# Image Steganography Analysis and Detection

Subalakshmi Shanthosi S

SSN College of Engineering

14-05-2019

# Presentation Outline

# Image Steganography



- Steganography is the process of hiding a secret message within a larger one in such a way that someone can not know the presence or contents of the hidden message
- Aim - To develop a detection system which is capable of detecting the alteration in image both its format and signature thereby predicting the actual type of forged file

# Challenges in Forged File Discovery



- Images without watermarking as digital signatures can be easily manipulated.
- With the advent of new photo editing software - hiding critical informations are easy and unpercievable
- Task to detect mix of scaled or compressed images as one is difficult
- Incorporating machine learning techniques for feature analysis and decision making to classify the image to be forged or not
- Tamper detection to check for change in the file format extension

# Presentation Outline

# Image Steganography Based on Modified LSB Substitution Method and Data Mapping [1]



- Steganographic method working on the principle of Modified LSB Technique with specific intend of reducing the number of 1's in the secret data
- Methods:Each pixel value of host image is changed if value of secret bit is 1 otherwise the LSB of each pixel value will remain unchanged
- Limitation :
  - Less secure
  - Limited pixel quality
- State-of-the-art methods in terms of PSNR,SSIM
- Future work:Better data mapping mechanism for reduced storage and computational performance

# A Review on Deep Learning based Image Steganalysis [2]



- Steganalysis based on deep learning approach
- Classified as the following categories:
  - Spatial Image Steganalysis
  - JPEG Domain Steganalysis
- Deep Learning based Steganalysis
  - Spatial Domain Steganography Steganalysis based on Deep Neural Network Design
    - Spatial Rich model(SRM)
    - Steganalysis Based on Fusion Approach
    - Steganalysis methods based on Learning Strategy
  - Jpeg Domain Steganography Steganalysis based on Deep Learning
    - Convolutional Neural Network(CNN) with 20 layers
    - CNN with 32 layers combined with SCA-GFR
    - CNN with four 5 x 5 high pass filters, which include a KV filter, a point filter, and 2 Gabor filters, are used to detect stego noise introduced by JPEG-domain embedding scheme

# A Review on Deep Learning based Image Steganalysis [2]



- Limitation and Mitigation:
- Acquisition and representation of statistical characteristics: Using Generative Adversarial Network(GAN)
- Low payload steganographic image detection: Combination of neural neteork design and various other techniques like training sample creation and learning
- Generalization of steganalysis: Combine Transfer Learning and Deep Learning
- Quantitative and locating image steganalysis based on deep learning
- Future work: Challenges resolution by adapting new learning and training sample techniques

# Steganalysis based on Steganography Pattern Discovery[3]



- SPD Approach:
  - Evolutionary method Steganalysis to extract the signature of stego images against clean images via fuzzy ifthen rules
  - Blind steganalysis on the discovered knowledge, suitable trained models for steganalysis can be employed and stego images will be detected with high accuracy
  - Using SPD, we can predict the type of steganography method from a stego image [ Employing SPD can enhance the approaches, which assume that a special steganography method is used
  - The effect of SPD before applying steganalysis methods has been investigated by some steganography and steganalysis techniques and it has been validated using some image databases
  - The second stage is a compound deep CNN network in which the model parameters are learned in the training procedure

