

Supplementary Materials: HideMIA: Hidden Wavelet Mining for Privacy-Enhancing Medical Image Analysis

Anonymous Authors

1 MORE RELATED WORKS

In the field of medical image privacy protection, federated learning (FL) [2, 11] has been extensively studied, allowing different data centers to securely share training data while ensuring that no participating center can access sensitive data from others [8, 9]. Federated learning, through distributed computing and strategies to maintain data locality, prompts collaboration in medical image analysis and reduces the risk of data breaches [5, 6, 10].

Another type of privacy protection method, unlearnable examples (UEs), as a specific type of data poisoning attacks [1], are designed to protect the datasets from unauthorized model training by injecting protective perturbation to images from the training dataset [3]. Models trained on datasets protected by UEs tend to perform random guessing on clean test images. Lin et al. [7] propose medical-prior-aware perturbation generation methods to effectively and imperceptibly protect training datasets.

As we conclude in Table 1, federated learning and unlearnable examples are designed for the protection of training data from being seen and being used for unauthorized model training, respectively. These two techniques cannot be adopted to protect test images after the model deployment. Unlike FL and UEs, deformation-based, homomorphic encryption-based, encoding-based, and our image hiding-based methods can be used for safeguarding test images against attacks. Different from existing privacy protection methods for client-server MIA, we conceal medical images within natural images and directly perform MIA in the steganographic domain of the concealed images. This is less likely to attract the attention of attackers.

2 MINOR LIMITATIONS

In our main manuscript, we discuss the necessity of further improving imperceptibility and MIA performance. Here, we present some minor issues that we can deal with in the feature.

(1) *Multi-class segmentation.* Although our method supports multi-classification, it is currently limited to supporting only three-channel image hiding and extraction. This limitation stems from our framework’s design, which only allows the hiding and recovery of three-channel images.

(2) *Robustness against steganalysis.* The ability to resist steganalysis is another concern. Existing DIH (Deep Image Hiding) methods are somewhat vulnerable to steganalysis. If an insider were to leak our pipeline, it would be challenging to defend against attackers using steganography to extract patient images. We can use some encryption to enhance our DIH process.

3 MORE IMPLEMENTATION DETAILS

Competing Methods. For competing medical image segmentation methods, we either adopt the official hyperparameters or assign a set of better ones. For deep image hiding (DIH) models, *i.e.*, HiNet, HiDDeN, and DeepStega, we utilize their pretrained weights to start

the training. For all compared covert MIA methods, the training methodology is similar to HideMIA: first, the DIH network is frozen, and the MIA network is trained using the total loss \mathcal{L}_{total} from Eq. (8), followed by joint fine-tuning of the DIH and MIA networks using \mathcal{L}_{total} . Note that the original version of HiDDeN [12] is not designed for image hiding but rather for concealing binary information within images. We employ a reimplemented version¹ of HiDDeN from PUSNet [4], used for image hiding. Since the official Github repository of DeepStega lacks pretrained weights, we utilize weights reproduced by others², which exhibit similar performance to the officially claimed metrics.

HideMIA. In AsyWA, due to limited computational resources, we downsample the size of the feature maps to 1/2 using adaptive average pooling and reduce the number of channels by half using 1×1 convolution before inputting them into AsyWA. Subsequently, we apply interpolation and 1×1 convolution to upsample the features outputted by AsyWA to match the size of the input features.

