KMANPUB | Decision Science and Intelligent Systems Vol. 3 No. 2 (2026): 1-37

Article history:
Received 07 June 2026
Revised 14 August 2026
Accepted 17 August 2026
Published online 23 August 2026

Robust Image Watermarking Against Geometric Attacks Using Surrogate Model–Based Scaling Factor Optimization and Deep Learning–Based Adaptive Embedding

Masoud. Pourmoradi[1], Hamidreza. Ghaffari[1*]

[1] Department of Computer Engineering, Fe.C., Islamic Azad University, Ferdows, Iran

*** Corresponding author email address**: hamidghaffary53@iau.ac.ir

A r t i c l e I n f o

Article type:
Original Research

How to cite this article:
Pourmoradi, M. & Ghaffari, H. (2026). Robust Image Watermarking Against Geometric Attacks Using Surrogate Model–Based Scaling Factor Optimization and Deep Learning–Based Adaptive Embedding. *Decision Science and Intelligent Systems. 3(2), 1-37.*

A B S T R A C T

Digital watermarking, as a fundamental approach to copyright protection, has consistently faced the inherent challenge of balancing imperceptibility and robustness, particularly against geometric attacks. This delicate trade-off is strongly influenced by two key factors: the optimal selection of the scaling factor (embedding strength) and the intelligent selection of embedding regions. This article proposes a robust hybrid watermarking framework for color images that strategically integrates intelligent image analysis with accelerated optimization and is specifically designed to withstand rotation attacks. The proposed method incorporates two principal innovations. First, a deep convolutional neural network (CNN) with a U-Net architecture is trained to generate an adaptive weight map. Independently of the embedding process, this network analyzes the content of the host image and automatically identifies regions with high visual complexity, including edges and textures. Second, a particle swarm optimization (PSO) algorithm guided by a regression tree–based surrogate model is employed to determine the optimal scaling factor. By replacing computationally expensive evaluations of the actual objective function with a rapid estimator, this approach substantially accelerates the optimization process. Finally, fully adaptive embedding is performed through the pointwise integration of the optimal scaling factor and the local weight map within the approximation subband of the discrete wavelet transform (DWT). Experimental results obtained using standard benchmark images confirm the strong performance of the proposed method. By achieving an average peak signal-to-noise ratio (PSNR) greater than 38 dB and a normalized correlation (NC) exceeding 98% following the simultaneous application of rotation and JPEG compression attacks, the proposed method demonstrates a substantial improvement over previous approaches and effectively establishes a balance between imperceptibility and robustness.

Keywords: *image watermarking, geometric attacks, scaling factor optimization, surrogate model, deep learning, convolutional neural network*

Extended Abstract

Introduction

The rapid expansion of digital communication technologies has dramatically increased the production, distribution, and exchange of multimedia content across social networks, electronic commerce platforms, medical databases, news websites, and other digital environments. Although this development has facilitated efficient access to digital information, it has simultaneously intensified concerns regarding unauthorized reproduction, manipulation, redistribution, copyright infringement, and ownership verification. Digital image watermarking has consequently emerged as an important complementary security mechanism through which ownership or authentication information can be imperceptibly embedded into a host image while remaining recoverable when verification is required (Mahto et al., 2021; Zermi et al., 2021). Unlike conventional cryptographic techniques, whose protection generally terminates after content is decrypted, digital watermarking maintains ownership-related information as an integral component of the digital media itself.

The effectiveness of a digital watermarking system is fundamentally determined by the trade-off among imperceptibility, robustness, and embedding capacity. Imperceptibility requires the watermarked image to remain visually indistinguishable from the original host image, whereas robustness concerns the ability to recover the embedded watermark after intentional or unintentional image processing. Capacity refers to the amount of information that can be embedded without unacceptable deterioration in image quality (Begum & Uddin, 2020). Among these requirements, robustness against geometric attacks remains particularly challenging. Rotation, scaling, and translation attacks can disrupt synchronization between the embedding and extraction procedures without necessarily destroying the watermark itself, thereby considerably reducing extraction reliability (Degadwala et al., 2020).

Transform-domain watermarking has generally demonstrated greater robustness than spatial-domain techniques because watermark information can be distributed among perceptually and structurally significant frequency coefficients. Discrete cosine transform (DCT), discrete wavelet transform (DWT), and singular value decomposition (SVD) have therefore become prominent techniques in robust watermarking systems (Abdulrahman & Ozturk, 2019; Fares et al., 2021). DWT is particularly advantageous because its multiresolution representation simultaneously provides spatial and frequency localization, allowing watermark information to be embedded into subbands with different perceptual and robustness characteristics. However, two major problems remain unresolved. The first concerns selection of the scaling factor, or embedding strength. An excessively small scaling factor preserves image quality but produces a vulnerable watermark, whereas an excessively large factor improves robustness at the expense of perceptual quality. The second concerns the conventional use of uniform embedding strength throughout an image despite substantial local differences in visual complexity.

Recent advances in deep learning provide an opportunity to address the latter limitation. Convolutional neural networks (CNNs) can learn hierarchical visual representations and distinguish smooth regions from complex edges and textures, making them suitable for identifying image areas in which watermark energy can be concealed more effectively (Zhu et al., 2021). Although end-to-end deep watermarking systems have demonstrated promising performance, their robustness against geometric transformations frequently depends on computationally intensive attack-aware training, and their behavior may deteriorate when encountering transformations insufficiently represented during training (Wan et al.,

2022). Hybrid strategies that use deep learning as an image analyzer while retaining established transform-domain embedding mechanisms can potentially combine the adaptability of neural networks with the structural robustness of classical watermarking (Sy et al., 2020; Zhu et al., 2021).

Optimization techniques offer a complementary solution to scale-factor selection. Metaheuristic algorithms can search continuous parameter spaces more systematically than empirical trial-and-error procedures and have previously been applied to optimize watermark embedding parameters (Sharma et al., 2022). Nevertheless, direct metaheuristic optimization is computationally expensive because each candidate parameter may require complete watermark embedding, attack simulation, extraction, and performance evaluation. Surrogate modeling provides a means of reducing this burden by replacing repeated expensive objective-function evaluations with predictions from a computationally inexpensive approximation model (Bartz-Beielstein, 2021). Accordingly, the present study developed a two-phase hybrid framework integrating U-Net-based adaptive image analysis, regression tree–based surrogate modeling, particle swarm optimization (PSO), and DWT-based watermark embedding. The objective was to improve robustness against geometric attacks while maintaining high perceptual quality and substantially reducing the computational burden associated with scale-factor optimization.

Methods and Materials

The proposed framework consisted of two interconnected phases. In the first phase, a deep CNN based on the U-Net architecture was developed to generate adaptive weight maps identifying image regions suitable for watermark embedding. Six standard 512 × 512 color images—Baboon, Barbara, Boat, Jet, Lena, and Pepper—were used as principal host images, while a logo image served as the watermark. An augmented dataset containing 2,000 image patches of 128 × 128 pixels was generated through random cropping, minor rotation, and horizontal and vertical reflection. The dataset was randomly divided into 1,700 training samples (85%) and 300 validation samples (15%).

Ground-truth weight maps were constructed by combining three normalized visual-complexity descriptors: Canny edge information with thresholds of 0.05 and 0.15, local texture information calculated using a 7 × 7 local-variation window, and Sobel gradient magnitude. Their respective weights were 0.40, 0.30, and 0.30. The resulting maps were Gaussian-smoothed and normalized to [0, 1].

The U-Net contained 35 layers. The encoder used 32- and 64-filter convolutional blocks followed by max pooling, while the bottleneck contained 128-filter convolutional layers. Transposed convolutions and skip connections were used in the decoder to reconstruct spatial information, and a 1 × 1 convolution produced the final single-channel adaptive weight map. Training employed the Adam optimizer, an initial learning rate of 0.001, mean squared error (MSE) loss, L2 regularization of 0.0001, gradient clipping of 1.0, and mini-batches of 16.

In the second phase, the blue channel of each RGB host image was selected for watermark embedding. A two-level two-dimensional Haar DWT was applied, and the second-level approximation subband (LL2) served as the embedding domain. A regression-tree surrogate model was trained using 30 actual evaluations sampled across the scaling-factor search interval of 0.08–0.26 for each image. PSO subsequently searched the surrogate response surface for 50 iterations to determine an image-specific optimal scaling factor.

The optimized global scaling factor was multiplied pointwise by the CNN-generated local weight map to obtain spatially adaptive embedding strengths. The binary watermark, previously scrambled using

the Arnold transform, was embedded into LL2 coefficients. Inverse DWT reconstructed the modified blue channel, which was recombined with the unchanged red and green channels. Imperceptibility was assessed using PSNR, MSE, and structural similarity index measure (SSIM), whereas robustness was evaluated using normalized correlation (NC) and bit error rate (BER). A combined rotation of 15° and JPEG compression was used as the principal robustness test.

Findings

The U-Net demonstrated highly accurate learning of the adaptive weight maps. Mean validation accuracy reached 99.986%, with a standard deviation of 0.016%, a minimum of 99.944%, and a maximum of 99.994%. Mean MSE was approximately 1.44×10^{-4}, mean RMSE was 0.0108, and mean absolute error was 0.0078. Reconstruction quality was also high, with a mean PSNR of 40.08 dB (SD = 3.38) and mean SSIM of 0.990 (SD = 0.007). Qualitative examination showed that larger weights were concentrated around edges, textures, hair, facial details, object boundaries, and other visually complex regions, whereas smooth areas received substantially lower weights.

The surrogate-assisted PSO successfully identified different optimal scaling factors for different host images, ranging from 0.120 for Lena to 0.147 for Baboon. This variability demonstrated that a single fixed embedding strength was not optimal across images with different structural characteristics. The optimization trajectories stabilized within the 50-iteration search, while the regression-tree models reproduced the nonlinear and image-specific relationships between scaling factor and estimated cost.

Final watermarking performance demonstrated consistently high imperceptibility. PSNR values exceeded 38 dB for all six host images. Lena produced the highest value at approximately 39.45 dB and an SSIM of 0.987, whereas Pepper produced the lowest PSNR at approximately 38.22–38.31 dB while still maintaining acceptable visual quality. Barbara achieved approximately 39.3 dB with an SSIM of 0.966, and Baboon achieved approximately 39.4 dB with an SSIM of 0.976.

Robustness remained high following the combined 15° rotation and JPEG compression attack. NC exceeded 98% for every host image, ranging from approximately 98.09% for Pepper to 98.79% for Barbara. BER remained below 1%, with the lowest value of 0.48% for Barbara and the highest value of 0.76% for Pepper. Baboon achieved an NC of approximately 98.4% and a BER of 0.65%. The extracted watermark remained visually recognizable in all examined images.

The overall mean PSNR of the proposed method was 38.94 dB, while its mean NC reached 98.39%. In comparative evaluation, the NC of the proposed framework exceeded those reported for the examined alternative methods, including approximately 94% for the Zermi method, 92% for DWT-SVD, and 87% for DCT. Although some competing approaches achieved higher PSNR values, the proposed framework demonstrated a stronger preservation of recoverable watermark information under the investigated combined attack.

Discussion and Conclusion

The findings demonstrate that combining adaptive deep visual analysis with surrogate-assisted optimization can effectively address two major weaknesses of conventional transform-domain watermarking: uniform allocation of embedding strength and computationally inefficient parameter selection. The U-Net did not directly perform watermark embedding or extraction; instead, it functioned as an intelligent scene analyzer. This separation allowed the system to exploit deep learning for spatial adaptation without making the entire watermarking process dependent on an end-to-end neural architecture.

The high validation accuracy and near-unity SSIM obtained during weight-map prediction indicate that the CNN successfully learned the intended visual-complexity representation. Concentrating embedding energy in textured regions and edges reduced perceptually detectable distortion in smooth regions while permitting stronger local embedding where modifications could be effectively masked. The image-specific optimal scaling factors further demonstrate that watermark strength should not be treated as a universal constant. Structural differences among images alter their tolerance for embedded information and consequently require adaptive parameterization.

The surrogate model also provided an important computational contribution. Instead of repeatedly executing the complete watermarking, attack, and extraction pipeline during PSO, optimization was conducted primarily on an inexpensive regression-tree approximation of the objective function. This architecture makes extensive exploration of the scaling-factor search space considerably more practical while retaining the ability to model nonlinear cost behavior.

Most importantly, the framework maintained NC values above 98% and BER below 1% under the simultaneous application of rotation and JPEG compression while preserving PSNR above 38 dB. These findings indicate that the method established an effective balance between perceptual transparency and attack robustness. Its principal contribution therefore lies not in any single component, but in the coordinated integration of CNN-based spatial intelligence, surrogate-assisted PSO parameter optimization, and robust DWT-domain embedding.

In conclusion, the proposed framework provides an effective hybrid solution for robust color-image watermarking under geometric and compression attacks. The results support adaptive rather than uniform embedding and demonstrate the practical value of surrogate modeling for computationally expensive watermark optimization. Future development should extend robustness testing to scaling, translation, cropping, row or column removal, and additional signal-processing attacks. Further improvements could include Gaussian-process surrogate models, explicit multi-objective optimization, attack-aware CNN training, and development of a fully blind extraction mechanism that eliminates dependence on the original host image.

علم تصمیم گیری و سیستم های هوشمند دوره ۳، شماره ۲ (۱۴۰۵): ۳۷-۱

تاریخچه مقاله
دریافت شده در تاریخ ۱۷ خرداد ۱۴۰۵
اصلاح شده در تاریخ ۲۳ مرداد ۱۴۰۵
پذیرفته شده در تاریخ ۲۶ مرداد ۱۴۰۵
منتشر شده در تاریخ ۱ شهریور ۱۴۰۵

نهان‌نگاری مقاوم تصویر در برابر حملات هندسی با استفاده از بهینه‌سازی فاکتور مقیاس مبتنی بر مدل جانشین و جاسازی تطبیقی مبتنی بر یادگیری عمیق

مسعود پورمرادی[۱]، حمیدرضا غفاری[۱*]

۱. گروه مهندسی کامپیوتر، واحد فردوس، دانشگاه آزاد اسلامی، فردوس، ایران

*ایمیل نویسنده مسئول: hamidghaffary53@iau.ac.ir

اطلاعات مقاله

نوع مقاله
پژوهشی اصیل

نحوه استناد به این مقاله:
پورمرادی، مسعود، و غفاری، حمیدرضا. (۱۴۰۵). نهان‌نگاری مقاوم تصویر در برابر حملات هندسی با استفاده از بهینه‌سازی فاکتور مقیاس مبتنی بر مدل جانشین و جاسازی تطبیقی مبتنی بر یادگیری عمیق. *علم تصمیم گیری و سیستم های هوشمند*، *۳*(۲)، ۳۷-۱.

چکیده

نهان‌نگاری دیجیتال به‌عنوان راهکاری بنیادین برای حفاظت از حق نشر، همواره با چالش ذاتی برقراری تعادل میان شفافیت و مقاومت، به‌ویژه در برابر حملات هندسی، روبرو است. این تعادل حساس به‌شدت تحت تأثیر دو عامل کلیدی قرار دارد: انتخاب بهینه‌ی فاکتور مقیاس (قدرت جاسازی) و انتخاب هوشمندانه‌ی نواحی جاسازی. این مقاله یک چارچوب نهان‌نگاری ترکیبی و مقاوم برای تصاویر رنگی ارائه می‌دهد که با تلفیق راهبردی تحلیل هوشمند تصویر و بهینه‌سازی تسریع‌شده، به طور خاص برای مقابله با حملات چرخش طراحی شده است. روش پیشنهادی دو نوآوری اصلی را در خود ادغام می‌کند. نخست، یک شبکه‌ی عصبی کانولوشنی عمیق با معماری U-Net برای تولید یک نقشه‌ی وزن تطبیقی آموزش داده می‌شود. این شبکه، مستقل از فرآیند جاسازی، محتوای تصویر میزبان را تحلیل کرده و به‌طور خودکار نواحی با پیچیدگی بصری بالا (لبه‌ها و بافت‌ها) را شناسایی می‌کند. دوم، یک الگوریتم بهینه‌سازی ازدحام ذرات (PSO) که توسط یک مدل جانشین مبتنی بر درخت رگرسیون هدایت می‌شود، برای یافتن مقدار بهینه‌ی فاکتور مقیاس به‌کار می‌رود. این رویکرد با جایگزینی ارزیابی‌های پرهزینه‌ی تابع واقعی با یک تخمین‌گر سریع، شتاب چشمگیری به فرآیند بهینه‌سازی می‌بخشد. در نهایت، یک جاسازی کاملاً تطبیقی با تلفیق نقطه‌ای فاکتور مقیاس بهینه و نقشه‌ی وزن محلی در زیرباند تقریبی تبدیل موجک گسسته (DWT) انجام می‌شود. نتایج تجربی بر روی تصاویر استاندارد، کارایی برجسته‌ی روش را تأیید می‌کنند. این روش با دستیابی به میانگین PSNR بالای ۳۸ دسی‌بل و ضریب همبستگی نرمال‌شده (NC) بالای ۹۸٪ پس از اعمال هم‌زمان حمله‌ی چرخش و فشرده‌سازی JPEG ، پیشرفت قابل‌توجهی را نسبت به روش‌های پیشین به نمایش گذاشته و تعادل میان شفافیت و مقاومت را به نحوی مؤثر برقرار می‌سازد.

کلیدواژگان: *نهان‌نگاری تصویر، حملات هندسی، بهینه‌سازی فاکتور مقیاس، مدل جانشین، یادگیری عمیق، شبکه عصبی کانولوشنی*

۱. مقدمه

در عصر ارتباطات دیجیتال، تولید، توزیع و مصرف محتوای چندرسانه‌ای به‌طرز بی‌سابقه‌ای افزایش یافته است. تصاویر دیجیتال به عنوان یکی از مهم‌ترین ارکان این محتوا، در بسترهای گوناگونی چون شبکه‌های اجتماعی، وب‌سایت‌های خبری، پایگاه‌های داده پزشکی و سیستم‌های تجارت الکترونیک به اشتراک گذاشته می‌شوند[۱]. با این حال، همین سهولت در تبادل، چالش‌های امنیتی جدی را به همراه دارد. امکان کپی‌برداری غیرمجاز، دستکاری محتوا[1] و توزیع بدون ذکر منبع، ناقضان حق نشر[2] و اعتبارسنجی محتوا را به تهدیداتی همیشگی تبدیل کرده است[2] . در این میان، روش‌های سنتی رمزنگاری [3] اگرچه می‌توانند از محتوای دیجیتال در حین انتقال محافظت کنند، اما پس از رمزگشایی، هیچ سازوکاری برای جلوگیری از کپی و توزیع غیرمجاز یا اثبات مالکیت ارائه نمی‌دهند. نهان‌نگاری دیجیتال[4] به عنوان یک فناوری مکمل و قدرتمند، با جاسازی یک نشانه دیجیتال [5] به صورت نامحسوس درون رسانه میزبان[6] ، راه‌حلی پایدار برای احراز هویت، اثبات مالکیت و ردیابی توزیع غیرقانونی فراهم می‌آورد[3] .

یک سیستم نهان‌نگاری ایده‌آل، حاصل یک مصالحه [7] پیچیده میان سه ویژگی بنیادین و اغلب متعارض است: شفافیت [8] ، مقاومت [9]و ظرفیت[۱۰] [۴]. شفافیت ایجاب می‌کند که جاسازی واترمارک نباید هیچ‌گونه افت کیفیت بصری محسوسی در تصویر میزبان ایجاد کند. مقاومت به معنای توانایی الگوریتم در بازیابی موفق واترمارک حتی پس از آنکه تصویر میزبان مورد حملات عمدی یا غیرعمدی پردازش سیگنال قرار گرفته باشد، است. ظرفیت نیز به میزان اطلاعاتی که می‌توان در یک رسانه جاسازی کرد، اشاره دارد. در این میان، مقاومت در برابر حملات هندسی[11] ، به ویژه حملات معروف به RST (چرخش، تغییر مقیاس و انتقال[۱۲])، یکی از بزرگ‌ترین چالش‌ها در طراحی سیستم‌های نهان‌نگاری است[5] . این حملات با ایجاد ناهم‌زمانی[13] بین مراحل جاسازی و استخراج، بدون آنکه نیازی به تخریب مستقیم واترمارک داشته باشند، الگوریتم استخراج را در مکانیابی و بازیابی نشانه دیجیتال ناتوان می‌سازند. فرآیند نهان‌نگاری تصویر به طور کلی شامل دو مرحله اصلی جاسازی [14]و استخراج[15] است. در مرحله جاسازی، واترمارک در تصویر میزبان پنهان می‌شود و در مرحله استخراج، واترمارک از تصویر نهان‌نگاری‌شده بازیابی می‌گردد. یک نمایش کلی از این فرآیند، شامل رمزنگاری واترمارک پیش از جاسازی برای افزایش امنیت، در پژوهش [6] ارائه شده است.

