Al-Farabi Journal of Engineering Sciences Volume (4), Issue (2) December (2025)

ISSN: 2617-5517 (issn.org)
Al-Farabi Journal of Engineering Sciences
https://iasj.rdd.edu.iq/journals/journal/view/97
مجلة الفارابي للعلوم الهندسية تصدرها جامعة الفارابي

A Comparative Analysis of Image Steganalysis Methods: Handcrafted, Deep Learning, and Hybrid Paradigms

Saif Salahaldin Affat[1, *], Ismael Taha Ahmed[2]

[1]College of Computer Science & IT, University of Anbar, Anbar, Iraq, sai24c1007@uoanbar.edu.iq
[2]College of Computer Science & IT, University of Anbar, Anbar, Iraq, ismail.taha@uoanbar.edu.iq

Abstract: Over the last decade, image steganalysis, the science of detecting hidden information in digital images, has undergone a fundamental transformation, from manually engineered statistical feature methods to sophisticated deep learning architectures that can automatically extract features. This survey presents a systematic and detailed review of image steganalysis techniques, which can be grouped into four main categories: 1) Traditional intensity and texture-based methods including threshold-based methods and Grey Level Co-occurrence Matrix (GLCM) features; 2) Handcrafted feature-based methods such as the Spatial Rich Model (SRM), SPAM, DCTR, and GFR with ensemble classifiers; 3) Machine learning and deep learning methods including CNN-based architectures such as Xu-Net, Ye-Net, Yedroudj-Net, and SRNet; and 4) Hybrid fusion-based frameworks combining handcrafted and learned features to enhance detection robustness. The survey also includes common evaluation criteria, popular benchmark datasets such as BOSSbase, ALASKA II and IStego100K, and a comparison of stated detection accuracies in terms of methodologies and datasets. Key findings indicate that deep learning approaches consistently outperform traditional methods, with reported accuracies of up to 99% on some benchmarks, but important challenges remain unaddressed, such as the cover-source mismatch, low-payload detection, and the generalization gap between controlled laboratory conditions and real-world social media transmission. This review highlights these outstanding difficulties and discusses the future research directions that need to be followed in order to produce universal, resilient and computationally efficient steganalysis systems that can work efficiently in realistic deployment situations.

Keywords— Image Steganalysis, Steganography, Deep Learning, Convolutional Neural Network, Handcrafted Features, Hybrid Methods.

1. Introduction

In the contemporary digital age, with the incredible rise of the Internet and information technology, a huge amount of data is produced every day, highlighting its vital role in influencing human experiences. Protection of communication Data has always been important for privacy, integrity and confidentiality. There are many ways of securing data which are commonly known as Information Hiding Techniques. Steganography is the art and science of writing hidden messages in such a way that they appear to be normal objects. Steganography derives from two Greek words: "steganos" meaning "secret" or "cover" and "graphia" meaning "writing". During the Greek period several different ways were used to perform steganography such shaving heads, wooden blocks, invisible ink, microdots and tattoos. Modern steganography, since the advent of computers, digitization and the fast transmission of information via the Internet, has used a variety of computer files and protocols to hide information. Multimedia files (films, audio, photographs, internet protocols such as TCP/IP) can be used to hide information, making it invisible to the human visual system.

Protection of information is done primarily by two techniques i.e. data concealing and encryption. Steganography is an important data hiding method which hides information in carriers to realize clandestine communication. Therefore, steganography may safeguard data efficiently because of its unique hiding properties[1]. Figure 1 shows the basic steganography process.

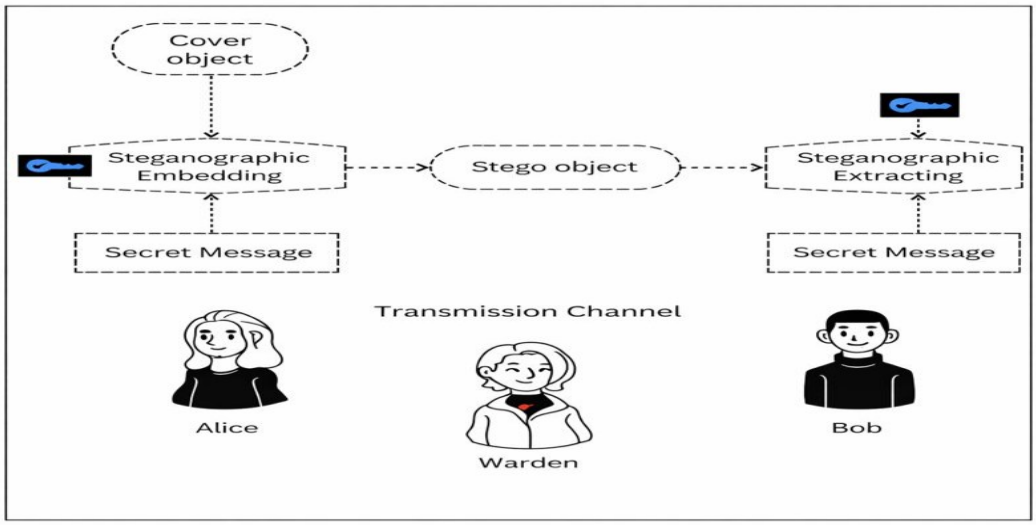

Figure 1*: Steganography process illustrated by the prisoner's problem*.

Steganalysis of image steganography is concerned with the analysis of statistical features of the carrier picture, detection of the existence of the embedded data, discovery of the embedded data and estimation of the information capacity. It extracts steganographic content in an optimal way[2]. Study of steganalysis is important for the protection of sensitive data, combating terrorism and criminal activities, and Internet security[3]. Steganalysis usually consists of three basic steps: First, to determine whether the target image contains steganographic information, i.e., to classify it as a steganographic image or a traditional image; Second, to determine the steganographic algorithm used. Used for the steganographic image includes the position of the hidden information and the capability for embedded cyphers. Finally recovering the secret information from the steganographic image.

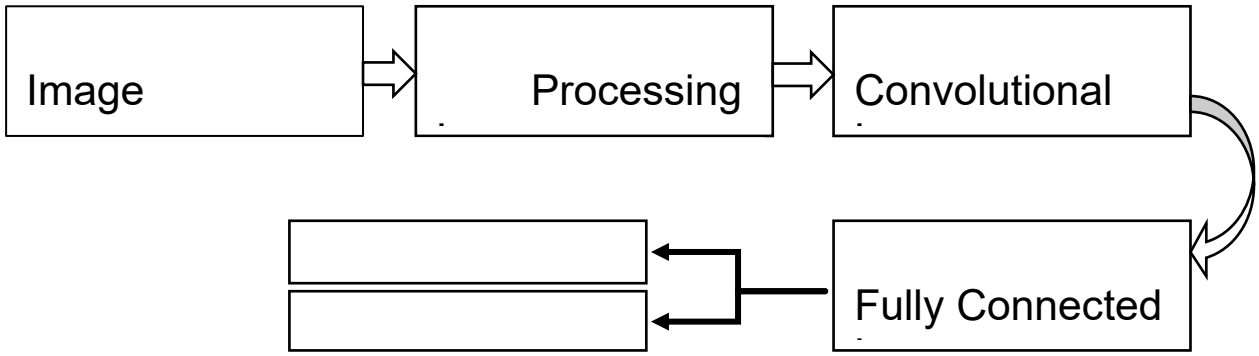

Figure 2**:** *The framework of steganalysis algorithm*.

