Ai XIAO, Zhi LI, Guomei WANG, Long ZHENG, Haoyuan SUN, 2025. SRIS-Net: a robust image steganography algorithm based on feature score maps. *Frontiers of Information Technology & Electronic Engineering*, 26(6):930-945. https://doi.org/10.1631/FITEE.2400069

SRIS-Net: a robust image steganography algorithm based on feature score maps

Key words: Image steganography; Robustness; Undetectability; Dual-task discriminator

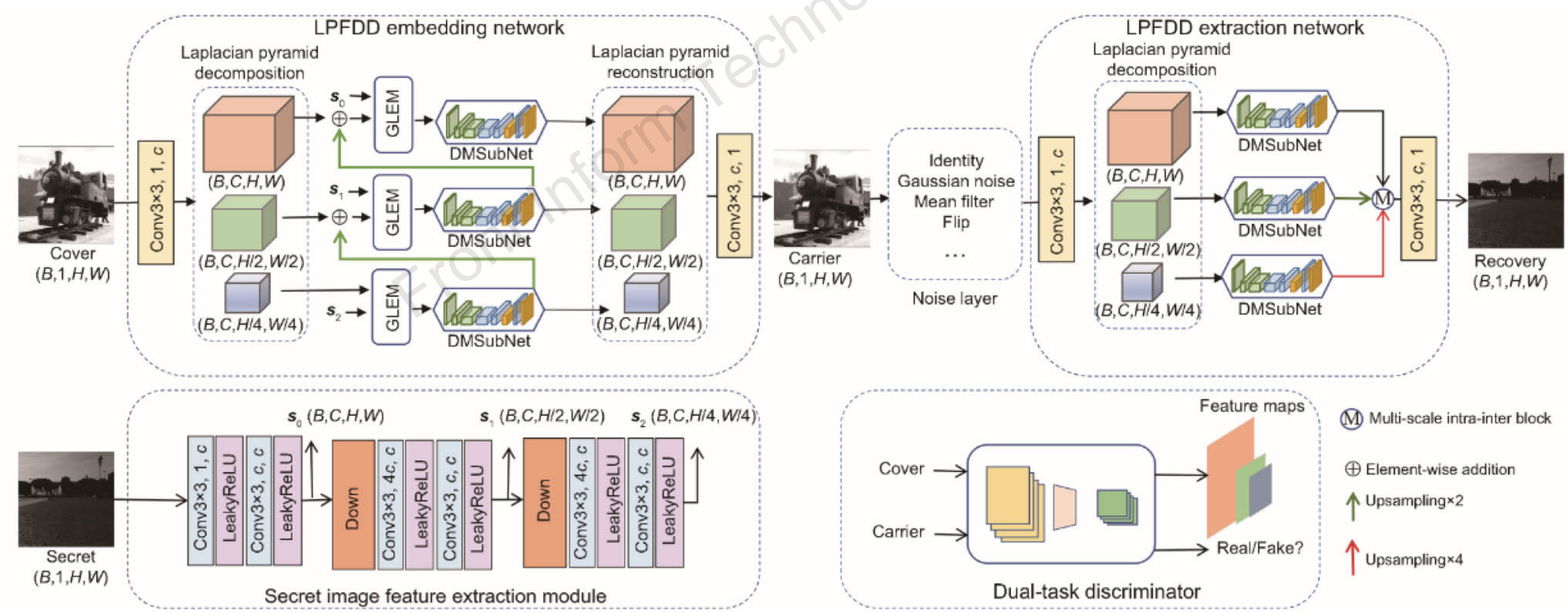
Corresponding author: Zhi LI
E-mail: zhili@gzu.edu.cn
ORCID: https://orcid.org/0000-0001-9813-4979

Motivation

- The rapid development of the Internet has posed significant challenges to the security of multimedia information during storage and transmission.
- Deep learning-based steganography methods have achieved promising results, but most focus on optimizing a single performance metric (e.g., imperceptibility, capacity, or robustness), failing to achieve a good balance among them.
- Existing models often rely on either spatial- or frequency-domain features. However, single-domain features are insufficient to comprehensively represent image content, leading to poor performance in multi-task steganographic scenarios.

