights assigned to different polarization components, thus effectively accommodating a wide range of scattering conditions. The proposed architecture incorporates a Gating Network (GTN) that efficiently integrates multiple polarization features and dynamically determines the suitable polarization information for various scenarios. Experimental result demonstrates that the network exhibits robust generalization capabilities across continuous scattering conditions.

© 2023 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

## 1. Introduction

Deep learning has become one of the important methods for solving complex imaging through scattering media [1–3]. In comparison to commonly used light-intensity images, polarization, being a stronger stability of light attributes in scattering media, plays a pivotal role in imaging through scattering media [4–6]. Utilizing polarization information enables imaging or detection capabilities that go beyond the reach of light-intensity information in specific environments like underwater [5,7,8], biological tissues [9–12], atmosphere [13,14], and so on [15,16]. Thanks to the rapid developments of deep learning algorithms driven by polarization datasets, which have also emerged and shown extraordinary results. Leveraging the unique properties of polarization, these algorithms exhibit outstanding performance and heightened stability in fields including underwater imaging [17,18], dehazing [19], denoising [20–22], etc. The integration of polarization datasets into deep learning frameworks shows their potential to enhance outcomes and improve generalization. Deep learning based imaging, not only needs to perform high-quality imaging of different targets under a single scattering condition, but also needs to extend the scope of application to a wider range of scattering environments. This is the pursuit of a greater degree of one-to-many scenario adaptability, that is, accommodating more scenarios and variables within a given environment. Although many methods have been proposed [23–25], they can only show optimal performance when the scattering conditions in the test data match the training data well. Therefore, in order to optimize such Expert Networks (EN), the availability of training data needs to have sufficient prior knowledge of scattering media and sufficient stability within the scattering environment. Polarization information exhibits distinctive stability within scattering

media, but there is a pertinent challenge lies in effectively selecting polarization features allowing the input feature to maximize the characterization of different scattering environments and targets' prior knowledge. For instance, in the research [26], various polarizations were analyzed using different pipes to attain complete effective characteristics. Although the final experimental outcomes demonstrated good performance in the same environment, the reconstructed results for different-material targets still had some challenges. Another approach is to train a single generic network (GN) with a larger dataset containing different scattering conditions. However, this approach is generally less performant due to the need to extract generalizable features in different scattering situations [2,27]. What's more, polarization cannot be directly detected, necessitating the adoption of various expression methods to characterize polarization information. Like the commonly used Stokes vector, Jones vector, and the indices of polarimetric purity (IPPs) [9,11,12] based on the Muller Matrix (MM). These methods enable the quantification and representation of polarization attributes in a reliable and meaningful manner. However, it is important to note that different expression methods exhibit varying degrees of sensitivity to different scenarios and objectives, which makes the most suitable polarization characteristics different for different scenes and different targets. For instance, Stokes vectors tend to offer a comprehensive representation of polarization states [28,29], and Jones vectors are particularly advantageous in situations where precise control over the polarization state is required [30]. Most of the existing data-driven methods using polarization datasets usually directly use ( $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ ,  $135^\circ$ ) as datasets [31] and other methods use specific parameters or combinations of parameters within the Stokes vector as datasets [24,32]. While by inputting the multi-dimensional polarization dataset, the information dimension will be increased and the feature expression can be enhanced, in which it is important to consider the issue of information redundancy and its associated overfitting. In addition, we cannot guarantee that existing representations are optimal for all scenarios and targets. Furthermore, incorporating targeted and advanced structures into the network model serves as a crucial approach to enhance the generalization capabilities of the model [24,33]. However, it is essential to consider that excessively large models present a significant challenge in terms of computational resources. Therefore, considering these limitations, in this paper, we propose a dynamic polarization fusion network (DPFN) to select polarization information to adapt to the imaging needs of multiple scenes, so as to realize the effective selection of polarization features to cope with different scattering scenarios and ensure the generalization of the network framework to a wider range of scenes while achieving high-quality imaging.

Our work is inspired by the framework of the mixture of experts (MoE) [34,35]. The unique properties of our DPFN include multiple polarization feature representations of the inputs and dynamic adjustment of fusion parameters, both of which are adjusted simultaneously to achieve scene-specific feature selection. The DPFN also alleviates the limitations of model switching between fixed expert DNNs, which makes DPFN more versatile and scalable. The DPFN is capable of synthesizing features in continuous, high-dimensional feature spaces to provide optimal performance under different scattering conditions. The DPFN's adaptability stems from the interaction between the Gating Network (GTN) and the Deep Neural Network (DNN) combination of polarization experts. Each EN in the DPFN extracts certain polarization features to provide different input representations. According to the guidance of centralized target features and the feedback of the training process, the GTN dynamically and intelligently fuses the extracted features, to synthesize the feature representation suitable for the current input. Finally, the target information is restored by the decoder. We verify the effectiveness of our proposed method through different validation experiments and ablation experiments and provide the possibility to realize the use of a single DNN architecture under a wide range of scattering conditions. Here, we selected the scattering environment of a milk-water mixture (by adding 7-10 ml milk into water tank) as a representative example to validate the effectiveness of our method.

The structure of this paper is arranged as follows. Section 2 introduces the physical basis and the principles of the framework and the network design of our proposed method. Section 3 presents experimental results and discussions. Finally, in Section 4, we summarize our work.

