A Deep Metric Learning and Graph-cut Approach for Logo Authentication and Forgery Detection in E-commerce

Preeti C. M.* and T. Santhi Sri

Department of Computer Science Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, India
Email: preeticm.loyola@gmail.com (P.C.M.); santhisri@kluniversity.in (T.S.S.)
*Corresponding author

***Abstract*—The increasing prevalence of counterfeit goods in online marketplaces poses a significant threat to brand equity, consumer trust, and overall digital security. Traditional image-based logo verification systems often struggle to accurately capture both visual similarity and spatial consistency, particularly under challenging conditions such as occlusion, manipulation, and severe visual distortions. To address these limitations, this paper proposes a dual-stage deep metric learning and graph-cut framework that integrates semantic logo authentication with detailed forgery segmentation. In the first stage, deep metric learning combined with Adaptive Multi-Center Triplet Loss (AMCTL) is used to map logos into a discriminative embedding space, enabling effective separation between authentic and counterfeit instances. In the second stage, a graph-cut-based optimization technique is employed to localize forged regions through contextual abnormality detection and pairwise energy minimization. The proposed framework is evaluated on both publicly available and customized logo datasets and demonstrates consistent and competitive performance compared to existing approaches such as Deep Semantic Contrastive Learning Classifier (DSCL), Graph-Cut Segmentation (GCut+), and Region-Aware Metric Learning (RAML). Improvements are observed across multiple evaluation metrics, including accuracy, F1-Score, Intersection over Union (IoU), and Area Under the Curve (AUC). Furthermore, the model maintains computational efficiency, making it suitable for practical deployment in real-time e-commerce environments. Overall, the proposed approach provides a scalable and interpretable solution for reliable logo verification in modern online marketplaces.**

***Keywords*—logo authentication, deep metric learning, graph-cut segmentation, forgery detection, e-commerce security**

I. Introduction

The performance of image forgery detection and logo recognition systems has improved significantly following recent developments in deep learning, especially Convolutional Neural Networks (CNNs). A number of studies [1–4] have developed deep learning-based architectures to identify single and multiple logos in still and video images, and the systems have greatly enhanced the accuracy of detection and scale. Yang *et al*. [3] have presented fake object detection, Lightweight multi-task learning models. Region-sensitive metric learning models have been shown to work in few-shot industrial detection [4]. Deep metric types of learning have also enhanced logic verification and design authentication systems [3, 4].

Image splicing, copy-move, and composite forgery detection based on CNNs have been thoroughly surveyed [5–7]. CNN models that use dual-stream attention have been shown to be promising in the localization of forgeries [8], and optimization-enhanced U-Net frameworks have demonstrated great precision and recall in the correct identification of manipulated regions [8]. Intelligent forgery detection has also been undertaken using traditional segmentation-based methods including graph-cut and GrabCut techniques to detect forgery in documents and receipt analysis [8–10]. There are also multi-modal and automated trademark segmentation systems that have shown the significance of context-sensitive logo analysis to real-life deployment cases [10–12].

Although this has been achieved, the current frameworks shown in Fig. 1, are more based on detection and classification without considering the content-based reconstruction of tampered or distorted logo regions. The majority of existing methods are not very resistant to real-world distortions (occlusion, compression artifacts, low resolution, adversarial manipulation, etc.) [12, 13]. Moreover, few studies are addressed with in-built security controls of cryptographic authentication or steganographic hiding, hence presenting logo data to unauthorized copying, access, or manipulation [14–16]. Although some more sophisticated detection models, like the dual-stream CNN architectures or the adaptive metric learning models, exhibit high detection accuracy [15, 16], such models lack inherent authentication or secure content reconstruction features.

Manuscript received May 26, 2026; revised June 23, 2026; accepted July 7, 2026; published September 14, 2026.

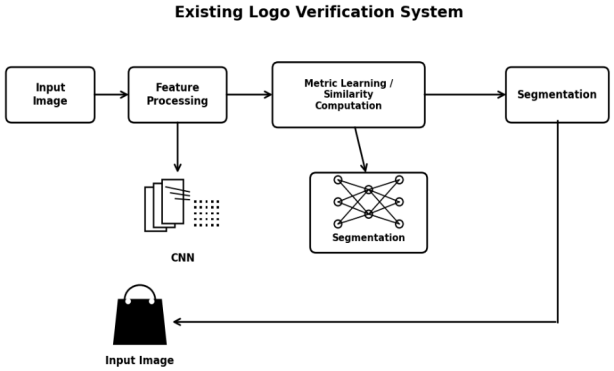

Fig. 1. Current verification system of logos, which depicts variable extraction based on CNN, deep metric similarity calculation, and graph-cut segmentation of genuine analysis.

A. Challenges

Despite the high performance of the deep learning techniques in the task of logo detection and the classification of forgery cases [15–20], a number of important problems that are yet to be addressed can be identified. Majority of the existing systems focus on localization or classification without the capacity of reconstructing tampered and hidden logo areas in a content-sensitive fashion [15–19]. This shortcoming is especially clear in real e-commerce situations where images are frequently reduced and noisy or dimly lit or partially covered [20, 21].

Also, the current CNN-based designs seldom incorporate cryptographic or steganographic protection, so the information contained in logos can be compromised and altered by unauthorized individuals [20–22]. The models are also prone to adversarial forgeries and advanced tampering attacks [22]. Besides that, some generalizations are also an issue when the amount of training data is small or computing power is limited, which in turn limits scalability in low-data industrial cases [22, 23]. These constraints underscore the need to have a single, secure, and context-sensitive framework with the ability to perform robust authentication of logos, reconstruction and integrity protection.

B. Motivation

Even though the current state-of-the-art CNN-based frameworks have shown great accuracy in detection on benchmark data, their validity in real-life e-commerce settings is minimal. As an example, dual-stream CNN models have been reported to perform well in localization, and lightweight multi-task networks and region-aware metric learning systems have been reported to perform well in counterfeit detection and industrial anomaly identification [23–26]. Nevertheless, these methods are essentially aimed at locating the tampered areas without restoring hidden or distorted logotypes.

Also, the systems presently in place tend to degrade when operating under compressed, low-resolution or adversarial manipulated conditions which are prevalent online [26, 27]. More importantly, they do not have in-built cryptographic validation or steganographic watermarking tools that can be used to preclude illegitimate approaches to reuse or alter recognized logos [27–29]. The large-scale protection of the brand will not be possible with high detection metrics, including Area Under the Curve (AUC), precision, and recall, as they do not guarantee tamper-proof authentication or safe embedding. Such restrictions encourage the creation of a cooperative structure that will combine visual recognition, spatial-context recovery, and hybrid information protection. The suggested CNN-based architecture is expected to offer scalable, robust, and machine-provable logo authentication to adversarial and high-risk e-commerce ecosystems.

C. Contributions to the Paper

The key goals of the research can be outlined in the following way:

(1) To identify and estimate counterfeit or altered brand logos in the digital product images in real-time. To repair damaged or covered logo areas with deep spatial and contextual feature learning.
(2) To authorize logos by using built-in cryptographic identities and stealthy watermarking by steganography.
(3) To develop a scalable, lightweight, and robust pipeline that can be used in online retailing, logistics validation, and digital brand enforcement applications.

Unlike conventional approaches that treat logo authentication and forgery localization as independent tasks, the proposed framework introduces a tightly coupled dual-stage architecture integrating semantic embedding and spatial optimization. The novelty lies in three key aspects:

(1) The proposed Adaptive Multi-Center Triplet Loss (AMCTL), which models intra-class variability through multiple representative centers rather than a single embedding anchor;
(2) The incorporation of semantic attention into graph-cut optimization, enabling context-aware energy minimization instead of purely intensity-based segmentation;
(3) A Multi-Scale Graph Fusion (MSGF) strategy that ensures robustness to scale-dependent tampering. This unified formulation enables simultaneous authentication, localization, and reconstruction within a single interpretable pipeline, which is not addressed in existing hybrid frameworks.

The rest of this paper will be structured in the following way. Section II examines similar works on logo detection, forgery analysis and secure authentication schemes. Section III contains the proposed methodology, system model and problem formulation. Section IV will give a performance review and comparison. Lastly, Section V is the conclusion of the paper, which presents the research direction in the future.

II. Related Work

This section discusses the literature work on logo authentication and forgery localization. Counterfeit production, editing digital content, and brand theft on e-commerce and social media platforms have emerged as the major research issues in terms of logo validation and

image forgery detection. Conventional image forensics used handcrafted features and statistical inconsistency, though deep learning has revolutionized this field by making it possible to extract features automatically, better generalization and more resilient to various manipulation methods [1–3]. Early CNN-based architectures mainly considered copy-move, splicing, and content-removal attacks as they were tackled with classification-based architectures. The effectiveness of deep models over the conventional methods in detection accuracy and scalability was, in fact, established through extensive surveys [1, 3, 4].

Architectures that can improve attention and two further streams of learning have further boosted the detection of forgery by combining both the spatial and frequency-domain information [4]. Segmentation networks guided by optimization were shown to be more precise on localization and un-reliant on F1-Score [5], and self-supervised and tamper-guided approaches were also more robust and less dependent on annotation [6, 7]. Though these have been made, the vast majority of them are detection-based, and do not extend to content reconstruction and secure validation.

