Int. J. Sci. R. Tech., 2026 3(9)
A Multidisciplinary peer-reviewed Journal
www.ijsrtjournal.com [ISSN: 2394-7063]

Original Article

A Comparative Benchmark Of Handcrafted Local Descriptors And Deep Metric Learning Architectures For Image-Based Counterfeit Identification

Harini B. R.*

Dept of Computer Applications, Sri K Puttaswamy First Grade College, VVCE Campus, Gokulam, Mysore-02

ABSTRACT

Counterfeit products and documents are becoming increasingly difficult to identify because modern copies may contain only small visual differences from genuine items. Image-based computer vision provides a practical approach for detecting such differences without relying entirely on manual inspection. This paper presents a comparative framework for studying traditional handcrafted image descriptors and modern deep learning approaches for counterfeit identification. Three local feature methods, namely Scale-Invariant Feature Transform (SIFT), Speeded-Up Robust Features (SURF), and Oriented FAST and Rotated BRIEF (ORB), are considered alongside deep convolutional models such as ResNet and MobileNet. A Siamese Neural Network with contrastive loss is also included to support image verification when only a limited number of counterfeit samples are available. The proposed framework considers recognition performance, computational cost, robustness to image variations, data requirements, and practical deployment. The study provides a common methodology for understanding the strengths and limitations of different approaches and can support the development of efficient counterfeit identification systems.

Keywords: Counterfeit Detection, Image Forensics, Image Matching, SIFT, SURF, ORB, CNN, ResNet, MobileNet, Siamese Neural Network, Deep Metric Learning.

INTRODUCTION

Counterfeit products, documents, certificates, currency, labels, and commercial packages create significant problems for organizations and consumers. Detecting counterfeit items is challenging because high-quality replicas may differ from genuine items only in small patterns, printing details, texture, alignment, or surface characteristics. Traditional inspection generally depends on human expertise or specialized equipment, which may not always be practical for large-scale verification.

Computer vision and image processing provide an alternative by allowing visual characteristics to be extracted and compared automatically. Earlier image-matching systems mainly relied on handcrafted local descriptors. SIFT, SURF, and ORB are widely used because they can identify distinctive points and represent local image regions. These methods can perform effectively when important visual characteristics remain visible under changes in scale, rotation, or viewpoint.

Deep learning provides another approach. Convolutional Neural Networks (CNNs) can automatically learn hierarchical image representations from training data. Architectures such as ResNet provide strong image representations, while MobileNet is designed with lower computational requirements and is suitable for resource-constrained devices.

However, counterfeit detection often faces a data limitation because obtaining a large and balanced collection of genuine and counterfeit samples can be difficult. Deep metric learning provides a useful alternative by learning whether two images are similar rather than requiring every counterfeit type to be represented as a separate class. Siamese Neural Networks are particularly suitable for this type of image verification.

Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

This study therefore proposes a common framework for comparing handcrafted descriptors, deep CNN features, and Siamese metric learning for image-based counterfeit identification.

1.1 Problem Statement

Counterfeit identification systems must handle differences caused by illumination, scale, rotation, blur, image quality, printing variation, and physical damage. Handcrafted methods can be computationally efficient but may not capture complex visual patterns. Deep CNN models can learn powerful representations but generally require greater computational resources and suitable training data.

Another challenge is that counterfeit categories can change over time. A classification system trained on fixed categories may not easily recognize previously unseen counterfeit patterns. A verification-based metric learning approach can address this limitation by comparing the similarity between a test image and a reference image.

Therefore, a systematic comparison of these approaches is required to understand their accuracy, computational requirements, robustness, and suitability for real-world counterfeit verification.

2. Research Contributions

The major contributions of this study are:

1. A common framework for comparing handcrafted, deep feature-based, and metric learning approaches.
2. Evaluation of SIFT, SURF, and ORB for local feature-based image matching.
3. Investigation of ResNet and MobileNet as deep feature extractors.
4. Integration of a Siamese Neural Network for image-pair verification.
5. Use of contrastive loss to learn similarity between genuine and counterfeit image pairs.
6. Consideration of accuracy, computational cost, robustness, data requirements, and deployment suitability.

3. Literature Review

Lowe introduced SIFT for detecting distinctive image features that remain relatively stable under scale and rotation changes [1]. SIFT has been widely applied to image matching and object recognition. Bay et al. proposed SURF as a faster alternative using efficient feature computation and integral images [2]. Rublee et al. developed ORB by combining FAST keypoint detection with a rotated BRIEF descriptor, providing a computationally efficient alternative to SIFT and SURF [3].