Input Stage The process begins with sending the image for examination. It is called in the graphic "Stego Image" but in reality it is the "tested image" that the system has to decide if it contains concealed data (Stego) or is an unmodified image (Cover). Processing Layer This is the most important phase to detect steganography. This layer does not read the image as colors but as the "noise" or "signatures" created by the data steganography process, using filters such as high-pass filters to cut down the features of the original image. The idea is to make hidden patterns more apparent to the AI. Convolutional Layers - (Extraction of Features) This is where the Convolutional Neural Network (CNN) comes into play. These layers are feature extractors themselves. They look for very little patterns and changes in the pixel values that are not visible to the human eye, indicating tampering or buried data. Classification Dense Layer This layer is the "decision-maker" once the important features are extracted in the previous phase. It aggregates all collected information and qualities, does statistical analysis and calculates the final likelihood of the image category. Result Finally the algorithm gives its final judgement (classification), assigning the image to one of two categories: Cover Image: An unaltered, genuine image without any hidden info. Stego Image: An image used to conceal covert data or communications for malicious reasons.

Image steganalysis is an important field of information security dealing with detection of concealed information in image files[4],[5]. Steganalysis is a countermeasure to steganography, which is the art of hiding secret information within public carriers, such as photos, music or video. The goal of steganalysis is to detect whether a cover medium contains hidden information[6],[7]. The goal is to analyses statistics of the carrier pictures, to reveal secret data and to find out whether any additional information has been embedded[6]. This approach is important to stop the leaking of sensitive information and to combat criminal activity[6]. Steganalysis techniques have been developed considerably from traditional machine learning techniques to advanced deep learning techniques[5]. Contemporary, complex steganography algorithms were a substantial barrier for conventional systems that focused on manually engineered feature extraction[6]. Current techniques that embed hidden information in regions of an image with high complexity are not addressed by traditional approaches, therefore modelling and detecting the hidden data is difficult[6]. Deep learning based models, in particular those that use Convolutional Neural Networks (CNNs), have overcome many barriers by automating feature extraction and boosting the detection accuracy[5],[7].

Modern steganalysis involves numerous important steps. First, it classifies the target image as a steganographic image or a normal image, so as to determine whether the target image contains hidden information. Secondly, it is aimed to detect the used steganography technique and to find the hidden data. Eventually, it may attempt to extract the sensitive information on its own[6]. These techniques work well with large-scale pre-labeled image datasets for training, but recent advances in active learning and deep reinforcement learning (DRL) are addressing the problem of collecting these expensive datasets by improving the model performance with less labelled data[4].

Steganalysis is the science of finding hidden information in digital material. According on the prior knowledge of the embedding technique, it may be roughly classified into two major categories: targeted (or specific) steganalysis and blind (or universal) steganalysis[8],[9].

Targeted steganalysis, sometimes called as targeted steganalysis, is based on the pre-knowledge of the steganographic technique or instrument used to disguise the secret message[8],[9]. This method uses known defects or vulnerabilities of a particular steganographic technique to determine its presence [5]. Jessica Fridrich's Breaking F5 is one of several specialized techniques that detect a certain embedding methodology[10]. These methods have shown high accuracy and low false alarm rate. However, they can only be applied practically against the specific algorithms they are designed for[9]. In contrast, blind steganalysis, also known as universal steganalysis, is independent of the embedding procedure and does not require prior knowledge of the exact steganographic method used[7],[8]. In this approach, steganographic detection is formulated as a classification problem, sometimes using machine learning to discover high-dimensional features that separate clean from stego content[9]. We seek to identify similar trends or statistical anomalies that may indicate buried information, regardless of the technique of embedding used[8],[7]. This improves the versatility of blind steganalysis and allows its application in real-world forensic scenarios where the steganography instrument is often unknown[7]. The main difference between the two is the amount of prior information needed. Targeted steganalysis is successful, but it is limited because it requires knowledge of the embedding procedure. Blind or universal steganalysis has a broader scope as it aims to detect any steganographic content without prior knowledge of the technology used and so it is more ideal for general purpose and forensic investigations[8],[7].

The organization of the remaining part of this paper is as follows: Traditional methods of image processing are described in Section 2. Section 3 describes handcrafted feature based approaches. Section 4 is dedicated to methods based on machine learning. 5. Deep Learning Based Approaches. Section 6 discusses hybrid or fusion based approaches. 7. Performance metrics and assessments. Datasets in Section 8 . Section 9 presents the available image steganalysis algorithms . Finally, Section 10 discusses the obstacles and future directions.

2. Traditional Image Processing-Based Methods

2.1. Threshold-Based Methods in Steganalysis

Texture and intensity-based approaches are important in steganalysis, the art of detecting hidden information in digital material. These techniques are often based on thresholding to better detect small statistical changes due to steganographic embedding. One notable methodology is the Threshold Local Binary Pattern (TLBP) which is an improvement of the Local Binary Pattern (LBP) method that is more versatile in the detection of embedding abnormalities[11]. The TLBP operator represents the magnitude relationship between a center pixel and its neighboring pixels, indicating their relationship within a certain distance[11]. Another approach is to quantify data dependency by using the co-occurrence matrix (CM) of differences between neighboring picture pixels, which indicates the noise level of the image. The method controls the absolute difference values by use of a statistical threshold, truncating them at a certain value T[12]. Enhanced thresholding methods lead to better detection, especially in few-shot learning settings with few training data. A normalized feature salience threshold can be used to discard the less apparent stego aspects of the instances, thus selecting a group of samples with more robust representative qualities for model training[13]. Instead, we adopt another way to obtain the probability maps. We produce the probability maps using a Variational Autoencoder (VAE) and then we apply Otsu's thresholding technique to divide each map into two areas, which then drives the embedding simulation to generate pseudo-stego samples[13]. Another technique is to use Structural Similarity Index Measure (SSIM) to create a threshold to detect stego or cover image[8].