## 2. Theory and methodology

### 2.1. Polarization information

Polarization characteristics refer to the asymmetric properties exhibited by the direction of vibration perpendicular to the propagation direction. In order to facilitate the analysis and utilization of polarization information, researchers have been actively investigating various representations of polarization, including the Stokes vector, Jones vector, and IPPs. These representations enable a more comprehensive and precise characterization of polarization properties, allowing for advanced applications in the field of polarization optics. By employing a polarizer positioned in front of the detector, the current detection methodology enables the measurement of polarization information at specific orientations, including  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ , and  $135^\circ$ , among others. Subsequently, the obtained data from these measurements is utilized to calculate essential parameters such as the Stokes vector [28] (Eq.1) and Degree of Polarization (DoP), etc.

$$S = \begin{bmatrix} S_1 \\ S_2 \\ S_3 \\ S_4 \end{bmatrix} = \begin{bmatrix} \langle E_{0x}E_{0x}^* + E_{0y}E_{0y}^* \rangle \\ \langle E_{0x}E_{0x}^* - E_{0y}E_{0y}^* \rangle \\ \langle E_{0x}E_{0y}^* + E_{0y}E_{0x}^* \rangle \\ i \langle E_{0x}E_{0y}^* - E_{0y}E_{0x}^* \rangle \end{bmatrix} = \begin{bmatrix} I_{0^\circ} + I_{90^\circ} \\ I_{0^\circ} - I_{90^\circ} \\ I_{45^\circ} - I_{135^\circ} \\ I_{R^\circ} - I_{L^\circ} \end{bmatrix}, \quad (1)$$

In Stokes vectors,  $S_1$  is the total light intensity;  $S_2$  is the difference between horizontal and vertical components;  $S_3$  is the difference between  $45^\circ$  and  $135^\circ$  components, and  $S_4$  is the difference between right-handed and left-handed components. Given the limitations of current detectors, direct acquisition of circular polarization information is not feasible. Therefore, our analysis primarily focuses on linear polarization instead. Then, the degree of linear polarization (DoLP) represents the ratio of the linear polarization component to the total light intensity:

$$DoLP = \frac{\sqrt{S_2^2 + S_3^2}}{S_1}, \quad (2)$$

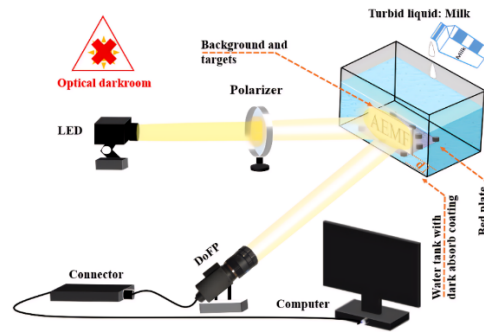
The angle of polarization (AoP) of the linear polarization can be expressed as:

$$AoP = \frac{1}{2} \tan^{-1} \left( \frac{S_3}{S_2} \right), \quad (3)$$

After extensive research in the field of polarization information, it has been observed that different representations of polarization information can capture distinct aspects of target characteristics. For instance,  $S_1$ , one of the parameters in the Stokes vector, can effectively mitigate backscatter interference [29], while DoLP can offer detailed target characteristics [36]. Furthermore, the MM's elements correspond to distinct physical properties of the target [37]. Consequently, it becomes imperative to judiciously select the appropriate polarization feature or a combination of multiple polarization characteristics, based on the specific environmental conditions and unique attributes of the target being analyzed. This selection process is fundamental in enhancing the accurate portrayal and understanding of the target's intrinsic information. Hence, in this article, we attempt to employ a suitable framework to facilitate an effective selection of polarization characteristics.

## 2.2. Measurement system

Currently, publicly available polarization datasets for broad utilization are limited. Therefore, we have established a specialized setup within a controlled laboratory environment to capture underwater polarization datasets. The acquired underwater polarization data obtained from DoFP validate the efficacy of our proposed methodology. Figure 1 illustrates a schematic diagram of the setup employed in this study. To achieve precise control over the captured target information, we systematically introduce varying quantities of milk into the water. We use a transparent glass container measuring 340 mm x 190 mm x 140 mm to hold the clean water. To minimize the effects of ambient light interference, black shields are affixed to three side walls and the bottom of the glass tank. For active illumination, an LED light source is selected, with a linear polarizer situated in front of it, generating horizontally polarized light ( $S = (I, I, 0, 0)^T$ ) [38]. To capture pixel-level corresponding underwater object datasets, we utilize a set of building blocks as the bed plate and supporting connectors fixed onto the bed plate. The targets are affixed to these connectors, thereby ensuring their precise positioning and providing a physical basis for subsequent acquisition of pixel-level corresponding datasets.



**Fig. 1.** Experimental setup.

We systematically capture label images and blurry images in both clear and murky water environments following a consistent sequence. To create a multi-concentration underwater polarization dataset, we acquire fuzzy polarization images under controlled conditions using an increasing concentration of milk-water mixture at a predetermined scattering imaging distance (SID), that is,  $d = 9$  cm. Additionally, we have analyzed the scattering and absorption coefficients of prepared milk-water mixtures [39], and it exhibits a relatively higher scattering coefficient and lower absorption coefficient. In such an environment, the characteristics of the targets' information can be retrieved well after it transfers a long distance, so the target information can be reconstructed by our proposed method. The target is made mostly of iron. We use the commercial DoFP (division of focal plane) polarization camera (LUCID, PHX055S-PC) with pixel counts of  $2048 \times 2448$  to capture images, which placed four polarizers in front of the detector. So, we need to separate four different polarization orientations of  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ , and  $135^\circ$  respectively, in which the size of each polarization image is  $1024 \times 1224$ . Then the corresponding Stokes vector is calculated by Eq.1. We take 110 groups of polarization images at every milk concentration, each group with four polarization directions ( $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ ,  $135^\circ$ ). In addition, we expand every set of them to 2000 to get the training set. The input image size of the neural network in our study is limited to  $256 \times 256$  pixels due to hardware memory constraints and limitations in computational power required for training neural networks.

