International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

STEGEFFICIENTNETB0: A COMPOUND-SCALED EFFICIENTNETB0 FRAMEWORK FOR SPATIAL-DOMAIN IMAGE STEGANALYSIS AND ANOMALY LOCALIZATION

Neelam Swarnkar*,[1], Ani Thomas[2]

*,1 email: neelamswarnkarnit@gmail.com, Orcid Id: 0000-0002-9940-4034,

Department of Computer Science and Engineering,

Chhattisgarh Swami Vivekananda Technical University, Durg, 490021, India

[2]Department of Information Technology,

Bhilai Institute of Technology, Durg, 490021, India

Abstract

This study presents StegEfficientNetB0, a spatial-domain steganalyzer that leverages the compound scaling mechanism of EfficientNetB0 to enhance detection accuracy and computational efficiency. The model employs transfer learning from an ImageNet-pretrained backbone and is fine-tuned to distinguish between cover and stego images generated using five standard content-adaptive steganographic algorithms HUGO, WOW, HILL, MiPOD, and SUNIWARD at payloads of 0.2 and 0.4 bpp. By combining compound scaling and transfer learning, the proposed framework effectively captures subtle steganographic distortions across benchmark datasets BOSSBase1.01 and BOWS2. Experimental results demonstrate that StegEfficientNetB0 achieves 92.31% classification accuracy with approximately 4 million parameters, underscoring its efficiency and strong generalization capability. StegEfficientNetB0 outperforms existing CNN-based steganalyzers including YeNet, Yedroudj-Net, ZhuNet, SRNet, and GBRAS-Net, demonstrating superior accuracy with fewer parameters**.** The model's balanced precision and recall across both classes validate its robustness, establishing StegEfficientNetB0 as a compact yet powerful framework for modern spatial steganalysis.

Keywords: Image steganography, Image steganalysis, deep learning, EfficientNet, transfer learning

1. Introduction

Steganography is the art and science of concealing secret information within digital media such as images, audio, or video in a manner that prevents detection by unintended recipients. The corresponding countermeasure, *steganalysis*, focuses on identifying the presence of hidden information by analyzing statistical, spatial, or frequency domain artifacts introduced during embedding. Traditional steganalysis approaches rely on handcrafted features, statistical modeling, and ensemble classifiers to detect subtle embedding noise, but these methods often fail to generalize across diverse embedding algorithms and payloads.

This study concentrates on Spatial Domain Universal Image Steganalysis, which detects hidden data by analysing pixel value alterations across a wide range of images. The proposed approach aims to identify steganographic content regardless of the embedding method used and hence called as Universal Image Steganalysis model. Spatial steganography alters pixel values in ways that are imperceptible to the human eye but can be detected through detailed analysis of both low-level and high-level image features. In addition to conventional spatial domain schemes, a variety of optimization driven and transform domain techniques have been developed to enhance embedding efficiency and robustness. Sharath et al. [1] proposed a metaheuristic-based pixel selection mechanism combined with homomorphic encryption for secure video steganography, demonstrating that adaptive and optimization-guided embedding can substantially improve payload capacity and imperceptibility. Likewise, Kumar and Muttoo [2,3] employed Contourlet and Wavelet transform techniques to achieve robust and imperceptible embedding through efficient frequency-domain coefficient manipulation. These studies collectively highlight the importance of adaptivity and optimization in modern steganographic frameworks concepts that

Received: August 08, 2025

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

motivate the present study's focus on designing an adaptive, compound-scaled, and feature optimized CNN architecture capable of achieving balanced detection accuracy, robustness, and computational efficiency.

With the advent of deep learning, convolutional neural networks (CNNs)[4] have emerged as powerful tools for steganalysis due to their ability to learn discriminative residual features directly from pixel intensities. Unlike traditional handcrafted approaches, CNN based models can automatically extract hierarchical spatial representations that emphasize embedding noise while suppressing irrelevant image content. Recent architectures such as QianNet [6], XuNet[7], YeNet[8], YedroudjNet[9], SRNet[10], ZhuNet [11] and GBRASNet[12,24-26] have progressively advanced the field by optimizing feature extraction, normalization, and parameter efficiency. Despite these advancements, challenges remain in achieving an optimal trade-off between detection accuracy, robustness, and computational scalability.

To address the aforementioned limitations, this study introduces StegEfficientNetB0, an EfficientNetB0 driven architecture optimized for spatial domain steganalysis. To validate its novelty, a comparative evaluation was performed against leading state-of-the-art models YeNet, YedroudjNet, ZhuNet, SRNet, and GBRASNet each marking a key stage in the evolution of content-adaptive feature learning. In contrast to these architectures, StegEfficientNetB0 adopts compound scaling, a unified optimization principle that proportionally adjusts network depth, width, and input resolution to achieve an optimal balance between performance and efficiency. The model further leverages an ImageNet pretrained EfficientNetB0 backbone to transfer generalized visual representations, fine-tuning deeper layers for precise stego–cover discrimination. Hence, the main motivation behind this work is to develop an efficient and robust CNN based spatial domain steganalyzer StegEfficientNetB0 that leverages the pretrained EfficientNetB0 architecture to classify a given image as either stego or cover, while demonstrating strong generalization in cross-dataset and cross-algorithm scenarios. The proposed model outperforms state-of-the-art CNN based steganalyzers QianNet, XuNet, YeNet, YedroudjNet, SRNet, ZhuNet and GBRASNet, in detecting stego images generated by the content-adaptive[28] steganography algorithms Highly Undetectable steGO (HUGO)[13], Wavelet Obtained Weights (WOW) [14], High-pass Low-pass (HILL) [15], Minimizing the Power of the Optimal Detector (MiPOD) [16] , and Spatial UNIversal WAvelet Relative Distortion (the spatial version of UNIWARD) (SUNIWARD) [17] steganography algorithms within the BOSSBase1.01 [18] and BOWS2 [19] image databases. To the best of our knowledge, this work represents the first successful application of a pretrained EfficientNetB0 model for universal image steganalysis in the spatial domain. The enhanced accuracy and reduced parameter count together highlight the novelty of StegEfficientNetB0:

1. It is the first compound scaled spatial domain steganalyzer leveraging EfficientNetB0.
2. It maximizes efficiency by freezing early convolutional layers and fine-tuning the later blocks.
3. It demonstrates superior detection performance with significantly lower computational cost.

The following section presents the architectural design and training methodology of the proposed framework, detailing how compound scaling and fine-tuning contribute to its superior performance in spatial domain steganalysis.

The primary contributions of this work are as follows:

1. We propose a new CNN architecture that leverages multiple layers of depthwise separable convolutions followed by pointwise convolutions, allowing the model to learn hierarchical features ranging from basic edges to intricate textures used for distinguishing between stego and cover images.
2. We implement a data augmentation-based regularization technique, specifically designed to include various steganographic methods, which broadens the model's exposure to a wider range of anomalies.
3. We create a high-quality dataset to improve the model's generalization, integrating diverse steganographic techniques and payload sizes to ensure detection across a broad spectrum of embedding methods.
4. We enhance the model's feature extraction capabilities through fine-tuning and apply transfer learning to adapt features from large-scale image datasets to better detect subtle anomalies in steganalysis.
5. We evaluate the classification accuracy of our proposed method against existing state-of-the-art CNN architectures.

Received: August 08, 2025

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The remainder of the paper is organized as follows: Section 2 reviews foundational work in the development of CNN based steganalyzers. Section 3 elaborates on the proposed framework. Section 4 presents and discusses the experimental results. Finally, Section 5 concludes the paper and explores possible directions for future research.