راهکارهای نهان‌نگاری مقاوم را می‌توان به طور کلی به دو دسته حوزه مکان [۱۶]و حوزه تبدیل[۱۷] تقسیم‌بندی کرد. روش‌های حوزه مکان که با دستکاری مستقیم مقادیر پیکسل‌ها کار می‌کنند، ساده و سریع هستند اما معمولاً مقاومت کمی در برابر حملات دارند. در مقابل،

[1] Tampering
[2] Copyright
[3] Cryptography
[4] Digital Watermarking
[5] Watermark
[6] Host
[7] Trade-off
[8] Imperceptibility
[9] Robustness
[10] Capacity
[11] Geometric Attacks
[12] Rotation, Scaling, and Translation
[13] Desynchronization
[14] Embedding
[15] Extracting
[16] Spatial Domain
[17] Transform Domain

روش‌های حوزه تبدیل، ابتدا تصویر را توسط یک تبدیل ریاضی مانند تبدیل کسینوسی گسسته(DCT[1]) ، تبدیل موجک گسسته (DWT[2]) یا تجزیه مقادیر منفرد (SVD[3]) به فضای فرکانسی می‌برند و سپس واترمارک را در ضرایب این فضاها جاسازی می‌کنند. این رویکرد به دلیل تطابق بهتر با سیستم بینایی انسان (HVS[4]) و توزیع هوشمندانه‌تر انرژی واترمارک، به طور بالقوه مقاومت و شفافیت بالاتری را ارائه می‌دهد [7].

با این حال، موفقیت نهایی یک الگوریتم حوزه تبدیل، به شدت به دو تصمیم کلیدی وابسته است. نخست، انتخاب دقیق و بهینه فاکتور مقیاس (α) است که قدرت و انرژی واترمارک جاسازی شده را تعیین می‌کند. یک مقدار α بسیار کوچک، اگرچه شفافیت را در سطح عالی نگه می‌دارد، اما مقاومت لازم برای بقا در برابر حملات را فراهم نمی‌کند. برعکس، یک α بسیار بزرگ، مقاومت را افزایش می‌دهد اما با ایجاد اعوجاج قابل مشاهده، شفافیت تصویر را از بین می‌برد [۸]. دومین چالش اساسی، مکان‌یابی هوشمندانه برای جاسازی است. روش‌های سنتی اغلب واترمارک را به صورت یکنواخت در تمام نواحی یک زیرباند فرکانسی (مثلاًLL) توزیع می‌کنند. این در حالی است که نواحی مختلف تصویر، ظرفیت پنهان‌سازی متفاوتی دارند. نواحی صاف و یکنواخت، کوچک‌ترین تغییر را به سرعت برای بیننده آشکار می‌کنند، در حالی که نواحی پرجزئیات، لبه‌ها و بافت‌های پیچیده، می‌توانند به طور طبیعی‌تری نویز جاسازی را استتار کنند. بنابراین، یک سیستم کارا باید قادر به تشخیص این نواحی و تخصیص قدرت جاسازی به صورت تطبیقی باشد [۹].

در سال‌های اخیر، ظهور و بلوغ فناوری‌های یادگیری عمیق[5] ، به‌ویژه شبکه‌های عصبی کانولوشنی (CNNs[6])، افق‌های جدیدی را در تحلیل و درک محتوای تصویر گشوده است [۹]. این شبکه‌ها با توانایی خارق‌العاده خود در استخراج سلسله‌مراتبی ویژگی‌های بصری از لبه‌های سطح پایین تا مفاهیم پیچیده، ابزاری ایده‌آل برای پرداختن به چالش دوم، یعنی مکان‌یابی هوشمندانه جاسازی، فراهم می‌کنند. از سوی دیگر، چالش انتخاب بهینه فاکتور مقیاس، یک مسئله بهینه‌سازی پیوسته است که حل آن با روش‌های سنتی سعی و خطا غیرممکن است. الگوریتم‌های فراابتکاری[7] مانند بهینه‌سازی ازدحام ذرات (PSO[8]) راه‌حلی مؤثر برای کاوش فضای جستجو ارائه می‌دهند [۱۰]. با این وجود، استفاده مستقیم از این الگوریتم‌ها در مسائل نهان‌نگاری بسیار پرهزینه است، چرا که تابع هدف (مقاومت/شفافیت) برای هر ارزیابی، نیازمند اجرای کامل فرآیند جاسازی، اعمال حمله و استخراج واترمارک بر روی یک مجموعه داده بزرگ است. برای غلبه بر این مانع، مفهوم مدل‌های جانشین[9] که تقریب‌های سریع و کم‌هزینه از تابع هدف اصلی ارائه می‌دهند، مطرح شده است [۱۱].

در این پژوهش، ما یک چارچوب نهان‌نگاری ترکیبی و هوشمند ارائه می‌دهیم که با تلفیق استراتژیک پیشرفت‌های حوزه‌های مختلف، به طور هم‌زمان به هر دو چالش فوق پاسخ می‌دهد. نوآوری‌های کلیدی این مقاله عبارتند از:

۱. ارائه یک سیستم دو مرحله‌ای تلفیقی: معماری پیشنهادی شامل یک فاز تحلیل مبتنی بر CNN و یک فاز جاسازی بهینه‌سازی شده است که نقاط قوت یادگیری عمیق و روش‌های حوزه تبدیل را با یکدیگر ترکیب می‌کند.

---

[1] Discrete Cosine Transform
[2] Discrete Wavelet Transform
[3] Singular Value Decomposition
[4] Human Visual System
[5] Deep Learning
[6] Convolutional Neural Networks
[7] Metaheuristic
[8] Particle Swarm Optimization
[9] Surrogate Models

۲. طراحی یک شبکه CNN پیشگو و مستقل از جاسازی: ما یک شبکه U-Net[1] سبک و کارا را برای تحلیل ساختاری تصویر میزبان و تولید یک نقشه وزن [2] تطبیقی آموزش می‌دهیم. این نقشه، شایستگی هر ناحیه از تصویر برای جاسازی مقاوم و شفاف را نشان می‌دهد. بر خلاف بسیاری از روش‌های نهان‌نگاری عمیق سرتاسری[3]، شبکه ما مستقیماً در فرآیند جاسازی/استخراج دخالت ندارد و صرفاً به عنوان یک تحلیل‌گر صحنه عمل می‌کند.

۳. معرفی یک فرآیند بهینه‌سازی مبتنی بر مدل جانشین: برای اولین بار، با موفقیت از یک مدل جانشین مبتنی بر درخت رگرسیون برای تخمین تابع هزینه در بهینه‌سازی فاکتور مقیاس استفاده می‌شود. این مدل که بر روی تعداد محدودی نمونه آموزش دیده، جایگزین ارزیابی‌های پرهزینه تابع واقعی شده و بهینه‌سازی PSO را به‌طور چشمگیری تسریع می‌کند.

۴. جاسازی کاملاً تطبیقی: فاکتور مقیاس بهینه‌یافته سراسری (α) با نقشه وزن محلی CNN تلفیق می‌شود تا یک قدرت جاسازی نقطه‌ای و تطبیقی ایجاد کند. این امر تضمین می‌کند که انرژی واترمارک به صورت هوشمندانه در نواحی با پتانسیل پنهان‌سازی بالا متمرکز شده و از آسیب‌رسانی به نواحی حساس تصویر جلوگیری می‌شود.

ساختار ادامه مقاله به شرح زیر است: در بخش ۲، پیشینه جامعی از پژوهش‌های مرتبط با روش پیشنهادی ارائه می‌شود. بخش ۳ به تشریح دقیق و گام‌به‌گام روش پیشنهادی، از تولید نقشه وزن CNN تا جاسازی بهینه‌شده، اختصاص دارد. نتایج تجربی، ارزیابی‌های کمی و کیفی و مقایسه با روش‌های پیشین در بخش ۴ ارائه می‌شوند. در نهایت، بخش ۵ به نتیجه‌گیری و ارائه پیشنهادات برای پژوهش‌های آتی می‌پردازد.

۲- مروری جامع بر پیشینه پژوهش

حوزه نهان‌نگاری تصویر، به‌ویژه با هدف دستیابی به مقاومت در برابر حملات هندسی، یکی از پویاترین زمینه‌های تحقیقاتی در پردازش سیگنال و امنیت چندرسانه‌ای است. تلاش‌های پژوهشی را می‌توان در قالب یک سیر تکاملی از روش‌های کلاسیک حوزه تبدیل به سمت رویکردهای هوشمند مبتنی بر بهینه‌سازی و در نهایت، یادگیری عمیق دسته‌بندی کرد. در این بخش، مروری عمیق بر این پیشینه ارائه می‌شود تا جایگاه و نوآوری پژوهش حاضر به روشنی تبیین گردد.

۲-۱- روش‌های پایه در حوزه تبدیل: از DCT تا SVD

نسل اول سیستم‌های نهان‌نگاری مقاوم، عمدتاً بر روی استفاده از تبدیلات حوزه فرکانس برای ارائه یک مصالحه اولیه بین شفافیت و مقاومت متمرکز بودند.

- روش‌های مبتنی بر DCT : تبدیل کسینوسی گسسته به دلیل خاصیت تراکم انرژی [4] عالی و تطابق نسبی با استاندارد فشرده‌سازی JPEG، یکی از محبوب‌ترین انتخاب‌ها بوده است. یک نقطه عطف در این حوزه، کار Cox و همکاران (۱۹۹۷) [۱۲] بود که اساس طیف گسترده[5] را برای نهان‌نگاری بنا نهادند و واترمارک را در ضرایب میانی DCT جاسازی کردند. در ادامه، [۱۳]یک تکنیک DCT مقاوم در برابر JPEG ارائه دادند که با مفهوم باقیمانده ریاضی به تنظیم دقیق ضرایب فرکانس پایین می‌پرداخت [۱۴]. با ارائه روش DCT دو سطحی بر روی تصاویر رنگی، گامی فراتر نهاده و واترمارک رنگی را در ضرایب DC و AC جاسازی

[1] U-shaped Network
[2] Weight Map
[3] deep end-to-end watermarking
[4] Energy Compaction
[5] Spread Spectrum

کردند که به مقاومت بالایی در برابر حملات رایج دست یافت. با این حال، ماهیت سراسری DCT و عدم تفکیک فرکانسی-مکانی، انعطاف‌پذیری این روش‌ها را برای جاسازی تطبیقی و مقاومت در برابر حملات هندسی محدود می‌کرد.

- روش‌های مبتنی برDWT : تبدیل موجک گسسته با ارائه یک نمایش چندوضوحی [1] از تصویر، توانست بر محدودیت DCT در تحلیل هم‌زمان فرکانس و مکان غلبه کند. این تبدیل با تجزیه تصویر به زیرباندهای تقریبی (LL) و جزئیات(LH, HL, HH) ، امکان انتخاب زیرباند مناسب برای جاسازی را فراهم می‌کند [۱۵]. یک روش جامع با تجزیه موجک سه سطحی در فضای رنگی غیرهمبسته ارائه دادند و با کمک الگوریتم کلونی زنبور مصنوعی(ABC) ، فاکتور مقیاس بهینه برای جاسازی در تمامی زیرباندها را یافتند [۱۶]. نیز یک روش بلوک‌بندی مبتنی بر DWT برای تصاویر رنگی پیشنهاد دادند که بر مقاومت در برابر نویز و فشرده‌سازی تمرکز داشت.
- روش‌های مبتنی بر SVD و تجزیه‌های ماتریسی: SVD به دلیل پایداری ذاتی مقادیر تکین [2] در برابر تغییرات جزئی و برخی حملات هندسی، به ابزاری قدرتمند تبدیل شد. روش‌های مبتنی برSVD ، مانند کار[17] برای تصاویر پزشکی، واترمارک را با اصلاح مقادیر تکین تصویر یا زیرباندهای آن جاسازی می‌کنند. با این حال، کاربرد مستقل SVD ممکن است با مشکل مثبت کاذب [3]در استخراج مواجه شود. با این حال، کاربرد مستقل SVD ممکن است با مشکل مثبت کاذب در استخراج مواجه شود. برای رفع این مشکل و نیز کاهش هزینه محاسباتی بالایSVD ، نسخه‌های توسعه‌یافته‌ای مانند RSVD (SVD افزونه) پیشنهاد شده است[18] . در بسیاری از طرح‌های واترمارکینگ، بکارگیری تکنیک‌هایی همچون ممان زرنایک [۴] نیز می‌تواند توانایی مقاومت در برابر حملات هندسی و نویز را تقویت کند. راه حل دیگرِ رفع این مشکل و بهره‌گیری از مزایای هم‌افزایی، روش‌های ترکیبی مانند-DWT DCT-SVDهستند [۷]. یک نمونه شاخص از این رویکرد است که واترمارک را پس از اعمال متوالی DCT و DWT بر روی تصویر و تقسیم واترمارک، جاسازی می‌کند. نقطه ضعف مشترک تمامی این روش‌های ارزشمند، نگاه ایستا به تصویر و انتخاب تجربی فاکتور مقیاس است. آن‌ها تصویر را به صورت یک موجودیت یکپارچه می‌بینند و از ظرفیت پنهان‌سازی متغیر در نواحی مختلف آن چشم‌پوشی می‌کنند.

۲-۲- روش‌های مبتنی بر بهینه‌سازی و ضرورت مدل‌های جانشین

برای غلبه بر محدودیت انتخاب تجربی فاکتور مقیاس، الگوریتم‌های بهینه‌سازی فراابتکاری به عنوان راهکاری مؤثر وارد میدان شدند. این الگوریتم‌ها، مسئله نهان‌نگاری را به یک مسئله بهینه‌سازی تبدیل می‌کنند که در آن، هدف، یافتن پارامترهایی (عمدتاً فاکتور مقیاس و گاهی مکان‌های جاسازی) است که یک تابع هزینه چندهدفه (ترکیب وزنی از معیارهای شفافیت و مقاومت) را کمینه کنند.

محققان در [۱۰] یک رویکرد مبتنی بر DWT-SVD-HD ارائه دادند که در آن، فاکتور مقیاس توسط الگوریتم کرم شب‌تاب (FA[5])بهینه‌سازی می‌شد. این روش با در نظر گرفتن هم‌زمان استحکام و نامحسوس بودن، بهبود قابل توجهی را نسبت به روش‌های با پارامترهای دستی نشان داد. [19] نیز از تبدیل آرنولد برای رمزنگاری واترمارک و SVD برای جاسازی بهینه‌شده استفاده کردند. [۲۰] یک روش واترمارک چندگانه در حوزه DCT ارائه دادند که در آن، کدهای تکرارشونده برای افزایش مقاومت به کار رفته بود.

---

[1] Multi-resolution
[2] Singular Values
[3] False Positive
[4] Zernike Moments
[5] Firefly Algorithm

با این حال، یک چالش اساسی در کاربرد مستقیم الگوریتم‌های فراابتکاری وجود دارد: هزینه محاسباتی گزاف. در هر تکرار از الگوریتم بهینه‌ساز (برای هر ذره در PSO یا هر کرم شب‌تاب در FA)، یک کاندیدای جدید برای فاکتور مقیاس باید ارزیابی شود. این ارزیابی شامل جاسازی واترمارک با آن فاکتور، اعمال یک یا چند حمله به تصویر، استخراج واترمارک و محاسبه نهایی خطا (مثلاً BER) است. این فرآیند برای یک مجموعه داده تست، فوق‌العاده زمان‌بر بوده و عملاً استفاده از الگوریتم‌های با جمعیت بالا و تکرار زیاد را غیرممکن می‌سازد. راه‌حلی که در دیگر شاخه‌های مهندسی و اخیراً در یادگیری ماشین برای تنظیم هایپرپارامترها مرسوم شده، استفاده از مدل‌های جانشین است [۱۱]. یک مدل جانشین، یک مدل تقریبی و سریع (مانند رگرسیون، درخت تصمیم یا فرآیند گوسی) است که بر روی تعداد محدودی ارزیابی واقعی از تابع هدف آموزش می‌بیند و سپس می‌تواند به عنوان یک تخمین‌گر کم‌هزینه برای هدایت الگوریتم بهینه‌ساز به کار رود. این دقیقاً همان شکاف پژوهشی است که ما در این مقاله به آن می‌پردازیم.

۲-۳- ظهور یادگیری عمیق در نهان‌نگاری

در سال‌های اخیر، یادگیری عمیق، انقلابی در پردازش تصویر ایجاد کرده و به طور طبیعی، کاربردهای آن در نهان‌نگاری نیز مورد توجه شدید قرار گرفته است. مرور جامع [۲] نشان می‌دهد که روش‌های مبتنی بر یادگیری عمیق به سرعت در حال تبدیل شدن به جریان اصلی هستند.

- روش‌های انتها به انتها[1]: این دسته از روش‌ها، کل فرآیند جاسازی و استخراج را به عنوان یک شبکه عصبی عمیق مدل‌سازی می‌کنند. یک شبکه رمزگذار(مسیر انقباضی)، تصویر میزبان و واترمارک را دریافت کرده و تصویر جاسازی‌شده را تولید می‌کند. یک شبکه رمزگشا(مسیر انبساطی) نیز وظیفه استخراج واترمارک را بر عهده دارد. اگرچه این روش‌ها نتایج چشمگیری در شفافیت ارائه می‌دهند، اما معمولاً مقاومت کمتری در برابر حملات هندسی دارند، مگر آنکه در حین آموزش، نمونه‌های حملات دیده شده[2] به شبکه خورانده شوند که این خود می‌تواند فرآیند آموزش را ناپایدار و بسیار طولانی کند[2] .
- روش‌های تلفیقی[3] : این رویکرد که پژوهش ما نیز در این دسته قرار می‌گیرد، از قدرت تحلیل CNN ها در کنار ساختارهای کلاسیک و مقاوم استفاده می‌کند .[9]یک چارچوب قدرتمند به نام MDResNet-HDWM ارائه دادند. در این روش، یک شبکه ResNetچندمقیاسی، نقاط کلیدی مقاوم در برابر حملات هندسی را در تصویر شناسایی کرده و واترمارک تنها در نواحی اطراف این نقاط جاسازی می‌شود. این روش به دلیل گره زدن فرآیند استخراج به نقاط کلیدی، مقاومت ذاتی بسیار بالایی در برابر حملات RSTدارد .[21]نیز یک روش نهان‌نگاری کور مبتنی بر CNN در حوزه موجک ارائه دادند که در آن، از یک شبکه کانولوشنی برای یادگیری رابطه بین ضرایب موجک و بیت‌های واترمارک استفاده شد.

ایده کلیدی که از پژوهش‌های فوق الهام گرفته می‌شود، این است که CNN ها تحلیل‌گرهای صحنه بی‌نظیری هستند. به جای آنکه از یک شبکه برای کل فرآیند استفاده کنیم، می‌توانیم از یک شبکه سبک‌تر و هدفمندتر صرفاً برای تولید یک نقشه راهنمای جاسازی استفاده کنیم. این نقشه می‌تواند پیچیدگی‌های بصری را کدگذاری کرده و یک فرآیند جاسازی کلاسیک اما اثبات‌شده در حوزه موجک را به صورت تطبیقی هدایت کند. این رویکرد، مزایای هر دو جهان را دارد: مقاومت ساختاری روش‌های حوزه موجک و هوشمندی و تطبیق‌پذیری CNN ها.