### 2.3. DPFN framework

In order to achieve an effective selection of polarization characteristics, we design a DPFN, the specific structure of which can be found in Fig. 2. In our study, we demonstrate the utility of our DPFN framework in restoring invisible targets in turbid water. Among the available polarization features, we carefully handpicked three features,  $S_0$ ,  $S_1$ , and DoLP, which possess relatively distinct characterization significance, to serve the purpose of removing scattering effects while also considering computational efficiency. The input to the network is the picture of polarization affected by turbid water, each corresponding to a polarization expert network (PEN). The DPFN is trained to select the appropriate high-dimensional features from the input multi-dimensional polarization picture to fuse to achieve the high-quality recovery of the target information. Firstly, each branch encoder independently extracts a set of multiscale feature maps from the corrupted input. Since each polarization representation focuses on the expression of the target information differently, the combined feature mapping set provides feature representations of different dimensions for removing scattering effects to restore the target. To intelligently utilize these multiscale features to process arbitrary scattering conditions, the GTN is used to reason about the predicted fusion weights to calculate the linear weighted sum of the extracted feature maps. This operation can effectively optimize the feature map to obtain a more general representation under different scattering conditions. Finally, the synthesis features of the encoder are processed by a decoder to achieve the recovery of the target.

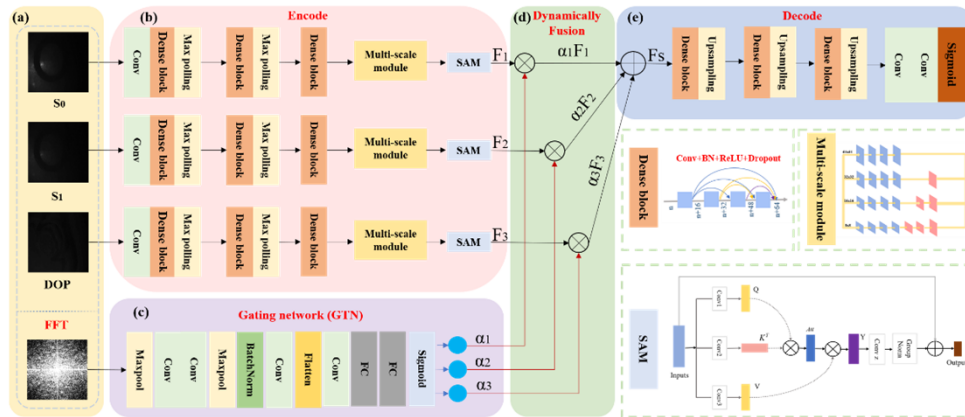
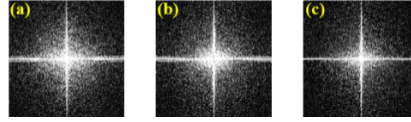


Fig. 2. The schedule of proposed DPFN.

The GTN provides a synthetic method for dynamic fusion by predicting weights. These weights serve as coefficients for the polarization features generated by the PEN. To do that, we design a GTN that extracts centralized target features from the Fourier transform image of a matched fuzzy image, and then generates three coefficients, namely  $\{\alpha_1, \alpha_2, \alpha_3\}$ , which collectively sum up to 1. In the GTN, we leverage the Fourier transform image of the fuzzy target image as input to enhance the precision of target information during weight generation. The Fourier transform feature of an image characterizes the gradient changes in the image pixels and captures target features at the frequency domain level. Notably, the overall characteristics of the transformed image remain relatively consistent across different scattering conditions, as shown in Fig. 3. This characteristic information that centralizes the expression of differences is useful for guiding the generation of fusion weights. The GTN performs a comprehensive analysis of the input Fourier information. It can adaptively blend features extracted by expert encoders and can stably adapt to different scattering scenarios due to its unique properties. Improve the stability and adaptability of the overall framework. During training, all polarization feature expert

encoders, and decoders in DPFN are trained with GTN. In this way, GTN learns how to optimize polarization characteristics for different underwater scattering effects.



**Fig. 3.** Fourier transform of intensity images at different concentrations. (a) 8 ml; (b) 9 ml; (c) 10 ml.

#### 2.4. Network design

The backbone network of the PEN, as shown in Fig. 2(b), is based on our previous work [24]. Since the polarization information itself belongs to high-latitude information, the existing detection method can only be expressed in a two-dimensional way. As a result, we incorporate sampling layers of varying sizes within the backbone network to extract polarization features at different scales. This facilitates a multi-scale and high-dimensional analysis for comprehensive extraction of polarization information. Subsequently, the extracted multi-scale information is integrated by a self-attention mechanism (SAM). The SAM enables the aggregation of effective polarization features by establishing interactions among feature information across different channels and the most useful features for the final goal recovery will be given higher weights [24,40]. The overall feature extraction part is composed of dense blocks, and the down-sampling is realized by the  $2 \times 2$  max-pooling. In the multi-scale module, features are processed into different scales by down-sampling of different sizes, and the high-latitude extraction of features is realized by  $3 \times 3$  convolutional layers.

The GTN follows the VGG structure [41] to predict the synthetic weight  $\alpha_i$ , which is shown in Fig. 2(c). Among them, the convolutional layer with a convolution kernel of  $3 \times 3$  is used, and the max pooling layer with a size of  $2 \times 2$  is used. The synthetic weights are used to blend features extracted from three polarization expert encoders. The final decoding part is still to use dense blocks to analyze the high-dimensional synthetic features, then use up-sampling to restore the features, and finally output the image results of  $256 \times 256$ . The activation function employed in the network structure is a rectified linear unit (ReLU). The overall parameter and FLOPs of DPFN are 23183.57 M and 1.65 M respectively. In our work, we use the Mean squared error (MSE) as the loss function to drive the interaction of polarization features within the network.

$$MSE = \frac{1}{MN} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [G(i,j) - P(i,j)]^2, \quad (4)$$

where  $P(i, j)$  represents the pixel of the reconstructed image,  $G(i, j)$  represents the pixel of the original target, and  $M$  and  $N$  represent the size of the image.

We trained the model in an image processing unit (NVIDIA RTX 3090) using a Pytorch framework with Python 3.6. To get the best optimal model, we trained 200 epochs. The optimizer is the Adam (Add Momentum Stochastic Gradient Descent) with a learning rate of 0.0001. Meanwhile, we train the model in the computational environment that Windows Server 10 (Version 21H1) Intel Core i7.9750 H CPU @2.60 Hz 2.59 GHz, and 16.0 GB of RAM, after training the DPFN just requires about 0.03592s for reconstructing a new test image.

