StegNet

Anonymous CVPR submission

Paper ID 6.689

Abstract

Need abstract

1. Introduction

Hidden or encrypted messages have a storied past in the history of the human kind. The first known example of cryptography dates back to 1900 BC, where Egyptians used ciphertext with hieroglyphs [9]. However, if one were to intercept an encrypted message from Julius Caesar to his generals, the adversary would know there is a hidden message. Using the simple substitution cipher, the message contains word like structures eliciting decryption attempts. Ideally, an interceptor would be deceived to think there is no message at all. Thus, steganography is born.

Steganography has an equally illustrious past. The oldest fable is one told by the ancient Greek historian Herodotus. He chronicled the tale of a man who killed a hare and sent it along with a hunter to be delivered to a third party. Both the messenger and message container aid in hiding the transmission. Steganography naturally extends to images. The first use was in World War II when the Germans reduced images to the size of a period, which would be inserted on the inside of envelope flaps. Ironically, this images images where photographs of written text [7].

With thousands of pixels spatially arranged in an image, one could imagine the vast ways of encoding information. Images that conceal a hidden message are referred to as a stego-image. Ironically, the amount of information in an image is both a feature and a bug. On the other end of the double-edged sword, pictures contain a vast amount of information constructed in a specific manner. Consequently, images are particularly difficult to conceal in a container. The solution: hide an image in an image.

We have two given images: a mask and a secret image. Each contains hundreds of pixels with multiple channels making for thousands of inputs. We also have an objective to minimize perceptual recognition of the secret image within the mask. The task of finding the perfect combination of thousands of variables to satisfy a given task is a quintessential task for neural networks.

outline paper and describe novelty

2. Related Work

Goodfellow et al. [6] proposed the idea of a Generative Adversarial Network (GAN) in 2014. Since then GAN's have regularly been employed in the context of steganography.

One of the first applications of this was in "Generative Adversarial Networks for Image Steganography", Volkhonskiy et al. [4] which first proposed the idea of employing a cover image to hide the secret image during transit. These methods were improved upon by Zhang et al. [10] in 2018 with "Invisible steganography via generative adversarial networks". This paper aimed to improve both the 'invisibility' of the secret image and the improve on the quality of the retrieved decoded image.

Although the methods presented above deliver excellent results, they are limited by the restriction to greyscale secret images. In real world applications, encrypting RGB images may often be necessary, for instance in medical heatmaps, where key information is encoded with color.

These limitations are addressed in Baluja's [3] take on the encoder-decoder network, which allows for a three channel, color image to be hidden within another, equal sized color image.

Additionally, in some cases it could be beneficial to encode multiple secret images with the same cover image. Such a system was proposed by Das et al [1]. in 2021. This method utilizes multiple decoders to retrieve the secret images.

Also of note are works such as "Hide and speak: Deep neural networks for speech steganography", Kreuk et al [8]. which examines the analogous problem of steganography in audio. Allowing for the encoding of multiple audio signals onto a 'carrier' signal.

Space permitting, could be worth expanding on some of methodological differences in each of these. describing the architechtural differences between them and the specific improvements in each.

3. Approach

As previously mentioned, hiding images is difficult because the vast information an image contains. A natural first step is to reduce the dimensionality while retaining the most critical information. As such, our network structure parallels an autoencoder [2]. We utilize the three-component network model introduced in [3]. We use a preparation network, a hiding network, and a reveal network all trained in concert.

The preparation network performs the rudimentary task of scaling the secret image to the size of the cover image. Additionally, it converts the image into three useful channels: an edge detector, FIND OUT OTHER TWO CHANNELS AND INSERT IMAGE. The preparation network outputs a concatenation of the cover images' RGB channels as well as the three manufactured channels.

finish sections describing networks

Unique to the prep-hide-reveal neural network is its idiosyncratic loss function across components. We begin our exploration with the loss function defined in Baluja (2017) shown in Equation 1:

$$\mathcal{L}(c, e, s, d) = \|c - e\| + \beta\|s - d\| \tag{1}$$

where c is the cover image, s is the secret image, e is the encrypted message, and d is the decrypted message. Since both c and e are an input and the output of the preparation and hiding network,the first term of the loss function, $||c - e||$ only affects those networks. Conversely, s and d are an input and the output of the entire network and thus $||s - d||$ influence the entire network. Note that we have a dual objective: minimize the difference between the mask and the encrypted image, and minimize the difference between the secret and decoded image. This naturally necessitates a weighting parameter, β. We explore varying parameter values in Section 4.