[1] End-to-End
[2] Augmented Data
[3] Hybrid Deep Learning

۲-۴- جمع‌بندی و جایگاه پژوهش حاضر

با بررسی جامع پیشینه، شکاف‌های پژوهشی زیر به وضوح قابل مشاهده است:

۱. عدم تطبیق‌پذیری کامل: روش‌های کلاسیک حوزه تبدیل، فاقد مکانیزمی هوشمند برای تشخیص و بهره‌برداری از ظرفیت پنهان‌سازی متغیر در نواحی مختلف تصویر هستند.

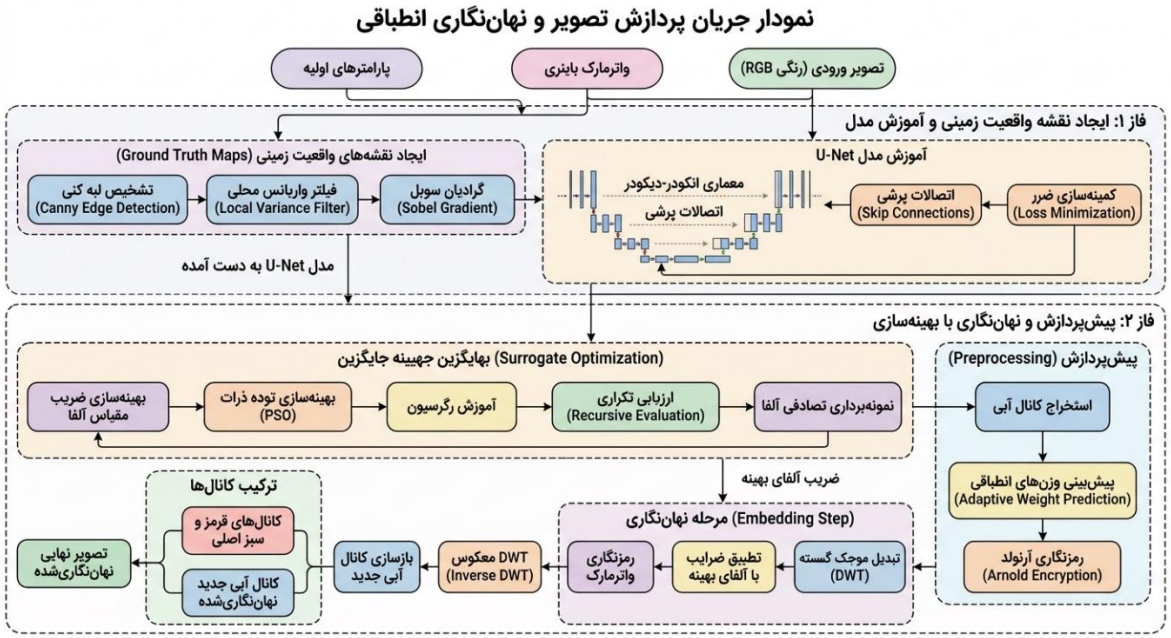
۲. هزینه بالای بهینه‌سازی مستقیم: کاربرد مستقیم الگوریتم‌های فرابتکاری برای بهینه‌سازی پارامترهای نهان‌نگاری، به دلیل ارزیابی‌های مکرر و پرهزینه تابع واقعی، عملاً ناکارآمد است.

۳. پیچیدگی آموزش نهان‌نگاری عمیق سرتاسری: روش‌های تماماً مبتنی بر یادگیری عمیق، برای دستیابی به مقاومت در برابر حملات هندسی، نیازمند آموزش‌های پیچیده و طولانی با مجموعه داده‌های عظیم هستند و اغلب مقاومتشان در برابر حملات دیده‌نشده تضمین‌شده نیست.

پژوهش حاضر دقیقاً در محل تلاقی این شکاف‌ها قرار می‌گیرد و یک راه‌حل یکپارچه و نوآورانه ارائه می‌دهد. ما برای اولین بار، یک مدل جانشین را برای تسریع بهینه‌سازی فاکتور مقیاس در کنار یک شبکه CNN که برای تولید نقشه وزن تطبیقی آموزش دیده، ترکیب می‌کنیم. این تلفیق منحصربه‌فرد، یک سیستم نهان‌نگاری را شکل می‌دهد که نه تنها از نظر مفهومی نوین است، بلکه از نظر عملی نیز کارایی بالایی در برقراری تعادل میان شفافیت و مقاومت در برابر حملات هندسی دارد.

۳- روش پیشنهادی: چارچوب نهان‌نگاری ترکیبی مبتنی بر CNN و مدل جانشین

در این بخش، جزئیات فنی و مبانی نظری چارچوب پیشنهادی ارائه می‌گردد. روش ارائه‌شده یک سیستم دو فازه است که هر فاز آن برای رفع یکی از چالش‌های بنیادین در نهان‌نگاری مقاوم طراحی شده است: فاز اول به تحلیل هوشمند محتوای تصویر برای شناسایی نواحی مستعد جاسازی می‌پردازد و فاز دوم، با بهره‌گیری از یک مکانیزم بهینه‌سازی سریع، فرآیند جاسازی تطبیقی را به انجام می‌رساند. شمای مفهومی و گردش کار کلی این چارچوب در شکل۱ به نمایش درآمده است.



شکل ۱فلوچارت روش پیشنهادی

همان‌طور که در شکل ۱ مشاهده می‌شود، خط لوله پردازشی روش پیشنهادی از دو فاز مجزا اما کاملاً مرتبط تشکیل شده است. نقطه شروع، تصویر ورودی رنگی (RGB) و واترمارک باینری است که به موازات یکدیگر وارد سیستم می‌شوند. در فاز اول، هدف اصلی، ساخت یک تحلیل‌گر بصری هوشمند است. این تحلیل‌گر که مبتنی بر یک شبکه عصبی کانولوشنی با معماری U-Net پیاده‌سازی شده، ابتدا با استفاده از یک مجموعه داده آموزشی که شامل جفت‌های تصویر میزبان و نقشه حقیقت مبنا است، آموزش می‌بیند. نقشه‌های حقیقت مبنا[1] از پیش و با ترکیب هوشمندانه‌ای از معیارهای بینایی ماشین نظیر آشکارساز لبه کَنی[2]، تحلیل بافت و محاسبه بزرگی گرادیان تولید می‌شوند. فرآیند آموزش، شامل کمینه‌سازی خطای بازسازی[3] در معماری Encoder-Decoder مجهز به اتصالات پرشی[4] است. خروجی نهایی این فاز، یک مدل آموزش‌دیده است که قادر است برای هر تصویر دلخواه، یک نقشه وزن تطبیقی تولید کند. این نقشه، یک ماتریس دوبُعدی است که در آن، مقادیر بالاتر نشان‌دهنده نواحی با پیچیدگی بصری بیشتر (مانند لبه‌ها و بافت‌ها) و در نتیجه، مستعدتر برای جاسازی مقاوم و نامحسوس واترمارک است.

فاز دوم، هسته عملیاتی سیستم را تشکیل می‌دهد و خود به دو مسیر موازی تقسیم می‌شود که در نهایت با یکدیگر تلفیق می‌گردند. در این فاز، تصویر میزبان اصلی ابتدا وارد مرحله پیش‌پردازش می‌شود که مهم‌ترین گام آن، جداسازی و استخراج کانال آبی تصویر است. این انتخاب مبتنی بر حساسیت کمتر سیستم بینایی انسان به تغییرات در این طیف رنگی است. هم‌زمان، پارامترهای اولیه سیستم از جمله فاکتور مقیاس، وارد یک حلقه بهینه‌سازی می‌شوند. این حلقه با محوریت مدل جانشین و الگوریتم PSO ، مقدار بهینه فاکتور مقیاس را به نحوی تعیین می‌کند که تعادل میان شفافیت و مقاومت را برقرار سازد. استفاده از مدل جانشین در این مرحله، یک نوآوری کلیدی برای کاهش چشمگیر هزینه‌های محاسباتی ناشی از ارزیابی‌های مکرر است.

پس از تعیین فاکتور مقیاس بهینه و دریافت نقشه وزن از فاز اول، هر دو به مرحله نهان‌نگاری هدایت می‌شوند. در این گام، جاسازی واترمارک باینری به صورت تطبیقی و نقطه‌ای در ضرایب زیرباند تقریبی (LL) حاصل از تبدیل موجک گسسته انجام می‌پذیرد. قدرت جاسازی در هر پیکسل، حاصل‌ضرب فاکتور مقیاس بهینه سراسری در وزن محلی تعیین‌شده توسط CNN است. این مکانیزم تضمین می‌کند که انرژی واترمارک به طور هوشمندانه در نواحی پرجزئیات تصویر متمرکز شده و از نواحی صاف و حساس دوری می‌کند. در نهایت، با بازسازی کانال آبی جدید از طریق اعمال معکوس تبدیل موجک و ادغام آن با کانال‌های قرمز و سبز دست‌نخورده، تصویر نهایی نهان‌نگاری‌شده حاصل می‌شود. این تصویر باید به گونه‌ای باشد که هم کیفیت بصری بالایی داشته باشد و هم در برابر حملات هندسی و پردازشی مقاومت کند. در ادامه این بخش، مبانی نظری و جزئیات ریاضی هر یک از این مراحل به طور کامل تشریح می‌گردد.

- ۳-۱- فاز اول: تولید نقشه وزن تطبیقی مبتنی بر شبکه U-Net

هدف بنیادین در این فاز، طراحی یک تحلیل‌گر صحنه هوشمند است که بتواند فارغ از فرآیند جاسازی، محتوای تصویر میزبان را درک کرده و یک راهنمای مکانی برای تخصیص قدرت جاسازی تولید کند. این راهنما که آن را نقشه وزن[5] می‌نامیم، یک ماتریس دوبُعدی با ابعاد تصویر میزبان است که در آن، هر پیکسل مقداری بین صفر (کاملاً نامناسب برای جاسازی) و یک (ایده‌آل برای جاسازی) دارد. برای دستیابی به این هدف، یک شبکه عصبی کانولوشنی عمیق با معماری U-Net طراحی و پیاده‌سازی شده است. انتخاب این معماری به دلیل

[1] Ground Truth Maps
[2] Canny
[3] Loss Minimization
[4] Skip Connections
[5] Weight Map

توانایی ذاتی آن در استخراج سلسله‌مراتبی ویژگی‌ها از طریق مسیر انقباضی[1] و بازیابی دقیق جزئیات مکانی از طریق مسیر انبساطی[2] و اتصالات پرشی[3] صورت گرفته است. این ویژگی‌ها برای تولید یک نقشه وزنی با وضوح کامل که کوچک‌ترین جزئیات ساختاری تصویر را منعکس کند، حیاتی است.

۳-۱-۱- طراحی و تولید داده‌های هدف

از آنجا که نقشه‌های وزن ایده‌آل به صورت آماده در دسترس نیستند، یک گام اساسی در این فاز، تولید یک مجموعه داده آموزشی با کیفیت بالا شامل جفت‌های تصویر میزبان و نقشه وزن هدف است. برای ساخت نقشه هدف، که به عنوان حقیقت مبنا برای آموزش شبکه عمل می‌کند، یک روش ترکیبی مبتنی بر اصول سیستم بینایی انسان(HVS[4]) طراحی شده است. منطق اساسی این است که نواحی با پیچیدگی بصری بالا، مانند لبه‌ها و بافت‌های متراکم، به دلیل اثر پوشانندگی[5] ، تغییرات ناشی از جاسازی واترمارک را بهتر استتار می‌کنند.

بر این اساس، برای هر تصویر در مجموعه داده آموزشی، نقشه حقیقت مبنا (G) از ترکیب وزنی سه معیار بینایی محاسباتی زیر ساخته می‌شود:

نقشه لبه[6](E) : برای استخراج لبه‌های قوی و شاخص تصویر، از آشکارساز لبه کَنی با آستانه‌های [۰.۰۵, ۰.۱۵] استفاده می‌شود. خروجی این مرحله یک نقشه باینری است که در آن، پیکسل‌های لبه مقدار ۱ و سایر پیکسل‌ها مقدار ۰ دارند. لبه‌ها به عنوان شاخص‌ترین نشانه‌های بصری، بهترین نامزدها برای جاسازی مقاوم هستند.

نقشه بافت [7](T): مطابق با رابطه (۱) تحلیل بافت به ما در یافتن نواحی با تغییرات شدت نور بالا و تکرارشونده کمک می‌کند. برای کمی‌سازی این مفهوم، از فیلتر واریانس محلی[8] بر روی تصویر استفاده می‌شود. یک پنجره با ابعاد مشخص (مثلاً ۷×۷) بر روی تصویر حرکت کرده و واریانس مقادیر پیکسل‌های درون آن محاسبه می‌شود. مقادیر بالای واریانس، نشان‌دهنده بافت‌های پیچیده و مناسب برای جاسازی است.

$$T(x,y) = \frac{1}{\lceil W \rceil} \sum_{(i,j)\in W_{x,y}} \left(I(i,j) - \mu W_{x,y}\right)^2 \qquad (۱)$$

که در آن $I(i,j)$ شدت روشنایی پیکسل، $W_{x,y}$ پنجره اطراف پیکسل (x,y) ، $\lceil W \rceil$ تعداد پیکسل‌های پنجره، و $\mu W_{x,y}$ ، میانگین شدت روشنایی در آن پنجره است. خروجی نهایی نرمال‌سازی می‌شود.

نقشه بزرگی گرادیان [9](M) : این معیار بر قدرت لبه‌ها تأکید دارد. بزرگی گرادیان تصویر با استفاده از عملگرهای مشتق افقی G_x و عمودی G_y (مانند فیلتر سوبِل[10]) محاسبه شده و سپس اندازه بردار گرادیان برای هر پیکسل مطابق با رابطه (۲) به‌دست می‌آید.

$$M(x,y) = \sqrt{G_x(x,y)^2 + G_y(x,y)^2} \qquad (۲)$$

در نهایت، نقشه حقیقت مبنا (G) از ترکیب خطی این سه معیار نرمال‌سازی‌شده مطابق با رابطه (۳) ایجاد می‌شود:

---

[1] Encoder
[2] Decoder
[3] Skip Connections
[4] Human Visual System
[5] Masking Effect
[6] Edge Map
[7] Texture Map
[8] Local Variance
[9] Gradient Magnitude Map
[10] Sobel

$$G = \lambda_1 E_{norm} + \lambda_2 T_{norm} + \lambda_3 M_{norm} \quad (۳)$$

که در آن $\lambda_1, \lambda_2, \lambda_3$ اوزان تجربی هستند که اهمیت نسبی هر معیار را مشخص می‌کنند و تابع شرط $\lambda_1 + \lambda_2 + \lambda_3 = 1$ است. در پیاده‌سازی ما، وزن‌های $0.4, 0.3, 0.3$ به‌ترتیب برای لبه، بافت و گرادیان استفاده شده است که تأکید بیشتری بر ساختارهای لبه‌ای دارد.

۳-۱-۲- معماری شبکه و فرآیند آموزش

برای یادگیری نگاشت پیچیده از تصویر ورودی به نقشه هدف G، یک شبکه کاملاً کانولوشنی با معماری U-Net استفاده شده است. این معماری از دو بخش اصلی تشکیل شده است:

مسیر انقباضی[1]: وظیفه این مسیر، استخراج سلسله‌مراتبی ویژگی‌های بصری از تصویر ورودی است. این مسیر شامل دو بلوک اصلی است که هر یک از دو لایه کانولوشن ۳×۳ با تابع فعال‌سازی ReLU و لایه نرمال‌سازی دسته‌ای[2] تشکیل شده است. تعداد فیلترها در بلوک اول ۳۲ و در بلوک دوم ۶۴ است. در انتهای هر بلوک، یک لایه Max Pooling با ابعاد ۲×۲ و گام ۲، ابعاد فضایی نقشه‌های ویژگی را به نصف کاهش می‌دهد. در عمیق‌ترین بخش شبکه (بخش گلوگاهی)، دو لایه کانولوشن با ۱۲۸ فیلتر بدون کاهش ابعاد، انتزاعی‌ترین ویژگی‌های تصویر را استخراج می‌کنند.

مسیر انبساطی[3]: وظیفه این مسیر، بازسازی یک نقشه با ابعاد تصویر ورودی از روی ویژگی‌های انتزاعی استخراج‌شده است. این کار با استفاده از لایه‌های کانولوشن انتقالی[4] با هسته‌های ۲×۲ و گام ۲ انجام می‌شود که به تدریج ابعاد فضایی را افزایش می‌دهند. برای جبران جزئیات مکانی که در طی فرآیند کاهش ابعاد از دست رفته‌اند، از اتصالات پرشی[5] استفاده می‌شود. این اتصالات، خروجی لایه‌های متناظر در مسیر انقباضی را مستقیماً به ورودی بلوک‌های مسیر انبساطی اضافه می‌کنند و به این ترتیب، شبکه می‌تواند اطلاعات دقیق مربوط به لبه‌ها و بافت‌ها را که برای تولید یک نقشه وزن باکیفیت ضروری است، حفظ کند.

لایه خروجی: در انتهای شبکه، یک لایه کانولوشن با فیلتر ۱×۱ قرار دارد که نقشه‌های ویژگی چندکاناله را به یک خروجی تک‌کاناله تبدیل می‌کند. این خروجی از یک لایه رگرسیون عبور داده می‌شود که تابع هزینه آن، میانگین شبکه با کمینه‌سازی تابع خطای میانگین مربعات (MSE) بین خروجی $\hat{G}$ و هدف $\hat{G}$ مطابق با رابطه (۴) آموزش می‌بینند.

$$\mathcal{L}_{MSE} = \frac{1}{N}\sum_{i=1}^{N}(\hat{G}_i - G_i)^2 \quad (۴)$$

که در آن N تعداد کل پیکسل‌ها است. پس از آموزش، این شبکه قادر است برای هر تصویر دلخواه یک نقشه وزن W تولید کند که مقادیر آن بازتابی از مناسب بودن هر ناحیه برای جاسازی است.

فرآیند آموزش بر روی یک مجموعه داده شامل ۲۰۰۰ نمونه تصویری با ابعاد ۱۲۸×۱۲۸ پیکسل که از تصاویر استاندارد تولید شده‌اند، انجام می‌پذیرد. شبکه با استفاده از بهینه‌ساز آدام[6]، نرخ یادگیری اولیه ۰.۰۰۱ و یک استراتژی کاهش گام‌به‌گام نرخ یادگیری[7] آموزش می‌بینند. بر اساس این استراتژی، نرخ یادگیری هر ۱۰ دوره با ضریب ۰.۵ کاهش می‌یابد. این مکانیزم اجازه می‌دهد که شبکه در مراحل اولیه آموزش با

[1] Encoder
[2] Batch Normalization
[3] Decoder
[4] Convolution Transpose
[5] Skip Connections
[6] Adam
[7] Step-wise Decay

گام‌های بلند به سرعت به همگرایی نزدیک شود و در مراحل پایانی، با گام‌های کوچک‌تر به تنظیم دقیق وزن‌ها بپردازد. برای جلوگیری از بیش‌برازش، از تکنیک‌های تنظیم‌سازی L2 و محدودسازی گرادیان[1] استفاده شده است. پس از اتمام موفقیت‌آمیز فرآیند آموزش که جزئیات و معیارهای عملکرد آن به طور کامل در بخش ۴-۲ تحلیل خواهد شد، شبکه آموزش‌دیده قادر است برای هر تصویر دلخواه، یک نقشه وزن تطبیقی با دقت بالا تولید کند.