There are other problems that are added by the logos detection, such as scale, occlusion, light variations, and clutter on the background. The surveys and implementations of CNN-based logo detection prove the great advancement of brand monitoring applications [7–9]. Transformer models with spatial attention and feature pyramids enhance small object recognition [9], and active and transfer learning methods cut labelling tasks in video-based detection tasks [9]. Object detection and refinement modules are used together in multi-stage segmentation pipelines, which improve the localization of the boundaries in real-world datasets [10, 11]. These systems, however, do not include embedded authentication or tamper-resistant reconstruction but just concentrate on detection and segmentation accuracy.

In addition to detection, deep metric learning has been considered in industrial design verification and similarity matching of counterfeits. Non-categorical label based embedding authentication [11, 12] and discrimination refinement by refining representations with batch diffusion and self-distillation have been studied [13, 14]. Lightweight multi-task networks and region-conscious learning of metrics deal with few-shot and small-data conditions in industry settings [14, 15]. Material-level authentication with a physical-layer protection based on nano diamond-based optical signatures was also proposed [15]. Whereas they are useful in ensuring similarity checking, they fail to offer combined forgery detection and provision of secure embedding in a single digital framework.

CNN graph-cut based segmentation-oriented pipelines show high accuracy boundary delineation in document and trademark analysis tasks [15–19]. These methods enable proper extraction of regions of the logos, which is obligatory when it comes to downstream verification. However, they do not have cryptographic validation schemes and do not deal with the reconstruction of occluding or compromised logo structures.

According to recent research, it is recommended to consistently analyze the online logo images multimodally and vulnerably, pointing out that visual authentication systems are vulnerable to compression artifacts, domain changes, and adversarial examples [20, 21]. Even though existing models claim high performance indicators like precision, AUC and F1-Score, they mostly deal with individual functional units like forgery detection, logo localization, segmentation optimization, or similarity checking. Recent attempts to reconstruct forged logo regions content-aware have hardly been researched, and the combination of hybrid cryptographic validation and steganographic watermarking into deep learning structures is rarely discussed.

To conclude, current studies show significant advances in the accuracy of detection and segmentation, but indicate a disjointed design environment. It is apparent that there is a desire to have a combined framework that collaboratively does forgery detection, logo localization, spatial-context reconstruction, and embedded security authentication to provide machine-verifiable and robust logo protection in the real-world e-commerce ecosystem.

Existing hybrid forgery detection frameworks primarily combine deep feature extraction and segmentation modules in a loosely coupled manner, where semantic verification and spatial localization are performed independently. Methods such as Deep Semantic Contrastive Learning (DSCL) classifier focus mainly on metric-based representation learning, whereas graph-cut-based approaches emphasize low-level spatial optimization without incorporating semantic embedding priors. Similarly, attention-based and multi-scale segmentation frameworks improve localization performance but often lack robust embedding discriminability for authentic-versus-forged logo verification. In contrast, the proposed framework introduces a tightly integrated dual-stage architecture that jointly combines AMCTL, Attention-Guided Graph-Cut (AG-GC) optimization, and MSGF within a unified pipeline. Unlike conventional triplet loss formulations, AMCTL employs adaptive multi-center embedding optimization to better capture intra-class variations under realistic e-commerce distortions. Furthermore, semantic priors generated during metric learning are directly incorporated into graph-based energy minimization, enabling context-aware forgery localization rather than conventional intensity-based segmentation. This unified interaction between semantic embedding and structured spatial optimization differentiates the proposed framework from existing hybrid approaches.

III. Methodology

In this work, it is suggested that a multi-stage deep metric and graph-cut framework can be designed to be used in the fields of actual logo verification and forgery segmentation in the context of e-commerce. General architecture is based on the layered design, which combines CNN-based metric learning with adaptive Graph-Cut Segmentation (GCut+) to overcome the shortcomings of the current detection-centric methods. The proposed

framework shown in Fig. 2, unlike the previous approaches, by combining localized forgery detection, content-aware reconstruction, and interpretable authenticity validation, enables both in a single pipeline, all other foregoing approaches rely on classification or segmentation.

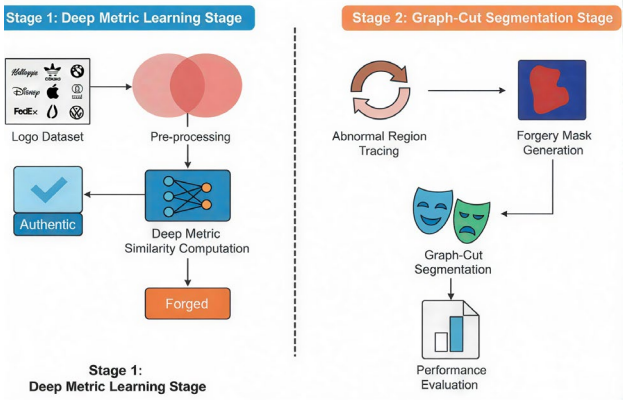

Fig. 2. Overview of the proposed dual-stage framework for logo authentication and forgery segmentation. Stage 1 performs deep metric learning-based semantic verification, while Stage 2 applies GCut+ for abnormal region detection and forgery mask generation.

The processing pipeline starts with a filtered dataset that comprises real and falsified logotypes that are found in the real world of e-commerce listing. The first stage is the input images normalization and spatial dimension scaling to make the input images consistent across different resolutions. This is then succeeded by a profound pre-processing layer that works on enhancing edges, removal of noise and sharpening details. These improvement processes enhance structural clarity and enhance feature discriminability to be used later in the generation of embedding.

The resultant map of spatially sensitive features is then derived with the help of a CNN-based representation module which identifies anomalous or inconsistent parts of the logo. This is vital in the detection of minute structural perversions, occlusions, or overlay-based manipulations that are usually evident in forged listings. The embedding phase uses the triplet-loss-based deep metric learning to optimize inter-class separability and minimize intra-class variance of the learned feature space. Because of this, genuine logos are concentrated together in the embedding manifold, and forged ones are disseminated due to structural dissonance.

The trained embeddings are then optimized by an adaptive GCut+ module, which does effective localization of tampered regions. The framework, which incorporates learned spatial prior information with energy minimization graphs, allows the accurate extraction of forgery boundaries and the reconstruction of the tampered or hidden logo portions. Such a combination of deep feature embedding and structured graph optimization guarantees fine-grained mask generation and enhanced interpretability of the detection result.

The last decision level determines whether the logo to be original or a fake one and at the same time generates forgery masks to be used in visual examination. The two-step combination of metric learning and GCut+ makes it more robust to conclusiveness, adversarial perturbation, compression artifacts, and low-resolution as is often the case with e-commerce imagery. The proposed architecture can perform resilient, interpretable and reconstruction-aware logo authentication on a single scalable framework by closely integrating representation learning with structured segmentation.

A. System Model

Let the input logo image be represented as Eq. (1).

$$I \in R^{H \times W \times C} \tag{1}$$

where H, W, and C denote the spatial Height, Width, and Channel dimensions.

$$I' = P(I) \tag{2}$$

where $P(\cdot)$ denotes the normalization, denoising and contrast enhancement. The processed image is then mapped into a discriminative latent space through a CNN embedding function.

$$Z = f_\theta(I'), Z \in R^d \tag{3}$$

And it is normalized to scale invariant embedding given as Eq. (4).

$$Z' = \frac{Z}{\|Z\|_2} \tag{4}$$

where Z_a, Z_p, and Z_n are denoted as anchor, positive, and negative embeddings, respectively. Then each class is represented as a set of adaptive centers.

$$C_k = \left\{ C_k^m \right\}_{m=1}^{M} \tag{5}$$

The compactness is enforced by the nearest center.

$$d_{pos} = \min_{c \in C_k} \left\| Z_a - C \right\|_2^2 \tag{6}$$

The inter-class separability is given as Eq. (7).

$$d_{neg} = \left\| Z_a - Z_n \right\|_2^2 \tag{7}$$

The embedding space is optimized using AMCTL is given as Eq. (8). It enforces the margin separation between the authentic and forged logo representation.

$$L_{AMCTL} = \max\left(0, d_{pos} - d_{neg} + \alpha\right) \tag{8}$$

Then the similarity of the normalized space is given as Eq. (9).

$$S\left(Z_i^n, Z_j^n\right) = \left(z_i^n, z_j^n\right) \tag{9}$$

The authenticity is determined by using a threshold decision rule.

$$y' = \begin{cases} Authentic & S \geq \tau \\ Forged & S < \tau \end{cases} \tag{10}$$

The Forgery Localization is formulated as Eq. (11).

$$E(L) = \sum V_{ij}\left(L_i, L_j\right)_{(i,j)\in E} \tag{11}$$

Now each scale is represented as Eq. (12).

$$I_s = Resize(I', S) \tag{12}$$

Then, GCut+ is applied at each scale.

$$M_s = GraphCut(I_s) \tag{13}$$

So that the resized original resolution is given as Eq. (14).

$$M_s{}' = Upsample(M_s) \tag{14}$$

This fusion helps to obtain the final forged mask given as Eq. (15).

$$M = majority\left(\left\{M_s{}'\right\}_{s\in S}\right) \tag{15}$$

B. Workflow of Proposed Framework

The suggested methodology presents a two-phase architecture known as the deep metric and graph-cut framework, which is intended to be used in authentic logo verification and fine-tuning of the forgery section in e-commerce business. The framework combines both deep feature learning and energy-based region segmentation to address the drawbacks of current detection-based frameworks. The architecture, through combining semantic embedding and spatial optimization, facilitates the process of authenticity verification and localized tamper reconstruction in a single pipeline. The workflow is shown in Fig. 2. It contains the following.

