Advancing skin cancer detection integrating a novel unsupervised classification and enhanced imaging techniques

Rahman, Md Abdur; Fahad, Nur Mohammad; Raiaan, Mohaimenul Azam Khan; Jonkman, Mirjam; De Boer, Friso; Azam, Sami

Published in:
CAAI Transactions on Intelligence Technology

DOI:
10.1049/cit2.12410

E-pub ahead of print: 01/02/2025

Document Version
E-pub ahead of print

Received: 30 May 2024 | Revised: 16 August 2024 | Accepted: 26 September 2024

CAAI Transactions on Intelligence Technology

DOI: 10.1049/cit2.12410

ORIGINAL RESEARCH

Advancing skin cancer detection integrating a novel unsupervised classification and enhanced imaging techniques

Md. Abdur Rahman[1] | **Nur Mohammad Fahad**[1,2] | **Mohaimenul Azam Khan Raiaan**[1,2] | **Mirjam Jonkman**[2] | **Friso De Boer**[2] | **Sami Azam**[2]

[1]Department of Computer Science and Engineering, United International University, Dhaka, Bangladesh

[2]Faculty of Science and Technology, Charles Darwin University, Casuarina, Northern Territory, Australia

Correspondence

Sami Azam.

Email: Sami.Azam@cdu.edu.au

Abstract

Skin cancer, a severe health threat, can spread rapidly if undetected. Therefore, early detection can lead to an advanced and efficient diagnosis, thus reducing mortality. Unsupervised classification techniques analyse extensive skin image datasets, identifying patterns and anomalies without prior labelling, facilitating early detection and effective diagnosis and potentially saving lives. In this study, the authors aim to explore the potential of unsupervised learning methods in classifying different types of skin lesions in dermatoscopic images. The authors aim to bridge the gap in dermatological research by introducing innovative techniques that enhance image quality and improve feature extraction. To achieve this, enhanced super-resolution generative adversarial networks (ESRGAN) was fine-tuned to strengthen the resolution of skin lesion images, making critical features more visible. The authors extracted histogram features to capture essential colour characteristics and used the Davies–Bouldin index and silhouette score to determine optimal clusters. Fine-tuned k-means clustering with Euclidean distance in the histogram feature space achieved 87.77% and 90.5% test accuracies on the ISIC2019 and HAM10000 datasets, respectively. The unsupervised approach effectively categorises skin lesions, indicating that unsupervised learning can significantly advance dermatology by enabling early detection and classification without extensive manual annotation.

KEYWORDS

computer vision, histograms, unsupervised learning

1 | INTRODUCTION

Skin cancer presents major health risks because of its high prevalence and aggressive nature, particularly in the case of melanoma, the most deadly form of skin cancer [1]. One concerning aspect is its ability to metastasise and spread to vital organs such as the lungs, liver, and brain [2, 3]. Skin cancer is one of the leading causes of brain metastases which are ultimately lethal [4–6]. Globally, millions are diagnosed with skin cancer each year [7].

Early detection is critical in improving prognosis [8], with early-stage melanoma boasting a 95% cure rate [9]. However, reliable identification remains challenging due to the subjective nature of skin lesion evaluations and variations in lesion appearance [1]. Traditional methods such as the ABCD (asymmetry, border, colour, and diameter) rule and seven-point checklist provide guidelines, although they may not consistently lead to the detection of malignancies at an early stage. Technologies that can analyse dermatoscopic images with high precision could potentially surpass human diagnostic capabilities. They may play a crucial role in the early detection and treatment of skin cancer in the future, ultimately saving lives and reducing the burden of this disease [1, 10].

Unsupervised classification, in comparison to supervised classification, offers advantages in real-world scenarios [11, 12]. Algorithms can autonomously identify patterns and groupings

within datasets without labelled examples [12]. This is beneficial in skin lesion classification where nuances are intricate. By clustering similar lesions based on intrinsic characteristics such as colour and texture, unsupervised algorithms can effectively categorise lesions. This approach can be valuable in distinguishing between various skin conditions, including melanoma, nevus, and basal cell carcinoma, and aid clinicians in accurate diagnosis and treatment planning.

Previous studies [13–21] have explored deep learning and machine learning models for skin cancer classification but there is a notable absence of unsupervised methods for classifying different types of skin lesions. One of the primary reasons is the difficulty in extracting discriminative features from raw image data without labelled information [13]. Skin lesion images exhibit significant variability in colour, texture, and contextual details, which complicates the extraction of meaningful features. Unsupervised classifiers often struggle with distinguishing subtle differences in lesions, leading to the inclusion of unnecessary features that do not contribute to accurate classification [14]. Furthermore, the complexity and variability of skin lesion images make it challenging to define clear boundaries between different lesion types, further complicating unsupervised classification tasks.

These technical challenges emphasise the necessity for innovative methodologies that can effectively handle the inherent complexity and variability in skin lesion images in order to advance learning approaches in this domain without supervision. To address this gap, our research presents a novel approach to skin lesion classification through unsupervised learning, uniquely combining image enhancement to improve feature visibility with an innovative feature extraction module. By optimising image fidelity and developing a distinctive feature extraction technique, we streamline the clustering process to effectively enable the classification of unlabelled skin lesion images.

The following are the primary contributions of our study:

- Firstly, to address the challenge of extracting significant features from skin lesion images that display a wide range of colours, textures, and contextual details, we employ vast ablation experiments to fine-tune and optimise the architecture of the enhanced super-resolution generative adversarial network (ESRGAN).
- Secondly, a novel histogram-based feature extraction approach is proposed to address the complexity of distinguishing differences and focus on discriminative features in skin lesion images. This method captures distinctive patterns and essential information about colour, texture, and pixel-level characteristics through various ablation experiments.
- Additionally, we address the issue of high-dimensional feature spaces complicating clustering tasks. To mitigate this challenge, we apply dimensionality reduction techniques such as the principal component analysis (PCA) and t-distributed stochastic neighbour embedding (t-SNE) to transform the feature space into a lower-dimensional representation, facilitating more effective clustering.
- Moreover, determining the optimal number of clusters amidst the variability of skin lesions poses another significant challenge. The study introduces a systematic approach utilising metrics such as the Davies–Bouldin index and silhouette score to determine the appropriate number of clusters.
- Finally, we implement the k-Means clustering algorithm for unsupervised clustering using Euclidean distance in the feature space created by histogram feature maps.

The rest of the sections of this study are organised as follows: Section 2 critically examines the existing literature, analysing seminal works and pertinent research in order to establish a contextual framework. Section 3 elucidates the methodology employed in this study, providing the detailed research design and utilised approach. Section 4 provides the results and outcomes of the research. Section 5 explores the implications and limitations of the study. Section 6 provides the conclusions of the research.

2 | RELATED WORK

This section reviews related work, which is categorised into two themes. The first covers recent advancements and diverse algorithms in unsupervised learning for clustering tasks. The second explores innovative models and their contributions to unsupervised image classification.

2.1 | Deep clustering

Various novel deep clustering algorithms have appeared in recent literature, demonstrating notable advancements in unsupervised learning for clustering tasks. For instance, Sinaga et al. [15] introduced an unsupervised k-means approach (U-k-Means), eliminating the need for initialisation and parameter selection by incorporating an entropy penalty term. Similarly, the fully unsupervised probabilistic c-means (FU-PCM) algorithm proposed by Yang et al. [16] avoids the need for parameter selection and initialisations. Murphy et al. [17] addressed unsupervised clustering challenges with the spectral–spatial diffusion learning (DLSS) technique, combining geometric estimation of class modes with a diffusion-inspired labelling method. Mrukwa et al. [18] introduced DiviK, a scalable algorithm for automated unsupervised learning in high-dimensional datasets. Zhang et al. [19] proposed the RFA-LCF algorithm for high-dimensional data clustering, addressing noise sensitivity and optimising locality. Polk et al. [20] introduced the D-VIC algorithm for unsupervised material clustering in hyperspectral images. Han et al. [21] presented a bias-avoidant clustering approach by training image representations from scratch using self-

supervised learning. Yan et al. [22] proposed the DERC framework, which combines image embedding, dimensionality reduction, and clustering to address challenges in unsupervised image clustering tasks. On the other hand, Zhang et al. [23] discussed the challenge of unsupervised image clustering by investigating methods such as the bag-of-visual-words and self-supervised and transfer learning.

2.2 | Unsupervised models

In unsupervised image classification, various innovative models have been proposed, each addressing distinct challenges. For instance, Wang et al. [24] proposed the self-supervised topology clustering network (STCN) employing a transformation-invariant network and a self-supervised maximum modularity clustering algorithm for the autonomous classification of unlabelled medical images without predefined class numbers. They achieved an accuracy of 80.6%. Khan et al. [25] explored the application of optimal transport-driven CycleGAN (OT-CycleGAN) for artefact removal, addressing the challenge of the limited number of paired high-quality images for supervised training. They investigated two variations of OT-CycleGAN—one with partial knowledge of image degradation and one without. Sivanesan et al. [26] introduced a method that utilises generative adversarial networks and edge detection techniques, an unsupervised semantic medical image segmentation approach. Yu et al. [27] proposed an unsupervised deep region competition (DRC) algorithm that uses a mixture of experts framework, learnt pixel re-assignment, and energy-based priors to accurately extract foreground objects from images through expectation maximisation. Cohn et al. [28] presented an unsupervised machine learning system for image classification utilising the pre-trained VGG16 network and principal component analysis. The approach leveraged k-means clustering to classify images without labelled training data. Furthermore, Ge et al. [29] introduced a dual adversarial autoencoder (Dual-AAE) designed for unsupervised clustering in complex data structures. This approach combines adversarial autoencoders and variational inference, maximising the likelihood function and mutual information. Zhou et al. [30] proposed saliency-CCE, a dual-task framework for skin lesion saliency detection and segmentation in biomedical images. It incorporates a colour contextual extractor module for salient object detection and an improved adaptive threshold for automated image segmentation. Innani et al. [31] presented efficient-GAN (EGAN), an adversarial learning framework focused on skin lesion segmentation using an unsupervised generative network with unique architectural elements. Li et al. [14] proposed USL-Net, an innovative uncertainty self-learning network tackling unsupervised skin lesion segmentation challenges through contrastive learning for feature extraction, generation of class activation maps (CAMs) as pseudo-labels for lesion regions, and identification of uncertainty regions to mitigate potential labelling errors. A CNN optimising label assignment based on similarity, spatial continuity, and a large number of unique labels was developed by Kim et al. [32]. It featured a novel end-to-end network, a spatial continuity loss function, and extensions for user-input scribbles and unseen image segmentation. Unsupervised models are also utilised in the agricultural field. For instance, Nazki et al. [33] extended the concept of unsupervised image translation with AR-GAN, a GAN-based pipeline addressing challenges such as class imbalance and data acquisition, incorporating an activation reconstruction loss function to improve the quality and semantics of generated images. They achieved an accuracy of 68% per pixel. Finally, Zhang et al. [34] introduced a collaborative unsupervised domain adaptation algorithm (CoUDA), leveraging rich labelled data from related domains to enhance learning in the target task with a specific focus on transferability-aware adaptation and collaborative mitigation of label noise, resulting in an accuracy of 87%.

