Artificial Intelligence and Robotics Research 人工智能与机器人研究, 2026, 15(2), 420-428
Published Online March 2026 in Hans. https://www.hanspub.org/journal/airr
https://doi.org/10.12677/airr.2026.152041

Hans汉斯

FADG：频域感知的自适应中间域生成方法用于载体源失配的隐写分析

郑　涵

北京印刷学院信息工程学院，北京

收稿日期：2026年1月22日；录用日期：2026年2月14日；发布日期：2026年3月3日

摘　要

随着数字媒体的广泛应用，信息隐藏技术在保密通信中扮演着重要角色，但同时也被恶意利用进行非法信息传播，对网络安全构成严重威胁。隐写分析作为检测和识别隐蔽信息的关键技术，在维护信息安全方面具有重要价值。近年来，深度学习的引入显著提升了隐写分析的检测性能，众多基于卷积神经网络的隐写分析模型取得了优异成果。然而，这些隐写分析模型面临一个严峻的挑战：当训练数据(源域)与实际检测数据(目标域)来源不一致时，模型性能急剧下降，这一现象被称为载体源失配(Cover Source Mismatch, CSM)问题。CSM问题严重制约了深度隐写分析模型的实用化进程。本文围绕CSM问题展开研究，提出了一种有效的跨域隐写分析方法，频域感知自适应中间域生成方法(Frequency-Aware Adaptive Domain Generation, FADG)。FADG通过光谱残差理论识别图像中频域能量波动大的区域，进而计算源域和目标域图像的频域隐写嵌入概率图，接着利用源域和目标域的嵌入概率图得到混合权重图用于生成中间域，从而实现从源域到目标域的平滑过渡。通过这种自适应生成中间域的方法，从而有效缓解了跨域隐写检测性能下降的问题。

关键词

隐写分析，载体源失配，中间域，光谱残差

FADG: Frequency-Aware Adaptive Domain Generation for Cover Source Mismatch in Steganalysis

Han Zheng

School of Information Engineering, Beijing Institute of Graphic Communication, Beijing

Received: January 22, 2026; accepted: February 14, 2026; published: March 3, 2026

文章引用: 郑涵. FADG: 频域感知的自适应中间域生成方法用于载体源失配的隐写分析[J]. 人工智能与机器人研究, 2026, 15(2): 420-428. DOI: 10.12677/airr.2026.152041

Abstract

With the widespread application of digital media, information hiding technology plays an important role in secure communication, but it is also maliciously exploited for illegal information dissemination, posing serious threats to network security. Steganalysis, as a key technology for detecting and identifying covert information, holds significant value in maintaining information security. In recent years, the introduction of deep learning has significantly improved the detection performance of steganalysis, with numerous steganalysis models based on convolutional neural networks achieving excellent results. However, these steganalysis models face a severe challenge: when the training data (source domain) and the actual detection data (target domain) originate from different sources, model performance drops sharply, a phenomenon known as the Cover Source Mismatch (CSM) problem. The CSM problem severely restricts the practical application of deep steganalysis models. This paper focuses on the CSM problem and proposes an effective cross-domain steganalysis method, Frequency-Aware Adaptive Domain Generation (FADG). FADG identifies regions with large frequency-domain energy fluctuations in images through spectral residual theory, then calculates frequency-domain steganographic embedding probability maps for both source and target domain images. Subsequently, it utilizes the embedding probability maps of the source and target domains to obtain blending weight maps for generating intermediate domains, thereby achieving a smooth transition from the source domain to the target domain. Through this adaptive intermediate domain generation method, the problem of performance degradation in cross-domain steganalysis detection is effectively alleviated.

Keywords

Steganalysis, Cover Source Mismatch (CSM), Intermediate Domain, Spectral Residual

1. 引言

随着数字媒体的普及，隐写术作为一种秘密通信技术，能够将秘密信息嵌入图像、音频、视频等数字载体中来实现隐蔽通信的技术[1] [2]。隐写术在保障信息秘密传输的同时，对于外来信息的入侵则构成了信息安全潜在的威胁[2]。隐写分析旨在检测数字媒体中是否存在隐藏信息，图像隐写分析则是要检测出图像中通过隐写术轻微修改过的痕迹，往往这些痕迹当中隐藏着秘密信息。如今图像隐写分析是保障网络空间安全的关键技术之一。

