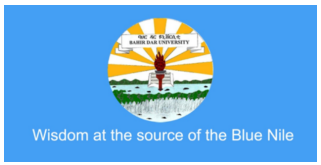
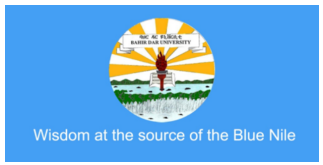
Powerful Generative Steganography With Generative Adversarial Networks

Nathan Melaku

BahirDar Institute Of Technology

December 4, 2019





Introduction
Literature Review
Problem Statement
Objectives
Methodology
Scope
Significance Of The Study
Work Plan
Budget
References

Introduction

Literature Review

Problem Statement

Objectives

Methodology

Scope

Significance Of The Study

Work Plan

Budget

References

Introduction

- Steganography
  - Hiding Information in plain sight.
  - simmon's prisoners. (Simmons, 1984)

Introduction

- Steganography
  - Hiding Information in plain sight.
  - simmon's prisoners. (Simmons, 1984)

- GAN

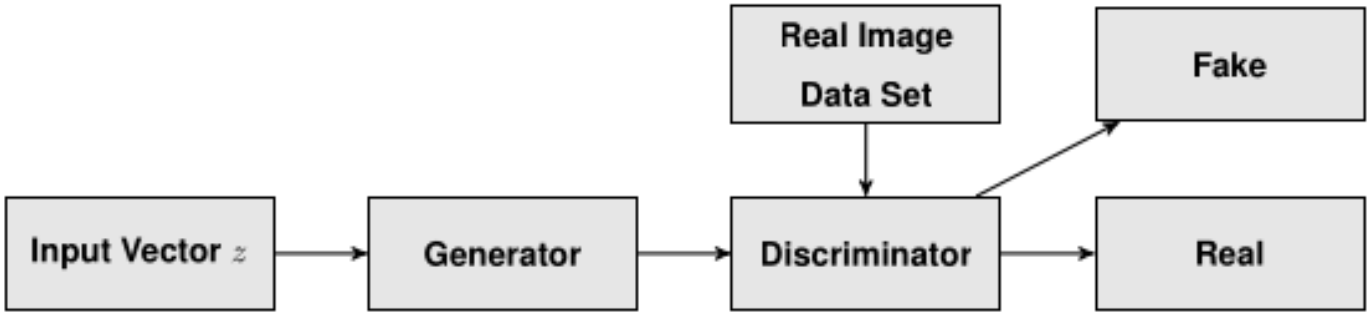


Figure: GAN

$$\min_G \max_D V(D,G) = E_{x\ p_{data}(x)}[logD(x)] + E_{z\ p_z(z)}[log(1-D(G(z)))].$$

Requirements for steganography

- Imperceptibility
- Robustness
- payload capacity

Requirements for steganography

- Imperceptibility
- Robustness
- payload capacity

Requirements for steganography

- Imperceptibility
- Robustness
- payload capacity

Literature Review

Different approaches to steganography

- cover modification
- cover synthesis: (Fridrich, 2009)

GAN in Steganography.

- (Volkhonskiy et al., 2017)
- (Shi et al., 2017)
- And some more others (Tang et al., 2017),(Hayes & Danezis, 2017)

GAN in Steganography.

- (Volkhonskiy et al., 2017)
- (Shi et al., 2017)
- And some more others (Tang et al., 2017),(Hayes & Danezis, 2017)

GAN in Steganography.

- (Volkhonskiy et al., 2017)
- (Shi et al., 2017)
- And some more others (Tang et al., 2017),(Hayes & Danezis, 2017)