In reviewing the existing literature, we have identified notable gaps that our research aims to fill. Many studies overlook the importance of implementing image enhancement strategies, limiting the clarity and quality of visual data analysis. Additionally, the inconsistent use of unsupervised learning techniques across studies hinders the extraction of discriminative features. This is especially challenging given the variations in the image colour, texture, and contextual details. Consequently, unsupervised classifiers struggle with this variability, resulting in the inclusion of unnecessary features and a decrease in classification accuracy.

Our research introduces a novel classification approach to address these gaps. We utilise a modified enhanced super-resolution generative adversarial network architecture, vastly fine-tuning it to optimise performance specifically for skin lesion classification. By comparing our approach with other GAN-based enhancers, we demonstrate the superior efficiency of our method. Because of the extensive fine-tuning applied to find the proper configuration, our study contributes vastly to image enhancement techniques. Furthermore, we explore various configurations of histogram feature extraction modules to capture subtle contextual features. This level of detail has been overlooked in previous studies. We implement standardisation and dimensionality reduction techniques to handle the high-dimensional histogram features. Combining these approaches missing in state-of-the-art studies, we address these gaps and provide a novel system. This effectively projects these features onto a more manageable subspace, enhancing the performance of k-means clustering and surpassing other tested algorithms. Through these innovations, our research addresses the existing gaps and establishes a new benchmark for the effectiveness of image classification methods and unsupervised classification in skin lesion classification. Table 1 provides an overview of the current studies.

TABLE 1 Comparison of state-of-the-art studies with the proposed model. The findings of our study are highlighted in bold.

References	Dataset	Image enhancement	Unsupervised learning	Colour histogram	Classification	Performance
[24]	ISIC 2018	X	X	X	✓	80.6% (acc.)
[26]	1. ISIC 2018	X	✓	X	X	90.40%
	2. Renal ultrasound images					90% (acc.)
[31]	ISIC 2018	X	✓	X	X	94.5% (acc.)
[14]	1. ISIC18 (&17)	X	✓	X	X	88.5% (acc.)
	2. PH2					92.4% (acc.)
[32]	1. PASCAL VOC 2012	X	✓	X	X	29.51%
	2. BSD500					36.47% (mIOU)
[34]	Colon dataset	X	✓	X	✓	87.75% (acc)
[21]	1. CIFAR10	X	X	X	✓	89.70%
	2. CIFAR100					68.30%
	3. OmniGlot					89.10%
	4. ImageNet					82.50%
[33]	Cityscapes	X	✓	X	✓	86.1% (acc)
[20]	1. Salinas A	X	✓	X	X	0.089
	2. Jasper ridge					0.865
	3. Indian pines					0.445 (acc)
[23]	ImageNet	X	X	X	✓	92% (within cluster consistency)
[17]	1. Indian pines	X	✓	X	✓	82%
	2. Pavia					83%
	3. Salinas A					90%
	4. Kennedy space centre					73%
[19]	1. UMIST, CMU PIE	X	✓	X	✓	41.58%–48.83%
	2. COIL100, ETH80					49.05%–67.59%
	3. HWDB1.1-D & L					38.71%–39.85%
[25]	sUS image	X	✓	X	X	CR = 17.21, CNR = 2.90, GCNR = 0.9620 (SLAmatrix)
[30]	1. ISIC-2016	X	✓	X	X	85.91% (acc.)
	2. ISIC- 2017					83.87% (acc.)
	3. SCISC datasets					67.80% (mIoU)
[27]	1. Caltech-UCSD	X	✓	X	✓	56.4% IoU
	2. Stanford dogs					71.7% IoU
	3. Stanford cars					72.4% IoU
	4. CLEVR					84.7% IoU
	5. TM-dSprits					78.8% IoU
[29]	1. MNIST	X	X	X	✓	97.86% acc
	2. HHAR					86.79% acc
	3. STL 10					89.15% acc
	3. REUTERS					81.45% acc

TABLE 1 (Continued)

References	Dataset	Image enhancement	Unsupervised learning	Colour histogram	Classification	Performance
[15]	1. UCI machine learning repository	X	✓	X	X	90.48% acc
	2. Yale face					34% acc
	3. CIFAR-10					33.8% acc
	4. Multi medical dataset					92% acc
[18]	1. OSCC dataset	X	✓	X	X	0.7372 dice
	2. 3D mouse kidney					0.750 dice
[22]	1. MNIST-test	X	✓	X	X	97.50%
	2. USPS					97.70%
	3. FRGC					51.30%
	4. YTF					65.80%
	5. CMU-PIE					97.90%
Ours	**1. ISIC2019**	✓	✓	✓	✓	**87.77%**
	2. HAM10000					**90.50%**

3 | METHODOLOGY

This section presents the methodology employed in our unsupervised image clustering approach. We utilise a combination of enhanced super-resolution generative adversarial networks (ESRGAN) [35] for enhanced clarity and critical feature visibility in skin lesion images, a histogram feature map for extracting distinctive patterns and shapes of skin lesions, determination of the optimal number of clusters, and the application of unsupervised clustering using the k-means clustering algorithm. Figure 1 depicts the architecture of our system. The system consists of six steps, and all of them are sequentially displayed to improve the compression.

3.1 | ESR-generative adversarial networks

ESR-generative adversarial networks (ESRGAN) are employed to enhance the image quality within our dataset. ESRGANs are built upon the traditional generative adversarial network (GAN) architecture. The generator, denoted as G, aims to generate high-resolution images from low-resolution inputs, whereas the discriminator, denoted as D, distinguishes between real and generated images. The training of ESRGANs involves minimising the following adversarial loss function:

$$\mathcal{L}_{\mathrm{GAN}}(G,D) = E_{\mathrm{real}}[\log D(I)] + E_{\mathrm{fake}}[\log(1 - D(G(I)))] \tag{1}$$

where E_{real} and E_{fake} represent the expectations over real and generated images, and I represents the input low-resolution image. Additionally, to ensure perceptual similarity between the generated high-resolution image and the ground truth, a content loss is employed:

$$\mathcal{L}_{\mathrm{content}}(\mathrm{G}) = \frac{1}{\mathrm{N}} \sum_{\mathrm{i}=1}^{\mathrm{N}} | \Phi(\mathrm{I_i}) - \Phi(\mathrm{G}(\mathrm{I_i})) |_1 \tag{2}$$

where $\boldsymbol{\Phi}$ is a pre-trained feature extractor, and N is the number of images. By minimising both the adversarial loss and the content loss during training, the clarity of the critical features and patterns of lesion images is enhanced. Figure 2 shows the effectiveness of ESRGAN in preserving essential features, resulting in a clearer representation of the shapes, patterns, and structures of skin lesions. This enhancement is crucial for our approach as it enables better subsequent feature extraction and clustering, leading to a better analysis of skin lesion images.

3.2 | Histogram feature map

Following the enhancement of image quality, our methodology incorporates a histogram feature map technique to extract distinctive features from the skin lesion images. This involves quantifying the distribution of pixel intensities in each image through a colour histogram, resulting in histogram feature maps denoted as 'H'. These maps provide a foundational representation capturing essential information about the colour characteristics of skin lesion images. The colour histogram for a given image 'I' is computed as follows [36]:

$$\mathrm{H}(\mathrm{I}) = \sum_{\mathrm{i}=1}^{\mathrm{N}} \left| \frac{1}{Num\_Pixels} \sum_{\mathrm{Pixels}} \mathrm{B_i} \right| \tag{3}$$

here, N represents the number of bins in the histogram, *Num_Pixels* is the total number of pixels in the image, and B_i is the colour value in bin i. These histogram feature maps, as shown in Figure 3, provide a concise and informative representation of the colour characteristics of each class for both ISIC 2019 and HAM10000 datasets.

To ensure the normalisation and consistency of the resulting histograms, we follow a standardisation process. The

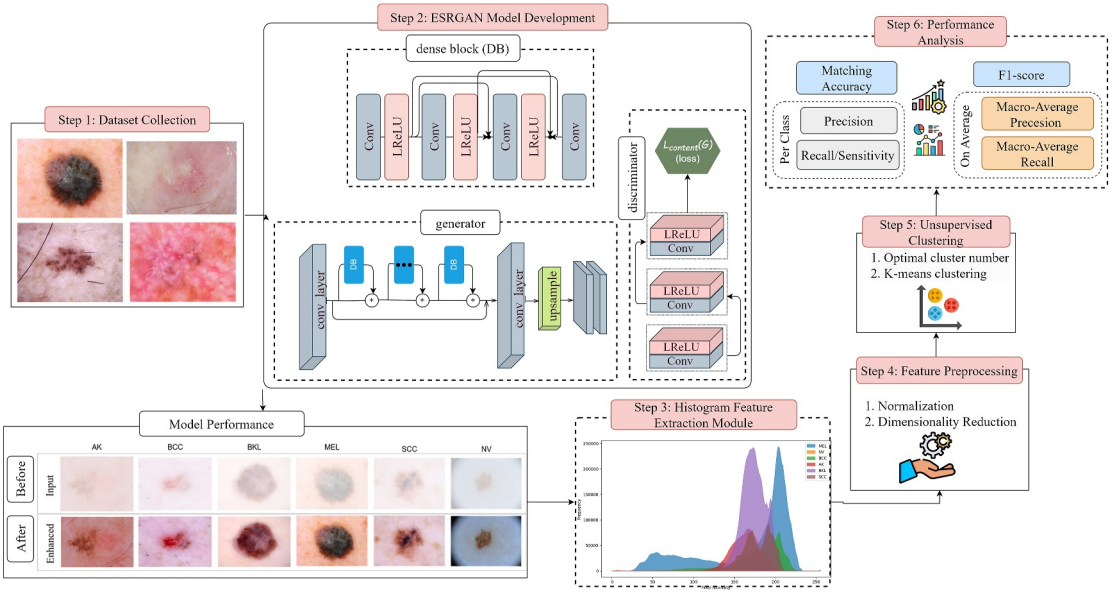


FIGURE 1 Flow of the proposed work.

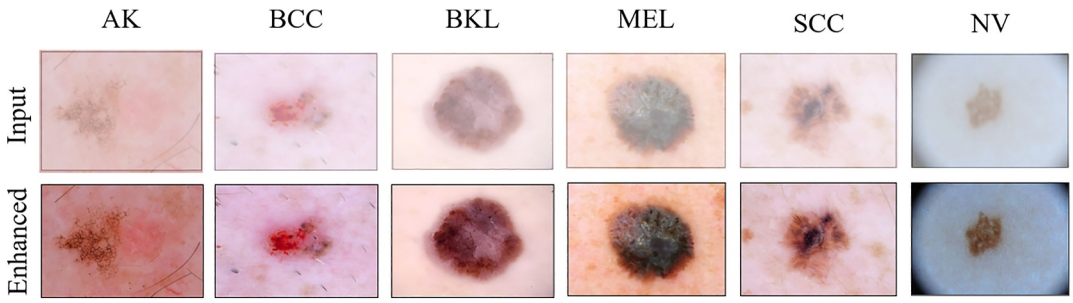


FIGURE 2 ESRGAN output. ESRGAN, enhanced super-resolution generative adversarial network.

