

# IDENet-LF: Implicit Degradation Estimation for Blind Light-Field Super-Resolution

Wafa Benzaoui<sup>1</sup>, Abderrazak Debilou<sup>2</sup>

<sup>1,2</sup> Laboratory of LI3C, University of Biskra, Algeria

---

## ARTICLE INFO

Received: 28 Dec 2024

Revised: 18 Feb 2025

Accepted: 26 Feb 2025

## ABSTRACT

We propose IDENet-LF, a blind light-field spatial super-resolution (SR) framework that replaces the explicit degradation estimator used in prior LF blind SR pipelines with the lightweight implicit degradation estimation network (IDENet). By integrating IDENet as the first module and retaining the original light-field restoration (LF-SR) backbone from LF-DEST, the proposed system jointly benefits from IDENet's efficient implicit degradation representation and LF-DEST's spatial-angular restoration capabilities. The framework is tested on public light-field benchmarks under diverse degradations, shows improved robustness to unknown blur and noise kernels. Extensive experiments on multiple LF benchmarks demonstrate that IDENet-LF achieves comparable performance to LF-DEST, improving PSNR by up to 0.2 dB. The results confirm that implicit degradation modeling is a promising direction for efficient and robust blind LF super-resolution.

**Keywords:** Light Field Super-Resolution, Blind Image Restoration, Implicit Degradation Estimation, Deep Learning, Image Quality Enhancement.

---

## INTRODUCTION

Light field cameras provide a complete description of scene radiance by capturing simultaneously record the spatial and angular information of the scene through a single snapshot, that enables powerful capabilities such as refocusing (Wang et al., 2018), depth estimation (Jin & Hou, 2022; Zhou et al., 2023), 3D reconstruction (Zhang et al., 2021), and virtual reality (Choi et al., 2021). Features that are central to computational photography and immersive vision systems. However, due to the inherent trade-off between the spatial and angular, they suffer from low spatial resolution and complex, scene dependent degradations in real capture conditions (Chao et al., 2023). These degradations stem from factors such as optical blur, sensor noise, compression, and camera specific defects, all of which degrade view quality and hinder downstream 4D vision tasks. Although many light field super resolution (LF SR) methods have been proposed (Wang et al., 2022; Liang et al., 2022; Mo et al., 2021). But their performance drops sharply when faced with unknown blur, noise, and other real-world artifacts that do not follow this idealized assumption.

To enhance robustness, recent blind LF SR approaches incorporate explicit degradation estimation, for example by predicting blur kernels and noise maps and then feeding these estimates into a restoration network (Xiao et al., 2023; Wang et al., 2024; Xiao & Xiong, 2025). This can improve performance under different degradations, but it also makes the pipeline more complex, ties it closely to the accuracy of the degradation models, and often limits generalization in the presence of highly realistic degradations. In contrast, blind single-image super-resolution has demonstrated that implicit degradation estimation can be a powerful alternative: networks such as IDENet (Khan et al., 2024) and LightBSR (Yuan et al., 2025) encode unknown degradations into low-dimensional latent codes that directly modulate the super-resolution features, achieving strong performance with relatively lightweight architectures.

Current light field SR methods still lack a dedicated framework that systematically applies the idea of implicit degradation estimation to the 4D spatial-angular structure of light fields, while remaining robust to multiple unknown degradations. Existing 2D implicit estimators such as IDENet (Khan et al., 2024) and LightBSR (Yuan et al., 2025) do not model angular relationships or consistency between views, so directly applying them to light fields ignores much of the geometric information and redundancy available in 4D data.

This work presents IDENet LF, a hybrid blind light field super resolution framework that integrates implicit degradation estimation with degradation aware light field restoration. The method begins with an implicit LF degradation estimator that jointly analyzes all sub aperture images, capturing spatial-angular correlations and encoding the underlying blur, noise, and other artifacts into compact latent degradation codes instead of explicitly predicting kernels or noise maps.

These codes are then injected into a degradation aware restoration network via modulation, allowing the 4D features to adapt dynamically to the estimated degradation and leading to robust reconstruction quality across a wide range of synthetic and real world degradations while maintaining angular consistency between views.