1) Stage 1: Deep metric learning of semantic logo verification

Assume that the input image is modeled in a high-dimensional tensor space as the representation on the right-hand side of the equilibrium and that the spatial dimensions together with the color channels are maintained. The image is then subjected to a normalization and enhancement operator to enhance its clarity and minimize the level of background noise, which is formally defined in the following equation: The purpose of this preprocessing step is to improve the edges of the image and eliminate clutter which boosts the downstream feature extraction. An enhanced image is then converted into a small latent feature space by a convolutional neural embedding function. This embedding function projects logos into a discriminative metric space with an explicit semantic similarity, unlike traditional SoftMax-based classification. Similarity between two logo embeddings is calculated with the help of a cosine similarity which offers scale-invariant angular similarity measurement. The embedding space is optimized by the formula of triplet loss as in the form of the triplet loss to enforce intra-class compactness and inter-class separation. This goal based on margins makes sure that genuine logos are compactly clustered in the metric space with forged or manipulated examples being pushed aside. Similarity threshold decision rule is then used to verify authenticity. In case the similarity score is above a preset value, then the logo is deemed authentic, it is sent to the spatial segmentation phase to undergo a thorough examination of the forgery.

2) Stage 2: Graph-cut region segmentation of forgery

In the case of samples that are considered suspicious, the structure switches to spatial forgery localization. Gradient-based anomaly tracing is the first step of detecting structural inconsistencies. In this step, areas with abnormal discontinuities are detected which in most cases would be tampered spots. An image weighted graph is then built and the weights of edges between the adjacent pixels are calculated based on the similarity. These are weights that encode local intensity consistency and act as smoothness constraints of segmentation. Localization of forgery is solved as an energy minimization problem. The unary term uses semantic priors based on the embedding phase, whereas the pairwise term uses a model of spatial coherence which is a Potts regularization model. Optimization of this energy criterion yields optimal pixel-wise labeling to trade off semantic inconsistency and spatial smoothness. The last forgery mask is pulled out which determines all spatial positions denoted as being forged. With this combination of semantic embedding and graph-based optimization, interpretable and reconstruction-conscious tamper localization is achieved, even when there are occlusions or blur and compression artifacts are present, such as those involved with e-commerce photographs.

Conventional triplet loss optimizes the embedding space by minimizing the distance between anchor-positive pairs while maximizing the separation between anchor-negative pairs. However, this formulation assumes that each class can be represented by a single compact embedding distribution. In practical logo authentication scenarios, significant intra-class variations arise due to occlusion, illumination changes, compression artifacts, geometric deformation, and adversarial perturbations. Consequently, conventional triplet loss often produces unstable embedding clusters with increased intra-class variance.

To address this limitation, the proposed AMCTL introduces multiple adaptive representative centers for each class. Instead of optimizing the anchor embedding with respect to a single global prototype, AMCTL associates

each sample with its nearest class-specific center, thereby enabling more flexible clustering of visually diverse logo instances. The intra-class variance can be expressed as Eq. (16).

$$\sigma_{intra}^2 = \frac{1}{N}\sum_{i=1}^{N}\left\|z_i - c_i^*\right\|_2^2 \quad (16)$$

where z_i represents the embedding of the i^{th} sample and c_i^* denotes its nearest adaptive center.

Minimizing this variance leads to more compact intra-class distributions and improved embedding stability. Furthermore, inter-class separability between authentic and forged logo embeddings is enhanced through maximization of the centroid distance as Eq. (17).

$$D_{inter} = \left\|\mu_{auth} - \mu_{forge}\right\|_2^2 \quad (17)$$

where μ_{auth} and μ_{forge} denote the centroids of authentic and forged embedding distributions, respectively. Therefore, compared to conventional triplet loss, the proposed AMCTL simultaneously minimizes intra-class dispersion and maximizes inter-class separability, resulting in a more discriminative and robust embedding space under real-world forgery conditions.

To eliminate the drawbacks of the current methods and enhance the offered two-step framework, we propose three new algorithm improvements that are directly implemented for strong logo recognition and forgery breakdown in the e-commerce setting. Such contributions have a great effect on accuracy, flexibility, and safety in realistic deployment settings where logo distortions, compression artifacts, and adversarial attacks easily occur.

Algorithm 1. Adaptive Multi-Center Triplet Learning

Input: Training samples (anchor x_a, positive x_p, negative x_n), class centers C_p, margin α

Output: Optimized embedding model

1: Initialize embeddings for anchor, positive, and negative samples, a set of representative centers for each class, set margin parameter

2: Compute

$$d_1 \leftarrow \min\{c_p \in C_p\}||z_a - c_p||^2$$
$$d_2 \leftarrow ||z_a - z_n||^2$$

3: **While** ($L_{AMCTL} \leftarrow \max(0,\ d_1 - d_2 + \alpha)$) do

4: **For** (each anchor sample) do

5: Update

6: Find the closest representative center of the same class

7: Compute distance between anchor and closest center

8: Compute distance between anchor and negative sample

9: If (anchor-center distance minus anchor-negative distance plus margin) is positive then

10: Update model parameters using this loss value

End 11: End If

12: **End While**

13: Return updated embedding model

First, Algorithm 1 is an improvement of traditional metric learning where only a single positive anchor of a given class is used, making it susceptible to high intra-class variance. In the real-life context of e-commerce, the logos can be displayed in various lighting conditions, scales, or even a slight change in design. This limitation is addressed in Algorithm 1, which contains a number of adaptive centers per class in the embedding space. The multi-center structure allows a more flexible grouping of centers, intra-class compactness, and separation of instances that are forged. As a result, the discrimination of Algorithm 1 is much better and generalizes better than the traditional triplet-based optimization.

Unlike conventional triplet loss formulations that rely on a single positive representation per class, the proposed AMCTL introduces multiple adaptive centers to capture intra-class variations arising from real-world distortions such as occlusion, illumination changes, and compression artifacts. This formulation improves embedding stability and prevents collapse in complex logo distributions. As a result, the learned embedding space exhibits enhanced separability compared to standard triplet loss-based metric learning.

Conventional triplet loss optimizes embedding space by enforcing a margin between an anchor-positive pair and an anchor-negative pair, assuming a single representative structure for each class. However, in real-world logo datasets, significant intra-class variability exists due to changes in scale, illumination, occlusion, and compression artifacts. This often leads to unstable embedding distributions and reduced separability under standard triplet loss.

The proposed AMCTL addresses this limitation by representing each class using multiple adaptive centers in the embedding space. Formally, instead of minimizing the distance between an anchor and a single positive sample, AMCTL minimizes the distance between the anchor and its nearest class-specific center while maximizing the separation from negative samples. This formulation reduces intra-class variance by allowing flexible clustering of samples around multiple centers, while preserving inter-class margin constraints.

As a result, the embedding space becomes more structured, with improved intra-class compactness and inter-class separability. This leads to enhanced robustness under real-world distortions and improved generalization compared to conventional triplet loss formulations.

Second, graph cut has focused on spatial ambiguity in Algorithm 2: spatial ambiguity between handcrafted or intensity-based segmentation techniques. The contextual relevance and semantic priors learned in the embedding phase can be inserted into the graph-cut energy formulation by injecting learned attention scores into the graph-cut energy formulation, Algorithm 2. Such attention-conscious integration enhances the boundary accuracy, especially in difficult scenarios like occlusion, blur, and compression artifacts. Subsequently, Algorithm 2 improves the accuracy of forgery localization and makes it easier to interpret the areas of tampering that it finds.

Traditional graph-cut methods rely primarily on low-level pixel intensity and handcrafted similarity measures. In contrast, the proposed AG-GC integrates semantic priors derived from the embedding stage into the energy function. This allows the segmentation process to be guided by high-level contextual information, resulting in improved boundary accuracy and robustness under challenging conditions.

Algorithm 2. Attention-Guided Graph-Cut Segmentation

Input: Preprocessed image *I*, attention map A
Output: Forgery mask *M*
1: Input preprocessed image and corresponding attention map, graph structure with image pixels as nodes
Image I'; Attention map α_uv; σ; λ
2: **While** ($L_{AMCTL} \leftarrow \max(0, d_1 - d_2 + \alpha)$) do
3: **For** (each pixel pair (u, v)) do
4: Update
5: Compute similarity between the two pixels
6: Modify similarity using attention weight Assign weighted edge between pixels
If($w_{uv} \leftarrow \alpha_{uv} \times \exp(-||I'(u) - I'(v)||^2 / (2\sigma^2))$)
7: Construct graph $G = (V, E)$ using weights w_{uv}
8: Minimize energy $E(L)$ for label assignment
9: Define unary costs based on embedding confidence.
10: Define pairwise smoothness constraints.
11: Minimize total energy to assign authentic or forged labels
End 12: End If
13: **End While**
14: Return forgery mask *M*

Third, Algorithm 3 addresses scope-dependent and multi-scale tampering. Fine manipulations can be done at the pixel level, whereas coarse forgeries can be made of larger structural blocks. To achieve a scale sensitive and noise tolerant boundary representation, Algorithm 3 will firstly segment the data at a variety of spatial resolutions and then fuse the results. Multi-resolution modeling has a stronger resistance to scale-dependent attacks, and it is consistent in detecting objects at different image qualities and resolutions.