2.2. Statistical Texture Features in Steganalysis

Statistical texture features are of vital importance in the steganalysis to detect the slight changes due to embedding hidden data into pictures. Grey Level Co-occurrence Matrix (GLCM) is one of the important feature extraction techniques. GLCM is a probability distribution function which describes the texture of an image by analyzing the spatial correlations of pixel pairs. It measures the texture of the image by calculating the occurrence of pixel pairs having certain grey levels at a specific distance and orientation. The GLCM can be used to calculate different texture properties as contrast, energy, homogeneity, entropy and correlation[14],[11],[15],[8]. Blind steganalysis is the task of detecting steganography without any knowledge of the embedding algorithm used. Many works rely on characteristics extracted from GLCM for blind steganalysis. One way is to generate four GLCMs at different angles (0°, 45°, 90° and 135°) and then average them[8]. From this average GLCM low dimensional characteristics can be extracted using Principle Component Analysis (PCA). The higher order statistical aspects called the Haralick features (such energy, contrast, entropy and correlation) are also computed[8]. The relationship between the magnitude of the center pixel and its surrounding pixels is described using the Threshold Local Binary Pattern (TLBP) method to obtain picture residuals[11]. The generated TLBP image is then used for extracting the textural features using GLCM, Discrete Wavelet Transform (DWT) and Contourlet Transform (CT). The combination of several statistical and multi-resolution feature sets has been shown to be very accurate for classification[11]. The statistical features are then input to machine learning classifiers such as deep neural networks (ANNs), convolutional neural networks (CNNs) or support vector machines (SVMs) to distinguish clean and stego photos [14],[15],[8]. The underlying assumption is that steganographic embedding, no matter how tiny, changes the statistical properties of a picture and these changes are reflected in the GLCM and its derived features. The GLCM of a clean image is usually diagonally symmetric, something that could be damaged after LSB encoding[8]. Training models with these feature sets makes steganalysis systems obtain higher detection accuracy[14],[8].

2.3. Limitations of Intensity and Texture-Based Steganalysis Methods

The sensitivity and texture-based steganalysis techniques are very limited due to the difficulty in dealing with picture noise and poor discrimination capabilities, especially for complex images. Images with a lot of detail and complex textures have a higher level of intrinsic noise. The high baseline noise could easily mask the small additive stego noise introduced when embedding a message and hence make it difficult to identify[12]. The statistical features adopted by these methods, including those extracted from the Grey Level Co-occurrence Matrix (GLCM), may not be sensitive enough to capture the small changes introduced by steganography in complex images. The statistical features used by these methods, such as those computed from the Grey Level Co-occurrence Matrix (GLCM), may not be sensitive enough to detect the small changes introduced by steganography in complex images. This leads to not enough discernability between cover (original) and stego (updated) images. This leads to poor discriminating and low detection accuracy. The experiments reveal that the detection accuracy of approaches based on full model features such as SRMQ1, for example, can be as low as 55.52% for images with intricate textures, which is somewhat better than the random guess. The features retrieved from this complicated photos are modified very little after embedding, thus, this feature type is not sensitive to this kind of images. However, the images with less sophisticated information and smoother regions have less initial noise strength, and therefore the stego noise is more obvious and the detection accuracy is improved. Thus, the performance of these methods depends strongly on the texture complexity of the image. This is a serious limitation when judging photos with complicated noisy textures[12].

3. Handcrafted Feature-Based Methods

Traditional approaches for detecting hidden information in digital photos are based on handcrafted feature-based steganalysis algorithms. These approaches work by extracting artificially designed statistical features that are sensitive to steganographic activity, then using an ensemble classifier to determine whether there is any hidden information in an image[16],[17]. Typically, this process follows a pre-defined sequence, which consists of the feature extraction techniques to get high-dimensional features from the images and the training of an ensemble classifier. The trained classifier (usually a set of sub-classifiers) then votes on the test image whether it is categorized as 'clear' or 'stego'[16].

Key Feature Sets

Several well-known handcrafted feature sets are used in spatial domain steganalysis, including:

1- The Spatial Rich Model (SRM) is obtained by merging many sub-models built from the joint distributions of neighboring samples in quantized image noise residuals. These residuals are due to linear and nonlinear high-pass filters[16]. SRM is regarded as a significant paradigm for evaluating the security of spatial steganography[11].

2- Subtractive Pixel Adjacency Matrix (SPAM): It is a commonly used feature set in the state-of-the-art LSB steganalysis[17].

3- Threshold Local Binary Pattern (TLBP): This method collects features from the second order co-occurrence matrices by high order differential filters to get the image residuals and then uses the threshold LBP operator[16]. It is an improved version of the Local Binary Pattern (LBP) and is considered to be flexible for detection of embedding abnormalities[11].

4- Co-occurrence Markov Features: These features characterize the tiny embedding distortions by describing the transitions between pixel values. The strategy is to compute the derivatives in several directions and then threshold to highlight the distortions. The feature set is extracted using the co-occurrence matrices of these derived features[17].

We carry out the extraction of textural features using Grey Level Co-occurrence Matrix (GLCM) and integrate them with multi-resolution data obtained by using transforms such as Discrete Wavelet Transform (DWT) and Contourlet Transform (CT), to construct a comprehensive feature vector[11].

For JPEG images, which are in the frequency domain, distinct set of features are used:

1- GFR (Gabor Filter Residual): This approach uses the 256 Gabor filter kernels with varying sizes and orientations[18].
2- DCTR (Discrete Cosine Transform Residual): This approach is a residual image gotten by convolving the image with a DCT filter kernel[18].
3- CC-JRM (Cartesian Calibration - JPEG Rich Model): this method creates several sub-models by analyzing DCT coefficients and spatial correlations[18].

Classification Feature extraction and training of an ensemble classifier (i.e. Fisher Linear Discriminant (FLD) classifier) for detection[16],[17]. This classifier has numerous base learners or sub-classifiers[16],[17]. In the detection process, each sub-classifier votes for picture “clear” (without secret data) or “stego” (with secret data). The decision is decided according to the votes count. If the count of 'stego' votes is more than 'clear' votes, it is concluded that the image contains hidden data[16]. Handcrafted feature-based approaches are a core approach in steganalysis, which is focused on manually designing and extracting statistical features sensitive to the subtle changes introduced by steganography. These techniques

work well but often lead to high-dimensional feature sets, and have mostly been replaced or supplemented by deep learning algorithms that automate the feature extraction process[16],[18].
Strengths and Weaknesses of Handcrafted Feature-Based Methods:

a. Strengths

Targeted Sensitivity: One of the strengths is their ability to use purposely designed statistical features that are sensitive to the steganographic activities[16]. This allows them to catch certain types of interruptions and distortions due to data embedding.
Established Models: This method has resulted in the creation of known and effective feature sets. Common features in spatial domain steganalysis are SPAM, SRM (Spatial Rich Model) and TLBP[16],[11]. The SRM is an important paradigm to evaluate the security of spatial steganography[11].

b. Weaknesses

One of the major drawbacks is the manual extraction of attributes. This is different from deep learning approaches, where this is automated and often leads to better results[16]. High Dimensionality and Redundancy: Such approaches generally provide high-dimensional feature sets, often comprising numerous duplicate and conflicting characteristics. Redundant features increase the computational dimensions, time complexity and storage needs and competing features may decrease the detection accuracy[18]. Limited Flexibility Custom solutions may not be able to cope with the growing complexity of the cutting-edge steganography techniques, especially those based on Generative Adversarial Networks (GANs). The dependence on predefined features limits their adaptability to different and evolving data hiding techniques, and therefore their end-to-end learning ability[19]. High Computational Cost: The retrieved features are complex and high dimensional, which may lead to a huge time overhead and storage requirement, thereby making the method computationally expensive[18].