现如今，随着深度学习技术的引入彻底改变了隐写分析的研究方向，使得端到端的自适应特征学习模型成为主流方向。2014 年，Tan 等人[3]首次尝试将卷积神经网络应用于隐写分析提出了 TanNet 图像隐写分析网络，该网络采用了三个卷积层和两个全连接层的设计在图像隐写检测中取得了很好的效果。2015 年，Qian 等人提出了 GNCNN (Gaussian-Neuron Convolutional Neural Network) [4]，GNCNN 进一步丰富了网络结构并通过引入高斯激活函数增强对隐写噪声的敏感性从而加快了网络的收敛速度。2016 年，Xu 等人[5]提出了经典的 XuNet 架构，该网络在输入层引入高通滤波器(HPF)作为预处理层，有效抑制了图像内容对隐写信号的干扰，通过绝对值激活函数(ABS)保留隐写噪声的幅度信息。实验表明，XuNet 在检测主流隐写算法上取得了显著效果。2017 年，Ye 等人[6]提出了 YeNet，将 SRM 滤波器组嵌入网络的

预处理层，结合截断线性单元(TLU)和批归一化技术，进一步提升了检测精度。2018 年，Yedroudj 等人提出了 YedroudjNet 网络[7]，该网络融合了 XuNet 和 YeNet 的设计优势，构建了更加高效的隐写分析架构。YedroudjNet 由预处理模块、五个卷积模块和全连接模块组成，在预处理阶段采用 SRM 的 30 个高通滤波器对输入图像进行滤波以提取噪声残差，同时继承了 XuNet 中的绝对值激活层和批归一化层的设计理念。通过这种架构融合，YedroudjNet 在多个基准数据集上的检测性能均超越了 XuNet 和 YeNet，展现出良好的特征提取和判别能力。同年，Boroumand 等人提出了 SRNet (Steganalysis Residual Network) [8]，SRNet 是首次采用了深度残差卷积神经网络的经典隐写分析模型，首次将传统隐写分析的核心思想与深度学习成功结合，其性能也超越了 YeNet。SRNet 这一设计模式为后续隐写分析网络设计提供了宝贵的经验。

然而，这些隐写分析模型的一个基本假设是模型训练和测试所用的载体图像都来源于同一个数据集。在实际应用中，这一假设往往不成立。例如，一个在标准隐写分析数据集上训练的隐写分析模型，当用于检测有着不同空间统计特性的数据集时，其性能会发生急剧衰退[9]。这种由于载体图像的空间统计特性不一致导致的性能下降现象，被称为载体源失配[10] (Cover-Source Mismatch, CSM)。载体源失配(CSM)时，隐写分析模型的检测准确率会显著下降。是因为分类器容易过拟合训练数据的分布特性，当测试数据的分布特性与训练数据的分布特性不同时，模型的检测能力会受到严重影响。

2. 相关工作

时至今日，基于深度学习的隐写分析方法虽然在同分布的数据上表现优异，但在面对载体源失配时往往表现出较差的泛化能力。目前大多数算法都采用基于深度学习领域的无监督域适应(UDA) [11]方法来解决 CSM 问题，这是因为载体源失配(CSM)和无监督域适配(UDA)之间存在相似的应用场景。CSM 和 UDA 都面临同一个本质问题：当训练数据和测试数据来自不同的分布，导致模型在测试时性能下降，其次两者都需要在没有目标域标签的情况下，使模型适应新的数据分布。针对这一问题，研究者们提出了多种基于 UDA 的解决方案：张等人提出了一种无监督域自适应的方法 J-Net [12]，通过最小化源域和目标域之间的联合最大均值差异(Joint Maximum Mean Discrepancy, JMMD)来执行域对齐、于等人进行了重要改进提出的 RCDD [13]，用可靠隐写标注机制来替代传统 UDA 方法中不可靠的聚类伪标签生成方式，并且通过类感知的域对齐策略将源域和目标域中相同类别(载体或隐写)的样本在特征空间中拉近，将不同类别的样本在特征空间中推远，实现了更精细的类级别对齐，显著提升了跨域检测性能。这类方法逐步深化了对隐写分析特性的理解，在保持检测准确性的同时有效增强了模型的跨域泛化能力，为解决载体源失配问题提供了可行的技术路径。

近几年为了解决 CSM 的问题，研究者提出了多种方法，其中中间域生成(Intermediate Domain Generation)策略，通过在训练集和测试集之间构建过渡域，能够有效缓解域偏移问题。通过构建中间域解决 CSM 问题的典型代表是 ISNet [14]和 GDNet [15]。具体而言，ISNet 通过局部特征级混合相关的补丁技术(LFMP)和域因子，生成多样化的中间域，从而在源域和目标域之间构建桥梁；GDNet 则通过骨干网络提取的深层特征确定重要性高的判别性区域，并通过像素级混合源域和目标域图像来构建判别性中间域，实现域间的有效适应。ISNet 和 GDNet 的成功表明，构建中间域是缓解 CSM 问题的有效途径。