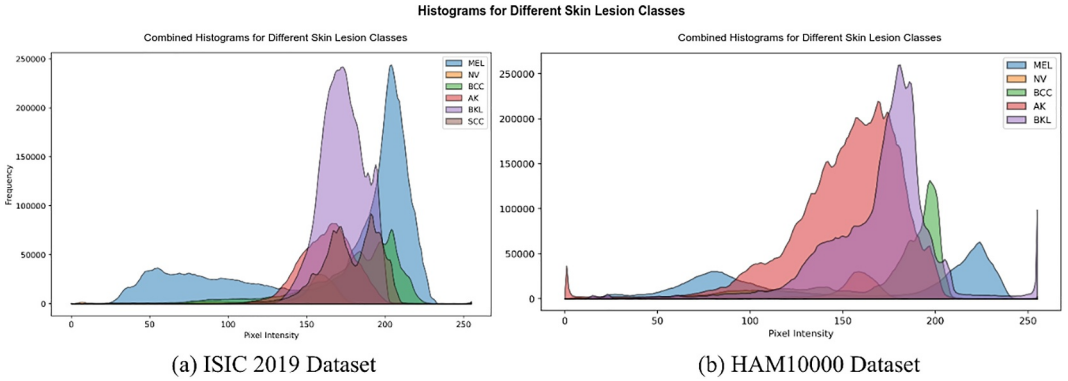


FIGURE 3 Compacted histogram representation.

standardised histograms H_{std} are obtained by subtracting the mean μ_H and dividing by the standard deviation σ_H of the original histograms H:

$$H_{\text{std}} = \frac{H - \mu_H}{\sigma_H} \tag{4}$$

We employed this standardisation procedure to maintain consistent scales across diverse images, thereby preventing biases introduced by varying scales. This procedure ensures that each feature contributes equally to the subsequent analysis, even if these features have different scales or units. The transformation centres the mean of each feature at 0 and scales the standard deviation to 1. Recognising the high-dimensional nature of the feature space, we employed the principal component analysis (PCA) to reduce dimensionality but retain essential characteristics of the data. The transformation of the set of standardised histograms H using PCA can be expressed as follows:

$$\text{PCA(H)} = \text{H} \cdot \text{W} \tag{5}$$

where W represents the matrix of principal components. Herein, W is comprised of eigenvectors corresponding to the eigenvalues of the covariance matrix of the standardised histograms. Each column of W signifies a principal component, capturing the directions in the original feature space where the data exhibits the most significant variations. This transformation allows for projecting the data into a lower-dimensional space while retaining essential information about the colour characteristics of skin lesion images. By reducing the number of dimensions in our feature space, PCA handles the challenges associated with high dimensionality, enhancing the computational efficiency of subsequent analyses. Simultaneously, the process ensures that the essential information about the colour characteristics of skin lesion images is retained by selecting principal components that encapsulate the maximum variance.

Furthermore, we incorporate t-distributed stochastic neighbour embedding (t-SNE) as a complementary technique for additional dimensionality reduction, which is particularly adept at preserving local structures within the data, making it a valuable tool for revealing potential clusters within the feature space. The dimensionality reduction process for t-SNE is expressed as follows:

$$\text{t-SNE(H)} = \text{t} - \text{SNE(PCA(H))} \tag{6}$$

PCA initially reduces the dimensionality while retaining essential characteristics, and t-SNE further refines this representation by focusing on preserving local structures.

3.3 | Optimal cluster number

Determining the optimal number of clusters is a critical aspect of our unsupervised clustering method. Our approach involves an analysis of multiple metrics, that is, the Davies–Bouldin index (DBI), and the silhouette score. The (DBI), calculated with Equation (7), is used to assess the compactness and separation of clusters.

$$\text{DBI} = \frac{1}{\text{K}} \sum_{\text{i}=1}^{\text{K}} \max_{j \neq i} \left(\frac{\text{avg}(R_i) + \text{avg}(R_j)}{\text{dist}\left(\mu_i, \mu_j\right)} \right) \tag{7}$$

here K represents the number of clusters, R_i is the set of distances from the i^{th} cluster centre to all points in the cluster, μ_i is the centroid of the i^{th} cluster, and dist (μ_i, μ_j) is the distance between centroids. In addition to the Davies–Bouldin index, the silhouette score was considered as another crucial metric for evaluating cluster quality. The silhouette score is computed for each data point and provides a measure of how well-separated the clusters are. The formula for the silhouette score is given as follows:

$$\text{Silhouette Score} = \frac{1}{\text{N}} \sum_{\text{i}=1}^{\text{N}} \frac{\text{b}_\text{i} - \text{a}_\text{i}}{\max(\text{b}_\text{i}, \text{a}_\text{i})} \tag{8}$$

here N is the total number of data points, ai is the average distance from the i^{th} point to other points in the same cluster, and bi is the average distance from the i^{th} point to points in the nearest cluster (different from the one to which i belongs).

The silhouette score ranges from −1 to +1. A score close to one indicates highly similar and well-matched instances within each cluster. A low DBI score indicates well-separated clusters. Our silhouette and DBI scores satisfy both criteria which means that there is evidence of a strong similarity within our clusters. Figure 4 illustrates that our process leads to the identification of an optimal value for k in our subsequent unsupervised clustering for both datasets. The optimal number

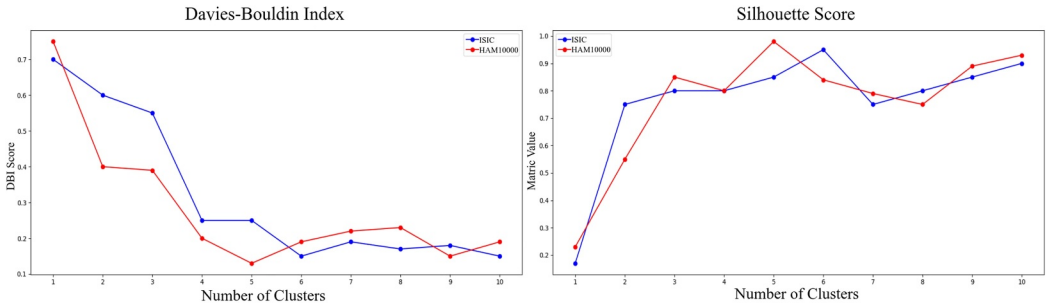


FIGURE 4 Optimal cluster.

of clusters was shown to be six for the ISIC2019 dataset and five for the HAM10000 dataset.

3.4 | Unsupervised clustering

The unsupervised clustering process is executed using the k-means clustering algorithm. This was employed to cluster images based on their similarity in the feature space created by the histogram feature maps. The assignment of images to clusters relies on the Euclidean distance metric:

$$\text{Cluster}(x_i) = \text{argmin}\left\|h_{x_i} - h_{c_j}\right\| \quad (9)$$

here, $\|.\|$ represents the Euclidean distance between the feature vector of image x_i (h_{x_i}) and the feature vector of a centroid in cluster c_j (*hcj*). Utilising the Euclidean distance metric in the algorithm facilitates the measurement of similarity or dissimilarity between feature vectors. The proposed algorithm is as follows:

Algorithm 1 Skin Lesion Image Analysis:

```
Input: Skin Lesion Images X = {x1, . . ., xi, . . ., xn}
Key Parameters: X, ESRGAN, X′, H, P CA, t − SNE, k
for all images x in X do
  G ← img_generation (x)
  Upsampled ← enhance_image (G[x])
  D ← img_discrimination (Upsampled[x]) (1)
  L ← calc_loss (Upsampled[x], D[x]) (2)
  X′ ← ESRGAN (G[x], D[x], L[x])
end for
H ← [ ]
for all img in X′ do
  Hcol ← calc_colour_histogram (img) (3)
  Hnorm ← normalise_histogram (Hcol)
  H [ ] ← Hnorm
end for
Hflat ← standardise_histograms(H [ ]) (4)
PCAmap ← PCA(Hflat) (5)
tSNEmap ← tSNE(PCAmap) (6)
k ← DBIandSilhouette(tSNEmap) (7) and (8)
cluster ← kMeans(tSNEmap, k) (9)
Output: Optimised Skin Lesion Image Clusters
```

Table 2 depicts the abbreviation for the Algorithm 1.

4 | EXPERIMENTAL RESULTS

4.1 | Dataset

The dataset utilised in this study comprises two primary sources: the ISIC 2019 challenge dataset [37] and the HAM10000 dataset [38]. The ISIC 2019 dataset includes data from previous years (2018 and 2017) and initially consisted of eight classes: melanoma (MEL), melanocytic nevus (NV), basal cell carcinoma (BCC), actinic keratosis (AK), benign keratosis lesion (BKL), cermatofibroma (DF), vascular lesion (VASC), and squamous cell carcinoma (SCC). The HAM10000 dataset initially had seven classes, the same classes as ISIC2019, except squamous cell carcinoma (SCC). Upon initial examination, both datasets exhibited imbalanced distributions across classes. Table 3 shows the detailed breakdown of the class distribution for both the ISIC 2019 and HAM10000 datasets.

Because of the limited number of samples in the dermatofibroma (DF) and vascular lesion (VASC) classes in both datasets, less than 1% in the ISIC2019 dataset and slightly over 1% in the HAM10000 dataset, we opted to exclude these classes from our analysis.

The final dataset comprises six classes from the ISIC2019 dataset with 23,400 images and five classes from the HAM10000 dataset with 5000 images. To address the imbalance and enhance the dataset's quality, an augmentation strategy was employed by integrating ESRGAN with traditional operations such as rotation, flipping, cropping, and zooming. These measures were implemented to distribute the classes more uniformly and ensure a more balanced representation within the respective datasets.

Note that our final ISIC2019 dataset is larger than the HAM10000 dataset. This choice was made to enable the

TABLE 2 Algorithm abbreviation.

Variables	Meanings
G	Generator
D	Discriminator
L	Content loss
X'	Enhanced images
H_{col}	Extracted colour histogram
H_{norm}	Normalised histogram
H_{flat}	Flattened histogram

TABLE 3 Distribution of classes.

Dataset	ISIC2019				HAM10000			
Type	Dermatosopic				Dermatosopic			
Availability	Public				Public			
	Original		Final		Original		Final	
Class	Image	Ratio %	Image	Ratio %	Image	Ratio %	Image	Ratio %
MEL	4522	17.85%	5000	21.37%	1113	11.11%	1000	20%
NV	12,875	50.83%	6000	25.64%	6705	66.95%	1200	24%
BCC	3323	13.12%	4000	17.10%	514	5.13%	900	18%
AK	867	3.42%	2500	10.68%	327	3.27%	800	16%
BKL	2624	10.36%	3400	14.53%	1099	10.97%	1100	22%
DF	239	0.94%	Discarded		115	1.15%	Discarded	
VASC	253	1%	Discarded		142	1.42%	Discarded	
SCC	628	2.48%	2500	10.68%	-	-	-	-
Total images	25,331	100%	23,400	100%	10,015	100%	5000	100%

evaluation of our framework across datasets of varying sizes and distributions.

4.2 | Implementation details

To implement the proposed skin lesion image analysis methodology, we combine image enhancement techniques, feature extraction, dimensionality reduction, and unsupervised clustering. The initial step involved addressing challenges associated with distortion and blurriness in skin lesion images. ESRGAN was employed for general image quality improvement, specifically to correct distortions and blurriness. Utilising a GAN architecture, ESRGAN effectively enhanced previously unclear features, contributing to an overall improvement in image quality. We then resized the images to 224 × 224 pixels for consistency and extracted histogram features to capture the colour characteristics of the skin lesion images. The histograms were normalised to ensure consistency across diverse images. To manage the high-dimensional nature of the feature space, we applied the principal component analysis (PCA) and t-distributed stochastic neighbour embedding (t-SNE) for dimensionality reduction. For unsupervised clustering, the k-means clustering algorithm was employed, using the Euclidean distance as a metric to measure proximity in the reduced feature space. The Davies–Bouldin index (DBI) and silhouette score were utilised to determine the optimal number of clusters. The DBI provides a measure of cluster compactness and separation, whereas the silhouette score assesses the cohesion and separation of clusters. After evaluating these metrics, the optimal number of clusters was determined to be six and five for the ISIC2019 and the HAM10000 datasets, respectively.

