LI *Journal of Engineering and Applied Science* (2025) 72:206
https://doi.org/10.1186/s44147-025-00785-8

Journal of Engineering and Applied Science

RESEARCH Open Access

Imbalanced learning using the area under the curve and proximal support vector machine for image steganalysis

Xinpan LI[1*]

*Correspondence:
Xinpan LI
xinpanli67@gmail.com
[1]Department of Information Engineering, Hebei Chemical & Pharmaceutical College, Shijiazhuang City, Hebei Province 050000, China

Abstract
The proposed research introduces a novel steganalytic tactic termed the Imbalanced Maximizing-AUC Proximal Support Vector Machine (PSVM). This method strengthens detection performance in the presence of imbalanced datasets by integrating AUC maximization into the PSVM framework. In doing so, it directly addresses one of the major challenges in steganalysis—class imbalance—while reducing the reliance on extensive hyperparameter tuning, thereby improving model performance when imbalance exists. Theoretically, the approach retains the key advantages of PSVMs, including fast incremental updates, making it well-suited for scenarios requiring rapid and adaptive adjustments. In parallel, an alternative version of the Differential Evolution (DE) scheme is introduced, featuring a novel mutation scheme based on k-means clustering to ensure effective hyperparameter optimization. This mechanism provides resilience and adaptability across diverse conditions. Empirical evaluation on standard databases—BossBase 1.01 and BOWS-2—reveals substantial improvements, achieving F-measure scores of 89.86% and 91.55%, respectively, surpassing existing steganalysis methods. Overall, the proposed approach marks a significant advancement in addressing class imbalance and optimizing detection efficiency, establishing a strong benchmark for future research in image steganalysis.

Keywords Image steganalysis, PSVM, Area under the curve (AUC), K-means clustering, DE

Introduction

The World Wide Web contains images that represent an electronic daybook of personal experience and convey feelings and fulfill various interests. Regrettably, image steganography technologies have given criminals a powerful tool for embedding sensitive messages into these harmless pictures. They are crafted to look like accidental noise from camera sensors and electronics, and are impossible to spot with the naked eye [1]. This significantly hampers public safety. Researchers are creating high-tech image steganalysis tactics to determine and disable these concealed communications and, thus, minimize the threats these pose to civilization [2].

Springer Open

© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Image steganography hides information inside an image, thus performing an important role in safe communication [3]. It uses an embedding technique employing an advanced algorithm for the insertion of secret messages into visual content with minimal degradation or visual artifacts. It assured the privacy of the concerned data. Hidden data gets secretly embedded with images, assuring the normal look for an image. Such a facility needs to be provided for safe data transmission and digital watermarking, where the confidentiality of sensitive data is paramount, signifying its importance in today's information security framework [4].

Steganalysis tactics are specifically crafted to identify, alter, or erase concealed data within a stego object. These tactics are broadly classified into two groups [5]: One targets recognition of the presence or absence of concealed data within digital media, and second, decoding the actual content of this concealed data. Functionally, there are four categories of steganalysis tactics: statistical analysis [6, 7], transform domain examination [8, 9], blind/universal steganalysis [10, 11], and spread-spectrum steganalysis [12, 13]. Statistical analysis seeks to identify the statistical attributes of an image, including pixel dispersion and pixel-to-pixel correlation, to detect anomalies typically found in steganography. Transform domain investigation searches for variations in image domains, that are altered through the Fourier or Discrete Cosine Transforms [14] to identify image coefficient shifts due to steganographic processes. Blind/universal steganalysis utilizes machine learning (ML) and statistical tactics for detecting steganographic content based on coverage without prior knowledge of embedding procedures, providing a universal solution to identify a wide variety of steganographic modifications. Ultimately, spread-spectrum steganalysis scrutinizes images where data gets spread over a large bandwidth of frequencies. It identifies minute alterations to spectral characteristics, an undertaking hindered by low signal-to-noise ratios (LSNR). They all possess different means to reveal concealed data within digital pictures, thus enhancing digital communication security measures [15].

Image steganalysis employs different tactics, including sophisticated tactics through DL and reinforcement learning [2, 16]. A major problem facing this area is data imbalance, where image databases lack equivalent numbers of steganalytic and non-steganalytic images. Imbalance results in biased outcomes and compromises the effectiveness of steganalytic detection. To obtain a balance, random resampling and SMOTE [17] are deployed. They, at times, have to be cautious with critical information and make the scheme fit the training data. While SMOTE is well-liked, its performance at boosting accuracy and, significantly, kernel-based schemes is inconsistent. Another option, cost-sensitive learning, differentiates between the weights of different classes based on misclassification error severity. However, creating an efficient cost-sensitive framework is not simple because of the complexity of cost definitions. More sophisticated measures, including AUC, perform well with imbalances than accuracy measures. Recent research focuses on creating classifiers optimized to AUC, effectively solving imbalances, and minimally requiring extensive parameter adjustment of cost-sensitive approaches. These sophisticated tactics are vital in steganalysis detection, where model performance must be strong and consistent.

PSVM distinguishes itself with its capacity for quick and sustained updating, an essential capability that enables it to adjust to changing patterns and data shifts. Such adaptability is necessary within environments where adaptability to changing data in real-time

is needed, as well as for managing large databases with changing values. The design of PSVM allows for rapid changes and incremental updates devoid of full retraining with each new batch of data, which sits well with dynamic situations. Its robustness to handle perturbations and adjust for non-linear patterns only adds to its adaptability in diverse adaptive learning situations. The quick processability and versatile functionality of PSVM are especially beneficial for tasks requiring quick and precise classifications. Besides, its compatibility with other optimization tactics only advances its versatility and effectiveness for various purposes [18].

Finding optimal hyperparameters for image steganalysis detection schemes is an essential problem [19]. Various ways, including grid search and GAs, are employed to find these important settings [20]. While grid search is suitable for small hyperparameters, it needs help with complex machine-learning procedures. Conversely, GAs perform well for traversing extensive and elaborate hyperparameter spaces, but generally demand the transformation of hyperparameters into discrete representations for efficient optimization [21]. Hyperparameter optimization is a key strategy for reducing learning errors, with metaheuristics—especially the DE scheme—being crucial. DE performs an intelligent search across large, complex spaces, evolving solutions over generations by exploiting differences between randomly selected individuals. This allows it to reach global optima without discretizing hyperparameters. Such capabilities enhance image steganalysis methods, making them more robust and accurate through efficient exploration and exploitation [22].

DE functions through three main steps: mutation, crossover, and selection. In mutation, new solutions are generated by modifying and scaling differences between existing solutions. During crossover, these mutated vectors combine with candidate solutions, introducing new genetic diversity. Finally, selection evaluates the performance of the new solutions against existing ones, retaining the most efficient candidates. Mutation is pivotal because it introduces new variations, which refresh the pool of solutions and avoid the static behavior of the optimization process. It is important to maintain the adaptability and capability of navigating the complex solution spaces of the algorithm [23].

A new image steganalysis technique, ImAUC-PSVM, improves upon the classical PSVM by favoring the maximum AUC to account for problems in class imbalance during classification. The approach uses a variable PSVM, aiding in better regulation of parameters in different applications, including fraud detection. The setting of hyperparameters is carried out based on DE, and then optimized based on k-means clustering to select its best candidate solutions, thus improving upon the effectiveness of tuning its hyperparameters. The main contributions of this technique are detailed below:

- ImAUC-PSVM for Imbalance Classification: The ImAUC-PSVM enhances detection in imbalanced databases by optimizing AUC, improving accuracy in distinguishing steganographic content amidst uneven class dispersions common in steganalysis.
- Using DE for Hyperparameter Optimization: DE is crucial in refining the scheme by accurately tuning hyperparameters, thereby boosting overall effectiveness and ensuring optimal classification accuracy.
- Improved DE with HMS Integration: Enhanced DE integrates Human Mental Search tactics to optimize hyperparameter tuning efficiently, accelerating optimal solution convergence and reducing computational load.

Research gaps and novelties

A main challenge in picture steganalysis is the issue of class imbalance inherent in popular databases, causing overcounts of non-steganographic images relative to a comparative undercount of steganographic ones. Class imbalance causes models to perform unevenly, as traditional detection methods are biased toward majority classes and often fail to identify minority classes, especially steganographic ones. Modern approaches, such as resampling and cost-sensitive learning, often deliver inconsistent results and require high computational effort. They also demand extensive manual hyperparameter tuning, making them unsuitable for real-time or dynamic environments where steganographic methods evolve rapidly. Effectively handling imbalanced databases remains a major challenge in robust image steganalysis.

This exploration proposes ImAUC-PSVM, which maximizes AUC during training to improve performance on imbalanced databases while reducing reliance on heavy hyperparameter tuning. PSVM is enhanced with an advanced DE optimization method featuring a novel k-means-based mutation strategy. It tackles both class imbalance and hyperparameter optimization, boosting accuracy without the computational burden of traditional techniques. Overall, ImAUC-PSVM significantly enhances detection accuracy and computational efficiency in image steganalysis.

Paper organization

The structure of this article is: Segment 2 offers an in-depth review of existing literature on the subject. Segment 3 describes our recommended methodology and the core tactics utilized, while Segment 4 discusses the outcomes of our empirical experiments and their analyses. Segment 5 provides a recapitulation of our findings and offers guidance for future exploration.

Related works

This part is partitioned into two key parts: the first concentrates on the study of steganography, and the second reviews advancements in steganalysis.

Steganography

Image steganography practices embed covert communications into visuals, maintaining a normal appearance. Even when modifications are slightly noticeable, the underlying concealed content remains undetected, preserving the ordinary photo illusion. Presently, image steganography encompasses three main types: basic steganography [24, 25], which utilizes straightforward embedding tactics; adaptive steganography [26, 27], adjusting the embedding tactics according to the visual characteristics, and AI-enhanced steganography [28, 29], applying sophisticated AI tactics to refine the embedding procedures and better mask the information.

Basic steganography approaches, while straightforward and engaging on digital media, can cause discernible image alterations. Essa et al. [30] introduced an advanced variant of classic steganography known as Multibit Least Significant Bit Matching (MLSBM) Steganography. This strategy enhances the LSBM technique by inserting multiple data bits into each pixel of every channel. MLSBM Steganography retains the construction of the initial image, thus lowering the discernibility of the data versus other tactics that insert a comparable volume of data. This improvement tackles the typical balance

between data capacity and the risk of detection in image steganography. Nashat et al. [31] recommended an enhanced image steganography algorithm containing additional data and enhancing the stego image's quality. Combining an LSB adjustment with the Hough Transform improves data concealment effectiveness. A sophisticated edge-detection filter first identifies regions around image borders and applies LSB reversal to preserve image integrity, then modifies LSBs in smooth areas to boost data embedding stealth and capacity. This approach aims to maximize payload while minimizing visual distortion.

Adaptive steganography, valued for its superior masking, embeds secret messages in complex image regions using advanced methods, including Syndrome Trellis Codes (STCs) to mix data seamlessly with visual content. Pramanik et al. [32] suggested an improved steganographic process that mixes spatial flexible and wavelet-oriented tactics for embedding data into digital media. It employs an edge intensity criterion and high-frequency sub-bands of wavelet transforms, which are boosted with a genetic algorithm (GA), to identify ideal embedding coefficients. It boosts the peak signal-to-noise ratio (PSNR) of the stego image to make it resistant to different types of steganalysis attacks. Xie et al. [33] suggested an unparalleled post-processing technique for flexible image steganography known as post-cost-enhancement, which adjusts embedding cost without training for generative adversarial networks (GANs). It employs a pre-trained CNN to produce a gradient map for a cover image, which is subsequently smoothed and employed to refine the steganography cost map, optimally creating stego images. Zhang et al. [34] developed an innovative technique called the Joint Adaptive Robust Steganography Network (JARS-Net). This network uses a hierarchical attentive invertible (HAI) mechanism for flexible attribute tuning across multiple depths and scales, combined with adaptive key learning (AKL) to produce responsive keys that enhance secret recovery even under distortions. This method improves the secret image retrieval from stego images, enhancing both concealment and resilience of the communication.