REFERENCES

- [1] Marco Barreno, Blaine Nelson, Anthony D Joseph, and J Doug Tygar. 2010. The security of machine learning. *Machine Learning* (2010).
- [2] Hao Guan, Pew-Thian Yap, Andrea Bozoki, and Mingxia Liu. 2024. Federated learning for medical image analysis: A survey. *Pattern Recognition* (2024), 110424.
- [3] Hanxun Huang, Xingjun Ma, Sarah Monazam Erfani, James Bailey, and Yisen Wang. 2021. Unlearnable Examples: Making Personal Data Unexploitable. In *Proceedings of the International Conference on Learning Representations*.
- [4] Guobiao Li, Sheng Li, Zicong Luo, Zhenxing Qian, and Xinpeng Zhang. 2024. Purified and Unified Steganographic Network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*.
- [5] Qiushi Li, Wenwu Zhu, Chao Wu, Xinglin Pan, Fan Yang, Yuezhi Zhou, and Yaoxue Zhang. 2020. InvisibleFL: federated learning over non-informative intermediate updates against multimedia privacy leakages. In *ACM MM*. 753–762.
- [6] Xiaoxiao Li, Yufeng Gu, Nicha Dvornek, Lawrence H Staib, Pamela Ventola, and James S Duncan. 2020. Multi-site fMRI analysis using privacy-preserving federated learning and domain adaptation: ABIDE results. *Medical Image Analysis* 65 (2020), 101765.
- [7] Xun Lin, Yi Yu, Song Xia, Jue Jiang, Haoran Wang, Zitong Yu, Yizhong Liu, Ying Fu, Shuai Wang, Wenzhong Tang, and Alex Kot. 2024. Safeguarding Medical Image Segmentation Datasets against Unauthorized Training via Contour- and Texture-Aware Perturbations. *arXiv* 2403.14250 (2024).
- [8] Ming Y Lu, Richard J Chen, Dehan Kong, Jana Lipkova, Rajendra Singh, Drew FK Williamson, Tiffany Y Chen, and Faisal Mahmood. 2022. Federated learning for computational pathology on gigapixel whole slide images. *Medical Image Analysis* 76 (2022), 102298.
- [9] Adnan Qayyum, Kashif Ahmad, Muhammad Ahtazaz Ahsan, Ala Al-Fuqaha, and Junaid Qadir. 2022. Collaborative federated learning for healthcare: Multi-modal covid-19 diagnosis at the edge. *IEEE Open Journal of the Computer Society* 3 (2022), 172–184.
- [10] Micah J Sheller, Brandon Edwards, G Anthony Reina, Jason Martin, Sarthak Pati, Aikaterini Kotrotsou, Mikhail Milchenko, Weilin Xu, Daniel Marcus, Rivka R Colen, et al. 2020. Federated learning in medicine: facilitating multi-institutional collaborations without sharing patient data. *Scientific reports* 10, 1 (2020), 12598.
- [11] Liping Yi, Gang Wang, Xiaoguang Liu, Zhuan Shi, and Han Yu. 2023. Fedgh: Heterogeneous federated learning with generalized global header. In *ACM MM*. 8686–8696.
- [12] Jiren Zhu, Russell Kaplan, Justin Johnson, and Li Fei-Fei. 2018. HiDDeN: Hiding Data With Deep Networks. In *Proceedings of the European Conference on Computer Vision*, Vol. 11219. 682–697.

¹<https://github.com/albblgb/Hiding-images-within-images>

²<https://github.com/arnoweng/PyTorch-Deep-Image-Steganography/tree/master>

Table 1: Characteristics for privacy-enhancing techniques for MIA. “FL” and “UEs” are short for federated learning and unlearnable examples. Others are privacy-enhancing server-client frameworks for MIA.

Characteristics		FL	UEs	Deformation-based	HE-based	Encoding-based	Our HideMIA
Protected subjects	Data sharing during training	✓	×	×	×	×	×
	Unauthorized model training	×	✓	×	×	×	×
	Testing data after Deployment	×	×	✓	✓	✓	✓
Exposing transmitted images as medical images		-	-	✓	×	×	×
Degree of normal distribution modification		-	-	Medium	High	High	Low
Time complexity		-	-	Low	High	Low	Low
Difficulty in arousing attackers’ suspicion		-	-	Medium	High	High	Low

Table 2: Details of each MIA dataset.