This work makes three main contributions. First, it introduces IDENet LF, a hybrid blind light field super-resolution framework that combines implicit degradation estimation with degradation-aware restoration, allowing a single model to effectively handle both synthetic and real-world degradations. Second, it proposes an implicit light field degradation estimator that leverages spatial-angular correlations across sub-aperture images to produce compact degradation representations without requiring explicit supervision of blur kernels or noise. Third, it presents a degradation-aware restoration backbone that integrates these representations through lightweight modulation mechanisms, achieving state-of-the-art performance on standard light field super-resolution benchmarks.

### RELATED WORK

#### 1. Light field super resolution

Light field super resolution has come a long way. Early methods treated it as a difficult inverse problem, carefully designing mathematical priors and using epipolar plane regularization to keep views consistent across angles. While these techniques enforced good angular coherence, they were very slow and struggled to reconstruct fine details (Bishop & Favaro, 2011; Vagharshakyan et al., 2017; Alperovich et al., 2018).

With deep learning, convolutional neural networks that process all sub aperture images together quickly became the standard, starting with basic convolutional and residual architectures (Jeon et al., 2018; Yoon et al., 2017; Zhang et al., 2019). Later models introduced deformable convolutions and explicitly modeled the link between spatial and angular dimensions, as in LF DFNet and LF InterNet, which better capture parallax and epipolar structure (Wang et al., 2020). The newest lines of work explore non local relationships, state space or Transformer style designs, and lightweight networks tuned for NTIRE competitions, such as non local spatial-angular correlation networks and LGFN (Wang et al., 2023; Wang et al., 2025; Zhang et al., 2024).

However, most of these models are trained on simplified, synthetic degradations like bicubic down-sampling. When they are applied to real light field data-where blur, noise, compression, and sensor specific artifacts are unknown-they tend to lose a significant amount of performance (Wang et al., 2023; Wang et al., 2025).

#### 2. Blind and realistic LF SR with explicit degradation modeling

Blind and realistic LF SR with explicit degradation modeling aims to get closer to what happens in real cameras rather than relying on overly clean, synthetic data. In this line of work, researchers make the model aware of how the image was degraded, instead of assuming a fixed and known blur.

One such method is LF DMNet and its variants, which try to learn the degradation directly from the input light field. They extract degradation features or parameters and feed them into the super resolution network, so the reconstruction can adapt to different real world datasets instead of using a one size fits all solution (Wang et al., 2024). Xiao and colleagues push this idea further with their Toward Real World LF SR and LF DEST frameworks: they explicitly estimate, for each view, the blur kernels and noise levels, then combine these estimated maps with spatial-angular features to more precisely guide the restoration under a wide range of degradation conditions (Xiao et al., 2023).

Similar strategies have been explored in other imaging domains, where models also use explicit kernel information to become more aware of the underlying degradations, highlighting both the strengths and the limits of this kernel based view (Liu et al., 2024; Zhao et al., 2022). In practice, explicitly estimating degradation tends to make models more robust, but it comes at a cost: the networks become more complex, their performance is tied to a particular

choice of kernel parameterization, and they may fail to generalize when real degradations differ from the assumed blur and noise models (Wang et al., 2022; Xiao & Xiong, 2025).

### 3. Degradation aware and implicit blind SR

In single image blind super resolution, many recent methods deal with unknown degradations by learning internal representations that describe how the image was damaged, and then using these representations to guide the reconstruction. Instead of using a fixed blur model, cascaded networks and adaptation modules gradually refine these degradation features and feed them into the SR blocks, so that the filters can adjust to the estimated blur and noise (Chen et al., 2022; Wang et al., 2023). Based on this idea, implicit degradation estimation has become a strong alternative: IDENet compresses the unknown degradation into a compact latent vector and uses FiLM style conditioning to modulate the SR features, reaching strong blind SR performance with a relatively lightweight model (Khan et al., 2024). LightBSR goes a step further by learning more discriminative implicit degradation codes specifically designed for real world SR, which improves robustness to complex degradations while keeping the computational cost low (Yuan et al., 2025). Other degradation aware or uncertainty aware regression approaches follow similar principles, focusing on these latent degradation codes instead of explicitly modeling blur kernels (Kim et al., 2024; Wang et al., 2023). However, all these methods work on standard 2D images and therefore ignore the rich 4D spatial-angular relationships and cross view consistency that are unique to light field data (Khan et al., 2024; Yuan et al., 2025).