DL in steganography is divided into four distinct areas: (1) synthesis-driven tactics [35, 36], (2) creation of probability maps [37, 38], (3) evasion of CNN-oriented steganalysis [39], and (4) tripartite game strategies [40]. Li et al. [35]crafted an innovative technique in coverless image steganography that implants a color secret into another color stego of identical size without utilizing covering layers. This tactic is powered by two sub-networks: SynthesisNet, which translates between images, and RevealNet, which recovers the hidden image. SynthesisNet bridges cross-domain gaps between source and target images, tailoring stego imagery to specific semantics and styles, thus boosting its genuineness and lowering detection probabilities. Huang et al. [36]recommended a novel synthesis-mapping architecture named Steganography Without Embedding (SWE), which addresses prevalent challenges. It integrates a disentanglement auto-encoder for generating images with a mapping module that leverages block statistics hash matching to conceal compression flaws, decreasing reliance on multiple carrier images, and enhancing the precision of hidden message retrieval. This tactic melds synthesis and mapping tactics, honing payload capacity and stealth. Huo et al. [37]revealed the Chaotic Mapping-Enhanced Image Steganography Network (CHASE), a pioneering method for embedding several color images into a single grayscale image. Utilizing chaotic mapping and image permutation, this strategy reduces perceptual variances between container and cover images, thus fortifying steganographic protection. CHASE integrates

GANs to keep high image integrity, balance security with image quality, and exhibit robust defense against steganalysis. Rubaie et al. [38]introduced dual steganography tactics—lossless and lossy—by merging cryptographic practices with steganography, employing three unique chaotic maps to boost data protection. In the cryptographic stage, data is encrypted using two chaotic maps, and during the steganographic phase, another map facilitates embedding this encrypted data into the trailing bits of a double-precision image's pixels. This tactic deploys the large bit redundancy in 64-bit per pixel grayscale images to ensure secure, high-capacity data hiding. Amit et al. [39]developed Adversarial Embedding (ADV-EMB) to advance image steganography against ML-based steganalysis. This tactic delicately alters image pixels through gradients back-propagated from a CNN steganalyzer, designed to confound it while preserving data integrity. ADV-EMB modulates pixel changes utilizing the probability of reversing gradient directions, producing steganographic images that skillfully avoid detection by sophisticated CNN-oriented steganalysis, boosting the security of the data against refined analytical assaults. Kumar and Soundrapandiyan et al. [40]recommended improving video steganography through a cooperative game theory model. This approach divides cover videos, preprocesses secret images, and integrates these images into Regions of Interest (ROI) pinpointed by motion vectors. A cooperative three-player game theory model, deploying the Iterative Elimination of Strictly Dominant Strategies (IESDS) technique, is used to balance the trade-offs among invisibility, payload capacity, and robustness, substantially elevating the video steganography system's quality and security.

Steganalysis

Recent innovations in DL and NNs have transformed the field of image steganalysis, enhancing the ability to detect hidden messages. Unlike conventional tactics that rely on manually chosen attributes, CNN-oriented tactics schemes include backpropagation to automatically extract attributes. Bravo-Ortiz et al. [41] created a convolutional vision transformer specifically for spatial domain image steganalysis. This advanced scheme mixes convolutional layers with attention mechanisms to effectively identify localized and widespread steganographic anomalies. The process includes three key steps: preprocessing, which employs trainable and constant filters to emphasize steganographic irregularities; noise analysis, boosted with Squeeze-and-Excitation blocks to increase sensitivity; and a categorization step that deploys both convolutional and transformer structures to meticulously assess steganographic patterns. Zhou et al. [42] designed a specialized NN for medical image steganalysis, which is especially vulnerable during transfer between Picture Archiving and Communication Systems (PACS) or telemedicine sessions. It starts with extracting image region-dependent specific global texture attributes, followed by multi-head self-attention and deep convolutional blocks to uncover hidden steganographic content, hence achieving enhanced convergence and improved accuracy, even at low embedding rates. Jeyaprakash et al. [43] suggested a three-phase steganalysis algorithm to identify stego images and cover images with visual appearance not compromised. The approach begins with a preprocessing step to boost image clarity, then dimensionality reduction via attribute derivation with the Mustard Honey Bee enhancement scheme, and finishes with an HSVGG-based CNN for classification. The technique harnesses state-of-the-art DL technologies to address steganography's persistent challenges and those faced by steganalysis. Ntivuguruzwa et al. [44]

presented an innovative CNN design to boost the hidden data recognition in digital media and to undertake regular training for various image types. The algorithm is split into three steps: preprocessing, which employs spatial rich scheme filters to boost noise signals of concealed data; attribute derivation, which employs 2D depthwise separable convolutions to increase SNR, and traditional convolutions to capture attributes locally; and categorization, which employs multi-scale average pooling to efficiently pool attributes, trailing three fully linked layers to map these attribute maps into class likelihood via softmax function. The organized technique significantly improves detection accuracy and reduces training time. Vijjapu et al. [45] an in-depth steganalysis method to counter advanced steganography deployed in illegal activities, such as terrorism. Their approach effectively extracts and analyzes hidden data within digital images, identifying concealed content irrespective of the steganography technique. It was rigorously tested against tools like Xiao and OpenPuff, showing strong performance for both grayscale and color images. Liu et al. [46] introduced an advanced image steganalysis method combining attention mechanisms and transfer learning to boost attribute retrieval and recognition, especially for images with low embedding densities. The tactic uses CNN with transposed convolution, regular convolution, fully connected, and preprocessing layers. An efficient channel attention module follows the regular convolution layer to concentrate on steganographic regions and adaptively modify weights for fine-tuning features. Transfer learning is utilized to transfer attributes with high embedding densities to low embedding densities with an optimization of steganalysis. Butora et al. [1] reviewed extensions to a DL detector for image steganography identification with emphasis on how the structures of an image are impacted by JPEG compression. From their research, it was observed that, under controlled conditions, the logits of a classifier follow a Gaussian dispersion with a linear dependence upon image size. With the removal of padding inside the convolutional NN, it was possible to stabilize the mean of the dispersion of logs for different image sizes. The development sets a theoretically stable false positive threshold for image dimensions and significantly improves the detector's performance for non-adaptive as well as content-adaptive steganographic tactics.

Several explorations have identified the potential of Deep Reinforcement Learning (DRL) for steganalysis [3]. Sun et al. [47]recommended an advanced image steganalysis scheme incorporating dilated convolution tactics with a Mutual Learning-based Artificial Bee Colony (ML-ABC) framework and reinforcement learning. The scheme employs convolutional NNs for attribute derivation and applies reinforcement learning to correct databases' imbalances. The ML-ABC technique is employed to optimize initial weights' training, thereby increasing potential solution quality. Al-Obaidi et al. [48]introduced a sophisticated image steganalysis method incorporating enhanced reinforcement learning strategies and real-time data augmentation. Their scheme attributes triple parallel dilated convolutions that improve and integrate feature vectors for categorization. The refined reinforcement learning framework adeptly manages database imbalances by viewing data samples as states in decision-making sequences, with the network operating as the decision agent. Moreover, through GAN-generated images and a regulatory mechanism, data augmentation helps overcome challenges associated with Generative Adversarial Networks (GANs), thereby enhancing the classification process. Wang et al. [9] created an RL technique specifically for linguistic steganalysis that was meant to tackle execution degradation resulting from dispersion shifts. It positions an agent (step

analyzer) in a GloVe observation space and utilizes Actor and Critic modules. The initial training step improves these modules with the primary database, while the subsequent refining step applies reinforcement training on modified data to improve the attribute derivation capabilities and provide an autonomous training framework.

Despite advances in image steganalysis, issues like data imbalance and sensitivity to hyperparameter settings remain prevalent. An innovative solution is proposed that capitalizes on both the strength of the PSVM and the adaptability of the DE scheme. The approach is specifically designed for accurate hyperparameter adjustment to make existing schemes more effective and to provide a sophisticated solution for various real-world problems.

The recommended scheme

In this exploration, an advanced image steganalysis technique based on AUC maximization learning is recommended that can efficiently handle class imbalances and incorporates an enhanced DE scheme for accurate hyperparameter optimization.

For several convincing reasons, the PSVM classifier was selected for the AUC maximization learning approach. In contrast to traditional SVMs, which have to solve a computationally expensive quadratic programming (QP) problem, the PSVM solves a simpler and more stable convex optimization problem. This simplification lends itself to an easier analytical solution, with low computational overhead and without compromise in predictive accuracy. Previous research [49, 50] has also noted PSVM's capability to enable incremental learning with streaming data, a precious attribute for learning under evolving environments. This advantage stems from PSVM not requiring frequent, heavy matrix inversions, enabling efficient and dynamic learning. Its architecture complements the AUC measure, with a surrogate loss function that closely estimates AUC, allowing effective AUC-maximizing training.

Using DE for hyperparameter optimization is justified by its strengths. DE efficiently navigates smooth parameter spaces, crucial for fine-tuning complex hyperparameters. Its ability to explore wide, challenging regions is key in image steganalysis, where hyperparameter choices strongly affect performance. DE also avoids common pitfalls, including getting stuck in local optima, and adapts during evolution, making it an excellent tool to boost the effectiveness and resilience of the detection scheme in real-world scenarios.

ImAUC-PSVM

AUC is identified as a more consistent measure than accuracy for measuring performance in imbalanced class dispersion databases. AUC is not heavily impacted by the skewness of class proportions, giving a more consistent measure of model performance. Statistically, AUC is the probability that a haphazardly chosen positive instance (x^+) of the majority group will be ranked higher than a negative instance (x^-) of the minority group. It is a good measure of how well the classifier can discriminate between classes, and thus its importance is regarded in situations where clear class discrimination is of utmost importance. The formula for AUC computation is:

$$AUC\,(s) = E_{x^+ \sim D^+} E_{x^- \sim D^-} 1(s\left(x^+\right) > s(x^-)) \tag{1}$$

Here, D^+ is used to refer to the dispersion of the majority class and D^- for minority class dispersion. The scoring function employed by the classifier, $s\,(x)$, is generally

expressed as $s(x) = \begin{pmatrix} \phi(x) \\ 1 \end{pmatrix}$ w′, with $\phi(x)$ converting input attributes and w' being the classifier's weight vector. The term E stands for mathematical expectation, and $1(\cdot)$ is the indicator function, which returns one if a specified condition is met and 0 otherwise. Due to the discontinuous nature of the indicator function, it is replaced with an ongoing and convex surrogate loss function, given by $(1 - \acute{w}^T(\begin{pmatrix} \phi(x^+) \\ 1 \end{pmatrix} - \begin{pmatrix} \phi(x^-) \\ 1 \end{pmatrix}))^2$. This substitution facilitates the creation of an empirical version, termed R $R_{AUC}(s)$ for $AUC(s)$, based on the selected scoring function.