Name	Modality	Subject	Type	Number of Images	Number of Classes
BUSI	Ultrasound	Breast	Segmentation	612	2
Kvasir-SEG	Endoscope	Intestine	Segmentation	1,000	2
ChildDental	X-ray	Teeth	Segmentation	1,510	2
SIPaKMed	Microscope	Cell	Classification	4,049	5
DermaMNIST	Dermatoscope	Skin	Classification	10,015	7
ChestCT	CT	Lung	Classification	1,000	4

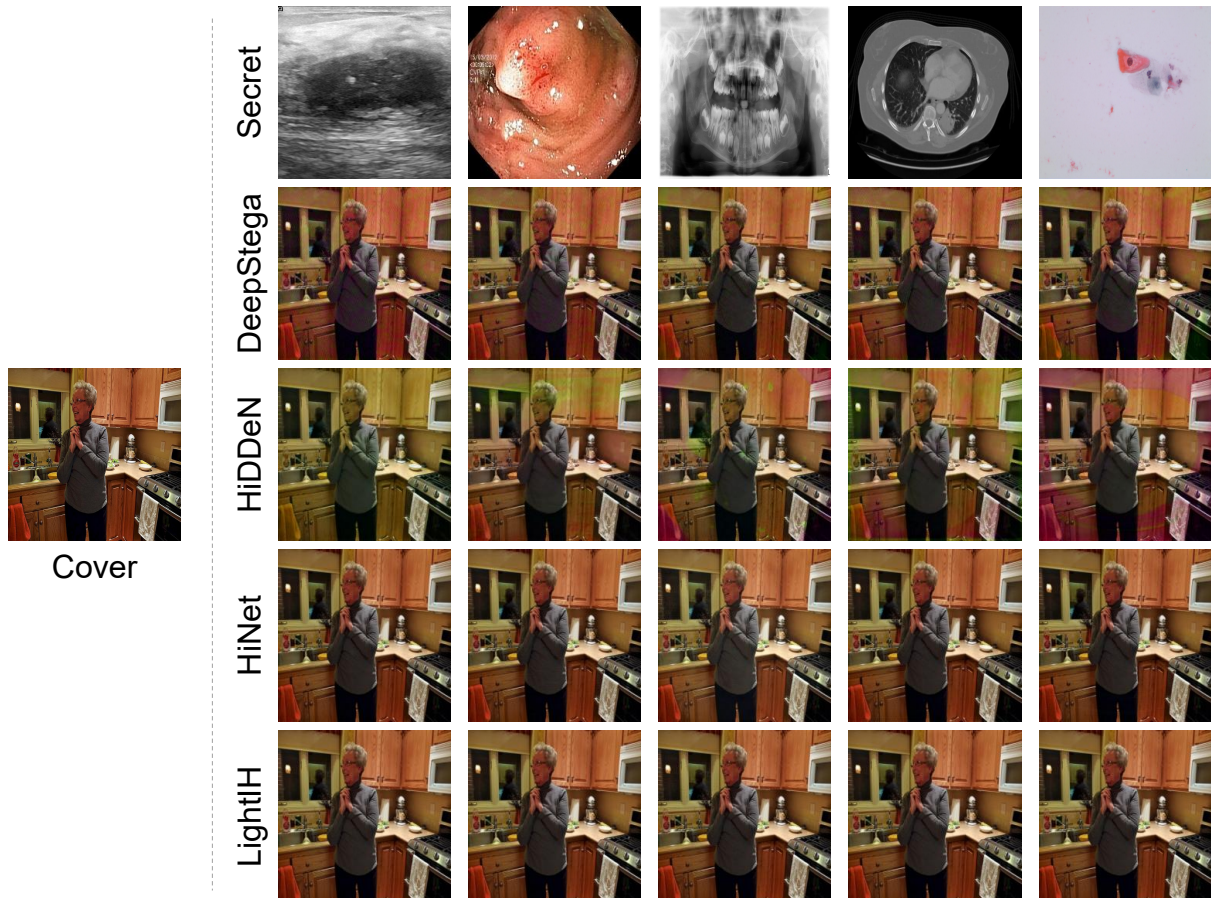


Figure 1: Visual comparisons for different DIH networks (x_{stega})