#### 2.5. Imaging quality

In this paper, to assess the quality of the output of the network, we adopt several evaluation metrics, i.e., Pearson Correlation Coefficient (PCC) and Peak Signal-to-Noise (PSNR) [42]. The

PCC with the value between 0-1, is a way to measure the similarity of images. It can be expressed as [43]:

$$PCC = \frac{\sum_{i=1}^w \sum_{j=1}^h (P(i,j) - P_1)(G(i,j) - G_1)}{\sqrt{\sum_{i=1}^w \sum_{j=1}^h (P(i,j) - P_1)^2} \sqrt{\sum_{i=1}^w \sum_{j=1}^h (G(i,j) - G_1)^2}}, \quad (5)$$

where  $P(i, j)$  represents the pixel of the reconstructed image,  $G(i, j)$  represents the pixel of the original target,  $G_1$  and  $P_1$  represent the mean of the original target and the network reconstruction image respectively, and  $M$  and  $N$  represent the size of the image. The PSNR is defined as:

$$PSNR = 10 \times \log_{10} \frac{MAX^2}{MSE}, \quad (6)$$

where  $MAX$  is the maximal value in the image.

### 3. Results and discussions

During the training process, both the expert encoders and decoder and the GTN are trained together. The training set consists of several underwater polarization datasets that contain varying degrees of scattering. Specifically, a series of polarization data sets were obtained by adding 7 ml, 8 ml, 9 ml, and 10 ml of milk into clean water to create a continuously tunable scattering environment. During training, corresponding Fourier transform diagrams are fed into the GTN. We initialized all the expert encoders, decoder, and the GTN with random weights and trained the optimal model.

#### 3.1. Performances of the DPFN

##### 3.1.1. Different targets

In this section, in order to verify the effectiveness of DPFN, we conduct tests on different targets which all are not included in the training set (the scattering environment is the same as the training set), and there are two types of testing targets, the same type as the training set (not seen in the training set) and the different types from the training set (Alphabetical targets and Chinese character targets). Both types of test images are captured in the same experimental environment as the training set. The model's test results are presented in Fig. 4 (We only show the results of the scattering of the maximum concentration in the training set, i.e. 10 ml of milk, in order to refine the results with pronouncement). For the results of the same type as the training set, there is no doubt that they are complete and excellent. Not only that, but the test results of different types from the training set are also excellent. Even Chinese character characters with significantly higher complexity than the training targets can still be fully recovered with high contrast. Based on the results, it concluded that DPFN can effectively recover different types of targets that are not in the training set.

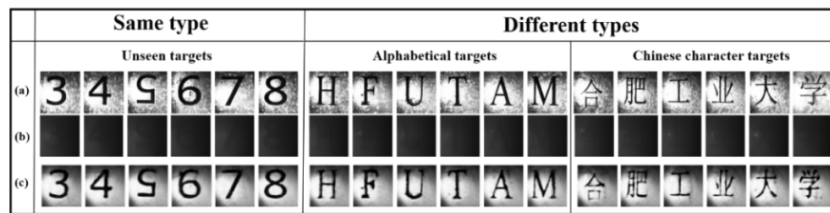


Fig. 4. The results of different types of targets recovered by DPFN.

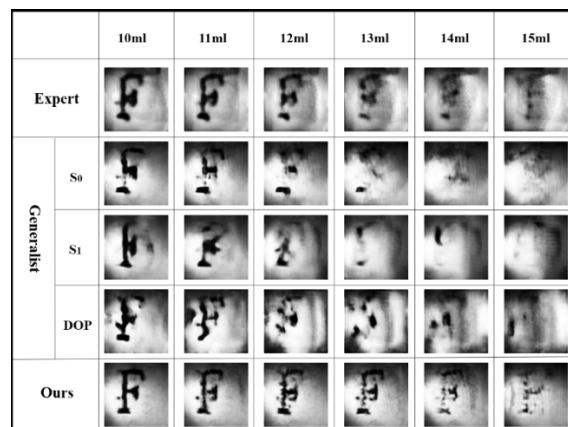
In addition, we calculated the corresponding evaluation indicators. From Table 1, our recovery results are more than 80% similar, and the PSNR is also satisfactory. This shows that our method can not only reconstruct the target with high quality but also does not carry excessive additional noise.

**Table 1. The average PCC and PSNR of the different types of targets**

	Unseen targets	Alphabetical targets	Chinese character targets
PCC	0.8529	0.8388	0.8175
PSNR	14.5494	14.6446	13.8806

### 3.1.2. Compare with EN and GN

In order to further analyze the characteristics of the proposed framework that is more stable, we trained the EN and the GN separately (the training set is the same as the DPFN). The EN is trained with 3D data (including  $S_0$ ,  $S_1$ , and DoLP), obtained under the scattering environment of 10 ml of milk, as a training set. The GN is trained by multi-concentration data in different polarization representation methods ( $S_0$ ,  $S_1$ , DoLP, separately). The network structure used by both EN and GN is composed of expert decoders and decoders as shown in Fig. 2(b) and Fig. 2(c) respectively. The test data are obtained by increasing the milk concentration starting from 10 ml milk as a baseline, and the test target never appeared in the training set. As depicted in Fig. 5, the results under different scattering conditions can be seen. The EN works best for concentrations that are closest to the training set. As the concentration increases, the target result becomes unrecognizable. Despite inputting multi-dimensional polarization information, and based on the stability of the polarization information, the expert model lacks a comprehensive and rational analysis of polarization characteristics beyond the scattering concentration range covered by the training set. The performance of GN is influenced by both the quality of the data and the architecture of the networks. Therefore, in limited circumstances, the GN does not produce better results. From Fig. 5, when polarization information is incorporated into the training set, it is observed that GN exhibits improved performance. However, its stability is particularly poor, which suggests that improving the generalization of GN requires more data and the ability to generalize the characteristics of the corresponding networks.