基于这一观察，本文提出了一种频域感知自适应中间域生成方法(Frequency-Aware Adaptive Domain Generation, FADG)。该方法从频域的视角分析图像特性，利用光谱残差(Spectral Residual)检测图像中能量波动显著的区域。这些能量波动往往与隐写嵌入引起的统计变化密切相关，因此可作为隐写敏感区域的指示器。在此基础上，FADG 采用自适应混合策略，对隐写敏感区域赋予更高的权重，生成更具判别力的中间域样本，从而有效提升模型在载体源失配场景下的检测性能。本文的主要贡献包括：

(1) 提出了基于频域光谱残差的隐写敏感区域检测方法，通过对图像进行傅里叶变换，计算频域光谱残差，能够有效识别能量波动剧烈的区域。这些区域通常对应着纹理复杂、边缘丰富的图像内容，正是隐写嵌入引起统计扰动最显著的地方从而为中间域生成提供了新的指导信息。

(2) 设计了自适应混合策略，根据隐写敏感度动态计算每个像素位置的混合权重，对隐写敏感区域(如纹理复杂区域)赋予更高的混合权重。这种差异化策略使得生成的中间域样本在保留判别性特征的同时，更有效地弥合了域间差距。

3. 提出方法

如图 1 所示是 FADG 网络的整体架构。FADG 是专门为解决隐写分析中由于训练数据集和测试数据集的图像由于空间特性分布差异所造成的检测性能严重下降的问题，这也就是载体源失配的问题。其核心在于从频域分析的本质机制出发，通过光谱残差理论识别图像中频域能量波动大的区域，进而计算源域和目标域图像的频域隐写嵌入概率图，接着利用源域和目标域的嵌入概率图得到混合权重图用于生成中间域。最终通过这种频域感知自适应生成中间域的方法构建从源域到目标域的“桥梁”。下面将详细介绍其方法。