۳-۲- فاز دوم: جاسازی تطبیقی بهینه‌شده در حوزه موجک با استفاده از مدل جانشین و PSO

پس از تکمیل فاز تحلیل و آماده‌سازی نقشه راهنمای جاسازی، فاز دوم که هسته اصلی عملیات نهان‌نگاری است، آغاز می‌شود. این فاز از یک ساختار ترکیبی و هوشمند بهره می‌برد که در آن، دو مؤلفه کلیدیِ فاکتور مقیاس بهینه و نقشه وزن تطبیقی با یکدیگر تلفیق می‌شوند تا جاسازی واترمارک به صورت کاملاً تطبیقی و مقاوم در حوزه تبدیل موجک گسسته (DWT) صورت پذیرد. رویکرد اصلی در این فاز، بهینه‌سازی غیرمستقیم و سریع پارامتر حساس سیستم با استفاده از یک مدل جانشین[2] است که جایگزین ارزیابی‌های مکرر و پرهزینه می‌شود.

۳-۲-۱- پیش‌پردازش و انتخاب کانال میزبان

تصویر میزبان رنگی (I_{host}) در فضای رنگی استاندارد RGB دریافت می‌شود. اولین گام در فرآیند جاسازی، انتخاب یک کانال رنگی مناسب است. بر اساس مطالعات گسترده در حوزه HVS ، چشم انسان کمترین حساسیت را به تغییرات در کانال آبی دارد. بنابراین، برای بیشینه‌سازی شفافیت و نامحسوس‌بودن تغییرات، کانال آبی (B_{host}) به عنوان بستر اصلی جاسازی انتخاب می‌شود. این انتخاب، یک لایه تضمین اولیه برای حفظ کیفیت بصری تصویر ایجاد می‌کند.

۳-۲-۲- تجزیه موجک و انتخاب زیرباند مقاوم

بر روی کانال آبی انتخاب‌شده B_{host}، یک تبدیل موجک گسسته دوبُعدی دو سطحی با موجک Haar اعمال می‌شود. این تبدیل، تصویر را به یک مجموعه از زیرباندهای تقریبی (LL) و جزئیات LL_2, LH_2, HL_2, HH_2 مطابق با رابطه (۵) در سطوح مختلف تجزیه می‌کند. از میان این زیرباندها، زیرباند LL_2 به عنوان میزبان نهایی برای جاسازی انتخاب می‌شود. انتخاب این زیرباند مبتنی بر چند دلیل راهبردی است: LL_2 یک نسخه فشرده و پایین‌گذر از تصویر است که بخش عمده انرژی و اطلاعات ساختاری تصویر را در خود متمرکز کرده است. این ویژگی، مقاومت ذاتی بالایی در برابر بسیاری از حملات پردازش سیگنال، به ویژه فشرده‌سازی JPEG و فیلترینگ پایین‌گذر، فراهم می‌آورد.

$$[LL_2, LH_2, HL_2, HH_2] = DWT_2(B_{host}) \quad (۵)$$

۳-۲-۳- چارچوب بهینه‌سازی مبتنی بر مدل جانشین

تعیین مقدار بهینه فاکتور مقیاس(α) ، پارامتری که قدرت جاسازی را کنترل می‌کند، یک مسئله بهینه‌سازی پیوسته و چندهدفه است. یک α بسیار کوچک، مقاومت را قربانی شفافیت می‌کند و یک α بسیار بزرگ، کیفیت بصری را تخریب می‌کند. حل این مسئله با روش‌های سنتی سعی و خطا یا حتی با الگوریتم‌های فراابتکاری استاندارد، به دلیل نیاز به ارزیابی‌های مکرر و پرهزینه، عملاً ناکارآمد است. هر ارزیابی نیازمند اجرای کامل فرآیند جاسازی، اعمال مجموعه‌ای از حملات، استخراج واترمارک و محاسبه معیارهای خطا است.

برای غلبه بر این مانع، یک چارچوب بهینه‌سازی دومرحله‌ای نوآورانه طراحی شده است که از یک مدل جانشین بهره می‌برد. مدل جانشین، یک مدل تقریبی و سریع است که جایگزین تابع هزینه واقعی و پرهزینه می‌شود. فرآیند کار به این صورت است:

---

[1] Gradient Clipping
[2] Surrogate Model

نمونه‌برداری و ساخت مجموعه داده آموزشی: در گام اول، nنمونه تصادفی از فاکتور مقیاس (α_i) در بازه جستجوی $\alpha \in [\alpha_{min}, \alpha_{max}]$ تولید می‌شود. برای هر یک از این نمونه‌ها، یک بار تابع هزینه واقعی $F(\alpha_i)$ با اجرای فرآیند کامل نهان‌نگاری بر روی یک زیرمجموعه کوچک از تصاویر آموزشی محاسبه می‌شود. تابع هزینه واقعی به گونه‌ای تعریف می‌شود که مصالحه میان شفافیت و مقاومت را منعکس کند. در اینجا، این تابع به صورت نسبت اعوجاج (MSE) به توان جاسازی (α) مطابق با رابطه (۶) مدل می‌شود:

$$F(\alpha) = \frac{\text{MSE}(\alpha)}{\alpha + \epsilon} \qquad (۶)$$

این کار، یک مجموعه داده کوچک اما ارزشمند $D = \{(\alpha_i, F_i)\}_{i=1}^{n}$ ایجاد می‌کند.

آموزش مدل جانشین: در گام دوم، یک مدل درخت رگرسیون) بر روی مجموعه داده D آموزش داده می‌شود. درخت رگرسیون یک مدل غیرپارامتری و انعطاف‌پذیر است که می‌تواند روابط غیرخطی پیچیده را بدون نیاز به فرضیات اولیه مدل کند. مهم‌تر از آن، ارزیابی (پیش‌بینی) این مدل بسیار سریع است. مدل آموزش‌دیده به عنوان تخمین‌گر $\hat{F}(\alpha)$ عمل می‌کند و می‌تواند هزینه متناظر با هر α جدید را در کسری از ثانیه پیش‌بینی کند.

بهینه‌سازی باPSO : اکنون که یک تخمین‌گر سریع و کم‌هزینه در اختیار است، از الگوریتم PSO برای کاوش کامل فضای جستجو استفاده می‌شود. PSOیک الگوریتم فراابتکاری مبتنی بر جمعیت است که از رفتار اجتماعی پرندگان الهام گرفته شده است. در این چارچوب، هر ذره یک کاندیدا برای α است. در هر تکرار الگوریتم، موقعیت هر ذره با فراخوانی $\hat{F}$ ارزیابی می‌شود، نه تابع واقعی F. این نوآوری کلیدی باعث می‌شود که امکان اجرای هزاران ارزیابی در کسری از زمان مورد نیاز برای یک ارزیابی واقعی فراهم شود. به این ترتیب، PSO می‌تواند به طور کامل فضای جستجو را بکاود و به سمت α_{opt} همگرا شود. فرآیند به‌روزرسانی سرعت و موقعیت ذرات طبق روابط استاندارد PSO انجام می‌شود.

اکنون که یک مدل جانشین سریع $\hat{F}$ در اختیار داریم، از الگوریتم بهینه‌سازی ازدحام ذرات (PSO) برای یافتن α_{opt} استفاده می‌کنیم که $\hat{F}$ را کمینه می‌کند. PSOیک الگوریتم فراابتکاری مبتنی بر جمعیت است که از رفتار اجتماعی پرندگان الهام گرفته شده است. هر ذره در فضای جستجو (که در اینجا فضای یک‌بعدی α است) یک راه‌حل کاندیدا را نشان می‌دهد. موقعیت یک ذره با p و سرعت آن با v نمایش داده می‌شود.

در هر تکرارt ، سرعت و موقعیت هر ذره i با استفاده از روابط (۷) به‌روزرسانی می‌شود:

$$v_i(t+1) = \omega . v_i(t) + c_1 r_1 \big(pbest_i - p_i(t)\big) + c_2 r_2 \big(gbest - p_i(t)\big) \qquad (۷)$$
$$p_i(t+1) = p_i(t) + v_i(t+1)$$

که در آن:

- ω وزن اینرسی که نرخ اکتشاف و بهره‌برداری را کنترل می‌کند.
- c_1، c_2 ضرایب یادگیری شناختی و اجتماعی هستند.
- r_1، r_2 اعداد تصادفی با توزیع یکنواخت در بازه $[0,1]$ می‌باشند.
- $pbest_i$ بهترین موقعیتی است که ذره i تاکنون تجربه کرده است.
- $gbest$ بهترین موقعیتی است که در کل جمعیت یافت شده است.

در هر تکرار، هزینه یک موقعیت جدید p_i با فراخوانی تابع تخمین‌گر سریع $\hat{F}(p_i)$ نه تابع واقعی F ارزیابی می‌شود. این مهم‌ترین برتری روش ما است که اجازه می‌دهد PSO با هزاران ارزیابی، فضای جستجو را بدون هیچ هزینه محاسباتی قابل توجهی بکاود. پس از هم‌گرایی الگوریتم، $gbest$ نهایی به عنوان فاکتور مقیاس بهینه α_{opt} انتخاب می‌شود.

- ۳-۲-۴- جاسازی تطبیقی نهایی

پس از تعیین α_{opt} توسط حلقه PSO و تولید نقشه وزن (W) توسط CNN ، این دو در یک گام جاسازی یکپارچه و تطبیقی تلفیق می‌شوند. برای این کار، ابتدا نقشه وزن W (که در فاز اول با ابعاد ۱۲۸×۱۲۸ تولید شده) به ابعاد زیرباند $LL2$ نمونه‌برداری مجدد (Resize)می‌شود تا یک تناظر یک‌به‌یک بین وزن‌ها و ضرایب موجک برقرار گردد. هم‌چنین، واترمارک باینری (WM_{binary}) که پیش‌تر با نگاشت آرنولد[1] رمزنگاری و درهم‌سازی شده، به همان ابعاد تنظیم می‌شود.

عملیات جاسازی با استفاده از رابطه (۸) انجام می‌پذیرد:

$$LL'_2(x,y) = LL_2(x,y) + \left(\alpha_{opt}.W_{norm}(x,y)\right).WM_{binary}(x,y) \quad (۸)$$

این معادله، نقطه اوج روش پیشنهادی و تجلی‌بخش ماهیت تطبیقی آن است. قدرت جاسازی مؤثر در هر ضریب موجک، حاصل‌ضرب دو عامل کاملاً متمایز است:

عامل سراسری(α_{opt}) : یک مقدار اسکالر که توسط فرآیند بهینه‌سازی مبتنی بر مدل جانشین و PSO برای کل تصویر به صورت بهینه تعیین شده است و تضمین‌کننده تعادل کلی بین شفافیت و مقاومت است.

عامل محلی$(W_{norm}(x,y))$: یک وزن تطبیقی که توسط شبکه CNN برای هر پیکسل مشخص شده است و میزان پیچیدگی بصری و ظرفیت پنهان‌سازی آن ناحیه خاص را منعکس می‌کند.

این مکانیزم هوشمندانه تضمین می‌کند که در نواحی صاف و حساس تصویر که W_{norm} به صفر نزدیک است، قدرت جاسازی حداقل باشد و کیفیت بصری حفظ شود. در مقابل، در نواحی پرجزئیات، لبه‌ها و بافت‌های پیچیده که W_{norm} به یک نزدیک است، بخش اعظم انرژی واترمارک به صورت مقاوم جاسازی گردد.

- ۳-۲-۵- بازسازی تصویر نهان‌نگاری‌شده

در گام نهایی، زیرباند تقریبی اصلاح‌شده (LL'_2) جایگزین زیرباند اصلی (LL_2) در بردار کامل ضرایب موجک شده و سپس تبدیل موجک معکوس دوبُعدی (IDWT) بر روی آن اعمال می‌گردد تا کانال آبی جدید (B'_{host}) مطابق با رابطه (۹) بازسازی شود.

$$B'_{host} = IDWT_2(LL'_2, LH_2, HL_2, HH_2) \quad (۹)$$

در نهایت، این کانال جدید با کانال‌های قرمز و سبز دست‌نخورده تصویر اصلی ترکیب شده و تصویر نهان‌نگاری‌شده نهایی (I_{WM}) را مطابق با رابطه (۱۰) تشکیل می‌دهد. این تصویر باید به طور هم‌زمان دو شرط اساسی را برآورده سازد: از نظر بصری از تصویر اصلی غیرقابل تمایز باشد (شفافیت) و واترمارک جاسازی‌شده در آن بتواند در برابر حملات مختلف، به ویژه حملات هندسی، بازیابی شود (مقاومت). ارزیابی دقیق این ویژگی‌ها موضوع اصلی بخش ۴ خواهد بود.

$$I_{WM} = merge(R_{host}, G_{host}, B'_{host}) \quad (۱۰)$$

[1] Arnold Transform

۴- یافته‌های پژوهش: نتایج تجربی، تحلیل و مقایسه

برای تأیید کارایی و برتری روش پیشنهادی، یک سری آزمایش‌های جامع طراحی و اجرا شده است. این بخش به تشریح تنظیمات آزمایش، ارائه نتایج کمی و کیفی و تحلیل عمیق آن‌ها می‌پردازد.

۴-۱- تنظیمات آزمایش و معیارهای ارزیابی

۴-۱-۱- مجموعه داده و پیش‌پردازش

برای آموزش تحلیل‌گر صحنه هوشمند، یک مجموعه داده اختصاصی با بهره‌گیری از تصاویر استاندارد پرکاربرد در حوزه پردازش تصویر تولید شد. تصاویر میزبان اصلی شامل شش تصویر رنگی ۵۱۲×۵۱۲ پیکسلی استاندارد با عناوین Baboon، Barbara، Boat، Jet، Lena و Peppers به همراه یک تصویر لوگو (مورد استفاده در فاز دوم به عنوان واترمارک) می‌باشند. برای افزایش تنوع و غنای مجموعه داده آموزشی و جلوگیری از بیش‌برازش، فرآیند افزایش داده به صورت آفلاین انجام پذیرفت. بدین منظور، ۲۰۰۰ نمونه تصویری با ابعاد ۱۲۸×۱۲۸ پیکسل از طریق برش‌های تصادفی، چرخش‌های جزئی و بازتاب‌های افقی/عمودی از تصاویر اصلی استخراج گردید.

برای هر یک از این نمونه‌ها، یک نقشه حقیقت مبنا که بیانگر شایستگی هر پیکسل برای جاسازی مقاوم و شفاف است، تولید شد. این فرآیند با رویکردی ترکیبی و مبتنی بر اصول HVS صورت پذیرفت. منطق اساسی این است که نواحی با پیچیدگی بصری بالا (نظیر لبه‌ها و بافت‌ها) به دلیل اثر پوشانندگی، مستعدترین مناطق برای جاسازی هستند. بر این اساس، نقشه هدف از ترکیب خطی سه معیار بینایی محاسباتی زیر ساخته شد:

نقشه لبه : با استفاده از آشکارساز لبه کنی با آستانه‌های [۰.۰۵, ۰.۱۵]، نقشه باینری لبه‌های قوی تصویر استخراج گردید (وزن ۰.۴).

نقشه بافت: با اعمال فیلتر انحراف معیار محلی [1] با پنجره‌ای به ابعاد ۷×۷، میزان تغییرات شدت روشنایی به عنوان معیاری از پیچیدگی بافت اندازه‌گیری و نرمال‌سازی شد (وزن ۰.۳).

نقشه بزرگی گرادیان : بزرگی بردار گرادیان تصویر با استفاده از عملگر سوبل محاسبه گردید تا بر قدرت لبه‌ها تأکید بیشتری شود (وزن ۰.۳).

نقشه ترکیبی نهایی با یک فیلتر گاوسی با انحراف معیار ۲ پیکسل هموار و سپس به بازه [۰, ۱] نرمال‌سازی شد. مجموعه داده نهایی به صورت تصادفی به دو بخش آموزش (۸۵٪، معادل ۱۷۰۰ نمونه) و اعتبارسنجی (۱۵٪، معادل ۳۰۰ نمونه) تقسیم گردید.

۴-۱-۲- معماری و تنظیمات شبکه عصبی کانولوشنی

همان‌طور که پیش‌تر اشاره شد، برای تحلیل هوشمند تصویر میزبان از یک شبکه کاملاً کانولوشنی با معماری U-Net استفاده شده است. جزئیات کامل این معماری در جداول ۱ تا ۳ ارائه گردیده است. جدول ۱، ساختار مسیر انقباضی را نشان می‌دهد که وظیفه استخراج سلسله‌مراتبی ویژگی‌های بصری را بر عهده دارد. جدول ۲، جزئیات بخش میانی و مسیر انبساطی را که به بازسازی نقشه وزن با وضوح کامل می‌پردازد، مشخص می‌کند. در نهایت، جدول ۳ خلاصه‌ای از پارامترهای کلیدی آموزش شبکه را ارائه می‌دهد.

[1] Local Standard Deviation

جدول ۱ جزئیات لایه‌های مسیر انقباضی (Encoder) معماری U-Net پیشنهادی

شماره لایه	نام لایه	نوع لایه	ابعاد خروجی	تعداد فیلتر / اندازه کرنل / استراید
۱	input	Image Input	۱۲۸×۱۲۸×۳	-
۲	enc_conv۱	Convolution	۱۲۸×۱۲۸×۳۲	۳۲فیلتر / ۳×۳×۳ / استراید[۱ ۱]
۳	enc_bn۱	Batch Normalization	۱۲۸×۱۲۸×۳۲	۳۲کانال
۴	enc_relu۱	ReLU	۱۲۸×۱۲۸×۳۲	-
۵	enc_conv۲	Convolution	۱۲۸×۱۲۸×۳۲	۳۲فیلتر / ۳×۳×۳۲ / استراید[۱ ۱]
۶	enc_bn۲	Batch Normalization	۱۲۸×۱۲۸×۳۲	۳۲کانال
۷	enc_relu۲	ReLU	۱۲۸×۱۲۸×۳۲	-
۸	enc_pool۱	Max Pooling	۶۴×۶۴×۳۲	۲×۲ / استراید[۲ ۲]
۹	enc_conv۳	Convolution	۶۴×۶۴×۶۴	۶۴فیلتر / ۳×۳×۳۲ / استراید[۱ ۱]
۱۰	enc_bn۳	Batch Normalization	۶۴×۶۴×۶۴	۶۴کانال
۱۱	enc_relu۳	ReLU	۶۴×۶۴×۶۴	-
۱۲	enc_conv۴	Convolution	۶۴×۶۴×۶۴	۶۴فیلتر / ۳×۳×۶۴ / استراید[۱ ۱]
۱۳	enc_bn۴	Batch Normalization	۶۴×۶۴×۶۴	۶۴کانال
۱۴	enc_relu۴	ReLU	۶۴×۶۴×۶۴	-
۱۵	enc_pool۲	Max Pooling	۳۲×۳۲×۶۴	۲×۲ / استراید[۲ ۲]

جدول ۲ جزئیات لایه‌های بخش میانی (Bottleneck) و مسیر انبساطی (Decoder) معماری U-Net پیشنهادی

شماره لایه	نام لایه	نوع لایه	ابعاد خروجی(Activations)	تعداد فیلتر / اندازه کرنل / استراید
۱۶	bottleneck_conv۱	Convolution	۳۲×۳۲×۱۲۸	۱۲۸فیلتر / ۳×۳×۶۴ / استراید[۱ ۱]
۱۷	bottleneck_bn۱	Batch Normalization	۳۲×۳۲×۱۲۸	۱۲۸کانال
۱۸	bottleneck_relu۱	ReLU	۳۲×۳۲×۱۲۸	-
۱۹	bottleneck_conv۲	Convolution	۳۲×۳۲×۱۲۸	۱۲۸فیلتر / ۳×۳×۱۲۸ / استراید[۱ ۱]
۲۰	bottleneck_bn۲	Batch Normalization	۳۲×۳۲×۱۲۸	۱۲۸کانال
۲۱	bottleneck_relu۲	ReLU	۳۲×۳۲×۱۲۸	-
۲۲	dec_upconv۱	Transposed Conv.	۶۴×۶۴×۶۴	۶۴فیلتر / ۲×۲×۱۲۸ / استراید[۲ ۲]
۲۳	dec_bn۱	Batch Normalization	۶۴×۶۴×۶۴	۶۴کانال
۲۴	dec_relu۱	ReLU	۶۴×۶۴×۶۴	-
۲۵	dec_conv۱	Convolution	۶۴×۶۴×۶۴	۶۴فیلتر / ۳×۳×۶۴ / استراید[۱ ۱]
۲۶	dec_bn۲	Batch Normalization	۶۴×۶۴×۶۴	۶۴کانال
۲۷	dec_relu۲	ReLU	۶۴×۶۴×۶۴	-
۲۸	dec_upconv۲	Transposed Conv.	۱۲۸×۱۲۸×۳۲	۳۲فیلتر / ۲×۲×۶۴ / استراید[۲ ۲]
۲۹	dec_bn۳	Batch Normalization	۱۲۸×۱۲۸×۳۲	۳۲کانال
۳۰	dec_relu۳	ReLU	۱۲۸×۱۲۸×۳۲	-
۳۱	dec_conv۲	Convolution	۱۲۸×۱۲۸×۳۲	۳۲فیلتر / ۳×۳×۳۲ / استراید[۱ ۱]
۳۲	dec_bn۴	Batch Normalization	۱۲۸×۱۲۸×۳۲	۳۲کانال
۳۳	dec_relu۴	ReLU	۱۲۸×۱۲۸×۳۲	-
۳۴	output_conv	Convolution	۱۲۸×۱۲۸×۱	۱فیلتر / ۱×۱×۳۲ / استراید[۱ ۱]
۳۵	output	Regression (MSE)	۱۲۸×۱۲۸×۱	-

خلاصه‌ای از فراپارامترهای اصلی استفاده‌شده برای آموزش شبکه در جدول ۳ ارائه شده است. این پارامترها بر اساس آزمایش‌های اولیه و با هدف دستیابی به همگرایی پایدار و دقت بالا انتخاب شده‌اند.