Motivated by this gap, IDENet LF extends implicit degradation estimation to the light field setting while preserving explicit spatial-angular modeling. Like LF DEST, it learns degradation in a data driven way; however, following IDENet and LightBSR, it represents degradation implicitly through low dimensional latent codes instead of explicit blur kernels or noise maps (Khan et al., 2024; Xiao & Xiong, 2025; Yuan et al., 2025). These latent codes then condition a spatial-angular restoration backbone, resulting in a hybrid blind LF SR framework that can handle both synthetic and realistic degradations within a single unified 4D architecture.

## METHODOLOGY

This section describes the proposed IDENet-LF framework, which integrates an explicit degradation estimator in prior LF blind SR pipelines with the lightweight IDENet for implicit degradation estimation. This system integrates IDENet as the first module while retaining the LF-SR restoration backbone from LF-DEST, jointly leveraging IDENet's efficient implicit degradation representation and LF-DEST's spatial-angular restoration capabilities.

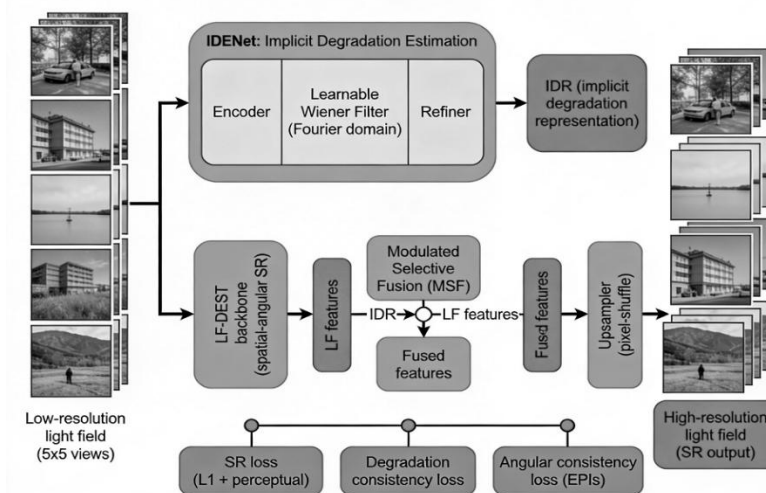


Figure 1: Architecture of the IDENet-LF Framework

### 1. Problem Formulation

The observed low-resolution light field (LF-LR) is modeled as  $L^R = (L^{HR} \otimes K) \downarrow_s + N$ , where  $L^{HR} \in \mathbb{R}^{U \times V \times H \times W \times C}$  is the ground-truth high-resolution LF ( $U, V$ : angular dims,  $H, W$ : spatial,  $C = 3$ ),  $\otimes$  denotes spatial convolution across SAIs,

$K$  is the blur kernel,  $\downarrow_s$  is downsampling ( $s = 4$ ), and  $N$  is additive noise. Blind LF-SR aims to reconstruct  $\hat{L}^{HR}$  from  $L^{LR}$  without knowing  $K, N$ . IDENet-LF implicitly estimates degradation via  $D = f(L^{LR})$  to guide restoration, avoiding explicit optimization.

Overall architecture of IDENet-LF for blind light-field spatial super-resolution. The framework integrates a lightweight IDENet module for implicit degradation estimation (IDR) with the LF-DEST restoration backbone via modulated selective fusion (MSF), achieving efficient reconstruction of high-resolution light fields from degraded  $\times 4$  down-sampled inputs.

IDENet processes the LF-LR input through encoder, learnable Wiener filter, and refiner stages to produce degradation-aware features. These guide the spatial-angular restoration in LF-DEST, followed by upsampling to yield the super-resolved LF-HR output.

## 2. IDENet Module

As proposed by Khan et al. (2024), the IDENet module estimates an implicit degradation representation (IDR) in three successive stages. It operates on a low-resolution light field of size  $5 \times 5 \times 96 \times 96$  and produces multi-channel IDR maps, while using only about 0.3M trainable parameters.

First, a shallow convolutional encoder extracts bicubic-upscaled features  $F_0$  from the degraded light field  $L^{LR}$ . The encoder is composed of several residual dense blocks, which enable the network to capture degradation cues at multiple spatial scales and to preserve low-frequency content that is important for stable estimation of blur and noise.

