symmetry

MDPI

Article

Diversified Cover Selection for Image Steganography

Xinran Li [1,*], Daidou Guo [2] and Chuan Qin [2]

[1] Business School, University of Shanghai for Science and Technology, Shanghai 200093, China
[2] School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China; 211240061@st.usst.edu.cn (D.G.); qin@usst.edu.cn (C.Q.)
* Correspondence: xinranli@usst.edu.cn

Abstract: This paper proposes a new cover selection method for steganography. We focus on the scenario that the available images for selection contain diversified sources, i.e., nature images and metaverse images. For the scenario, we design a targeted strategy to evaluate the suitability for steganography of a candidate image, which selects images according to the undetectability against steganalytic tools symmetrically. Firstly, steganalytic features of the candidate images are extracted. Then, the features are fed on a steganalytic classifier, and the possibility of carrying secret data is calculated for cover selection. As a result, the selected images are the best candidates to resist steganalysis. Experimental results show that our method performs better than existing cover selection schemes when checked by steganalytic tools.

Keywords: steganography; cover selection; diversified sources; digital image

1. Introduction

The field of steganography aims to discreetly embed confidential information within regular media, enabling its transmission without arousing suspicion [1]. To achieve this objective, modifications must be made to the content of cover images, which serve as carriers for the concealed data. Steganalysis, as an adversarial approach to steganography, focuses on discerning the presence of hidden data by scrutinizing the statistical attributes of stego images—images that bear concealed information [2]. Both techniques have witnessed substantial advancements over the past two decades [3]. Currently, cutting-edge steganographic techniques focus on minimizing the alterations' impact to enhance their concealment against steganalysis. This objective is accomplished by employing advanced methods like STC (syndrome trellis coding) [4] or SPC (steganographic polar codes) [5], which utilize predefined distortion functions. For spatial images, several distortion functions have been proposed, including WOW (wavelet obtained weights) [6], SUNIWARD (spatial universal wavelet relative distortion) [7], HILL (high-pass, low-pass, and low-pass) [8], and MiPOD (minimizing the power of optimal detector) [9]. Regarding JPEG images, researchers have developed UED (uniform embedding distortion) [10], UERD (uniform embedding revisited distortion) [11], and GUED (generalized UED) [12].

In certain steganographic scenarios, the sender is presented with multiple options in the form of candidate images (cover images) for embedding secret data, subsequently transmitting the resulting stego images. The sender's objective in such cases is to strategically choose the most appropriate images for steganography, thereby minimizing the risk of exposure. As depicted in Figure 1, the available image pool comprises content sourced from social networks, the Internet, and captured using a camera. All these images are potential candidates for steganographic purposes. Given the communication channel's limited capacity, it is essential for the sender to carefully choose a subset of images suitable for data embedding. Consequently, the process of cover selection becomes crucial in determining the optimal images that guarantee undetectability when subjected to steganalysis.

check for updates

Citation: Li, X.; Guo, D.; Qin, C. Diversified Cover Selection for Image Steganography. *Symmetry* **2023**, *15*, 2024. https://doi.org/10.3390/sym15112024

Academic Editors: Zichi Wang and Alexander Shelupanov

Received: 26 September 2023
Revised: 24 October 2023
Accepted: 25 October 2023
Published: 6 November 2023

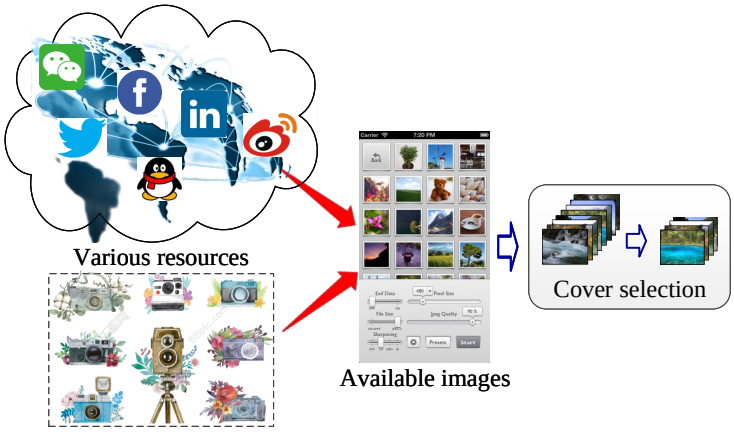

Figure 1. Applications of cover selection.

In current daily life, digital images can be obtained from diversified sources such as nature images and metaverse images. As shown in Figure 2, the statistical properties of these two kinds of images are quite different. Nature images are the objective description of the real world, while metaverse images are the self-definition of virtual space. Existing cover selection schemes are not appropriate for nature images and metaverse images simultaneously. In view of this, we focus on the scenario that the available images for selection contain not only nature images but also metaverse images. We propose a new cover selection method that is designed for the image set that contains both nature images and metaverse images. Specifically, according to the undetectability against steganalytic tools, we designed a strategy to evaluate the suitability for steganography of a candidate image. Steganalytic features of the candidate image are extracted and fed on a steganalytic classifier. Then, the possibility of carrying secret data is calculated for cover selection. In this way, the selected images are the best candidates to resist steganalysis.

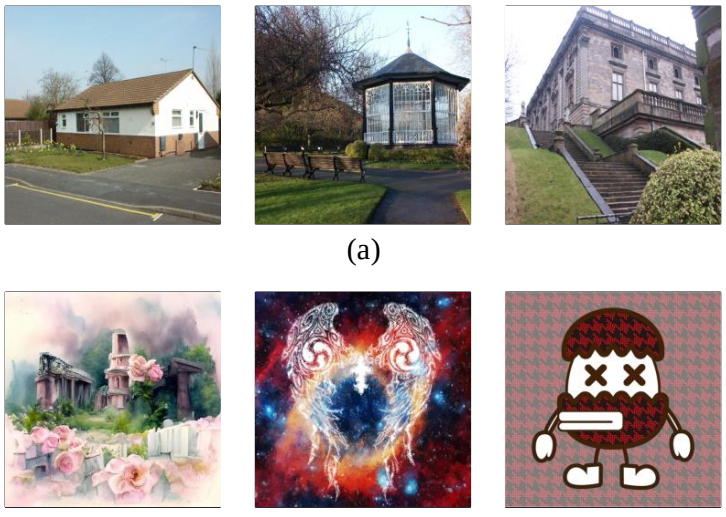

Figure 2. Nature images and metaverse images: (**a**) nature images; (**b**) metaverse images.