$$\begin{aligned} R_{AUC}(s) =& \frac{1}{N^+}\sum_{i=1}^{N^+}\left(\acute{w}^T\left(c^+ - \left(\begin{smallmatrix}\phi(x_i^+)\\1\end{smallmatrix}\right)\right)\right)^2 + \frac{1}{N^-}\sum_{j=1}^{N^-}\left(\acute{w}^T\left(\left(\begin{smallmatrix}\phi(x_j^-)\\1\end{smallmatrix}\right) - c^-\right)\right)^2 \\ &+ \frac{1}{2N^+}\sum_{i=1}^{N^+}\left(\acute{w}^T\left(1 - \left(\left(\begin{smallmatrix}\phi(x_i^+)\\1\end{smallmatrix}\right) - c^-\right)\right)\right)^2 \\ &+ \frac{1}{2N^-}\sum_{j=1}^{N^-}\left(\acute{w}^T\left(1 - \left(c^+ - \left(\begin{smallmatrix}\phi(x_j^-)\\1\end{smallmatrix}\right)\right)\right)\right)^2 \end{aligned} \tag{2}$$

$$c^+ = E_{x^+\sim D^+}\begin{pmatrix} \phi(x^+) \\ 1 \end{pmatrix} \tag{3}$$

$$c^- = E_{x^-\sim D^-}\begin{pmatrix} \phi(x^-) \\ 1 \end{pmatrix} \tag{4}$$

$$N = N^+ + N^- \tag{5}$$

N^+ and N^- indicate the overall count of samples in the majority and minority groups, accordingly. c^+ and c^- represent the expected values of the feature mappings for samples from the minority and majority groups, accordingly. Utilizing the previously described surrogate loss function, the goal of maximizing $AUC(s)$ is transformed into a task of minimizing $R_{AUC}(s)$ [51]. Similarly, AUC(s) is a primary focus, $R_{AUC}(s)$ is integrated into our precision-driven formula, leading to the following equation [52]:

$$\min_{\acute{w}} \frac{1}{2}\acute{w}^T\acute{w} + \gamma R_{AUC}(s) \tag{6}$$

In this formulation, γ displays the chosen balancing hyperparameter. By integrating Eq. 2 into Eq. 6, the enhancement issue for the categorizer is established below:

$$\begin{aligned} &\min_{\acute{w}} \frac{1}{2}\acute{w}^T\left(I + \frac{2\gamma}{N^+}\sum_{i=1}^{N^+}\left(c^+ - \left(\begin{smallmatrix}\phi(x_i^+)\\1\end{smallmatrix}\right)\right)\left(c^+ - \left(\begin{smallmatrix}\phi(x_i^+)\\1\end{smallmatrix}\right)\right)\right)^T \\ &+ \left(\frac{2\gamma}{N^-}\sum_{j=1}^{N^-}\left(\left(\begin{smallmatrix}\phi(x_j^+)\\1\end{smallmatrix}\right) - c^-\right)\left(\left(\begin{smallmatrix}\phi(x_j^+)\\1\end{smallmatrix}\right) - c^-\right)\right)^T \\ &\acute{w} + \frac{\gamma}{2N^+}\sum_{i=1}^{N^+}(1 - \acute{w}^T(\left(\begin{smallmatrix}\phi(x_i^+)\\1\end{smallmatrix}\right) - c^-))^2 + \frac{\gamma}{2N^-}\sum_{j=1}^{N^-}(1 - \acute{w}^T(c^+ - \left(\begin{smallmatrix}\phi(x_j^-)\\1\end{smallmatrix}\right))^2 \end{aligned} \tag{7}$$

I denotes the identity matrix, and T signifies the transpose operation. Using the following formulas:

$$A = I + \frac{2\gamma}{N^-}\sum_{j=1}^{N^-}\left(\begin{pmatrix}\phi\left(x_j{}^-\right)\\1\end{pmatrix} - c^-\right)\left(\begin{pmatrix}\phi\left(x_j{}^-\right)\\1\end{pmatrix} - c^-\right)^T + \frac{2\gamma}{N^+}\sum_{i=1}^{N^+}\left(c^+ - \begin{pmatrix}\phi\left(x_i{}^+\right)\\1\end{pmatrix}\right)\left(c^+ - \begin{pmatrix}\phi\left(x_i{}^+\right)\\1\end{pmatrix}\right)^T \tag{8}$$

$$\left(\begin{pmatrix}\acute{\phi}\ (x_i)\\0\end{pmatrix}, \acute{y}_i\right) = \begin{cases}\left(\sqrt{\frac{1}{2p}}\left(\begin{pmatrix}\phi\,(x_i)\\1\end{pmatrix} - c^-\right), \sqrt{\frac{1}{2p}}\right), & y_i = +1\\ \left(\sqrt{\frac{1}{2\,(1-p)}}\left(\begin{pmatrix}\phi\,(x_i)\\1\end{pmatrix} - c^+\right), -\sqrt{\frac{1}{2\,(1-p)}}\right), & y_i = -1\end{cases} \tag{9}$$

$$p = \frac{N^+}{N^+ + N^-} \tag{10}$$

Equation 7 is rewritten as follows:

$$\min_{\acute{w}} \frac{1}{2}\,\acute{w}^T A\,\acute{w} + \frac{\gamma}{2N}\sum_{i=1}^{N}(\acute{y}_i - \acute{w}^T\begin{pmatrix}\acute{\phi}\ (x_i)\\0\end{pmatrix})^2 \tag{11}$$

Where $\acute{y}_i$ is the original label of the i-th sample. The solution is derived straightforwardly:

$$\acute{w} = \acute{Z}^T(\acute{Z}\acute{Z}^T + \frac{N}{\gamma}A)^{-1}\acute{Y} \tag{12}$$

$$\acute{Z} = \left[\left(\phi'\,(x_1)^T, 0\right); (\phi'\,(x_2)^T, 0), \ldots, (\phi'\,(x_N)^T, 0)\right] \tag{13}$$

$$\acute{Y} = [y_1', y_2', \ldots, y_N'] \tag{14}$$

Where $\acute{Z}$ and A can be precomputed. If $\phi\ (x_i)$ is not directly known, the kernel matrix K is produced using the training database. By decomposing K, i.e., $K = (\lambda^{\frac{1}{2}}\,\mathrm{p})^T(\lambda^{\frac{1}{2}}\,\mathrm{p})$, solutions for $\phi\ (x_i)$, $(i = 1, \ldots, N)$ are identified, and then determine $\phi\prime\ (x_i)$, $\acute{Z}$ and A For any test sample x, its decision function is defined below:

$$\begin{aligned} f\,(x) &= \begin{pmatrix}\phi\ (x)\\1\end{pmatrix}^T\acute{Z}^T\left(\acute{Z}\acute{Z}^T + \frac{N}{\gamma}A\right)^{-1}\\ \acute{Y} &= \begin{pmatrix}\phi\ (x)\\1\end{pmatrix}^T\left(\begin{pmatrix}\acute{\phi}\ (x_1)\\0\end{pmatrix}, \begin{pmatrix}\acute{\phi}\ (x_2)\\0\end{pmatrix}, \ldots, \begin{pmatrix}\acute{\phi}\ (x_N)\\0\end{pmatrix}\right)\left(\acute{Z}\acute{Z}^T + \frac{N}{\gamma}A\right)^{-1}\\ \acute{Y} &= (\phi\ (x)^T\acute{\phi}\ (x_1), \phi\ (x)^T\acute{\phi}\ (x_2), \ldots, \phi\ (x)^T\acute{\phi}\ (x_N))\left(\acute{Z}\acute{Z}^T + \frac{N}{\gamma}A\right)^{-1}\acute{Y}\end{aligned} \tag{15}$$

Each $\phi\ (x)^T\acute{\phi}\ (x_i)(i = 1, \ldots, N)$ is gauged using Eq. 16, mixing the kernel function.

$$\phi\ (x)^T \acute{\phi}\ (x_i) = \begin{cases} \sqrt{\frac{1}{2p}} \begin{pmatrix} \phi\,(x) \\ 1 \end{pmatrix}^T \left(\begin{pmatrix} \phi\,(x_i) \\ 1 \end{pmatrix} - \left(\frac{\phi\left(x_1^-\right)+\ldots+\phi\left(x_{N-}^-\right)}{N_1^-} \right) \right), & y_i \in\ D^+ \\ \sqrt{\frac{1}{2\,(1-p)}} \begin{pmatrix} \phi\,(x) \\ 1 \end{pmatrix}^T \left(\begin{pmatrix} \phi\,(x_i) \\ 1 \end{pmatrix} - \left(\frac{\phi\left(x_1^+\right)+\ldots+\phi\left(x_{N-}^+\right)}{N_1^+} \right) \right), & y_i \in\ D^- \end{cases} \quad (16)$$

$$= \begin{cases} \sqrt{\frac{1}{2p}} \left(K\,(x, x_i) - \frac{K\left(x,x_1^-\right)+\ldots+K\left(x,x_N^-\right)}{N^-} \right), & y_i \in\ D^+ \\ \sqrt{\frac{1}{2\,(1-p)}} \left(K\,(x, x_i) - \frac{K\left(x,x_1^+\right)+\ldots+K\left(x,x_N^+\right)}{N^+} \right), & y_i \in\ D^- \end{cases}$$

In the recommended scheme, the variable x displays an image that serves as the input to the PSVM classifier. The output from the PSVM classifier, which assigns a class label to the input image, is defined by Eq. 15.

Hyperparameter optimization

Optimizing hyperparameters is crucial for ML schemes because it significantly affects their performance and efficiency. Hyperparameters are the configurable settings or parameters not learned directly from the training data. Instead, they are established before training starts and can significantly influence the behavior and success of learning algorithms. Proper hyperparameter tuning is essential for achieving the balance between under- and overfitting. Underfitting occurs when a model is too simplistic, marked by insufficient parameters, leading to poor performance on training and test data. Overfitting, in contrast, arises when a scheme is overly complex, capturing noise instead of helpful information from the training data, degrading its performance on new, unseen data. Furthermore, hyperparameter optimization impacts the speed at which a scheme learns and reaches a solution. Variables, including batch size and the number of training epochs determine the training speed and efficiency. Inappropriately selected hyperparameters can result in prolonged training times or hasten convergence to a suboptimal solution. Proper tuning of these parameters is key to enhancing model training and ensuring optimal performance [53].

Random key (RK)

This article utilizes RK [54] for hyperparameter enhancement for ImAUC-PSVM, which is selected specifically because it easily handles both categorical and continuous hyperparameters. It simplifies how different types of parameters are incorporated into one optimization process, thus enabling more consistent and efficient tunability. It also allows for an equitable treatment approach to optimize searching through different dimensions of the parameter space, and therefore, for intricate schemes where there is a lot of parameter interaction.