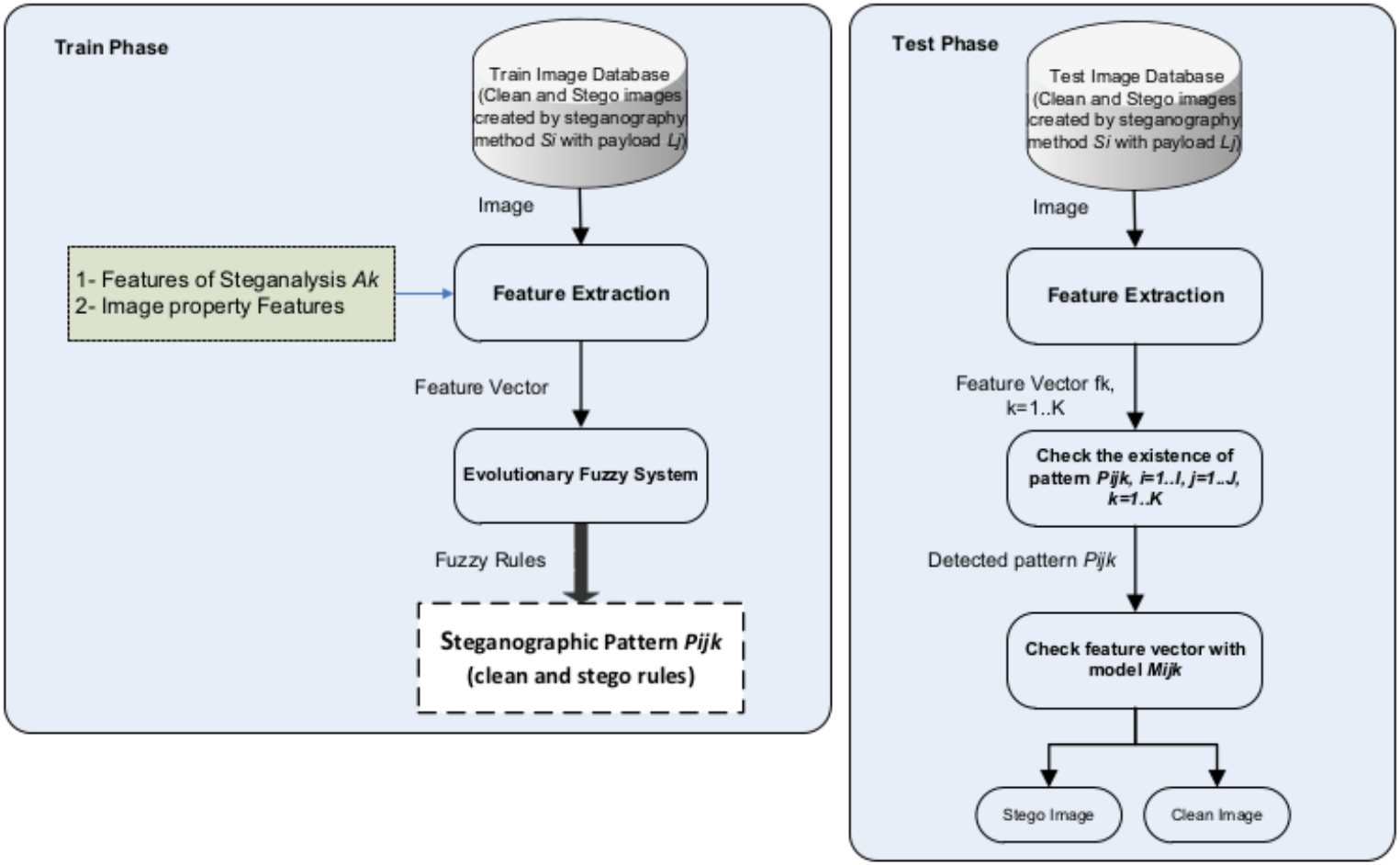
# Steganalysis based on Steganography Pattern Discovery[3]



- Steps carried out in SPD :
  - Image Feature Selection, Two Groups as methods:
    - Filtering :Select feature subsets independently from the learning classifiers and do not include learning
    - Wrapping :Wrap around a certain learning algorithm that can assess the selected feature subsets in terms of estimated classification errors and then build the final classifiers
  - Fuzzy rule generation : Iterative Rule Learning approach, each individual codes one rule and in each iteration of Genetic Algorithm (GA) a new rule is adapted and added to the rule set, iteratively

# The block diagram of Steganography pattern discovery [3]





# Steganalysis based on Steganography Pattern Discovery[3]



- Feature vector generation
  - 274 dimension Steganalyser
  - 324 dimension Feature Vector - First order and second order histograms.
  - Wavelet based Steganalysis
  - 14 dimensional Feature Vector

# Steganalysis based on Steganography Pattern Discovery[3]



- Fuzzy rule generation - Evolutionary method searches for a relatively smaller if-then-rules
- Certinity Factor for Fuzzy Rules :
  - Calculate the compatibility of each training sample
  - For clean and stego images, calculate the relative sum of the compatibility grades of training samples with rule $R_j$ [
  - The grade of certainty $CF_j$ for clean images [
- Evolutionary Fuzzy Algorithm
  - Initiation.
  - Generation.
  - Replacement.
  - Inner Cycle Termination Test.
  - Outer Cycle Termination Test.
  - Weight Adjustement.