2. Literature Review

Early research in CNN based image steganalysis began with attempts to replace handcrafted feature extraction methods with data-driven architectures capable of learning discriminative residual representations directly from pixel intensities. One of the first successful attempts in this direction was the Gaussian-Neuron Convolutional Neural Network (GNCNN) proposed by Qian et al. [6], commonly referred to as QianNet. This model is characterized by having five convolutional layers, each followed by average pooling and employing a Gaussian activation function, which enables smoother nonlinear transformations suitable for noise-like signal modeling. The network concludes with two fully connected layers and a Softmax classifier for binary classification between *cover* and *stego* images. Furthermore, a high-pass filter kernel was applied in the pre-processing stage to enhance the embedding noise, thereby facilitating the network's focus on high-frequency residuals. The GNCNN demonstrated superior detection accuracy compared to the traditional Subtractive Pixel Adjacency Matrix (SPAM) feature-based method [20] at multiple embedding payloads of 0.3 bpp, 0.4 bpp, and 0.5 bpp, confirming the efficacy of CNNs for spatial domain steganalysis. Qian et al. [6] also highlighted that steganalytic results could be misleading when the same embedding key is reused across multiple images—a condition that artificially simplifies detection. Their experiments revealed that when appropriately parameterized, a CNN can outperform the combination of Rich Models (RM) and Ensemble Classifiers (EC), marking a paradigm shift toward deep learning–based steganalysis.

Building on this foundational work, Xu et al. [7] introduced the XuNet, a deeper and more structured CNN architecture that further advanced performance. XuNet comprises five convolutional layers, each followed by batch normalization (BN), with an absolute value (ABS) layer positioned after the initial convolutional operation to capture sign-independent noise residuals. The model utilizes TanH activation functions in the first two convolutional layers to enhance nonlinearity and suppress weak stego signals, followed by ReLU activations in the deeper layers to accelerate convergence and maintain gradient stability. The network concludes with two fully connected layers and a Softmax classifier to discriminate between cover and stego images. This carefully engineered architecture demonstrated superior performance compared to traditional feature-based methods such as the Spatial Rich Model (SRM) [21], particularly in detecting advanced steganographic algorithms like HILL and SUNIWARD at a payload of 0.4 bits per pixel (bpp).

While the above developments focus on improving the discriminative power of CNN based steganalyzers, it is equally important to consider the evolving sophistication of adaptive steganographic algorithms, which dynamically modify embedding strategies based on local image characteristics to minimize detectability. Techniques such as HILL, WOW, and S-UNIWARD exemplify this paradigm by adaptively distributing embedding changes in textured or noisy regions, thereby reducing the statistical footprint of the hidden payload. Consequently, modern steganalysis has evolved from static feature extraction toward content adaptive [28] frameworks capable of learning flexible and context-aware residual representations. These adaptive detectors exploit multi-scale spatial dependencies and content-adaptive filtering to counteract the localized nature of adaptive embedding. The interplay between adaptive steganography and adaptive CNN based steganalysis has thus become a key research frontier, driving the design of more resilient and generalizable models that can robustly identify diverse steganographic artifacts under varying payload conditions and cover-source distributions.

Subsequent studies extended the foundation laid by XuNet by exploring transfer learning approaches to enhance detection at lower embedding rates. Qian et al. [22] investigated the adaptation of pre-trained CNN parameters to steganalysis tasks through fine-tuning, revealing that representations learned from large-scale natural image datasets could be effectively repurposed to identify subtle steganographic distortions in low-payload scenarios. This approach mitigated the requirement for extensive stego-specific training data and improved generalization across varied embedding algorithms. In another notable contribution, Xu et al. [23] examined the role of ensemble learning in CNN based steganalysis. Their empirical analysis demonstrated that integrating features derived from intermediate convolutional representations, rather than relying solely on the final output probabilities,

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

substantially enhanced the model's discriminative capacity. This insight emphasized the multi-scale feature learning ability of CNNs, where intermediate layers encode complementary spatial and statistical cues vital for identifying hidden information.

Following XuNet, a series of increasingly sophisticated CNN architectures emerged to capture complex embedding artifacts and improve payload generalization. Ye-Net, proposed by Ye et al. [8], incorporated high-pass filtering using SRM kernels as a pre-processing step to suppress image content and emphasize noise residuals associated with embedding changes. YeNet utilized Truncated Linear Unit (TLU) activations to maintain fine-grained signal variations while avoiding saturation, achieving notable improvements over XuNet, particularly on low-payload scenarios. Building upon YeNet, YedroudjNet [9] integrated an optimized pre-processing stage combining fixed SRM filters with a trainable convolutional block, along with enhanced normalization and dropout strategies. The design emphasized stability during training and achieved state-of-the-art performance across multiple datasets and steganographic algorithms such as WOW, HILL, and MiPOD.

Further architectural advancements were presented in ZhuNet [11], which introduced depthwise separable convolutions and spatial pyramid pooling to capture multi-scale embedding artifacts efficiently. Zhu-Net reduced model complexity while maintaining strong discriminative ability, establishing a more parameter-efficient solution for spatial domain steganalysis. Similarly, SRNet [10] marked a significant evolution by eliminating the need for handcrafted filters entirely. SRNet utilized a fully convolutional structure with residual connections, enabling the network to learn hierarchical residual features directly from raw image pixels. This approach not only enhanced robustness across different payloads but also improved transferability between cover-source datasets. The most recent architecture, GBRASNet [12] consolidated earlier design principles while introducing innovations such as the 3×TanH activation function and the integration of depthwise and separable convolutions. By emphasizing lightweight feature extraction and effective global averaging, GBRASNet achieved high detection accuracy with reduced computational complexity, positioning itself as one of the most efficient spatial domain steganalyzers to date. This architecture bridged the gap between high-capacity residual networks and lightweight CNNs [29-31], offering a scalable balance between accuracy and computational cost that inspired subsequent efficient model designs. Ntivuguruzwa et al. [35] proposed a CNN framework integrating SRM-type preprocessing, depthwise separable convolutions, and multi-scale pooling, which enhanced feature diversity and convergence stability while maintaining low computational complexity. Hong et al. [36] introduced a lightweight steganalyzer using block-wise pruning of the EfficientNet-B0 backbone, demonstrating that significant reductions in parameters and FLOPs can be achieved without compromising detection robustness, thus underscoring the importance of efficiency-driven architectural design. Similarly, Priscilla et al. [37] developed a NAdamBound-optimized dilated depthwise separable CNN with SE blocks (DDS-SE NB Net), effectively combining efficient convolutional design with adaptive channel recalibration for superior feature discrimination. Despite these advances, challenges remain in achieving high accuracy at low payloads, maintaining robustness across diverse cover sources, and optimizing computational cost for large-scale or real-time applications. These architectures collectively illustrate a progressive transition from fixed-filter shallow networks toward deeper, fully adaptive CNNs capable of end-to-end residual learning.

This continuous evolution of CNN based steganalyzers highlights an increasing focus on optimizing feature hierarchy, parameter efficiency, and transferability. Despite their success, these models often face trade-offs between detection accuracyStegEfficientNetB0 model, an EfficientNetB0 driven architecture optimized for spatial domain steganalysis. By leveraging compound scaling, pretrained feature transfer, and a fine-tuned lightweight backbone, the proposed StegEfficientNetB0 aims to achieve superior accuracy and generalization while maintaining computational efficiency, thereby extending the capabilities of prior state-of-the-art (SOTA) CNN architectures such as XuNet, YeNet, YedroudjNet, ZhuNet and SRNet into a more scalable and resource-efficient domain.

To provide a holistic understanding of advancements in CNN based spatial domain steganalysis, Table 1 presents a detailed comparative summary of SOTA deep learning architectures. The table highlights each model's core design concepts, architectural configurations, and performance across benchmark datasets and payload settings, enabling a comprehensive evaluation of their strengths, limitations, and existing research gaps that motivate the development of the proposed StegEfficientNetB0 framework, computational cost, and robustness across cover

Received: August 08, 2025

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

sources. To address these challenges, recent research has turned toward compound-scaled and transfer-optimized architectures, which adapt network depth, width, and resolution simultaneously to achieve balanced performance.