**Fig. 5.** The schedule of the results from EN, GN and the proposed DPFN.

Notably, the test results of the DPFN model are shown in Fig. 5, compared with the above traditional training methods, our results not only have high-quality and high-contrast imaging



results but also have strong stability. Specifically, even at 15 ml of milk concentration, it is still possible to recover a target that is identifiable and distinguishable. At 13 ml of milk, the EN and GN are no longer able to reconstruct the target, while our method is still able to achieve high-quality imaging without distortion of the target itself. In addition, from the evaluation indicators in Table 2, it can be seen that all the results of our method can reach a similarity of more than 70%, and the value of the PSNR is higher than that of other methods. Furthermore, it is worth noting that the overall numerical level of the results remains relatively stable throughout the experiments, without exhibiting any abrupt or drastic declines.

**Table 2. The average PCC and PSNR of the different training methods within continuously tunable scattering environments**

		10ml	11ml	12ml	13ml	14ml	15ml
EN	PCC	0.6717	0.6607	0.6038	0.5415	0.4102	0.3737
	PSNR	12.4446	12.2671	11.7323	11.1494	10.1712	9.8527
GN-S <sub>0</sub>	PCC	0.6562	0.5319	0.4130	0.4130	0.3274	0.2466
	PSNR	11.7566	10.2889	9.2960	9.4137	9.0605	8.8130
GN-S <sub>1</sub>	PCC	0.7198	0.6774	0.5980	0.4854	0.4305	0.4013
	PSNR	12.7452	12.0436	11.3043	10.0116	12.4979	9.1861
GN-DoP	PCC	0.7547	0.7365	0.7107	0.6563	0.5760	0.5601
	PSNR	13.1379	12.9723	12.5195	11.8034	10.8383	10.6590
Ours	PCC	<b>0.8682</b>	<b>0.8375</b>	<b>0.7969</b>	<b>0.7243</b>	<b>0.7086</b>	<b>0.6989</b>
	PSNR	<b>15.7714</b>	<b>13.9905</b>	<b>13.7269</b>	<b>12.7352</b>	<b>12.0718</b>	<b>11.9817</b>

### 3.2. Performances of DPFN on adapting to a wide range of scattering conditions

#### 3.2.1. Effects of GTN

The starting point for our proposed framework is based on effectively integrating polarization characteristics to enhance adaptability to a broader range of scattering environments. So, in this section, we investigate the adaptation of our framework to handle additional scattering environments, and we conduct ablation experiments to validate its necessity. To do that, we train two models, which have dynamic fusion parts and non-dynamic fusion parts, named DFF and NDF, respectively. We obtain a series of targets with different types of images from the training set under scattering conditions that differ from the training scattering conditions to obtain polarization scattering images as the test data. We mainly obtained the test set in a more intense scattering environment, i.e. adding more volume of milk than training to the water. Then we input the test data into the DFF and NDF to get the recovery result, as shown in Fig. 6(c) and Fig. 6(a).

The experimental results demonstrate that the DFF is relatively more stable when dealing with a series of continuous scattering environments. Moreover, the reconstructed images exhibit completeness without excessive background noise. The NDF already has fringe-like noise when 12 ml of milk, and the subsequent results are incomplete, even unable to recognize the target shape. From Table 3, it can be seen that the DFF has higher values of similarity and PSNR than the NDF, and the change of them is more stable. These highlight that it is very effective to generate scenario-specific weights through the GTN, resulting in a significant 50% increase in the generalization capability beyond the maximum concentration of training conditions.

In addition, we output the results of the middle part of DFF and NDF, which poured milk is 12 ml (because at this point the results of the comparison start to change a lot), visualizing the dynamic fusion part of DPFN, as shown in Fig. 7. It can be seen that the intermediate output of NDF has low feature richness and most of the details are missing compared with the

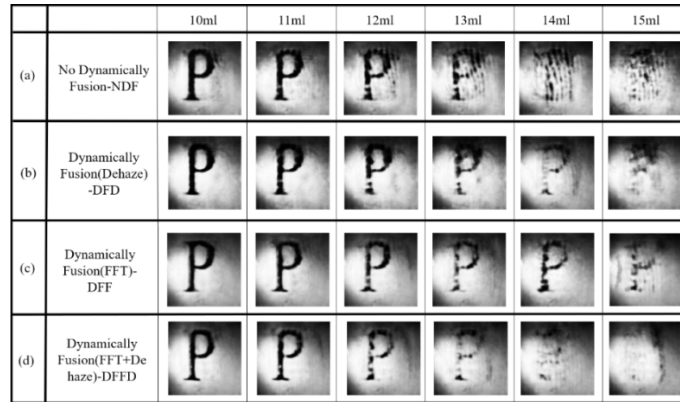


Fig. 6. The result was recovered by different models.

Table 3. The average PCC and PSNR of the different models

		10ml	11ml	12ml	13ml	14ml	15ml
No Dynamically Fusion-NDF	PCC	0.7844	0.7579	0.7154	0.6803	0.5050	0.4176
	PSNR	12.8924	12.7881	12.7201	12.2486	10.4414	9.7618
Dynamically Fusion <sup>a</sup> -DFD	PCC	0.8387	0.8054	0.7834	0.7159	0.5973	0.5088
	PSNR	14.9852	13.7371	13.7447	12.5998	10.2829	10.0987
Dynamically Fusion <sup>b</sup> -DFF	PCC	<b>0.8682</b>	<b>0.8375</b>	<b>0.7969</b>	<b>0.7243</b>	<b>0.7086</b>	<b>0.6989</b>
	PSNR	<b>15.7714</b>	<b>13.9905</b>	<b>13.7269</b>	<b>12.7352</b>	<b>12.0718</b>	<b>11.9817</b>
Dynamically Fusion <sup>c</sup> -DFFD	PCC	0.8661	0.8134	0.7909	0.7203	0.6002	0.5567
	PSNR	13.2551	13.3929	14.0737	12.0134	11.3609	10.1876

<sup>a</sup>The AoP-based dehazing images are input into GTN.