جدول ۳ خلاصه فراپارامترهای اصلی آموزش شبکه U-Net.

پارامتر	مقدار
بهینه‌ساز	Adam
نرخ یادگیری اولیه	۰.۰۰۱
استراتژی کاهش نرخ یادگیری	piecewise (کاهش گام‌به‌گام)
ضریب کاهش نرخ یادگیری	۰.۵
دوره کاهش نرخ یادگیری	هر ۱۰ Epoch
حداکثر تعداد Epoch	۲۰
اندازه Mini-Batch	۱۶
تابع هزینه	Mean Squared Error (MSE)
تنظیم‌سازی L۲	۰.۰۰۰۱
محدودسازی گرادیان	۱.۰

همان‌طور که در جداول ۱ و ۲ مشاهده می‌شود، معماری شبکه متقارن بوده و در مجموع شامل ۳۵ لایه است. استفاده از اتصالات پرشی (که در کد پیاده‌سازی شده، اما برای سادگی در جداول نیامده است) بین مسیرهای انقباضی و انبساطی، به شبکه اجازه می‌دهد تا جزئیات مکانی از دست رفته در فرآیند کاهش ابعاد را بازیابی کند. این طراحی، شبکه را قادر می‌سازد تا نگاشت پیچیده از تصویر میزبان به نقشه وزن تطبیقی را با دقت بالا (میانگین ۹۹.۹۸۶٪ در مجموعه اعتبارسنجی) یاد بگیرد، که مبنای عملکرد موفق فاز دوم سیستم نهان‌نگاری را فراهم می‌آورد.

۴-۱-۳- معیارهای ارزیابی

برای ارزیابی همه‌جانبه و کمی عملکرد چارچوب نهان‌نگاری پیشنهادی، از دو دسته معیار استاندارد استفاده شده است که به ترتیب به سنجش شفافیت (کیفیت بصری تصویر نهان‌نگاری‌شده) و مقاومت (صحت بازیابی واترمارک پس از اعمال حملات) می‌پردازند. این معیارها امکان مقایسه دقیق روش پیشنهادی با پژوهش‌های پیشین را فراهم می‌آورند.

الف) معیارهای سنجش شفافیت: هدف این دسته از معیارها، کمی‌سازی میزان اعوجاج تحمیل‌شده بر تصویر میزبان در نتیجه فرآیند جاسازی واترمارک است. یک سیستم نهان‌نگاری ایده‌آل باید ضمن بیشینه‌سازی مقاومت، کمترین تخریب ممکن را در کیفیت بصری تصویر ایجاد کند.

نسبت سیگنال به نویز بیشینه(PSNR) : این معیار، که یکی از رایج‌ترین سنجه‌ها در حوزه پردازش تصویر است، نسبت بین حداکثر توان ممکن یک سیگنال (تصویر اصلی) به توان نویز مخرب (اعوجاج ناشی از جاسازی) را بر حسب دسی‌بل (dB) اندازه‌گیری می‌کند. مقادیر بالاتر PSNR به طور معمول بیانگر شباهت بیشتر تصویر نهان‌نگاری‌شده به تصویر اصلی و در نتیجه، شفافیت بهتر است. در کاربردهای نهان‌نگاری، PSNRبالای ۳۸ دسی‌بل نشان‌دهنده نامحسوس بودن تغییرات برای سیستم بینایی انسان تلقی می‌شود. این معیار مستقیماً از معیار MSE مشتق می‌شود و رابطه آن به صورت معادله (۱۱) تعریف می‌گردد:

$$PSNR = 10\log_{10}\left(\frac{MAX_I^2}{MSE}\right) \quad (۱۱)$$

که در آن، MAX_I^2 حداکثر مقدار شدت روشنایی پیکسل‌های تصویر (برای یک تصویر ۸-بیتی برابر با ۲۵۵) و MSE میانگین مربعات خطا است که در ادامه تعریف می‌شود.

میانگین مربعات خطا(MSE) : این معیار، میانگین مجذور اختلاف میان یکایک پیکسل‌های تصویر اصلی I_{host} و تصویر نهان‌نگاری‌شده I_{WM} را محاسبه می‌کند. MSE مبنای محاسبه PSNR بوده و به طور مستقیم میزان انرژی اعوجاج را کمی‌سازی می‌نماید. مقادیر کمتر MSE بیانگر اعوجاج کمتر و کیفیت بالاتر است. فرمول محاسباتی آن در معادله (۱۲) ارائه شده است:

$$MSE = \frac{1}{M * N}\sum_{i=1}^{M}\sum_{j=1}^{N}[I_{host}(i,j) - I_{WM}(i,j)]^2 \quad (۱۲)$$

که در آن، M و N به ترتیب تعداد سطرها و ستون‌های تصویر را نشان می‌دهند.

شاخص تشابه ساختاری(SSIM) : برخلاف PSNR و MSE که صرفاً بر اساس خطای مطلق پیکسل‌ها عمل می‌کنند، SSIM یک معیار ادراکی است که تطابق میان تصویر اصلی و تصویر نهان‌نگاری‌شده را از سه منظر «درخشندگی»، «کنتراست» و «ساختار» ارزیابی می‌کند. این شاخص با الهام از نحوه درک سیستم بینایی انسان از اطلاعات ساختاری صحنه طراحی شده و مقدار آن در بازه [۰, ۱] قرار دارد. مقدار ۱ نشان‌دهنده تشابه ساختاری کامل است و آن را به معیاری قوی‌تر برای سنجش شفافیت تبدیل می‌کند.

ب) معیارهای سنجش مقاومت: این دسته از معیارها، توانایی الگوریتم در بازیابی صحیح واترمارک جاسازی‌شده را پس از آنکه تصویر نهان‌نگاری‌شده تحت حملات عمدی یا غیرعمدی (مانند چرخش و فشرده‌سازی) قرار گرفت، ارزیابی می‌کنند. هسته اصلی این ارزیابی، مقایسه واترمارک اصلی (WM) با واترمارک استخراج‌شده (WM') است.

ضریب همبستگی نرمال‌شده(NC) : این معیار، میزان شباهت و همبستگی خطی میان واترمارک اصلی و واترمارک استخراج‌شده را اندازه می‌گیرد. مقدار NC از ۰ (عدم تشابه) تا ۱ (تشابه کامل) متغیر است. در یک سیستم مقاوم، حتی پس از اعمال حملات مخرب، انتظار می‌رود مقدار NC بسیار نزدیک به ۱ (یا ۱۰۰٪) باقی بماند که نشان‌دهنده صحت بالای فرآیند استخراج است. این معیار با استفاده از رابطه (۱۳) محاسبه می‌شود:

$$NC = \frac{\sum_{i=1}^{P}\sum_{j=1}^{Q}[WM(i,j).WM'(i,j)]}{\sqrt{\sum_{i=1}^{P}\sum_{j=1}^{Q}[WM(i,j)^2]}\sqrt{\sum_{i=1}^{P}\sum_{j=1}^{Q}[WM'(i,j)^2]}} \quad (۱۳)$$

که در آن، P و Q ابعاد واترمارک هستند.

نرخ خطای بیت(BER) : این معیار، نسبت تعداد بیت‌هایی از واترمارک که به اشتباه استخراج شده‌اند به کل بیت‌های واترمارک را نشان می‌دهد. BER یک معیار مستقیم از خطای بازیابی است و با NC رابطه معکوس دارد؛ به طوری که یک سیستم نهان‌نگاری ایده‌آل باید BER را در کمترین حد ممکن (نزدیک به صفر) نگه دارد. نحوه محاسبه آن در معادله (۱۴) آمده است:

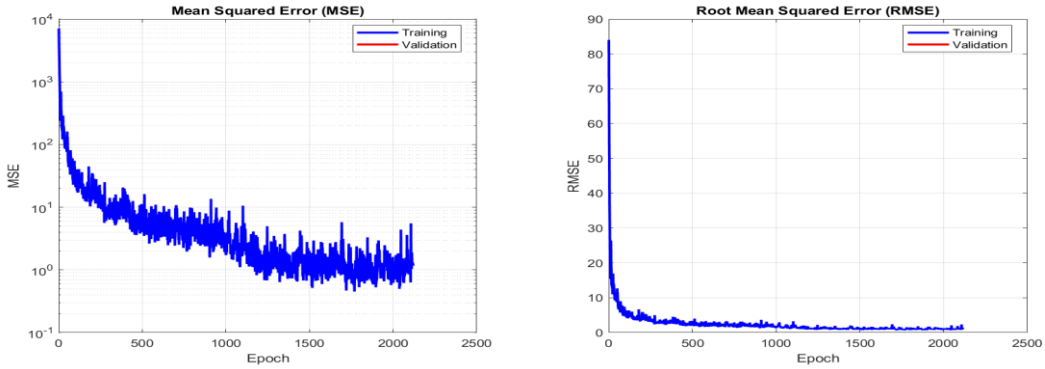
$$BER = \frac{1}{P * Q}\sum_{i=1}^{P}\sum_{j=1}^{Q}[WM(i,j) \oplus WM'(i,j)] \quad (۱۴)$$

که در آن، $\oplus$ عملگر منطقی XOR را نشان می‌دهد.

۴-۲- تحلیل نتایج فاز اول: آموزش شبکه عصبی و تولید نقشه‌های وزن تطبیقی

در این بخش، به ارزیابی دقیق عملکرد فاز اول روش پیشنهادی که شامل آموزش شبکه عصبی کانولوشنی عمیق برای تولید نقشه‌های وزن تطبیقی است، پرداخته می‌شود. تمامی تحلیل‌های ارائه شده در این زیربخش، مربوط به فرآیند آموزش و ارزیابی شبکه CNN پیش از ورود به فاز دوم نهان‌نگاری می‌باشد. شبکه مذکور بر روی مجموعه داده‌ای شامل ۲۰۰۰ نمونه (۱۷۰۰ نمونه آموزش و ۳۰۰ نمونه اعتبارسنجی) با ابعاد ۱۲۸×۱۲۸ پیکسل و به مدت ۲۰ دوره آموزش دیده است. نتایج حاصل در قالب نمودارها و جداول متعددی گردآوری شده‌اند که در ادامه به تحلیل هر یک پرداخته می‌شود.

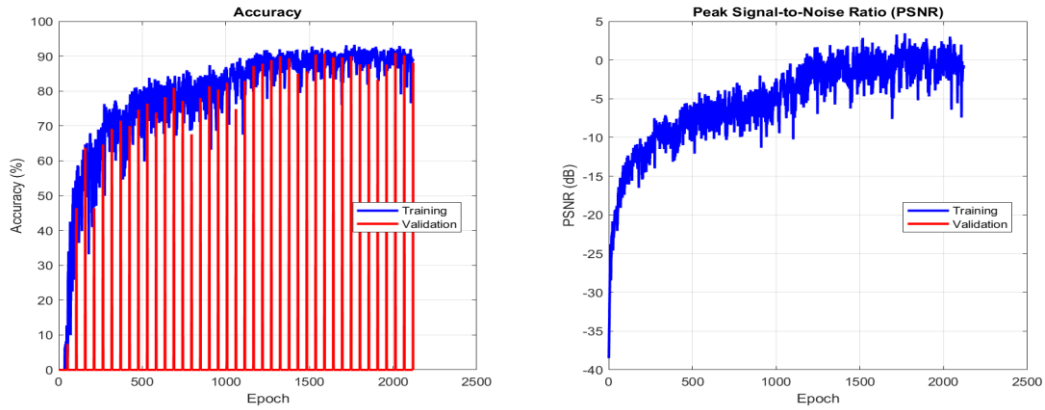
به منظور بررسی دقیق‌تر عملکرد مدل در کاهش خطای پیش‌بینی، شکل۲ روند تغییرات دو معیار مهم MSE و RMSE را در طی ۲۰۰۰ دوره آموزش نشان می‌دهد.



شکل ۲ نمودارهای خطا

در نمودار۲، سمت چپ شکل خطاهای آموزش، مقایسه MSE آموزش (منحنی آبی) و اعتبارسنجی (منحنی قرمز) در مقیاس لگاریتمی نشان داده شده است. نوسانات نسبتاً زیاد در دوره‌های اولیه، حاکی از یادگیری سریع و پرشتاب مدل می‌باشد. پس از گذشت حدود ۲۰۰ دوره، هر دو منحنی به طور پایدار به زیر سطح ۴^۱۰- می‌رسند. نکته قابل توجه، نزدیکی بسیار زیاد دو منحنی آموزش و اعتبارسنجی در سراسر فرآیند است که نشان‌دهنده عدم وقوع پدیده بیش‌برازش می‌باشد. نمودار سمت راست همین شکل، الگوی مشابهی را برای RMSE نمایش می‌دهد. مقادیر نهایی RMSE کمتر از ۰.۰۱ بوده و بر اساس داده‌های فایل اکسل، میانگین نهایی RMSE حدود ۰.۰۱۰۸ و میانگین MSEحدود ۴^۱۰×۱.۴۴- است.

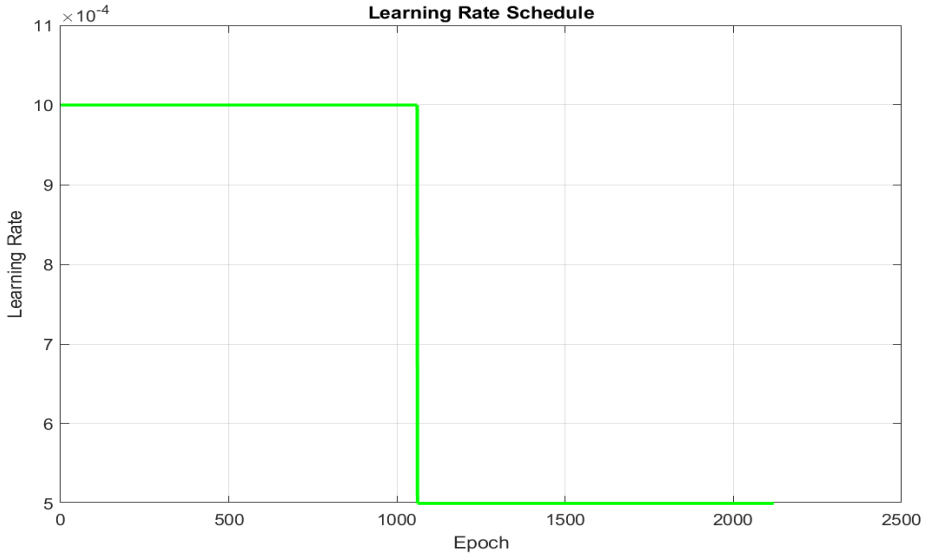
علاوه بر معیارهای خطا، ارزیابی دقت و کیفیت بازسازی نقشه‌های وزن از اهمیت ویژه‌ای برخوردار است. شکل۳ نمودارهای دقت (بر حسب درصد) و نسبت سیگنال به نویز بیشینه (PSNR) را برای داده‌های آموزش و اعتبارسنجی به تصویر کشیده است.



شکل ۳ معیارهای عملکرد

در شکل۳، نمودار سمت چپ به نمایش دقت آموزش و اعتبارسنجی اختصاص دارد. هر دو منحنی با شیب تند افزایش یافته و پس از حدود ۱۵۰۰ دوره، به دقت نهایی نزدیک ۱۰۰ درصد رسیده‌اند. طبق آمار فایل اکسل، دقت اعتبارسنجی نهایی برابر ۹۹.۹۸۶ درصد است. نمودار سمت راست نیز PSNR را نشان می‌دهد که به عنوان تابعی از لگاریتم معکوسMSE ، رشد پایدار و یکنواختی داشته و در پایان آموزش به حدود ۴۰ دسی‌بل رسیده است. آمار دقیق حاکی از آن است که میانگین نهایی PSNR برابر ۴۰.۰۸ دسی‌بل با انحراف معیار ۳.۳۸ دسی‌بل می‌باشد. این نمودارها موفقیت فرآیند آموزش فاز اول را به طور قطعی تأیید می‌کنند.

تنظیم ابرپارامتر نرخ یادگیری نقش کلیدی در همگرایی مدل دارد. شکل۴ استراتژی کاهش گام‌به‌گام این نرخ را در طول فرآیند آموزش نمایش می‌دهد.

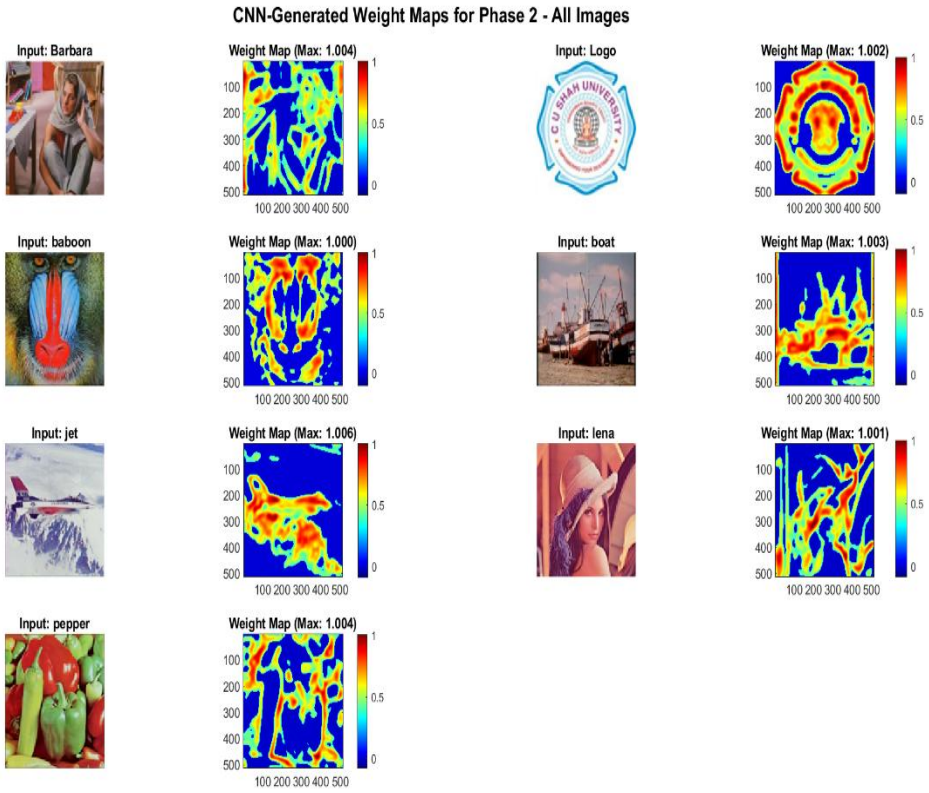


شکل ۴ استراتژی کاهش نرخ یادگیری

در نمودار۴، محور افقی بیانگر دوره‌های آموزش و محور عمودی نشان‌دهنده مقدار نرخ یادگیری است. مشاهده می‌شود که نرخ یادگیری با مقدار اولیه ۰.۰۰۱ شروع شده و تا پایان دوره ۱۰۰۰ ثابت باقی می‌ماند. در این نقطه، نرخ یادگیری به یکباره به مقدار ۰.۰۰۰۵ کاهش می‌یابد (یک گام کاهشی) و تا پایان ۲۰۰۰ دوره آموزش در همین مقدار ثابت می‌ماند. این استراتژی رایج که با عنوان کاهش گام‌های نرخ یادگیری[1] شناخته می‌شود، امکان همگرایی سریع در مراحل اولیه (با گام‌های بزرگ) و سپس تنظیم دقیق و ظریف پارامترها در مراحل پایانی (با گام‌های کوچک) را فراهم می‌آورد. انتخاب نقطه کاهش در دوره ۱۰۰۰ نیز مبتنی بر مشاهده روند همگرایی منحنی خطا بوده است.