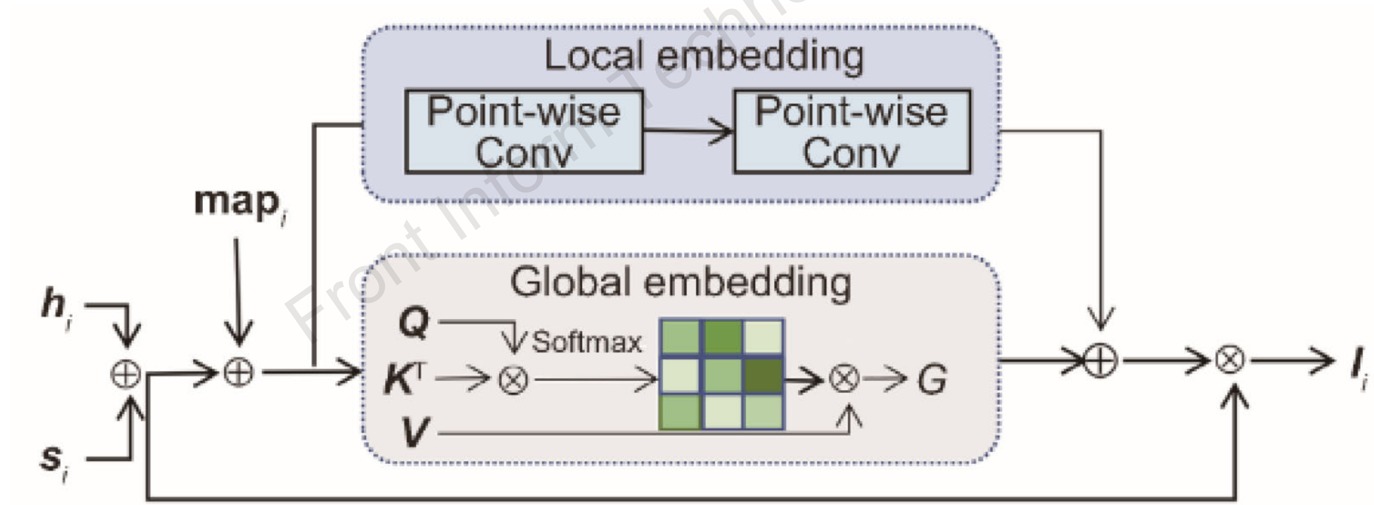
1) SRIS-Net framework

- SRIS-Net is a novel steganographic framework that integrates a dual-task discriminator to generate feature score maps as feedback. It performs Laplacian pyramid-based frequency domain decomposition (LPFDD) in the spatial domain and achieves secret image hiding through a progressive embedding and reconstruction strategy across multiple frequency sub-bands.



2) Global–local embedding module (GLEM)

- GLEM embeds secret information adaptively by considering both the overall image structure and local details.
- The use of global and local branches enables adaptive embedding, improving both robustness and security.



3) Dual multi-scale aggregation sub-network (DMSubNet)

- DMSubNet incorporates multi-scale intra-blocks (MSIBs) and performs hierarchical down-sampling and up-sampling operations to extract features at different scales.
- The selective fusion module (SFM) avoids direct addition of shallow and deep features, which may lead to reconstruction failure due to interference from carrier image information.