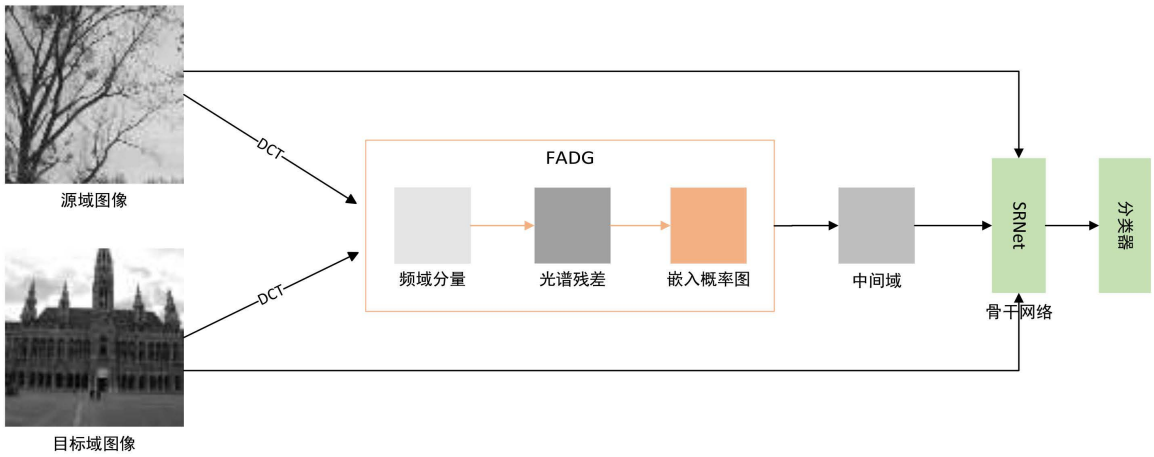

Figure 1. The overall architecture of the FADG network
图 1. FADG 网络的整体架构

3.1. 基于光谱残差计算隐写嵌入概率

现代隐写算法通常会选择纹理复杂和图像边缘等肉眼难以观察的区域进行隐写信号的嵌入。然而频域可以代表图像的能量分布，从图像的频域上分析这些隐写信号嵌入的敏感区域，可以发现这些区域的能量都是集中在中高频。所以量化出源域和目标域这种频域特性的差异尤为关键。如图 2 所示，通过光谱残差计算频域内位置适合嵌入隐写信息的概率。首先通过二维傅里叶变换也就是余弦变换，将一个高为 H 宽为 W 的图像 $I(x,y)$ 变换到频域 $F(u,v)$ 如公式(1)。得到的频率分量 $F(u,v)$ 包含幅度的能量和相位的结构信息，紧接着提取 $F(u,v)$ 的幅度频谱 $A(u,v)$ 如公式(2)，其中 Re 代表实部 Im 代表虚部，这一步主要是为了提取能量信息，也就是频率的强度，通常低频能量大，高频能量小。因为低频数值远高于高频数值，所以后续残差计算就会被低频主导，这样就无法有效分析高频信息。因此在这里我们根据频谱 $A(u,v)$ 计算出对数频谱 $L(u,v)=\log\left[A(u,v)+\varepsilon\right]$ 这里 ε 是一个常数。为了计算一个异常检测的基准，所以这里我们通过对对数频谱进行平均池化得到期望频谱 $\overline{L}(u,v)$ 如公式(3)，表示某个频率位置周围邻域的平均值，这里邻域是 3×3 区域内。

$$F(u,v)=\sum_{x=0}^{H-1}\sum_{y=0}^{W-1}I(x,y)\cdot\left[\cos\left(2\pi\left(\frac{ux}{H}+\frac{vy}{W}\right)\right)-i\sin\left(2\pi\left(\frac{ux}{H}+\frac{vy}{W}\right)\right)\right] \tag{1}$$

$$A(u,v)=\left|F(u,v)\right|=\sqrt{\left[\mathrm{Re}(F)\right]^2+\left[\mathrm{Im}(F)\right]^2} \tag{2}$$

$$\bar{L}(u,v)=\frac{1}{9}\sum_{i=-1}^{1}\sum_{j=-1}^{1}L(u+i,v+j) \tag{3}$$

根据对数频谱和期望频谱我们就可以计算出光谱的残差 $R(u,v)$ 如公式(4)，这也是 FADG 中最重要的一步。光谱残差 $R(u,v)$ 就代表该频域的能量异常程度。这里异常能量高的区域也就是隐写算法嵌入代价最小的地方，从而就能代表隐写信号嵌入概率高的区域。随后通过逆傅里叶变换返回得到空域光谱残差 $R(x,y)$，根据空域光谱残差得到嵌入概率图 $P(x,y)$ 如公式(5)，$P(x,y)$ 是归一化后的归一化后的概率表示，P_{max} 为最大概率、P_{mix} 为最小概率。

$$R(u,v)=\left|L(u,v)-\bar{L}(u,v)\right| \tag{4}$$

$$P(x,y)=\frac{\left(1-\mathrm{e}^{-R(x,y)}\right)-P_{min}}{P_{max}-P_{min}} \tag{5}$$

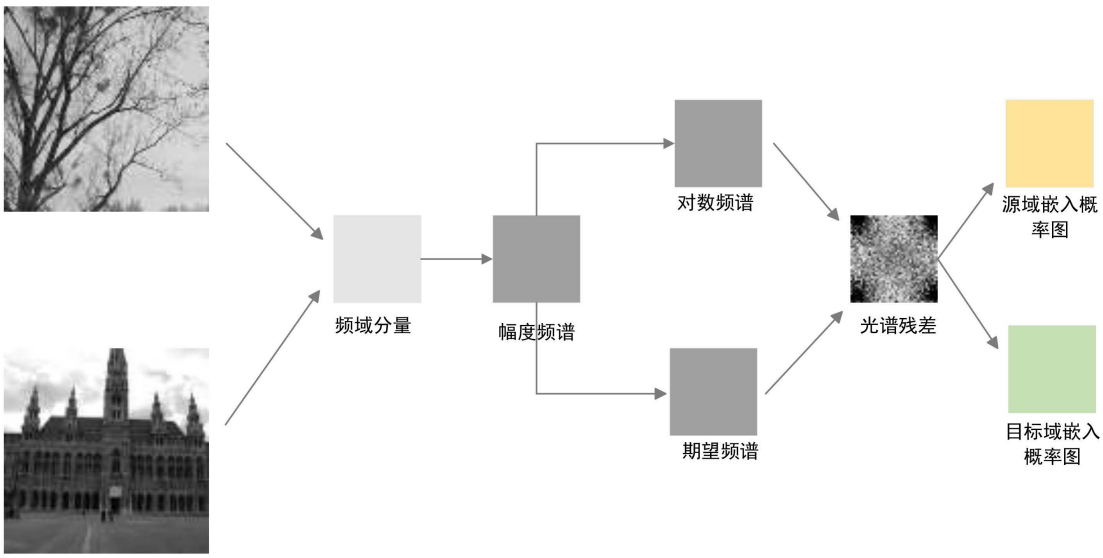

Figure 2. Probability graph for steganographic embedding calculation based on spectral residuals
图 2. 基于光谱残差计算隐写嵌入概率图