# Steganalysis based on Steganography Pattern Discovery[3]



- Evolutionary Method :
  - Extracts signature of stego images against clean images using Fuzzy if-then-rules statements
  - The Steganalyzer trained to detect only one steganography method at once
- Limitations :
  - Using Fuzzy rules increases computational complexicity
  - Only 4 Class Feature Classification - Limited Features

# A Novel Image Steganography Method via Deep Convolutional Generative Adversarial Networks[4]



- A novel Steganography without Embedding (SWE), which does not need to modify the data of the carrier image, appeared to overcome the detection of machine-learning-based steganalysis algorithms
- SWE method based on deep convolutional generative adversarial networks.
  - Generative Adversarial Network (GAN)
    - GAN and discriminative model [
    - The henerative model deceives the discriminative model via generated images that appear like real images while the discriminative model judges whether the images are real or unreal.
  - Deep Convolutional Generatice Adversarial Network (DGAN)
    - Deep convolutional generative adversarial networks (DCGANs) are an extension of GANs in which the models are deep convolutional networks
    - Currently, GANs are widely used for the following works:

# A Novel Image Steganography Method via Deep Convolutional Generative Adversarial Networks[4]



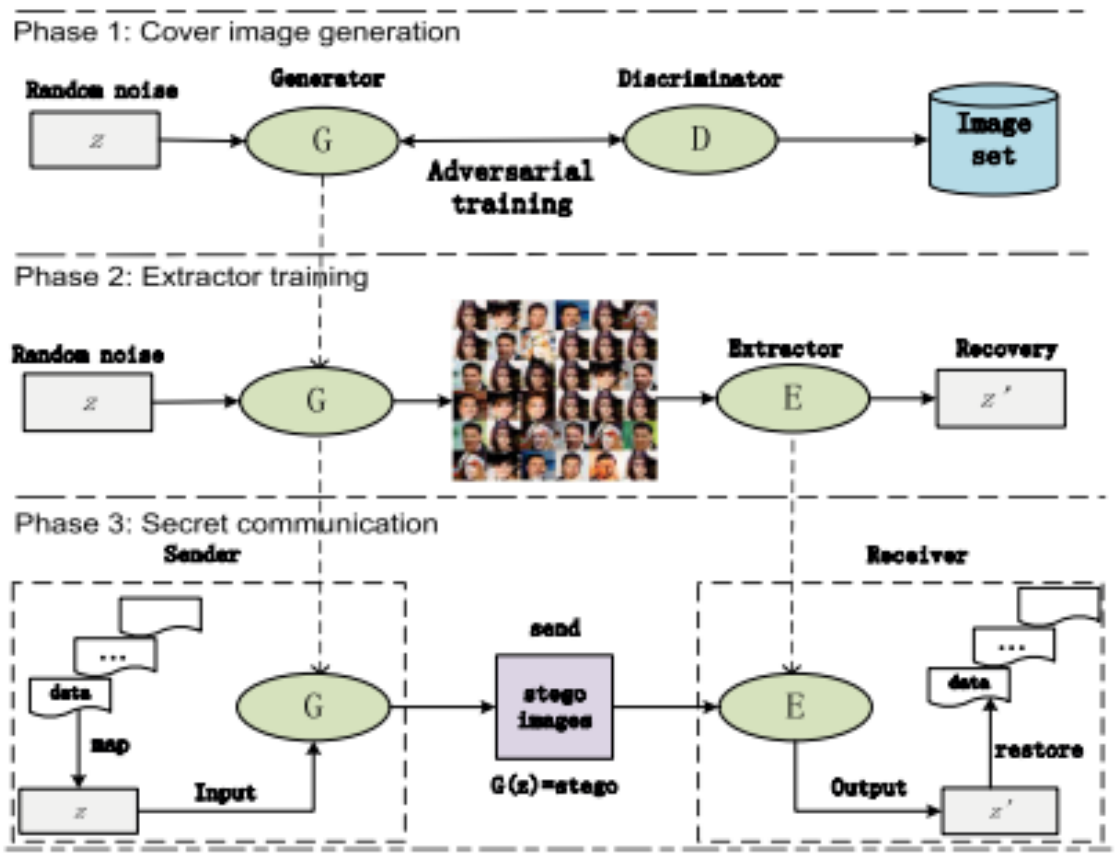



FIGURE 1. The proposed steganography framework using DCGANs for SWE.

Figure: Steganography Framework using DCGAN and SWE.

# A Novel Image Steganography Method via Deep Convolutional Generative Adversarial Networks[4]



- The Proposed Image Steganography without embedding:
  - Train DCGANs on an image set and obtain generator G after DCGANs convergence.
  - Train a CNN's model, called the extractor E, based on the recovery errors from a large number of random noise vectors.
  - The sender and the receiver hold the network and parameters of G and E, respectively.
- Cover Image Generation
  - Secret message is segmented $S_i$ and then map each segment $S_i$ to noise vector $z_i$ .
  - Generate a cover image stegoi from the noise vector $z_i$ with the help of DCGANs
- Training of the Extractor
  - We design the CNNs, called the extractor E, to recover the secret data from stego images generated by G .
  - Has four convolutional fully connected layer .