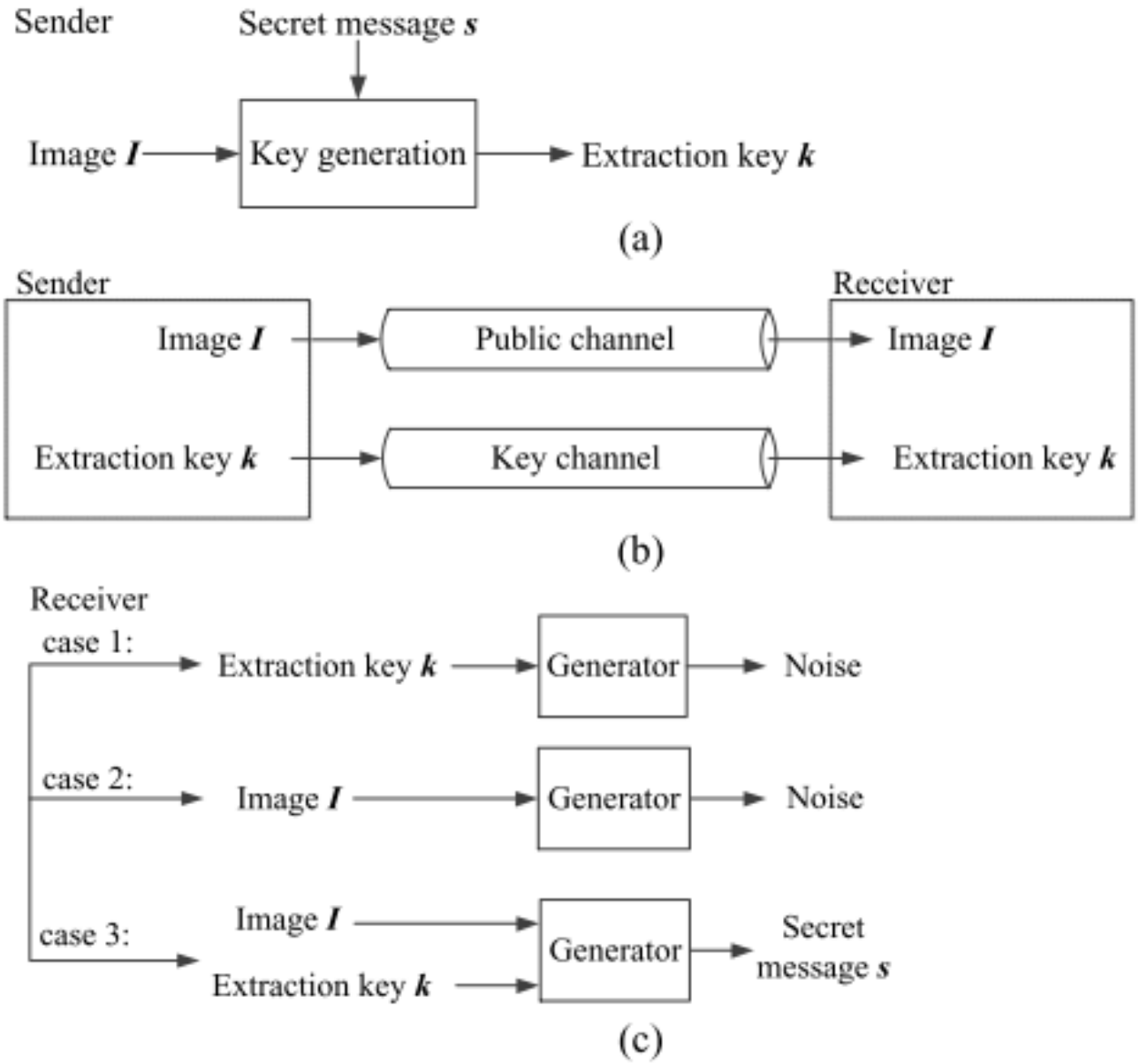


Figure: Proposed System in (Ke et al., 2017)