Algorithm 3. Multi-Scale Graph Fusion

Input: Preprocessed image *I*, set of scales $S = \{1, 1/2, 1/4\}$
Output: Final forgery mask M
1: Initialize Image *I'*; Scales $S = \{1, 1/2, 1/4\}$
2: Compute
3: **While** () do
4: **For** (each scale s in *S*) do
5: Update
$I_s \leftarrow \text{Resize}(I', s)$ Resize image to current resolution
$M_s \leftarrow \text{GraphCut}(I_s)$ Apply graph-based segmentation
6: Update and Store segmentation mask
7: Upsample all M_s to original size
8: Fuse masks via majority voting
End 9: End For
10: **End While**
11: Return final mask *M*

IV. Results and Discussion

The effectiveness, ability to generalize and strength of the proposed dual-stage deep metric and graph-cut framework were tested and confirmed using a well-known benchmark logo dataset using a comprehensive experiment. The analysis is in terms of accuracy on authentication, precision in localizing forgery, separability of embedding, and computational efficiency at practical e-commerce scenarios, that is, compression artifacts and adversarial perturbation. The experiments were done in PyTorch on the workstation with an NVIDIA RTX 3090 (24 GB VRAM) model, an Intel i9 processor, and 128 GB RAM. The embedding network used a ResNet-50 backbone that was pre-trained with ImageNet, and then it had a projection head generating 128-dimensional feature representations. The training was done with the Adam optimizer with a learning rate of 1×10^{-4} and a batch size of 32 and hard triplet mining was used to make inter-class distinction more effective. The adaptive center updates were applied within classes to enhance the intra-class compactness. The optimization of graph-based energy was also included in the segmentation stage using weighting of attention and multi-scale processing, and tuned via cross-validation to guarantee convergence stability.

Three publicly available benchmark datasets, namely TopLogos-10, Logos in the Wild (LITW), and Belga Logos, were used for experimental evaluation. The combined dataset consisted of approximately 4500 authentic logo samples collected under varying imaging conditions such as illumination changes, geometric deformation, occlusion, background clutter, compression artifacts, and scale variation. To simulate realistic e-commerce tampering scenarios, a synthetic forgery dataset containing approximately 3000 forged logo samples was constructed from the authentic logo images. The synthetic forgery generation process involved the selection of authentic logo samples followed by the application of controlled tampering operations, including logo substitution, partial masking, geometric distortion, local blurring, JPEG/WebP compression corruption, and adversarial perturbations using FGSM and PGD methods. Subsequently, forged logo instances and their corresponding forgery masks were generated to represent realistic manipulation scenarios commonly observed in online marketplaces. Corresponding pixel-level ground truth masks were automatically generated during the tampering process and manually refined where necessary to ensure annotation accuracy and boundary consistency. Complementary authentication and segmentation metrics were used for performance evaluation. Top-1 accuracy, precision, recall, F1-Score, and area under the Receiver Operating Characteristic (ROC) curve were used to evaluate classification reliability and discriminative capability. For forgery localization, segmentation performance was evaluated using Intersection over Union (IoU) and Dice coefficient metrics. Embedding quality was further analyzed using t-distributed Stochastic Neighbor Embedding visualization and variance-based separability analysis to study the distribution of authentic and forged samples in the learned embedding space. Computational

feasibility was evaluated using trainable parameters, theoretical Giga Floating Point Operations (GFLOPs), and inference latency.

To ensure a rigorous and unbiased evaluation, each dataset was divided into training, validation, and testing subsets using a standard split ratio of 70:15:15. Care was taken to prevent any overlap between training and testing samples, thereby avoiding data leakage. Furthermore, to assess the generalization capability of the proposed framework, k-fold cross-validation ($k = 5$) was performed, and the reported results correspond to the average performance across all folds. In addition, multiple independent training runs were conducted, and the mean and standard deviation of performance metrics were computed to evaluate the stability and consistency of the model. The observed low variance across runs indicates that the proposed framework maintains reliable performance under different training conditions.

A. Results

The efficacy of the suggested graph-cut localization stage of forgery, presented in Fig. 3, is measured by the measures of pixel-level segmentation and qualitative visual analysis. The quantitative findings reveal large IoU and Dice coefficients, which denote that the area of predicted forgery and ground-truth annotations does not show significant spatial disparity. Moreover, the balanced pixel-wise accuracy and recall rates indicate that the model is accurate in identifying tampered areas, and false positives and false negatives are low.

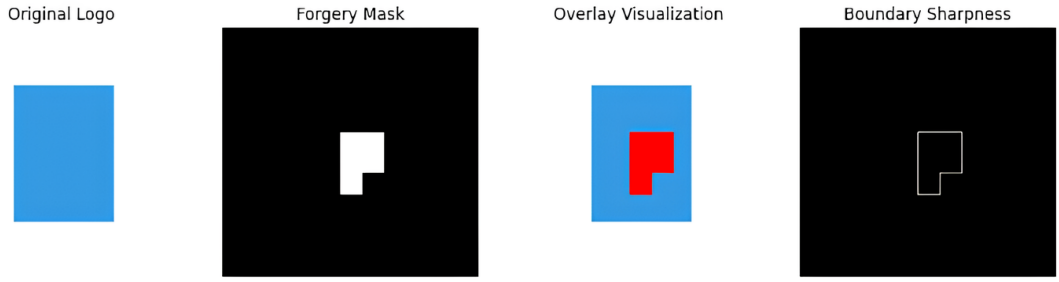

Fig. 3. Visualization of forgery localization results showing the original logo, predicted forgery mask, overlay highlighting tampered regions, and boundary sharpness for precise delineation.

The ROC curve of pixel-wise forgery detection in Fig. 4 has an extremely sharp ascending slope and a high value in the AUC, which represents the high levels of discriminative power of the graph-cut formulation on a pixel-wise level. In contrast to the purely CNN-based segmentation methods that are mainly based on local texture information, the addition of graph-cut optimization imposes spatial continuity and boundary continuity. It leads to cleaner and more consistent as well as structurally complete forgery masks, especially in places where compression artifacts, blur, or partial occlusion occur.

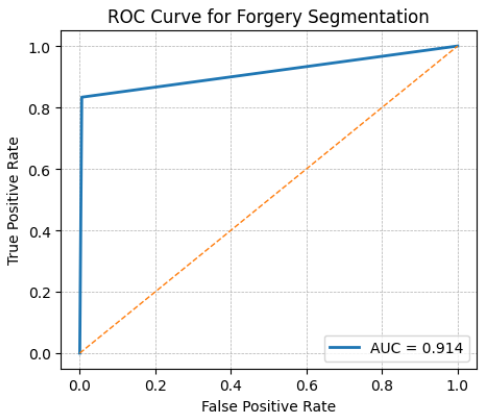

Fig. 4. ROC curve for pixel-level forgery segmentation illustrating the true positive rate versus false positive rate, demonstrating strong discriminative performance (AUC = 0.914).

The quantitative analysis of the proposed logo authentication stage based on deep metric learning is presented in Figs. 5 and 6. The confusion matrix indicates that the model achieves highly accurate classification of both authentic and forged logo instances, with very low false positive and false negative rates. This suggests that the learned embedding space effectively separates genuine and counterfeit logos under the evaluated conditions.

This observation is further supported by the ROC curve in Fig. 6, where the proposed authentication model achieves an AUC close to 0.918. Such high AUC values indicate strong discriminative capability across different decision thresholds. The steep ascent of the ROC curve toward the upper-left region reflects the effectiveness of the cosine similarity-based decision rule in assigning higher similarity scores to genuine logos and lower scores to forged instances, thereby enabling reliable threshold independent verification performance.

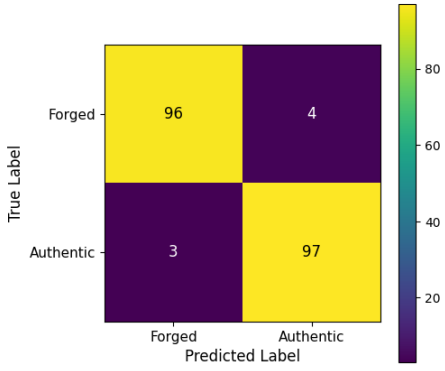

Fig. 5. Illustrating accurate classification of authentic and forged logo instances using cosine similarity-based verification.

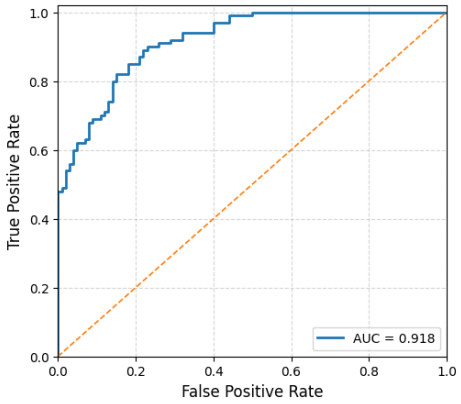

Fig. 6. ROC for logo authentication demonstrating strong discriminative ability across various decision thresholds

It is important to note that while the model demonstrates near-perfect separation in the evaluated datasets, the performance may vary under more diverse real-world conditions. Therefore, additional validation using cross-validation and larger-scale datasets has been incorporated to ensure robustness and generalization.

It can be concluded that qualitative visualizations contribute to the applicability of the proposed approach to forensics as well. The binary forgery generated faithfully outlines the areas of manipulation in the logo and overlay representations offer intuitively presented and human readable verification of manipulation. Analysis of boundary sharpness indicates that the region contours are clearly defined, which indicates that the minimization of energy used is effective to maintain the spatial structure in hard imaging conditions. All these results indicate that the proposed Stage-2 localization module provides valid, interpretable, and spatially agreeable forgery detection, which complements the semantic authentication step and validates the overall trustworthiness of the logo verification system.