4. Machine Learning-Based Methods

Machine learning (ML), and in particular deep learning (DL), has become the major factor in modern steganalysis, the process of detecting hidden data in digital material. Traditional steganalysis methods relied on manually designed features and statistical models, but they were often inadequate for high-end, content adaptive steganography algorithms. Deep learning techniques, in particular those based on Convolutional Neural Networks (CNNs), can automate the feature extraction and classification process, thus improving the detection performance. In pre-processing, these networks often apply a high-pass filter to augment the weak stego signal as input to the trainable network layers. Many CNN architectures have been proposed, e.g., GNCNN, Xu-Net, Ye-Net, Yedroudj-Net, and SRNet . Some improvements include dedicated activation functions, e.g., Truncated Linear Unit (TLU), residual learning to preserve weak signals, and the use of selection-channel aware (SCA) knowledge[9],[20],[21]. Apart from traditional based CNNs, different ML paradigms are also under research. Recurrent neural networks (RNNs) and LSTM (Long Short-Term Memory) networks are used to analyze sequential data like text and audio. Deep Belief Networks (DBNs) and Automatic encoders (AEs) for extracted features and typical noise removal[9]. Deep Reinforcement Learning (DRL) is one of the advanced approaches used for training autonomous analysis of digital data. Active learning, often used with deep reinforcement learning (DRL), addresses the problem of needing large amounts of labelled data by intelligently choosing the most informative unlabeled samples to be annotated, reducing annotation costs and improving training efficiency[22]. An alternative, Green Learning (GL) is a modular, theoretically sound, and energy-efficient option to Deep Learning (DL) for detecting anomalies due to steganographic embedding, using unsupervised representation learning and supervised feature learning[21]. These different Machine learning approaches have taken the field a long way from manually engineered characteristics to automatic, robust and ever more adaptive detection systems[9],[20].
Steganalysis classifiers: After extracting features from a digital media, several classifiers can be applied to decide whether it contains hidden information. Classifiers are a basic part of the machine learning techniques which often involve complex feature extraction and classification. The accuracy of these classifiers is determined by the features supplied. For the deep learning applications, the neural networks employed for classification tasks are encountered with the Support Vector Machine (SVM) and ensemble classifiers[23]. The Support Vector Machine (SVM) is one of the most widely used classifiers for binary classification problems in steganalysis, especially when the aim is to distinguish between a 'cover' (original) image and a 'stego' (modified) image[23],[15]. Support Vector Machine (SVM) comes from the idea of constructing an ideal hyperplane for classifying the data points in classes. Good performance in high dimensional areas is vital, which is generally the case when you are working with a considerable number of recovered features. Evolutionary optimization techniques such as Particle Swarm Optimization (PSO) can be applied to improve SVM performance[15]. Other classifiers as referred to literature are ensemble classifiers and more recently, deep neural networks based classifiers. Ensemble classifiers, such as AdaBoost, combine multiple weak classifiers to create a powerful and robust model, known for its simplicity and outstanding efficiency[11]. Deep learning approaches, especially Convolutional Neural Networks (CNN) integrate feature extraction and classification in a single end-to-end process. The previous literature has used K-Nearest Neighbors (KNN), a popular classification technique, but none of the studies discusses its application for steganalysis feature categorization[23],[15],[11],[19]. The given texts do not mention the Random Forest as a classification tool, similarly[23],[15],[19]. Support Vector Machines (SVM) and ensemble classifiers are well-known classifiers for the

recovered features in conventional steganalysis. Support Vector Machines (SVM) are known for their good performance on high-dimensional datasets but deep learning techniques are a paradigm shift since they combine the feature extraction and classification steps.

4.1. Machine Learning in Steganalysis: Feature Selection and Classification

In general, the steganalysis process of detecting hidden information in digital images using machine learning is a two-step process, feature extraction/selection and subsequent classification[15],[19]. The approach is designed to differentiate between stego images (images with concealed data) and cover images (original images)[15]. However, traditional methods are sometimes ineffective, especially against advanced adaptive steganography, where features need to be extracted from many perspectives, resulting in very high dimensional and complex feature sets[18]. The retrieved features are high dimensional and may contain redundant or contradictory information, hence feature selection is needed to reduce detection accuracy and computational burden. Features are picked to reduce dimensions but maintaining accuracy. For this, specialist or universal techniques have been developed (applicable to most steganalysis methodologies). The Fisher criterion is a widely used criterion to evaluate the discriminability of features. It measures the intraclass aggregation and interclass dispersion[18]. The Spatial Rich Model (SRM) is a popular traditional approach that generates a high-dimensional feature vector from noisy residuals and inputs it to a classifier[21]. Once a collection of features is picked, a classifier is trained to take the ultimate decision.

5. Deep Learning-Based Methods

Deep learning methods have emerged as a leading approach for steganalysis with significant gains in detection accuracy over traditional feature engineering approaches[24]. These techniques apply deep neural networks (DNNs) to learn and extract subtle high-frequency noise patterns and texture fluctuations indicating the presence of hidden data automatically[24]. Such automatic feature learning obviates the need of manual feature generation and improves detection accuracy. However, challenges are still there, e.g., focusing on small steganographic changes and reducing information loss in training[9].

5.1 Common Architectures and Models:

Deep learning steganalysis networks are primarily composed of three components:
Preprocessing module, feature extraction module and classifier module[25],[26].

1- Preprocessing: This first stage often applies high-pass filters from the Spatial Rich Model (SRM) to minimize visual content and raise the steganographic noise-to-signal ratio[27],[25]. Some newer networks (e.g. SRNet) learn the filters directly instead of preset pre-filters, which uses a bigger training dataset[25].

2- Feature extraction: Feature extraction is usually conducted by a convolutional neural network (CNN), which is good to find latent patterns in images[9]. Architectures such as Xu-Net, Ye-Net, Yedroudj-Net and SRNet have had a substantial impact on the area, each proposing different configurations of convolutional layers, activation functions (including TLU or ReLU) and pooling layers to improve performance[25],[9]. Recent models have proposed advanced components such as attention algorithms which allow them to focus on the regions more sensitive to steganography, and lightweight DenseNet (Dense Lite) architectures which help to better retain features and lose less information[24].

3- Classification: The last module is usually a fully connected neural network that takes the retrieved features and classifies the image as a “cover” (original) image or a “stego” (containing hidden data) image. The last layer of the module is SoftMax function, which produces the probabilities of the classes[25].