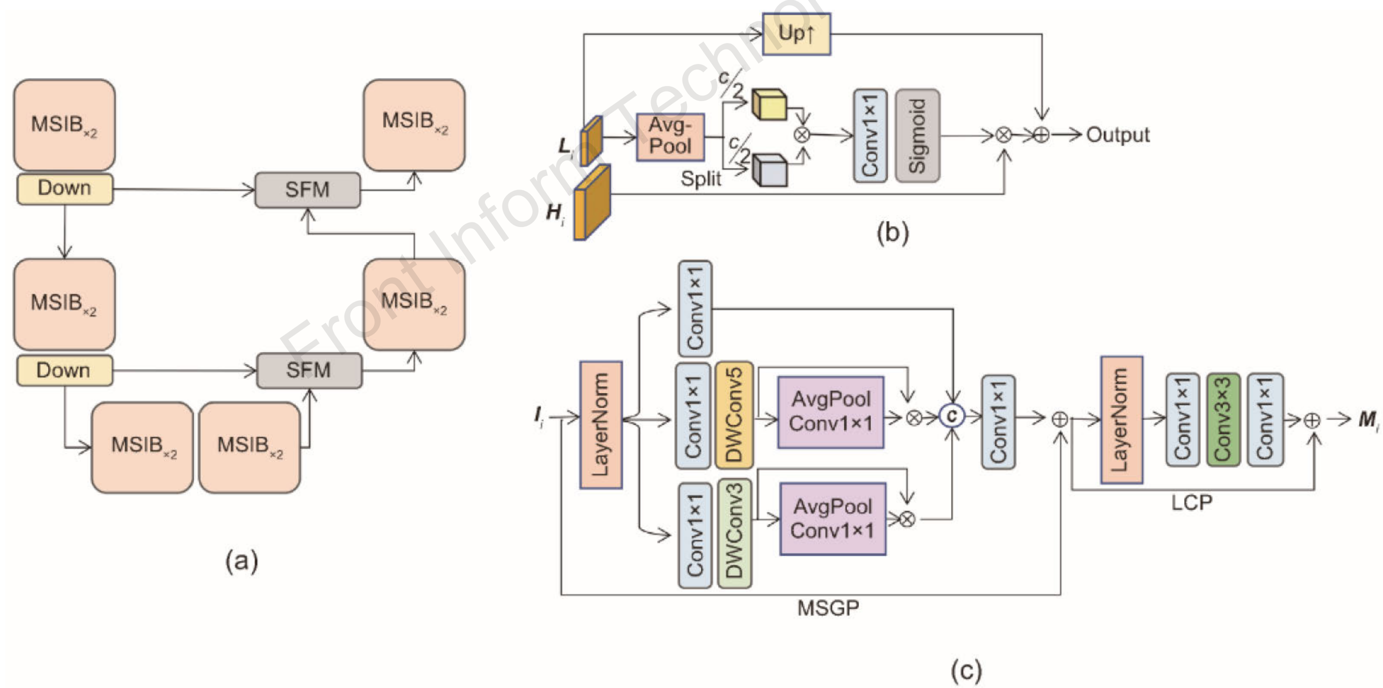
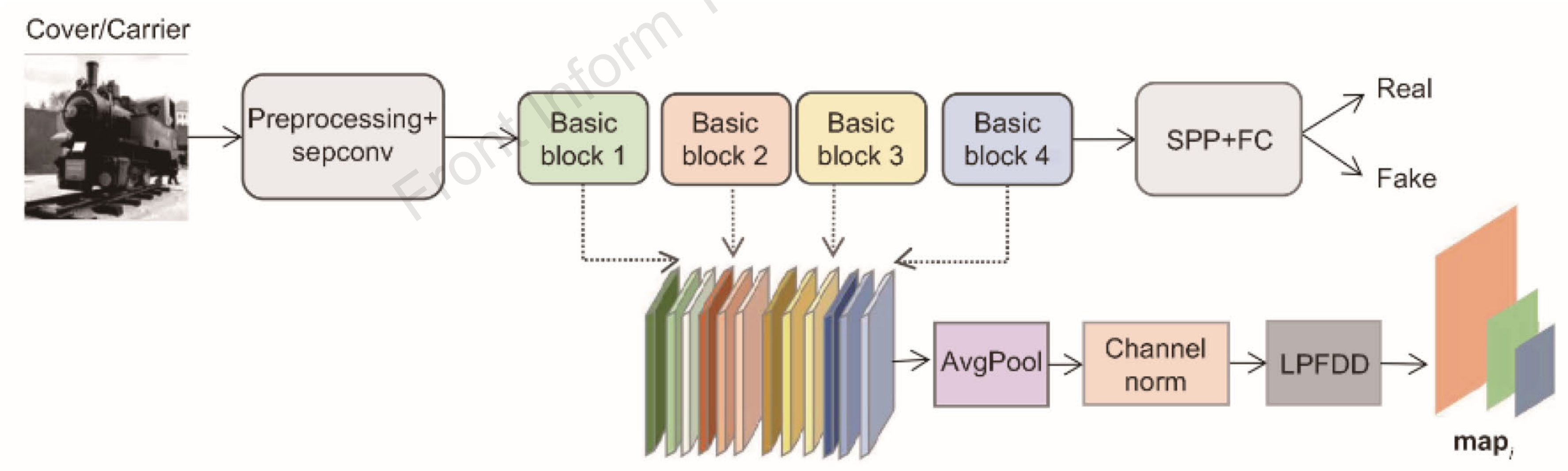