3.2. 生成自适应的中间域

根据以上方法我们可以得到源域嵌入概率图 $P_s(x,y)$ 和目标域嵌入概率图 $P_t(x,y)$。根据得到的嵌入概率图计算出混合权重图 M，具体计算方式如公式(6)所示，我们将目标域和源域的嵌入概率图的差值也作为混合权重比例的参考，并且以 0.3 的权重比例和目标域的嵌入概率图相加得到最终的混合权重图 M，为了混合生成中间域的有效性将混合权重的值进行了限制 $M(i,j)\in[0.1,0.8]$。最后利用混合权重图生成中间域 I_{MIX} 如公式(7)，其中 I_{TARGET} 为目标域图像，I_{SOURCE} 为源域图像。

$$M=P_t(x,y)+0.3\cdot\left[P_t(x,y)-P_s(x,y)\right] \tag{6}$$

$$I_{MIX} = M \cdot I_{TARGET} + (1-M) \cdot I_{SOURCE} \tag{7}$$

3.3. 损失函数设计

对于混合生成的中间域样本的损失 L_{mix} 如公式(8)所示，其中中间域样本的预测值是 y_{mix}， y^s 表示源域的真实标签， y^t 表示目标域的伪标签。

$$L_{mix} = \alpha \cdot L_{cls}\left(y_{mix}, y^t\right) + (1-\alpha) \cdot L_{cls}\left(y_{mix}, y^s\right) \tag{8}$$

最终总损失 L_{total} 设计为源域损失 $L_s = L_{cls}\left(y_s, y^s\right)$ 加上一定比例的中间域损失 L_{mix}，表示为公式(9)

$$L_{total} = L_s + 0.8L_{mix} \tag{9}$$

4. 实验结果与分析

4.1. 实验设置

实验使用了三个公开数据集：Alaska (A) [16]包含 80,000 张各种格式的图像(实验中选择 256 × 256 像素的 PGM 格式灰度图像 10,000 张)，BOSSBase1.01 (B) [17]包含 10,000 张 PGM 格式灰度图像，MIRFlickr 25k (M) [18]包含 25,000 张不同尺寸的彩色 JPEG 图像。

在空间域实验中，从每个数据集随机选择 10,000 张图像，共 30,000 张图像通过 MATLAB 的双线性插值重采样至 256 × 256 像素。使用四种内容自适应隐写方法 S-UNIWARD [19]、HILL [20]、MIPOD [21]和 WOW [22]在指定负载下分别生成 10,000 对图像数据，本实验是以 0.4 bpp 和 0.2 bpp 负载为例，每种隐写方法对应三个数据集，分别是 Alaska (A)、BOSSBase1.01 (B)、MIRFlickr 25k (M)，这样就一共会产生 6 种载体源适配(CSM)的场景。

在 JPEG 域中，BOSSBase1.01 的 10,000 张图像首先用质量因子 QF [23]为 75、85、95 进行压缩生成三个不同的压缩数据集，每个包含 10,000 张载体图像。在 0.4 bpnc 负载下使用 J-UNIWARD [19]和 UERD [24]生成 6 组载体/隐写图像对。

实验中使用 SRNet 作为 FADG 的骨干网络。在预训练阶段，每 10,000 对载体和隐写图像按 8:2 的比例划分为训练集和测试集，训练参数与 SRNet 原文献相同。在域适应训练阶段，特征提取部分参数用预训练的 SRNet 权重初始化，从训练集和测试集中随机选择 500 对图像分别代表源域和目标域。源域和目标域的批量大小均设为 32，学习率设为 0.0005。整个训练过程共进行 100 个 epoch。所有实验在配备 NVIDIA 3090 的 Pytorch 2.4.0 环境中实现。

4.2. 空域实验结果

对于空间域，表 1 展示了每个载体源失配(CSM)场景下四种隐写分析算法在 0.4 bpp 负载下的准确率。对于 S-UNIWARD、WOW、HILL 和 MIPOD 隐写算法，FADG 的整体性能有所提高。仅在少数情况下略低于 RCDD。此外，与 J-Net、RCDD 相比，FADG 在 0.4 bpp 负载下的平均性能分别取得了约 6%~0.5%的提升。

