

Foveation Improves Payload Capacity in Steganography

Lifeng Qiu Lin¹, Henry Kam,² Qi Sun,² Kaan Akşit¹

¹University College London ²New York University



Problem

Steganography lets images convey messages beyond their visible content. Applications such as provenance watermarking and scene labelling require larger payload capacity to embed sufficiently long messages.

Inspired by the human visual system, we ask: **can perception be leveraged to expand steganographic payload?**

Related Work

Recent methods use deep learning for improved quality. The most relevant open-source work in this direction is perha RoSteALS, which can embed 100 bits.

Foveation rendering is a promising proxy for human gaze. Because it focuses on the foveated area and is more tolerant in the periphery, it could allow higher payload capacity.

Our Solution

We leverage the trained latent space of autoencoders to accelerate learning of the embedding process. A CNN is sufficient to identify embedded bits. Finally, foveation is incorporated into the objective via a Metameric Loss.

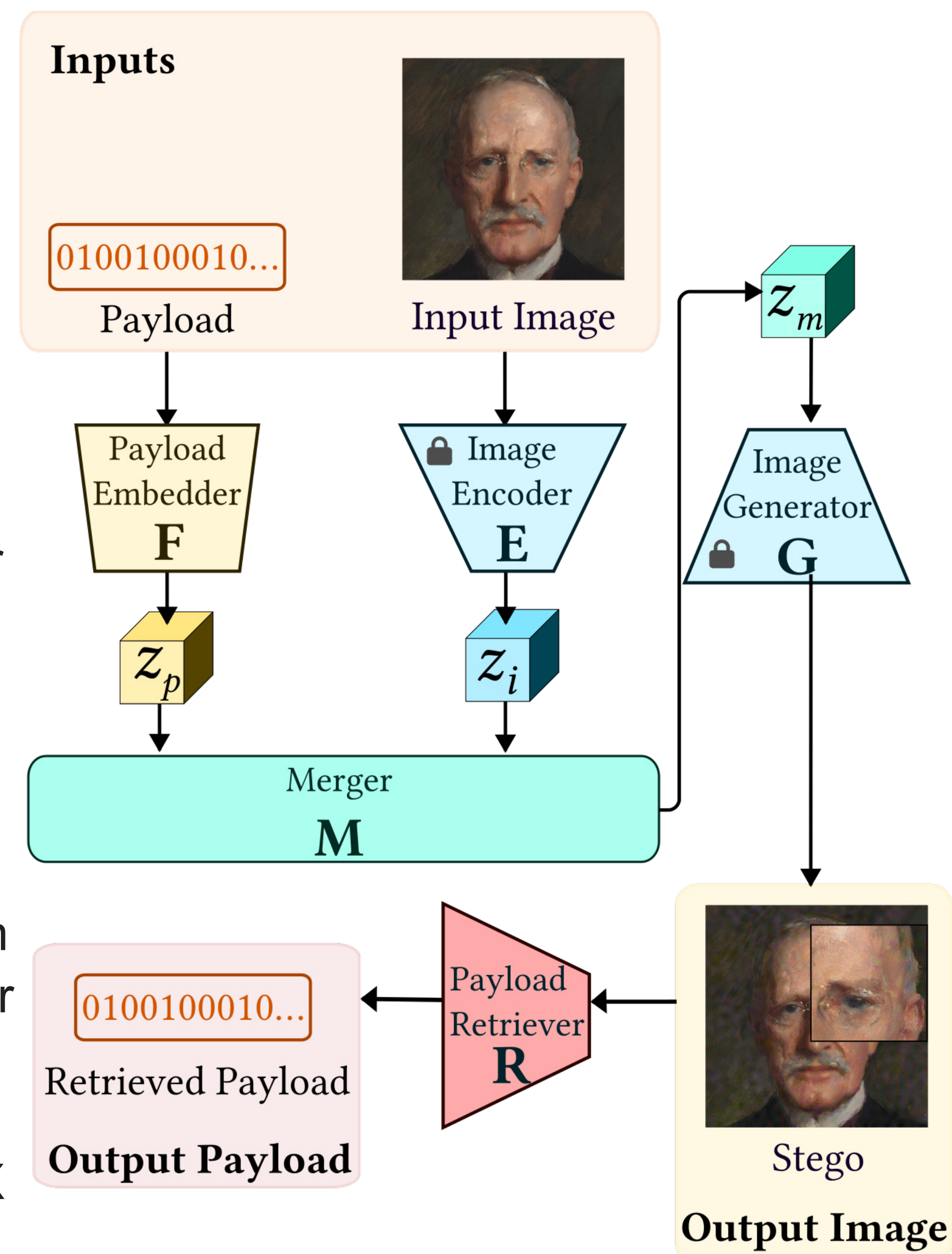
References

- Tu Bui, Shruti Agarwal, Ning Yu, and John Collomosse. 2023. Rosteals: Robust steganography using autoencoder latent space. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 933–942.
- David R Walton, Koray Kavakli, Rafael Kuffner Dos Anjos, David Swapp, Tim Weyrich, Hakan Urey, Anthony Steed, Tobias Ritschel, and Kaan Akşit. 2022. Metamericvarifocal holograms. In 2022 IEEE Conference on Virtual Reality and 3D User Interfaces (VR). IEEE, 746–755.

Method

The full pipeline is shown in the right-hand figure. Leveraging latent representations from the LDM–VQGAN family, the original image and message are encoded, merged, and used to generate the stego image. A trained payload retriever then recovers the message from the stego image.

These models are trained on only 2,000 images from MetFaces and CLIC, fewer than most methods, and training converges within two hours on a single RTX 4090 GPU in the minimal setup.



Results

We first assessed a baseline built from several preliminary experiments. Compared to RoSteALS, the baseline increases payload capacity and accuracy but at a substantial expense to image quality.

Next, we replaced the image-loss term (MSE) with a Metameric Loss. This change increased payload capacity from 100 bits to 200 bits while achieving visual fidelity comparable to—or better than—RoSteALS, and maintaining high bit accuracy. This demonstrates the potential of perceptual rendering for next-generation steganography..

Experiment	Resolution	Capacity	Bit Accuracy	MSE	PSNR	SSIM	LPIPS	Metameric Loss
Benchmark Baseline (RoSteALS)	256	100	0.9942	0.0009	32.16	0.8971	0.0780	0.0016
Vanilla Baseline	256	100	1	0.0015	28.90	0.8833	0.1621	0.0072
Vanilla Baseline	256	200	1	0.0018	28.37	0.8681	0.2047	0.0072
Metameric Baseline	256	200	0.9998	0.0010	31.47	0.8871	0.1288	0.0017
Vanilla Baseline	256	500	0.9997	0.0021	27.56	0.8348	0.2613	0.0077
—	—	—	—	—	—	—	—	—
Vanilla Baseline	128	500	0.4997	0.0007	33.11	0.9061	0.0570	0.0023