Our motivation encompasses two primary aspects. Firstly, instead of developing a new embedding algorithm, we aim to enhance the undetectability of steganography by carefully selecting appropriate cover images. This approach holds practical value as senders typically have access to a diverse range of available images obtained through methods such as shooting, downloading, and generation. Secondly, our focus lies in the scenario where the available images encompass both nature-themed and metaverse-themed content, reflecting the current abundance of digital images from various sources in

everyday life. Through our proposed method, we significantly elevate the undetectability of steganography, which stands as a crucial metric within the field. The contributions offered by this paper are delineated as follows:

(1) **New scenario:** We examine a practical scenario where the selection pool of available images comprises both nature-themed and metaverse-themed content. It is a significant scenario since digital images can be obtained from diversified sources. However, existing cover selection schemes are not appropriate for nature images and metaverse images simultaneously. To this end, we propose to select cover images according to undetectability against steganalytic tools. In this way, the selected images are the best candidates to resist steganalysis.

(2) **New methodology:** Our method is based on the theory of ensemble learning, which is widely used in steganalysis. The theoretical justification of our method is more practical since it resists steganalytic tools directly. Based on the ensemble learning theory, a targeted cover selection method has been devised to identify the most suitable candidates for resisting steganalysis. Our proposed approach surpasses existing cover selection schemes in terms of performance.

The following sections of this manuscript follow a specific structure: Section 2 presents a comprehensive review of the pertinent literature. In Section 3, we outline our proposed method for cover selection. To validate the effectiveness of our approach, we conduct a series of experiments in Section 4. Lastly, Section 5 provides a concise summary of the comprehensive conclusions derived from this study.

2. Related Work

This section presents a comprehensive overview of relevant literature, encompassing established methodologies for cover selection in steganography as well as contemporary steganalytic tools.

2.1. Cover Selection Schemes

Since we aim to select suitable cover images for steganography, the most related work of our method are other cover selection schemes designed for steganography. Currently, there exist a variety of cover selection techniques utilized in the field of steganography. These techniques encompass different methodologies, such as empirical selection based on changeable DCT (discrete cosine transform) coefficients [13], visual quality assessment [14], ratio of similar blocks [15], content complexity, and texture regions [16,17]. However, it is important to acknowledge that these empirical approaches do not demonstrate optimal undetectability when subjected to steganalysis. Wu et al. [18] conducted work involving the modeling of digital images using the Gaussian mixture model, utilizing the real value derived from the Fisher information matrix as a measure of image suitability. Alternatively, Wang et al. [19] presented a scheme specifically designed to counter pooled steganalysis. This approach focuses on maintaining the MMD (maximum mean discrepancy) distance between arbitrary and stego image sets below a predetermined threshold during the cover selection process. In a similar vein, the authors in [20] proposed an innovative strategy to circumvent the use of processed images as covers. They introduced a scheme that effectively identifies and eliminates processed images from the pool of candidates. In [21], the embedding distortion of individual candidate images was directly computed using a distortion function. Consequently, the selection process favored cover images with lower embedding distortions. Additionally, in the work presented by Wang et al. [22], an alternative approach was proposed to avoid the selection of similar images for steganography purposes. This led to a reduction in the provision of less relevant samples for steganalysis. Building upon this premise, image similarity and embedding distortion were combined as joint metrics to evaluate the suitability of each available image.

Most existing cover selection schemes are based on the theory of embedding distortion minimization, which selects suitable covers according to the theoretical distortion. Different from existing cover selection schemes, our method is based on the theory of ensemble

learning, which is widely used in steganalysis. The theoretical justification of our method is more practical since it resists steganalytic tools directly. In addition, the application scenario of our method is also different from existing cover selection schemes. We focus on the scenario that the available images for selection contain not only nature images but also metaverse images. Since the candidate images contain diversified sources, we design a targeted strategy to evaluate the suitability for steganography, which select images according to their undetectability against steganalytic tools. In this way, the selected images are the best candidates to resist steganalysis.

2.2. Digital Steganalysis

Steganalysis endeavors to detect the presence of steganography by scrutinizing the statistical characteristics exhibited by stego images. The underlying assumption is that stego images exhibit substantial disparities compared to unaltered ones, attributable to the alterations incurred during data embedding [23]. Contemporary steganalysis methodologies leverage supervised machine learning techniques to discriminate between clear image models and stego image models. A variety of feature extraction methodologies have been employed in the existing literature for training a comprehensive steganalytic classifier, which subsequently enables the discrimination between clear images and stego images [24]. Noteworthy examples include SPAM (subtractive pixel adjacency matrix) [25], SRMQ1 (SRM with single quantization step) [26], maxSRMd2 (selection-channel-aware variant of SRM) [27], and TLBP (threshold local binary pattern) [28].