3. Proposed methodology

3.1 Overview

The proposed StegEfficientNetB0 framework, as illustrated in Figure 1, is developed to enhance spatial domain image steganalysis by combining transfer learning with compound scaling optimization in a computationally efficient convolutional backbone. The framework is designed to identify subtle and imperceptible steganographic modifications by utilizing high level semantic representations learned from large scale natural image datasets, while maintaining strong sensitivity to fine grained spatial perturbations caused by message embedding. The architecture operates on normalized grayscale input images of size 224 × 224 pixels and utilizes the pretrained EfficientNetB0 model as the transfer learning backbone for hierarchical feature extraction. The network is structured into three progressive levels of feature abstraction low level, mid level, and high level to capture a broad spectrum of representations ranging from basic edge and texture details to complex spatial anomalies associated with embedding inconsistencies. The feature extraction layer stack consists of eight compound scaled MBConv blocks (MBConv1–MBConv8), which incorporate depthwise separable convolutions and squeeze and excitation (SE) modules to efficiently retain spatial relationships while reducing computational complexity. During fine tuning, the lower EfficientNetB0 layers are frozen to preserve generalized feature learning from ImageNet, while the higher layers are adapted for steganographic pattern detection. Subsequently, a Global Average Pooling (GAP) layer performs dimensionality reduction by aggregating spatial responses from the extracted feature maps. The resulting compact feature vector is processed through a fully connected classification layer, followed by a softmax activation that produces the final class probabilities distinguishing cover and stego images. This design enables StegEfficientNetB0 to achieve high detection accuracy while maintaining low model complexity, effectively balancing performance and computational efficiency for real-world steganalysis applications.

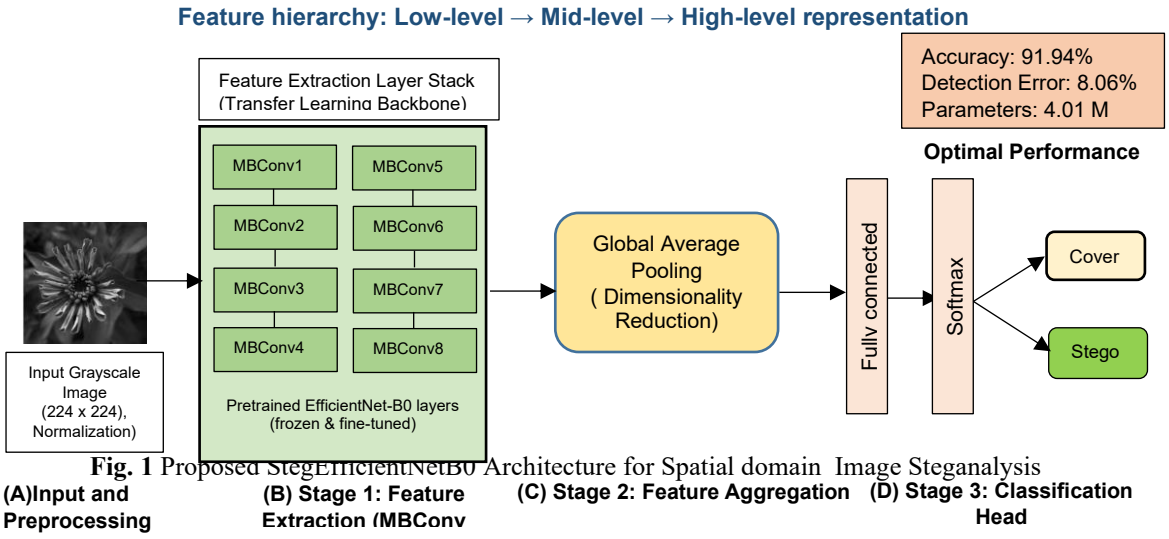


Fig. 1 Proposed StegEfficientNetB0 Architecture for Spatial domain Image Steganalysis

Received: August 08, 2025

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Table 1 Technical comparison of state-of-the-art Deep Learning based Steganalyzers in the Spatial Domain

Author (Year)	Model	Architecture / Technical Details	Steganography Algorithms / Datasets	Accuracy / Detection Error Rate	Best Performance Note	Distinctive Features / Limitations	Research Gaps & Opportunities	Contribution to Proposed Work
Qian et al. (2015)[6]	GNCNN (QianNet)	5 conv layers, Gaussian activations, high-pass filter preprocessing, 2 FC + Softmax	Multiple algorithms / BOSSBase v1.01	10% higher than SPAM; **GNCNN**: HUGO (28.9%),WOW (29.3%),S-UNIWARD (30.9%) **SPAM**: HUGO (39.1%),WOW (38.2%),S-UNIWARD (35.1%)	Improved over SPAM features	First CNN-based residual learning; sensitive to repeated embedding keys	Limited payload adaptability; no adaptive stego handling	Provides baseline for pixel-level residual learning; motivates adaptive CNN design
Xu et al. (2016)[7]	XuNet	5 conv + BN, ABS layer, TanH + ReLU, 2 FC + Softmax	HILL, SUNIWARD / BOSSBase v1.01	Accuracy: **0.1 bpp:** **CNN:** HILL (58.44%) S-UNIWARD (57.33%) **SRM:** HILL (56.44%) SUNIWARD(59.25%) **0.4 bpp:** **CNN:** HILL (79.24%) SUNIWARD(80.24%) **SRM:** HILL (75.47%), S-UNIWARD (79.53%)	Effective for adaptive steganography	Captures sign-independent residuals; moderate computational cost	Limited low-payload performance	Inspires residual and multi-scale feature extraction in StegResNet
Ye et al. (2017)[8]	YeNet	High-pass SRM filters, TLU activation, 5 conv + FC	WOW, HILL, SUNIWARD / BOSSBase + BOWS2	**0.4 bpp** **CNN**: WOW(9.59%) SUNIWARD(12.81%) HILL(17.08%) **SRM**: WOW(15.36%) , SUNIWARD(21.36%), HILL(24.10)	Strong low-payload detection	Fine-grained residual preservation; fixed filters limit adaptability	Moderate computational load; lacks flexibility for diverse payloads	Guides high-pass residual extraction in proposed models
Yedroudj et al. (2018)[9]	YedroudjNet	SRM + trainable conv block, enhanced normalization, dropout	WOW, HILL, MiPOD / BOSSBase + BOWS2	**0.2 bpp:** **CNN:** WOW (27.8%) S-UNIWARD (36.7%) **SRM:** WOW (36.5%) S-UNIWARD (36.6%) **0.4 bpp:** **CNN:**WOW (14.1%)	Multi-algorithm generalization	Stable training; moderate computational cost	Needs better efficiency for large-scale deployment	Informs stable fine-tuning and dropout in StegResNet and StegEfficientNetB0