4.3 | Evaluation criteria

To evaluate the effectiveness of our skin lesion image analysis methodology for a multi-class task, we utilise a set of evaluation criteria. These criteria include accuracy, macro-average precision, macro-average recall, and F1-Score, which collectively assess the model's capabilities.

The accuracy of our framework is determined by calculating accuracy (Equation 10), the ratio of correctly predicted samples to the total number of samples across all classes. This measure assesses the system's ability to cluster samples from all classes correctly.

$$\text{Accuracy} = \frac{\sum_{i=1}^{n}(TP_i + TN_i)}{\sum_{i=1}^{n}(TP_i + TN_i + FP_i + FN_i)} \tag{10}$$

Next, we evaluate macro-average precision (M.A.P.), which focuses on the accuracy of positive predictions across all classes. Equation (11) quantifies the proportion of correctly identified positive instances out of the predicted positives, averaged across all classes.

$$\text{M.A.P.} = \frac{1}{n}\sum_{i=1}^{n}\frac{TP_i}{TP_i + FP_i} \tag{11}$$

Similarly, macro-average recall (M.A.R.), also known as sensitivity or the true positive rate, is calculated (Equation 12) by finding the ratio of true positives to the sum of true positives and false negatives, averaged across all classes.

$$\text{M.A.R.} = \frac{1}{n}\sum_{i=1}^{n}\frac{TP_i}{TP_i + FN_i} \tag{12}$$

To ascertain the impact of both precision and recall, the F1-Score is computed, representing the harmonic mean of these two metrics.

$$\text{F1} - \text{Score} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \tag{13}$$

The F1-Score always lies between precision and recall, that is, $\min(P,R) \le \text{F1} - \text{Score} \le \max(P,R)$. In the Equations (10), (11), (12), and (13) TP_i, TF_i, FP_i, FT_i, represent the corresponding values for each class i. These values are calculated based on the model's performance in classifying instances from the datasets, where n is the total number of classes in our data. To comprehensively evaluate the performance of our image enhancement techniques, we utilised a range of metrics that collectively address different aspects of image quality:

First, we assess the peak signal-to-noise ratio (PSNR) to measure the overall fidelity of the enhanced images compared to the originals. PSNR quantifies the amount of distortion introduced by the enhancement process. It is computed as follows:

$$\text{PSNR} = 10\log_{10}\left(\frac{\text{R}^2}{\text{MSE}}\right) \tag{14}$$

where R is the maximum pixel value and MSE is the mean squared error between the original and enhanced images. Next, we evaluate the structural similarity index (SSIM) to understand how well the structural content of the enhanced image matches that of the original. SSIM considers luminance, contrast, and texture which provides insight into how well these features are preserved. The SSIM is calculated as follows:

$$\text{SSIM}(x,y) = \frac{\left(2\mu_x\mu_y + C_1\right)\left(2\sigma_{xy} + C_2\right)}{\left(\mu_x{}^2 + \mu_y{}^2 + C_1\right)\left(\sigma_x{}^2 + \sigma_y{}^2 + C_2\right)} \tag{15}$$

where μ_x and μ_y are the means, $\sigma_x{}^2$ and $\sigma_y{}^2$ are the variances, and σ_{xy} is the covariance of the original and enhanced images. Following this, we compute the Perceptual Index (PI) to evaluate how visually different the enhanced image is from the original, capturing perceptual distortions. For original images x, and enhanced images y, the PI is defined using the following equation:

$$\text{PI} = \exp(-\text{mean}(\text{abs}(\text{x} - \text{y}))) \tag{16}$$

Subsequently, to understand how well spatial structures are preserved after enhancement, we determine the spatial

correlation coefficient (SCC) to measure the correlation between the spatial features of the original and enhanced images. For x and y to be the flattened arrays of the original and enhanced images, respectively, SCC is calculated as follows:

$$\mathrm{SCC} = \mathrm{pearsonCC}(x[\,], y[\,]) \tag{17}$$

Finally, to assess the naturalness of the enhanced images, we calculate the naturalness image quality evaluator (NIQE). NIQE evaluates how natural the enhanced image appears by analysing statistical properties. It is given using the following equation:

$$\mathrm{NIQE} = 0.5(|\mathrm{kurtosis} - 3| + |\mathrm{skewness}|) \tag{18}$$

where kurtosis and skewness measure the distribution of image intensities. Lower NIQE values indicate that the enhanced image is perceived as more natural.

To further evaluate the effectiveness of our representations, we used normalised mutual information (NMI) and adjusted Rand index (ARI). NMI measures the amount of information shared between the true data labels and the clustering results. It is calculated as follows:

$$\mathrm{NMI} = \frac{I(C;L)}{\sqrt{H(C)\,.\,H(L)}} \tag{19}$$

where $I\,(C;\,L)$ is the mutual information between clusters C and labels L, and $H\,(C)$. $H\,(L)$ are the entropies of the clusters and labels, respectively.

Finally, correcting for chance, to assess the agreement between the clustering result and the true labels, for $\mathrm{R_{index}}$ to be the Rand index that quantifies the similarity between two data clusters, we incorporate the adjusted Rand index (ARI) matric which is computed as follows:

$$\mathrm{ARI} = \frac{\mathrm{R_{index}} - \mathrm{Expected}(\mathrm{R_{index}})}{\mathrm{Max}(\mathrm{R_{index}}) - \mathrm{Expected}(\mathrm{R_{index}})} \tag{20}$$

4.4 | Performance analysis

4.4.1 | ESRGAN ablation study

To improve the quality of the skin lesion images, we conducted an extensive ablation study on the ESRGAN model. The ESRGAN architecture consists of a generator network responsible for upscaling the input images and a discriminator network that aims to distinguish between the super-resolved images and ground-truth images. We systematically explored various configurations of the generator and discriminator networks, as well as the impact of different hyperparameters, to identify the optimal setup for enhancing the skin lesion image quality. The generator network was based on the residual-in-residual dense block (RRDB) architecture, and we varied the number of RRDB blocks from 8 to 32 to analyse the effect of the model depth on the reconstruction performance. Additionally, we experimented with different residual scaling factors, ranging from 0.1 to 0.3, to control the magnitude of the residual connections within the generator. The initialisation variance of the generator network was also adjusted, with values from 0.01 to 0.05, to investigate the impact of the weight initialisation on the convergence and final results. For the discriminator network, we explored different configurations by varying the number of convolutional layers from 5 to 32 to identify the most effective discriminator architecture to guide the generator towards generating more super-resolved images. We also evaluated the influence of different optimisers, including Adam, RMSProp, and stochastic gradient descent (SGD), along with their corresponding learning rates (ranging from 1.00E-04 to 2.00E-04). To quantify the performance of the various ESRGAN configurations, we employed several evaluation metrics, including peak signal-to-noise ratio (PSNR), structural similarity index (SSIM), perceptual index (PI), Spearman correlation coefficient (SCC), and natural image quality evaluator (NIQE).

The results of the ESRGAN ablation study, as presented in Table 4, reveal that the best overall performance was achieved in Experiment 7, which utilised a generator network with 28 RRDB blocks, a residual scaling factor of 0.15, and an initialisation variance of 0.03. The discriminator network in this configuration consisted of 14 convolutional layers. This setup was trained using the Adam optimiser with a learning rate of 1.00E-04 and employed an adversarial loss weight of 0.01 and a content loss weight of 0.03. Figure 5 illustrates the visual differences in the reconstructed skin lesion images across the various experimental configurations.

4.4.2 | Fine-tuned ESRGAN comparison with state-of-the-art methods

After conducting a vast ablation study, we proposed a fine-tuned version of ESRGAN that can effectively classify skin lesions. A comparative analysis to evaluate the performance of our optimised ESRGAN model against several pre-trained GAN-based image enhancement models, including super-resolution generative adversarial networks (SRGAN), BigGAN, high dynamic range GAN (HDR-GAN), StyleGAN, and DeblurGAN. The results, as shown in Table 5, demonstrate that our ESRGAN configuration obtained the highest PSNR, SSIM, and SCC values, along with the lowest PI and NIQE scores, outperforming all the pre-trained models.

The consistent and substantial improvements observed across multiple evaluation metrics suggest that the fine-tuned ESRGAN model can effectively enhance the quality of skin lesion images and outperform the pre-trained GAN-based image enhancement models.

4.4.3 | Histogram feature extraction ablation study

To validate the effectiveness of our histogram feature extraction module, we conducted a series of experiments with

TABLE 4 ESRGAN hyperparameters configuration results. Experiment values highlighted in bold present the best overall performances.

	Exp #	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Exp 7 (best)	Exp 8	Exp 9	Exp 10
Generator	**RRDB**	8	16	24	14	32	12	**28**	16	32	16
	Residual scaling	0.1	0.2	0.15	0.3	0.2	0.1	**0.15**	0.3	0.1	0.2
	Initialisation variance	0.01	0.05	0.03	0.01	0.025	0.015	**0.03**	0.01	0.015	0.02
Discriminator	**Conv layer**	5	7	12	8	9	28	**14**	10	24	32
Hyperparameters	**Optimiser**	Adam	Adam	Adam	RMSProp	SGD	SGD	**Adam**	RMSProp	RMSProp	SGD
	Learning rate	1.00E−04	1.00E−04	1.00E−04	2.00E−04	2.00E−04	1.00E−04	**1.00E−04**	2.00E−04	2.00E−04	1.00E−04
Loss	**Adversarial loss**	0.37	0.21	0.05	0.29	0.19	0.09	**0.01**	0.12	0.02	0.08
	Content loss	0.25	0.19	0.11	0.23	0.17	0.07	**0.03**	0.13	0.07	0.09
ISIC2019	**PSNR (dB)**	31.8	33.49	36.08	33.01	38.56	34.48	**40.10**	39.01	40.01	37.09
	SSIM	0.7912	0.8401	0.7416	0.8699	0.9322	0.9187	**0.9641**	0.9557	0.9618	0.9417
	PI	0.4551	0.3983	0.2513	0.3517	0.1901	0.1696	**0.1457**	0.1503	0.1935	0.1733
	SCC	0.7709	0.8402	0.9003	0.8811	0.9326	0.9786	**0.9978**	0.9819	0.9912	0.9666
	NIQE	0.2988	0.2561	0.2156	0.2411	**0.1913**	0.2137	0.1948	0.2419	0.2316	0.2012
HAM10000	**PSNR**	32.80	36.91	35.48	38.56	37.09	38.24	**39.01**	39.51	36.46	38.72
	SSIM	0.6812	0.7591	0.8524	0.9111	0.8965	0.9518	**0.9715**	0.9611	0.9701	0.9418
	PI	0.3681	0.2796	0.2037	0.1633	0.1499	0.2316	0.1401	0.1936	**0.1309**	0.1463
	SCC	0.8107	0.8697	0.9107	0.8922	0.9766	0.971	**0.9866**	0.9801	0.9691	0.9631
	NIQE	0.2852	0.2651	0.2149	0.2496	0.1967	0.1878	**0.1528**	0.1687	0.1599	0.2112

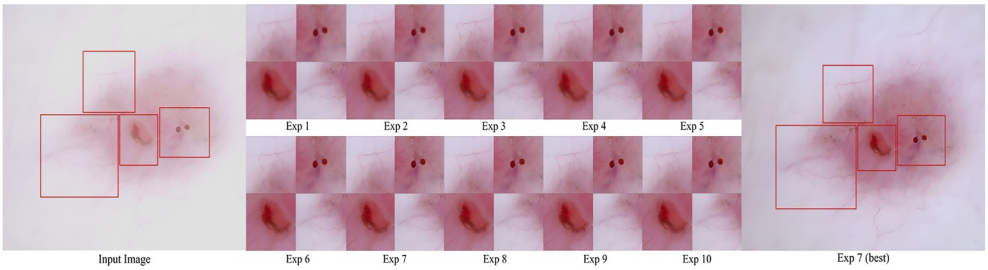


FIGURE 5 Different ESRGAN configured images. ESRGAN, enhanced super-resolution generative adversarial network.