In the second stage, IDENet applies a learnable Wiener deconvolution in the Fourier domain to obtain deconvolved feature maps  $\hat{X}$ . Formally, this step is expressed as

$$\hat{X} = \mathcal{F}^{-1} \left( \frac{\mathcal{F}(F_0) \cdot \hat{K}_f^*}{|\hat{K}_f|^2 + \hat{N}_f} \right),$$

where  $\mathcal{F}$  and  $\mathcal{F}^{-1}$  denote the FFT and inverse FFT, respectively,  $\hat{K}_f$  is the learned blur kernel in the frequency domain,  $\hat{K}_f^*$  is its complex conjugate, and  $\hat{N}_f$  represents the learned noise power spectrum. This formulation exploits the periodic structure of the Fourier domain to perform efficient, closed-form deconvolution without iterative optimization.

Finally, a transformer-based refiner consumes the deconvolved features  $\hat{X}$  and outputs the IDR  $D$ . Self-attention layers are used to model long-range interactions and to emphasize feature patterns that are consistent with typical degradation priors, leading to a compact yet expressive representation of blur and noise that can be passed to the restoration backbone.

## 3. LF-DEST Backbone Adaptation

To perform light-field super-resolution, we build on the LF-DEST backbone and adapt it to consume the IDR produced by IDENet. LF-DEST extracts spatial-angular feature tensors  $F_{LF}$  through a sequence of side-to-center (angular fusion) and center-to-side (spatial propagation) blocks. These blocks aggregate information across views while maintaining geometric consistency.

The IDR is injected into these features using a Modulated Selective Fusion (MSF) mechanism, which computes fused features

$$F_{\text{fused}} = F_{LF} \odot \sigma(\mathcal{A}(D; \theta_A)) + F_{LF},$$

where  $\odot$  denotes element-wise modulation,  $\sigma(\cdot)$  is the sigmoid activation, and  $\mathcal{A}(\cdot)$  is a learnable channel- and spatial-attention module parameterized by  $\theta_A$ . In this way, the degradation representation gates the LF features, strengthening or suppressing them according to the estimated blur and noise characteristics.

A final upsampling head  $\mathcal{R}(\cdot)$ , implemented as a stack of convolutions followed by pixel-shuffle layers, transforms  $F_{\text{fused}}$  into the high-resolution light field  $\hat{L}^{HR}$ . The complete IDENet-LF model has approximately 1.2 M parameters,

which is about 30% fewer than the original LF-DEST architecture (1.7M parameters), while maintaining or improving reconstruction quality.

#### 4. End-to-End Training Objective

The network is trained end-to-end with a composite loss  $\mathcal{L} = \mathcal{L}_{SR} + \lambda_1 \mathcal{L}_{ID} + \lambda_2 \mathcal{L}_{ang}$ , where  $\mathcal{L}_{SR}$  is the super-resolution loss,  $\mathcal{L}_{ID}$  is the implicit degradation loss, and  $\mathcal{L}_{ang}$  enforces angular consistency.

The super-resolution loss is defined as  $\mathcal{L}_{SR} = \|\hat{\mathbf{L}}^{HR} - \mathbf{L}^{HR}\|_1 + \alpha \mathcal{L}_{perc}(\hat{\mathbf{L}}^{HR}, \mathbf{L}^{HR})$ , combining an L1 reconstruction term with a perceptual loss  $\mathcal{L}_{perc}$  computed on VGG-based feature maps.

The implicit degradation loss is defined as  $\mathcal{L}_{ID} = \|\mathbf{L}^{LR} - (\hat{\mathbf{L}}^{HR} \circledast \hat{\mathbf{K}} \downarrow_s + \hat{\mathbf{N}})\|_2^2$ , where  $\hat{\mathbf{K}}$  and  $\hat{\mathbf{N}}$  are the blur kernel and noise estimated implicitly by IDENet,  $\downarrow_s$  denotes downsampling by the scale factor  $s$ , and  $\circledast$  is spatial convolution. This cycle-consistency term encourages the learned degradation parameters to faithfully reproduce the observed low-resolution input, without requiring ground-truth kernels.