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Author (Year)	Model	Architecture / Technical Details	Steganography Algorithms / Datasets	Accuracy / Detection Error Rate	Best Performance Note	Distinctive Features / Limitations	Research Gaps & Opportunities	Contribution to Proposed Work
				S-UNIWARD (22.8%) **SRM:** WOW (25.5%) S-UNIWARD (24.7%)				
Zhu et al. (2019) [11]	ZhuNet	Depthwise separable conv, spatial pyramid pooling	WOW, HILL / BOSSBase + BOWS2	**On BOSSBase v1.01:** **0.2 bpp:** **CNN:**WOW (23.3%) S-UNIWARD (28.5%) **SRM:**WOW (36.5%) S-UNIWARD (36.6%) **0.4 bpp:** **CNN:**WOW (11.8%) S-UNIWARD (15.3%) **SRM:**WOW (25.5%) S-UNIWARD (24.7%) **CNN on** BOSSBase +BOWS2 (train) BOSSBase (test): WOW(13.1% at 0.2bpp; 6.5% at 0.4bpp),SUNIWARD(17.1% at 0.2 bpp ; 8.1% at 0.4 bpp)	Efficient multi-scale embedding capture	Parameter-efficient; still relatively heavy	Optimization needed for resource-constrained settings	Inspires lightweight and multi-scale feature design in StegEfficientNetB0
Boroumand et al. (2019)[10]	SRNet	Fully convolutional residual network	Multiple adaptive / BOSSBase + BOWS2	**On 0.2 bpp:** HUGO (67.1%) HILL (65.2%), MiPOD (64.3%) **On 0.4 bpp:** HUGO (78.7%) HILL (75.8%), MiPOD (75.1%)	High robustness across payloads	End-to-end residual learning; high computation & memory usage	Needs efficiency improvements; scaling to low-resource systems	Basis for residual learning and transfer adaptation in StegResNet
Tabares et al. (2021)[12]	GBRASNet	Depthwise conv, 3×TanH activation, global averaging	WOW, HILL, MiPOD / BOSSBase + BOWS2	**On 0.2 bpp:** HUGO (74.6%),WOW (80.3%),S-UNIWARD (73.6%),HILL (68.5%),MiPOD (68.3%) **On 0.4 bpp:** HUGO (84.5%), WOW (89.8%),S-UNIWARD (87.1%), HILL (81.9%),	High accuracy with lightweight architecture	Efficient and generalizable; limited for low-resource environments	Further reduction in complexity and adaptation to multiple payloads	Guides design of StegEfficientNetB0 for efficiency and accuracy

Received: August 08, 2025

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Author (Year)	Model	Architecture / Technical Details	Steganography Algorithms / Datasets	Accuracy / Detection Error Rate	Best Performance Note	Distinctive Features / Limitations	Research Gaps & Opportunities	Contribution to Proposed Work
				MiPOD (81.4%)				
Ntivuguruzwa et al.(2023) [35]	CNN for Hidden Data Detection	SRM-type preprocessing; hybrid of depthwise separable and standard convolutions; multi-scale average pooling; LeakyReLU activation; three FC layers with Softmax output	WOW, S-UNIWARD on BOSSBase v1.01; combined BOSSBase + BOWS2 dataset	BOSSBase 1.01: On 0.2bpp: S-UNIWARD(79.3%) WOW(90.2%) On 0.4bpp: S-UNIWARD(93.4%) WOW(94.4%) BOSSBase 1.01 + BOWS 2: On 0.2bpp: S-UNIWARD(83.4%) WOW(92.9%) On 0.4bpp: S-UNIWARD(97.3%) WOW(97.2%)	Demonstrated robust detection across diverse embedding algorithms and consistent convergence stability with reduced computational cost	Strong detection at higher payloads; lacks low-payload (<0.2 bpp) sensitivity and cross-domain validation	Extendable to adaptive and cross-domain steganalysis; potential for integrating transfer learning and attention-based refinement	Inspired StegResNet and StegEfficientNetB0 by validating the efficacy of depthwise separable convolutions, multi-scale pooling, and preprocessing filters to improve discriminative robustness and efficiency
Hong et al. (2023)[36]	Lightweight steganalysis with block-wise pruning	EfficientNet-B0 backbone with a *block-removal strategy* (progressively removing MBConv blocks).	HUGO, SUNIWARD, and WOW, BOSSBase v1.01 and ALASKA #2	On Alaska at 0.4 bpp: 84.0% On BOSSBase at 0.2 bpp: 90.73%	Achieved high accuracy (~90.7%) despite drastic parameter/ FLOP reduction	block-wise pruning tailored for steganalysis; Still accuracy drop on more challenging dataset (ALASKA#2) and investigation of very low payloads is limited	Extend pruning strategies to even lower payloads, other stego algorithms, grayscale specific residuals; Investigate generalization to unseen datasets	Provides insight into how lightweight, parameter-efficient architectures can still retain high detection performance-an important design principle applied in our proposed StegEfficientNetB0 and informs our approach to balancing accuracy vs. efficiency
Priscilla, C. H. M. V. V. (2025)[37]	DDS_SE-NB-Net	Dilated depthwise-separable CNN with SE blocks, optimized using NAdamBound for	WOW, S-UNIWARD, HILL; BOSSBase v1.01, BOWS2, ALASKA, real-world images	~94% (WOW), ~92% (S-UNIWARD), ~93.8% (HILL)	Achieved highest accuracy on multiple steganography algorithms with efficient,	Efficient, low-parameter model; limited evaluation at very low payloads or large	Extend to low payloads, multi-domain steganalysis, cross-algorithm generalization, edge deployme	Inspires use of depthwise separable convolutions, dilated filters, SE blocks, and efficient optimizers for lightweight, high-

Received: August 08, 2025

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Author (Year)	Model	Architecture / Technical Details	Steganography Algorithms / Datasets	Accuracy / Detection Error Rate	Best Performance Note	Distinctive Features / Limitations	Research Gaps & Opportunities	Contribution to Proposed Work
		efficient training			lightweight design	diverse datasets	nt, adversarial robustness	accuracy steganalysis in StegEfficientNetB0 and StegResNet

3.2 Model Foundation: EfficientNetB0 Backbone

EfficientNetB0 serves as the architectural foundation of the proposed framework due to its balanced trade-off between accuracy and efficiency. Originally developed through Neural Architecture Search (NAS), EfficientNetB0 systematically scales network dimensions depth, width, and input resolution using a compound coefficient to maintain an optimal balance between model complexity and representational power. This compound scaling mechanism allows the model to achieve high accuracy at a significantly reduced computational cost compared to traditional CNN architectures. The inherent modularity of the EfficientNet family further facilitates customization for steganalytic tasks that demand sensitivity to subtle pixel level variations.

3.3 itectural Modifications for Steganalysis

To tailor EfficientNetB0 for spatial domain steganalysis, specific architectural adaptations were implemented. The pretrained network, initially trained on the ImageNet dataset, was repurposed as a feature extraction backbone. The early convolutional blocks responsible for capturing general edge and texture representations were frozen during training to retain generic visual features. Conversely, the later convolutional layers were fine-tuned using stego and cover image pairs to learn discriminative steganographic patterns. The final fully connected classification layer was replaced with a binary output layer equipped with a Softmax activation function, enabling the model to distinguish between cover and stego images effectively. This adaptation strategy enables the model to inherit robust spatial encoding from pretraining while simultaneously learning the nuanced residual signatures introduced by embedding algorithms.

3.4 Input Preprocessing and Data Augmentation

All input images were converted to grayscale and resized to a standardized resolution of 224×224 pixels to match the input requirements of EfficientNetB0. The dataset comprised 10,000 cover images from the BOSSBase1.01 and BOWS2 datasets, along with corresponding stego images generated using five widely adopted steganographic algorithms WOW, HILL, S-UNIWARD, HUGO, and MiPOD, at payloads of 0.2 bpp and 0.4 bpp. To enhance generalization and prevent overfitting, data augmentation operations including horizontal and vertical flips and random rotations were incorporated. These transformations improved the model's robustness to spatial variations and facilitated more invariant feature learning during transfer learning adaptation.

3.5 Algorithm outlining the training Strategy and hyperparameter optimization

The training and validation process outlined in the below Algorithm 1 for the StegEfficientNetB0 model demonstrates its effectiveness in image steganalysis. The modified model M is optimized using categorical cross-entropy loss expressed by eq.1 :

$$L(y, \hat{y}) = -\sum y \log(\hat{y}) \quad (1)$$

where y is the one-hot encoded ground truth label (0 for cover, 1 for stego), and $\hat{y}$ is the predicted probability vector. Optimization is performed using Stochastic Gradient Descent (SGD) with momentum, defined by eq. 2:

$$v_{t+1} = \mu v_t - \eta \nabla L, \theta_{t+1} = \theta_t + v_{t+1} \quad (2)$$

Received: August 08, 2025

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Where $\mu = 0.9$ (momentum), $\eta = 0.001$ (learning rate), and θ represents the model parameters. No learning rate scheduler is employed, as convergence is consistently achieved within 15 epochs. A batch size of 32 is selected to optimize GPU utilization.