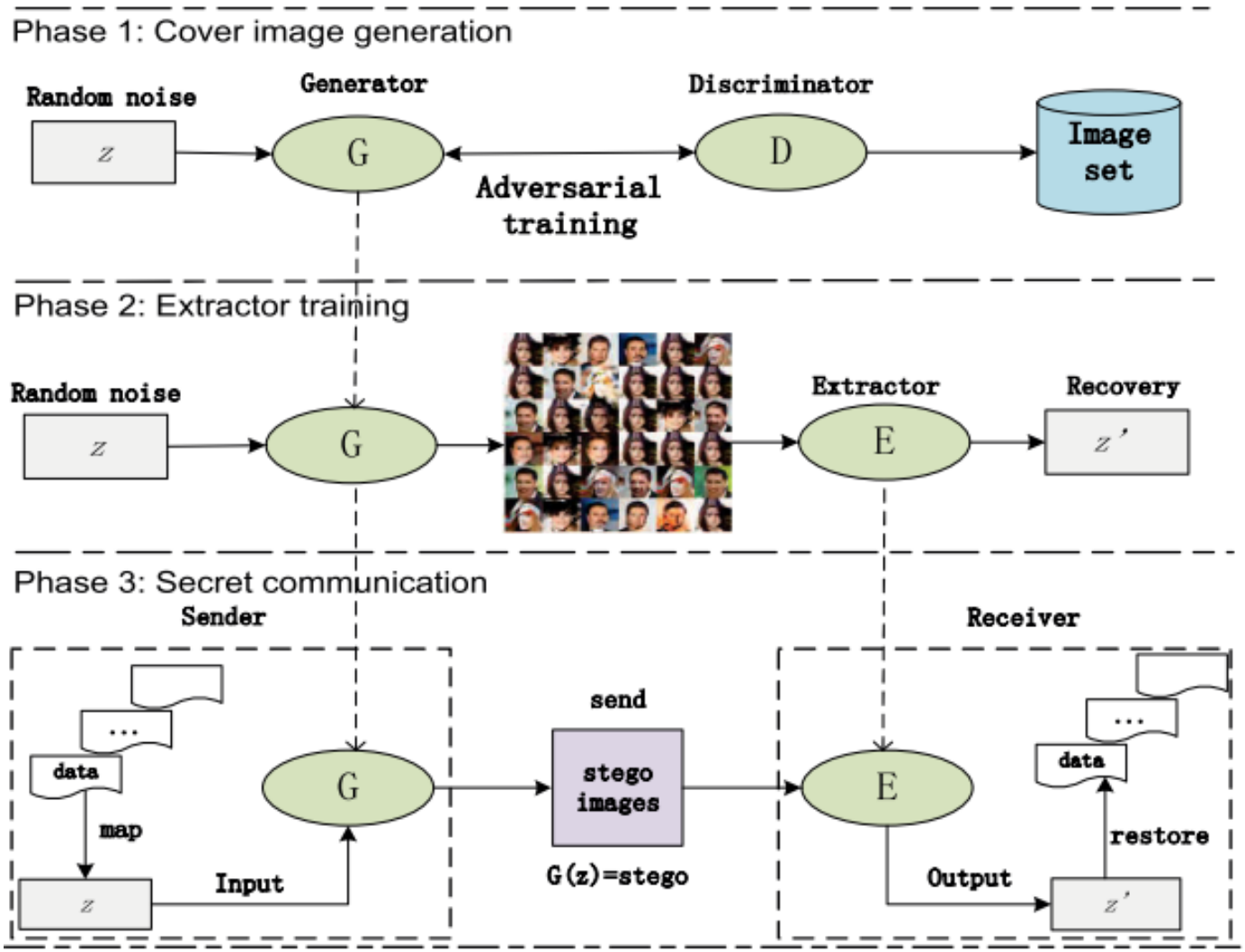


Figure: Proposed System in (Hu et al., 2018)