Random Key employs an encoding scheme with T numeric vectors, D dimensions, p_1, p_2, ..., p_T. These vectors represent a population with each member being a potential solution with an association with scheme hyperparameters via a mapping function referred to as the random key. For C hyperparameters to be optimized (as displayed in Table 1), each hyperparameter $c = 1, \ldots, C$ is represented by D_c positions in the vector. When a hyperparameter is continuous, D_c is set to 1, leading to a total dimensionality D expressed as $D = \sum_{c=1}^{C} D_c$. Each vector p_i is split into C segments, each entailing D_c positions corresponding to a hyperparameter's potential values. For categorical hyperparameters, the Random Key method applies a mapping from the numeric segment D_c of

Table 1 Outline of the hyperparameters adjusted during the investigation, with a range of values based on precedents from existing image steganalysis research

Hyperparameter	Range	Best
Iteration	32 to 1056	255
Batch size	8 to 512	35
Kernel	Polynomial, Linear, Sigmoid, RBF	RBF
Feature scaling	None, Standard, Min-Max, Robust	Standard
γ	0.01 to 1	0.25

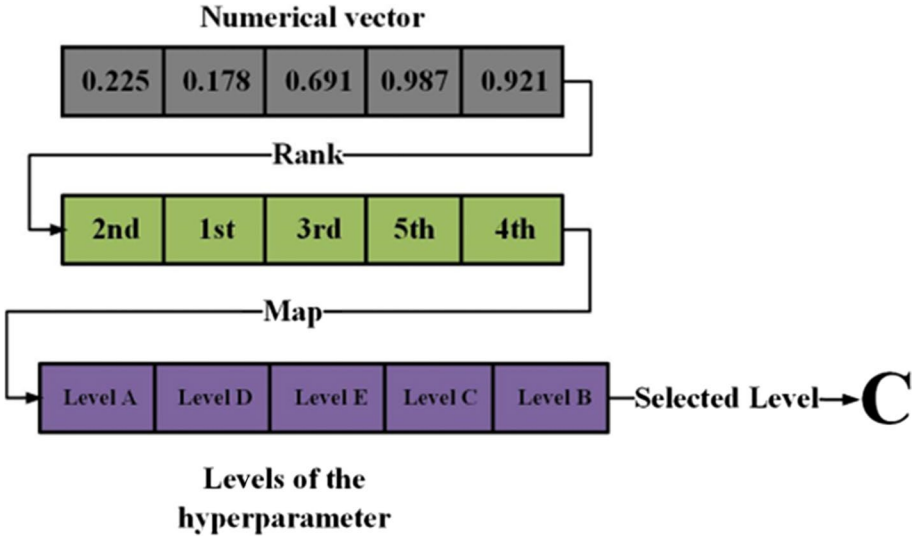

Fig. 1 Example of the random key method with $D_c = 5$

vector p_i to a corresponding categorical vector MAP_c, which enumerates the options available for the c^{th} hyperparameter. This mapping is achieved by sorting the elements within the D_c segment and deploying the position of the highest-ranking element as an index into the MAP_c vector to choose the proper categorical value. This guarantees that the outcomes of evolutionary operations like mutation, crossover, and selection, when deployed to the numeric vectors p_i, are consistently translatable back into configurations supporting categorical and uninterrupted hyperparameter values. An example of this is demonstrated in Fig. 1, where $D_c = 5$. Ultimately, RK is a vector of real numeric values that, upon ranking, plays a crucial role in aligning with a predefined array of options. This arrangement allows more significant attributes to rise to higher values within the key, while less significant attributes fall to lower ranks. Systematically laid out key structures and attributes according to priority, yielding a clear and compelling landscape for optimization.

Differential evolution

The DE scheme is leveraged to boost the effectiveness of the Random Key approach with its tested capability for global optimization. DE performs particularly well with complex, high-dimensional search spaces, a domain that fits well with our model's rich and complex hyperparameter configurations. Its capability to avert local optima and converge quickly to global optima makes DE ideal for fine-tuning our model and achieving optimal results across diverse test cases.

DE operates in three key phases: mutation, crossover, and selection. The first and most crucial phase, mutation, introduces new genetic diversity into the solution set. It modifies existing solutions using components from other population members through arithmetic and geometric operations—typically transforming a base solution by adding a weighted difference between two others to create a new variant. These mutations maintain diversity, uncover promising solutions, and prevent convergence to suboptimal points. The efficacy of DE largely relies on the mutation's capability to produce a diverse, high-quality collection of resolutions to guide the algorithm toward the optimum.

In the DE framework, the mutation process generates a new vector as follows [23]:

$$\overrightarrow{v}_{i,g} = \overrightarrow{x}_{r_1,g} + F\left(\overrightarrow{x}_{r_2,g} - \overrightarrow{x}_{r_3,g}\right) \tag{17}$$

Here, $\overrightarrow{x}_{r_1,g}$, $\overrightarrow{x}_{r_2,g}$, and $\overrightarrow{x}_{r_3}$ are three haphazardly chosen different solutions out of the existing group, with F being a scaling factor that regulates how the difference between two of them affects the third one. Subsequent to mutation, there follows a crossover stage, which unites aspects from this newly mutated vector and those of the existing target vector. It is usually done through a Binomial crossover technique:

$$u_{i,j,g} = \begin{cases} v_{i,j,g} \; if \; rand\,(0,\,1) \leq \; CR \, or \, j \; = j_{rand} \\ x_{i,j,g} \qquad\qquad\qquad\qquad otherwise \end{cases} \tag{18}$$

CR displays the crossover rate, and j_{rand} is an integer haphazardly chosen from the range $\{1,\,2,\,...,\,D\}$, with D being the dimensionality of the solution. Then, selection occurs, where the original target vector and the newly formed trial vector from the crossover are evaluated against each other. The algorithm then selects the better-performing solution to advance in the evolutionary process. This phase is crucial as it promotes the continuous enhancement of solutions within the population, ensuring only the most effective solutions are retained.

The enhanced DE variant introduces a new mutation mechanism that draws on recent advancements in the field, as detailed in the literature [55]. Initially, k-means clustering is applied to the current population to determine distinct groups within the search domain. This tactic segments the group into several clusters, with the number of clusters, k, being haphazardly determined to fall between $[2,\ \sqrt{N}]$. Focus is then directed towards the cluster that displays the smallest average objective function value, marking it as the target for deeper analysis. Figure 2 illustrates this approach, depicting 18 potential solutions divided into three clusters.

A novel mutation operator influenced by clustering is introduced, described by the following equation:

$$\overrightarrow{v^{clu}}_i = \overrightarrow{win}_g + F\left(\overrightarrow{x}_{r_1,g} - \overrightarrow{x}_{r_2,g}\right) \tag{19}$$

In this formula, $\overrightarrow{x}_{r_1,g}$ and $\overrightarrow{x}_{r_2,g}$ are two haphazardly selected candidate solutions from the group, and $\overrightarrow{win}_g$ displays the ideal resolution within a particularly promising cluster. It is important to note that $\overrightarrow{win}_g$ may not necessarily be the desired resolution across the entire group. This clustering-oriented mutation mechanism is implemented repeatedly for M iterations. Subsequently, the population is processed through several stages as outlined by the Generic Population-Based Algorithm (GPBA) [56]:

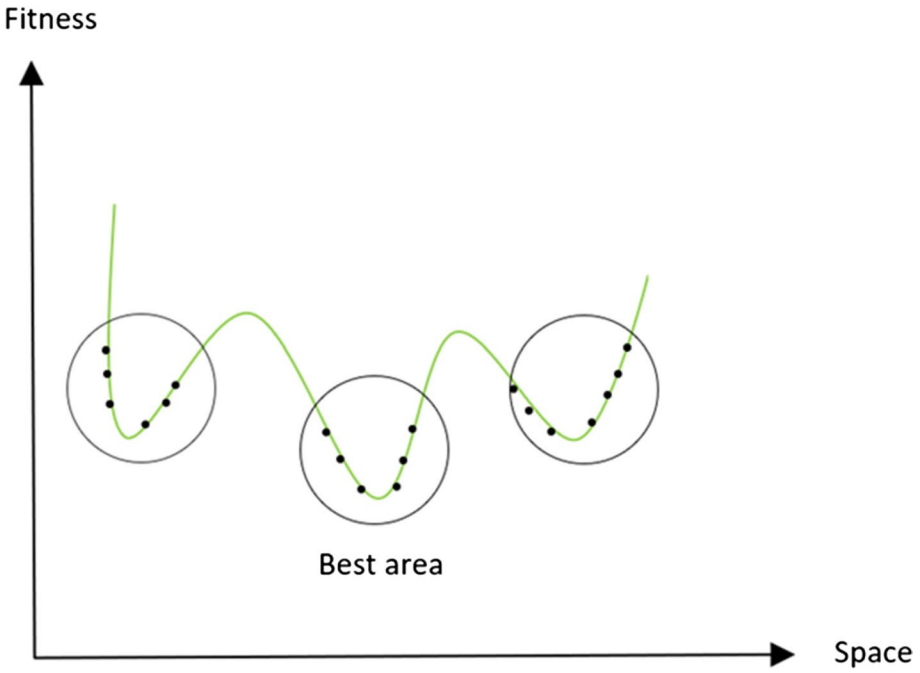

Fig. 2 Illustration of the enhanced DE variant with k-means clustering applied to the population, dividing 18 potential solutions into three clusters

- Selection: k candidate solutions are haphazardly chosen to serve as the initial centroids for the k-means clustering.
- Generation: M new candidate resolutions are produced through mutation, forming the group v^{clu}.
- Replacement: M candidates are haphazardly selected from the existing group to create a set BBB.
- Update: The top M performing candidates from the combination of v^{clu} and B are selected to form the new set $B\prime$. The updated group is then formed by mixing $(P - B) \cup B'$.

Empirical evaluation

Data

The recommended scheme employs two prominent databases: BOSSbase 1.01 [57] and BOWS-2 [58]. Each database entails 10,000 8-bit grayscale images, all standardized at a resolution of 512 × 512 pixels. The images from each database were partitioned into two groups to streamline the training and testing processes. For Class 1, 60% of the images, totaling 6,000, are allotted for training, while the remaining 40%, which equates to 4,000 images, are sllotted for testing. Class 2 follows a slightly different dispersion, with 70% of the images, or 7,000, dedicated to training and the remaining 30%, 3,000, earmarked for testing. These databases are vital for evaluating the performance of steganalysis tactics, as they offer diverse hurdles about image content and complexity that mirror real-world conditions. The BOSSbase database, specifically developed for steganalysis research, contains different natural images generally employed to compare the strength of steganographic tactics against detection, while the BOWS-2 database, designed for watermarking competitions, contains images generally employed to investigate watermarks' resistance to different digital manipulations and attacks. All algorithmic validations,

both model training and testing, were performed under MATLAB 2018a, selected for its rich collection of included functions and tools well-suited for image manipulation and ML processes, hence ensuring efficient application of recommended tactics.

Metrics

Various performance evaluation measures, including accuracy, F-measure, G-means, and AUC, were applied to assess models' performances. These measures give a detailed description of classification effectiveness, which is necessary when trying to strike a trade-off between sensitivity and specificity in difficult cases, including steganalysis. Although accuracy gives a broad description of performance, it is not effective in cases of class imbalancing; as such, it is supported by the F-measure, which includes accuracy together with recall. The F-measure is of relevance in applications, including steganalysis, where false negatives are of higher consequence than false positives. G-means, as a geometric mean of sensitivity as well as specificity, ensures a model's effectiveness in both classes, in particular where a minority class is of greater importance. The AUC looks at a scheme's capability to discern between classes at diverging thresholds, where higher values of the AUC are better performance, in particular in imbalanced collection cases.

The mathematical descriptions for the Accuracy, F-Measure, and G-means metrics are as follows:

$$Accuracy = \frac{True\ Positives\ (TP) + True\ Negatives\ (TN)}{Total\ observations} \tag{20}$$

$$F - Measure\ (F1\ Score) = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{21}$$

With

$$Precision = \frac{TP}{TP + False\ Positives\ (FP)} \tag{22}$$

And

$$Recall = \frac{TP}{TP + False\ Negatives\ (FN)} \tag{23}$$

$$G - means = \sqrt{Sensitivity \times\ Specificity} \tag{24}$$

Outcomes

A 5-fold stratified cross-validation technique was used to evaluate methodologies' accuracy and robustness. The cross-validation is stratified such that there is a minimum of an equivalent sample allocation from every class for every fold, offsetting bias in imbalanced data. The approach provides fair estimates by providing appropriate class presence in tests and training. The optimization in 5-fold is also maximized in computational efficiency by optimizing data use, reducing variance, and optimizing a model's generalization. The approach is well-suited in situations where there is limited data, optimizing its use without overfitting, and is directly applicable in fields, including radiology imaging and steganalysis.