Fig. 3 DMSubNet architecture (a), SFM (b), and MSIB (c)

4) Dual-task discriminator

- Guided embedding via feature score maps: Beyond traditional real/fake classification, the discriminator generates feature score maps that reflect embedding suitability across regions of interest.
- Enhanced security through discriminative feedback: By analyzing distinguishing features between cover and carrier images, the discriminator provides fine-grained feedback across layers.



Main results

- **Steganographic quality**: Our method achieved state-of-the-art performance across all five evaluation metrics during the steganography stage. While PRIS leveraging a reversible architecture and demonstrating superior performance in secret information extraction, it falls short in terms of robustness and security compared to our proposed approach.

Table 2 Image hiding quality results

Method	Cover/Carrier MSE↓	PSNR (dB)↑	SSIM↑	MS-SSIM↑	SCC↑
SimultaneousCNN	0.002 15	25.63	0.9420	0.9774	0.9790
ISGAN	0.011 20	18.95	0.8863	0.9036	0.9637
StegNet	0.004 12	22.38	0.9069	0.9554	0.9082
U-Net structure	0.001 21	27.76	0.9430	0.9805	0.9795
Huang	0.009 74	19.97	0.9023	0.9142	0.9528
HCRGAN	0.009 95	18.29	0.7754	0.8961	0.7470
Encoder–decoder	0.000 51	31.38	0.9099	0.9891	0.8565
Baluja	0.000 35	33.26	0.9338	0.9923	0.9024
SteganoCNN	0.000 32	33.47	0.9568	0.9922	0.9126
Liu	0.000 30	33.89	0.9721	0.9916	0.9534
Improved Xception	0.000 20	35.86	0.9651	0.9959	0.9482
StegGAN	0.000 41	33.19	0.9574	0.9834	0.9154
DBPSNet	0.000 25	34.26	0.9922	0.9921	0.9811
DAH-Net	0.000 62	32.35	0.9920	0.9900	0.9797
PRIS	<u>0.000 10</u>	<u>40.88</u>	<u>0.9982</u>	<u>0.9978</u>	<u>0.9964</u>
Ours	**0.000 06**	**43.55**	**0.9990**	**0.9990**	**0.9979**

Method	Secret/Recovery MSE↓	PSNR (dB)↑	SSIM↑	MS-SSIM↑	SCC↑
SimultaneousCNN	0.002 00	25.33	0.9034	0.9492	0.9632
ISGAN	0.006 21	21.09	0.8887	0.9353	0.9316
StegNet	0.002 93	24.21	0.9134	0.9599	0.9756
U-Net structure	0.001 39	26.69	0.9276	0.9707	0.9699
Huang	0.005 26	21.86	0.9052	0.9391	0.9533
HCRGAN	0.010 33	16.14	0.6933	0.8448	0.7097
Encoder–decoder	0.000 92	28.70	0.8972	0.9747	0.9207
Baluja	0.000 37	31.38	0.9310	0.9798	0.9577
SteganoCNN	0.000 67	30.88	0.9651	0.9878	0.9804
Liu	0.000 32	32.64	0.9398	0.9847	0.9726
Improved Xception	0.001 58	26.42	0.9502	0.9788	0.9828
StegGAN	0.000 43	31.54	0.9293	0.9485	0.9667
DBPSNet	0.000 23	36.31	0.9902	<u>0.9966</u>	<u>0.9957</u>
DAH-Net	0.000 68	32.21	0.9892	0.9817	0.9805
PRIS	**0.000 11**	**40.58**	**0.9981**	**0.9977**	**0.9964**
Ours	<u>0.000 13</u>	<u>39.78</u>	<u>0.9979</u>	0.9956	0.9952

The best results are in bold; the second best results are underlined. ↑ means the larger the value, the better the result; ↓ means the smaller the value, the better the result

Main results (Cont'd)

- **Security**: Steganalysis was conducted using both traditional methods (SRM and CSR) and a classical deep learning-based steganalyzer (XuNet). Our proposed method achieved the best performance, attaining the lowest detection error rates in all cases.