The ensemble classifier, a prevalent choice for steganalytic classification, is utilized to assess the characteristics of extracted feature sets [29]. This ensemble classifier comprises numerous base learners generated via FLD (Fisher linear discriminant) learners. Each FLD learner undergoes training on an independently selected subspace within the comprehensive feature space, subsequently producing a binary decision, namely "clear" or "stego", for each input image. The final decision of the ensemble classifier is determined by aggregating the votes cast by all FLD learners. The performance evaluation metric for the ensemble classifier is denoted as the minimal total error, represented as P_{E}, achieved on the testing sets with equal prior probabilities. This minimal total error is defined mathematically in Equation (1).

$$P_{\mathrm{E}} = \min_{P_{\mathrm{FA}}}\left(\frac{P_{\mathrm{FA}} + P_{\mathrm{MD}}}{2}\right). \tag{1}$$

In this context, the parameters P_{FA} and P_{MD} represent the false alarm rate and missed detection rate, respectively. A higher value of P_{E} indicates diminished accuracy in steganalysis, resulting in a heightened level of undetectability within steganography.

Recently, researchers developed a number of deep learning-based steganalytic networks, in which the operations of feature extraction and classifier training are joined together. The detection accuracy of deep learning-based steganalysis clearly exceeds that of handcrafted feature-based steganalysis in [30] by the utilization of high-pass filters. Further, the authors in [31] employed the residual network to capture the modification details of steganography. In [2], the authors aimed to detect adversarial steganography using a two-stream CNN steganalyzer that leverages confidence artifacts and pixel artifacts. In [32], spatial attention was used to exploit the texture information of the image itself. Currently, the interpretability of deep learning-based steganalysis has not matured yet.

In the experimental phase, a subset of the aforementioned steganalytic tools will be utilized to evaluate the level of undetectability exhibited by our proposed method, along with other cover selection schemes.

3. Proposed Method

This study presents a novel approach for cover image selection in a scenario where the available image pool includes both nature and metaverse images. In this context, we propose a method that prioritizes the undetectability of cover images through direct steganalysis.

3.1. Framework

As discussed in Section 1, when employing steganography, the sender typically possesses multiple candidate images. However, due to limitations in channel capacity, only a subset of these images can be selected for data embedding. This paper specifically addresses a scenario where the available image pool consists of both nature and metaverse images. Existing cover selection approaches do not adequately cater to both image types simultaneously. The primary objective of cover selection is to ensure the undetectability of steganographic content through steganalysis. To achieve this goal, we propose a direct approach for selecting cover images based on their undetectability against steganalysis. By adopting this method, we can develop a cover selection technique that accommodates diverse image sources.

Figure 3 depicts the architectural framework employed in our proposed approach for cover selection. We begin with a collection of n candidate images, represented as $\{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n\}$. The assessment of each image's suitability for steganography is based on its ability to remain undetected by steganalysis techniques. Consequently, the resulting suitability values are denoted as $\{D_1, D_2, \ldots, D_n\}$ and arranged in descending order as $\{D_{f(1)}, D_{f(2)}, \ldots, D_{f(n)}\}$. Subsequently, we select a subset of k images, specifically $\{\mathbf{X}_{f(1)}, \mathbf{X}_{f(1)}, \ldots, \mathbf{X}_{f(k)}\}$, which correspond to the highest suitability values $\{D_{f(1)}, D_{f(2)}, \ldots, D_{f(k)}\}$, where $k \in \{1, 2, \ldots, n\}$. These selected images serve as covers for steganography. Further elucidation regarding this process is provided below.

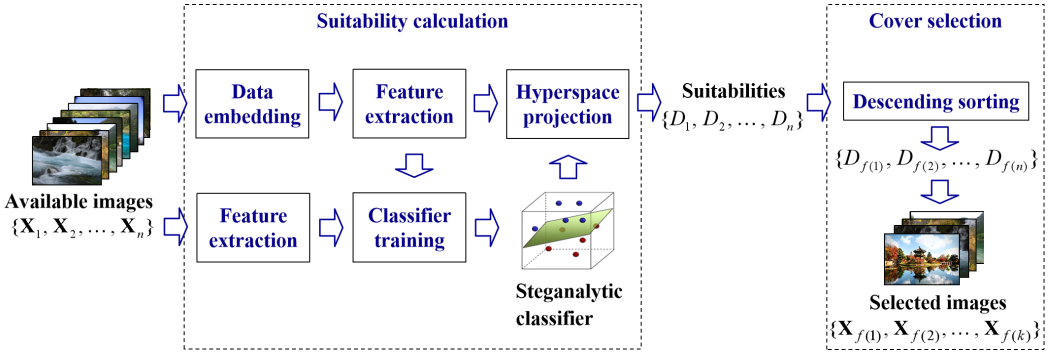

Figure 3. Architecture of proposed method.

3.2. Suitability Calculation

To assess the appropriateness of an image candidate, steganalytic features are first extracted and subsequently inputted into a steganalytic classifier. Following this, the feasibility of embedding concealed information is evaluated through hyperspace projection, serving as a measure of suitability.

Given n candidate images $\{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n\}$, the corresponding feature vectors $\{\mathbf{f}_1^c, \mathbf{f}_2^c, \ldots, \mathbf{f}_n^c\}$ are, respectively, extracted using a steganalytic feature extraction algorithm, e.g., SPAM, SRMQ1, maxSRMd2, or TLBP. Meanwhile, secret data are respectively embedded into the original images $\{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n\}$ with a given payload to obtain the corresponding n stego images using an embedding algorithm, e.g., HILL or MiPOD. Subsequently, the same feature extraction algorithm is applied to these stego images, allowing the extraction of their corresponding feature vectors $\{\mathbf{f}_1^s, \mathbf{f}_2^s, \ldots, \mathbf{f}_n^s\}$. Both HILL and MiPOD are the embedding algorithms based on the embedding distortion minimization framework achieved by STC or SPC. The differences between them are the steps to calculate the embedding cost of each pixel. The four feature extraction schemes SPAM, SRMQ1, maxSRMd2, and TLBP are used to capture the discrepancy of clear and stego images. The differences between them are the strategies to find the modification trace. For example, SPAM employs the transition probability of pixel residuals, SRMQ1 designs a number of high-pass filters, maxSRMd2 proposes to use selection channels of steganography, and the local binary pattern is used in TLBP.