Table 1. Accuracy rates of six steganalysis algorithms in each CSM scenario at a 0.4 bpp load
表 1. 每个 CSM 场景下六种隐写分析算法在 0.4 bpp 负载时的准确率

	Method	A→B	A→M	B→A	B→M	M→A	M→B	AVG
SUN	Backbone	75.60	68.00	55.10	53.50	59.50	80.90	65.43
	J-Net	82.40	70.00	54.60	55.90	59.93	85.00	67.97

续表

	RCDD	76.40	70.00	64.00	70.40	64.70	79.90	70.90
	FADG	83.20	69.80	64.00	70.90	64.00	86.60	73.08
HIL	Backbone	83.78	64.90	55.10	64.30	65.10	73.50	68.68
	J-Net	84.00	70.30	61.60	58.70	65.10	85.10	70.80
	RCDD	83.00	72.10	64.80	73.50	63.40	82.50	73.22
	FADG	84.20	72.20	66.00	72.40	67.80	87.20	74.97
MIP	Backbone	81.30	65.50	63.90	62.88	63.23	74.48	68.55
	J-Net	82.70	71.10	65.70	67.90	64.80	82.30	71.43
	RCDD	81.40	73.20	67.80	72.50	64.80	72.10	71.97
	FADG	82.60	71.80	67.40	70.30	64.60	86.60	73.88
WOW	Backbone	84.33	68.50	54.65	51.58	59.43	86.90	67.57
	J-Net	87.30	71.60	58.60	63.20	64.80	86.00	71.92
	RCDD	76.20	69.40	64.00	69.10	62.80	73.40	69.15
	FADG	84.80	72.00	64.30	74.00	65.40	89.00	74.92

为了验证 FADG 在低负载下仍然有效，表 2 对比了四种隐写分析算法在使用 S-UNIWARD 算法 0.2 bpp 负载时的实验结果。FADG 相比 J-Net 提高了约 8%，FADG 相比 RCDD 提高了约 2%。

Table 2. Accuracy rates of four steganalysis algorithms at a 0.2 bpp load of the S-UNIWARD algorithm
表 2. 四种隐写分析算法在 S-UNIWARD 算法 0.2 bpp 负载时的准确率

	Method	A→B	A→M	B→A	B→M	M→A	M→B	AVG
SUN	Backbone	70.40	56.78	52.10	51.43	54.05	68.38	58.86
	J-Net	73.20	59.30	53.10	53.30	54.60	64.50	59.67
	RCDD	73.20	60.20	59.50	60.10	58.30	72.80	64.02
	FADG	74.00	61.20	59.00	62.80	60.00	78.00	65.83

4.3. JPEG 域实验结果

为了验证 FADG 在 JPEG 域的优越性，我们在 JPEG 域的 CSM 场景将其与 Stega-SL [25]进行比较，其中 Stega-SL 是专门为 JPEG 域中的 CSM 场景设计的。实验中 JPEG 压缩质量因子 QF 分别有 75、85、95。表 3 展示了 J-UNIWARD 和 UERD 在 0.4 bpnc 负载下的比较结果。具体来说，与经典的 Stega-SL 方法相比，FADG 在 J-UNIWARD 压缩算法中平均性能提升了约 16%，在 UERD 压缩算法中平均性能提升了约 13%。

Table 3. Accuracy rate of JPEG domain (J-UNIWARD and UERD) at a 0.4 bpnc load
表 3. JPEG 域(J-UNIWARD 和 UERD)在 0.4 bpnc 负载时的准确率

	Method	75→85	75→95	85→75	85→95	95→75	95→75	AVG
JUN	Backbone	50.10	50.10	50.10	50.80	50.00	50.00	50.18
	Stega-SL	51.80	50.90	51.60	52.70	52.40	52.40	51.97
	FADG	80.10	60.00	84.50	63.70	62.20	72.00	70.42
UER	Backbone	50.85	50.00	50.03	50.03	50.00	50.00	50.15
	Stega-SL	59.70	50.80	51.80	51.60	52.00	53.40	53.22
	FADG	84.40	56.60	79.80	63.00	56.80	63.30	67.32