TABLE 5 Comparison with pre-trained GAN-based image enhancer models. Bold values indicate the best results.

		ISIC2019					HAM10000				
Models	**Type**	**PSNR (dB)**	**SSIM**	**PI**	**SCC**	**NIQE**	**PSNR**	**SSIM**	**PI**	**SCC**	**NIQE**
SRGAN [39]	Pre-trained	38.05	0.9306	0.1701	0.9799	0.2197	37.67	0.9703	0.1593	0.9817	0.1667
BigGAN [39]	Pre-trained	37.91	0.9188	0.2391	0.8983	0.2569	32.19	0.9119	0.2119	0.9153	0.2257
HDR-GAN [39]	Pre-trained	38.57	0.9441	0.1611	0.9618	0.2115	35.78	0.9697	0.1638	0.9667	0.1879
StyleGAN [39]	Pre-trained	33.70	0.8914	0.2001	0.9367	0.2639	31.09	0.9319	0.2019	0.9173	0.2117
DeblurGAN [39]	Pre-trained	35.17	0.9217	0.1864	0.9701	0.2399	35.19	0.9611	0.1789	0.9337	0.1896
Our fine-tuned ESRGAN	**Experimented**	**40.10**	**0.9641**	**0.1457**	**0.9978**	**0.1948**	**39.01**	**0.9715**	**0.1401**	**0.9866**	**0.1528**

varying key parameters including the number of histogram bins, the colour space, and the use of dimensionality reduction techniques such as PCA and t-SNE. The clustering performance was evaluated using the Davies–Bouldin index (DBI) and the silhouette score (SS), with a lower DBI and higher SS indicating better clustering performance. The results, as shown in Table 6, indicate that the choice of these parameters significantly impacts the clustering outcomes.

Our analysis of the experimental results reveals several key findings. The best overall performance was achieved in Experiment 8, which utilised a bin size of (8, 8, 8) in the RGB colour space, combined with both PCA and t-SNE. This configuration resulted in the lowest DBI of 0.0588 for the ISIC2019 dataset and 0.0497 for the HAM10000 dataset, along with the highest silhouette scores of 0.9303 and 0.9515, respectively. The consistent improvement in clustering performance when employing PCA and t-SNE suggests that dimensionality reduction and non-linear feature transformations can effectively capture the underlying structure of the skin lesion data, leading to better clustering outcomes. Furthermore, it is observed that the RGB colour space generally outperformed the HSV colour space, particularly when combined with PCA and t-SNE.

Moreover, while increasing the bin size from (8, 8, 8) to (16, 16, 16) did not consistently improve the performance for the ISIC2019 dataset, Experiment 16 with the (16, 16, 16) bin size in the RGB colour space, combined with PCA and t-SNE, achieved the lowest DBI of 0.0441 for the HAM10000 dataset. However, apart from this exception, the (8, 8, 8) bin size with the RGB colour space and applying PCA and t-SNE (Experiment 8) outperformed all other configurations across both datasets. This finding suggests that the (8, 8, 8) bin size may be more appropriate for capturing the relevant colour information in the skin lesion images for most cases. Thus, Experiment 8 is chosen for the subsequent stages of the analysis.

4.4.4 | Clustering performance

Our k-means clustering approach demonstrates notable performance across the diverse skin lesion categories within the ISIC2019 and HAM10000 datasets. The precision is 88.04% for the ISIC2019 dataset and 90.43% for the HAM10000 dataset. The clustering algorithm also exhibits a good recall, illustrating its sensitivity in correctly identifying true positive cases. The mean average recall for the ISIC2019 dataset is 86.84%, whereas for the HAM10000 dataset, it reaches 90.38%. A micro-average precision of 88.3% and 90.43% was achieved for the ISIC2019 and HAM10000 datasets, respectively. Table 7 provides detailed performance metrics for each class for both datasets.

The macro-average recall was 88.0% for the ISIC2019 dataset and 90.38% for the ISIC2019 and HAM10000 datasets. In some classes, as seen in Table 7, the precision values are slightly lower than the corresponding F1-Score. This occurs because for these classes; while effectively identifying a high proportion of true positive instances, our classifier includes some false positive instances leading to slightly lower precision. For instance, in the ISIC2019 dataset, the class

TABLE 6 Performance with different histogram feature extraction configurations. The highest scores for the respective metrics are highlighted in bold, with Experiment 8 providing the best overall configurations.

					ISIC2019		HAM10000	
Exp#	**BINS**	**Colour space**	**PCA**	**t-SNE**	**DBI**	**SS**	**DBI**	**SS**
Exp1	(8, 8, 8)	HSV	X	X	0.2978	0.3983	0.2403	0.4201
Exp2	(8, 8, 8)	HSV	X	✓	0.1771	0.5417	0.1417	0.5899
Exp3	(8, 8, 8)	HSV	✓	X	0.2199	0.6591	0.1713	0.7013
Exp4	(8, 8, 8)	HSV	✓	✓	0.1011	0.8179	0.0701	0.8427
Exp5	(8, 8, 8)	RGB	X	X	0.2359	0.4101	0.1993	0.4591
Exp6	(8, 8, 8)	RGB	X	✓	0.1107	0.5996	0.0997	0.6577
Exp7	(8, 8, 8)	RGB	✓	X	0.1917	0.6417	0.1574	0.7012
Exp8	**(8, 8, 8)**	**RGB**	✓	✓	**0.0588**	**0.9303**	0.0497	**0.9515**
Exp9	(16, 16, 16)	HSV	X	X	0.3517	0.3485	0.2978	0.3532
Exp10	(16, 16, 16)	HSV	X	✓	0.2301	0.4733	0.2190	0.5416
Exp11	(16, 16, 16)	HSV	✓	X	0.1119	0.5101	0.0981	0.5751
Exp12	(16, 16, 16)	HSV	✓	✓	0.1003	0.7970	0.0619	0.8721
Exp13	(16, 16, 16)	RGB	X	X	0.2599	0.4017	0.2113	0.4337
Exp14	(16, 16, 16)	RGB	X	✓	0.1507	0.3913	0.1283	0.4510
Exp15	(16, 16, 16)	RGB	✓	X	0.2317	0.4218	0.1891	0.4701
Exp16	(16, 16, 16)	RGB	✓	✓	0.0757	0.8419	**0.0441**	0.8891

TABLE 7 Classification performance: The best results for the respective classes are highlighted in bold.

		ISIC2019			HAM10000		
Algorithms	**Classes**	**Precision (%)**	**Recall (%)**	**F1-score (%)**	**Precision (%)**	**Recall (%)**	**F1-score (%)**
Cluster-based similarity partitioning	MEL	73.22	78.30	75.67	60.59	83.30	70.15
	NV	72.02	83.44	77.31	70.11	63.34	66.55
	BCC	64.69	58.85	61.63	58.73	65.00	61.71
	AK	79.25	**87.20**	83.04	71.83	64.38	67.90
	BKL	81.44	70.97	75.85	83.89	63.34	72.18
	SCC	58.55	40.96	48.20	-	-	-
	Mean Avg. (%)	71.53	69.96	70.28	69.03	67.87	67.70
	Accuracy (%)	72.19			67.92		
Invariant information clustering	MEL	83.98	82.28	83.12	**86.93**	**94.40**	**90.51**
	NV	73.80	80.25	76.89	80.14	68.25	73.72
	BCC	68.93	68.28	68.60	71.45	77.56	74.38
	AK	73.04	84.96	78.55	84.96	**96.00**	90.14
	BKL	81.76	74.32	77.86	**94.03**	86.49	90.10
	SCC	74.55	60.12	66.56	-	-	-
	Mean Avg. (%)	76.01	75.04	75.26	83.50	84.54	83.77
	Accuracy (%)	76.13			83.78		
Gaussian mixture models	MEL	64.88	52.24	57.88	55.09	63.90	59.17
	NV	62.78	59.40	61.04	66.60	60.34	63.32
	BCC	51.07	58.40	54.49	56.41	57.67	57.03
	AK	52.17	69.44	59.58	72.09	72.63	72.36
	BKL	74.53	58.28	64.41	66.92	62.53	64.65
	SCC	45.60	53.96	49.43	-	-	-
	Mean Avg. (%)	58.50	58.62	57.80	63.42	63.41	63.31
	Accuracy (%)	59.13			63.14		
k-means clustering	MEL	**87.15**	**90.20**	**88.65**	86.84	91.10	88.81
	NV	**86.64**	**89.90**	**88.24**	**92.95**	**92.25**	**92.60**
	BCC	**89.92**	**88.80**	**89.35**	**87.20**	**89.34**	**88.26**
	AK	**92.20**	84.72	**88.30**	**92.02**	89.38	**90.68**
	BKL	**85.93**	**85.32**	**85.62**	93.12	**89.82**	**91.44**
	SCC	**86.40**	**82.12**	**84.20**	-	-	-
	Mean Avg. (%)	88.04	86.84	87.40	90.43	90.38	90.36
	Accuracy (%)	87.77			90.50		

MEL has a precision and recall of 87.15% and 90.2%, respectively. The F1-Score, being the harmonic mean of them, balances the two metrics and stands at 88.65%. However, since both the F1-Score and recall value are higher than the precision value, it indicates that the classifier has included some false positive instances as well, which lowers the precision to 87.15%. This pattern is also observed in other classes. As it is seen, in the HAM10000 dataset, the class BCC has a precision of 87.2% and a recall of 89.34%, resulting in an F1-Score of 88.26%. Here, the higher recall value indicates the model's sensitivity in capturing true positives, but the inclusion of false positive instances lowers the precision. The overall balance between precision and recall, as reflected in the strong F1-Score, underscores the model's effectiveness in maintaining a high true positive rate while managing false positives. Finally, the overall accuracy measured through cluster matching was 87.77% for the ISIC2019 dataset and 90.5% for the HAM10000 dataset.