In the evaluative step, the recommended scheme was subjected to a stringent comparison against eight separate DL schemes: CVTStego-Net [41], GTSCT-Net [42], HSVGG [43], MSP-CNN [44], Yedroudj [45], LEA [46], MLR [47], and RL-GAN [48]. Additionally, the scheme was compared to two variants: one substituting SVM and another substituting PSVM for ImAUC-PSVM, termed Recommended with SVM and Recommended with PSVM, respectively, and a third variant excluding hyperparameter optimization, referred to as Recommended w/o HO. The outcomes of these experiments for BossBase 1.01 and BOWS-2 databases are displayed in Fig. 3.

For the BossBase 1.01 database, performance reveals a definite edge of the recommended scheme over conventional schemes, including CVTStego-Net and GTSCT-Net, which are differentially designed for steganography and lack robustness for challenging image analysis tasks. The recommended scheme performs better than these by an estimated 15% in F-Measure and around 10% for AUC, reflecting a superior management of the challenges of the database. Schemes, including HSVGG and MSP-CNN,

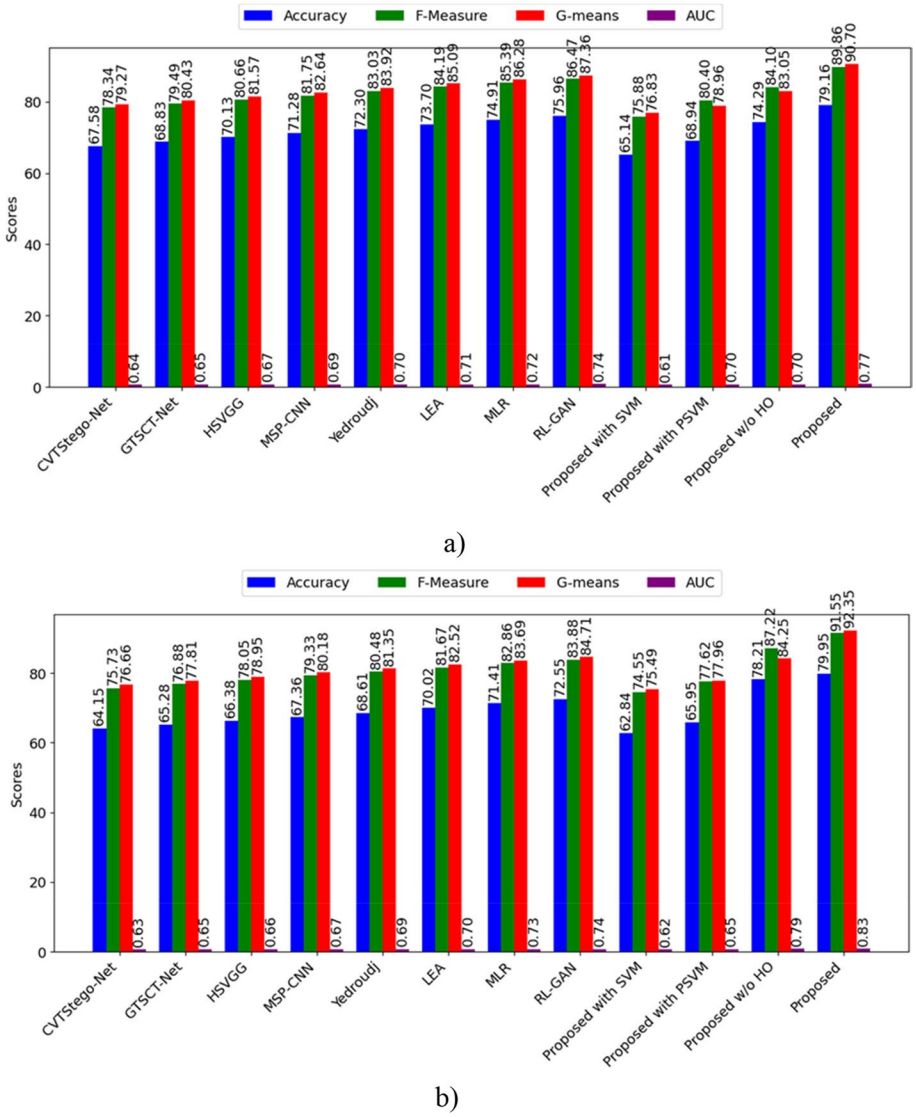

Fig. 3 Comparative performance of the recommended scheme with existing schemes on (**a**) BossBase 1.01 and (**b**) BOWS-2 databases a) b)

which are optimized for high-dimensional data, perform well but lack the recommended scheme's capability to generalize well under various conditions, exhibiting a circa 5–8% performance degradation at critical measures, including G-means. Comparing the recommended scheme to its variants on this database, the baseline version (without modifications) retains a strong edge, especially over the Recommended w/o HO, which incurs a reduction of around 5% under accuracy and similar reductions under other measures, reflecting how much optimized hyperparameters contribute to upper-performing performance.

For purposes of the BOWS-2 database, the scheme proves to be robust and flexible, performing well above baseline schemes like LEA and MLR, which are lagging by a margin of 20% AUC and 10–15% F-Measure. This indicates a strong ability of the recommended scheme to deal with diverse and possibly more complex imaging conditions present in the BOWS-2 database. The variants of the recommended scheme, specifically those incorporating SVM and PSVM, show minimal deviation in performance, suggesting that while the core architecture is solid, the choice of classifier only partially influences the overall effectiveness. This is contrasted by the Recommended w/o HO variant, which exhibits a notable decrease in performance metrics by about 10%, emphasizing again the critical role that finely tuned hyperparameters play in maintaining the scheme's efficacy across different databases.

Statistical analysis applied to the BossBase 1.01 database proves that our recommended framework outperforms current methods. The use of paired t-tests applied to accuracy, F-measure, G-means, and AUC results in p-values smaller than 0.05 in every case, thus validating its significant improvement over a variety of performance measures. Smaller intervals of confidence reflect a robustness increase as well as an accuracy increase. The BOWS-2 database shows a similar behavior, where our recommended framework outperforms competing models as well, in a case where p-values are again smaller than 0.05, thus verifying its statistical relevance. These experiments highlight the robustness, relevance, and applicability of our recommended scheme in dealing with complex and heterogeneous databases.

Figure 4 shows a comparative evaluation of the computational efficiency of different steganalysis methods, as assessed by runtime as well as GPU utilization, based specifically upon BossBase 1.01 as well as BOWS-2 image databases. For BossBase 1.01, methods like MSP-CNN and Yedroudj showed prolonged runtimes; however, RL-GAN, as well as our recommended approach, showed boosted computational efficiency. Importantly, GPU utilization as a function of time was not linear, as LEA used maximum GPU levels despite a normal time runtime. For BOWS-2, RL-GAN as well as our approach, showed decreased runtimes; however, a rise in GPU levels was observed, indicating that the level of intricacy in data requires greater processing power. The approach suggested herein, in fact, balances effectively between time efficiency as well as GPU levels, such that it is adoptable in resource-poor environments.

Figure 5 indicates loss curves for validation and training for 250 iterations over BossBase 1.01 and BOWS-2 databases. For BossBase 1.01, validation and training loss reduce; however, larger validation loss variability indicates overfitting and an inability to generalize to new data. For BOWS-2, initial losses reduce a lot; however, larger validation loss variability is an indication of sensitivity to data complexity alongside possible overfitting.

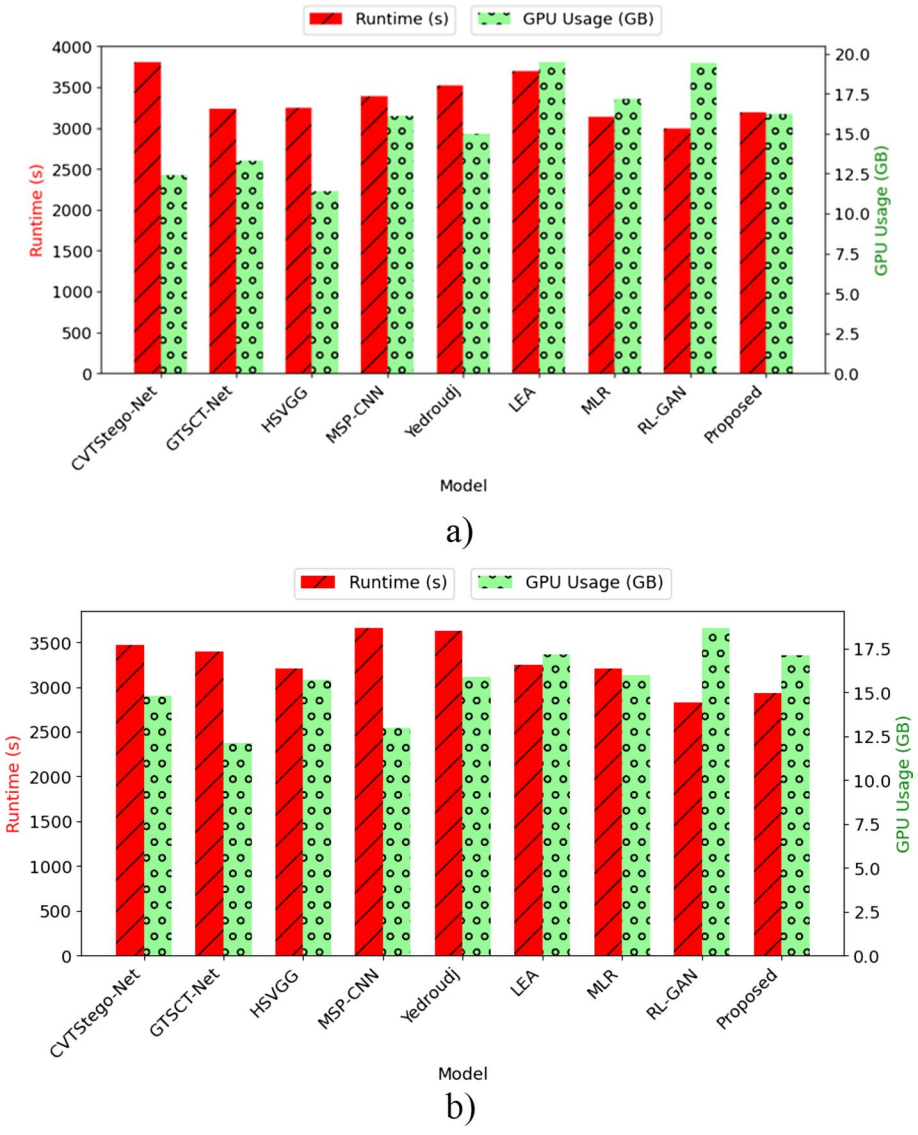

Fig. 4 Comparative runtime and GPU usage analysis of different steganalysis schemes for the (**a**) BossBase 1.01 and (**b**) BOWS-2 databases a) b)

These are indicators of a need to optimize hyperparameters well in an attempt to boost generalization.

Figure 6 shows the variation in decision times for the recommended approach over the BossBase 1.01 and BOWS-2 databases. The BossBase 1.01 histogram illustrates a tighter distribution in which the response time is between 100 and 140 milliseconds, which indicates higher processing efficiency based on relative ease in describing data. The BOWS-2 histogram, however, illustrates a larger range where durations are between 140 and 220 milliseconds, indicating reduced speed in decision-making due to higher data complexity. This result points out how efficient the approach is when handling fewer complex sets of data, yet concurrently indicates a point of improvement when handling larger complex sets of data, including BOWS-2, particularly in time-critical situations.