5.2 Advanced Techniques and Frameworks:

Researchers have studied more complex approaches to increase performance. We also propose an attention mechanism that allows the model to select and highlight the most relevant areas of an image, which helps to extract steganographic information. Another strategy incorporates multi-scale variables by employing a Spatial Pyramid Pooling (SPP) module to increase the sensitivity of the model to weak signals. Moreover, evolutionary-based algorithms frameworks have been provided to tune the network parameters in the training process which can improve detection accuracy and convergence time[24]. Deep learning has brought about a revolution in steganalysis as it enables the automatic extraction of complex features indicative of the presence of hidden information[26]. The first models, like Xu-Net and Ye-Net, proved the possibility of applying CNNs for this purpose, but later works proposed more sophisticated designs, with residual connections (SRNet), attention mechanisms, and new optimization strategies to improve the accuracy and robustness[24],[25],[9].

6. Hybrid / Fusion-Based Methods

6.1 combining Manual and Deep Features in Steganalysis

Hybrid or fusion-based steganalysis methods generally combine the hand-crafted features and the features obtained from deep learning models to improve the detection accuracy. Traditional steganalysis relied on the manual extraction of high-dimensional features, such as those based on the Spatial Rich Model (SRM), which were then used to train a classifier[28]. However, the rise of deep learning, particularly Convolutional Neural Networks (CNNs), has altered the field of steganalysis, enabling automatic feature extraction and achieving a significant enhancement in detection

performance[28]. Several approaches have been proposed to combine these two views. One method uses a hybrid deep learning framework that combines handcrafted features, such as those derived from noise residuals in complex models, with a deep neural network[29]. A framework can be a carefully constructed first stage compatible with the convolution and quantization/truncation phases of complex models and a second stage in the form of a composite deep neural network whose parameters are learned during training[29]. Another approach is to use a CNN with residual-based analysis to extract different handcrafted features to model the noise residuals and then feed them to a robust CNN for classification[9]. Information fusion methodologies can be applied to combine the results of multiple steganalysis methods, whether they are based on manual or deep features[30]. This can be done at several stages, e.g. pre-classification (concatenation of feature vectors) or post-classification (fusion of decision values)[30]. Fusion methods combine the advantages of several steganalysis to improve the accuracy of distinguishing between cover and stego images[30]. One of the powerful options for steganalysis is to mix the manually designed features of well-known rich models and the automatic feature learning ability of deep neural networks. This hybridization combines the domain knowledge of classical approaches with the extended capabilities of deep learning, resulting in more resilient and accurate detection systems.

6.2. Fusion Techniques in Steganalysis

Steganalysis can use information fusion techniques at several levels to combine the results of separate or generic steganalysis methods, which can improve the overall detection efficiency. The application of fusion consists of two primary steps pre-classification or feature level fusion and post-classification which consists of decision level fusion:

1- Fusion of features takes place at the pre-classification stage, wherein feature vectors extracted from multiple steganalysis methods are combined prior to the classifier making a decision[30].

2- Decision level fusion is used after the classification stages where the feature vectors are fed to various classifiers and the output is provided in the form of decision values or the posterior probability[30].

7. Performance and evaluation metrics

According to the sources, the main technique to steganalysis is the binary classification problem: the algorithm should distinguish between “cover” shots and “stego” images[31].

Various statistical and mathematical criteria are used by sources to evaluate the effectiveness of the algorithms in this context, ensuring accuracy and reliability.

The performance and assessment metrics given in the sources can be concisely stated as follows:

Basic detection metrics

• ***Classification Accuracy (PA):*** The main criterion, which shows the ratio of correctly classified samples (both visible and hidden) to the total number of samples[32].

• ***Confusion Matrix:*** A basic visual tool that efficiently displays the results in four categories: True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN)[32].

Reliability and balance measures

• ***Precision:*** The ratio of correctly identified hidden photos by the model to the total count of images labelled as hidden [32].

• ***Specificity:*** Measures the ability to correctly identify “cover” images (i.e., images without any data)[5].

• ***F1 (F1-Score):*** It is the geometric mean of accuracy & recall . It is used to balance between precision and recall especially for imbalanced data[33].

Operational efficiency measures

• ***Area Under the Curve (AUC):*** To assess the capacity of the model to distinguish between classes at various classification criteria, offering a holistic perspective on the performance[4].

• ***Weighted AUC:*** A special statistic, used for example in ALASKA II Challenge, that provides more weight to the areas with low false alarm rates, which is important to real-world security scenarios[34].

• ***G-means:*** calculates the geometric mean of sensitivity and specificity such that it performs equally well for the dominating groups in imbalanced datasets[4].

Image quality and distortion measurements

While masking analysis is concerned with statistical rather than visual features, some researchers use quality measures to evaluate the degree of distortion introduced by masking:

• ***Peak Signal to Noise Ratio (PSNR):*** Used to measure the distortion suffered by the masking in the image[9].

• ***Structural Similarity Index (SSIM):*** Quantifies the preservation of the structure of the original image in the “hidden” image[9].

Additional considerations in the evaluation

• ***Payload capacity:*** Performance of the meter is strongly affected by the “embedding rate”, studies show that small loads (e.g. 0.05 bpp) are hard to detect and have higher mistake rates than large loads[31].

• ***Computational complexity:*** The effectiveness of current models, especially deep learning architectures, is judged by training time, number of parameters, and floating-point operations (FLOPs) to evaluate their feasibility in resource-constrained scenarios[21].

• ***Transferability:*** Measures how well a model trained on a particular dataset can maintain its performance when tested on data from different sources (Cover-Source Mismatch)[31].

8. Datasets

Standard databases provide a necessary basis for evaluation and comparison of steganalysis algorithms, which makes results quantitative and reproducible[31]. Data available sources reveal that databases of different types are used, from small traditional databases to large-scale databases that support deep learning technologies.

The most prominent databases and their associated trends can be summarized in the following points:

1- The most common standard databases

- ***BOSSbase (Break Our Steganographic System):*** is regarded as the most renowned and extensively utilized database in this domain, particularly version(BOSSbase 1.01) [32],[35]. It typically contains 10,000 grayscale images at 512×512 pixels in PGM format[31],[36],[21]. It was explicitly developed for information hiding analysis competitions to evaluate the efficacy of algorithms in relation to content-adaptive concealing[37].
- ***BOWS2 (Break Our Watermarking System 2):*** a popular database of 10,000 grey scale images[32],[38]. They are typically used as extra data for BOSSbase to increase the amount of the training sample and fight the "overfitting" of the deep learning models[36],[39],[23].
- ***ALASKA and ALASKA II:*** These collections are designed to display “real world” settings (Into the Wild). ALASKA II has 80,005 photographs (up to 305,000 in some classifications) in a variety of sizes and formats (color and greyscale)[31],[2]. This database is aimed to integrate laboratory research with real applications providing images from various cameras and multiple processors[31],[20].