To validate our approach, we conducted experiments with other clustering algorithms, as presented in Table 7. One of the

methods we explored is cluster-based similarity partitioning (CBSP), which aims to partition data by grouping similar data points based on a similarity metric. While CBSP achieves a high recall of 87.20% for the AK class in the ISIC2019 dataset, its performance across other classes and the HAM10000 dataset is suboptimal. The primary challenge with this algorithm is its difficulty in handling high intra-class variance and the visual similarity of lesions from different classes. For instance, the algorithm achieves only 58.55% precision and 40.96% recall for the SCC class in the ISIC2019 dataset. Similar trends are observed in the HAM10000 dataset, where CBSP's performance is inconsistent and generally lower than other methods.

Our experiments show that invariant information clustering (IIC) performs better than CBSP but still has notable limitations. IIC's mean average precision and recall across all classes in the ISIC2019 dataset are 76.01% and 75.04%, respectively, resulting in a mean F1-Score of 75.26%. In the HAM10000 dataset, IIC's performance improves with a mean average precision of 83.50%, recall of 84.54%, and an F1-Score of 83.77%. Although IIC demonstrates better consistency than CBSP, it still struggles with certain classes and fails to match the performance of our proposed algorithm. For example, whereas IIC achieves a recall of 96.00% for the AK class and precision of 94.03% for the BKL class in the HAM10000 dataset, its performance for other classes in both datasets and overall accuracy (76.13% for ISIC2019 and 83.78% for HAM10000) indicates room for improvement.

The Gaussian mixture models (GMM) algorithm exhibits the lowest performance among the experimented algorithms. In the ISIC2019 dataset, GMM's precision ranges from 45.60% to 74.53%, and recall ranges from 52.24% to 69.44%. This results in a mean average precision of 58.50%, recall of 58.62%, and an F1-Score of 57.80%. The primary drawback of GMM is its assumption of a Gaussian distribution, which may not hold for the complex and varied distribution of skin lesion data. For instance, GMM achieves only 52.24% recall for the MEL class and 58.40% precision for the BCC class in the ISIC2019 dataset.

In contrast, our proposed k-means clustering algorithm demonstrates superior performance across both datasets for all classes. For the ISIC2019 dataset, k-means achieves a mean average precision of 88.04%, recall of 86.84%, and an F1-Score of 87.40%. In the HAM10000 dataset, the performance is even more impressive, with a mean average precision of 90.43%, recall of 90.38%, and an F1-Score of 90.36%. The overall accuracy of k-Means is 87.77% for the ISIC2019 dataset and 90.50% for the HAM10000 dataset, outperforming all the experimented algorithms.

4.4.5 | Comparison with contrastive learning techniques

To assess the performance of our proposed histogram-based unsupervised feature extraction approach, we compared it against several supervised and self-supervised contrastive learning techniques.

The unsupervised contrastive learning methods, swapping assignments between views (SwAV), momentum contrast (MoCo), and information noise-contrastive estimation (InfoNCE), exhibited relatively poor results across both the ISIC2019 and HAM10000 datasets. On the ISIC2019 dataset, their normalised mutual information (NMI) scores ranged from 0.8281 for InfoNCE to 0.8598 for MoCo, and their adjusted Rand index (ARI) scores were between 0.8134 for InfoNCE and 0.8338 for MoCo. The performance of these unsupervised methods was similarly weak on the HAM10000 dataset, with NMI scores from 0.8219 for InfoNCE to 0.8604 for MoCo, and ARI scores from 0.8032 for InfoNCE to 0.8297 for MoCo.

Among the self-supervised contrastive learning techniques, simple contrastive learning for visual representations (SimCLR) achieved the highest NMI score of 0.9407 on the ISIC2019 dataset. However, its ARI score of 0.8993 on the ISIC2019 dataset was lower than our proposed histogram-based method's ARI of 0.9072. Additionally, SimCLR's performance on the HAM10000 dataset was not as strong, with an NMI of 0.9201 and an ARI of 0.9013, which were lower than our histogram-based method's results. On the other hand, the Barlow Twins method stood out by achieving the highest ARI score of 0.9301 on the HAM10000 dataset. However, its NMI score of 0.9487 on the HAM10000 dataset, while still high, was lower than our histogram-based method's score of 0.9507.

Table 8 illustrates that, in contrast to the mixed results of the contrastive learning techniques, our proposed histogram-based unsupervised feature extraction method consistently outperformed the competition across both datasets and metrics. On the ISIC2019 dataset, our approach achieved an NMI of 0.9311 and an ARI of 0.9072, surpassing the best-performing contrastive learning method, SimCLR, on the ARI metric. On the HAM10000 dataset, our method obtained the highest NMI of 0.9507 and the second-highest ARI of 0.9219, outperforming all the contrastive learning techniques in terms of NMI.

4.4.6 | Comparison with state-of-the-art transfer learning models

To benchmark against established techniques, we employed several state-of-the-art (SOTA) models for feature extraction:

TABLE 8 Comparison with the contrastive learning techniques. The NMI and ARI scores with the highest results are presented in bold.

		ISIC2019		HAM10000	
Models	**Category**	**NMI**	**ARI**	**NMI**	**ARI**
SwAV [40]	Unsupervised	0.8779	0.8491	0.8863	0.8626
SimCLR [41]	Self-supervised	**0.9407**	0.8993	0.9201	0.9013
BYOL [42]	Self-supervised	0.9167	0.8805	0.8954	0.9177
MoCo [43]	Unsupervised	0.8598	0.8338	0.8604	0.8297
InfoNCE [44]	Unsupervised	0.8281	0.8134	0.8219	0.8032
Barlow twins [45]	Self-supervised	0.9244	0.8876	0.9487	**0.9301**
Ours	Unsupervised	0.9311	**0.9072**	**0.9507**	0.9219

VGG16, VGG19, ResNet50, ResNet101, EfficientNet-B7, InceptionV3, NASNet, AlexNet, Xception, and DenseNet121 [35]. Each model has a unique architecture and capabilities for feature extraction from skin lesion images. We evaluated the clustering performance of our approach against these models and compared accuracy, precision, recall, and F1-Scores. Among the models considered, EfficientNet-B7 had the highest precision and recall rates, a precision of 84.96% and a recall of 83.14% for the ISIC2019 dataset and a precision of 87.10% and a recall of 87.63% for the HAM10000 dataset. NASNet exhibited the lowest precision and recall rates, with 65.69% precision and 67.52% recall for the ISIC2019 dataset and 70.38% precision and 70.79% recall for the HAM10000 dataset. Models such as AlexNet and VGG16 demonstrated a moderate performance, whereas models such as ResNet101 and DenseNet121 had higher precision and recall rates. was in between. However, our proposed method outperformed all of these state-of-the-art models in terms of precision and recall. Table 9 provides a comparison of the performance of our model with the various SOTA transfer learning models.

Our approach also achieved superior accuracy for both datasets. Although various SOTA models demonstrated respectable accuracies, ranging from 68.87% to 84.96% for the ISIC2019 dataset and 71.41%–86.09% for the HAM10000 dataset, our proposed method outperformed them with accuracies of 87.77% for the ISIC2019 dataset and 90.5% for the HAM10000 dataset. Figure 6 shows the accuracies for both datasets.

4.4.7 | Confusion matrix result

The confusion matrices for both the ISIC2019 and HAM10000 datasets provide valuable insight into the performance of our unsupervised clustering approach in classifying skin lesion images. For the ISIC2019 dataset, out of 5000 samples of melanoma, 4510 instances were clustered into the same group. For the nevus and basal cell carcinoma, 5394 out of 6000 and 3552 out of 4000 samples, respectively, were clustered in the same group. Our approach achieved greater success in the HAM10000 dataset, where out of 1000 samples of melanoma, 911 instances were correctly clustered. Likewise, for nevus and basal cell carcinoma, 1107 out of 1200 and 804 out of 1000 samples, respectively, were successfully clustered in the same group. Although the model performs well in most classes, there are some instances of misclassification, particularly in classes like actinic keratosis and benign keratosis lesion. The confusion matrix, shown in Figure 7, provides a visual representation of the performance of our clustering method on both datasets.

We also examined the confusion matrices for the other clustering algorithms. As shown in Figure 8, the CBSP algorithm struggles to accurately cluster the diverse skin lesion classes, with significant overlap between classes like NV and BCC. Even though the IIC algorithm exhibits more consistent performance, it still has difficulty distinguishing certain classes, particularly BCC and AK. However, with a high degree of misclassification across multiple classes, the GMM algorithm showed the poorest performance.

4.4.8 | Feature map visualisation

The histogram feature map is a representation of the distribution of pixel intensities within an image. It provides valuable insight into the overall characteristics of an image, showing the prevalence of different intensity levels. In our analysis, we utilise histograms to capture the diversity of pixel intensities in images of various skin lesions. Figure 9 depicts examples of images and their corresponding histograms for the different classes: MEL, NV, BCC, AK, BKL, and SCC.

The histograms' horizontal axis spans pixel intensities from 0 to 255, capturing the full spectrum of values. The vertical axis

TABLE 9 Comparison with state-of-the-art transfer learning models. The best scores are presented in bold.

	ISIC2019				HAM10000			
Feature extraction models	**Avg. Time per epoch**	**Precision**	**Recall**	**F1-score**	**Avg. Time per epoch**	**Precision**	**Recall**	**F1-score**
VGG16	62 s	71.96	70.10	70.02	57 s	78.06	77.12	77.59
VGG19	59 s	73.40	72.92	73.16	63 s	81.11	80.42	80.76
ResNet50	48 s	78.01	80.16	79.07	48 s	82.14	83.21	82.67
ResNet101	54 s	81.87	81.66	81.77	62 s	84.39	85.47	84.93
EfficientNet-B7	64 s	84.96	83.14	84.04	67 s	87.10	87.63	87.86
InceptionV3	46 s	76.32	76.68	76.5	58 s	78.93	78.36	78.64
NASNet	62 s	65.69	67.52	66.60	41 s	70.38	70.79	70.58
AlexNet	56 s	73.21	73.95	73.58	59 s	76.81	76.33	76.43
Xception	45 s	83.32	83.88	83.6	52 s	86.96	87.16	87.06
DenseNet121	67 s	83.29	82.86	83.07	71 s	85.62	86.59	86.10
Proposed	-	**86.84**	**87.4**	**87.12**	-	**90.38**	**90.36**	**90.37**

represents the frequency of different intensities. Examining these histograms reveals critical patterns for distinguishing features in skin lesion images.

4.4.9 | Clustering result

Our clustering strategy effectively grouped skin lesion images into six distinct clusters. Utilising PCA and t-SNE for dimensionality reduction, particularly with the colour histogram feature and mean-based methods for clustering refinement, our approach demonstrated the ability to discern inherent patterns and structures within the skin lesion datasets.

Figure 10 provides a visual representation of how our algorithm clustered similar skin lesions. Each cluster corresponds to a specific category of skin lesion conditions. While emphasising the algorithm's discriminating capacity in distinguishing different types of skin lesions, inter-cluster dissimilarity is well-maintained as well. Figure 11 shows the distribution of instances for the clusters.