It is shown in Fig. 7 that, for BossBase 1.01 and BOWS-2 databases, the performance of ImAUC-PSVM is better than Standard PSVM after 250 epochs. For BossBase 1.01,

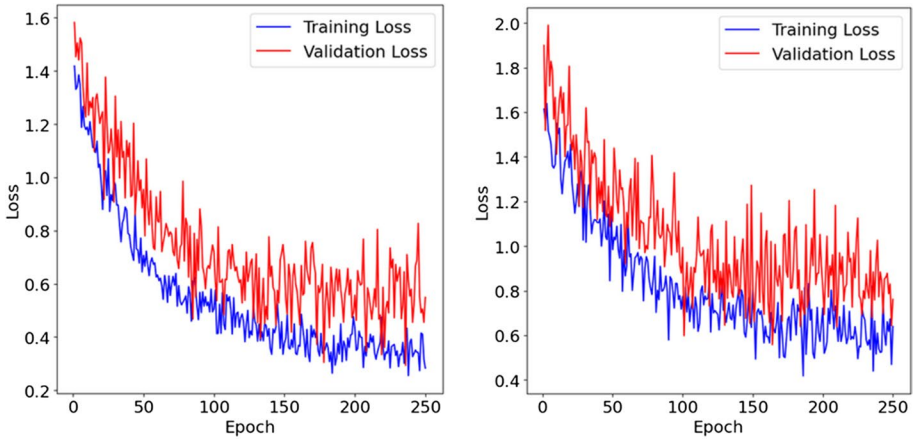

Fig. 5 Loss curves for 250 epochs during training and validation phases for (**a**) BossBase 1.01 and (**b**) BOWS-2 databases

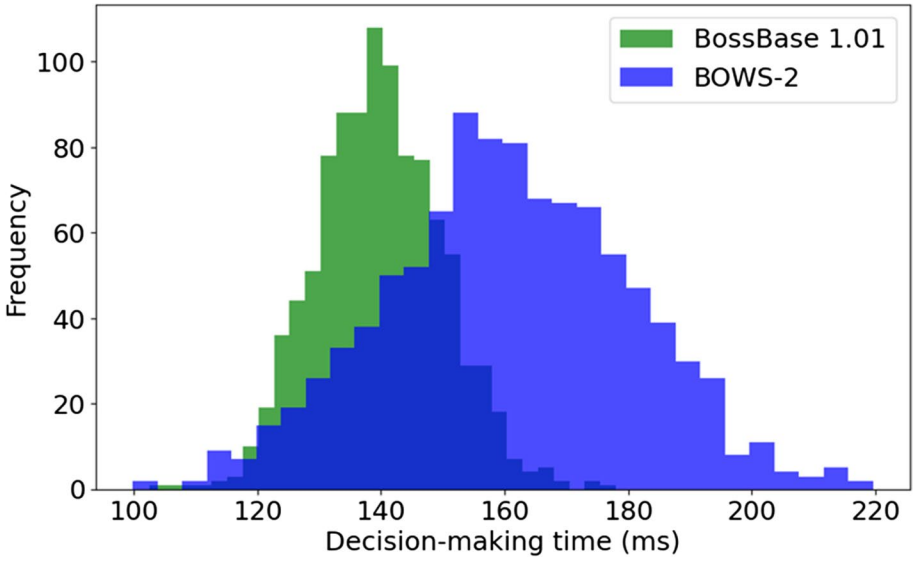

Fig. 6 Decision-making time dispersion of the recommended scheme for the BossBase 1.01 and BOWS-2 databases

ImAUC-PSVM shows a faster and steadier loss reduction, achieving a faster lower loss plateau compared to Standard PSVM, such that its adaptability and generalization are enhanced. For BOWS-2, ImAUC-PSVM is superior to Standard PSVM with a faster loss reduction, such that its ability in dealing with noisy and complex data is shown. The comparative performance gap proves ImAUC-PSVM's remarkable capabilities for rapid adaptability and fast decision-making in complex steganalysis problems.

Analysis of the recommended DE

Figure 8 presents an overall performance comparison of the new recommended DE scheme versus several well-established metaheuristic optimization tactics based on the BossBase 1.01 and BOWS-2 databases. The tactics being used for comparison include Human mental search (HMS) [59], Salp swarm algorithm (SSA) [60], cuckoo

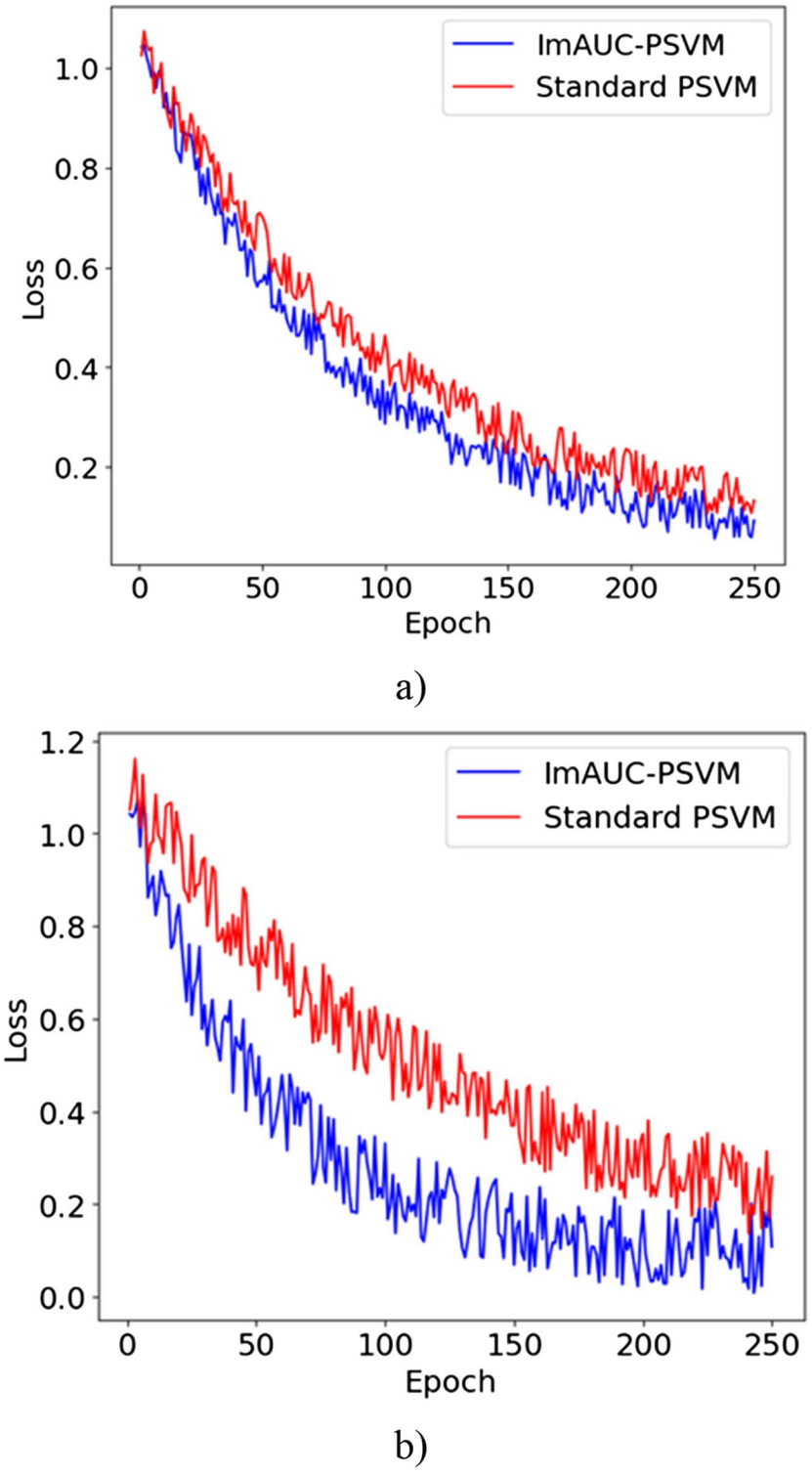

Fig. 7 Comparative loss paths of ImAUC-PSVM and regular PSVM across 250 epochs for (**a**) BossBase 1.01 and (**b**) BOWS-2 databases a) b)

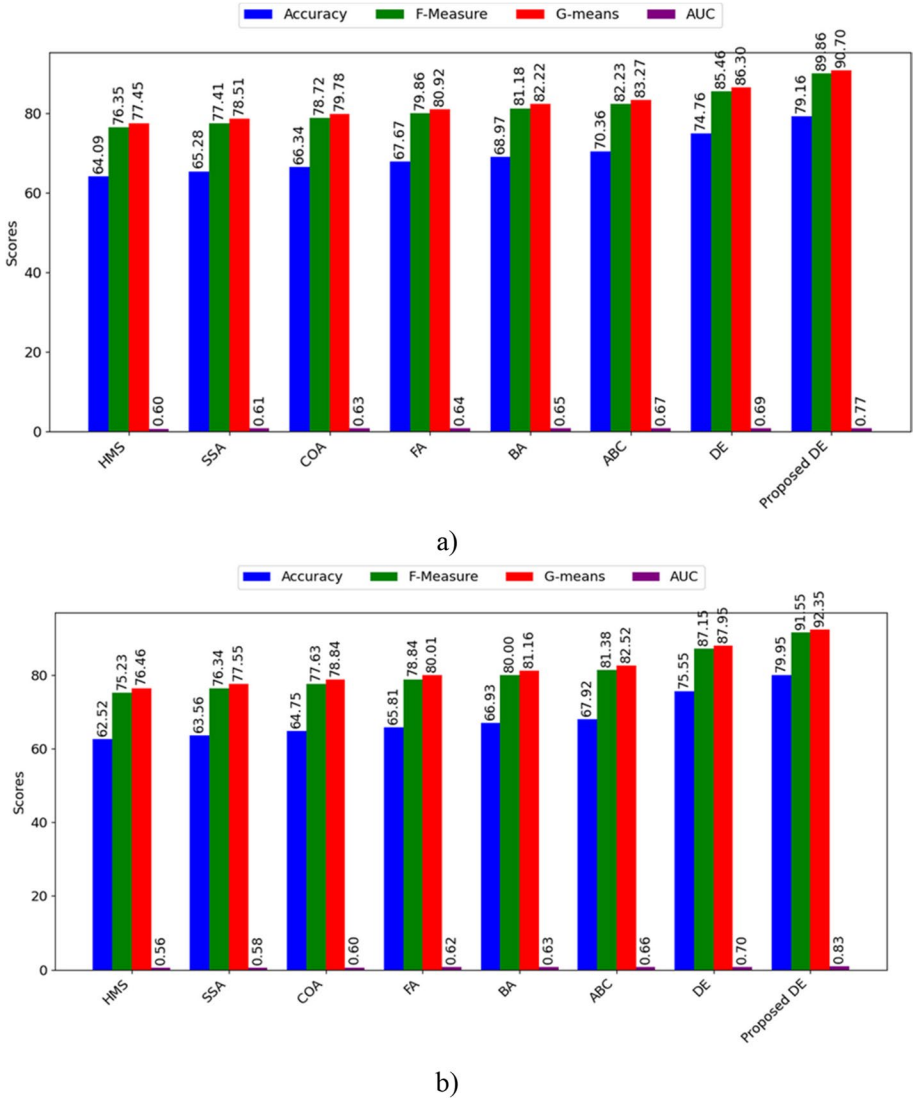

Fig. 8 Comparison of the performance of the recommended DE scheme with others for (**a**) BossBase 1.01 and (**b**) BOWS-2 databases a) b)

optimization algorithm (COA) [61], Firefly algorithm (FA) [62], Bat algorithm (BA) [63], ABC [64], and original DE. On both databases, the recommended DE scheme outperforms others, especially for Accuracy, F-Measure, G-means, and AUC measures, being pivotal for practical hyperparameter tuning of complex optimization problem domains, including image steganalysis.

The recommended DE outperforms other methods on both the BossBase 1.01 and BOWS-2 databases, showing significant improvements across all metrics. Its advanced mutation strategies and effective selection enable precise hyperparameter optimization, handling complex data structures efficiently. The performance gap, particularly versus conventional DE, highlights the effectiveness of this version, supporting its use for optimization in machine learning and data science.