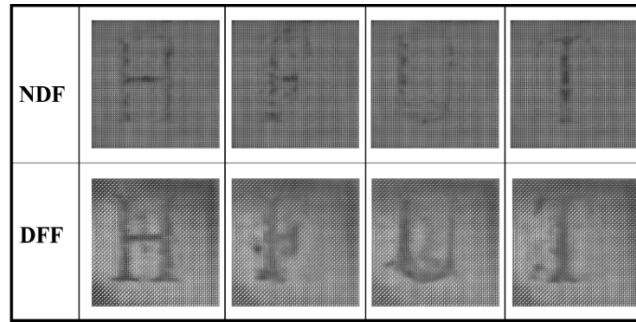
<sup>b</sup>The FFT images are input into GTN.

<sup>c</sup>The AoP-based dehazing images and the FFT images both are input into GTN.

middle output of DFF. The significant differences between the test and training sets make it challenging for the EN model to adapt to these variations. Although direct feature fusion can increase the feature expression, it is still insufficient in addressing the changes in the environment and compensating for the quality of subsequent reconstruction. DFF is optimized simultaneously due to the interaction of expert decoders and the GN. Due to the stability of the polarization information in the scattering media, the optimization of EN can capture generalized polarization features in different scattering environments, and it can be known from our previous work [23]. Then, the GTN can generate suitable weights according to different scattering conditions to fuse polarization features extracted by the EN. These make it possible to prepare features of the decoder, which, despite the increased scattering difference, still has rich features for reconstructing the target (The output of the middle part of the network during concentration increase will be discussed later.).

### 3.2.2. Effects of FFT

A key point of our proposed DPFN is its capacity to process polarization features extracted by EN with appropriate weights for different environments. Therefore, in addition to optimizing the weight generation through the feedback of the loss function during training, the quality of weight generation can also be improved by selecting suitable data input for GTN. So, in this section, we discuss choosing a data type suitable for guiding weight generation. As weight serves



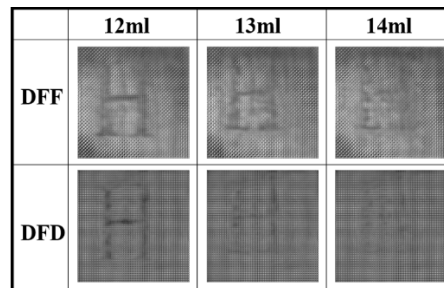
**Fig. 7.** The middle out of NDF and DFF.

to guide the effective fusion of polarization features, it is critical to leverage information that highlights target features to guide polarization features toward a fusion direction more conducive to accurate target reconstruction. The network parameters are also updated and optimized in the direction more suitable for polarization features.

First, we use an AoP-based dehazing method to process the image as the input of the GTN, which is an algorithm with a better effect on turbid scattering [44]. This method aims to directly generate the available target features as input, thereby facilitating the generation of effective fusion weights. In addition, we use a direct processing method, that is, using Fourier transform to process the image as the GTN input. This approach aims to unify target characteristics by centralizing the relevant features, ultimately facilitating the effective generation of fusion weights. We use two methods to train the model to get dynamic networks based on dehazing data and dynamic networks based on Fourier features, named DFD and DFF. Test images obtained under the scattering media with continuously increasing concentrations are input to both models. The results are shown in Fig. 6(b) and Fig. 6(c). The image data after dehazing treatment will show obvious target characteristics to a certain extent, and serve as a GTN input target priori to guide the weight that is conducive to reconstructing the target. Nevertheless, it is important to note that dehazing algorithms, including the one used in our study often relies on parameter estimation. The accuracy of parameter estimation plays a crucial role in determining the final dehazing result. However, as the degree of scattering increases gradually, the risk of estimation failure also becomes more prominent. Therefore, using this type of data as input to the GTN works well in a small range, as shown in Fig. 6(b), but when the concentration increases, the effect begins to drop sharply and cannot adapt to a wider range of scattering conditions. Furthermore, we employ the Fourier transform to directly process the image intensity and feed it as input to GTN. The transformed image expresses target features in a unified manner at the frequency domain representation, thereby facilitating effective guidance for weight generation, with the grayscale value representing the amplitude, and the distance from a certain point to the center representing the frequency. Even with different scattering concentrations, the Fourier transform image has a certain uniformity, from Fig. 3, which can be understood as a kind of generalization representation of the target priori. Therefore, the use of Fourier features can better guide the fusion of polarization features to adapt to different scattering scenarios, from Fig. 6(c). When the concentration of the milk increases, DFF can still maintain relatively stable imaging, and up to 15 ml of milk, it can still reconstruct a result sufficient to identify the target. In addition, from the evaluation indicators calculated in Table 3, on the whole, DFF has a more satisfactory value than DFD in a wider scattering environment in the future. Finally, we use the images of two processing methods with different emphases as input to the GTN, named DFFD, and the results are shown in Fig. 6(d). It can be seen that this method does not improve the imaging quality due to the input of more data, which also shows that the concise data feature expression

method is more suitable for guiding weight generation. From the evaluation data in the Table 3, The DFFD has better performance when the scattering degree is small, but when the scattering degree increases, the similarity of imaging effects decreases sharply. This also justifies our use of Fourier characteristics from the side.

In order to more clearly illustrate the rationality of Fourier's production of guidance weights. We also output the feature map of the middle part with the increase in concentration. From Fig. 8, it can be seen that with the increase in concentration, DFF has richer and more stable characteristics, and the features after dynamic fusion (DFF) have a clear target structure and do not show a large contrast. This can not only have good results in the reconstruction process, but also will not cause overfitting in the subsequent target feature upgrading process, and can adapt to more changing environments. The DFD features are smoother and not rich enough, and cannot provide sufficient differentiated information for subsequent high-dimensional feature extraction to drive network parameter updates, that is, there is no better adaptability to dynamic changes, especially when the milk concentration increases, the characteristic information begins to decrease significantly.