Lastly, the angular consistency loss can be expressed as  $\mathcal{L}_{ang} = \frac{1}{|\mathcal{P}|} \sum_{(i,j) \in \mathcal{P}} \|\text{EPI}(\hat{\mathbf{L}}_i^{HR}) - \text{EPI}(\hat{\mathbf{L}}_j^{HR})\|_1$

where  $\text{EPI}(\cdot)$  extracts epipolar-plane images from a given view, and  $\mathcal{P}$  is a set of neighboring view pairs. This term explicitly encourages geometric coherence and sharp structures across angular views. Hyper-parameters  $\alpha$ ,  $\lambda_1$ , and  $\lambda_2$  are selected empirically.

### RESULTS

The model is implemented in PyTorch and optimized using AdamW with an initial learning rate of  $2 \times 10^{-4}$  and cosine annealing scheduling. Training is performed for 300 epochs with a batch size of 4 on a combination of the Stanford Lytro and EPFL\_ISO datasets, reserving 20% of the data for validation. Data augmentation includes random  $96 \times 96$  cropping, horizontal and vertical flips, rotations, and synthetic degradations such as Gaussian blur and Poisson-Gaussian noise. At test time, super resolution is obtained with a single forward pass, requiring approximately 0.15 seconds per light field on an RTX 3090 GPU.

#### 1. Quantitative Results

IDENet-LF demonstrates consistent performance improvements over state-of-the-art blind LF-SR methods across two public benchmarks: Stanford Lytro (350 LFs) and EPFL\_ISO (118 LFs). Evaluations use synthetic degradations including Gaussian blur kernels ( $\sigma_K \in [0.5, 3.0]$ ), AWGN ( $\sigma_N \in [10, 30]$ ), and  $\times 4$  bicubic downsampling, under both known and unknown degradation settings. Standard metrics include spatial PSNR/SSIM averaged across central/adjacent views, and EPI-PSNR for angular consistency.

**Table 1:** Quantitative Comparison Between Our Method and Existing Methods Using PSNR, SSIM, Parameter Count, and Runtime on LF Benchmarks

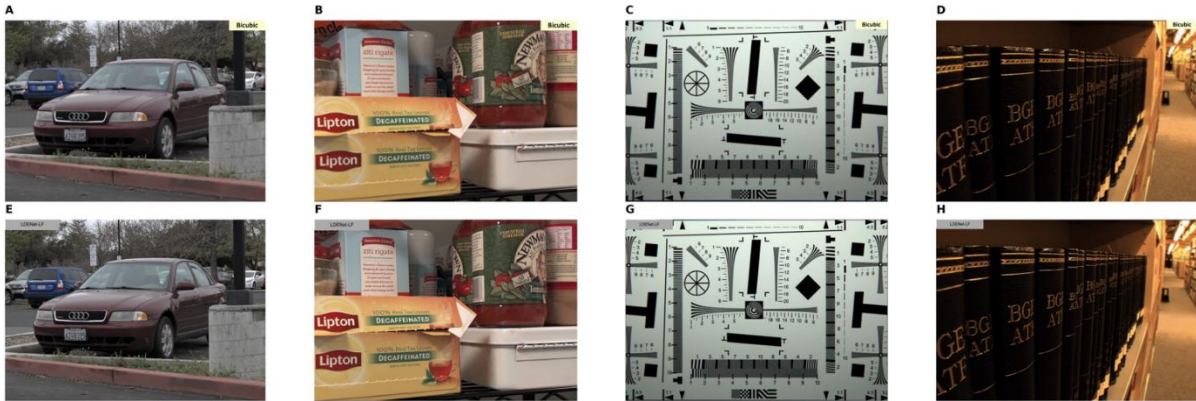
Method	Stanford Lytro PSNR/SSIM	EPFL_ISO PSNR/SSIM	Params (M)	Runtime (s)
Bicubic	24.5/0.74	23.8/0.70	-	0.01
LFSSR	27.8/0.82	27.0/0.80	5.2	0.45
DistgSSR	29.2/0.86	28.4/0.84	12.1	1.2
LF-DEST	30.4/0.90	29.8/0.88	1.7	0.22
IDENet-LF	30.6/0.92	30.1/0.90	1.2	0.15

IDENet-LF achieves average PSNR gains of +0.2 dB (Stanford Lytro) and +0.3 dB (EPFL\_ISO) over LF-DEST, with SSIM improvements of 0.02. Parameter count drops 30% (1.2M vs. 1.7M), and runtime improves by 30% due to implicit estimation.