After that, the obtained feature vectors $\{\mathbf{f}_1^c, \mathbf{f}_2^c, \ldots, \mathbf{f}_n^c\}$ and $\{\mathbf{f}_1^s, \mathbf{f}_2^s, \ldots, \mathbf{f}_n^s\}$ are used to train a steganalytic classifier. In this paper, the ensemble classifier is employed since it is widely used in steganalysis. As mentioned in Section 2.2, the ensemble classifier is the combination of a lot of FLD learners. Each FLD learner is subjected to training within a subspace, which is randomly selected from the entirety of the feature space. For each feature vector $\mathbf{f}_i^c$ or $\mathbf{f}_i^s$, following hyperspace projection, each FLD learner provides a binary decision of either "clear" or "stego", where $i \in \{1, 2, \ldots, n\}$. The generalized eigenvector fully characterizes each FLD learner during the hyperspace projection process.

$$\mathbf{v} = (\mathbf{S}_W + \lambda \mathbf{I})^{-1}(\boldsymbol{\mu}_c - \boldsymbol{\mu}_s). \tag{2}$$

where $\mathbf{I}$ is a unity matrix, and $\boldsymbol{\mu}_c$ and $\boldsymbol{\mu}_s$ are the means of class "clear" and "stego", respectively,

$$\boldsymbol{\mu}_c = \frac{1}{n}\sum_{i=1}^{n} \mathbf{f}_i^c, \qquad \boldsymbol{\mu}_s = \frac{1}{n}\sum_{i=1}^{n} \mathbf{f}_i^s. \tag{3}$$

$$\mathbf{S}_W = \sum_{i=1}^{n} (\mathbf{f}_i^c - \boldsymbol{\mu}_c)(\mathbf{f}_i^c - \boldsymbol{\mu}_c)^{\mathrm{T}} + (\mathbf{f}_i^s - \boldsymbol{\mu}_s)(\mathbf{f}_i^s - \boldsymbol{\mu}_s)^{\mathrm{T}} \tag{4}$$

is denoted as the within-class scatter matrix. The small positive value of $\lambda = 10^{-10}$ is used for augmentation. This augmentation ensures the positivity of the resulting matrix $\mathbf{S}_W + \lambda \mathbf{I}$, effectively addressing potential numerical instability issues that may arise in practical scenarios where $\mathbf{S}_W$ is singular or ill-conditioned. In the context of feature vector $\mathbf{y}$ classification, an FLD learner employs the computed projection $\mathbf{v}^{\mathrm{T}}\mathbf{y}$ and compares it against a predetermined threshold, previously established to achieve the desired performance, in order to reach a decision.

Denote the number of FLD learners of ensemble classifier as L, and there are l_i^c and l_i^s FLD learners give a "stego" decision for $\mathbf{f}_i^c$ and $\mathbf{f}_i^s$, respectively, $l_i^c \in \{0, 1, \ldots, L\}$, $l_i^s \in \{0, 1, \ldots, L\}$. Thus, the suitability D_i of $\mathbf{X}_i$ can be calculated as,

$$D_i = L - |l_i^c - l_i^s|. \tag{5}$$

A higher value of D_i stands for higher suitability of $\mathbf{X}_i$. Intuitively, it seems that a stego image is undetectable if few FLD learners make the "stego" decision. In other words, it seems that an image with a small value of l_i^s is suitable for steganography. However, a stego image is produced from its clear version, and thus an undetectable stego image should be as similar as possible to the clear version. Therefore, the logic behind Equation (5) is that an image is suitable for steganography if its stego version is indistinguishable from itself [33]. For this reason, we propose to select the images that show the number of FLD learners who make the "stego" decision on the stego version is close to that on the clear version.

3.3. Cover Selection Strategy

Considering the suitability values $\{D_1, D_2, \ldots, D_n\}$ obtained by evaluating undetectability against steganalysis using Equation (5), we can identify the most appropriate cover images for steganography. By arranging the suitability values in descending order as $\{D_{f(1)}, D_{f(2)}, \ldots, D_{f(n)}\}$, the cover image $\mathbf{X}_{f(1)}$ corresponding to the highest suitability value $D_{f(1)}$ is chosen. If the need arises to select multiple images (denoted by k), then the k images $\{\mathbf{X}_{f(1)}, \mathbf{X}_{f(1)}, \ldots, \mathbf{X}_{f(k)}\}$ are selected as covers, each corresponding to their respective suitability values $\{D_{f(1)}, D_{f(2)}, \ldots, D_{f(k)}\}$. Our proposal advocates selecting images with the highest suitability values from the available options, as greater suitability indicates stronger undetectability.

The procedure of cover selection is summarized as Algorithm 1. Firstly, the feature vectors $\{\mathbf{f}_1^c, \mathbf{f}_2^c, \ldots, \mathbf{f}_n^c\}$ of $\{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n\}$ are extracted using an existing steganalytic feature extraction schemes such as SPAM, SRMQ1, maxSRMd2, and TLBP (Step 1). Meanwhile, se-