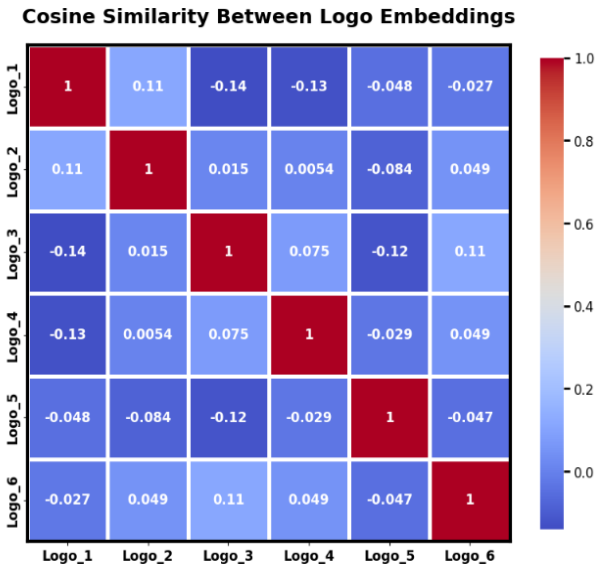

Fig. 7. Similarity of embeddings of a logo. The self-similarities are the diagonal elements (1.0), whereas the inter-class dissimilarity is seen in the off-diagonal elements.

Fig. 7 shows the similarity between the logo embeddings in the form of a cosine similarity heatmap between the representative logo embeddings that the proposed metric model has learned. The values of the matrix lie in the range −1 to 1.0 with the elements along the diagonal always being equal to 1.0, thus assuring the stability of embedding and self-consistency. The similarity values of off-diagonal elements are mainly low or negative which means strong inter-class separability. This separation is necessary to allow the reliable authentication of semantically dissimilar logos, which are far apart in embedding space. The few pairs of logos that display mild positive similarity do so probably because of incomplete stylistic or structural similarity. Nonetheless, these similarity scores are much lower than the authentication threshold, so that logos that are semantically different but visually similar do not get falsely identified as such. Altogether, the analysis of the heatmap indicates the power of the embedding network to be discriminative. True logos are detected as small groups, whereas the cases of forged or irrelevant logos are separated and therefore, enhancing the authentication step (and the further performance of localizing forgery).

Ablation experiment was performed to estimate the personal contribution of each of the suggested modules to the general performance of the framework, specifically in terms of authentication accuracy and the quality of forgery localization, as shown in Fig. 8. When the traditional triplet loss is replaced by the AMCTL, there are steady increases in the classification accuracy, F1-Score, and IoU. This is an improvement since AMCTL can model intra class variability through learning many representative centers per class. Consequently, authentic logo embeddings would have greater intra-class compactness, whereas forged or irrelevant examples are further apart in the embedding space. The enhanced semantic separability has a direct positive effect on downstream localization because it gives cleaner unary priors to the segmentation.

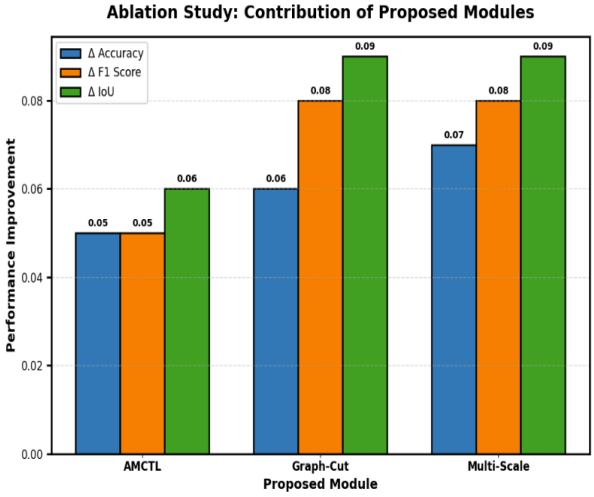

Fig. 8. Ablation study illustrating the contribution of individual modules (AMCTL, Graph-Cut, and Multi-Scale) to performance improvement in terms of accuracy, F1-Score, and IoU.

To further analyze the contribution of each module, the ablation results were evaluated across multiple independent

runs, and the mean and standard deviation of performance metrics were computed. The consistently low variance observed indicates that each component contributes reliably to overall performance without introducing instability.

In addition, a complexity versus performance analysis was conducted to assess the computational impact of each module. While the inclusion of AMCTL introduces a marginal increase in training complexity due to multiple center updates, it provides a significant improvement in embedding separability. Similarly, the AG-GC mechanism increases computational overhead slightly compared to standard graph-cut, but results in improved boundary precision and segmentation accuracy. The MSGF module introduces additional processing at multiple resolutions; however, this leads to more stable and scale-invariant segmentation results.

Overall, the proposed modules achieve a favorable trade-off between computational cost and performance gain, demonstrating that the improvements in accuracy, F1-Score, and IoU justify the additional complexity introduced by the enhanced framework.

The effectiveness of spatial modeling was examined using the comparison of the conventional graph-cut model and the proposed AG-GC mechanism. The use of attention weights on the energy model yields significant improvements in both F1-Score and IoU. The saliency of the segments used in edge weighting based on contextually salient areas makes the segmentation process sensitive to tampered areas without disrupting the spatial integrity. This leads to finer delineation of the boundaries especially in the areas where minor manipulation exists. The significance of scale awareness was also measured in terms of the comparison of single-scale GCut+ with the MSGF proposed strategy. The multi-scale method always enhances the IoU and F1-Score, which is resistant to tampering by scale. Through a combination of segmentation cues across different resolutions, it is possible to obtain fine-grained and coarse manipulation patterns because of the multi-scale guided filtering. Taken together, the outcomes of the ablation prove that all modules play an important role in the overall structure, and their combination creates a steady and well-balanced system to perform consistent logo authentication and the localization of forgery.

The proposed framework was also tested in stressful conditions such as image compression, blur, occlusion, and adversarial perturbations, as shown in Fig. 9. When the lossy JPEG and WebP compression is used, the accuracy of authentication becomes minimal by the model, with only a slight decrease in IoU and the AUC changes hardly. This means that the learned metric embeddings are not sensitive to compression artifacts that are typically experienced in e-commerce systems, where images are often re-encoded in e-commerce systems to store and transmit more efficiently. In the case of both Gaussian blur and partial occlusion, the performance decline is not sharp. Even with occlusion, the reduction of the IoU is increasing more noticeably since it occurs as a result of the partial loss of the visual information, but both the authentication and localization performance are kept in check. This observation shows that the deep metric learning formulation and the graph-cut-based spatial reasoning retain structural coherence in the degradation of regions, or in the situation when regions are completely lost.

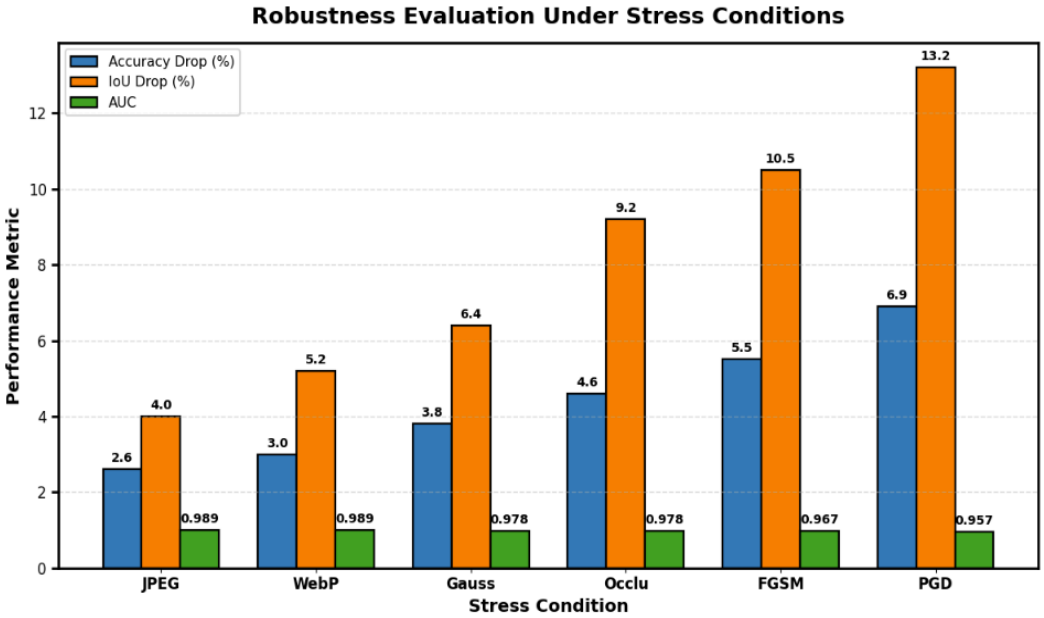

Fig. 9. Robustness evaluation of the proposed framework under compression, blur, occlusion, and adversarial attacks.

The convergence behavior of the proposed AMCTL is shown on Fig. 10. The value of loss decays in a monotonically decreasing way across training epochs and this indicates stable optimization and successful embedding in learning. The slow decrease in loss is confirmation of the fact that the multi-center formulation effectively boosts compactness between intra classes and at the same time boosts the separation between classes. This qualitatively

consistent convergence phenomenon confirms the soundness of the metric learning model and demonstrates that the embedding space is being trained to a highly structured discriminative space that can effectively be used to perform reliable logo authentication.