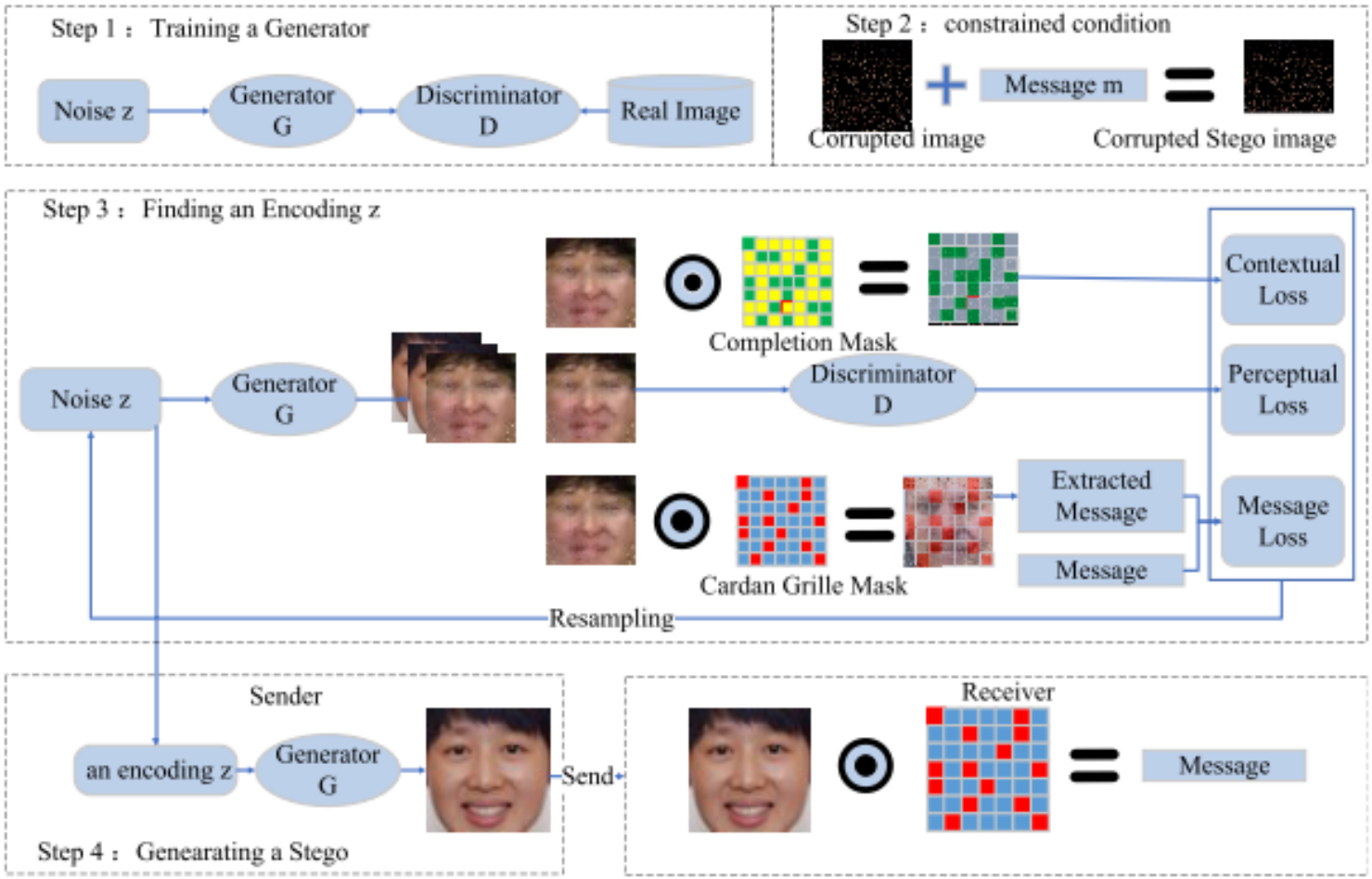


Figure: Proposed System in (Zhang & Yang, 2019)

Problem Statement

Observed problems:

- Instability of Training
- Slow convergence
- Inadequately realistic image generation

"This is mainly associated to the GAN technique used in the framework."

Problem Statement

Observed problems:

- Instability of Training
- Slow convergence
- Inadequately realistic image generation

“This is mainly associated to the GAN technique used in the framework.”

Objectives

Design and implement a powerful generative steganographic framework using a most suitable GAN.

Specific objectives:

- Improve stability of Training and convergence speed.
- Generate more realistic image.
- Maintain state of the art security.

Specific objectives:

- Improve stability of Training and convergence speed.
- Generate more realistic image.
- Maintain state of the art security.

Specific objectives:

- Improve stability of Training and convergence speed.
- Generate more realistic image.
- Maintain state of the art security.