Table 3 Assessment of steganography security

Method	Error detection rate (%)			Method	Error detection rate (%)		
	SRM	CSR	XuNet		SRM	CSR	XuNet
SimultaneousCNN	10.52	14.42	23	SteganoCNN	15.22	16.24	14
ISGAN	2.08	3.97	11	Liu	17.36	18.73	20
StegNet	12.10	14.39	19	Improved Xception	16.40	17.85	16
U-Net structure	12.90	15.03	23	StegGAN	15.97	18.60	16
Huang	4.08	6.49	12	DBPSNet	22.56	27.51	34
HCRGAN	2.18	3.69	4	DAH-Net	15.80	18.98	16
Encoder–decoder	5.26	8.39	14	PRIS	18.80	20.00	17
Baluja	9.06	11.27	17	Ours	**24.38**	**40.60**	**36**

The best results are in bold; the second best results are underlined

Main results (Cont'd)

- **Anti-distortion capability**: Our method exhibits superior robustness against typical image processing attacks, including Gaussian filtering, mean filtering, and sharpening, consistently outperforming other methods in resistance to these distortions.

Method	Mean filtering			Gaussian filtering			Sharpening		
	MSE↓	PSNR (dB)↑	SSIM↑	MSE↓	PSNR (dB)↑	SSIM↑	MSE↓	PSNR (dB)↑	SSIM↑
SimultaneousCNN	0.015 10	16.74	0.4503	0.045 66	13.71	0.3073	0.019 04	16.89	0.6210
ISGAN	0.028 76	13.65	0.2099	0.029 69	12.25	0.3391	0.008 57	19.26	0.8437
StegNet	0.021 09	17.20	0.2474	0.023 08	17.68	0.4810	0.017 00	18.35	0.6220
U-Net structure	0.014 94	16.56	0.4490	0.048 98	13.10	0.3267	0.016 02	17.80	0.6714
Huang	0.024 62	16.03	0.2432	0.039 28	14.76	0.3605	0.007 11	20.60	0.8529
HCRGAN	0.027 07	12.07	0.3507	0.021 95	13.01	0.4799	0.012 95	15.03	0.6092
Encoder–decoder	0.076 54	6.67	0.1190	0.065 29	5.11	0.2060	0.794 31	14.07	0.1600
Baluja	0.062 80	7.47	0.2420	0.057 99	11.25	0.3019	0.602 21	2.32	0.2266
SteganoCNN	0.016 67	18.22	0.4638	0.039 76	14.19	0.3861	0.002 14	26.29	0.8764
Liu	0.016 21	16.29	0.4466	0.038 97	14.69	0.3469	0.009 45	17.55	0.7869
Improved Xception	0.017 21	18.13	0.4153	0.019 63	18.75	0.5387	0.039 98	15.02	0.5305
StegGAN	0.143 70	18.38	0.4874	0.034 15	15.46	0.4956	0.007 92	19.27	0.8004
DBPSNet	0.000 68	31.30	0.9611	0.000 23	36.09	0.9916	0.000 26	35.86	0.9905
DAH-Net	0.001 22	29.72	0.9805	0.000 88	31.13	0.9864	0.000 45	34.05	0.9931
PRIS	0.000 61	32.44	0.9887	0.000 91	31.27	0.9855	0.000 16	39.07	0.9975
Ours	**0.000 13**	**36.62**	**0.9957**	**0.000 15**	**39.23**	**0.9976**	**0.000 15**	**39.29**	**0.9976**

Main results (Cont'd)

- **Anti-distortion capability**: In experiments involving Gaussian noise and image flipping, our method maintains superior robustness compared to existing methods.

