

# Supplementary Materials: Achieving Resolution-Agnostic DNN-based Image Watermarking: A Novel Perspective of Implicit Neural Representation

Anonymous Authors

In Section 1, we provide detailed information about Stage 1 and Stage 2 that was not mentioned in the experiment section. In Section 2, we provide some additional experimental results and comparisons based on the settings in our experiment section.

## 1 IMPLEMENTATION DETAILS

**Stage 1.** We create INR by splitting the RGB channels into three parts and fitting them separately. We use six hidden layers, each containing 256 perceptrons. We set the PSNR to stop training at 45dB to ensure that the fine-tuning stage does not affect the invisibility of INR [1, 5, 6].

**Stage 2.** We utilize the Conv-BN-ReLU blocks for both encoder and decoder. The encoder consists of 4 blocks, with 64 output channels, 2D kernel size 3, stride 1, and padding 1. The decoder has seven blocks with 64 output channels, followed by a block with  $n$  output channels, where  $n$  is the length of the secret message. Then, the message is recovered after applying an average pool and a linear layer. The attacks for pre-training watermark decoder are crop( $s = 0.25$ ), JPEG( $Q = 50$ ), Resize( $p = 0.5$ ), Resize( $p = 2$ ) and MF( $k_s = 7$ ). The non-differentiable noise JPEG is simulated by following [7].

## 2 MORE EXPERIMENTAL RESULTS

### 2.1 Evaluation on Specific Attack

In the paper, we mentioned that other baseline methods have poor robustness in at least one type of attack. Here, we compare the robustness of watermarking methods under different resolutions from the perspectives of two attacks, JPEG compression and crop.

As shown in Table 1 and 2, MBRS[3] can only watermark images with training resolution, thus MBRS is not tested under other resolutions. Under JPEG attack, the accuracies of HiDDeN [8], and TSDL [4] are around 60% and 55%, showing that they are not robust against JPEG compression. Under cropping attack, the accuracy of DWSF [2] is lower than 65% because the watermarked block is attacked, resulting in poor robustness.

### 2.2 Evaluation on Varied Resolutions

In this section, we show the watermarked images which are sampled from the watermarked INR on six different resolutions:  $256 \times 256$ ,  $384 \times 384$ ,  $512 \times 512$ ,  $480 \times 854$  (480p),  $720 \times 1280$  (720p) and  $1080 \times 1920$  (1080p). As shown in Figure 1, we resize the six watermarked images to the same height. We choose three different images to demonstrate the quality of our watermark images. The watermarks are distributed in the image’s high-frequency areas, allowing for good invisibility.

**Table 1: Accuracy (%) under JPEG attack. “/” denotes that MBRS is not applicable when the resolution varies.**

Resolution	HiDDeN	TSDL	MBRS	DWSF	RAIMARK
$256 \times 256$	60.77	54.30	97.43	95.83	<b>99.97</b>
$384 \times 384$	59.23	56.10	/	93.93	<b>99.97</b>
$512 \times 512$	57.97	55.13	/	93.93	<b>99.97</b>
$480 \times 854$	57.80	54.43	/	92.00	<b>99.93</b>
$720 \times 1280$	55.97	53.93	/	95.43	<b>100.00</b>
$1080 \times 1920$	57.13	54.03	/	95.83	<b>99.97</b>

**Table 2: Accuracy (%) under cropping attack. “/” denotes that MBRS is not applicable when the resolution varies.**

Resolution	HiDDeN	TSDL	MBRS	DWSF	RAIMARK
$256 \times 256$	98.53	86.73	58.80	51.33	<b>99.20</b>
$384 \times 384$	98.80	89.10	/	50.53	<b>99.33</b>
$512 \times 512$	99.23	87.73	/	51.20	<b>99.43</b>
$480 \times 854$	99.17	86.70	/	50.77	<b>99.80</b>
$720 \times 1280$	98.97	87.13	/	62.80	<b>99.73</b>
$1080 \times 1920$	98.97	84.83	/	61.77	<b>99.50</b>