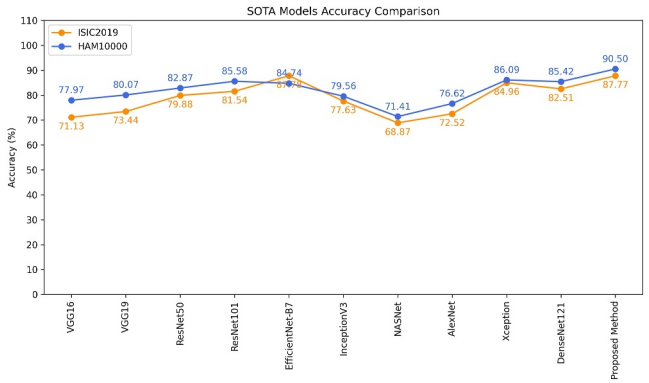
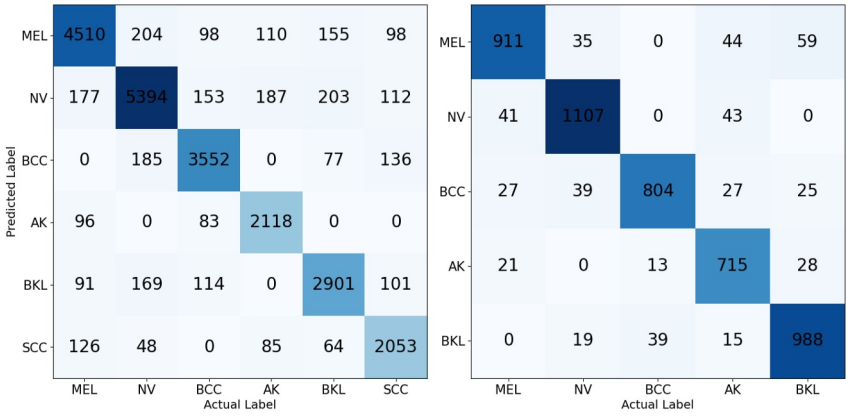


FIGURE 6 Accuracy comparison with several SOTA models. SOTA, state-of-the-art.

5 | DISCUSSION

Our research presents significant findings in medical image analysis through a comprehensive study of dermatoscopic images. It highlights the effectiveness of unsupervised learning methods, specifically, k-means clustering combined with colour histograms, and a modified generative adversarial network, for subtype classification. Through extensive experimentation, we systematically modified the ESRGAN architecture to achieve the best possible image fidelity and feature visibility for our skin lesion analysis task. By conducting benchmarking analyses against other pre-trained GAN-based image enhancement models using quantitative and qualitative assessments, we proved its ability to better preserve the structural and contextual information within the skin lesion images. Our methodology incorporates colour histograms as feature descriptors which captures important information representative colour, texture, and other contextual details. To further enhance the discriminability of our histogram features extraction module, we explored various configurations experimenting with different numbers of bins and colour spaces, which allowed us to extract highly informative representations. We conducted a comprehensive comparative study to assess the effectiveness of our histogram-based feature extraction approach against several transfer learning and contrastive learning models. Moreover, we incorporated normalisation and dimensionality reduction techniques to project the high-dimensional histogram features onto a more manageable subspace. This pre-processing step proved crucial, as it enabled the k-means clustering algorithm to more effectively identify the inherent groupings within the skin lesion data, and outperform other experimented clustering algorithms in our comprehensive comparative analysis. The effectiveness of our unsupervised approach is further emphasised by the comprehensive evaluations conducted across two skin lesion datasets, ISIC2019 and HAM10000. Through in-depth analysis of these benchmarks, we systematically assessed the performance of



(a) ISIC 2019 Dataset (b) HAM10000 Dataset

FIGURE 7 Confusion matrix for cluster matching.

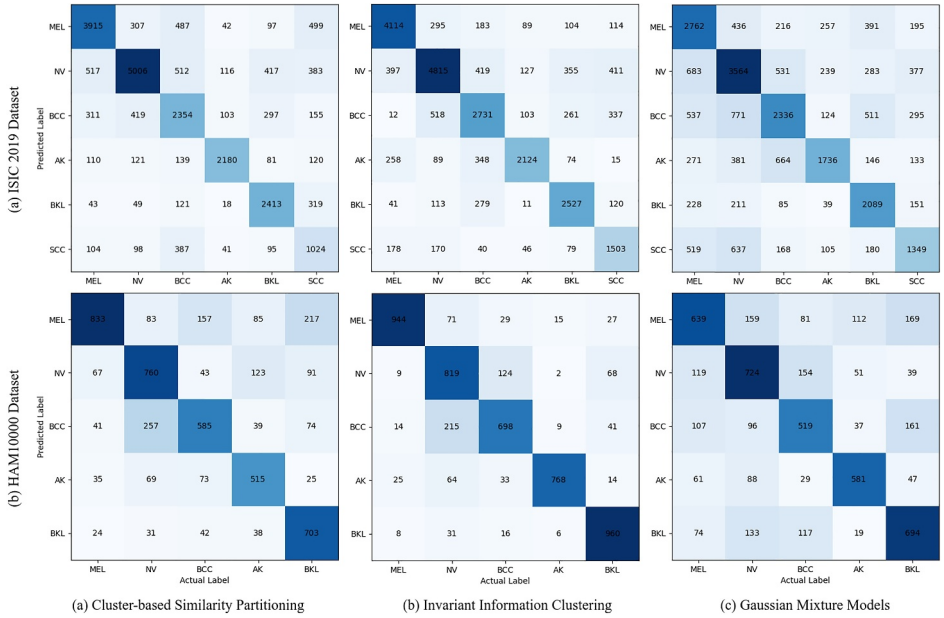


FIGURE 8 Confusion matrix for other clustering algorithms.

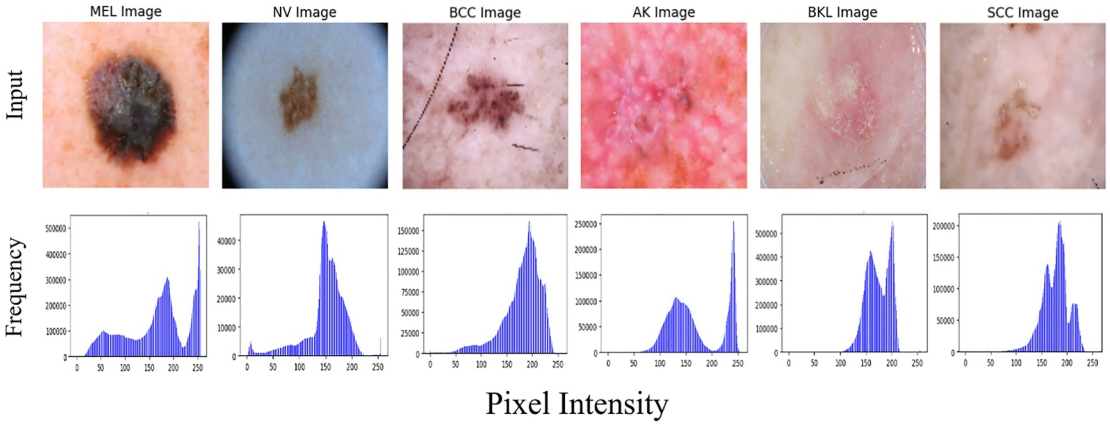


FIGURE 9 Histogram feature map.

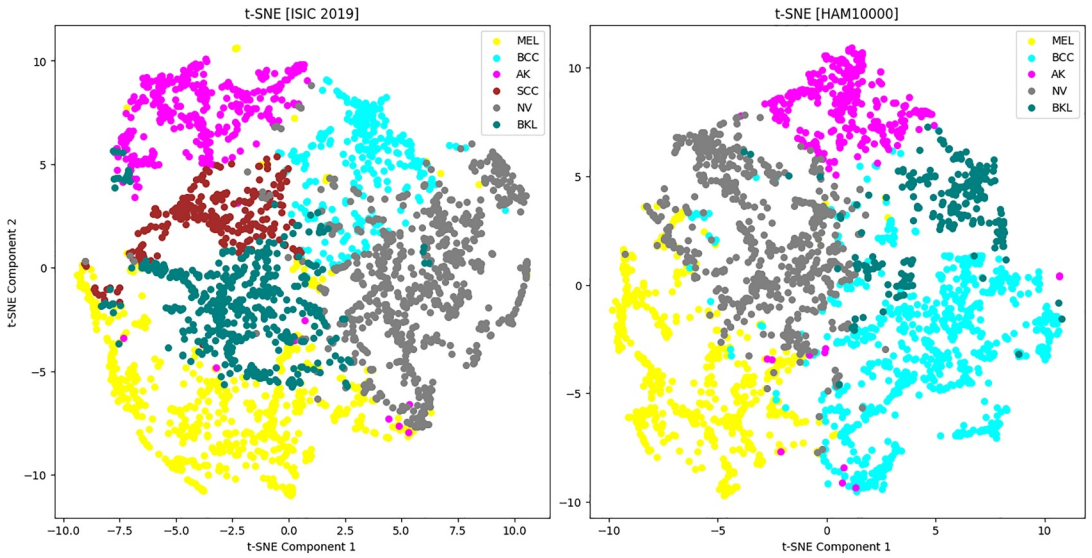


FIGURE 10 Cluster visualisation.

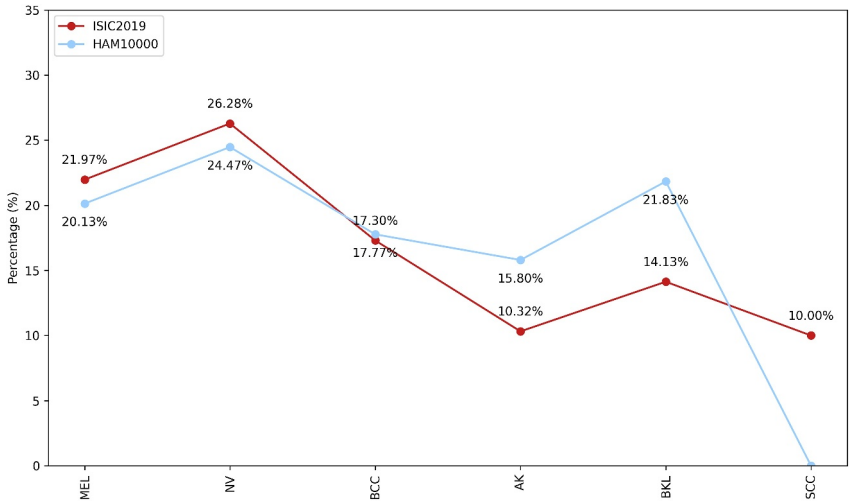


FIGURE 11 Distribution of clusters.

our novel methodology using a variety of evaluation metrics, with our results demonstrating impressive classification accuracies of 87.77% and 90.5% on the ISIC2019 and HAM10000 datasets, respectively.

Despite the significant advancements made in our study, certain limitations exist. One limitation is the inherent imbalance within the datasets, particularly evident in the disproportionately low representation of the dermatofibroma (DF) and vascular lesion (VASC) classes. These classes comprise only 0.94% and 1% of the ISIC2019 dataset and 1.15% and 1.42% of the HAM10000 dataset, respectively. This imbalance could impact the effectiveness of our classification model, potentially leading to biased outcomes and reduced overall performance. We have excluded these classes from our analysis to mitigate the risk of skewed results.

Our study primarily focuses on medical image datasets; future research will explore the inclusion of multi-domain image datasets. This expansion aims to evaluate the generalisability and scalability of our methodology beyond dermatoscopic images. This may provide valuable insight into the adaptability of our approach and its potential applications across different areas.