cret data are embedded into $\{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n\}$ with a fixed payload such as 0.5 bpp using an existing embedding algorithms such as HILL and MiPOD (Step 2). Then, the feature vectors $\{\mathbf{f}_1^s, \mathbf{f}_2^s, \ldots, \mathbf{f}_n^s\}$ of the n obtained stego images are extracted using the same feature extraction scheme extracting $\{\mathbf{f}_1^c, \mathbf{f}_2^c, \ldots, \mathbf{f}_n^c\}$ (Step 3). With $\{\mathbf{f}_1^c, \mathbf{f}_2^c, \ldots, \mathbf{f}_n^c\}$ and $\{\mathbf{f}_1^s, \mathbf{f}_2^s, \ldots, \mathbf{f}_n^s\}$, the steganalytic classifier can be trained. In our method, an ensemble classifier is employed (Step 4). With the trained ensemble classifier, all values of $\{l_i^c\}$ and $\{l_i^s\}$ are calculated by counting the numbers of FLD learners that give a "stego" decision for $\{\mathbf{f}_i^c\}$ and $\{\mathbf{f}_i^s\}$ (Step 5). Subsequently, the suitability scores $\{D_1, D_2, \ldots, D_n\}$ for each image $\{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n\}$ are computed using Equation (5) (Step 6). As higher values of D_i indicate greater suitability of $\mathbf{X}_i$, the images with the highest suitability scores are chosen as covers. To accomplish this, the suitability scores $\{D_1, D_2, \ldots, D_n\}$ are arranged in descending order as $\{D_{f(1)}, D_{f(2)}, \ldots, D_{f(n)}\}$ (Step 7). Ultimately, the images $\{\mathbf{X}_{f(1)}, \mathbf{X}_{f(2)}, \ldots, \mathbf{X}_{f(k)}\}$ with the topmost suitability scores $\{D_{f(1)}, D_{f(2)}, \ldots, D_{f(k)}\}$ are selected as covers.

It can be known that the computational cost of above steps mainly contains data embedding, feature extraction, and classifier training. Therefore, the computational cost of our method is determined by those of existing embedding algorithms, steganalytic feature extraction schemes, and ensemble classifiers.

Algorithm 1 Cover selection strategy

Input: Available images $\{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n\}$; Number of needed cover images: k
Output: Selected k images $\{\mathbf{X}_{f(1)}, \mathbf{X}_{f(2)}, \ldots, \mathbf{X}_{f(k)}\}$

(1) Extract the feature vectors $\{\mathbf{f}_1^c, \mathbf{f}_2^c, \ldots, \mathbf{f}_n^c\}$ of $\{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n\}$ using an existing steganalytic feature extraction scheme;
(2) Embed secret data into $\{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n\}$ with a fixed payload;
(3) Extract the feature vectors $\{\mathbf{f}_1^s, \mathbf{f}_2^s, \ldots, \mathbf{f}_n^s\}$ of the n stego images obtained by Step (2) using the same feature extraction scheme;
(4) Train an ensemble classifier using the obtained feature vectors $\{\mathbf{f}_1^c, \mathbf{f}_2^c, \ldots, \mathbf{f}_n^c\}$ and $\{\mathbf{f}_1^s, \mathbf{f}_2^s, \ldots, \mathbf{f}_n^s\}$;
(5) Calculate all values of $\{l_i^c\}$ and $\{l_i^s\}$ with the trained classifier;
(6) Calculate suitabilities $\{D_1, D_2, \ldots, D_n\}$ for all images $\{\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_n\}$ using Equation (5);
(7) Sort $\{D_1, D_2, \ldots, D_n\}$ in descending order into $\{D_{f(1)}, D_{f(2)}, \ldots, D_{f(n)}\}$;
(8) Select $\{\mathbf{X}_{f(1)}, \mathbf{X}_{f(2)}, \ldots, \mathbf{X}_{f(k)}\}$ as cover images for steganography.

4. Experimental Results

To validate the efficacy of our approach, a series of experiments were carried out within this section. Initially, we established the experimental settings and subsequently presented the outcomes pertaining to the unnoticeability aspect, as assessed through contemporary steganalytic techniques.

4.1. Experiment Setup

In our experimental investigations, the UCID image dataset [34] was utilized, encompassing a collection of 1338 nature images with dimensions of 512 × 384. Additionally, we incorporated 1000 metaverse images sourced from the Bigverse website [35], which exhibited diverse sizes. To ensure uniformity, all images were resized to dimensions of 512 × 512 and employed as the pool of available images (n = 2338) for cover selection.

Utilizing our proposed methodology, a selection of cover images suitable for steganographic purposes can be made. In the subsequent experimental phase, varying quantities of cover images were chosen from the pool of 2338 available images using our approach, specifically 200, 400, 600, 800, and 1000 images (k = 200, 400, 600, 800, and 1000). For the purpose of comparison, we also employed four cover selection schemes introduced in [19–22], selecting an equal number of images from the same pool of available images. Subsequently, the HILL and MiPOD embedding schemes, both widely acknowledged, were

employed to embed data at a payload of 0.5 bits per pixel (bpp), thereby producing the corresponding stego images.

To assess the undetectability of steganography in the obtained stego images, steganalytic tools described in Section 2.2 were employed. Four widely used feature extraction algorithms, namely SPAM, SRMQ1, maxSRMd2, and TLBP, were utilized to extract the feature sets from both cover and stego images. Afterwards, an ensemble classifier was employed to evaluate the performance of these feature sets. The training phase involved utilizing half of the cover and stego feature sets, while the remaining sets were used for testing. The undetectability of steganography was determined by measuring the minimal total error P_E achieved on the testing sets, assuming identical priors.

4.2. Undetectability

For steganography, the most important indicator is the undetectability measured by P_E. Other indicators for digital images such as PSNR aSSIM can be always satisfactory since the modification amplitude of steganography is +1 or −1 usually. Figures 4–7 presented the comparisons of P_E (minimal total error) achieved by our method and the cover selection schemes proposed in previous works [19–22]. In these graphs, a varying number of images (200, 400, 600, 800, and 1000) were selected from a pool of 2338 available images using either our method or the schemes described in [19–22]. The legends "HILL-Proposed" and "MiPOD-Proposed" refer to cover images selected by our method and subsequently embedded using HILL or MiPOD techniques. The legends "HILL-1" [19], "HILL-2" [20], "HILL-3" [21], "HILL-4" [22], "MiPOD-1" [19], "MiPOD-2" [20], "MiPOD-3" [21], and "MiPOD-4" [22] represent cover images selected by the schemes outlined in [19–22], followed by embedding using HILL or MiPOD methods.