# A Novel Image Steganography Method via Deep Convolutional Generative Adversarial Networks[4]



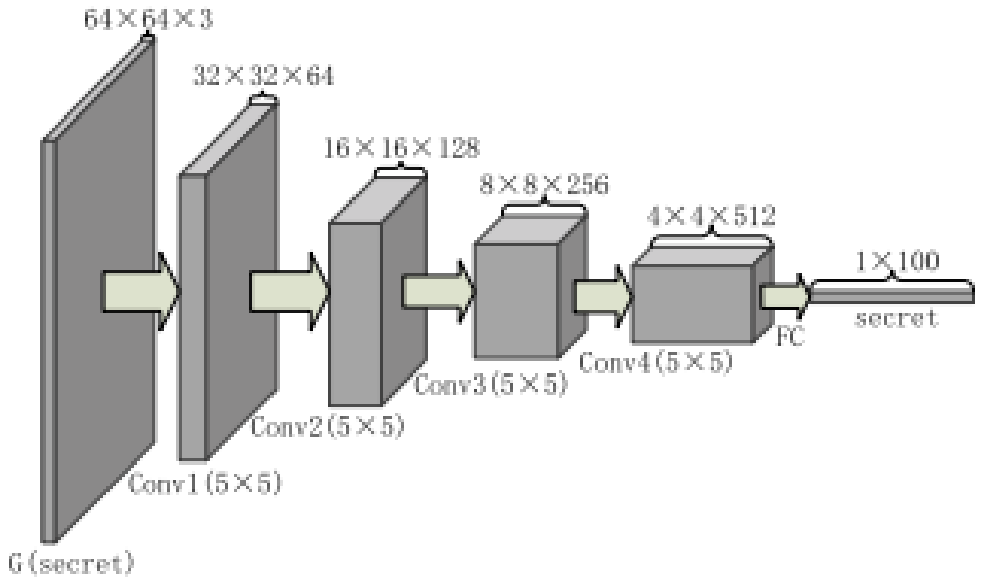



**FIGURE 3.** **The structure of extractor E.**

Figure: The structure of Extractor - E.

# A Novel Image Steganography Method via Deep Convolutional Generative Adversarial Networks[4]



- leak-Relu activation function and batch normalization in each layer with no pooling layer or dropout operation
- Afully connected layer is used after last convolutional layer
- Train E to extract information from the generated stego images from G
- The training procedure of the extractor is,

## Formula

$$L(E) = \sum_{i=1}^{n} (z - E(stego)^2) \tag{1}$$

# A Novel Image Steganography Method via Deep Convolutional Generative Adversarial Networks[4]



- Secret Communication
  - Sender holds the CNNs model G and the corresponding network parameters of G and the receiver holds the CNNs model E and the corresponding network parameters of E

# Steganography Algorithms Recognition based on Match Image and Deep Features Verification[5]



- Sub section of steganalysis.
- Match image is generated by **Gaussian Filtering** on testing image to remove the possible stego signal.
- CNN model is trained on test images to extract deep features from test and match images.
- Proposed system also works better with unknown dataset.
- Related work:
  - Residual filters in steganalysis : Stego data to be placed in highly noise area to make it undistinguishable from noise component.
  - Convolutional Neural Network for steganalysis: Convolution followed by deep inner product to extract features. Finding local features.

# Steganography Algorithms Recognition based on Match Image and Deep Features Verification[5]



- Motivation and Analysis of Cover Source Mismatch (CSM)
  - Difficult approach
  - Stego signal S obeys Gauss distribution

$$S \sim N(0, \sigma^2) \tag{2}$$

  - Variable-controlling approach is adopted.
  - Motivation :
    - Match image should not contain stego signal.
    - Match and test image should be similar.

# Steganography Algorithms Recognition based on Match Image and Deep Features Verification[5]



- Obtaining match image:
  - Obtained by reference.
  - Gaussian filter is used on test image to remove stego signal.
  - Two dimensional Gaussian distribution

$$G(x,y) = \frac{1}{2\pi\sigma^2} exp(-\frac{x^2+y^2}{2\sigma^2}) \quad (3)$$

  - Match image generation :

$$I_m = conv2(I_t, K_G) \quad (4)$$

- CNN-extracted deep features
- Using low dimensional representation of image.
- Modelled CNN for extracting features.
- Performance Analysis
- Feature Normalisation
- Weighted similarity and Inner product similarity.