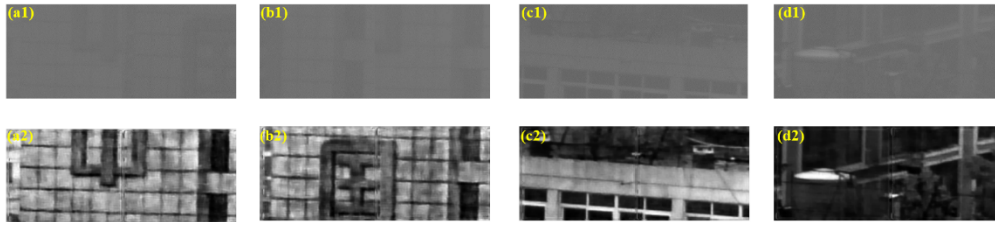
**Fig. 8.** The middle out of DFF and DFD.

### 3.3. Performance of DPFN on new scenarios

Since the training set was obtained in a laboratory environment, in order to reflect the adaptability of our method to different scenarios, we changed the type of scattering environment. Specifically, we collected a separate dataset by photographing outdoor scenes under foggy conditions during daylight. This foggy dataset was subsequently employed to evaluate the capability of our model, obtained from the above training, to handle hazy atmospheric environments. From Fig. 9, our method demonstrates significant dehazing performance. Although the target details are not fully restored, the main information of the picture is recognizable. This shows that the proposed method can indeed use polarization information to extract generalized features of the scattering environment and use Fourier features to effectively adjust dynamic parameters, so that a single model can be used in different scattering scenarios. In addition, we calculated the PSNR of the recovered images as shown in Table 4. And, we can also see the high adaptability of our proposed method to different scenarios from the training set scenarios. These results verify the ability of the framework we designed to adapt to new scenarios and the possibility of implementing multiple scenarios with a single framework.

**Table 4.** The PSNR of the recovery results of the new natural scenes.

	a	b	c	d
PSNR	12.2924	13.8639	12.8453	11.2457



**Fig. 9.** The recovery results of the new natural scenes.

Naturally, given the limited data, more work will be required to ensure that our method works in every scattering environment. This is a common issue for the scattering imaging based on the deep learning algorithm.

#### 4. Conclusion

In this paper, we propose a DPFN to achieve a reasonable selection of polarization characteristics and utilize a single model to adapt to different extensive scattering conditions. Through the interaction of encoder, decoder, and GTN, EN's ability to extract polarization information stability features and generalize target features can be optimized effectively, so that the DPFN can pay more attention to the more generalized expression of target features. Furthermore, The GTN is optimized by utilizing the unique target feature expression method inherent to Fourier features and the interactive feedback. This allows the GTN to generate polarization feature synthesis weights in a direction conducive to accurate target reconstruction. Through a series of ablation experiments, the rationality of our framework as a whole and the superiority of polarization characteristics are proved well. Our approach offers the possibility of multi-scenario use of a single network.

**Funding.** National Natural Science Foundation of China (61775050).

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

**Data availability.** Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

#### References

1. S. Li, M. Deng, J. Lee, *et al.*, "Imaging through glass diffusers using densely connected convolutional networks," *Optica* **5**(7), 803–813 (2018).
2. Y. Li, Y. Xue, and L. Tian, "Deep speckle correlation: a deep learning approach toward scalable imaging through scattering media," *Optica* **5**(10), 1181–1190 (2018).
3. S. Zhu, E. Guo, J. Gu, *et al.*, "Imaging through unknown scattering media based on physics-informed learning," *Photonics Res.* **9**(5), B210–B219 (2021).
4. Q. Xu, Z. Guo, Q. Tao, *et al.*, "Multi-spectral characteristics of polarization retrieve in various atmospheric conditions," *Opt. Commun.* **339**, 167–170 (2015).
5. Q. Xu, Z. Guo, Q. Tao, *et al.*, "Transmitting characteristics of the polarization information under seawater," *Appl. Opt.* **54**(21), 6584–6588 (2015).
6. W. Yu, S. Shah, D. Li, *et al.*, "Polarized computational ghost imaging in scattering system with half-cyclic sinusoidal patterns," *Opt. Laser Technol.* **169**, 110024 (2024).
7. K. Purohit, S. Mandal, and A. N. Rajagoplan, "Multilevel weighted enhancement for underwater image dehazing," *J. Opt. Soc. Am. A* **36**(6), 1098–1108 (2019).
8. B. Huang, T. Liu, H. Hu, *et al.*, "Underwater image recovery considering polarization effects of objects," *Opt. Express* **24**(9), 9826–9838 (2016).
9. D. Li, C. Xu, M. Zhang, *et al.*, "Measuring glucose concentration in a solution based on the indices of polarimetric purity," *Biomed. Opt. Express* **12**(4), 2447–2459 (2021).
10. R. Horstmeyer, H. Ruan, and C. Yang, "Guidestar-assisted wavefront-shaping methods for focusing light into biological tissue," *Nat. Photonics* **9**(9), 563–571 (2015).