### 2.3 Evaluation on Different Methods

In this section, we show the original images and the watermarked images generated by HiDDeN, TSDL, MBRS, DWSF, and our proposed method RAIMARK under  $256 \times 256$  resolution, which is shown in Figure 2. In MBRS and DWSF, apparent artifacts and densely distributed watermark regions can be observed, resulting in poor invisibility. Although HiDDeN and TSDL evenly distribute the watermark in the image, the high watermark intensity causes some color deviation from the original image in some areas. Our proposed method RAIMARK achieves the best invisibility by embedding the watermark in areas where pixel value changes significantly.

## REFERENCES

- [1] Malleshm Dasari, Arani Bhattacharya, Santiago Vargas, Pranjal Sahu, Aruna Balasubramanian, and Samir R Das. 2020. Streaming 360-degree videos using super-resolution. In *IEEE INFOCOM 2020-IEEE Conference on Computer Communications*. 1977–1986.
- [2] Hengchang Guo, Qilong Zhang, Junwei Luo, Feng Guo, Wenbin Zhang, Xiaodong Su, and Minglei Li. 2023. Practical Deep Dispersed Watermarking with Synchronization and Fusion. In *Proceedings of the ACM International Conference on Multimedia*. 7922–7932.
- [3] Zhaoyang Jia, Han Fang, and Weiming Zhang. 2021. MBRS: Enhancing Robustness of DNN-based Watermarking by Mini-Batch of Real and Simulated JPEG Compression. In *Proceedings of the ACM International Conference on Multimedia*. 41–49.
- [4] Yang Liu, Mengxi Guo, Jian Zhang, Yuesheng Zhu, and Xiaodong Xie. 2019. A Novel Two-stage Separable Deep Learning Framework for Practical Blind Watermarking. In *Proceedings of the ACM International Conference on Multimedia*.



Figure 1: Watermarked images with different resolutions. The first row: the watermarked images  $I_w$ . The second row: the residual images  $I_r$ . The third row: the normalized residual images  $I_m$ .