The employed steganalytic feature extraction algorithm in Figures 4–7 varied across different settings. In Figure 4, we utilized SPAM to extract features from the selected images, including those chosen by our method and the scheme presented in [19–22]. Additionally, SPAM was applied to extract features from the stego images generated from the selected cover images. For Figure 5, the feature extraction involved using SRMQ1 on both cover and stego images. Finally, maxSRMd2 and TLBP were employed for feature extraction in Figures 6 and 7, respectively.

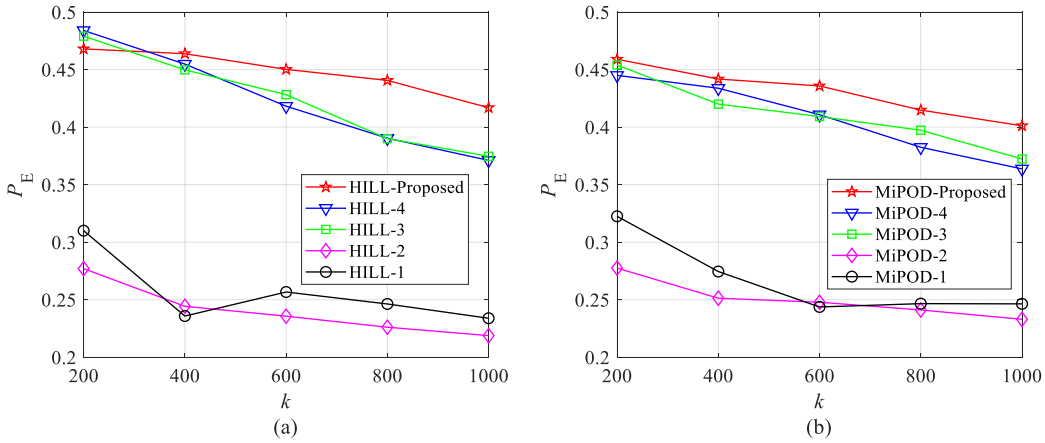

Figure 4. Undetectability comparisons of cover selection methods against steganalytic tool SPAM with embedding schemes (**a**) HILL and (**b**) MiPOD.

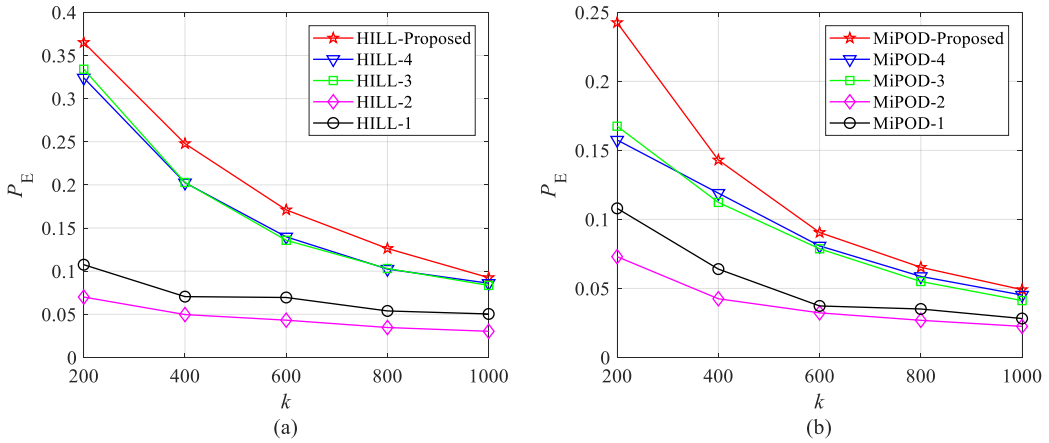

Figure 5. Undetectability comparisons of cover selection methods against steganalytic tool SRMQ1 with embedding schemes (**a**) HILL and (**b**) MiPOD.

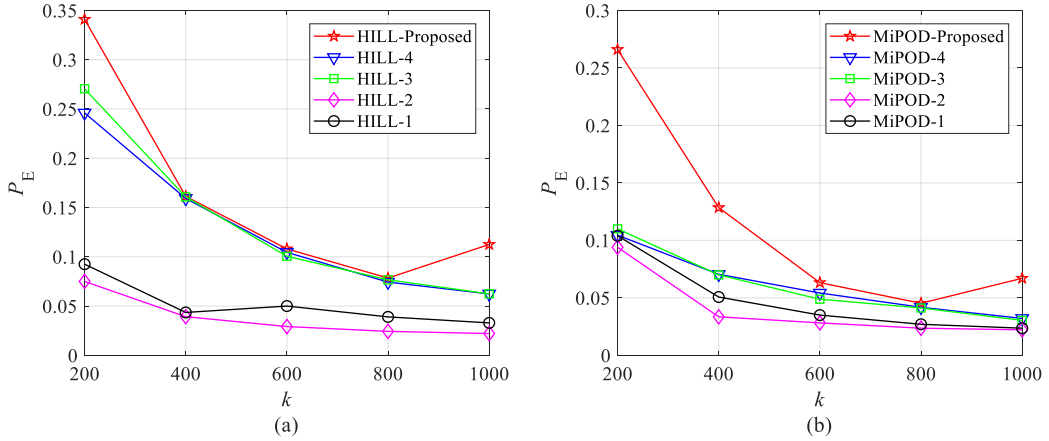

Figure 6. Undetectability comparisons of cover selection methods against steganalytic tool maxSRMd2 with embedding schemes (**a**) HILL and (**b**) MiPOD.