Figure 9 shows the loss minimization trajectories over 300 rounds of DE optimization on the BossBase 1.01 and BOWS-2 databases. For BossBase 1.01, the loss drops sharply at first, reflecting rapid parameter optimization, then stabilizes as the algorithm

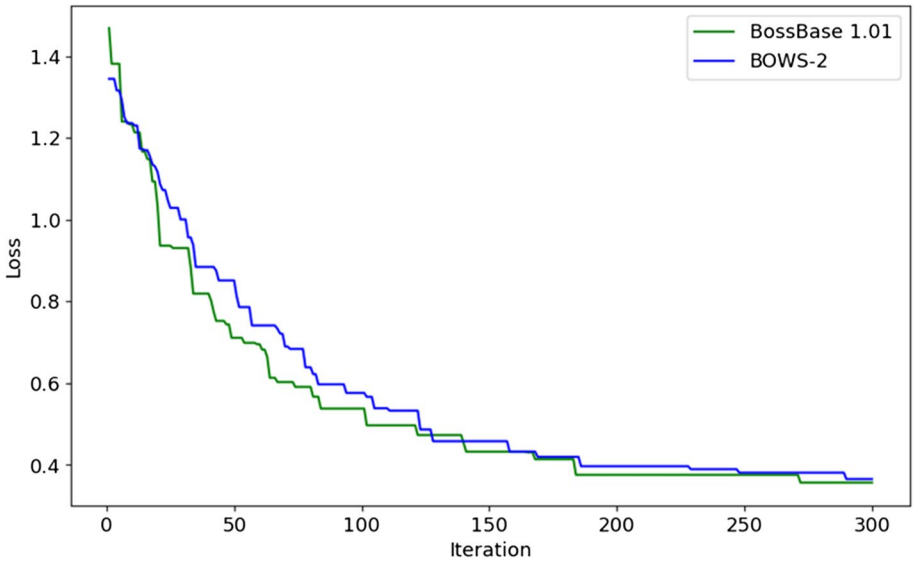

Fig. 9 Loss minimization paths of the DE optimization process for the BossBase 1.01 and BOWS-2 databases through 300 iterations

converges. In BOWS-2, the loss decreases gradually and steadily, suggesting a smoother optimization landscape or better initial settings. This comparison highlights DE's ability to refine parameters and achieve fast convergence, even with complex databases.

Figures 5 and 7, and 9 do not originate from the point (0,0) according to the characteristics of the data represented. These numbers specifically denote loss values and decision times during training and validation, wherein the model's beginning circumstances may not precisely correspond to zero. The plots commence at an elevated value due to the model initiating its learning from a non-zero initial state, indicative of performance throughout the initial training phases. The data points depicted in the pictures correspond to specific time intervals or epochs, with the starting point indicating the model's progress rather than the initial conditions. This is a prevalent method in performance assessment, particularly when addressing iterative learning models. The lack of origin in these graphs does not signify any problems with the model or data; instead, it is a consequence of the training and evaluation framework.

Discussion

This paper presents ImAUC-PSVM, an innovative image steganalysis model that optimizes detection for imbalanced databases. The scheme, for the first time, embodies AUC maximization into its operating system to promote effectiveness and adaptability. Incorporating ImAUC-PSVM into image steganalysis achieves dramatic improvements for dealing with imbalanced data dispersions, a prevalent problem for digital security domains. Comparing the recommended scheme's performance to several well-established algorithms indicates its superior performance, particularly at the BossBase 1.01 and BOWS-2 databases, and thus its promising potential for revolutionizing steganalysis significantly. This study provides strong evidence for the benefit of embedding AUC maximization into the SVM framework, addressing a key gap in steganalysis research affected by class imbalance. Incorporating DE for hyperparameter optimization enhances efficiency, yielding a system with fine control that achieves high accuracy

Table 2 F-measure comparison of the recommended scheme and State-of-the-Art methods on various databases

Method	F-measure (BossBase 1.01)	F-measure (BOWS-2)	F-measure (ICDAR2013)	F-measure (ICDAR2015)
Recommended Scheme	92.1%	91.5%	92.1%	91.5%
CVTStego-Net	87.2%	85.5%	87.2%	84.1%
GTSCT-Net	86.5%	83.4%	86.5%	82.0%
HSVGG	89.5%	87.1%	89.5%	85.8%
MSP-CNN	88.2%	86.4%	88.2%	84.7%
Yedroudj	86.9%	84.9%	86.9%	83.6%
LEA	79.8%	78.2%	79.8%	75.6%
MLR	81.3%	79.6%	81.3%	77.4%
RL-GAN	90.5%	88.9%	90.5%	87.0%

under various conditions. As Fig. 3 shows, the recommended method outperforms others in Accuracy, F-measure, G-means, and AUC, aligning with literature emphasizing the need for consistent, adaptive frameworks in security-sensitive applications. These results validate the method's effectiveness and establish a benchmark for future automated steganalysis research.

Comparing F-measure values with innovative approaches further demonstrates its superiority. On the ICDAR2013 database, the recommended scheme achieved 92.1% F-measure versus 87.2% for SPCNET. On ICDAR2015, it also outperformed SPCNET, which scored 74.1%. These findings confirm the model's enhanced detection accuracy compared to existing baselines.

Table 2 juxtaposes the F-measure values of the recommended strategy against various cutting-edge methods throughout the BossBase 1.01, BOWS-2, ICDAR2013, and ICDAR2015 databases. The results unequivocally demonstrate that the recommended strategy continuously surpasses all alternative methods, attaining the maximum F-measure across all databases. It significantly outperforms techniques, including SPCNET and GTSCT-Net, achieving an F-measure of 92.1% on the ICDAR2013 database. The higher performance is due to the integration of AUC maximization and DE optimization, which markedly augment the model's capacity to manage imbalanced data, resulting in improved accuracy in identifying steganographic images.

ImAUC-PSVM tackles the issue of unbalanced databases in image steganalysis, characterized by a predominance of non-steganographic images. Classical classifiers exhibit majority class bias, leading to inadequate detection of the minority class. ImAUC-PSVM enhances classification performance by integrating area under the curve maximization into its objective function, hence improving detection for both classes. This method mitigates sensitivity to class imbalance, providing a more reliable performance metric and enhancing the efficacy of the support vector machine in identifying concealed data inside images.

DE enhances the ImAUC-PSVM model by optimizing hyperparameters, including the penalty factor and support vector machine kernel parameters. This improves classification efficacy and guarantees the model generalizes effectively to novel data. An improved variant of DE utilizing k-means clustering for mutation has been introduced to boost performance. This enables DE to concentrate on promising areas of the search domain, circumventing local optima and accelerating convergence. The improved DE method offers an adaptive strategy for hyperparameter optimization, essential for dynamic steganalysis fields.

ImAUC-PSVM is relevant in domains, including fraud detection, medical diagnosis, and network security, where unbalanced data poses difficulties. Enhancing the identification of infrequent occurrences, including fraud and abnormalities, facilitates more effective preventive strategies. The use of DE guarantees the model's optimization as data progresses, rendering it resilient for practical applications.

The constraints of the recommended scheme are as follows:

- The ImAUC-PSVM and enhanced DE methodologies entail elevated computational expenses, particularly in extensive applications, owing to the ongoing assessment and modification of parameters, potentially constraining their applicability in resource-limited settings.
- Data sensitivity: ImAUC-PSVM exhibits significant sensitivity to data quality; the presence of noise or substandard data might impair performance, necessitating comprehensive pre-processing and meticulous attribute selection.
- The risk of overfitting arises from the simultaneous maximization of the area under the curve and hyperparameter optimization, especially in imbalanced databases, where the model may excessively adapt to the minority class and exhibit subpar performance on novel data.
- Adaptability to developing tactics: As steganography techniques progress, the system may necessitate regular retraining and recalibration to sustain efficacy, presenting issues in contexts with constant changes.

Conclusions

This study offers a substantial advancement in picture steganalysis by introducing an enhanced support vector machine model that optimizes the area under the curve, along with hyperparameter optimization using DE. The integration of area under the curve maximization within the support vector machine architecture improves model efficacy, especially for unbalanced databases, which are a prevalent issue in practical applications. The revised framework decreases computational complexity while maintaining the benefits of the traditional support vector machine, yielding an efficient and adaptable model for contemporary steganalysis. The advanced DE framework, incorporating a novel mutation process, facilitates dynamic and precise hyperparameter optimization, hence enhancing model efficacy. Experimental findings on the BossBase 1.01 and BOWS-2 databases validate the model's exceptional efficacy, especially in differentiating between unaltered and steganographic photos, as demonstrated by elevated F-measure scores. The findings highlight the capability of the enhanced support vector machine model to boost digital security by offering a reliable instrument for identifying hidden data in images.

Future endeavors may enhance the refined support vector machine model by integrating real-time adaptive learning to more effectively address the progression of steganographic methodologies. Furthermore, extending the model's application to other digital media, like video and audio files, could expand its use, improving security systems' capacity to identify concealed material across various communication platforms.

Abbreviations

AUC	Area Under the Curve
RR_AUC(s)	Regularized Risk AUC (scoring function)
s(x)	Scoring Function
φ(x)	Feature Mapping Function

w′	Classifier Weight Vector
1(·)	Indicator Function
N	Total Number of Samples
γ	Regularization Parameter
I	Identity Matrix
T	Transpose
φ′(x)	Transformed Feature Mapping
Z′	Design Matrix
Y′	Transformed Label Vector
K(x,x')	Kernel Function
p	Proportion of Majority Class
PSVM	Projected Support Vector Machine
f(x)	Decision Function
DE	Differential Evolution
$\vec{v}$_(i,g)	Mutant Vector in DE
$\vec{x}$_(r_k,g)	Random Vector in DE
F	Scaling Factor in DE
CR	Crossover Rate
u_(i,j,g)	Trial Vector

Acknowledgements
No individuals or organizations are to be acknowledged for contributions to this exploration.

Authors' contributions
XL: Writing-Original draft preparation, Conceptualization, Supervision, Project administration.

Funding
This investigation was not funded by a particular grant from any public, commercial, or charitable funding body.

Data availability
Data is obtainable upon request.

Declarations

Competing interests
The investigators assert no competing interests.