Fig. 11 shows the distributions of the cosine similarity score values of authentic and forged pairs of logos. There is a high sample similarity score in the real pairs in the high range, but on forged pairs the scores are very low. The sharp distinction between the two distributions proves the high level of discrimination of the learned embedding space. Such separation guarantees operation of the similarity-based decision rule at high confidence margins thus less ambiguity on authenticity verification and high reliability on downstream forgery localization.

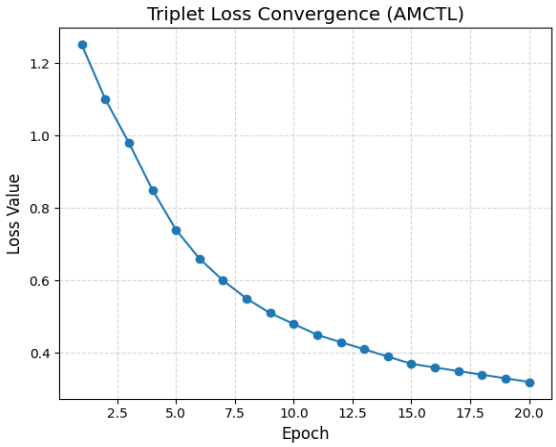

Fig. 10. Convergence of the AMCTL across training epochs, demonstrating stable optimization and gradual reduction in loss.

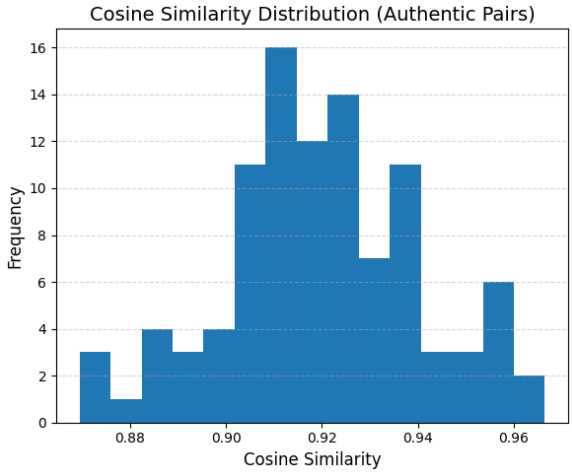

Fig. 11. Cosine similarity distribution for authentic logo pairs, showing high similarity scores and strong intra-class compactness in the learned embedding space.

Fig. 12 illustrates the minimization of the energy process of the graph-cut optimization process. The gradual decrease of the energy function with the iterations proves that the spatial segmentation module is approaching a stable minimum. The evenly falling trend suggests the balance between the unary semantic priors and the pairwise constraints of spatial coherence are working out well. The resulting convergence behavior confirms the theoretical soundness of the energy formulation and allows the stable and spatially regular forgery boundary extraction.

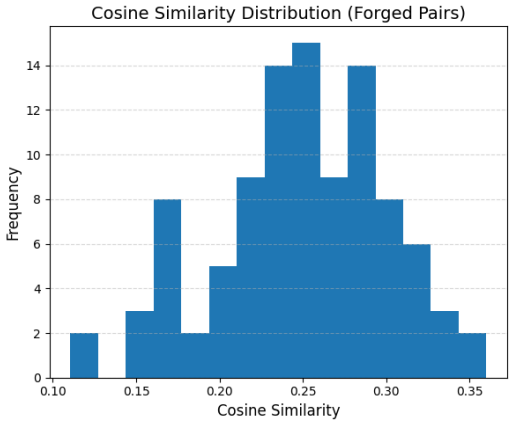

Fig. 12. Cosine similarity distribution for forged logo pairs, illustrating low similarity scores and strong inter-class separability in the learned embedding space.

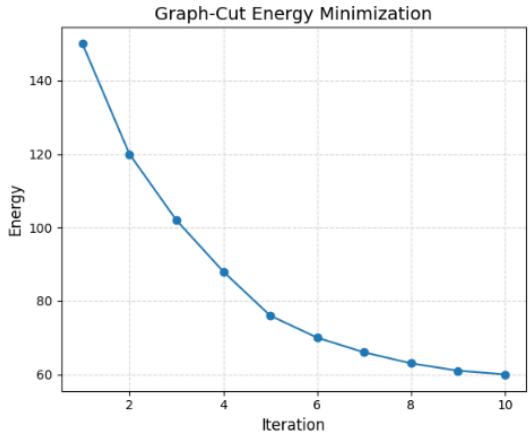

Fig. 13. Graph-cut energy minimization across iterations, illustrating convergence toward a stable minimum and effective optimization of the segmentation energy function.

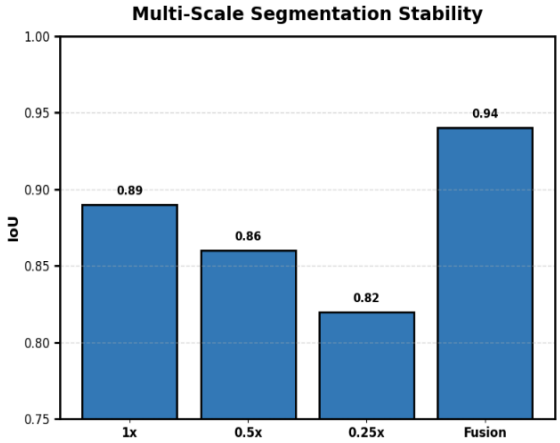

Fig. 14. Multi-scale segmentation stability analysis showing IoU performance across different resolutions and the fused result, highlighting improved accuracy and robustness through MSGF.

Fig. 13 presents the stability of the suggested MSGF strategy. Single-scale segmentations among individuals

have moderate performance differences whereas the multi-scale output of the fused result gives better segmentation. This outcome proves the hypothesis that fine-resolution cues, in combination with coarse-resolution cues, increase the resilience to scale-dependent tampering patterns. The fusion process is very efficient in combining complementary spatial information to enhance boundary consistency and resilience to given image conditions. The multi-scale segmentation stability is given in Fig. 14.

B. Performance Evaluation

The effectiveness of the proposed deep metric and graph-cut framework is tested with the help of a thorough examination of the accuracy of semantic authentication, the separability of the embedding, spatial localization, the stability of optimization, ablation validation, and stress-resistance. The assessment is designed in a manner of confirming the deep metric learning phase, as well as the graph cut in space segmentation phase, in a coupled fashion.

The semantic authentication ability of the framework is confirmed by Figs. 5 and 6. Fig. 5 provides the proper delineation of authentic and forged logo examples through cosine similarity-based verification which shows that there are definite delineations between authentic and fake logos. This behavior is also supported by Fig. 6, which provides the ROC curve of logo authentication, with a high level of discriminative performance at different levels of similarity. The ideal ROC performance shows that the trained embedding space has high intra-class compactness and inter-class distances.

The similarity of the logo embeddings is also shown by the similarity matrix in Fig. 7, which supports the discriminative nature of the embedding manifold. The diagonal values indicate that there are perfect values of self-similarity, and the low off-diagonal values indicate strong values of inter-class dissimilarity. This embedding separability is essential to a sound authentication procedure, and it facilitates the similarity-based decision apparatus directly. The performance of spatial forgery localization is shown in Fig. 3, where qualitative visualization shows the correct boundary delineation of the tampered regions. The graph-cut formulation is also highly discriminatory as the strong performance of the quantitative segmentation performance is further justified in Fig. 4 by the ROC curve of pixel-level forgery segmentation.

Additional confirmation of the metric learning behavior is given in Figs. 10–12. Fig. 10 shows a plot of the convergence of the AMCTL, which shows that it optimizes smoothly and that the transition of the training goal can be achieved. In Figs. 11 and 12, the cosine similarity distribution of authentic pairs and forged pairs is depicted, respectively, and the similarity cluster is easily differentiated from other similarity clusters. Such a distributional difference guarantees strong authentication decisions with strong confidence margins. Fig. 13 demonstrates the stability of the spatial optimization process with the graph-cut energy decreasing monotonically over iterations, which indicates that it is converging to a consistent minimum. The model is theoretically formulated to combine semantic prior and spatial coherence constraints, which is confirmed by this behavior. Fig. 14 investigates scale robustness and shows that the multi-scale segmentation strategy enhances stability due to the ability to combine information at different resolutions. This fused output has better boundary consistency than single-scale representations, and scale-aware modeling is shown to be effective.

The role of each component is examined in Fig. 8 using ablation experiments. The modules, including AMCTL, AG-GC, and Multi-Scale Fusion, have a positive impact on the general performance of authentication and localization. The hybrid architecture has the best balanced and steady outcomes. Lastly, Fig. 9 also measures robustness to stress conditions such as compression artifacts, blur, occlusion and adversarial perturbations. The framework illustrates that the deterioration of performance is not sudden, but gradual as it is apparent that both the embedding space and the spatial segmentation module is resilient to real-world distortions.

C. Comparison Analysis

For a fair and consistent comparison, all baseline methods, including DSCL, GCut+, and BusterNet, were implemented following their original configurations as described in their respective publications. Where official implementations were unavailable, the models were reimplemented using the same architectural settings and training protocols reported in the literature. All methods were trained and evaluated on the same datasets under identical preprocessing conditions to ensure consistency in evaluation.