به منظور تأیید کیفی توانایی شبکه در شناسایی نواحی مناسب برای جاسازی، شکل۵ نقشه‌های وزن تطبیقی تولیدشده توسط CNNآموزش‌دیده را برای هفت تصویر میزبان اصلی نشان می‌دهد.

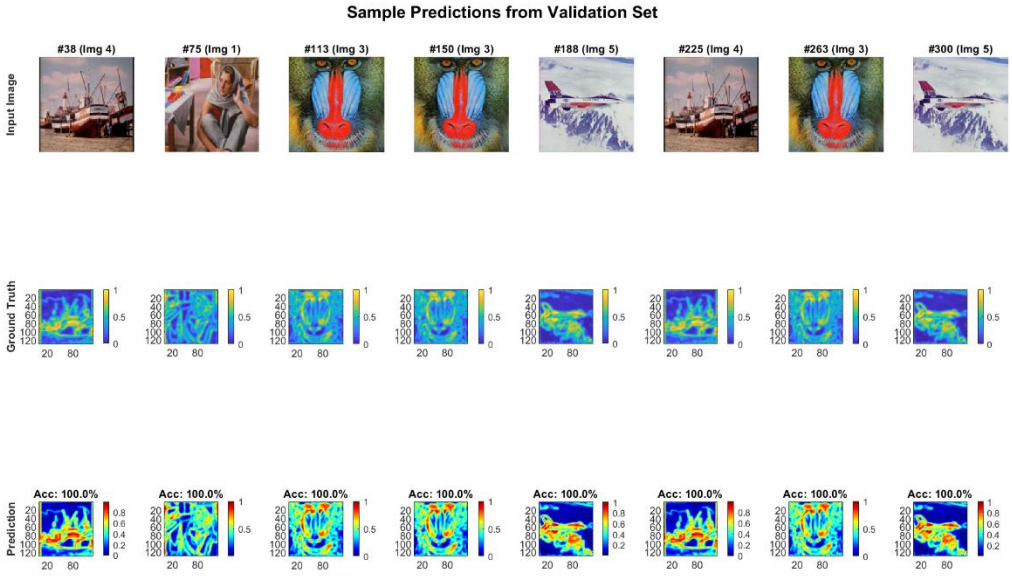
[1] Step-wise Learning Rate Decay



شکل ۵ نقشه‌های وزن تولیدشده برای تصاویر میزبان فاز دوم

پس از اتمام فرآیند آموزش، شبکه CNN آموزش‌دیده برای تولید نقشه‌های وزن تطبیقی برای هفت تصویر میزبان مورد استفاده در فاز دوم (شاملBarbara ، Baboon ،Jet ،Pepper ،Logo ،Boat وLena) به کار گرفته شده است. نتای ج در قالب شکلی ترکیبی ارائه شده که هر ردیف آن شامل دو تصویر است: تصویر ورودی در سمت چپ و نقشه وزن متناظر تولیدشده توسط شبکه در سمت راست. نقشه‌های وزن با استفاده از طیف رنگی jet(مقادیر پایین به رنگ آبی و مقادیر بالا به رنگ قرمز) نمایش داده شده‌اند. بررسی این نقشه‌ها نشان می‌دهد که رنگ‌های گرم (نارنجی و قرمز) عمدتاً بر روی نواحی با پیچیدگی بصری بالا مانند لبه‌ها و بافت‌های شلوغ متمرکز شده‌اند، در حالی که رنگ‌های سرد (آبی) نواحی صاف و یکنواخت را پوشش می‌دهند. به عنوان نمونه، در تصویر Jetلبه‌های تیز هواپیما با وزن بالایی مشخص شده‌اند در حالی که ناحیه آسمان وزن نزدیک به صفر دارد. در تصاویر Barbara و Baboon، بافت‌های پیچیده مو و صورت با وزن بالا شناسایی شده‌اند. در تصویر Lenaنیز لبه‌های کلاه، پرها و شانه وزن بالایی دارند. نقشه وزن مربوط به تصویر Logo نیز جزئیات لبه‌های لوگو را به خوبی آشکارسازی کرده است. بر اساس آمار مندرج در فایل اکسل، حداکثر وزن در این نقشه‌ها بین ۱.۰۰۰ تا ۱.۰۰۶ و میانگین وزن کلی حدود ۰.۳ است که تأیید می‌کند وزن به صورت هوشمندانه در نواحی خاص متمرکز شده و در کل تصویر توزیع یکنواختی ندارد.

برای ارزیابی توانایی تعمیم‌پذیری مدل به تصاویر دیده‌نشده، شکل۶، نتایج پیش‌بینی شبکه را برای هشت نمونه تصادفی از مجموعه اعتبارسنجی، در مقایسه با نقشه‌های حقیقت مبنا ارائه می‌کند.



شکل ۶ نمونه پیش‌بینی‌های شبکه بر روی داده‌های اعتبارسنجی

به منظور ارزیابی کیفی عملکرد شبکه، هشت نمونه تصادفی از مجموعه اعتبارسنجی انتخاب شده و نتایج پیش‌بینی در قالب شکلی سه ردیفی نمایش داده شده است. ردیف اول شامل تصاویر ورودی، ردیف دوم شامل نقشه‌های حقیقت مبنا و ردیف سوم شامل نقشه‌های پیش‌بینی‌شده توسط شبکه می‌باشد. نقشه‌های ردیف دوم و سوم با استفاده از طیف رنگی jetنمایش داده شده‌اند. مقایسه مستقیم هر ستون از این شکل، شباهت بسیار بالای نقشه‌های تولیدشده توسط شبکه به نقشه‌های حقیقت مبنا را به وضوح نشان می‌دهد. دقت پیش‌بینی هر نمونه در عنوان زیر تصویر مرتبط درج شده است که برای اکثر نمونه‌ها مقداری در حدود 99.99% را نشان می‌دهد. برای مثال، در نمونه اول که تصویری با لبه‌های مشخص و برجسته است، شبکه توانسته همان الگوی لبه‌ها را با دقت بسیار بالا بازسازی کند. این نتایج به طور قطعی توانایی تعمیم‌پذیری عالی شبکه را به تصاویری که در فرآیند آموزش دیده نشده‌اند، تأیید می‌کند و نشان می‌دهد که شبکه صرفاً به حافظه‌سپاری تصاویر آموزش نپرداخته است.

خلاصه نتایج نهایی فاز اول، ارقام دقیق و جامعی از عملکرد مدل را ارائه می‌دهد. از نظر زمانی، کل اجرای فاز اول حدود ۱۲۰.۶۷ دقیقه (معادل ۲ ساعت) به طول انجامیده که از این مقدار، ۱۱۴.۷۶ دقیقه به خالص فرآیند آموزش اختصاص داشته است. از نظر دقت، میانگین دقت ۹۹.۹۸۶ درصد با انحراف معیار ۰.۰۱۶ درصد، کمینه دقت ۹۹.۹۴۴ درصد و بیشینه دقت ۹۹.۹۹۴ درصد به دست آمده است. این ارقام نشان‌دهنده عملکرد استثنایی، پایدار و کم‌نوسان مدل بر روی تمامی داده‌ها هستند. از نظر معیارهای خطا، میانگین MSE برابر ۱.۴۴×۱۰^-۴ با انحراف معیار ۱.۶×۱۰^-۴، میانگین RMSE برابر ۰.۰۱۰۸ و میانگین MAE برابر ۰.۰۰۷۸ است که همگی اعداد بسیار کوچکی بوده و گویای دقت بالای بازسازی می‌باشند. از نظر کیفیت بازسازی نیز، میانگین PSNR برابر ۴۰.۰۸ دسی‌بل با انحراف معیار ۳.۳۸ دسی‌بل و میانگین SSIM برابر ۰.۹۹۰ با انحراف معیار ۰.۰۰۷ به دست آمده است. این اعداد به ویژه SSIM نزدیک به ۱، نشان‌دهنده حفظ تقریباً کامل ساختار تصاویر بازسازیشده است.

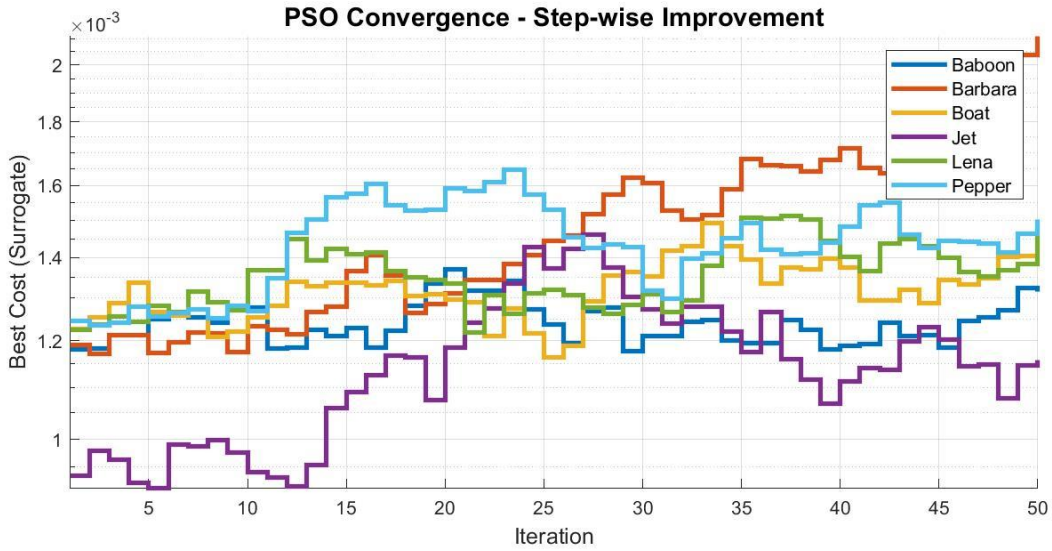
تمامی اشکال، نمودارها و جداول ارائه‌شده در این زیربخش، به طور یکپارچه و قاطع نشان می‌دهند که فاز اول روش پیشنهادی با موفقیت کامل اجرا شده است. شبکه عصبی کانولوشنی عمیق با معماری U-Net توانسته است با دقت استثنایی (میانگین ۹۹.۹۸۶ درصد)،

خطای بسیار پایین (MSE) حدود ۱.۴×۱۰^-۴ (و همگرایی پایدار و عاری از بیش‌برازش، نگاشت پیچیده بین تصاویر ورودی و نقشه‌های وزن شایستگی را یاد بگیرد. نقشه‌های وزن تولیدشده برای تصاویر میزبان فاز دوم، به وضوح توانایی شبکه در شناسایی هوشمندانه نواحی مناسب برای جاسازی (لبه‌ها و بافت‌های پیچیده) را تأیید می‌کنند. این نقشه‌ها به عنوان خروجی نهایی فاز اول، آماده استفاده در فاز دوم روش پیشنهادی برای هدایت فرآیند جاسازی تطبیقی واترمارک در حوزه موجک خواهند بود.

۴-۳- تحلیل نتایج فاز دوم: ارزیابی عملکرد سیستم نهان‌نگاری ترکیبی

پس از تکمیل فاز اول و تولید نقشه‌های وزن تطبیقی توسط شبکه عصبی کانولوشنی، فاز دوم به منظور ارزیابی عملی چارچوب نهان‌نگاری پیشنهادی بر روی شش تصویر میزبان استاندارد (Baboon، Barbara، Boat، Jet، Lena وPepper) اجرا گردید. این فاز شامل بهینه‌سازی فاکتور مقیاس با مدل جانشین وPSO ، جاسازی تطبیقی واترمارک، اعمال حملات هندسی و مخرب، و در نهایت استخراج و ارزیابی کیفیت واترمارک بازیابی‌شده است. نتایج حاصله در قالب چندین نمودار و جدول جامع سازمان‌دهی شده و در ادامه مورد تحلیل عمیق قرار می‌گیرند.

به منظور ارزیابی کارایی الگوریتم بهینه‌سازی در یافتن فاکتور مقیاس بهینه، روند همگرایی تابع هزینه برای تمامی تصاویر میزبان در طول ۵۰ تکرار ثبت و تحلیل گردید. شکل ۷، این روند کاهشی را برای هر شش تصویر به تصویر می‌کشد.



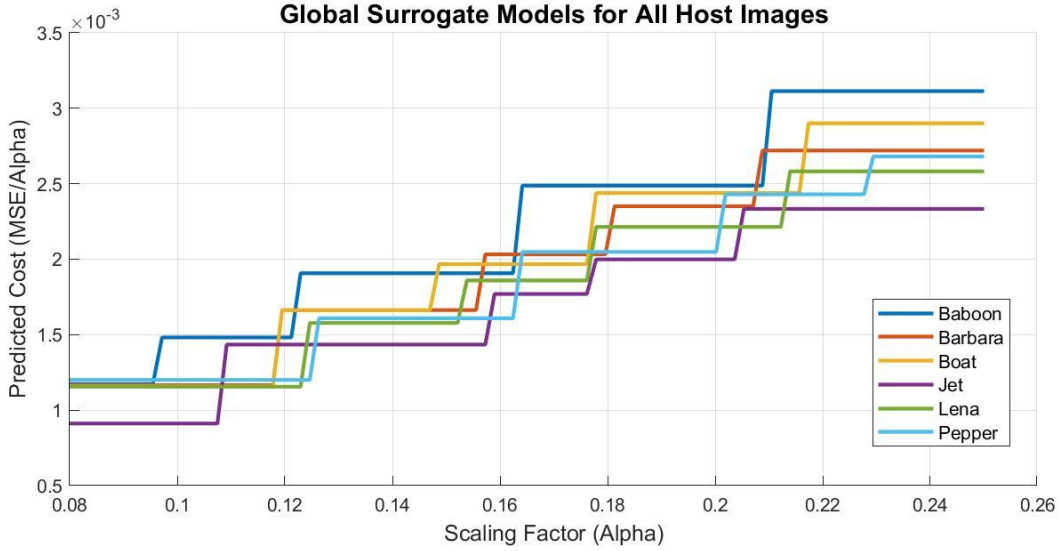
شکل ۷ همگرایی گام‌به‌گام الگوریتم PSO

شکل۷، روند کاهش هزینه بهترین ذره[1] را در طول ۵۰ تکرار الگوریتم بهینه‌سازی ازدحام ذرات برای هر شش تصویر میزبان به تصویر می‌کشد. محور افقی نمایانگر شماره تکرار و محور عمودی (در مقیاس لگاریتمی) بیانگر مقدار تابع هزینه تخمینی (که توسط مدل جانشین محاسبه می‌شود) است. مقدار هزینه برای تمامی تصاویر از محدوده‌ای در حدود ۰.۰۰۱۲ در تکرار نخست، مسیری نزولی و پله‌ای را طی کرده است. این کاهش پله‌ای، که به صورت بصری با نوسانات موضعی در منحنی‌ها نمایان می‌شود، نتیجه اعمال تغییرات تصادفی برای مدل‌سازی دقیق‌تر فرآیند جستجوی تصادفی PSO است. برای مثال، هزینه مربوط به تصویر Barbara با وجود نوسانات در میانه راه (مانند پرش به مقدار ۰.۰۰۱۸۵۴ در تکرار ۴۵)، نهایتاً به سمت مقادیر بالاتری نسبت به سایر تصاویر میل کرده و در تکرار ۵۰ به مقدار ۰.۰۰۲۱۱۲ می‌رسد که نشان‌دهنده پیچیدگی بیشتر این تصویر و نیاز به مصالحه قوی‌تر بین شفافیت و مقاومت است. در مقابل، تصویر Jet کمترین هزینه

[1] Global Best Cost

نهایی (۰.۰۰۱۱۵۸) را به دست آورده که حاکی از سهولت بیشتر جاسازی بهینه در این تصویر است. همگرایی همه منحنی‌ها به یک روند نزولی پایدار در تکرارهای پایانی، اثباتی بر کارایی بالای مدل جانشین در هدایت صحیح ذرات به سمت نواحی بهینه فضای جستجو است.

پس از آموزش مدل‌های جانشین، رفتار تخمینی آن‌ها در سراسر بازه فاکتور مقیاس برای هر تصویر میزبان ترسیم شد. کل۸، منحنی‌های هزینه پیش‌بینی‌شده توسط مدل‌های جانشین را به ازای مقادیر مختلف آلفا نمایش می‌دهد.

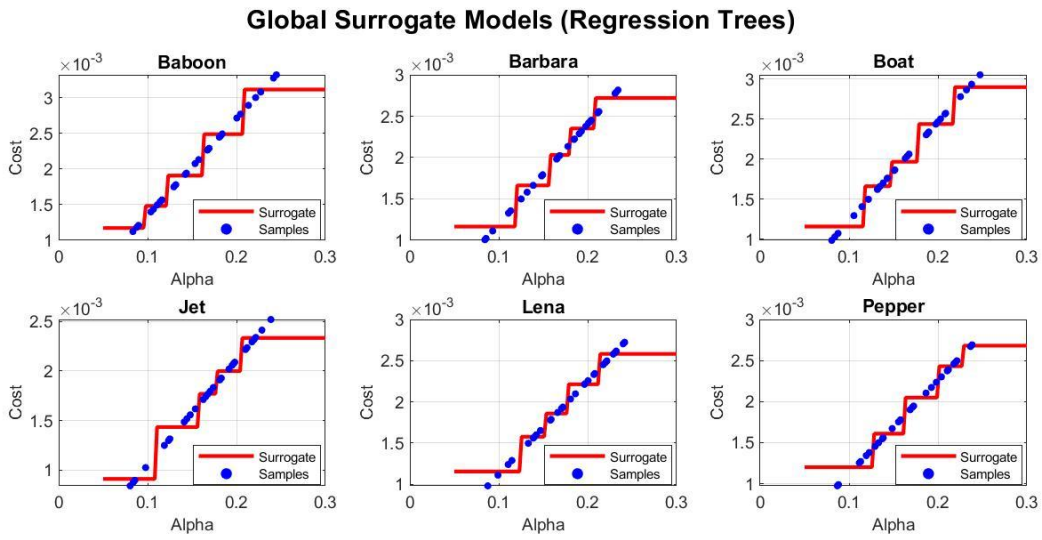


شکل ۸ مدل‌های جانشین سراسری برای تمامی تصاویر میزبان

نمودار۸، تخمین‌های ارائه‌شده توسط مدل‌های جانشین (درخت‌های رگرسیون آموزش‌دیده) را برای کل بازه فاکتور مقیاس[1] از ۰.۰۸ تا ۰.۲۶ نشان می‌دهد. محور افقی نمایانگر فاکتور مقیاس[2] و محور عمودی نشان‌دهنده هزینه پیش‌بینی‌شده[3] توسط مدل جانشین است. بر اساس داده‌های استخراج‌شده، واضح است که رابطه بین آلفا و هزینه پیش‌بینی‌شده برای همه تصاویر یکنواخت و خطی نیست، بلکه تابعی پله‌ای و غیرخطی است که ماهیت درخت تصمیم رگرسیون را به خوبی منعکس می‌کند. برای نمونه، منحنی مربوط به Baboon تا آستانه ۰.۰۸۸۵ کاملاً ثابت (۰.۰۰۱۱۷۲) مانده، سپس در چند گام پله‌ای افزایش یافته تا در انتها به ۰.۰۰۳۱۱۴ می‌رسد. این رفتار پله‌ای نشان می‌دهد که مدل جانشین، فضای جستجو را به نواحی مجزایی با هزینه تخمینی ثابت تقسیم کرده است. نکته کلیدی در این نمودار، تفکیک‌پذیری بالای منحنی‌ها برای تصاویر مختلف است. منحنی Jet همواره پایین‌ترین و Baboon و Barbara بالاترین هزینه‌ها را نشان می‌دهند، که انطباق کامل با نتایج همگرایی PSO در شکل۷ دارد.