2- Large and colorful databases

- ImageNet: The large dataset (almost 14 million photos) was originally developed for classifying things in computer vision, but academics have also used it to examine information hiding[2]. They are used for large scale training and to evaluate the models’ ability to generalize, with some study using millions of images to train convolutional neural networks (CNNs)[29],[40].
- IStego100K: Large database published in 2019, consisting of more than 200,000 color photos with a size of 1024×1024[31],[41]. to foster the development of Universal Steganalysis techniques[31].
- LSSD (Large Scale Steganalysis Database): A modern public collection of 2 million images (color and greyscale)[40]. RAW (Raw) photos were processed under controlled techniques to ensure diversity and reproducibility[40].

3- Other specialized databases

- COCO (Microsoft Common Objects in Context): Sometimes used as an alternative or additional database for testing hybrid models, its uniqueness is defined by the diversity of its content[42],[43].
- CIFAR-10: This data collection has been used to train neural networks to learn patterns in small images (32 by 32). It features 60,000 images that are divided into 10 different classes[44].
- UCID (Uncompressed Color Image Database): A collection of uncompressed color images, typically for testing and comparison[9],[15].

Table 1: Datasets commonly used in steganography and steganalysis research, along with their specific details

.ef		t
3][34] [35]	e	
0][31]	١	
9][38]	t 0	
8][39]	K 4	
][41]	s	
2]	0	rays
[15]	?	

9. Image steganalysis methods

This paper provides a summary of the most common methodologies for Image Steganalysis. This review covers deep learning methods, classical methods, and an overview of hybrid approaches combining the two. Table 2 summarizes each of the studies discussed.

Table 2: Summary of steganalysis Methodologies

			ce Measure (Accuracy)
urning (DNN) with PCA, and Haralick	analysis to distinguish between clean images and predict hidden data quantity	nain BOSS	(LFW) and 84.63%

NN (S-CNN)	steganography in images of arbitrary ut retraining parameters by exploiting onships between sub-regions.	main	1.01 and v2	(for S-UNIWARD t 0.4 bpp payload)
olution Feature Fusion NN)	tection via multi-scale features	main	1.01	or S-UNIWARD at 0.4 d)
Fusion CNN Deep Driven Feature-Based (using SRNet)	robust features that generalize across nbedding algorithms.	main	1.01	verage accuracy across d algorithms like HILL, RD, and WOW)
del (CNN + Bi-LSTM n)	ne detection of weak steganographic	main		(for S-UNIWARD t 0.4 bpp payload)
(Four-stage er)	e Cover-Source Mismatch (CSM) by labeled intermediate domain and fusion (CCF and SCF modules).	PEG	1.01, ALASKA-RFlickr	Average accuracy in d scenarios for S-D at 0.4 bpp)
rning (CNN) with Texture Features	e the automated detectability of hidden olor images by combining statistical correlations with deep learning on.	mages (RGB		
t DenseNet + ttention Mechanism	ve detection accuracy in sensitive nd reduce feature loss during on while maintaining a low parameter	main	1.01 and BOWS2	bpp 90.77%and S-D 0.4 bpp 88.69%
ResNet50 + Big Data BDM)	generic steganography recognition by high-dimensional feature matching idual learning architecture to detect bedding algorithms.	main	1.01 and Custom amples	Maximum accuracy using the combined Net approach)
Learning with Feature ion (VAE + Gabor	e detection of unknown steganographic when only a very small number of nples (Few-Shot) are available.	main Images	1.01 and v2	verage accuracy in a 5-nario for detecting lgorithms)
Net (Fuzzy-based Adaptive High-Pass	e detection of steganographic signals by fuzzy logic to handle pixel es and adaptive filtering for feature nt.	main	1.01, BOWS2, and v2	of 99.07% and an F1 .68%.
arning + Off-policy Differential Evolution	dependence on large labeled datasets by Reinforcement Learning (DRL) for data selection and DE for neter tuning.	main	1.01 and ataset	on BossBase 1.01 and n the BOWS-2 dataset
Information on + Feature tion Loss	enerative steganography (coverless) by intrinsic differences between phic AI-generated images and ordinary ed images.	ted Images	CelebA	
otimization (Pruning, on, Weight)	e model complexity (size and) for resource-constrained devices taining high detection accuracy.	main	1.01	
Ensemble Fisher criminant (FLD) with eatures	rm active/blind steganalysis by specific LSB-based algorithms using sional textural features.	main (LSB)	1.01	
Attention Mechanism	ze feature extraction and improve efficiency by focusing on the most eganographic regions in an image.	main Images	1.01 (20,000 d)	
Statistical is (SVM + PSO)	andom embedding in JPEG images by ration to estimate the original cover and classification via Particle Swarm on (PSO).	nain (Spatial &)	Image Database ed)	
Chain Transition Matrices + AutoML	feature vectors from JPEG images l-to-pixel transition probabilities and arious ML classifiers (Boosting, SVM, ystems).	ges (Transform	BOSSBase 1.01	
Learning (Pretrained ResNet, DenseNet,)	raining time and computational cost by pretrained deep learning models for images as "Cover" or "Stego".	Domain	ζ	
earning Sensitivity SRNet, Yedroudj-Net,	e the impact of different preprocessing M kernels), activation functions, and ategies on the sensitivity and accuracy sed steganalysis.	main Images	1.01	

10. Challenges and Future Directions

10.1 . Challenges.

According to sources, although the field of picture steganalysis has made tremendous progress with the help of deep learning techniques, it still suffers from many fundamental problems that limit its full efficacy in the “real world”. These issues include, but are not limited to, the following:

Cover-Source Mismatch (CSM); This is a highly intricate issue. The detection accuracy drops significantly if the statistical distributions of the training imagery are different from the test images[45]. This difference results from several factors, including:

1- Different imaging technologies, different sensors, different sensitivity settings (ISO)[25].

2- Different digital processing techniques including JPEG compression methods, size reduction operations and filtering[9].

3- The failure of laboratory-trained models to generalize when confronted with images from varied and unfamiliar sources[45].

4- Difficulty in detecting low payload capacities (Low Payload Detection)

When communications are concealed at exceedingly low rates (below 0.01 bits per pixel), the statistical impact of the concealment becomes markedly weak and challenging to differentiate from standard sensor noise. Sources indicate that decreasing the concealment rate is an efficient tactic utilized by hiders to make information imperceptible to contemporary analyzers[49].

5- The gap between laboratory environments and reality (robustness).

Most research is performed in controlled environments that presume an ideal transmission channel, however reality has the following challenges:

• ***Typical transformations:*** Social media images undergo "lossy" operations such as scaling, recurring JPEG compression, cropping and adding noise that destroy or alter masking elements
[49].

• ***Cover type:*** Most of the studies are based on the greyscale photos whereas the real application is the colored JPEG images which have a very different statistical properties[14].

10.2. Future Directions.

This research is designed to increase picture steganalysis by combining several feature domains to improve detection accuracy while minimizing the complexity.