Traditional image-processing methods remain useful when computational resources are limited. However, their performance can decrease when visual differences between genuine and counterfeit samples are complex.

The development of CNNs significantly improved image representation learning. Krizhevsky et al. demonstrated the effectiveness of deep CNNs for image classification [5]. He et al. introduced ResNet, which uses residual connections to enable the training of deeper networks [6]. Howard et al. proposed MobileNet as a lightweight CNN architecture suitable for mobile and embedded applications [9].

Siamese Neural Networks were introduced for learning similarity between pairs of inputs. Koch et al. applied Siamese networks to one-shot image recognition [7]. Chopra et al. demonstrated discriminative similarity learning for image verification [8]. These approaches are relevant to counterfeit verification because a system can determine whether two images are sufficiently similar without requiring a separate class for every possible counterfeit type.

4. Methodology

The proposed methodology consists of five main stages: image preparation, handcrafted feature extraction, deep feature extraction, metric learning, and similarity-based decision making.

4.1 Image Acquisition and Preprocessing

Images of genuine and counterfeit items are collected from suitable sources. Before feature extraction, images are resized to a common resolution and may undergo noise reduction, contrast adjustment, region-of-interest selection, and normalization. Controlled transformations such as rotation, scaling, illumination changes, and blur can also be applied to test robustness.

4.2 Handcrafted Feature Extraction

Three local descriptors are considered:

- **SIFT:** Detects scale- and rotation-invariant keypoints and generates descriptors for local image regions.
- **SURF:** Provides efficient local feature extraction using Hessian-based feature detection.
- **ORB:** Combines FAST keypoint detection with rotated BRIEF descriptors and offers relatively low computational cost.

The extracted descriptors can be matched using Brute-Force or FLANN-based matching. Ratio-based filtering can be used to remove unreliable matches.

4.3 Deep Feature Extraction

ResNet and MobileNet are used as deep feature extractors. Instead of relying only on the final classification layer, intermediate or embedding-level representations can be extracted and compared between images. ResNet provides strong feature representation, whereas MobileNet offers a lower computational footprint.

4.4 Siamese Metric Learning

The Siamese model consists of two identical neural network branches with shared weights. Two images are supplied to the branches, and each branch generates an embedding vector. The distance between the two embeddings indicates their similarity.

For two image representations, Euclidean distance is calculated as:

$$
D_W=\sqrt{\sum_i(W_i^{(1)}-W_i^{(2)})^2}
$$

The network can be trained using contrastive loss:

$$
L=(1-Y)\frac{1}{2}D_W^2+Y\frac{1}{2}[\max(0,m-D_W)]^2
$$

where (Y) represents the relationship between the image pair and (m) represents the margin. Similar images are encouraged to have smaller embedding distances, while dissimilar images are pushed farther apart.

4.5 Evaluation

The approaches can be evaluated using accuracy, precision, recall, and F1-score. For verification systems, False Acceptance Rate (FAR) and False Rejection Rate (FRR) are also important. Computational factors such as feature extraction time, inference time, memory usage, and model size should also be considered.

Robustness can be examined under changes in illumination, rotation, scale, blur, and image quality. Performance with limited training samples is particularly important because counterfeit datasets can be difficult to collect.

5. Research Gap

Existing studies generally focus on individual feature extraction or classification techniques. Direct comparison between handcrafted local descriptors, lightweight CNNs, deeper CNN architectures, and Siamese metric learning is less common within a single counterfeit-identification framework.

Another important gap is the limited availability of standardized counterfeit image datasets. This makes consistent comparison across different methods difficult. In addition, many conventional classifiers assume that all counterfeit categories are known during training, which may not represent practical situations where new counterfeit patterns appear.

The proposed framework addresses these issues by considering both classification-oriented and verification-oriented approaches under common evaluation criteria.

6. Objectives of the Study

The main objectives are:

- To study handcrafted image descriptors for counterfeit identification.
- To compare SIFT, SURF, and ORB based on matching capability and computational requirements.
- To examine ResNet and MobileNet as deep image feature extractors.
- To develop a Siamese-based image verification approach.
- To investigate the effect of illumination, scale, rotation, and image quality changes.
- To examine the usefulness of metric learning when training data are limited.
- To identify approaches suitable for practical and resource-constrained deployment.

7. Proposed Counterfeit Identification Framework

The proposed framework consists of the following layers.

7.1 Image Acquisition Layer

This layer collects genuine and counterfeit images from documents, products, labels, packaging, or other target objects. Images may be captured using cameras or obtained from prepared datasets.