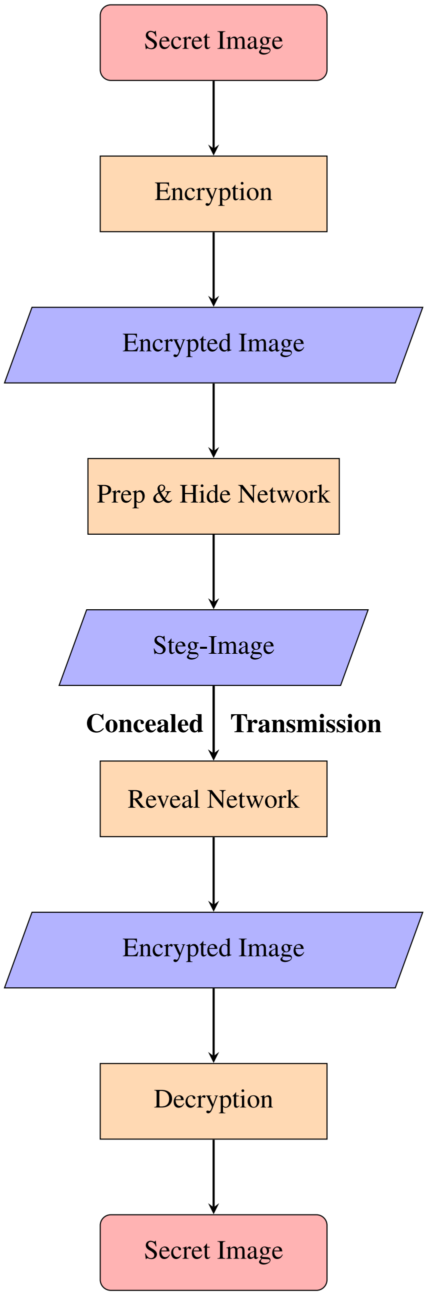
A limitation of Equation 1, is its unadorned mechanism for comparing images. The equation simply tries to limit individual pixel differences between individual two images. Thus, all pixels are equally weighted and mutually exclusive. One could imagine the problems that arise when pixel correlation and content structure are ignored. We

peak-signal-to-noise ratio (PSNR) as well as the structural similarity index measure (SSIM) [5]

elaborate on metrics below that work after training: https://up42.com/blog/tech/image-similarity-measures

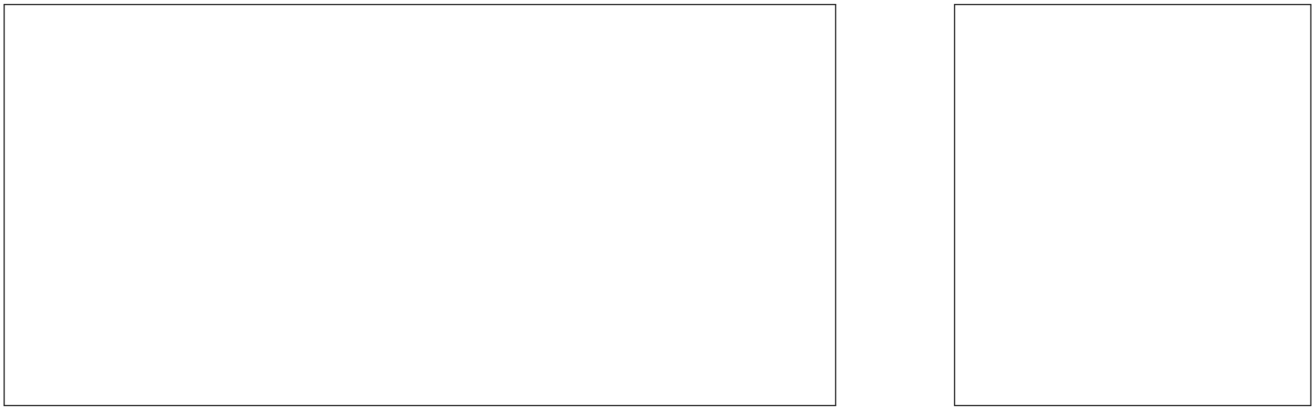
Additionally, we explore the idea of combining peripheral cryptology with deep steganography. Recall the purpose of steganopgraphy is to hide a message in plain sight, *in transit*. We postulate one is able to encrypt an image, generate a steg-image for transit, and on the back-end reveal and subsequently decrypt into the secret image. This multi-step process is illustrated in Figure 1.

Figure 1. Steganography with Peripheral Encryption



4. Experimental Results

5. Conclusion



(a) An example of a subfigure.

(b) Another example of a subfigure.

Figure 2. Example of a short caption, which should be centered.

References

[1] M. Anand Y. Rana A. Das, J. Singh Wahi. Multi-image steganography using deep neural networks. 2021. 1

[2] Pierre Baldi. Autoencoders, unsupervised learning, and deep architectures. In Isabelle Guyon, Gideon Dror, Vincent Lemaire, Graham Taylor, and Daniel Silver, editors, *Proceedings of ICML Workshop on Unsupervised and Transfer Learning*, volume 27 of *Proceedings of Machine Learning Research*, pages 37–49, Bellevue, Washington, USA, 02 Jul 2012. PMLR. 2

[3] Shumeet Baluja. Hiding images in plain sight: Deep steganography. *Advances in neural information processing systems*, 30, 2017. 1, 2

[4] E. Burnaev D. Volkhonskiy, B. Borisenko. Generative adversarial networks for image steganography. 2016. 1

[5] Alain Hore and Djemel Ziou. Image quality metrics: Psnr vs. ssim. In *2010 20th international conference on pattern recognition*, pages 2366–2369. IEEE, 2010. 2

[6] Mehdi Mirza Bing Xu David Warde-Farley Sherjil Ozair Aaron Courville Yoshua Bengio Ian J. Goodfellow, Jean Pouget-Abadie. Generative adversarial nets. *Advances in neural information processing systems*, 27, 2014. 1

[7] David Kahn. The history of steganography. In *International workshop on information hiding*, pages 1–5. Springer, 1996. 1

[8] Adi Y. Raj-B. Singh R. Kreuk, F. and J. Keshet. Hide and speak: Deep neural networks for speech steganography. 2019. 1

[9] Dwiti Pandya, Khushboo Ram, Sneha Thakkar, Tanvi Madhekar, and B Thakare. Brief history of encryption. *International Journal of Computer Applications*, 131(9):28–31, 2015. 1

[10] J. Liu R. Zhang, S. Dong. Invisible steganography via generative adversarial networks. *Multimed. Tools Appl.*, 78(7), 2019. 1