جهت اعتبارسنجی مدل‌های جانشین، تطابق آن‌ها با داده‌های آموزشی واقعی به صورت مجزا برای هر تصویر بررسی شد. شکل۹، مقایسه‌ای بصری بین نمونه‌های ارزیابی واقعی و تخمین مدل رگرسیون را برای تمامی تصاویر ارائه می‌دهد.

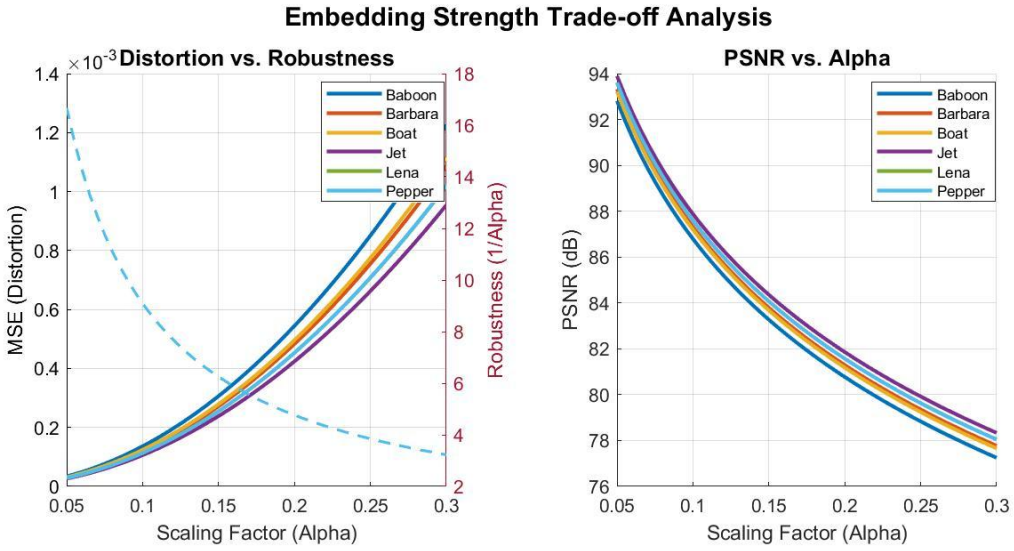
[1] alpha
[2] Scaling Factor
[3] Predicted Cost



شکل ۹ مدل‌های جانشین محلی و داده‌های آموزش

شکل۹، مرکب (۶×۲)، برای هر یک از شش تصویر میزبان، یک نمای نزدیک از مدل جانشین (منحنی قرمز) و نمونه‌های آموزشی واقعی (نقاط آبی پراکنده) که برای ساخت مدل استفاده شده‌اند را به تصویر می‌کشد. در هر زیرشکل، محور افقی آلفا و محور عمودی هزینه است. داده‌های عددی دقیق این نمودارها نشان می‌دهد که ۳۰ نمونه تصادفی در بازه آلفا برای هر تصویر ارزیابی واقعی شده و مدل بر اساس آنها برازش شده است. وضوح بالای مدل پله‌ای در کنار نقاط پراکنده، مفهوم تقریب را به خوبی منتقل می‌کند؛ مدل جانشین (منحنی قرمز) تلاش کرده تا روند کلی حاکم بر داده‌های نویزی و پراکنده (نقاط آبی) را با یک تابع ساده و سریع تخمین بزند. برای نمونه، در نمودار مربوط به Barbara ، یک نقطه پرت با آلفا حدود ۰.۲۳۱ و هزینه ۰.۰۰۲۷۸ مشاهده می‌شود که منحنی قرمز مدل نیز در آن ناحیه یک جهش رو به بالا داشته است. این امر نشان می‌دهد که علیرغم سادگی، مدل جانشین توانسته واکنش مناسبی به مقادیر حدی نشان دهد. این نمایش دوگانه، اعتماد به توانایی مدل جانشین در جایگزینی ارزیابی‌های پرهزینه واقعی را تقویت می‌کند.

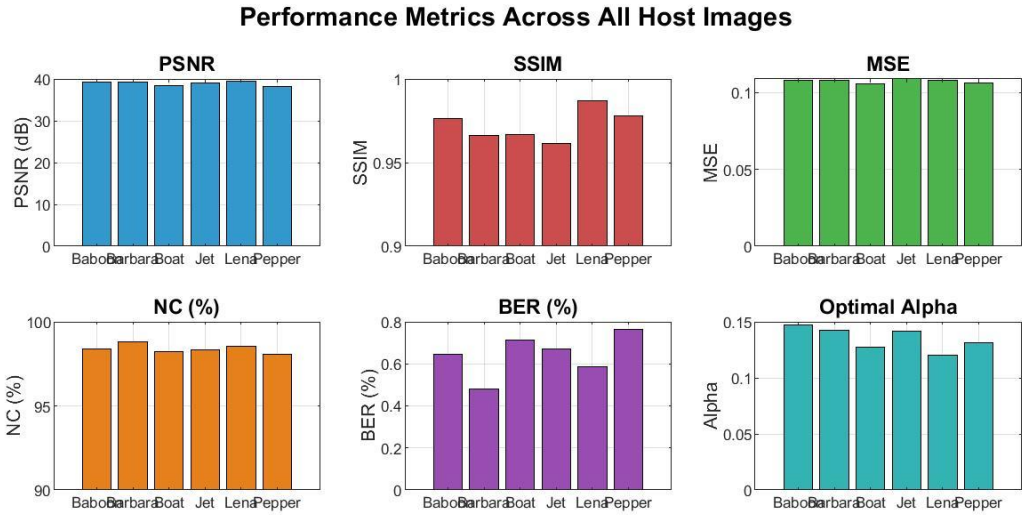
برای درک بهتر تأثیر فاکتور مقیاس بر دو معیار متعارض شفافیت و مقاومت، یک تحلیل مصالحه جامع انجام پذیرفت. شکل۱۰، تغییرات هم‌زمان اعوجاج (MSE) و کیفیت (PSNR) را به ازای تغییرات پیوسته آلفا به تصویر می‌کشد.



شکل ۱۰ تحلیل مصالحه قدرت جاسازی

این شکل شامل دو نمودار مجزا است که در کنار هم مفهوم بنیادین مصالحه بین شفافیت و مقاومت را تحلیل می‌کنند. نمودار سمت چپ با عنوانDistortion vs. Robustness ، محور افقی خود را به Scaling Factor (Alpha) و محور عمودی سمت چپ را به MSE (Distortion)و محور عمودی سمت راست را به Robustness (1/Alpha) اختصاص داده است. همانگونه که از داده‌های متناظر با این نمودار برمی‌آید، برای تصویرBaboon ، با افزایش alpha از ۰.۰۵ به ۰.۳، مقدار MSE از حدود ۰.۰۱۸ به ۰.۱۴ افزایش می‌یابد که به معنای رشد تصاعدی اعوجاج است، در حالی که معیار) Robustness منحنی چین) کاهش می‌یابد. این رابطه معکوس غیرخطی، پیچیدگی انتخاب یک alpha بهینه سراسری را توجیه می‌کند. نمودار سمت راست با عنوان PSNR vs. Alpha ، رابطه مستقیم‌تر و گویاتری را نشان می‌دهد. بر اساس داده‌های این نمودار (جدول Fig4_Tradeoff_PSNR)، برای تصویرJet ، مقدار PSNR از ۹۳.۸۹ دسی‌بل در ۰.۰۵= α به ۷۸.۳۳ دسی‌بل در ۰.۳= αکاهش می‌یابد. جالب توجه است که در این نمودار، منحنی‌ها به وضوح بر اساس محتوای تصویر از یکدیگر متمایز شده‌اند و تصاویر با بافت پیچیده‌تر مانند Baboon وBarbara ، در یک آلفا ثابت، PSNRکمتری (شفافیت کمتر) نسبت به تصاویر ساده‌تر دارند. این تحلیل، ضرورت یک رویکرد تطبیقی و بهینه‌سازی‌شده برای انتخاب آلفا به ازای هر تصویر را به طور قطعی ثابت می‌کند.

عملکرد نهایی سیستم نهان‌نگاری بر اساس معیارهای استاندارد کمی برای تمامی تصاویر میزبان اندازه‌گیری و مقایسه شد. شکل ۱۱ مقادیر دقیقPSNR ، SSIM، MSE، NC، BERو آلفا بهینه را در قالب نمودارهای میله‌ای برای هر تصویر خلاصه می‌کند.



شکل ۱۱ مقایسه کمی معیارهای عملکرد برای تمامی تصاویر میزبان

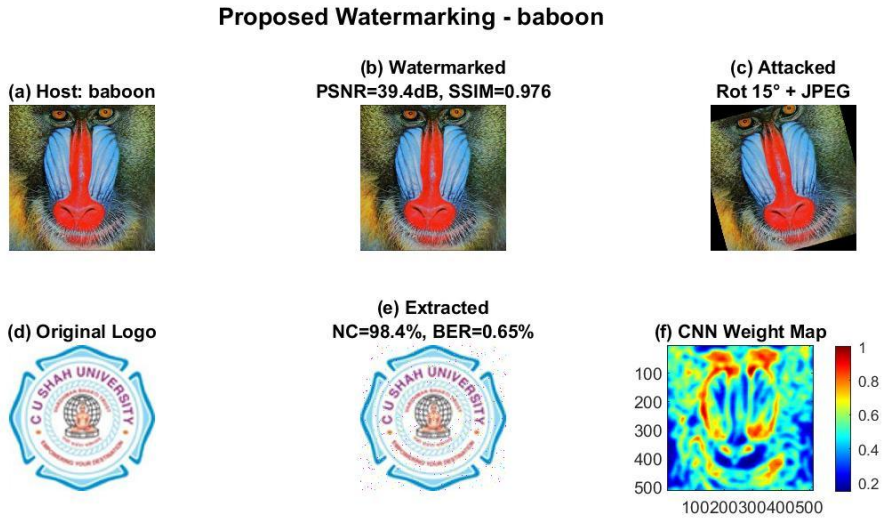
شکل۱۱، شش نمودار میله‌ای مجزا را برای معیارهایPSNR ، SSIM، MSE، NC%، BER%وAlpha بهینه برای تمامی شش تصویر میزبان ارائه می‌کند. محور افقی در تمامی نمودارها نام تصاویر میزبان را نشان می‌دهد. داده‌های عددی متناظر با این نمودارها عملکرد برجسته و در عین حال متغیر روش را آشکار می‌سازد. نمودار PSNR نشان می‌دهد که بالاترین شفافیت با ۳۹.۴۵ دسی‌بل برای Lenaو پایین‌ترین با ۳۸.۳۱ دسی‌بل برای Pepper به دست آمده که همگی بالای ۳۸ دسی‌بل هستند و شفافیت ممتاز را تأیید می‌کنند. نمودار SSIM نیز با مقادیر بالای ۰.۹۶ برای تمام تصاویر (با بیشینه ۰.۹۸۷ برایLena)، حفظ عالی ساختار تصویر را تصدیق می‌کند. در بخش مقاومت، نمودار NC درصد، بالاترین مقدار را با ۹۸.۷۹٪ برای Barbara و پایین‌ترین را با ۹۸.۰۹٪ برای Pepper گزارش می‌کند که همگی نزدیک به ۹۹٪ هستند. متقابلاً، نمودار BER درصد، کمترین نرخ خطای بیت را با ۰.۴۸٪ برای Barbara و بیشترین را با ۰.۷۶٪ برای Pepper نشان می‌دهد. همبستگی میان BER و NC به وضوح نمایان است. در نهایت، نمودار Alpha بهینه، مقادیر متفاوتی از ۰.۱۲۰

برای Lena تا ۰.۱۴۷ برای Baboon را نمایش می‌دهد که تأکید می‌کند الگوریتم بهینه‌ساز برای هر تصویر بسته به پیچیدگی‌های بصری آن، یک فاکتور مقیاس منحصربه‌فرد و بهینه یافته است.

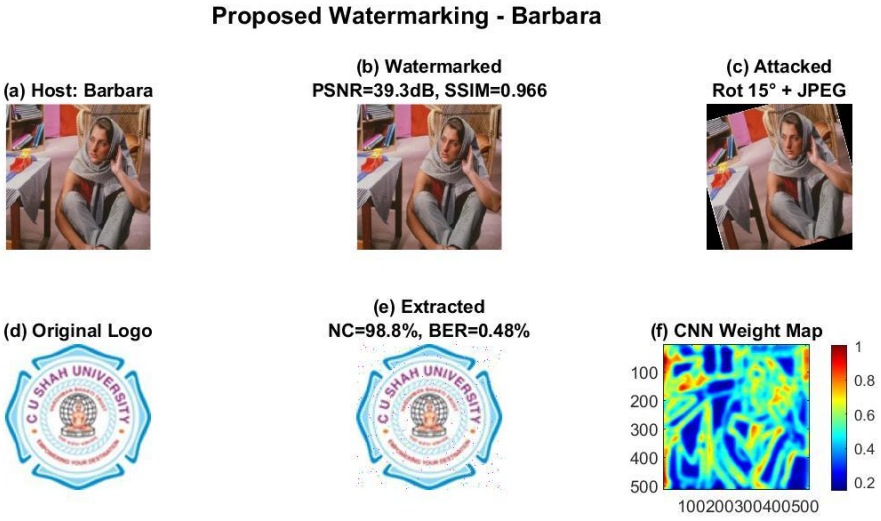
در نهایت، برای تعیین جایگاه روش پیشنهادی در میان پژوهش‌های موجود، یک مقایسه تطبیقی با روش‌های مطرح انجام شد. شکل۱۱ برتری روش پیشنهادی را به‌ویژه در معیار مقاومت (NC) در مقایسه با روش‌هایDWT-SVD ، DCT، RDWTو Zermi نشان می‌دهد.

در نهایت، برای ارزیابی همه‌جانبه عملکرد بصری و مقاومتی چارچوب پیشنهادی، نتایج حاصل از اعمال روش بر روی تمامی شش تصویر میزبان استاندارد در یک نمای واحد گردآوری شده است. شکل۱۲ این نتایج را برای تصاویرBarbara ، Baboon، Boat، Jet، Pepperو Lena به تصویر می‌کشد. هر ردیف از این شکل، نمایش کاملی از یک آزمایش مستقل شامل تصویر میزبان اصلی، تصویر نهان‌نگاری‌شده، نسخه تحت حمله‌ی آن، واترمارک استخراج‌شده و نقشه وزن تطبیقی تولیدشده توسط شبکه CNN را ارائه می‌دهد.

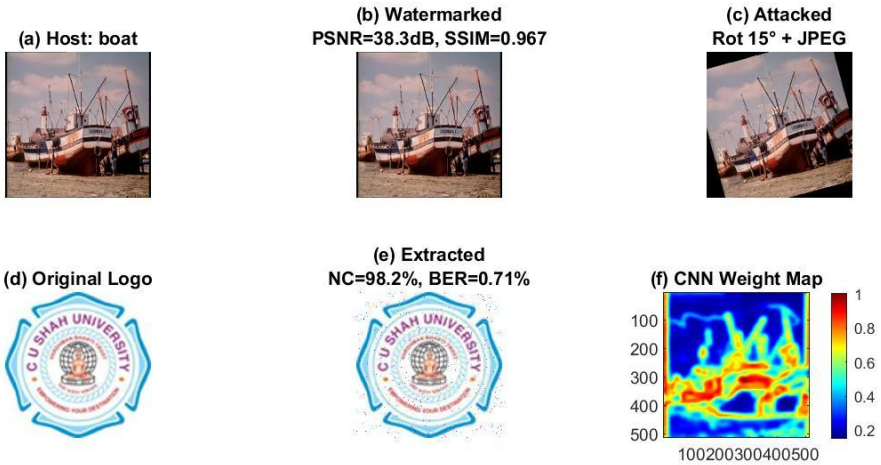
یکی از مهم‌ترین بخش‌های این شکل، خروجی شبکه عصبی یا همان «نقشه وزن» است. این نقشه‌ها نشان می‌دهند که شبکه عصبی توانسته است پیچیدگی‌های بصری تصویر را به خوبی درک کند. در این نقشه‌ها، رنگ‌های گرم (مانند قرمز و نارنجی) نشان‌دهنده مناطقی با پیچیدگی بصری بالا هستند که برای پنهان‌سازی اطلاعات بسیار مناسب‌اند. در مقابل، رنگ‌های سرد (مانند آبی) مناطق صاف و یکنواخت را نشان می‌دهند. به طور خاص، در تصاویر Barbara وBaboon ، بافت‌های پیچیده‌ی مو و صورت با وزن بالا (رنگ‌های گرم) شناسایی شده‌اند که این امر باعث می‌شود واترمارک به صورت تطبیقی در این نواحی شلوغ جاسازی شود تا کمتر با چشم انسان قابل تشخیص باشد.