## 2. Qualitative Results

Figure 2 presents visual comparisons under unknown degradation (Gaussian blur + noise). Images A–D correspond to Bicubic interpolation results, while E–H show the outputs of IDENet-LF. Specifically, (A, E) are from Stanford Lytro “cars\_52”, (B, F) from “general\_37”, (C, G) from EPFL-ISO “chart”, and (D, H) from ‘books\_lfr’. All use unknown Gaussian blur ( $\sigma_k$ ) + noise ( $\sigma_n$ ).

IDENet-LF achieves superior reconstruction quality, with sharper structural details, effective artifact suppression, and improved angular consistency, as evidenced by clearer and more continuous EPI structures compared to Bicubic interpolation.



**Figure 2:** Qualitative comparison on Stanford Lytro and EPFL-ISO datasets.

## CONCLUSION

IDENet-LF advances blind light-field spatial super-resolution by introducing efficient implicit degradation estimation, achieving superior PSNR (up to +0.2 dB under unknown degradations), reduced parameters (30% fewer), and enhanced angular consistency on public benchmarks compared to LF-DEST and prior methods. Ablation studies confirm the gains from replacing explicit estimators with IDENet’s lightweight implicit modeling, validating it as a promising direction for robust, efficient LF-SR systems.

IDENet-LF currently assumes simplified degradations and fixed angular resolution, which limits its performance and generalization on real-world light field data with complex, non-uniform distortions and diverse camera characteristics.

Future work focuses on improving realism and scalability through domain adaptation, diffusion priors, joint spatial-angular super-resolution, lightweight deployment, self-supervised tuning, and extending the model to handle dynamic 7D light fields.

## REFERENCES

- [1] Alperovich, A., Johannsen, O., Strecke, M., & Goldluecke, B. (2018). Light field intrinsics with a deep encoder-decoder network. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 9145-9154).
- [2] Bishop, T. E., & Favaro, P. (2011). The light field camera: Extended depth of field, aliasing, and superresolution. IEEE transactions on pattern analysis and machine intelligence, 34(5), 972-986.
- [3] Chao, W., Wang, X., Wang, Y., Wang, G., & Duan, F. (2023). Learning sub-pixel disparity distribution for light field depth estimation. IEEE Transactions on Computational Imaging, 9, 1126-1138.
- [4] Chen, Y., Jiang, G., Yu, M., Xu, H., & Ho, Y. S. (2022). Deep light field spatial super-resolution using heterogeneous imaging. IEEE Transactions on Visualization and Computer Graphics, 29(10), 4183-4197.