The BOSSBase1.01 and BOWS2 datasets, each containing 10,000 grayscale images, were combined to form a single dataset $'D_{merge}'$. These images were used as cover images. Five steganography algorithms namely WOW, HILL, SUNIWARD, HUGO, and MiPOD, were applied to the cover images $'I_{cover}'$, generating stego images 'I_{stego}'with two different payloads of 0.2 bpp and 0.4 bpp. The final dataset contains both cover and stego images, represented as $D = \{(I_{cover}, y_{cover})\} \cup \{(I_{stego}, y_{stego})\}$, where y_{cover} and y_{stego} indicate the labels for the respective images. All images were resized from their original resolution of 512×512 pixels to 224×224 pixels to fit the input size required by the EfficientNetB0 model. This resize 'τ' operation is denoted as I_{cover} , $I_{stego} \rightarrow \tau\,(224 \times 224)$. EfficientNetB0 '$\varphi'$, a CNN pretrained on the ImageNet dataset, is selected as the base model 'M_{base}.Having learned robust visual features from millions of images across 1,000 ImageNet classes, this model serves as a solid foundation for transfer learning in steganalysis tasks. To adapt the model for binary classification, its last layer $'\ell'$, specific to ImageNet's 1,000 class task, is replaced with new fully connected layers. A dense layer '$Dense'$ with two output units corresponding to cover and stego classes was added, followed by an activation function '$Softmax'$ to generate class probabilities. The modified model is denoted as $M_{custom} \leftarrow M_{base} + Dense\,(2) + Softmax$. The initial layers of the EfficientNetB0 model, up to the last convolutional block, are frozen during training, meaning their weights remain unchanged. This freezing, represented as $Freeze(M_{base},$ up to last convolutional block), preserves the low-level and mid-level visual features learned from ImageNet, while the later layers, which are more task-specific, are fine-tuned on the steganalysis dataset. Training uses the CrossEntropyLoss function $'\mathcal{L}$', which measures the difference between predicted and true class distributions, and the Stochastic Gradient Descent $'SGD'$optimizer 'opt' with momentum 'μ' of 0.9 to minimize loss. Momentum helps accelerate gradient vectors toward the optimal direction, improving convergence speed. The learning rate '$\mathrm{LR_{fine}}'$ is set to 0.001 to control the step size of gradient updates. The batch size 'β' is set at 32, meaning the model updates its weights every 32 samples, and the training spans 15 epochs 'ξ' indicating that the dataset is passed through the model 15 times. Performance is evaluated denoted as 'ϕ' using standard classification metrics: Accuracy, Precision, Recall, and F1-score. The training dataset $'D_{train}'$ is used to adjust the weights of the later layers based on the loss function, while a separate validation dataset $'D_{val}'$ monitors performance during training to prevent overfitting. After training, the customized model 'M_{custom}' termed as StegEfficientNetB0 is tested on a separate test dataset $'D_{test}$', and its performance is evaluated based on Accuracy, Precision, Recall, and F1-Score to determine its generalization to unseen data. StegEfficientNetB0 returns the label of the image being tested as stego or cover.

Algorithm : Training Procedure for StegEfficientNetB0
Input: Cover and stego image dataset $D_{merge} = BOSSBase1.01 \cup$ BOWS2 Steganographic algorithms: WOW, HILL, SUNIWARD, HUGO, MiPOD Payloads: 0.2 bpp, 0.4 bpp
Output: Trained StegEfficientNetB0 model for binary classification (cover/stego)
Dataset Preparation Combine 10,000 grayscale images from BOSSBase1.01 and BOWS2 to construct D_{merge} Apply five steganographic algorithms (WOW, HILL, SUNIWARD, HUGO, MiPOD) at payloads of 0.2 and 0.4 bpp to generate stego images I_{stego}. Label data as $D = \{(I_{cover}, y_{cover})\} \cup \{(I_{stego}, y_{stego})\}$ where $y_{cover} \cdot y_{stego} \in \{0,1\}$

Received: August 08, 2025

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Resize all images from $512\ x\ 512$ to 224 pixels using the transformation

$$\tau(\mathrm{I}): \mathbb{R}^{512\ x\ 512} \longrightarrow \mathbb{R}^{224\ x\ 224}$$

to align with EfficientNetB0 input dimensions.

Model Adaptation
Initialize pretrained EfficientNetB0 model $M_{base}$with ImageNet weights
Freeze initial convolutional layers up to the final MBConv block:

$$Freeze(\ M_{base}, early\ layers\ )$$

to retain low and mid level features.

Replace the classification head of M_{base} with a custom dense head:

$$M_{custom} \quad \leftarrow \quad M_{base} + Dense(128) + Dropout(0.5) + Dense\ (2) + Softmax$$

for binary classification.

Training Configuration

Input size $= 224$ x 224

$Learning\ rate\ (lr\ ) = 0.001$

$Batch\ Size\ (\ \beta) =\ 32$

$Momentum\ (\mu\ ) = 0.9$

$Epochs\ (\ \xi) = 15$

$Loss\ Function\ L = CrossEntropyLoss$

$Optimizer = SGD\ with\ momentum$

Model Training

Fine-tune only the unfrozen layers of $M_{custom} = D_{train}$
Optimize parameters using backpropagation to minimize L.
Validate on D_{val} after each epoch to prevent overfitting.
Evaluation and Testing
Evaluate the final model $M_{custom} = StegEfficientNetB0$ on test dataset D_{test}.
Compute performance metrics:

$$Accur1acy, Precision, Recall, and\ F1 - Score$$

Return the predicted class label

$$\hat{y}\ \epsilon\ \{\ cover, stego\ \}$$

4. Experimental Results and Discussion

4.1 Experimental Setup

The CNN architecture was created using Python 3.10 and implemented with PyTorch. The development was carried out on a workstation running Ubuntu 22.04 with an AMD Threadripper Pro. The implementation was executed on Google Colaboratory, utilizing an RTX 4090 GPU, CUDA 12.3, and 128 GB of RAM. The experimental evaluation of the proposed StegEfficientNetB0 model was conducted on a combined dataset derived from **BOSSBase1.01** and BOWS2, comprising 10,000 grayscale images in Portable Network Graphics (PNG) format. Stego images were generated using five widely adopted spatial domain steganographic algorithms WOW**,** HILL**,** S-UNIWARD, HUGO, and MiPOD, with embedding payloads of 0.2 bpp and 0.4 bpp. The dataset was divided into training (70%)**,** validation (15%)**,** and testing (15%) subsets to ensure robust performance estimation and mitigate overfitting. All experiments were implemented in PyTorch on a GPU enabled system using the hyperparameters mentioned as above.