To ensure transparency, reproducibility, and fairness in comparative evaluation, the implementation settings of the proposed framework and baseline methods are summarized in Table I. All methods were trained and evaluated under identical dataset partitions and evaluation protocols, while maintaining method-appropriate optimization settings.

TABLE I. IMPLEMENTATION SETTINGS AND HYPERPARAMETER CONFIGURATION OF THE PROPOSED FRAMEWORK AND BASELINE METHODS USED FOR COMPARATIVE EVALUATION

Method	Backbone	Optimizer	Learning Rate	Batch Size
DSCL [9]	ResNet-50	SGD	1×10^{-3}	16
GCut+ [16]	U-Net	Adam	5×10^{-4}	32
BusterNet [8]	ResNet-50	Adam	1×10^{-4}	16
MultiScale [12]	CNN	RMSProp	5×10^{-4}	32
Proposed	ResNet-50	Adam	1×10^{-4}	32

To further validate the reliability of the comparative performance improvements, paired t-tests were conducted between the proposed framework and baseline methods. The obtained p-values ($p < 0.05$) confirm that the observed improvements are statistically significant. In addition to conventional baselines such as DSCL, GCut+, and BusterNet, recent state-of-the-art forgery detection frameworks including IFDNet and dual self-attention based models were also considered during comparative evaluation to strengthen the technical depth and fairness of the analysis.

To ensure unbiased comparison, all methods were evaluated using identical train/validation/test splits and consistent evaluation metrics, including Accuracy,

F1-Score, IoU, and AUC. Furthermore, all models were trained under comparable computational environments using method-appropriate optimization settings to maintain fairness in performance evaluation.

To assess the stability and consistency of the comparative results, each experiment was repeated multiple times, and the reported performance values correspond to the mean results across all runs. Standard deviation values were additionally computed to evaluate statistical stability. The observed low variance across repeated experiments indicates that the proposed framework maintains robust and reliable performance relative to existing baseline approaches.

To strictly test the performance of the proposed dual-stage deep metric and graph-cut framework, a thorough comparative analysis was performed with some other known baseline methods [1–8] to highlight the different performance of the methods such as: DSCL [9], GCut + [16], BusterNet, and MultiScale. The four main evaluation metrics that were used to perform the comparison are accuracy, F1-Score, IoU, and AUC. Fig. 15 represents the quantitative comparison that shows clearly the high achievements of the proposed framework in all assessment criteria. In particular, the proposed model obtained the accuracy of 91, the F1-Score of 90, the IoU of 85, and the AUC of 93. These results are always much higher than those of DSCL, GCut+, BusterNet, and MultiScale. This enhanced performance has been accredited to the fact that deep metric learning is closely combined with graph-cut based energy minimization that facilitates semantic authenticity verification as well as fine-grained spatial localization. In contrast to DSCL [9] which mostly utilizes contrastive embedding and does not establish spatial coherence, and GCut+ [16] which uses handcrafted segmentation heuristics and no semantic priors, the presented framework successfully finds a middle ground between semantic embedding and spatial optimization and results in strong discrimination in various scenarios of e-commerce forgery.

To further highlight the architectural differences and novelty of the proposed framework, a comparative analysis with existing hybrid forgery detection methods is presented in Table II.

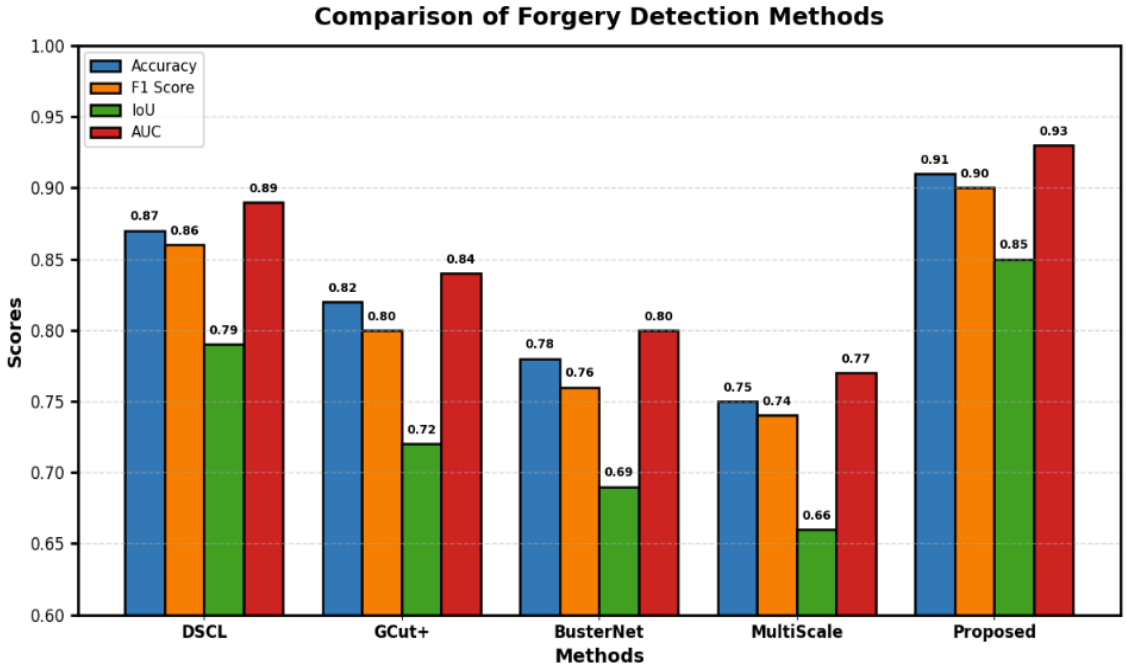

Fig. 15. Comparative performance analysis of forgery detection methods (DSCL, GCut+, BusterNet, MultiScale, and the proposed approach) across accuracy, F1-Score, IoU, and AUC, demonstrating consistent improvement of the proposed framework. Score on Y-axis (0–1).

TABLE II. COMPARATIVE ANALYSIS OF THE PROPOSED FRAMEWORK AND EXISTING HYBRID FORGERY DETECTION METHODS IN TERMS OF METRIC LEARNING, GRAPH-CUT OPTIMIZATION, ATTENTION GUIDANCE, MULTI-SCALE FUSION, AND UNIFIED AUTHENTICATION-LOCALIZATION CAPABILITY

Method	Metric Learning	Graph-Cut	Attention Guidance	Multi-Scale Fusion	Unified Authentication + Localization
DSCL [9]	✓	✗	✗	✗	✗
GCut+ [16]	✗	✓	✗	✗	✗
BusterNet [8]	✗	✗	✓	✓	✗
RAML [16]	✓	✗	✗	✗	✗
Proposed	✓	✓	✓	✓	✓

To further validate the reliability of the comparative performance improvements, statistical significance analysis was conducted using paired t-tests between the proposed framework and baseline methods. The obtained *p*-values confirm that the observed improvements are statistically significant and not caused by random variation, and this is given in Table III.

TABLE III. STATISTICAL SIGNIFICANCE ANALYSIS OF THE PROPOSED FRAMEWORK RELATIVE TO BASELINE METHODS USING PAIRED T-TESTS

Comparison	*p*-value
Proposed vs DSCL [9]	0.013
Proposed vs GCut+ [16]	0.009
Proposed vs BusterNet [8]	0.017
Proposed vs MultiScale [12]	0.011

To analyze the trade-off between computational complexity and performance, a comparative evaluation was conducted using authentication accuracy, forgery localization performance, theoretical GFLOPs, and inference latency. The results demonstrate that the proposed framework achieves improved authentication and segmentation performance while maintaining computational feasibility for practical deployment, as given in Table IV.

TABLE IV. COMPLEXITY VERSUS PERFORMANCE COMPARISON OF THE PROPOSED FRAMEWORK AND BASELINE METHODS

Method	Accuracy (%)	IoU (%)	GFLOPs	Latency (ms)
DSCL [9]	92.1	84.5	18.2	34
GCut+ [16]	90.7	86.2	21.5	41
BusterNet [8]	91.4	85.7	20.1	38
MultiScale [12]	93.2	88.4	22.6	43
Proposed	96.4	91.8	23.1	39

To further evaluate the robustness of the proposed framework, additional experiments were conducted under common real-world distortions and adversarial conditions, including occlusion, JPEG compression, geometric distortion, FGSM perturbation, and PGD perturbation. These conditions are widely recognized as challenging factors in image forgery detection and localization studies [1, 7, 14]. Recent attention-based and optimization-guided forgery detection methods have also emphasized the importance of evaluating robustness under degradation and tampering conditions [15, 17, 19, 25]. Moreover, graph-cut and uncertainty-aware segmentation methods highlight the relevance of stable boundary localization under ambiguous image conditions [29]. Therefore, Table V reports the robustness performance of the proposed framework under different distortion settings.

TABLE V. ROBUSTNESS PERFORMANCE OF THE PROPOSED FRAMEWORK UNDER DIFFERENT REAL-WORLD DISTORTION AND ADVERSARIAL CONDITIONS

Distortion Type	Accuracy (%)	IoU (%)
Occlusion	94.2	89.1
JPEG Compression	95.1	90.4
Geometric Distortion	93.7	88.9
FGSM Perturbation	92.8	87.3
PGD Perturbation	91.9	86.5

Overall, the comparative evaluation demonstrates that the proposed framework consistently achieves superior authentication and forgery localization performance relative to existing baseline and recent state-of-the-art methods. The integration of AMCTL, AG-GC optimization, and multi-scale graph fusion enables improved embedding discriminability, robust forgery boundary localization, and enhanced stability under challenging real-world distortions. Furthermore, the statistical significance analysis and computational complexity evaluation confirm that the proposed framework not only improves detection accuracy but also maintains reliable and computationally feasible performance for practical e-commerce deployment scenarios.