References

[1]Y. Cheng, Z. Luo, and Z. Yin, "Graphical Abstract Robust Steganography with Boundary-Preserving Overflow Alleviation and Adaptive Error Cor-rection Highlights Robust Steganography with Boundary-Preserving Overflow Alleviation and Adaptive Error Cor-rection Robust Steganography with Boundary-Preserving Overflow Alleviation and Adaptive Error Correction."

[2]W. M. Eid, S. S. Alotaibi, H. M. Alqahtani, and S. Q. Saleh, "Digital Image Steganalysis: Current Methodologies and Future Challenges," *IEEE Access*, vol. 10, pp. 92321–92336, 2022, doi: 10.1109/ACCESS.2022.3202905.

[3]N. Farooq and A. Selwal, "Image steganalysis using deep learning: a systematic review and open research challenges," *J. Ambient Intell. Humaniz. Comput.*, vol. 14, no. 6, pp. 7761–7793, Jun. 2023, doi: 10.1007/s12652-023-04591-z.

[4]S. A. R. Al-obaidi, M. Z. Lighvan, and M. T. Alnaseri, "Image steganalysis using reinforcement learning-based active learning and scope loss function," *Intelligent Decision Technologies*, vol. 19, no. 3, pp. 1703–1730, May 2025, doi: 10.1177/18724981241309450.

[5]N. J. D. La Croix, T. Ahmad, and F. Han, "Comprehensive survey on image steganalysis using deep learning," Jul. 01, 2024, *Elsevier B.V.* doi: 10.1016/j.array.2024.100353.

[6]X. Gao, J. Yi, L. Liu, and L. Tan, "A Generic Image Steganography Recognition Scheme with Big Data Matching and an Improved ResNet50 Deep Learning Network," *Electronics (Switzerland)*, vol. 14, no. 8, Apr. 2025, doi: 10.3390/electronics14081610.

[7]Y. Li, B. Ling, D. Hu, S. Zheng, and G. Zhang, "A Deep Learning Driven Feature Based Steganalysis Approach," *Intelligent Automation and Soft Computing*, vol. 37, no. 2, pp. 2213–2225, 2023, doi: 10.32604/iasc.2023.029983.

[8]B. R. Ghosh, S. Banerjee, A. Chakraborty, S. Saha, and J. K. Mandal, "A Deep Learning Based Image Steganalysis Using Gray Level Co-Occurrence Matrix," in *2022 2nd International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies, ICAECT 2022*, Institute of Electrical and Electronics Engineers Inc., 2022. doi: 10.1109/ICAECT54875.2022.9808013.

[9]H. Kheddar, M. Hemis, Y. Himeur, D. Megías, and A. Amira, "Deep learning for steganalysis of diverse data types: A review of methods, taxonomy, challenges and future directions," May 07, 2024, *Elsevier B.V.* doi: 10.1016/j.neucom.2024.127528.

[10]N. Bunzel, M. Steinebach, and H. Liu, "Cover-aware Steganalysis," *Journal of Cyber Security and Mobility*, vol. 10, no. 1, pp. 1–26, 2021, doi: 10.13052/jcsm2245-1439.1011.

[11]Proceedings of the 2nd International Conference on Smart Systems and Inventive Technology (ICSSIT 2019) : 27-29, November 2019. [IEEE], 2019.

[12]J. Lu, G. Zhou, C. Yang, Z. Li, and M. Lan, "Steganalysis of Content-Adaptive Steganography Based on Massive Datasets Pre-Classification and Feature Selection," *IEEE Access*, vol. 7, pp. 21702–21711, 2019, doi: 10.1109/ACCESS.2019.2896781.

[13]Z. Tu, Z. Wang, and X. Zhang, "Feature augmentation-based few-shot image steganalysis," *J. Electron. Imaging*, vol. 34, no. 03, May 2025, doi: 10.1117/1.jei.34.3.033028.

[14]M. T. Sultan and H. El Sayed, "A Deep Learning-based Steganalysis Model for Color Images Using Statistical Texture Features," in *2024 IEEE Global Conference on Artificial Intelligence and Internet of Things, GCAIoT 2024*, Institute of Electrical and Electronics Engineers Inc., 2024. doi: 10.1109/GCAIOT63427.2024.10833592.

[15]D. D. Shankar and A. S. Azhakath, "Random embedded calibrated statistical blind steganalysis using cross validated support vector machine and support vector machine with particle swarm optimization," *Sci. Rep.*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-29453-8.

[16]H. Cao, Z. Wang, and X. Zhang, "On improving steganalysis against cover selection steganography," *Cybersecurity*, vol. 8, no. 1, Dec. 2025, doi: 10.1186/s42400-025-00423-z.

[17]V. Thanasekaran and A. Selvaraj, "Low Dimensional Multi Class Steganalysis of Spatial LSB based Stego Images Using Textural Features," *The International Arab Journal of Information Technology*, vol. 21, no. 2, 2024, doi: 10.34028/iajit/21/2/6.

[18]X. Yu, Y. Ma, Y. Zhang, X. Li, and Y. Zhao, "Fast dominant feature selection with compensation for efficient image steganalysis," *Signal Processing*, vol. 220, Jul. 2024, doi: 10.1016/j.sigpro.2024.109475.

[19]N. J. D. La Croix, T. Ahmad, F. Han, and R. M. Ijtihadie, "HSDetect-Net: A Fuzzy-Based Deep Learning Steganalysis Framework to Detect Possible Hidden Data in Digital Images," *IEEE Access*, vol. 13, pp. 43013–43027, 2025, doi: 10.1109/ACCESS.2025.3546510.

[20]R. Cogranne, M. Chaumont, and P. Bas, "Steganalysis: Detection of hidden data in multimedia content," in *Multimedia Security 1: Authentication and Data Hiding*, Wiley Blackwell, 2022, pp. 247–287. doi: 10.1002/9781119901808.ch8.

[21]Y. Zhu, X. Wang, H.-S. Chen, R. Salloum, and C.-C. J. Kuo, "Green Steganalyzer: A Green Learning Approach to Image Steganalysis."

[22]L. Bohang *et al.*, "Image steganalysis using active learning and hyperparameter optimization," *Sci. Rep.*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-92082-w.

[23]A. Selvaraj, A. Ezhilarasan, S. L. J. Wellington, and A. R. Sam, "Digital image steganalysis: A survey on paradigm shift from machine learning to deep learning based techniques," *IET Image Process.*, vol. 15, no. 2, pp. 504–522, Feb. 2021, doi: 10.1049/ipr2.12043.

[24]Z. He, R. Wu, and X. Wang, "Image Steganalysis Based on an Adaptive Attention Mechanism and Lightweight DenseNet," *Computers, Materials and Continua*, vol. 85, no. 1, pp. 1631–1651, 2025, doi: 10.32604/cmc.2025.067252.