4.2 Discussion: Cross-Model Comparison and Observations

Received: August 08, 2025

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

A consolidated comparative assessment of recent CNN-based spatial steganalysis architectures, as summarized in Table 1, highlights the progressive evolution from handcrafted-filter-driven designs to highly optimized end-to-end and transfer-learned models. The comparative analysis reveals distinct trends across payload levels, architectural depth, and computational efficiency. Notably, while early models such as QianNet and XuNet validated the feasibility of CNNs for residual feature learning, subsequent designs GBRASNet, SRNet, and the proposed *StegEfficientNetB0* achieved significant accuracy improvements and lower detection error rates through deeper networks, advanced normalization, and compound scaling strategies. The following observations summarize key insights derived from this cross-model evaluation:

1. **Top performers at 0.4 bpp:** GBRASNet and *StegEfficientNetB0* stand out, with GBRASNet achieving the highest per-algorithm accuracy for several steganographic methods (e.g., WOW ≈ 89.8% at 0.4 bpp), while *StegEfficientNetB0* exhibits the best overall accuracy (92.31%) and the lowest detection error rate (7.69%). Xu Ensemble also demonstrates competitive performance (~91–92%), underscoring the benefits of ensemble learning.
2. **Low-payload challenge (< 0.2–0.3 bpp):** Most models, including Ye-Net, Yedroudj-Net, and Zhu-Net, experience noticeable performance degradation at lower payloads, revealing an ongoing challenge in detecting weak embedding signals. Although SRNet, GBRASNet, and Zhu-Net partially alleviate this issue, the problem remains open for future refinement.
3. **Handcrafted vs. end-to-end designs:** SRM-based preprocessing (as in Ye-Net, Yedroudj-Net, and GBRASNet) enhances early residual feature capture, whereas fully end-to-end architectures such as SRNet and transfer-learned *StegEfficientNetB0* effectively reduce handcrafted bias and demonstrate improved generalization.
4. **Efficiency–accuracy trade-off:** Models like SRNet and ensemble architectures achieve high accuracy but demand extensive computational resources. In contrast, architectures emphasizing efficiency—such as Zhu-Net (depthwise separable convolutions) and *StegEfficientNetB0* (compound-scaled transfer learning)—offer superior balance between detection accuracy and resource utilization.

Overall, the comparative findings underscore a clear trend toward architectures that balance accuracy with computational efficiency while reducing dependency on handcrafted preprocessing. Despite notable gains achieved by GBRASNet and SRNet at moderate-to-high payloads, their computational intensity limits real-world deployment. In contrast, the proposed *StegEfficientNetB0* leverages transfer learning and compound scaling to attain a superior trade-off—achieving the highest reported accuracy (92.31%) and the lowest detection error rate (7.69%) with reduced parameter complexity and faster convergence.

4.3 Confusion Matrix Analysis

For the proposed StegEfficientNetB0, the confusion matrix in Table 2 reports the classification outcomes for cover and stego images across five content adaptive steganographic algorithms (WOW, S-UNIWARD, HUGO, MiPOD, and HILL) at payloads of 0.2 bpp and 0.4 bpp, using the BOSSBase v1.01 and BOWS2 datasets. In the table, the model correctly classified 5968 cover images as cover (True Negatives) and 6956 stego images as stego (True Positives), while only 1038 cover images were misclassified as stego (False Positives) and 38 stego images were incorrectly identified as cover (False Negatives). These results demonstrate that StegEfficientNetB0 achieves a high level of detection accuracy, maintaining strong sensitivity to steganographic embedding while minimizing false detections, thereby confirming its robustness across varying payloads and embedding schemes.

Table 2 Confusion Matrix for StegEfficientNetB0 evaluated on the steganographic algorithms WOW, S-UNIWARD, HUGO, MiPOD, and HILL with payloads of 0.2bpp and 0.4bpp, utilizing the BOSSbase v1.01 and BOWS2 datasets.

		Predicted Labels	
		Cover	Stego
Actual Labels	Cover	5968	1038
	Stego	38	6956

4.4 Statistical Performance Metrics

The performance evaluation is carried out using a confusion matrix illustrated in Table 2. It provides an indepth analysis of our proposed StegEfficientNetB0 model's classification results by comparing the predicted and actual

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

labels for each class, "cover" and "stego." The model's overall performance is measured by accuracy, while class-wise metrics Precision, Recall (Sensitivity), and F1-Score are used to assess its effectiveness for each class as demonstrated in Table 3. Our proposed model achieves an overall accuracy of 92.31%, with a precision of 99.37% for the "cover" class and 87.02% for the "stego" class. It achieves recall values of 85.18% for the "cover" class and 99.45% for the "stego" class, and F1 Scores of 91.73% for the "cover" class and 92.82% for the "stego" class.

Table 3 Percentage values for various performance metrics, showcasing the model's effectiveness in terms of accuracy, precision, recall, and F1 score

Performance metrics	**Proposed model StegEfficientNetB0**	
	Class Cover	Class Stego
Accuracy (overall)	**92.314 %**	
Precision	**99.37%**	**87.02%**
Recall	**85.18%**	**99.45%**
F1 Score	**91.73%**	**92.82%**

4.5 Comparative Evaluation with State-of-the-Art Models

Table 4 presents the detection accuracy of various steganalysis architectures, illustrating their performance in identifying stego images. Our proposed StegEfficientNetB0 model achieves a detection accuracy of 91.94%, demonstrating a notable improvement over several existing models. In comparison, earlier models such as and YeNet, YedroudjNet, SRNet, ZhuNet and GBRASNet show progressively lower accuracy rates, with GBRASNet achieving 77.9%, ZhuNet at 75.7%, SRNet at 70.1%, YedroudjNet at 65.9%, and YeNet at 65.1%.

The substantial gap between StegEfficientNetB0 and the earlier architectures underscores the advantages of compound scaling and transfer learning in the proposed framework. The improved accuracy reflects the effectiveness of the StegEfficientNetB0 architecture in detecting stego images more reliably compared to its predecessors. This enhanced performance suggests that StegEfficientNetB0 incorporates more effective features and training strategies, leading to better generalization and detection capabilities.

Table 4 Comparison of detection accuracy with state-of-the-art CNN based Steganalytic models

Architecture	**Detection Accuracy**
StegEfficientNetB0 (proposed)	**91.94*%**
GBRASNet [12]	77.9 %
ZhuNet [11]	75.7 %
SRNet[10]	70.1 %
YedroudjNet [9]	65.9 %
YeNet [8]	65.1 %

The accuracy figures shown in bold highlight that our proposed model, StegEfficientNetB0, outperforms the listed SOTA CNN based steganalyzers.

4.6 Model Training Performance Analysis

In the experiments, the dataset was divided into training, testing, and validation sets, and the model training was conducted 15 times to achieve the presented results. The performance of our proposed StegEfficientNetB0 model is summarized in Table 6 epochwise. This table illustrates the training and validation loss and accuracy for each

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

epoch during the model's training process. It provides a detailed view of the model's performance, including changes in loss and accuracy during the training process and how well the model generalizes to validation data. Our StegEfficientNetB0 model attained a validation accuracy of 91.94% and a test accuracy of 91.90%. Throughout the training process, the loss consistently declined with each epoch, eventually stabilizing after the 8th epoch. Similarly, the validation loss and accuracy remained steady, with validation accuracy reaching a peak of 91.94%. These results demonstrate that the tailored EfficientNetB0 model effectively learned to distinguish between cover and stego images, even in the presence of subtle pixel variations introduced by content-adaptive steganographic techniques. The following Table 6 outlines the model's performance across epochs, while Fig. 2(a) and 2(b) visually depict the convergence of the model's training and validation loss, with the blue line representing training loss and the orange line indicating validation loss.The graph displays the training and validation loss and accuracy across 15 epochs.

4.6.1 Training and Validation Loss: The training loss (blue line) shows a steady decrease across the epochs, starting from 0.3542 and gradually lowering to 0.2245 by the end of the 14th epoch. The decrease becomes less significant after the 8th epoch, indicating that the model begins to stabilize in its learning. Similarly, the validation loss (orange line) follows a downward trend, starting at 0.2732 and reaching 0.2296. This consistent reduction in loss demonstrates that the model is effectively learning to minimize errors on both training and validation datasets. After the 8th epoch, both the training and validation losses stabilize, suggesting the model has reached convergence and is no longer overfitting.