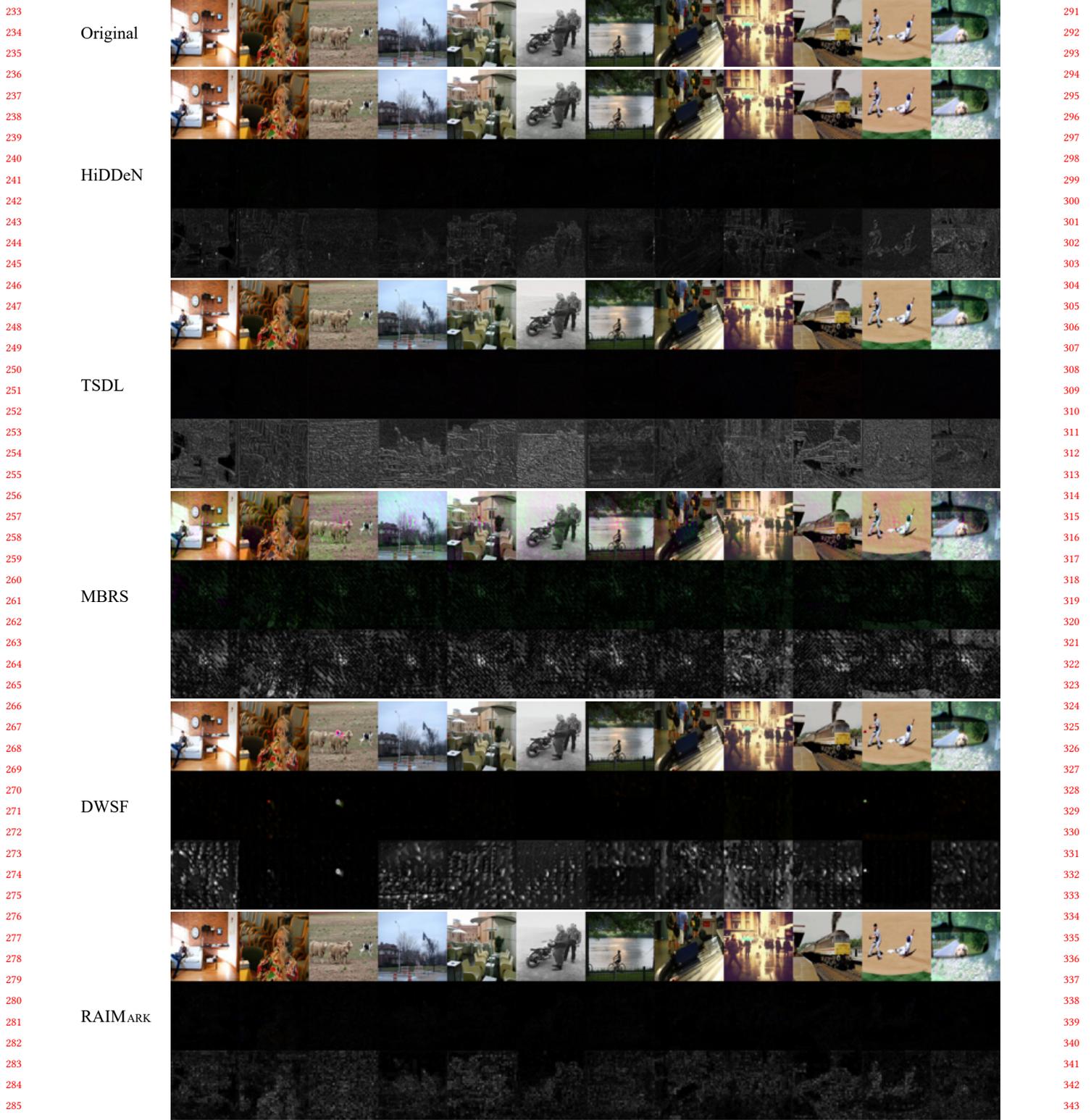


Figure 2: Watermarked images by different methods. The original images are shown on the top. The first row of each method: the watermarked images  $I_w$ . The second row of each method: the residual images  $I_r$ . The third row of each method: the normalized residual images  $I_m$ .

349	1509–1517.		
350	[5] Nikolaos Thomos, Nikolaos V Boulgouris, and Michael G Strintzis. 2005. Opti-	[7] Chaoning Zhang, Adil Karjauv, Philipp Benz, and In So Kweon. 2021. Towards	407
351	mized transmission of JPEG2000 streams over wireless channels. <i>IEEE Transac-</i>	Robust Deep Hiding Under Non-Differentiable Distortions for Practical Blind Wa-	408
352	<i>tions on image processing</i> 15, 1 (2005), 54–67.	termarking. In <i>ACM International Conference on Multimedia</i> . 5158–5166.	409
353	[6] Anlan Zhang, Chendong Wang, Bo Han, and Feng Qian. 2022. YuZu:Neural-	[8] Jiren Zhu, Russell Kaplan, Justin Johnson, and Li Fei-Fei. 2018. HiDDeN: Hiding	410
354	Enhanced volumetric video streaming. In <i>19th USENIX Symposium on Networked</i>	Data With Deep Networks. In <i>Proceedings of the European Conference on Computer</i>	411
355	<i>Systems Design and Implementation (NSDI 22)</i> . 137–154.	<i>Vision (ECCV)</i> . 682–697.	412
356			413
357			414
358			415
359			416
360			417
361			418
362			419
363			420
364			421
365			422
366			423
367			424
368			425
369			426
370			427
371			428
372			429
373			430
374			431
375			432
376			433
377			434
378			435
379			436
380			437
381			438
382			439
383			440
384			441
385			442
386			443
387			444
388			445
389			446
390			447
391			448
392			449
393			450
394			451
395			452
396			453
397			454
398			455
399			456
400			457
401			458
402			459
403			460
404			461
405			462
406			463
			464