7.2 Image Preprocessing Layer

Images are standardized through resizing, normalization, noise removal, contrast enhancement, and region selection. Data augmentation can be used to create realistic variations.

7.3 Handcrafted Feature Layer

SIFT, SURF, and ORB are independently applied to extract local keypoints and descriptors. Matching scores are calculated between reference and test images.

7.4 Deep Feature Extraction Layer

ResNet and MobileNet generate high-level feature representations. These embeddings can be compared using distance-based similarity measures.

7.5 Deep Metric Learning Layer

A Siamese Neural Network learns a similarity function between pairs of images. The model can be trained using genuine-genuine and genuine-counterfeit image pairs.

7.6 Similarity and Decision Layer

The extracted features or learned embeddings are compared. Based on a predefined similarity threshold, the system identifies an input as matching or non-matching with the reference image.

This layered design allows different approaches to be evaluated independently and provides flexibility for selecting a method according to accuracy, speed, hardware availability, and data constraints.

8. Expected Benefits

The proposed framework provides a systematic basis for selecting image-based counterfeit identification techniques. Handcrafted descriptors may be useful for lightweight systems where computational resources are limited. ResNet can provide strong visual representations when sufficient computational resources and training data are available. MobileNet can be considered for mobile and edge-based applications.

Siamese metric learning is particularly useful when new counterfeit patterns appear or when only a limited number of samples are available. The framework can support applications in document verification, product authentication, packaging inspection, financial security, and digital image forensics.

9. Future Research Directions

Future research can focus on creating larger and standardized counterfeit image datasets. More extensive experiments can be conducted under different lighting, rotation, blur, scale, and background conditions. Contrastive loss can also be compared with triplet loss and other metric learning strategies.

Recent architectures such as Vision Transformers and hybrid CNN-transformer models can be investigated. Model compression, quantization, and knowledge distillation may improve deployment on mobile and edge devices. Multispectral or infrared images may provide additional information that is not visible in ordinary RGB images. Explainable AI techniques can also be incorporated to show which regions of an image influenced the counterfeit decision.

CONCLUSION

This paper presents a comparative framework for image-based counterfeit identification using handcrafted local descriptors, deep CNN features, and deep metric learning. SIFT, SURF, and ORB provide traditional feature-based alternatives, while ResNet and MobileNet offer learned visual representations. A Siamese Neural Network with contrastive loss provides a verification-oriented approach that can be useful when counterfeit samples are limited or previously unseen.

Rather than assuming that one algorithm is suitable for every situation, the proposed framework emphasizes the relationship between recognition performance, computational requirements, robustness, available training data, and deployment environment. A systematic comparison of these approaches can help researchers and practitioners select suitable techniques for developing reliable and practical counterfeit identification systems.

REFERENCES

1. D. G. Lowe, “Distinctive image features from scale-invariant keypoints,” International Journal of Computer Vision, vol. 60, no. 2, pp. 91–110, 2004.
2. H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, “Speeded-Up Robust Features (SURF),” Computer Vision and Image Understanding, vol. 110, no. 3, pp. 346–359, 2008.
3. E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, “ORB: An efficient alternative to SIFT or SURF,” in Proceedings of ICCV, 2011, pp. 2564–2571.
4. R. C. Gonzalez and R. E. Woods, Digital Image Processing, 4th ed. Pearson, 2018.
5. A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” Communications of the ACM, vol. 60, no. 6, pp. 84–90, 2017.
6. K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in Proceedings of CVPR, 2016, pp. 770–778.
7. G. Koch, R. Zemel, and R. Salakhutdinov, “Siamese neural networks for one-shot image recognition,” in ICML Deep Learning Workshop, 2015.
8. S. Chopra, R. Hadsell, and Y. LeCun, “Learning a similarity metric discriminatively, with application to face verification,” in Proceedings of CVPR, 2005, pp. 539–546.
9. A. G. Howard et al., “MobileNets: Efficient convolutional neural networks for mobile vision applications,” arXiv preprint arXiv:1704.04861, 2017.
10. F. Schroff, D. Kalenichenko, and J. Philbin, “FaceNet: A unified embedding for face recognition and clustering,” in Proceedings of CVPR, 2015, pp. 815–823.

HOW TO CITE: Harini B. R.*, A Comparative Benchmark Of Handcrafted Local Descriptors And Deep Metric Learning Architectures For Image-Based Counterfeit Identification, Int. J. Sci. R. Tech., 2026, 3 (9), 501-505. https://doi.org/10.5281/zenodo.22939806