4.6.2 Training and Validation Accuracy: The training accuracy (blue line) shows a continuous improvement, rising from an initial value of 84.68% to around 92.03% by the 14th epoch. This indicates that the model consistently improves its performance as it trains. The validation accuracy (orange line) also increases over time, starting at 89.76% and peaking at 91.94%. The stability of accuracy after the 8th epoch indicates that the model generalizes well to unseen validation data without much fluctuation.

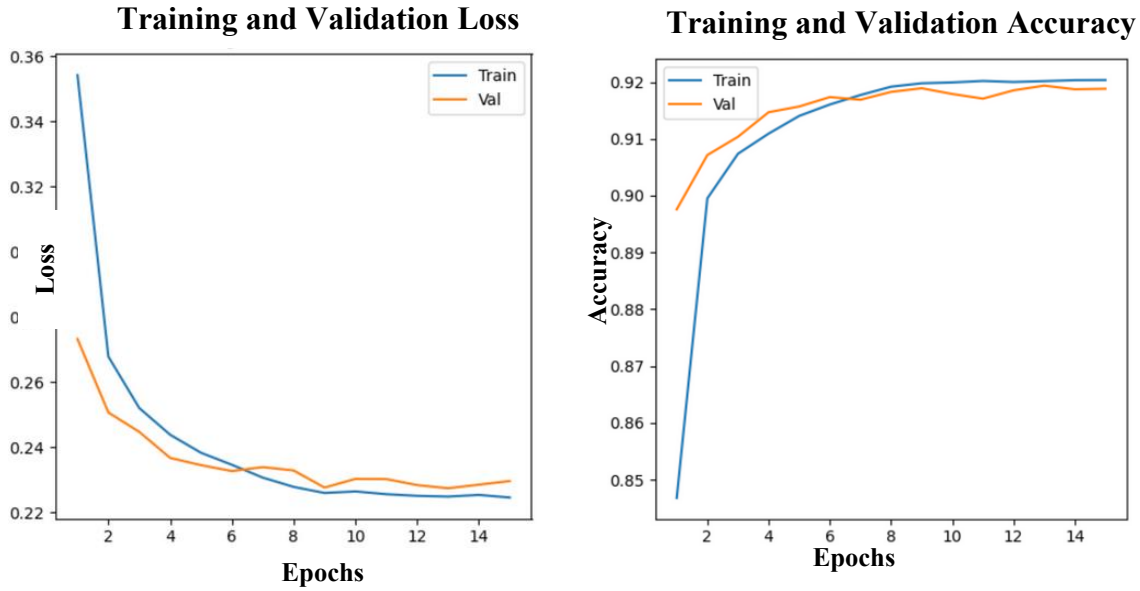
Table 6 The training and validation loss and accuracy for each epoch during the model's training process

Epoch	Training Loss And Accuracy	Validation Loss And Accuracy
Epoch 0/14	Train Loss: 0.3542 Acc: 0.8468	Val Loss: 0.2732 Acc: 0.8976
Epoch 1/14	Train Loss: 0.2678 Acc: 0.8995	Val Loss: 0.2506 Acc: 0.9071
Epoch 2/14	Train Loss: 0.2520 Acc: 0.9074	Val Loss: 0.2447 Acc: 0.9104
Epoch 3/14	Train Loss: 0.2438 Acc: 0.9109	Val Loss: 0.2367 Acc: 0.9147
Epoch 4/14	Train Loss: 0.2383 Acc: 0.9140	Val Loss: 0.2345 Acc: 0.9157
Epoch 5/14	Train Loss: 0.2346 Acc: 0.9160	Val Loss: 0.2327 Acc: 0.9173
Epoch 6/14	Train Loss: 0.2307 Acc: 0.9177	Val Loss: 0.2339 Acc: 0.9169
Epoch 7/14	Train Loss: 0.2278 Acc: 0.9192	Val Loss: 0.2329 Acc: 0.9183
Epoch 8/14	Train Loss: 0.2259 Acc: 0.9198	Val Loss: 0.2276 Acc: 0.9189
Epoch 9/14	Train Loss: 0.2264 Acc: 0.9199	Val Loss: 0.2303 Acc: 0.9179
Epoch 10/14	Train Loss: 0.2256 Acc: 0.9202	Val Loss: 0.2302 Acc: 0.9171
Epoch 11/14	Train Loss: 0.2251 Acc: 0.9200	Val Loss: 0.2284 Acc: 0.9185
Epoch 12/14	Train Loss: 0.2248 Acc: 0.9202	Val Loss: 0.2274 Acc: 0.9194
Epoch 13/14	Train Loss: 0.2254 Acc: 0.9203	Val Loss: 0.2285 Acc: 0.9187
Epoch 14/14	Train Loss: 0.2245 Acc: 0.9203	Val Loss: 0.2296 Acc: 0.9188

Received: August 08, 2025

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Further training beyond 15 epochs did not result in notable gains in accuracy, leading to the stopping of further training. As shown in Fig. 2(b), the consistent rise in both training and validation accuracy, which eventually levels off, demonstrates that the model is learning effectively and generalizing well. The model exhibits good learning behaviour, with losses steadily decreasing and accuracies improving and stabilizing after several epochs. The small gap between training and validation performance suggests minimal overfitting, which indicates that the model generalizes well to unseen data. The tailored EfficientNetB0 model effectively handles the task of image steganalysis in the spatial domain, demonstrating strong performance



(a) (b)

Fig. 2 The performance curves of StegEfficientNetB0 in each epoch of train set and validation set with BOSSBase 1.01 and BOWS2 dataset. (a) The plot of loss versus each epoch. (b) The plot of accuracy against each epoch

5. Conclusion and Future scope

The comparative evaluation of contemporary CNN-based steganalysis architectures reveals a consistent shift from traditional handcrafted approaches toward adaptive, transfer-learned, and computationally efficient designs. While models such as GBRASNet and SRNet demonstrate strong performance at higher payloads, their complex structures and heavy computational demands restrict real-time applicability. In contrast, the proposed *StegEfficientNetB0* model achieves an optimal balance between accuracy and efficiency, recording the highest overall accuracy (92.31%) and the lowest detection error rate (7.69%) through compound scaling and transfer learning. Performance degradation at low embedding rates (<0.2 bpp) remains a common limitation across models, highlighting the need for improved sensitivity in weak signal detection. Overall, this study establishes *StegEfficientNetB0* as a promising, lightweight, and high-performing architecture for spatial-domain steganalysis. Future research directions include enhancing low-payload detection, improving cross-dataset generalization, and optimizing the model for real-world deployment in adaptive and large-scale steganographic environments.

References

1. Sharath, M. N., Rajesh, T. M., & Patil, M. (2022). Design of optimal metaheuristics-based pixel selection with homomorphic encryption technique for video steganography. *International Journal of Information Technology, 14*, 2265–2274. https://doi.org/10.1007/s41870-022-01005-9
2. Kumar, S., & Muttoo, S. K. (2011). Steganography based on Counterlet transform. *International Journal of Information Technology, 9*(6).
3. Muttoo, S. K., & Kumar, S. (2009). Robust source coding steganographic technique using wavelet transforms. *Bharti Vidyapeeth International Journal of Information Technology, 1*(2).