4.4. 消融实验

为验证 FADG 中基于光谱残差的频域感的有效性，表 4 对比了在 J-UNIWARD 隐写算法、0.4 bpnc 嵌入率条件下，FADG 与采用随机生成掩码策略(w/o)的检测准确率。实验结果表明在不采用 FADG 的方法时性能下降了约 7%。

Table 4. Comparison of the accuracy rates of FADG and random mask generation strategy under 0.4 bpnc load for J-UNIWARD
表 4. J-UNIWARD 在 0.4 bpnc 负载下 FADG 与随机生成掩码策略的准确率比较

	Method	75→85	75→95	85→75	85→95	95→75	95→75	AVG
JUN	w/o	72.60	54.20	74.80	56.40	54.30	68.10	63.40
	FADG	80.10	60.00	84.50	63.70	62.20	72.00	70.42

5. 结论

本文针对隐写分析中的载体源失配问题，提出了一种频域感知的自适应中间域生成方法(FADG)。该方法的核心创新在于利用频域光谱残差分析图像的能量分布特性，识别隐写敏感区域，并基于此设计自适应混合策略，生成更具判别力的中间域样本。该方法的核心创新在于将频域分析理论与中间域生成策略相结合，通过构建从源域到目标域的“桥梁”，有效缓解了载体源失配的问题。实验结果表明，FADG 方法在检测 S-UNIWARD、WOW、HILL 和 MIPOD 等隐写算法时均取得了提升，验证了频域感知的自适应混合策略的有效性，为跨域隐写分析提供了新的研究思路。

参考文献

[1] Zhong, N., Qian, Z., Wang, Z., Zhang, X. and Li, X. (2021) Batch Steganography via Generative Network. *IEEE Transactions on Circuits and Systems for Video Technology*, **31**, 88-97. https://doi.org/10.1109/tcsvt.2020.2974884

[2] Verma, V., Muttoo, S.K. and Singh, V.B. (2022) Detecting Stegomalware: Malicious Image Steganography and Its Intrusion in Windows. In: Rao, U.P., Patel, S.J., Raj, P. and Visconti, A., Eds., *Security, Privacy and Data Analytics*, Springer, 103-116. https://doi.org/10.1007/978-981-16-9089-1_9

[3] Tan, S. and Li, B. (2014) Stacked Convolutional Auto-Encoders for Steganalysis of Digital Images. *Signal and Information Processing Association Annual Summit and Conference* (*APSIPA*), 2014 *Asia-Pacific*, Siem Reap, 9-12 December 2014, 1-4. https://doi.org/10.1109/apsipa.2014.7041565

[4] Qian, Y., Dong, J., Wang, W. and Tan, T. (2015) Deep Learning for Steganalysis via Convolutional Neural Networks. *SPIE Proceedings*, **9409**, Article ID: 94090J. https://doi.org/10.1117/12.2083479

[5] Xu, G., Wu, H. and Shi, Y. (2016) Structural Design of Convolutional Neural Networks for Steganalysis. *IEEE Signal Processing Letters*, **23**, 708-712. https://doi.org/10.1109/lsp.2016.2548421

[6] Ye, J., Ni, J. and Yi, Y. (2017) Deep Learning Hierarchical Representations for Image Steganalysis. *IEEE Transactions on Information Forensics and Security*, **12**, 2545-2557. https://doi.org/10.1109/tifs.2017.2710946

[7] Yedroudj, M., Comby, F. and Chaumont, M. (2018) Yedroudj-Net: An Efficient CNN for Spatial Steganalysis. 2018 *IEEE International Conference on Acoustics, Speech and Signal Processing* (*ICASSP*), Calgary, 15-20 April 2018, 2092-2096. https://doi.org/10.1109/icassp.2018.8461438

[8] Boroumand, M., Chen, M. and Fridrich, J. (2019) Deep Residual Network for Steganalysis of Digital Images. *IEEE Transactions on Information Forensics and Security*, **14**, 1181-1193. https://doi.org/10.1109/tifs.2018.2871749

[9] Cancelli, G., Doerr, G., Barni, M. and Cox, I.J. (2008) A Comparative Study of ± Steganalyzers. 2008 *IEEE* 10*th Workshop on Multimedia Signal Processing*, Cairns, 8-10 October 2008, 791-796. https://doi.org/10.1109/mmsp.2008.4665182

[10] Kodovský, J., Sedighi, V. and Fridrich, J. (2014) Study of Cover Source Mismatch in Steganalysis and Ways to Mitigate Its Impact. *SPIE Proceedings*, **9028**, Article ID: 90280J. https://doi.org/10.1117/12.2039693