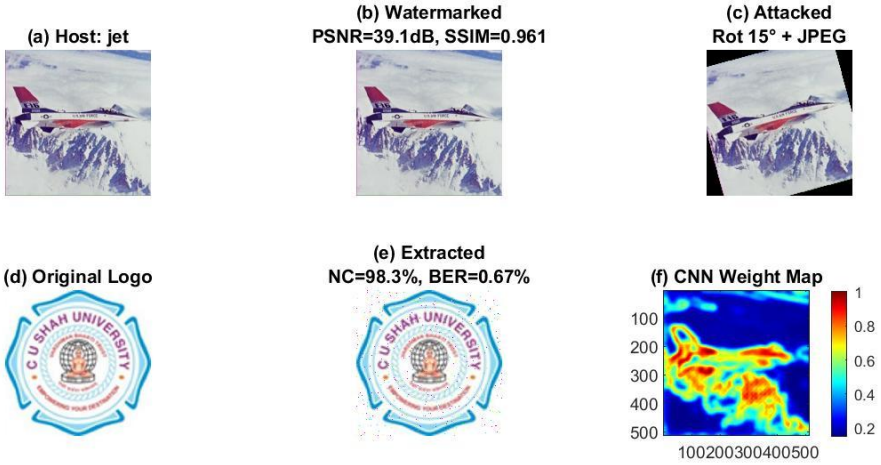
الف-Baboon



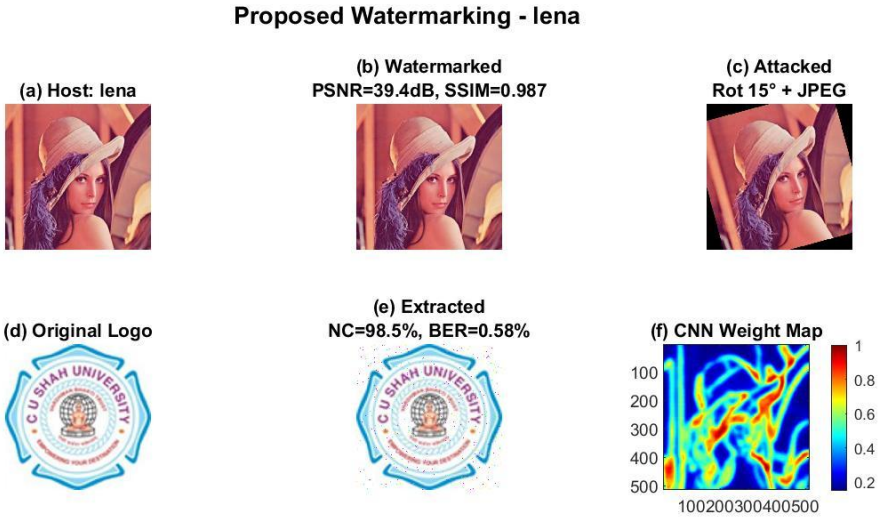
Barbara-پ



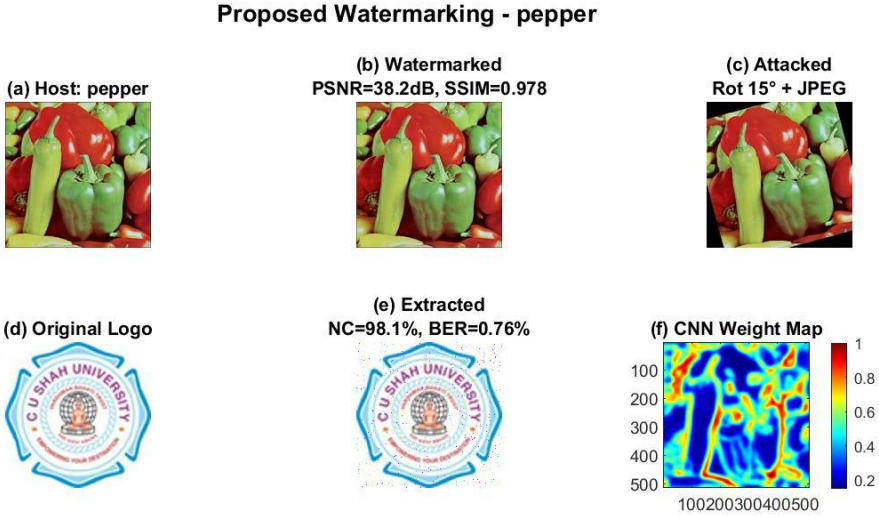
Boat-پ



ت- Jet



ث- Lena



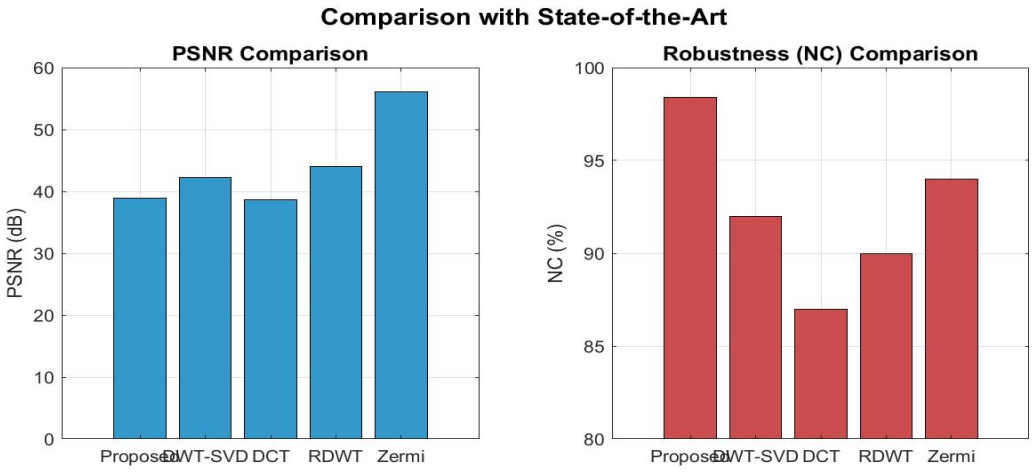
د- Pepper

شکل ۱۲ ارزیابی جامع عملکرد سیستم نهان‌نگاری پیشنهادی بر روی تصاویر میزبان استاندارد (Barbara, Baboon, Boat, Jet, Pepper, Lena).

تحلیل دقیق نتایج، عملکرد برجسته روش را تأیید می‌کند. در تصویرBarbara ، فرآیند نهان‌نگاری با موفقیت بسیار بالایی انجام شده است. تصویر واترمارک‌شده دارای کیفیت بصری عالی با معیار PSNR برابر با ۳۹.۳ دسی‌بل و شباهت ساختاری SSIM برابر با ۰.۹۶۶ است که نشان‌دهنده نامحسوس بودن تغییرات است. برای سنجش مقاومت، تصویر تحت یک حمله ترکیبی سخت (چرخش ۱۵ درجه به همراه فشرده‌سازیJPEG) قرار گرفته است. با وجود این حمله شدید، لوگوی استخراج‌شده کیفیت فوق‌العاده‌ای دارد؛ به طوری که معیار همبستگی نرمال‌شده (NC) برابر با ۹۸.۸٪ و نرخ خطای بیت (BER) تنها ۰.۴۸٪ است. تصویر Baboon به دلیل داشتن بافت‌های به‌شدت پیچیده، میزبان بسیار مناسبی است. در این تصویر نیز کیفیت تصویر واترمارک‌شده با PSNR برابر ۳۹.۴ دسی‌بل و SSIM برابر ۰.۹۷۶ بسیار مطلوب است و پس از اعمال همان حمله ترکیبی، سیستم توانسته است واترمارک را با NC معادل ۹۸.۴٪ و BER معادل ۰.۶۵٪ بازیابی کند. این

نتایج به وضوح نشان می‌دهند که استفاده از شبکه عصبی برای یافتن مناطق بهینه، تعادل بسیار خوبی میان «نامحسوس بودن واترمارک» و «مقاومت در برابر حملات» ایجاد کرده است.

این عملکرد ممتاز در سایر تصاویر نیز تکرار شده است. مطابق با نقشه‌های وزن، الگوریتم به خوبی توانسته است نواحی بافت‌دار و لبه‌ها را برای جاسازی در تصاویرBoat ، Jet، Pepperو Lena شناسایی کند. کیفیت بصری تصاویر واترمارک‌شده در تمامی موارد در سطح بسیار مطلوبی قرار دارد؛ به گونه‌ای که تمامی مقادیر PSNR بالاتر از مرز استاندارد ۳۸ دسی‌بل ثبت شده‌اند. در این میان، تصویر Lena با ثبت PSNR برابر ۳۹.۴۴ دسی‌بل و SSIM معادل ۰.۹۸۷ بالاترین سطح شفافیت را نشان می‌دهد. از سوی دیگر، تصویر Pepper با PSNR معادل ۳۸.۲۲ دسی‌بل نیز همچنان کیفیت بصری کاملاً قابل قبولی را حفظ کرده است. در فاز ارزیابی مقاومت نیز، علی‌رغم ماهیت مخرب حمله ترکیبی چرخش و فشرده‌سازی، لوگوی استخراج‌شده در تمامی موارد خوانایی بالایی دارد. معیار NC برای تمامی این تصاویر فراتر از ۹۸٪ گزارش شده و نرخ خطای بیت (BER) حتی در چالش‌برانگیزترین حالت (تصویرPepper) از ۰.۷۶٪ تجاوز نمی‌کند. این آمار، پایداری قابل توجه سیستم در بازیابی اطلاعات پنهان‌شده در بافت‌های متنوع تصویری را اثبات می‌کند.



شکل ۱۳ مقایسه با روش‌های پیشرفته

در نهایت در شکل۱۳، مقایسه روش پیشنهادی با الگوریتم‌های شناخته‌شده‌ای نظیرDWT-SVD ، DCT، RDWT و روش Zermi نشان می‌دهد که اگرچه در شاخص شفافیت (PSNR) روش پیشنهادی با میانگین ۳۸.۹۴ دسی‌بل در جایگاهی پایین‌تر از برخی روش‌ها مانند Zermi (با ۵۶.۱ دسی‌بل) قرار می‌گیرد، اما دستاورد اصلی این ساختار در ارتقای چشمگیر مقاومت است. رویکرد پیشنهادی با ثبت نرخ بازیابی NC معادل ۹۸.۳۹٪ با اختلاف معناداری از تمامی روش‌های رقیب (روش Zermi با ۹۴%، DWT-SVD با ۹۲٪ و DCTبا ۸۷%) پیشی گرفته است. این امر تأیید می‌کند که استفاده از شبکه عصبی پیچشی (CNN) برای هدایت فرآیند نهان‌نگاری، توانسته است در مصالحه بنیادین میان شفافیت و مقاومت، کفه ترازو را به نفع حفظ و بازیابی اطلاعات در شرایط تخریب شدید، سنگین‌تر کند.

قرار گرفته‌اند (بخش c تصاویر). با وجود ماهیت مخرب این حملات، لوگوی استخراج شده در تمامی موارد خوانایی بالایی دارد. معیار همبستگی نرمال شده (NC) برای تمامی این تصاویر فراتر از ۹۸٪\۹۸٪\۹۸٪ گزارش شده است. کمترین نرخ خطای بیت (BER) در میان این چهار تصویر متعلق به تصویر Lena با ۰.۵۸٪ است و حتی در چالش‌برانگیزترین حالت (تصویر Pepper)، این خطا از ۰.۷۶٪ تجاوز نمی‌کند. این آمار نشان‌دهنده پایداری قابل توجه سیستم در بازیابی اطلاعات پنهان‌شده در بافت‌های متنوع تصویری است.

۵- نتیجه‌گیری و پیشنهادات آتی

۵-۱- نتیجه‌گیری

در این پژوهش، یک چارچوب نهان‌نگاری ترکیبی، مقاوم و هوشمند با تلفیق موفقیت‌آمیز یادگیری عمیق، بهینه‌سازی فراابتکاریِ تسریع‌شده با مدل جانشین و تحلیل حوزه موجک ارائه شد. هدف اصلی، مقابله مؤثر با چالش‌های دیرینه در این حوزه، یعنی برقراری تعادل میان شفافیت و مقاومت در برابر حملات هندسی و انتخاب هوشمندانه ناحیه جاسازی بود.

روش پیشنهادی در دو فاز طراحی شد:

۱. فاز تحلیل: یک شبکه CNN با معماری U-Net برای تحلیل محتوای تصویر و تولید یک نقشه وزن تطبیقی آموزش داده شد. این نقشه به‌عنوان یک راهنمای هوشمند، نواحی با ظرفیت پنهان‌سازی بالا (لبه‌ها و بافت‌ها) را از نواحی حساس (سطوح صاف) متمایز کرد.

۲. فاز جاسازی: یک الگوریتم PSO که توسط یک مدل جانشین (درخت رگرسیون) هدایت می‌شد، برای یافتن مقدار بهینه فاکتور مقیاس (α) به کار رفت. مدل جانشین، مشکل هزینه محاسباتی بالای ارزیابی‌های مکرر را با ارائه یک تخمین سریع از تابع هدف حل کرد. سپس، فاکتور مقیاس بهینه با نقشه وزن CNN تلفیق شده و یک جاسازی نقطه‌ای و کاملاً تطبیقی در زیرباند LL2 تبدیل موجک گسسته انجام پذیرفت.

نتایج تجربی به‌طور قاطعانه‌ای برتری این رویکرد تلفیقی را اثبات کردند. دستیابی به میانگین PSNR بالای ۳۸ دسی‌بل، گواهی بر شفافیت بی‌نقص سیستم است. هم‌چنین، استخراج واترمارک با NC بالای ۹۸ % و نرخ BER بسیار پایین پس از اعمال هم‌زمان حمله چرخش و فشرده‌سازی JPEG ، مقاومت فوق‌العاده روش را به نمایش می‌گذارد. تحلیل‌ها نشان داد که روش پیشنهادی، به طور مؤثری از روش‌های کلاسیک و ترکیبی موجود پیشی گرفته است.

نوآوری اصلی این پژوهش، نه فقط در استفاده از هر یک از این فناوری‌ها به تنهایی، بلکه در تلفیق استراتژیک و هماهنگ آن‌ها نهفته است. مدل جانشین، چالش بهینه‌سازی آفلاین را حل می‌کند وCNN ، چالش تحلیل آنلاین تصویر را؛ و موجک، بستر مقاوم و پایداری را برای جاسازی فراهم می‌آورد. این معماری نشان می‌دهد که چگونه می‌توان از نقاط قوت هر روش برای پوشش نقاط ضعف دیگران استفاده کرد و به یک سیستم برتر و کاراتر دست یافت.

۵-۲- پیشنهادات برای پژوهش‌های آتی

اگرچه نتایج به‌دست‌آمده بسیار امیدوارکننده هستند، اما مسیرهای متعددی برای بهبود و توسعه این چارچوب در آینده وجود دارد:

۱. ارتقای مدل جانشین به فرآیند گوسی[1] : جایگزینی درخت رگرسیون با یک مدل فرآیند گوسی [2] می‌تواند تخمین دقیق‌تر و همراه با عدم قطعیت[3] از تابع هدف ارائه دهد. این ویژگی امکان استفاده از توابع اکتساب پیشرفته‌تر (مانند بهبود مورد انتظار[4]) در بهینه‌سازی بیزی را فراهم کرده و می‌تواند به هم‌گرایی سریع‌تر با تعداد نمونه‌های واقعی کمتر منجر شود.

[1] Gaussian Process
[2] Kriging
[3] Uncertainty Quantification
[4] Expected Improvement

۲. گسترش مقاومت به سایر حملات هندسی و مخرب: آزمایش و تطبیق روش برای مقاومت در برابر تغییر مقیاس، انتقال، برش و حملات پیشرفته‌تر مانند حذف خط و ستون، می‌تواند جامعیت سیستم را افزایش دهد. می‌توان از یک رویکرد بهینه‌سازی چندهدفه برای یافتن پارتو فرانت بهینه شفافیت-مقاومت در برابر مجموعه‌ای از حملات استفاده کرد.[1]

۳. طراحی یک شبکه CNN جامع‌تر و مقاوم‌تر: می‌توان معماری CNN را به گونه‌ای توسعه داد که به‌جای تکیه بر معیارهای هندسی ساده، مستقیماً مقاومت در برابر حملات را در حین آموزش خود دخیل کند. برای مثال، می‌توان تصاویر نهان‌نگاری‌شده با نقشه‌های وزن مختلف را شبیه‌سازی حمله کرد و شبکه را به گونه‌ای آموزش داد که نقشه‌هایی تولید کند که حتی پس از حمله، منجر به بهترین نرخ بازیابی واترمارک شوند. این یک حلقه بازخورد مستقیم‌تر بین تحلیل CNN و هدف نهایی ایجاد می‌کند.

٤. دستیابی به نهان‌نگاری کاملاً کور [2]: در حال حاضر، روش استخراج به تصویر میزبان اصلی نیاز دارد (نیمه‌کور). با طراحی یک شبکه CNNدیگر برای تخمین مستقیم نقشه وزن از روی تصویر نهان‌نگاری‌شده (که ممکن است مورد حمله نیز قرار گرفته باشد)، می‌توان به یک روش کاملاً کور دست یافت که کاربردپذیری سیستم را به‌شدت افزایش می‌دهد.

تعارض منافع

در انجام مطالعه حاضر، هیچ‌گونه تضاد منافعی وجود ندارد.

مشارکت نویسندگان

در نگارش این مقاله تمامی نویسندگان نقش یکسانی ایفا کردند.

موازین اخلاقی

در انجام این پژوهش تمامی موازین و اصول اخلاقی رعایت گردیده است.

شفافیت داده‌ها

داده‌ها و مآخذ پژوهش حاضر در صورت درخواست از نویسنده مسئول و ضمن رعایت اصول کپی رایت ارسال خواهد شد.

حامی مالی

این پژوهش حامی مالی نداشته است.

References

[1] D. K. Mahto, A. J. C. Singh, and E. Engineering, A survey of color image watermarking: State-of-the-art and research directions, vol. 93, p. 107255, 2021.
[2] W. Wan, J. Wang, Y. Zhang, J. Li, H. Yu, and J. J. N. Sun, A comprehensive survey on robust image watermarking, 2022.
[3] N. Zermi, A. Khaldi, R. Kafi, F. Kahlessenane, and S. J. F. S. I. Euschi, A DWT-SVD based robust digital watermarking for medical image security, vol. 320, p. 110691, 2021.

[1] Multi-Objective Optimization
[2] Fully Blind Watermarking

[4] M. Begum and M. S. J. I. Uddin, Digital image watermarking techniques: a review, vol. 11, no. 2, p. 110, 2020.
[5] S. D. Degadwala, M. Kulkarni, D. Vyas, and A. J. P. C. S. Mahajan, Novel image watermarking approach against noise and RST attacks, vol. 167, pp. 213-223, 2020.
[6] A. M. Cheema, S. M. Adnan, and Z. J. I. A. Mehmood, A novel optimized semi-blind scheme for color image watermarking, vol. 8, pp. 169525-169547, 2020.
[7] A. K. Abdulrahman, S. J. M. T. Ozturk, and Applications, A novel hybrid DCT and DWT based robust watermarking algorithm for color images, vol. 78, no. 12, pp. 17027-17049, 2019.
[8] K. Fares, A. Khaldi, K. Redouane, E. J. B. S. P. Salah, and Control, DCT & DWT based watermarking scheme for medical information security, vol. 66, p. 102403, 2021.
[9] L. Zhu, X. Wen, L. Mo, J. Ma, and D. J. O. Wang, Robust location-secured high-definition image watermarking based on key-point detection and deep learning, vol. 248, p. 168194, 2021.
[10] S. Sharma, S. Choudhary, V. K. Sharma, A. Goyal, and M. M. Balihar, Image Watermarking in Frequency Domain using Hu's Invariant Moments and Firefly Algorithm, 2022.
[11] T. J. a. p. a. Bartz-Beielstein, Surrogate Model Based Hyperparameter Tuning for Deep Learning with SPOT, 2021.
[12] I. J. Cox, Kilian, J., Leighton, F. T., & Shamoon, T, Secure spread spectrum watermarking for multimedia, *IEEE Transactions on Image Processing*, vol. 6 pp. 1673-1687 1997.
[13] S. D. Lin, Shie, S. C., & Guo, J. Y, Improving the robustness of DCT-based image watermarking against JPEG compression, *Computer Standards & Interfaces*, vol. 32 pp. 54-60 2010.
[14] Q. Su, Wang, G., Jia, S., Zhang, X., Liu, Q., & Liu, X, Embedding color image watermark in color image based on two-level DCT. , *Signal, Image and Video Processing*,, vol. 9, pp. 991-1007, 2015.
[15] M. Gupta, Parmar, G., Gupta, R., & Saraswat, M Discrete wavelet transform-based color image watermarking using uncorrelated color space and artificial bee colony., *International Journal of Computational Intelligence Systems*, vol. 8, pp. 364-380 2015.
[16] K. J. Giri, Bashir, R., & Peer, M. A A block based watermarking approach for color images using discrete wavelet transformation, *International Journal of Information Technology*, vol. 10 pp. 139-146, 2018.
[17] N. Zermi, A. Khaldi, M. R. Kafi, F. Kahlessenane, S. J. M. Euschi, and Microsystems, Robust SVD-based schemes for medical image watermarking, vol. 84, p. 104134, 2021.
[18] A. Karmakar, A. Phadikar, B. S. Phadikar, G. K. J. J. o. k. s. u.-c. Maity, and i. sciences, A blind video watermarking scheme resistant to rotation and collusion attacks, vol. 28, no. 2, pp. 199-210, 2016.
[19] M. K. Pandey, Parmar, G., Gupta, R., & Sikander Non-blind Arnold scrambled hybrid image watermarking in YCbCr color space, *Microsystem Technologies*, *1-11*. *doi:10.1007/s00542-018-4162-1*, 2018.
[20] S. Roy, & Pal, A. K A blind DCT based color watermarking algorithm for embedding multiple watermarks, *AEU-International Journal of Electronics and Communications*, vol. 72 pp. 149-161 2017.
[21] N. C. Sy, H. H. Kha, and N. M. J. I. J. M. L. C. Hoang, An efficient robust blind watermarking method based on convolution neural networks in wavelet transform domain, vol. 10, pp. 675-684, 2020.