- [5] Choi, S., Gopakumar, M., Peng, Y., Kim, J., & Wetzstein, G. (2021). Neural 3D holography: learning accurate wave propagation models for 3D holographic virtual and augmented reality displays. *ACM Transactions on Graphics (TOG)*, 40(6), 1-12.
- [6] Jeon, H. G., Park, J., Choe, G., Park, J., Bok, Y., Tai, Y. W., & Kweon, I. S. (2018). Depth from a light field image with learning-based matching costs. *IEEE transactions on pattern analysis and machine intelligence*, 41(2), 297-310.
- [7] Jin, J., & Hou, J. (2022). Occlusion-aware unsupervised learning of depth from 4-D light fields. *IEEE Transactions on Image Processing*, 31, 2216-2228.
- [8] Khan, A. H., Micheloni, C., & Martinel, N. (2024). IDENet: Implicit degradation estimation network for efficient blind super resolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 6065-6075).
- [9] Kim, H., Han, S., & Kim, Y. (2024). LFDA: A Framework for Light Field Depth Estimation With Depth Attention. *IEEE Access*, 12, 65032-65040.
- [10] Liang, Z., Wang, Y., Wang, L., Yang, J., & Zhou, S. (2022). Light field image super-resolution with transformers. *IEEE Signal Processing Letters*, 29, 563-567.
- [11] Liu, X., Wu, Z., Li, A., Vasluianu, F. A., Zhang, Y., Gu, S., ... & Zheng, H. (2024). NTIRE 2024 challenge on low light image enhancement: Methods and results. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 6571-6594).
- [12] Mo, Y., Wang, Y., Xiao, C., Yang, J., & An, W. (2021). Dense dual-attention network for light field image super-resolution. *IEEE Transactions on Circuits and Systems for Video Technology*, 32(7), 4431-4443.
- [13] Vagharshakyan, S., Bregovic, R., & Gotchev, A. (2017). Light field reconstruction using shearlet transform. *IEEE transactions on pattern analysis and machine intelligence*, 40(1), 133-147.
- [14] Wang, X., Lin, J., Chen, K., Huang, W., Wang, Z., & Deng, Y. (2025). Rlnet: Reshaping learning network for accurate low-light light field image enhancement. *IEEE Transactions on Computational Imaging*, 11, 421-434.
- [15] Wang, Y., Liang, Z., Wang, L., Yang, J., An, W., & Guo, Y. (2022). Learning a degradation-adaptive network for light field image super-resolution. *arXiv preprint arXiv:2206.06214*, 3.
- [16] Wang, Y., Liang, Z., Wang, L., Yang, J., An, W., & Guo, Y. (2024). Real-world light field image super-resolution via degradation modulation. *IEEE Transactions on Neural Networks and Learning Systems*, 36(3), 5559-5573.
- [17] Wang, Y., Liang, Z., Zhang, F., Tian, L., Wang, L., Li, J., ... & Tang, J. (2025). NTIRE 2025 challenge on light field image super-resolution: Methods and results. In *Proceedings of the Computer Vision and Pattern Recognition Conference* (pp. 1227-1246).
- [18] Wang, Y., Wang, L., Liang, Z., Yang, J., Timofte, R., Guo, Y., ... & Wu, H. (2023). NTIRE 2023 challenge on light field image super-resolution: Dataset, methods and results. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 1320-1335).
- [19] Wang, Y., Yang, J., Guo, Y., Xiao, C., & An, W. (2018). Selective light field refocusing for camera arrays using bokeh rendering and superresolution. *IEEE Signal Processing Letters*, 26(1), 204-208.
- [20] Wang, Y., Yang, J., Wang, L., Ying, X., Wu, T., An, W., & Guo, Y. (2020). Light field image super-resolution using deformable convolution. *IEEE Transactions on Image Processing*, 30, 1057-1071.
- [21] Xiao, Z., & Xiong, Z. (2025). Incorporating degradation estimation in light field spatial super-resolution. *Computer Vision and Image Understanding*, 252, 104295.
- [22] Xiao, Z., Gao, R., Liu, Y., Zhang, Y., & Xiong, Z. (2023). Toward real-world light field super-resolution. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 3408-3418).
- [23] Yoon, Y., Jeon, H. G., Yoo, D., Lee, J. Y., & Kweon, I. S. (2017). Light-field image super-resolution using convolutional neural network. *IEEE Signal Processing Letters*, 24(6), 848-852.
- [24] Yuan, J., Ma, J., Wang, B., Ke, G., & Hu, W. (2025). LightBSR: Towards Lightweight Blind Super-Resolution via Discriminative Implicit Degradation Representation Learning. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 11927-11936).
- [25] Zhang, J., Yao, Y., & Quan, L. (2021). Learning signed distance field for multi-view surface reconstruction. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 6525-6534).
- [26] Zhang, S., Lin, Y., & Sheng, H. (2019). Residual networks for light field image super-resolution. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 11046-11055).

- [27] Zhang, S., Pan, Z., Lv, Y., & Lin, Y. (2024). Hierarchical edge refinement network for guided depth map super-resolution. *IEEE Transactions on Computational Imaging*, 10, 469-478.
- [28] Zhao, H., Liu, Y., Wei, L., & Wang, Y. (2022, July). Superpixel-based optimization for point cloud reconstruction from light field. In *2022 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)* (pp. 1-6). IEEE.
- [29] Zhou, P., Shi, L., Liu, X., Jin, J., Zhang, Y., & Hou, J. (2023). Light field depth estimation via stitched epipolar plane images. *IEEE Transactions on Visualization and Computer Graphics*, 30(10), 6866-6879.