Methodology

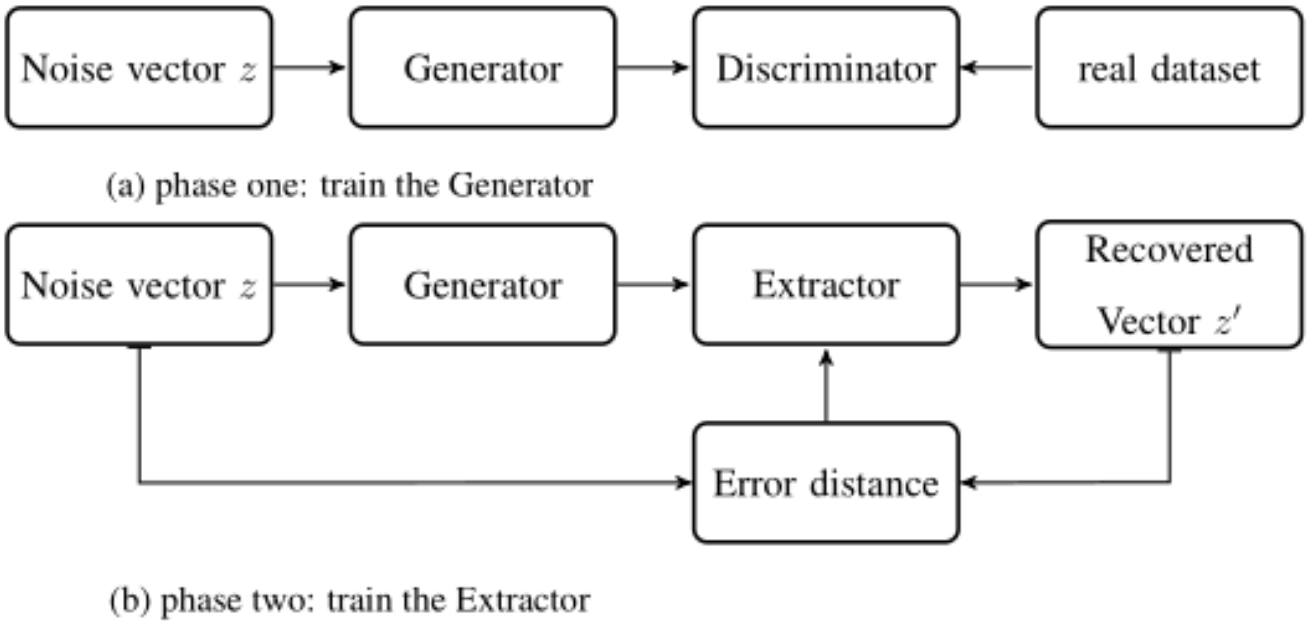


Figure: System Architecture

Choice of GAN:

- BEGAN
- RGAN
- WGAN-DIV
- WGAN-GP

Choice of dataset:

- CelebA
- LFW
- Food101
- MNIST

Choice of testing method:

- Visual instpection
- Inception score
- Visual Turing test
- MOS

Scope

- Robustness
- Passive Adversary
- Payload Capacity
- Imperceptibility

Scope

- Robustness
- Passive Adversary
- Payload Capacity
- Imperceptibility

Scope

- Robustness
- Passive Adversary
- Payload Capacity
- Imperceptibility

Scope

- Robustness
- Passive Adversary
- Payload Capacity
- Imperceptibility

Significance Of The Study

The main problem in generative steganography

Imperceptibility

Work Plan

Tasks	Month	December				January				February				March				April				May				June			
	Week	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
1. Literature Review																													
2. Implementation																													
2.1 With BEGAN																													
2.2 With WGAN-DIV																													
2.3 With WGAN-GP																													
2.4 With RGAN																													
2.5. Extractor																													
3. Train Implementations																													
4. Run Tests																													
5. Analysis																													
6. Documentation																													
7. Draft Report submission																													
8. Final Report submission																													

Figure: Proposed work plan

Budget

Main cost of the project:

- Powerful GPU for training.
  - From the university
  - Amazon EC2
  - Google Cloud

Reference I

Fridrich, J. (2009). *Steganography in digital media: principles, algorithms, and applications*. Cambridge University Press.

Hayes, J., & Danezis, G. (2017). Generating steganographic images via adversarial training. In *Advances in neural information processing systems* (pp. 1954–1963).

Hu, D., Wang, L., Jiang, W., Zheng, S., & Li, B. (2018). A novel image steganography method via deep convolutional generative adversarial networks. *IEEE Access*, *6*, 38303–38314. doi: 10.1109/ACCESS.2018.2852771

Ke, Y., Zhang, M.-q., Liu, J., Su, T.-t., & Yang, X.-y. (2017). Generative Steganography with Kerckhoffs' Principle based on Generative Adversarial Networks. *arXiv preprint arXiv:1711.04916*, 1–5.

Reference II

Shi, H., Dong, J., Wang, W., Qian, Y., & Zhang, X. (2017). Ssgan: secure steganography based on generative adversarial networks. In *Pacific rim conference on multimedia* (pp. 534–544).

Simmons, G. J. (1984). The prisoners' problem and the subliminal channel. In *Advances in cryptology* (pp. 51–67).

Tang, W., Tan, S., Li, B., & Huang, J. (2017). Automatic steganographic distortion learning using a generative adversarial network. *IEEE Signal Processing Letters*, *24*(10), 1547–1551.

Volkhonskiy, D., Nazarov, I., Borisenko, B., & Burnaev, E. (2017). Steganographic generative adversarial networks. *arXiv preprint arXiv:1703.05502*.

Reference III

Zhang, M., & Yang, X. (2019). Generative Steganography by Sampling. *IEEE Access*, *7*, 118586–118597. doi: 10.1109/ACCESS.2019.2920313

THANK YOU