Method	Gaussian noise			Flipping		
	MSE↓	PSNR (dB)↑	SSIM↑	MSE↓	PSNR (dB)↑	SSIM↑
SimultaneousCNN	0.007 72	20.44	0.5404	0.045 95	12.65	0.3541
ISGAN	0.053 60	11.81	0.1260	0.056 60	11.38	0.3282
StegNet	0.031 61	13.66	0.1994	0.050 03	12.75	0.2778
U-Net structure	0.034 53	14.29	0.1944	0.044 17	12.68	0.3418
Huang	0.046 25	12.94	0.1374	0.054 00	11.96	0.3096
HCRGAN	0.032 30	10.98	0.1866	0.047 89	9.84	0.3068
Encoder–decoder	0.017 56	17.15	0.3316	0.050 79	14.92	0.2109
Baluja	0.041 24	13.44	0.2256	0.047 58	12.48	0.2006
SteganoCNN	0.048 70	12.41	0.1522	0.044 67	12.72	0.3264
Liu	0.008 13	18.69	0.4536	0.046 83	12.65	0.3217
Improved Xception	0.035 32	14.45	0.1836	0.068 91	11.82	0.1982
StegGAN	0.007 78	19.27	0.4818	0.042 88	13.12	0.3787
DBPSNet	0.000 27	34.76	0.9870	0.003 91	13.80	0.5623
DAH-Net	0.001 14	29.80	0.9817	0.003 59	25.01	0.9420
PRIS	0.000 48	33.49	0.9921	0.049 56	13.29	0.1119
Ours	**0.000 23**	**36.77**	**0.9964**	**0.002 45**	**26.95**	**0.9635**

Main results (Cont'd)

- **Capacity**: We compare SRIS-Net with StegGAN by conducting experiments to embed 2, 3, and 4 images in one cover image, in order to verify the superiority of our method in terms of embedding capacity.

Table 6 Results for capacity of StegGAN and SRIS-Net

Number of images embedded	PSNR (dB)			
	Cover/Carrier		Secret/Recovery	
	StegGAN	Ours	StegGAN	Ours
2	33.21	41.96	21.41	34.65
3	31.35	41.56	18.56	32.33
4	29.87	41.46	16.18	30.85

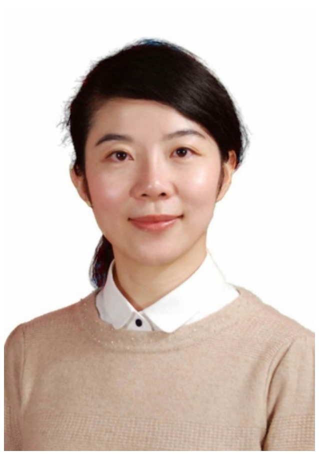
Conclusions

- We propose SRIS-Net, a robust image steganography algorithm based on feature score maps, which fuses spatial- and frequency-domain features and adopts a progressive assisted hiding strategy.

- The use of GLEM and DMSubNet enables multi-scale feature reconstruction and progressive embedding, reducing visual distortion and improving the quality of the image.

- A dual-task discriminator not only distinguishes real/fake images but also generates ROI feature score maps, guiding the embedding process for better imperceptibility and security.

Ai XIAO received her B.S. degree in software engineering from Chongqing University in 2016, and is currently pursuing an M.S. degree in computer technology at Guizhou University. Her current research interests include information security and image steganography, deep learning, and computer vision.



Zhi LI received her Ph.D. degree in computer science and software from Guizhou University in 2009. She is a Professor with the School of Computer Science and Technology, Guizhou University, Guiyang, China. Her research interests include medical imaging, image processing, pattern recognition, and machine learning.

Guomei WANG received her B.S. and M.S. degrees in computer science from Northwest A&F University in 1998 and 2001, respectively. She is currently an Associate Professor with the School of Computer Science and Technology, Guizhou University, Guiyang, China. Her research interests include computer vision, image steganography, and adversarial attacks and defenses.

Long ZHENG received his B.S. degree in software engineering in 2020, and is currently pursuing the Ph.D. degree in software engineering at Guizhou University. His current research interests include machine learning, deep learning, and computer vision.

Haoyuan SUN received his B.S. degree in computer science and technology (Innovation Class) from Guizhou University in 2025. He is currently an undergraduate student at the School of Computer Science and Technology, Guizhou University. His research interests include deep learning, video segmentation, and medical image analysis.