# Steganography Algorithms Recognition based on Match Image and Deep Features Verification[5]



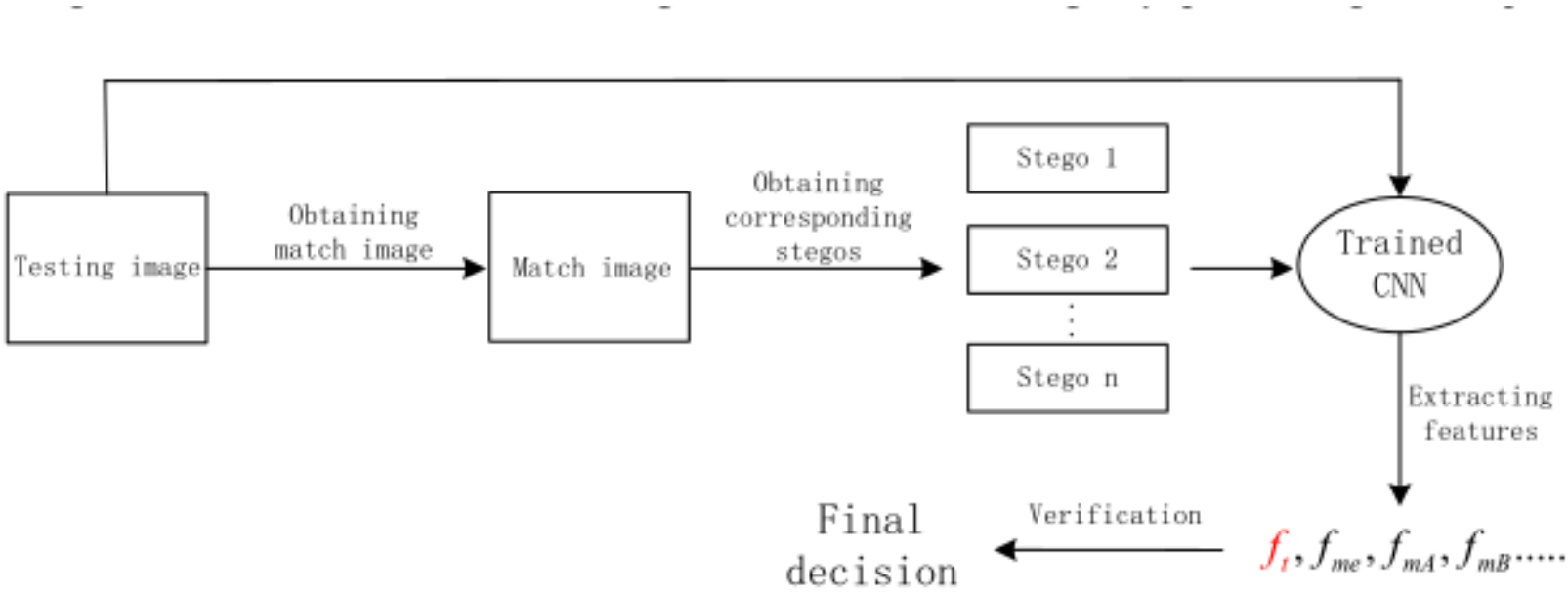



Figure: Framework of match steganalysis.

# Steganography Algorithms Recognition based on Match Image and Deep Features Verification[5]



- Works even with unseen cover images.
- Computational processing higher than that of the conventional method.
- Future work:
  - More practical Matching of images.
  - More accurate similarity function.
  - Reducing the computational complexcity.

# Presentation Outline

- **Digital Image Steganography Using Modified LSB and AES Cryptography[1]** , International Journal of Computer Science and Network Security(IJCSNS) ,Subhash Panwara , Shreenidhi Damanib , Mukesh Kumar DOI.

- **A Review on Deep Learning based Image Steganalysis [2].** , 2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC) ,Yong-he TANG, Lie-hui JIANG, Hong-qi HE, Wei-yu DONG , DOI.

- **Steganalysis based on steganography pattern discovery[3]** , Journal of Information Security and Applications , Hedieh Sajedi , DOI.

- **A Novel Image Steganography Method via Deep Convolutional Generative Adversarial Networks[4].** , IEEE Access , Donghui Hu 1 , Liang Wang1 , Wenjie Jiang , Shuli Zheng , Bin Li , DOI.

- **Steganography Algorithms Recognition based on Match Image and Deep Features Verification[5].** , Multimedia Tools and Applications Journal ,Xu Xiaoyu ,Sun Yifeng Wu, Jiang , Sun Yi DOI.