[11] Pan, S.J. and Yang, Q. (2010) A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering*,

22, 1345-1359. https://doi.org/10.1109/tkde.2009.191

[12] Zhang, X., Kong, X., Wang, P. and Wang, B. (2020) Cover-Source Mismatch in Deep Spatial Steganalysis. In: Wang, H., Zhao, X., Shi, Y., Kim, H. and Piva, A., Eds., *Digital Forensics and Watermarking*, Springer, 71-83. https://doi.org/10.1007/978-3-030-43575-2_6

[13] Yu, L., Weng, S., Chen, M. and Wei, Y. (2024) RCDD: Contrastive Domain Discrepancy with Reliable Steganalysis Labeling for Cover Source Mismatch. *Expert Systems with Applications*, **237**, Article ID: 121543. https://doi.org/10.1016/j.eswa.2023.121543

[14] Weng, S., Zhang, Z., Yu, L., Cao, P. and Cao, G. (2024) Universal Mismatched Steganalysis Equipped with Progressive Intermediate Domains. *IEEE Signal Processing Letters*, **31**, 800-804. https://doi.org/10.1109/lsp.2024.3374601

[15] Li, Y., Yu, L., Weng, S., Tian, H. and Cao, G. (2023) Discriminability-Aware Intermediate Domains for Mismatched Steganalysis. *Journal of LaTeX Class Files*, **14**, 1-8.

[16] Cogranne, R., Giboulot, Q. and Bas, P. (2019) The ALASKA Steganalysis Challenge: A First Step towards Steganalysis. *Proceedings of the ACM Workshop on Information Hiding and Multimedia Security*, Paris, 3-5 July 2019, 125-137. https://doi.org/10.1145/3335203.3335726

[17] Bas, P., Filler, T. and Pevný, T. (2011) "Break Our Steganographic System": The Ins and Outs of Organizing Boss. In: Filler, T., Pevný, T., Craver, S. and Ker, A., Eds., *Information Hiding*, Springer, 59-70. https://doi.org/10.1007/978-3-642-24178-9_5

[18] Huiskes, M.J. and Lew, M.S. (2008) The MIR Flickr Retrieval Evaluation. *Proceedings of the* 1*st ACM International Conference on Multimedia Information Retrieval*, Vancouver, 30-31 October 2008, 39-43. https://doi.org/10.1145/1460096.1460104

[19] Holub, V., Fridrich, J. and Denemark, T. (2014) Universal Distortion Function for Steganography in an Arbitrary Domain. *EURASIP Journal on Information Security*, **2014**, Article No. 1. https://doi.org/10.1186/1687-417x-2014-1

[20] Li, B., Wang, M., Huang, J. and Li, X. (2014) A New Cost Function for Spatial Image Steganography. 2014 *IEEE International Conference on Image Processing* (*ICIP*), Paris, 27-30 October 2014, 4206-4210. https://doi.org/10.1109/icip.2014.7025854

[21] Sedighi, V., Cogranne, R. and Fridrich, J. (2016) Content-adaptive Steganography by Minimizing Statistical Detectability. *IEEE Transactions on Information Forensics and Security*, **11**, 221-234. https://doi.org/10.1109/tifs.2015.2486744

[22] Holub, V. and Fridrich, J. (2012) Designing Steganographic Distortion Using Directional Filters. 2012 *IEEE International Workshop on Information Forensics and Security* (*WIFS*), Costa Adeje, 2-5 December 2012, 234-239. https://doi.org/10.1109/wifs.2012.6412655

[23] Golner, M.A., Mikhael, W.B., Krishnan, V. and Ramaswamy, A. (2000) Region Based Variable Quantization for JPEG Image Compression. *Proceedings of the* 43*rd IEEE Midwest Symposium on Circuits and Systems* (*Cat.No.CH*37144), Lansing, 8-11 August 2000, 604-607. https://doi.org/10.1109/mwscas.2000.952829

[24] Guo, L., Ni, J., Su, W., Tang, C. and Shi, Y. (2015) Using Statistical Image Model for JPEG Steganography: Uniform Embedding Revisited. *IEEE Transactions on Information Forensics and Security*, **10**, 2669-2680. https://doi.org/10.1109/tifs.2015.2473815

[25] Xue, Y., Yang, L., Wen, J., Niu, S. and Zhong, P. (2018) A Subspace Learning-Based Method for JPEG Mismatched Steganalysis. *Multimedia Tools and Applications*, **78**, 8151-8166. https://doi.org/10.1007/s11042-018-6719-5