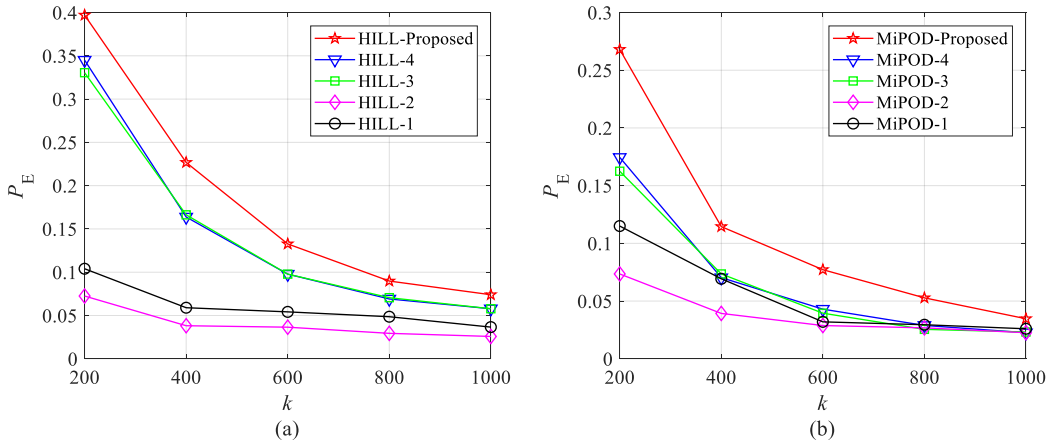

Figure 7. Undetectability comparisons of cover selection methods against steganalytic tool TLBP with embedding schemes (**a**) HILL and (**b**) MiPOD.

In most cases, our method exhibits a higher level of undetectability compared to other cover selection schemes. Compared with the scheme in [22], our method shows average improvements of 2.37% and 2.77% on P_E against SPAM and SRMQ1, respectively. Against maxSRMd2 and TLBP, the improvements are 4.21% and 3.94%. Similarly, when compared

to the scheme introduced in [21], our method demonstrates average improvements of 2.17%, 2.79%, 4.00%, and 4.20% on P_E against SPAM, SRMQ1, maxSRMd2, and TLBP, respectively. Furthermore, Compared with schemes in [19,20], our method achieves corresponding average improvements of 17.75%, 9.68%, 8.73%, 8.93%, 19.35%, 11.67%, 9.79%, and 10.74% against SPAM, SRMQ1, maxSRMd2, and TLBP, respectively. The observed improvements in our method are justified as it is specifically tailored for an image set comprising both nature images and metaverse images, distinguishing it from other methods. Furthermore, it is evident that our method, along with the schemes proposed in [21,22], exhibits significantly higher undetectability compared to the remaining two schemes. This distinction can be attributed to the fact that the schemes presented in [19,20] are primarily designed for specific scenarios.

Our method adopts a distortion minimization framework that employs pre-defined distortion functions, such as MiPOD, HILL, SUNIWARD, and WOW, which are widely recognized steganographic paradigms. These distortion functions represent the most popular approaches in the field. Furthermore, our approach exhibits substantial efficacy when employed in diverse steganographic frameworks that employ a consistent distortion function, thereby ensuring its adaptability across a wide range of scenarios.

5. Conclusions

This paper presents a novel approach to cover selection for steganography, targeting a practical scenario whereby available images consist of both nature and metaverse images. The proposed method focuses on assessing the suitability of a candidate image for steganography based on its undetectability against steganalysis. Experimental results demonstrate the effectiveness of this approach, as it surpasses existing cover selection schemes in terms of undetectability when evaluated using modern steganalytic tools. Future research endeavors could explore the development of more comprehensive cover selection frameworks that cater to different image types, such as spatial images and JPEG images, which vary in their data processing characteristics and require specialized approaches.

Author Contributions: Conceptualization, X.L.; methodology, X.L. and D.G.; software, X.L. and D.G.; validation, X.L. and D.G.; formal analysis, X.L. and C.Q.; investigation, X.L. and C.Q.; resources, C.Q.; data curation, D.G.; writing—original draft preparation, X.L.; writing—review and editing, C.Q.; visualization, X.L. and D.G.; supervision, C.Q.; project administration, X.L. and D.G.; funding acquisition, C.Q. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by Natural Science Foundation of China under Grants 62172280, U20B2051, 62376148, and 62002214 and supported in part by the Chenguang Program of Shanghai Education Development Foundation and Shanghai Municipal Education Commission under Grant 22CGA46.

Data Availability Statement: Not applicable.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Wang, Z.; Feng, G.; Qian, Z.; Zhang, X. JPEG steganography with content similarity evaluation. *IEEE Trans. Cybern.* **2023**, *53*, 5082–5093.
2. Hu, M.; Wang, H. Image steganalysis against adversarial steganography by combining confidence and pixel artifacts. *IEEE Signal Process. Lett.* **2023**, *30*, 987–991.
3. Wang, Z.; Feng, G.; Zhang, X. Repeatable data hiding: Towards the reusability of digital images. *IEEE Trans. Circuits Syst. Video Technol.* **2022**, *32*, 135–146.
4. Filler, T.; Judas, J.; Fridrich, J. Minimizing additive distortion in steganography using syndrome-trellis codes. *IEEE Trans. Inf. Forensics Secur.* **2011**, *6*, 920–935.
5. Li, W.; Zhang, W.; Li, L.; Zhou, H.; Yu, N. Designing near-optimal steganographic codes in practice based on polar codes. *IEEE Trans. Commun.* **2020**, *68*, 3948–3962.
6. Holub, V.; Fridrich, J. Designing steganographic distortion using directional filters. In Proceedings of the 2012 IEEE International Workshop on Information Forensics and Security (WIFS), Costa Adeje, Spain, 2–5 December 2012; pp. 234–239.