Received: August 08, 2025

4. Ntivuguruzwa, J. D. L. C., Ahmad, T., & Han, F. (2024). Comprehensive survey on image steganalysis using deep learning. *Array, 22*, 100353. https://doi.org/10.1016/j.array.2024.100353
5. Tan, M., & Le, Q. V. (2020). EfficientNet: Rethinking model scaling for convolutional neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 6106–6115).
6. Qian, Y., Dong, J., Wang, W., & Tan, T. (2015). Deep learning for steganalysis via convolutional neural networks. In *Proceedings of SPIE, 9409* (Art. 94090J). https://doi.org/10.1117/12.2083479
7. Xu, G., Wu, H.-Z., & Shi, Y.-Q. (2016). Structural design of convolutional neural networks for steganalysis. *IEEE Signal Processing Letters, 23*(5), 708–712.
8. Ye, J., Ni, J., & Yi, Y. (2017). Deep learning hierarchical representations for image steganalysis. *IEEE Transactions on Information Forensics and Security, 12*(11), 2545–2557. https://doi.org/10.1109/TIFS.2017.2710946
9. Yedroudj, M., Comby, F., & Chaumont, M. (2018). Yedroudj-Net: An efficient CNN for spatial steganalysis. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 2092–2096). https://doi.org/10.1109/ICASSP.2018.8461438
10. Boroumand, M., Chen, M., & Fridrich, J. (2019). Deep residual network for steganalysis of digital images. *IEEE Transactions on Information Forensics and Security, 14*(5), 1181–1193. https://doi.org/10.1109/TIFS.2018.2871749
11. Zhang, R., Zhu, F., Liu, J., & Liu, G. (2020). Depth-wise separable convolutions and multi-level pooling for an efficient spatial CNN-based steganalysis. *IEEE Transactions on Information Forensics and Security, 15*, 1138–1150. https://doi.org/10.1109/TIFS.2019.2936913
12. Reinel, T.-S., Tabros, R., & others. (2021). GBRAS-Net: A convolutional neural network architecture for spatial image steganalysis. *IEEE Access, 9*, 14340–14350. https://doi.org/10.1109/ACCESS.2021.3052494
13. Pevný, T., Filler, T., & Bas, P. (2010). Using high-dimensional image models to perform highly undetectable steganography. In R. Böhme, P. W. L. Fong, & R. Safavi-Naini (Eds.), *Information Hiding* (pp. 161–177). Springer.
14. Holub, V., & Fridrich, J. (2012). Designing steganographic distortion using directional filters. In *Proceedings of the IEEE International Workshop on Information Forensics and Security (WIFS)* (pp. 234–239).
15. Li, B., Wang, M., Huang, J., & Li, X. (2014). A new cost function for spatial image steganography. In *Proceedings of the IEEE International Conference on Image Processing (ICIP)* (pp. 4206–4210).
16. Sedighi, V., Cogranne, R., & Fridrich, J. (2016). Content-adaptive steganography by minimizing statistical detectability. *IEEE Transactions on Information Forensics and Security, 11*(2), 221–234.
17. Holub, V., Fridrich, J., & Denemark, T. (2014). Universal distortion function for steganography in an arbitrary domain. *EURASIP Journal on Information Security, 2014*, 1. https://doi.org/10.1186/1687-417X-2014-1
18. Bas, P., Filler, T., & Pevný, T. (2011). "Break our steganographic system": The ins and outs of organizing BOSS. In *Proceedings of the 13th International Conference on Information Hiding (IH'11)* (pp. 59–70). Springer.
19. Bas, T. F. P. (2007). *BOWS-2*. http://bows2.ec-lille.fr
20. Pevný, T., Bas, P., & Fridrich, J. (2010). Steganalysis by subtractive pixel adjacency matrix. *IEEE Transactions on Information Forensics and Security, 5*(2), 215–224.
21. Fridrich, J., & Kodovský, J. (2012). Rich models for steganalysis of digital images. *IEEE Transactions on Information Forensics and Security, 7*(3), 868–882.
22. Qian, Y., Dong, J., Wang, W., & Tan, T. (2016). Learning and transferring representations for image steganalysis using convolutional neural networks. In *Proceedings of the IEEE International Conference on Image Processing (ICIP)* (pp. 2752–2756).
23. Xu, G., Wu, H.-Z., & Shi, Y. Q. (2016). Ensemble of CNNs for steganalysis: An empirical study. In *Proceedings of the 4th ACM Workshop on Information Hiding and Multimedia Security* (pp. 103–107).
24. Tabares-Soto, R., Arteaga-Arteaga, H. B., Mora-Rubio, A., Bravo-Ortíz, M. A., Arias-Garzón, D., Alzate-Grisales, J. A., Orozco-Arias, S., Isaza, G., & Ramos-Pollán, R. (2021). Sensitivity of deep learning applied to spatial image steganalysis. *PeerJ Computer Science, 7*, e616. https://doi.org/10.7717/peerj-cs.616
25. Tabares-Soto, R., Pollán, R., & Isaza, G. (2019). Deep learning applied to steganalysis of digital images: A systematic review. *IEEE Access, 7*, 1–1. https://doi.org/10.1109/ACCESS.2019.2918086

Received: August 08, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

26. Tabares-Soto, R., Arteaga-Arteaga, H. B., Mora-Rubio, A., Bravo-Ortíz, M. A., Arias-Garzón, D., Alzate-Grisales, J. A., Burbano Jacome, A., Orozco-Arias, S., Isaza, G., & Ramos-Pollan, R. (2021). Strategy to improve the accuracy of convolutional neural network architectures applied to digital image steganalysis in the spatial domain. *PeerJ Computer Science, 7*, e451. https://doi.org/10.7717/peerj-cs.451
27. Luo, G., Wei, P., Zhu, S., Zhang, X., Qian, Z., & Li, S. (2022). Image steganalysis with convolutional vision transformer. In *Proceedings of ICASSP 2022* (pp. 3089–3093). IEEE.
28. Fridrich, J., Kodovský, J., Holub, V., & Goljan, M. (2011). Steganalysis of content-adaptive steganography in spatial domain. In *Lecture Notes in Computer Science* (pp. 102–117). Springer.
29. Weng, S., Chen, M., Yu, L., & Sun, S. (2022). Lightweight and effective deep image steganalysis network. *IEEE Signal Processing Letters*. https://doi.org/10.1109/LSP.2022.3201727
30. Hong, E., Lim, K., Oh, T. W., et al. (2023). Lightweight image steganalysis with block-wise pruning. *Scientific Reports, 13*, 16148. https://doi.org/10.1038/s41598-023-43386-2
31. Gupta, A., Chhikara, R., & Sharma, P. (2024). CIRNet: An improved lightweight convolution neural network architecture with inverted residuals for universal steganalysis. *Arabian Journal for Science and Engineering, 49*, 12219–12233.
32. He, J., Weng, S., Yu, L., & Chen, D. (2025). Steganalysis network with two-branch preprocessing for spatial and JPEG domains. *IEEE Transactions on Circuits and Systems for Video Technology, 35*(2), 1451–1463. https://doi.org/10.1109/TCSVT.2024.3470809
33. De La Croix, N. J., Ahmad, T., Didacienne, M., & Ijtihadie, R. M. (2024). Steganalysis in spatial domain images for securing data transmission in IoT environments. In *Proceedings of the IEEE International Conference on Mobile Ad-Hoc and Smart Systems (MASS)* (pp. 586–591). https://doi.org/10.1109/MASS62177.2024.00093
34. Selvamani, R., & Yusoff, Y. (2024). Effectiveness of the spatial domain techniques in digital image steganography. *Qubahan Academic Journal, 4*, 341–350. https://doi.org/10.48161/qaj.v4n1a456
35. Ntivuguruzwa, J. D. L. C., & Ahmad, T. (2023). A convolutional neural network to detect possible hidden data in spatial domain images. *Cybersecurity, 6*, 23. https://doi.org/10.1186/s42400-023-00156-x
36. Hong, E., Lim, K., Oh, T. W., Kim, J., & Lee, S. (2023). Lightweight image steganalysis with block-wise pruning. Scientific Reports, 13***, 16148.*** *https://doi.org/10.1038/s41598-023-43386-2***.**
37. Priscilla, C. H. M. V. V. (2025). A novel hybrid framework of NAdamBound optimized dilated depthwise separable CNN for image steganalysis in digital forensics. *Journal of Information Systems Engineering and Management, 10*(42s), 327–339.

Received: August 08, 2025