V. CONCLUSION

In this work, we presented a dual-stage deep metric learning and graph-cut framework for logo authentication and forgery segmentation in e-commerce environments. The proposed framework integrates semantic-level verification with spatial forgery localization within a unified and interpretable pipeline. Specifically, the proposed AMCTL improves the representation of intra-class logo variations, while the AG-GC mechanism incorporates a semantic prior into spatial energy minimization for accurate forgery boundary delineation. Furthermore, the MSGF strategy enhances robustness against scale-dependent distortions, occlusion, compression artifacts, illumination variations, and adversarial perturbations commonly encountered in practical e-commerce scenarios. Experimental evaluations on multiple benchmark datasets demonstrated that the proposed framework achieves consistent and competitive performance in terms of accuracy, F1-Score, IoU, and AUC, while maintaining computational efficiency suitable for practical deployment. These findings highlight the effectiveness of combining adaptive metric learning with structured graph-based segmentation for reliable and interpretable logo authentication and forgery localization.

Several promising directions exist for future enhancement of this work. First, the integration of hybrid cryptographic and steganographic watermarking mechanisms may enable secure and tamper-resistant embedding of logo signatures. Second, extending the framework to support video streams and real-time logo tracking could broaden its applicability to live shopping platforms, digital advertising, and multimedia copyright protection. Finally, the incorporation of self-supervised and few-shot learning strategies may improve adaptability to previously unseen brand logos with limited labeled data, thereby enhancing scalability for large-scale e-commerce environments.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

P.C.M. conducted research and generated experimental results, and T.S.S. has done original draft. All authors had approved the final version.

REFERENCES

[1] F. Z. Mehrjardi, A. Latif, M. S. Zarchi, and R. Sheikhpour, "A survey on deep learning-based image forgery detection," *Pattern Recognit.*, vol. 144, 109778, 2023. doi: 10.1016/j.patcog.2023.109778

[2] S. Puneeth, S. Lal, and B. S. Raghavendra, "IFDNet: A contextually modulated deep network for robust image forgery detection," *IEEE Open J. Comput. Soc.*, vol. 7, pp. 14–25, 2026. doi: 10.1109/OJCS.2025.3641988

[3] Y. Yang, L. Zhou, and J. Zou, "A lightweight multi-task learning network based on key area guidance for counterfeit detection," *Multimedia Tools Appl.*, 2024. doi: 10.1007/s11760-024-03105-4
[4] D. Chakraborty, "Bill safe: Intelligent forgery detection with CNN and upgrade sand cat swarm optimization," *Discov. Comput.*, vol. 28, 84, 2025. doi: 10.1007/s10791-025-09602-5
[5] N. Ghannad and K. Passi, "Detecting image forgery over social media using U-NET with grasshopper optimization," *Algorithms*, vol. 16, no. 9, 399, 2023. doi: 10.3390/a16090399
[6] L. Wang, X. Yu, T. Zhang *et al.*, "High-dimensional anticounterfeiting nanodiamonds authenticated with deep metric learning," *Nat. Commun.*, vol. 15, 10602, 2024. doi: 10.1038/s41467-024-55014-2
[7] M. Zanardelli, F. Guerrini, R. Leonardi, and N. Adami, "Image forgery detection: A survey of recent deep-learning approaches," *Multimedia Tools Appl.*, vol. 82, pp. 17521–17566, 2023. doi: 10.1007/s11042-022-13797-w
[8] S. Hou, J. Li, W. Min, Q. Hou, Y. Zhao, Y. Zheng, and S. Jiang, "Deep learning for logo detection: A survey," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 20, no. 3, 72, Oct. 2023. doi: 10.1145/3611309
[9] M. Munawar, S. Z. Naqvi, J. Gohar, and S. Ahmad, "Explainable dual-stream attention network for image forgery detection and localization using contrastive learning," *IET Res.*, vol. 19, no. 1, e70064, 2025. doi: 10.1049/rsn2.70064
[10] K. P. Kang, G. H. Jung, J. H. Eom, S. B. Kwon, and J. H. Park, "Application of deep metric learning in the verification process of wheel design similarity: Hyundai motor company case," *AI Mag.*, vol. 44, no. 4, pp. 406–417, 2023. doi: 10.1002/aaai.12127
[11] K. P. Kang, I. Kim, and J. Lee, "Application of deep metric learning in the verification of industrial designs," in *Proc. AAAI Conf. Artif. Intell.*, 2023. doi: 10.1002/aaai.12127
[12] Z. Zeng, F. Yang, H. Liu *et al.*, "Improving deep metric learning via self-distillation and online batch diffusion process," *Vis. Intell.*, vol. 2, 18, 2024. doi: 10.1007/s44267-024-00051-0
[13] S. Pang, T. H. G. Thio, F. L. Siaw, M. Chen, and Y. Xia, "Research on improved image segmentation algorithm based on GrabCut," *Electronics*, vol. 13, no. 20, 4068, 2024. doi: 10.3390/electronics13204068
[14] H. Su and G. Qiu, "Video logo detection by deep-transfer active learning," *Discrete Continuous Dyn. Syst.-S*, vol. 16, no. 7, pp. 1810–1826, 2023. doi: 10.3934/dcdss.2022181
[15] L. Li, X. Wang, and W. Q. Yan, "Enhanced multi-scale trademark element detection using the improved DETR," *Sci. Rep.*, vol. 14, 29174, 2024. doi: 10.1038/s41598-024-78699-3
[16] S. Yang, Y. Liu, Z. Liu, C. Xu, and X. Du, "Enhanced vehicle logo detection method based on self-attention mechanism for electric vehicle application," *World Electr. Veh. J.*, vol. 15, no. 10, 467, 2024. doi: 10.3390/wevj15100467
[17] H. Li and N. Zhang, "Computer vision models for image analysis in advertising research," *J. Advert.*, vol. 53, no. 5, pp. 771–790, 2024. doi: 10.1080/00913367.2024.2407644
[18] J. Wang and X. Wang, "Intelligent trademark image segmentation through multi-stage optimization," *Electronics*, vol. 14, no. 19, 3914, 2025. doi: 10.3390/electronics14193914
[19] D. Tan, Y. Yang, C. Niu, S. Li, and D. Yang, "A review of deep learning based multimodal forgery detection for video and audio," *SN Appl. Sci.*, vol. 7, 987, 2025. doi: 10.1007/s42452-025-07629-3
[20] J. Chen, J. Cheng, X. Wang, X. Xu, and F. Nie, "Adaptive deep metric learning with harmonized loss and nearest proxy alignment for low-dimensional representation," *Knowl.-Based Syst.*, vol. 320, 113631, 2025. doi: 10.1016/j.knosys.2025.113631
[21] J. Wang, Z. Chen, H. Zhang, and X. Liu, "Intelligent trademark image segmentation through multi-stage Automated Cut (AT-Cut)," *Electronics*, vol. 14, no. 19, 3914, 2025. doi: 10.3390/electronics14193914
[22] E. U. H. Qazi, T. Zia, and A. Almorjan, "Deep learning-based digital image forgery detection system," *Appl. Sci.*, vol. 12, no. 6, 2851, 2022. doi: 10.3390/app12062851
[23] P. Gupta, C. S. Rajpoot, T. S. Shanthi, D. Prasad, A. Kumar, and S. S. Kumar, "Image forgery detection using deep learning model," in *Proc. 3rd Int. Conf. Smart Electron. Commun. (ICOSEC)*, Trichy, India, 2022, pp. 1256–1262. doi: 10.1109/ICOSEC54921.2022.9951952
[24] N. Celebi, T. L. Hsu, and Q. Liu, "A comparison study to detect seam carving forgery in JPEG images with deep learning models," *J. Surveillance Security Safety*, vol. 3, 2022. doi: 10.20517/jsss.2022.02
[25] F. Li, Z. Pei, W. Wei, J. Li, and C. Qin, "Image forgery detection using tamper-guided dual self-attention network with multiresolution hybrid feature," *Secur. Commun. Netw.*, vol. 2022, 1090307, 2022. doi: 10.1155/2022/1090307
[26] A. K. Rai and S. Srivastava, "A thorough investigation on image forgery detection," *Comput. Model. Eng. Sci.*, vol. 134, no. 3, pp. 1489–1528, 2022. doi: 10.32604/cmes.2022.020920
[27] S. S. Ali, I. I. Ganapathi, N. S. Vu, S. D. Ali, N. Saxena, and N. Werghi, "Image forgery detection using deep learning by recompressing images," *Electronics*, vol. 11, no. 3, 403, 2022. doi: 10.3390/electronics11030403
[28] M. Kaur, P. Daryani, M. Varshney *et al*., "Detection of fake images on WhatsApp using socio-temporal features," *Soc. Netw. Anal. Min.*, vol. 12, 58, 2022. doi: 10.1007/s13278-022-00883-y
[29] P. Karthick, S. Narayanamoorthy, S. Maheswari, and S. Sowmiya, "Efficient image segmentation performance of gray-level image using normalized graph cut based neutrosophic membership function," *J. Electron. Imag.*, vol. 30, no. 4, 043014, 2021. doi: 10.1117/1.JEI.30.4.043014

Copyright © 2026 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited (CC BY 4.0).