Received: 30 April 2025 / Accepted: 24 October 2025
Published online: 04 November 2025

References

1. Butora J, Bas P (2024) Size-independent reliable CNN for RJCA steganalysis. IEEE Trans Inf Forensics Secur. https://doi.org/10.1109/TIFS.2024.3379849
2. De La Croix NJ, Ahmad T, Han F (2024) Comprehensive survey on image steganalysis using deep learning. Array. https://doi.org/10.1016/j.array.2024.100353
3. Farooq N, Selwal A (2023) Image steganalysis using deep learning: a systematic review and open research challenges. J Ambient Intell Humaniz Comput 14(6):7761–7793
4. Rustad S, Andono PN, Shidik GF (2023) Digital image steganography survey and investigation (goal, assessment, method, development, and dataset). Signal Process 206:108908
5. Agarwal S, Jung K-H (2024) Digital image steganalysis using entropy driven deep neural network. J Inf Secur Appl 84:103799
6. Shankar DD, Azhakath AS (2023) Random embedded calibrated statistical blind steganalysis using cross validated support vector machine and support vector machine with particle swarm optimization. Sci Rep 13(1):2359
7. Suryawanshi G, Mali SN, Patil PD (2023) Multiclass Discriminator for Blind Steganalysis Using Statistical Features of Digital Images, in *International Conference on Smart Trends in Computing and Communications*, Springer, 1:625–637. https://doi.org/10.1007/978-981-99-0769-4_54.
8. Shankar DD, Khalil N, Azhakath AS (2023) Moderate embed cross validated and feature reduced steganalysis using principal component analysis in spatial and transform domain with support vector machine and support vector machine-particle swarm optimization. Multimed Tools Appl 82(7):10249–10276
9. Wang Y, Zhang R, Liu J (2023) Rls-dts: reinforcement-learning linguistic steganalysis in distribution-transformed scenario. IEEE Signal Process Lett 30:1232–1236
10. Michaylov KD, Sarmah DK (2025) Steganography and steganalysis for digital image enhanced forensic analysis and recommendations. J Cyber Secur Technol 9(1):1–27
11. Qasim HA, Mohammed RS (2024) Lightweight steganography technique in the internet of things: a survey. Al-Salam J Eng Technol 3(1):97–108
12. Maulana MS, Widianto SR, Sasongko A (2023) Steganography based on the B217AN algorithm for secret messages on flip horizontal and resize image. World J Adv Eng Technol Sci 9(1):17–28
13. Kuznetsov O, Frontoni E, Chernov K (2024) Beyond traditional steganography: enhancing security and performance with spread spectrum image steganography. Appl Intell 54(7):5253–5277
14. Zainal N, Hoshi AR, Ismail M, Rahem AA-RT, Wadi SM (2024) A hybrid steganography and watermark algorithm for copyright protection by using multiple embedding approaches. Bull Electr Eng Inf 13(3):1877–1896

15. Kheddar H, Hemis M, Himeur Y, Megías D, Amira A (2024) Deep learning for steganalysis of diverse data types: A review of methods, taxonomy, challenges and future directions. Neurocomputing, 581:127528. https://doi.org/10.1016/j.neucom.2024.127528.
16. Kuznetsov A, Luhanko N, Frontoni E, Romeo L, Rosati R (2024) Image steganalysis using deep learning models. Multimed Tools Appl 83(16):48607–48630
17. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP (2002) SMOTE: synthetic minority over-sampling technique. J Artif Intell Res 16:321–357
18. Wu Q, Zhang H-C, Chiu Y-J (2023) Application of asymmetric proximal support vector regression based on multitask learning in the stock market. Expert Syst Appl 227:120208
19. Martyniak R, Czaplewski B (2024) Architectural Modifications to boost Steganalysis with Convolutional Neural Networks, in *International Conference on Computational Science*, Springer, pp. 50–67 https://doi.org/10.1007/978-3-031-63751-3_4.
20. Zhao Y, Zhang W, Liu X (2024) Grid search with a weighted error function: hyper-parameter optimization for financial time series forecasting. Appl Soft Comput 154:111362
21. Franchini G (2024) Greennas: a green approach to the hyperparameters tuning in deep learning. Mathematics 12(6):850
22. Jiang J, Wen Z, Mansoor A, Mian A (2024) Efficient hyperparameter optimization with adaptive fidelity identification, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 26181–26190
23. Moravvej SV, Mousavirad SJ, Oliva D, Schaefer G, Sobhaninia Z (2022) An improved de algorithm to optimise the learning process of a bert-based plagiarism detection model. Padua, Italy: in 2022 IEEE Congress on evolutionary computation (CEC). IEEE, pp 1–7. https://doi.org/10.1109/CEC55065.2022.9870280.
24. Van Schyndel RG, Tirkel AZ, Osborne CF (1994) A digital watermark, Austin, TX, USA: in *Proceedings of 1st international conference on image processing*, IEEE, pp. 86–90. https://doi.org/10.1109/ICIP.1994.413536.
25. Cogranne R, Zitzmann C, Fillatre L, Retraint F, Nikiforov I, Cornu P (2011) A cover image model for reliable steganalysis, in *Information Hiding: 13th International Conference, IH* Prague, Czech Republic, May 18–20, 2011, Revised Selected Papers 13, Springer, 2011, pp. 178–192. https://doi.org/10.1007/978-3-642-24178-9_13.
26. Sedighi V, Cogranne R, Fridrich J (2015) Content-adaptive steganography by minimizing statistical detectability. IEEE Trans Inf Forensics Secur 11(2):221–234
27. Holub V, Fridrich J (2013) Digital image steganography using universal distortion, in *Proceedings of the first ACM workshop on Information hiding and multimedia security*, Association for Computing Machinery New York, NY, United States, pp. 59–68
28. Shi H, Dong J, Wang W, Qian Y, Zhang X (2017) SSGAN: Secure steganography based on generative adversarial networks, in *Advances in Multimedia Information Processing–PCM 2017: 18th Pacific-Rim Conference on Multimedia, Harbin, China, September 28–29*, Revised Selected Papers, Part I 18, Springer, 2018, pp. 534–544
29. You W, Zhang H, Zhao X (2020) A siamese CNN for image steganalysis. IEEE Trans Inf Forensics Secur 16:291–306
30. Essa A, Wu I (2024) Multibit least significant bit matching: a novel approach to image steganography. in Multimodal image exploitation and learning 2024. SPIE, 13033:107–113
31. Nashat D, Mamdouh L (2023) A least significant bit steganographic method using hough transform technique. Journal of Networking and Network Applications 3(2):73–80
32. Pramanik S (2023) An adaptive image steganography approach depending on integer wavelet transform and genetic algorithm. Multimed Tools Appl 82(22):34287–34319
33. Xie G, Ren J, Marshall S, Zhao H, Li R (2023) A novel gradient-guided post-processing method for adaptive image steganography. Signal Process 203:108813
34. Zhang L, Lu Y, Li T, Lu G (2024) Joint adaptive robust steganography network. IEEE Trans Ind Inform. https://doi.org/10.1109/TII.2024.3388674
35. Li G, Feng B, He M, Weng J, Lu W (2023) High-capacity coverless image steganographic scheme based on image synthesis. Signal Process Image Commun 111:116894
36. Huang R, Lian C, Dai Z, Li Z, Ma Z (2023) A novel hybrid image synthesis-mapping framework for steganography without embedding. IEEE Access 11:113176–113188
37. Huo L, Chen R, Wei J, Huang L (2024) A high-capacity and high-security image steganography network based on chaotic mapping and generative adversarial networks. Appl Sci 14(3):1225
38. Al Rubaie SF, Al-Azawi MKM (2024) High capacity double precision image steganography based on chaotic maps. Bull Electr Eng Inform 13(1):320–331
39. Amit A, Agarwal S, Gupta M, Sanadhya A, Kumar, Singh A (2023) CNN-Based adversarial embedding for image steganograpy, in *International conference on information technology*, Springer, 1:275–285. https://doi.org/10.1007/978-981-99-5997-6_24.
40. Kumar S, Soundrapandiyan R (2024) A cooperative three-player game theory approach for designing an ideal video steganography framework. Imaging Sci J 72(7):898–925
41. Bravo-Ortiz MA et al (2024) CVTStego-Net: A convolutional vision transformer architecture for spatial image steganalysis. J Inf Secur Appl 81:103695
42. Zhou Z et al (2024) Global texture sensitive convolutional transformer for medical image steganalysis. Multimedia Syst 30(3):155
43. Jeyaprakash H, Chokkalingam BS, V. V, and, Mohan S (2023) Stego detection: image steganalysis using a novel hidden Stego visual geometry group–based CNN classification. J Appl Secur Res 18(4):979–999
44. Ntivuguruzwa JDLC, Ahmad T (2023) A convolutional neural network to detect possible hidden data in spatial domain images. Cybersecurity 6(1):23
45. Vijjapu A, Vinod YS, Murty S, Raju BE, Satyanarayana BVV, Kumar GP Steganalysis using Convolutional Neural Networks-Yedroudj Net, in (2023) Coimbatore, India: *International conference on computer communication and informatics (ICCCI)*, IEEE, 2023, pp. 1–7. https://doi.org/10.1109/ICCCI56745.2023.10128190.
46. Liu S, Zhang C, Wang L, Yang P, Hua S, Zhang T (2023) Image steganalysis of low embedding rate based on the attention mechanism and transfer learning. Electronics 12(4):969
47. Sun Y (2024) Enhancing image steganalysis via integrated reinforcement learning and dilated convolution techniques. Signal Image Video Process 18:1–16

48. Al-Obaidi SAR, Lighvan MZ, Asadpour M (2024) Enhanced image steganalysis through reinforcement learning and generative adversarial networks. Intell Decis Technol 18(2):1077–1100
49. Zhu Z, Zhu X, Guo Y, Ye Y, Xue X (2012) Inverse matrix-free incremental proximal support vector machine. Decis Support Syst 53(3):395–405
50. Lo S-Y, Oza P, Patel VM (2022) Adversarially robust one-class novelty detection. IEEE Trans Pattern Anal Mach Intell 45(4):4167–4179
51. Wang G, Wong KW, Lu J (2020) AUC-based extreme learning machines for supervised and semi-supervised imbalanced classification. IEEE Trans Syst Man Cybern Syst 51(12):7919–7930
52. Wang G, Teoh JY-C, Lu J, Choi K-S (2020) Least squares support vector machines with fast leave-one-out AUC optimization on imbalanced prostate cancer data. Int J Mach Learn Cybernet 11(8):1909–1922
53. Karl F et al (2023) Multi-objective hyperparameter optimization in machine learning—an overview. ACM Trans Evol Comput Optim 3(4):1–50
54. Bean JC (1994) Genetic algorithms and random keys for sequencing and optimization. ORSA J Comput 6(2):154–160
55. Mousavirad SJ, Schaefer G, Korovin I, Moghadam MH, Saadatmand M, Pedram M An enhanced differential evolution algorithm using a novel clustering-based mutation operator, in (2021) Melbourne, Australia: *IEEE international conference on systems, man, and cybernetics (SMC)*, IEEE, 2021, pp. 176–181. https://doi.org/10.1109/SMC52423.2021.9658743.
56. Deb K (2005) A population-based algorithm-generator for real-parameter optimization. Soft Comput 9:236–253
57. Ma Y, Yang Z, Li T, Xu L, Qiao Y (2023) Image steganalysis method based on cover selection and adaptive filtered residual network. Comput Graph 115:43–54
58. Swarnkar N, Thomas A, Selwal A (2024) A generalized image steganalysis approach via decision level fusion of deep models. Multimed Tools Appl 83(14):43513–43538
59. Mousavirad SJ, Schaefer G, Ebrahimpour-Komleh H (2021) The human mental search algorithm for solving optimisation problems. Enabling AI Appl Data Sci, pp. Springer, Cham 27–47. https://doi.org/10.1007/978-3-030-52067-0_2.
60. Elgamal ZM, Yasin NM, Sabri AQM, Sihwail R, Tubishat M, Jarrah H (2021) Improved equilibrium optimization algorithm using elite opposition-based learning and new local search strategy for feature selection in medical datasets. Computation 9(6):68
61. Yang X-S, Deb S (2009) Cuckoo search via Lévy flights, Coimbatore, India: in *2009 World congress on nature & biologically inspired computing (NaBIC)*, Ieee, pp. 210–214. https://doi.org/10.1109/NABIC.2009.5393690.
62. Yang X-S (2010) Firefly algorithm, stochastic test functions and design optimisation. Int J Bio-Inspir Comput 2(2):78–84
63. Yildizdan G, Baykan ÖK (2020) A novel modified bat algorithm hybridizing by differential evolution algorithm. Expert Syst Appl 141:112949
64. Vakilian S, Moravvej SV, Fanian A Using the artificial bee colony (ABC) algorithm in collaboration with the fog nodes in the Internet of Things three-layer architecture, in (2021) Tehran, Iran, Islamic Republic of: *29th Iranian conference on electrical engineering (ICEE)*, IEEE, 2021, pp. 509–513. https://doi.org/10.1109/ICEE52715.2021.9544399.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.