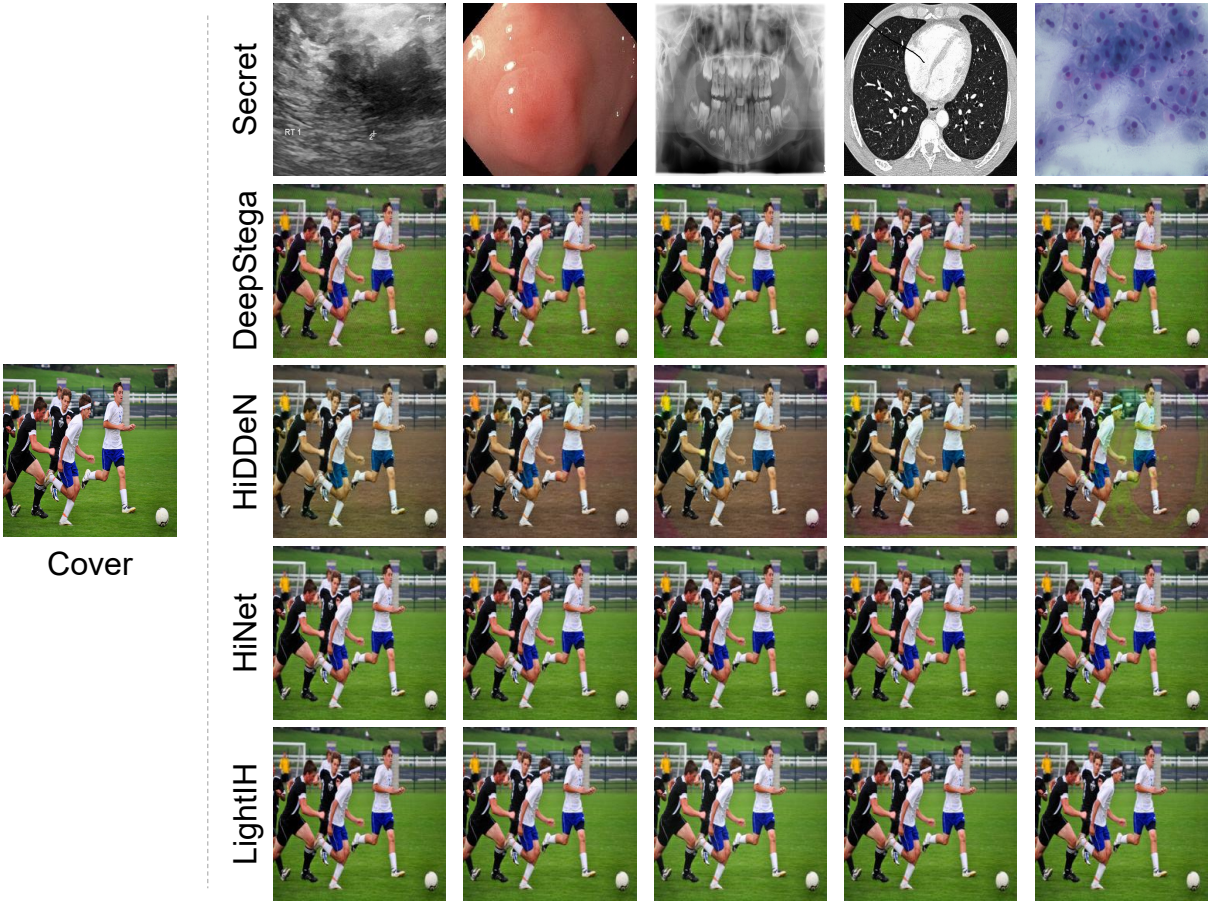


Figure 2: Visual comparisons for different DIH networks (x_{stega})