7. Holub, V.; Fridrich, J.; Denemark, T. Universal distortion function for steganography in an arbitrary domain. *EURASIP J. Inf. Secur.* **2014**, *2014*, 1.
8. Li, B.; Wang, M.; Huang, J.; Li, X. A new cost function for spatial image steganography. In Proceedings of the 2014 IEEE International Conference on Image Processing (ICIP), Paris, France, 27–30 October 2014; pp. 4206–4210.
9. Sedighi, V.; Cogranne, R.; Fridrich, J. Content-adaptive steganography by minimizing statistical detectability. *IEEE Trans. Inf. Forensics Secur.* **2016**, *11*, 221–234.
10. Guo, L.; Ni, J.; Shi, Y. Uniform embedding for efficient jpeg steganography. *IEEE trans. Inf. Forensics Secur.* **2014**, *9*, 814–825.
11. Guo, L.; Ni, J.; Su, W.; Tang, C.; Shi, Y. Using statistical image model for jpeg steganography: Uniform embedding revisited. *IEEE Trans. Inf. Forensics Secur.* **2015**, *10*, 2669–2680.
12. Su, W.; Ni, J.; Li, X.; Shi, Y. A new distortion function design for JPEG steganography using the generalized uniform embedding strategy. *IEEE Trans. Circuits Syst. Video Technol.* **2018**, *28*, 3545–3549.
13. Kharrazi, M.; Sencar, H.; Memon, N. Cover selection for steganographic embedding. In Proceedings of the 2006 International Conference on Image Processing, Atlanta, GA, USA, 8–11 October 2006; pp. 117–120.
14. Evsutin, O.; Kokurina, A.; Meshcheryakov, R. Approach to the selection of the best cover image for information embedding in JPEG images based on the principles of the optimality. *J. Decis. Syst.* **2018**, *27* (Suppl. 1), 256–264.
15. Sajedi, H.; Jamzad, M. Cover selection steganography method based on similarity of image blocks. In Proceedings of the IEEE 8th International Conference on Computer and Information Technology, Sydney, Australia, 8–11 July 2008; pp. 379–384.
16. Sajedi, H.; Jamzad, M. Using contourlet transform and cover selection for secure steganography. *Int. J. Inf. Secur.* **2010**, *9*, 337–352.
17. Subhedar, M.; Mankar, V. Curvelet transform and cover selection for secure steganography. *Multimed. Tools Appl.* **2018**, *77*, 8115–8138.
18. Wu, S.; Liu, Y.; Zhong, S.; Liu, Y. What makes the stego image undetectable? In Proceedings of the 7th International Conference on Internet Multimedia Computing and Service, Zhangjiajie, China, 19–August 2015; p. 47.
19. Wang, Z.; Zhang, X. Secure cover selection for steganography. *IEEE Access* **2019**, *7*, 57857–57867.
20. Wang, Z.; Zhang, X.; Qian, Z. Practical cover selection for steganography. *IEEE Signal Process. Lett.* **2020**, *27*, 71–75.
21. Wang, Z.; Zhang, X.; Yin, Z. Joint cover-selection and payload-allocation by steganographic distortion optimization. *IEEE Signal Process. Lett.* **2018**, *25*, 1530–1534.
22. Wang, Z.; Feng, G.; Shen, L.; Zhang, X. Cover selection for steganography using image similarity. *IEEE Trans. Dependable Secur. Comput.* **2023**, *20*, 2328–2340.
23. Liu, J.; Jiao, G.; Sun, X. Feature passing learning for image steganalysis. *IEEE Signal Process. Lett.* **2022**, *29*, 2233–2237.
24. Megıas, D.; Lerch-Hostalot, D. Subsequent embedding in targeted image steganalysis: Theoretical framework and practical applications. *IEEE Trans. Dependable Secur. Comput.* **2022**, 20, 1403–1421.
25. Pevny, T.; Bas, P.; Fridrich, J. Steganalysis by subtractive pixel adjacency matrix. In Proceedings of the 11th ACM Workshop on Multimedia and Security, Princeton, NJ, USA, 7–8 September 2009; pp. 75–84.
26. Fridrich, J.; Kodovsky, J. Rich models for steganalysis of digital images. *IEEE Trans. Inf. Forensics Secur.* **2012**, *7*, 868–882.
27. Denemark, T.; Sedighi, V.; Holub, V.; Cogranne, R.; Fridrich, J. Selection-channel-aware rich model for steganalysis of digital images. In Proceedings of the 2014 IEEE International Workshop on Information Forensics and Security (WIFS), Atlanta, GA, USA, 3–5 December 2014; pp. 48–53.
28. Li, B.; Li, Z.; Zhou, S.; Tan, S.; Zhang, X. New steganalytic features for spatial image steganography based on derivative filters and threshold LBP operator. *IEEE Trans. Inf. Forensics Secur.* **2017**, *13*, 1242–1257.
29. Kodovsky, J.; Fridrich, J.; Holub, V. Ensemble classifiers for steganalysis of digital media. *IEEE Trans. Inf. Forensics Secur.* **2012**, 7, 432–444.
30. Ye, J.; Ni, J.; Yi, Y. Deep Learning Hierarchical Representations for Image Steganalysis. *IEEE Trans. Inf. Forensics Secur.* **2017**, *12*, 2545–2557.
31. Boroumand, M.; Chen, M.; Fridrich, J. Deep Residual Network for Steganalysis of Digital Images. *IEEE Trans. Inf. Forensics Secur.* **2019**, *14*, 1181–1193.
32. Zhang, X.; Zhang, X.; Feng, G. Image Steganalysis Network Based on Dual-Attention Mechanism. *IEEE Signal Process. Lett.* **2023**, *30*, 1287–1291.
33. Cachin, C. An information-theoretic model for steganography. In *International Workshop on Information Hiding*; Springer: Berlin/Heidelberg, Germany, 1998; pp. 306–318.
34. Schaefer, G.; Stich, M. UCID: An uncompressed color image database. In *Storage and Retrieval Methods and Applications for Multimedia 2004*; International Society for Optics and Photonics: Bellingham, WA, USA, 2003; volume 5307, pp. 472–480.
35. NFTCN. 2021. Available online: https://www.nftcn.com/pc/#/index (accessed on 1 September 2023).