11. F. Shen, M. Zhang, K. Guo, *et al.*, "The Depolarization Performances of Scattering Systems Based on Indices of Polarimetric Purity," *Opt. Express* **27**(20), 28337–28349 (2019).
12. F. Shen, B. Zhang, K. Guo, *et al.*, "The depolarization performances of the polarized light in different scattering media systems," *IEEE Photonics J.* **10**(2), 1–12 (2018).
13. T. Hu, F. Shen, K. Wang, *et al.*, "Broad-band transmission characteristics of Polarizations in foggy environments," *Atmosphere* **10**(6), 342 (2019).
14. X. Wang, T. Hu, D. Li, *et al.*, "Performances of polarization-retrieve imaging in stratified dispersion media," *Remote Sens.* **12**(18), 2895 (2020).
15. C. Xu, D. Li, K. Guo, *et al.*, "Computational ghost imaging with key-patterns for image encryption," *Opt. Commun.* **537**(129190), 129190 (2023).
16. C. Xu, D. Li, X. Fan, *et al.*, "High-performance deep-learning based polarization computational ghost imaging with random patterns and orthonormalization," *Phys. Scr.* **98**(6), 065011 (2023).
17. X. Ding, Y. Wang, and X. Fu, "Multi-polarization fusion generative adversarial networks for clear underwater imaging," *Opt. Lasers Eng.* **152**, 106971 (2022).
18. H. Hu, Y. Zhang, X. Li, *et al.*, "Polarimetric underwater image recovery via deep learning," *Opt. Lasers Eng.* **133**(23-24), 106152 (2020).
19. Y. Shi, E. Guo, L. Bai, *et al.*, "Polarization-Based Haze Removal Using Self-Supervised Network," *Front. Phys.* **9**, 789232 (2022).
20. X. Li, H. Li, Y. Lin, *et al.*, "Learning-based denoising for polarimetric images," *Opt. Express* **28**(11), 16309–16321 (2020).
21. H. Liu, X. Li, Z. Cheng, *et al.*, "Pol2Pol: self-supervised polarimetric image denoising," *Opt. Lett.* **48**(18), 4821–4824 (2023).
22. H. Hu, H. Jin, H. Liu, *et al.*, "Polarimetric image denoising on small datasets using deep transfer learning," *Opt. Laser Technol.* **166**, 109632 (2023).
23. D. Li, B. Lin, X. Wang, *et al.*, "High-Performance Polarization Remote Sensing With the Modified U-Net Based Deep-Learning Network," *IEEE Trans. Geosci. Remote Sensing* **60**, 1–10 (2022).
24. B. Lin, X. Fan, and Z. Guo, "Self-attention module in a multi-scale improved U-net (SAM-MIU-net) motivating high-performance polarization scattering imaging," *Opt. Express* **31**(2), 3046–3058 (2023).
25. M. Lyu, H. Wang, G. Li, *et al.*, "Learning-based lensless imaging through optically thick scattering media [J]," *Adv. Photonics* **1**(03), 1 (2019).
26. X. Fan, B. Lin, K. Guo, *et al.*, "TSMPN-PSI: high-performance polarization scattering imaging based on three-stage multi-pipeline networks," *Opt. Express* **31**(23), 38097–38113 (2023).
27. Y. Li, S. Cheng, Y. Xue, *et al.*, "Displacement-agnostic coherent imaging through scatter with an interpretable deep neural network," *Opt. Express* **29**(2), 2244–2257 (2021).
28. G. G Stokes, *Mathematical and physical papers*, (Cambridge University, 1901).
29. J. S. T. Yo, M. P. Rowe, and E. N. Pugh, "Target detection in optically scattering media by polarization-difference imaging," *Appl. Opt.* **35**(11), 1855–1870 (1996).
30. M. Born, *Wolf. Principles of Optics*, (Pergamon, New York, 1975, (pp. 665–668)).
31. H. Liu, X. Li, Z. Cheng, *et al.*, "Polarization Maintaining 3-D Convolutional Neural Network for Color Polarimetric Images Denoising," *IEEE Trans. Instrum. Meas.* **72**, 1–9 (2023).
32. B. Lin, X. Fan, D. Li, *et al.*, "High-Performance Polarization Imaging Reconstruction in Scattering System under Natural Light Conditions with an Improved U-Net," *Photonics* **10**(2), 204 (2023).
33. W. Zhang, X. Li, S. Xu, *et al.*, "Underwater Image Restoration via Adaptive Color Correction and Contrast Enhancement Fusion," *Remote Sens.* **15**(19), 4699 (2023).
34. S. E. Yuksel, J. N. Wilson, and P. D. Gader, "Twenty years of mixture of experts," *IEEE Trans. Neural Netw. Learning Syst.* **23**(8), 1177–1193 (2012).
35. F. Agostinelli, M. R. Anderson, and H. Lee, "Adaptive multi-column deep neural networks with application to robust image denoising," In *Proc 26th International Conference on Neural Information Processing Systems*, 1493–1501 (2013).
36. J. Zhang, J. Shao, J. Chen, *et al.*, "PFNet: an unsupervised deep network for polarization image fusion," *Opt. Lett.* **45**(6), 1507–1510 (2020).
37. J. D. Laan, D. A. Scrymgeour, S. A. Kemme, *et al.*, "Detection range enhancement using circularly polarized light in scattering environments for infrared wavelengths," *Appl. Opt.* **54**(9), 2266–2274 (2015).
38. T. Treibitz and Y. Y. Schechner, "Active Polarization Descattering," *IEEE Trans. Pattern Anal. Mach. Intell.* **31**(3), 385–399 (2009).
39. S. Xu, Y. Xi, W. Liu, *et al.*, "Imaging Dynamics Beneath Turbid Media via Parallelized Single-Photon Detection," *Adv. Sci.* **9**(24), e2201885 (2022).
40. Z. Wang, N. Zou, D. Shen, *et al.*, "Non-local u-nets for biomedical image segmentation," in *Proceedings of the AAAI Conference on Artificial Intelligence* **34**(04), 6315–6322 (2020).
41. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv*, arXiv:1409.1556 (2014).
42. Q. Huynh-Thu and M. Ghanbari, "Scope of validity of PSNR in image/video quality assessment," *Electron. Lett.* **44**(13), 800 (2008).

43. A. Buda, "Life time of correlation between stocks prices on established and emerging markets," [arXiv](#), arXiv:1105.6272 (2011).
44. J. Liang, L. Ren, H. Ju, *et al.*, "Polarimetric dehazing method for dense haze removal based on distribution analysis of angle of polarization," *Opt. Express* **23**(20), 26146–26157 (2015).