[25]M. Chaumont, "Deep learning in steganography and steganalysis," in *Digital Media Steganography: Principles, Algorithms, and Advances*, Elsevier, 2020, pp. 321–349. doi: 10.1016/B978-0-12-819438-6.00022-0.

[26]Y. Ma, X. Zhang, J. Wang, R. Jin, R. Nasimov, and H. Zhang, "Digital image steganalysis network strengthening framework based on evolutionary algorithm," *Sci. Rep.*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-91390-5.

[27]P. Li, Y. Li, H. Wang, and C. Liu, "Research on steganalysis of digital image based on deep learning," in Proceedings - 2021 4th International Conference on Advanced Electronic Materials, Computers and Software Engineering, AEMCSE 2021, Institute of Electrical and Electronics Engineers Inc., Mar. 2021, pp. 528–534. doi: 10.1109/AEMCSE51986.2021.00114.

[28]M. Hu and H. Wang, "Mutual Information-Optimized Steganalysis for Generative Steganography," *IEEE Transactions on Information Forensics and Security*, vol. 20, pp. 1852–1865, 2025, doi: 10.1109/TIFS.2025.3539089.

[29]J. Zeng, S. Tan, B. Li, and J. Huang, "Large-scale JPEG steganalysis using hybrid deep-learning framework," Nov. 2017, doi: 10.1109/TIFS.2017.2779446.

[30]M. Kharrazi, H. T. Senear, and N. Memon, "Improving steganalysis by fusion techniques: A case study with image steganography," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Springer Verlag, 2006, pp. 123–137. doi: 10.1007/11926214_5.

[31]W. Luo *et al.*, "A Comprehensive Survey of Digital Image Steganography and Steganalysis," *APSIPA Trans. Signal Inf. Process.*, vol. 13, no. 1, 2024, doi: 10.1561/116.20240038.

[32]H. Li and S. Dong, "Image steganalysis algorithm based on deep learning and attention mechanism for computer communication," *J. Electron. Imaging*, vol. 33, no. 01, Jan. 2024, doi: 10.1117/1.jei.33.1.013015.

[33]. S. and M. Kaur, "A Comparative Analysis of Local Binary Pattern(LBP) Variants for Image Tamper Detection," Nov. 17, 2023. doi: 10.21203/rs.3.rs-3608580/v1.

[34]W. You, H. Zhang, and X. Zhao, "A Siamese CNN for Image Steganalysis," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 291–306, 2021, doi: 10.1109/TIFS.2020.3013204.

[35]Z. Wu and S. Wan, "An Image Steganalysis Algorithm Based on Multi-Resolution Feature Fusion," *International Journal of Information Security and Privacy*, vol. 18, no. 1, 2024, doi: 10.4018/IJISP.359893.

[36]A. Kuznetsov, N. Luhanko, E. Frontoni, L. Romeo, and R. Rosati, "Deep Learning Based Image Steganalysis," in *2022 IEEE 9th International Conference on Problems of Infocommunications Science and Technology, PIC S and T 2022 - Proceedings*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 364–368. doi: 10.1109/PICST57299.2022.10238549.

[37]A. Kumar Ranjan, R. Ranjan Sharma, S. Saurav, and M. Tripathi, "STEGANALYSIS WITH CONVOLUTIONAL NEURAL NETWORKS: DETECTING HIDDEN DATA," 2013.

[38]A. Kuznetsov, N. Luhanko, E. Frontoni, L. Romeo, and R. Rosati, "Deep Learning Based Image Steganalysis," in *2022 IEEE 9th International Conference on Problems of Infocommunications Science and Technology, PIC S and T 2022 - Proceedings*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 364–368. doi: 10.1109/PICST57299.2022.10238549.

[39]T. Wu, L. Wang, L. Zhai, C. Fang, and M. Zhang, "Progressive selection-channel networks for image steganalysis," *International Journal of Intelligent Systems*, vol. 37, no. 10, pp. 7444–7458, Oct. 2022, doi: 10.1002/int.22888.

[40]H. Ruiz, M. Yedroudj, M. Chaumont, F. Comby, and G. Subsol, "LSSD: a Controlled Large JPEG Image Database for Deep-Learning-based Steganalysis 'into the Wild,'" Jan. 2021, [Online]. Available: http://arxiv.org/abs/2101.01495

[41]I. Taha Ahmed, B. Tareq Hammad, and N. Jamil, "Image Steganalysis based on Pretrained Convolutional Neural Networks," in *2022 IEEE 18th International Colloquium on Signal Processing and Applications, CSPA 2022 - Proceeding*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 283–286. doi: 10.1109/CSPA55076.2022.9782061.

[42]S. Yang, X. Jia, F. Zou, Y. Zhang, and C. Yuan, "A novel hybrid network model for image steganalysis," *J. Vis. Commun. Image Represent.*, vol. 103, Aug. 2024, doi: 10.1016/j.jvcir.2024.104251.

[43]S. Yang, X. Jia, F. Zou, Y. Zhang, and C. Yuan, "An novel hybrid network model for image steganalysis ⋆," 2024. [Online]. Available: www.elsevier.com/locate/jvci

[44]K. Lichy, P. Lipinski, and M. Grzelak, "Deep Convolutional Network for Steganalysis of HUGO, WOW, and UNIWARD algorithms," in *16th IEEE International Conference on Control, Automation, Robotics and Vision, ICARCV 2020*, Institute of Electrical and Electronics Engineers Inc., Dec. 2020, pp. 421–427. doi: 10.1109/ICARCV50220.2020.9305354.

[45]S. Weng, Y. Li, L. Yu, G. Cao, and L. Chen, "A Labeled Intermediate Domain Aided Two-Domain Correlation Fusion for Mismatched Steganalysis," *IEEE Signal Process. Lett.*, vol. 32, pp. 2937–2941, 2025, doi: 10.1109/LSP.2025.3584296.

[46]G. Ferreira, M. H. da N. Marinho, V. Severo, and F. Madeiro, "Optimization Strategies Applied to Deep Learning Models for Image Steganalysis: Application of Pruning, Quantization and Weight Clustering," *Applied Sciences (Switzerland)*, vol. 15, no. 9, May 2025, doi: 10.3390/app15094632.

[47]A. V. Prokofieva and A. N. Shniperov, “A Markov Chain - Based Method for JPEG Image Steganalysis and Its Application in Combination with Various Machine Learning Algorithms,” *Vestnik NSU. Series: Information Technologies*, vol. 20, no. 4, pp. 61–75, Jun. 2023, doi: 10.25205/1818-7900-2022-20-4-61-75.

[48]R. Tabares-Soto *et al.*, “Sensitivity of deep learning applied to spatial image steganalysis,” *PeerJ Comput. Sci.*, vol. 7, pp. 1–27, 2021, doi: 10.7717/peerj-cs.616.

[49]O. A. Alrusaini, “Deep learning for steganalysis: evaluating model robustness against image transformations,” *Front. Artif. Intell.*, vol. 8, 2025, doi: 10.3389/frai.2025.1532895.