Despite these limitations, our research significantly addresses key challenges in dermatoscopic image analysis and introduces innovative methodologies. Specifically, our unsupervised learning approach offers a potential solution to the long-standing reliance on labelled data for classification tasks which previous studies have failed to adequately address. It bypasses the need for labour-intensive annotation processes and costly labelling efforts. This can reduce the burden of data labelling and ensure scalability and cost-effectiveness in medical image analysis methodologies. Furthermore, our method reduces dependence on complex pre-trained models, making it more accessible and interpretable for clinicians and researchers. Ultimately, our research aims to enhance diagnostic capabilities and patient care in dermatology by providing robust and accessible solutions for dermatoscopic image analysis.

6 | CONCLUSION

This study presents a novel approach to dermatoscopic image analysis by utilising unsupervised learning methods to classify different subtypes of skin lesions. We demonstrate the potential of unsupervised techniques to identify patterns within the data by employing k-means clustering with colour histograms for feature extraction. We have demonstrated the effectiveness of our proposed methodology, achieving accuracy rates of 87.77% and 90.5%, for the ISIC2019 and HAM10000 datasets, respectively. The use of unsupervised learning methods eliminates the need for time-consuming manual annotation and complex pre-trained feature extraction models, thus reducing computational overhead. By providing clinicians with advanced tools for accurate and efficient diagnosis without requiring prior knowledge of the dataset, our methodology may help to enable earlier disease detection and improve patient outcomes. The scalability and interpretability of our approach mean that it has the potential to be made adaptable to diverse clinical settings.

ACKNOWLEDGEMENTS

This study does not include any external funding. Open access publishing facilitated by Charles Darwin University, as part of the Wiley - Charles Darwin University agreement via the Council of Australian University Librarians.

CONFLICT OF INTEREST STATEMENT

The authors state that they have no known financial conflicts of interest or personal relationships that could have influenced the work presented in this paper.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in the ISIC2019 dataset at https://www.kaggle.com/datasets/andrewmvd/isic-2019 and the HAM10000 dataset at https://www.kaggle.com/datasets/kmader/skin-cancer-mnist-ham10000.

CODE AVAILABILITY

https://github.com/mak-raiaan/UnsupervisedMedicalImaging/.

ORCID

Sami Azam https://orcid.org/0000-0001-7572-9750

REFERENCES

1. Ali, K., et al.: Multiclass skin cancer classification using EfficientNets–a first step towards preventing skin cancer. Neuroscience Infor 2(4), 100034 (2022). https://doi.org/10.1016/j.neuri.2021.100034
2. Nordmann, N., et al.: Effect of gamma knife radiosurgery and programmed cell death 1 receptor antagonists on metastatic melanoma. Cureus 9(12) (2017). https://doi.org/10.7759/cureus.1943
3. Najafpour, Z., et al.: Cost-effectiveness of neuroimaging technologies in management of psychiatric and insomnia disorders: a meta-analysis and prospective cost analysis. J. Neuroradiol. 48(5), 348–358 (2021). https://doi.org/10.1016/j.neurad.2020.12.003
4. Chukwueke, U., Batchelor, T., Brastianos, P.: Management of brain metastases in patients with melanoma. J. Oncol. Pract. 12(6), 536–542 (2016). https://doi.org/10.1200/jop.2016.011882
5. Lekkala, M.R., Mullangi, S.: Malignant melanoma metastatic to the central nervous system. In: StatPearls [Internet]. StatPearls Publishing (2023)
6. Morais, S., et al.: Melanoma brain metastases presenting as delirium: a case report. Arch. Clin. Psychiatry 44(2), 53–54 (2017). https://doi.org/10.1590/0101-60830000000118
7. Rogers, H.W., et al.: Incidence estimate of nonmelanoma skin cancer (keratinocyte carcinomas) in the US population, 2012. JAMA dermatology 151(10), 1081–1086 (2015). https://doi.org/10.1001/jamadermatol.2015.1187
8. Fahad, N.M., et al.: SkinNet-8: an efficient CNN architecture for classifying skin cancer on an imbalanced dataset. In: 2023 International Conference on Electrical, Computer and Communication Engineering (ECCE), pp. 1–6. IEEE (2023)
9. Thörn, M., et al.: Clinical and histopathologic predictors of survival in patients with malignant melanoma: a population-based study in Sweden. JNCI: J. Natl. Cancer Inst. 86(10), 761–769 (1994). https://doi.org/10.1093/jnci/86.10.761
10. Stolz, W.: ABCD rule of dermatoscopy: a new practical method for early recognition of malignant melanoma. Eur. J. Dermatol. 4, 521–527 (1994)
11. Maturo, F., Verde, R.: Combining unsupervised and supervised learning techniques for enhancing the performance of functional data classifiers. Comput. Stat. 39(1), 239–270 (2024). https://doi.org/10.1007/s00180-022-01259-8
12. Song, M., Kim, Y.: Optimizing proportional balance between supervised and unsupervised features for ultrasound breast lesion classification. Biomed. Signal Process Control 87, 105443 (2024). https://doi.org/10.1016/j.bspc.2023.105443
13. Zhang, S., et al.: Generative adversarial and self-supervised dehazing network. IEEE Trans. Ind. Inf. 20(3), 4187–4197 (2023). https://doi.org/10.1109/tii.2023.3316180
14. Li, X., et al.: USL-Net: uncertainty self-learning network for unsupervised skin lesion segmentation. Biomed. Signal Process Control 89, 105769 (2024). https://doi.org/10.1016/j.bspc.2023.105769
15. Sinaga, K.P., Yang, M.-S.: Unsupervised K-means clustering algorithm. IEEE Access 8, 80716–80727 (2020). https://doi.org/10.1109/access.2020.2988796
16. Yang, M.-S., Chang-Chien, S.-J., Nataliani, Y.: A fully-unsupervised possibilistic c-means clustering algorithm. IEEE Access 6, 78308–78320 (2018). https://doi.org/10.1109/access.2018.2884956
17. Murphy, J.M., Maggioni, M.: Unsupervised clustering and active learning of hyperspectral images with nonlinear diffusion. IEEE Trans. Geosci. Rem. Sens. 57(3), 1829–1845 (2018). https://doi.org/10.1109/tgrs.2018.2869723
18. Mrukwa, G., Polanska, J.: DiviK: divisive intelligent K-means for hands-free unsupervised clustering in big biological data. BMC Bioinf. 23(1), 538 (2022). https://doi.org/10.1186/s12859-022-05093-z
19. Zhang, Z., et al.: Flexible auto-weighted local-coordinate concept factorization: a robust framework for unsupervised clustering. IEEE Trans. Knowl. Data Eng. 33(4), 1523–1539 (2019). https://doi.org/10.1109/tkde.2019.2940576
20. Polk, S.L., et al.: Unsupervised diffusion and volume maximization-based clustering of hyperspectral images. Rem. Sens. 15(4), 1053 (2023). https://doi.org/10.3390/rs15041053
21. Han, K., et al.: Autonovel: automatically discovering and learning novel visual categories. IEEE Trans. Pattern Anal. Mach. Intell. 44(10), 6767–6781 (2021). https://doi.org/10.1109/tpami.2021.3091944
22. Yan, Y., et al.: Image clustering via deep embedded dimensionality reduction and probability-based triplet loss. IEEE Trans. Image Process. 29, 5652–5661 (2020). https://doi.org/10.1109/tip.2020.2984360
23. Zhang, H., Peng, Y.: Image clustering: an unsupervised approach to categorize visual data in social science research. Socio. Methods Res., 00491241221082603 (2022)
24. Wang, D., et al.: Unlabeled skin lesion classification by self-supervised topology clustering network. Biomed. Signal Process Control 66, 102428 (2021). https://doi.org/10.1016/j.bspc.2021.102428
25. Khan, S., Huh, J., Ye, J.C.: Variational formulation of unsupervised deep learning for ultrasound image artifact removal. IEEE Trans. Ultrason. Ferroelectrics Freq. Control 68(6), 2086–2100 (2021). https://doi.org/10.1109/tuffc.2021.3056197
26. Sivanesan, U., et al.: Unsupervised medical image segmentation with adversarial networks: from edge diagrams to segmentation maps. arXiv preprint arXiv:1911.05140, (2019)
27. Yu, P., et al.: Unsupervised foreground extraction via deep region competition. Adv. Neural Inf. Process. Syst. 34, 14264–14279 (2021)
28. Cohn, R., Holm, E.: Unsupervised machine learning via transfer learning and k-means clustering to classify materials image data. Integr. Mater. Manuf. Innov. 10(2), 231–244 (2021). https://doi.org/10.1007/s40192-021-00205-8
29. Ge, P., et al.: Dual adversarial autoencoders for clustering. IEEE Transact. Neural Networks Learn. Syst. 31(4), 1417–1424 (2019). https://doi.org/10.1109/tnnls.2019.2919948
30. Zhou, X., et al.: Saliency-CCE: exploiting colour contextual extractor and saliency-based biomedical image segmentation. Comput. Biol. Med. 154, 106551 (2023). https://doi.org/10.1016/j.compbiomed.2023.106551
31. Innani, S., et al.: Generative adversarial networks based skin lesion segmentation. Sci. Rep. 13(1), 13467 (2023). https://doi.org/10.1038/s41598-023-39648-8
32. Kim, W., Kanezaki, A., Tanaka, M.: Unsupervised learning of image segmentation based on differentiable feature clustering. IEEE Trans. Image Process. 29, 8055–8068 (2020). https://doi.org/10.1109/tip.2020.3011269
33. Nazki, H., et al.: Unsupervised image translation using adversarial networks for improved plant disease recognition. Comput. Electron. Agric. 168, 105117 (2020). https://doi.org/10.1016/j.compag.2019.105117
34. Zhang, Y., et al.: Collaborative unsupervised domain adaptation for medical image diagnosis. IEEE Trans. Image Process. 29, 7834–7844 (2020). https://doi.org/10.1109/tip.2020.3006377
35. Wang, X., et al.: Esrgan: enhanced super-resolution generative adversarial networks. In: Proceedings of the European Conference on Computer Vision (ECCV) Workshops (2018)
36. Raiaan, M.A.K., et al.: A lightweight robust deep learning model gained high accuracy in classifying a wide range of diabetic retinopathy images. IEEE Access 11, 42361–42388 (2023). https://doi.org/10.1109/access.2023.3272228

37. Codella, N.C., et al.: Skin lesion analysis toward melanoma detection: a challenge at the 2017 international symposium on biomedical imaging (isbi), hosted by the international skin imaging collaboration (isic). In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), pp. 168–172. IEEE (2018)
38. Tschandl, P., Rosendahl, C., Kittler, H.: The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. Sci. Data 5(1), 1–9 (2018). https://doi.org/10.1038/sdata.2018.161
39. Esfahani, S.N., Latifi, S.: A survey of state-of-the-Art GAN-based approaches to image synthesis. In: Computer Science & Information Technology (CS & IT) Computer Science Conference Proceedings (CSCP), pp. 63–76 (2019)
40. Caron, M., et al.: Unsupervised learning of visual features by contrasting cluster assignments. Adv. Neural Inf. Process. Syst. 33, 9912–9924 (2020)
41. Chen, T., et al.: A simple framework for contrastive learning of visual representations. In: International Conference on Machine Learning, pp. 1597–1607. PMLR (2020)
42. Grill, J.-B., et al.: Bootstrap your own latent-a new approach to self-supervised learning. Adv. Neural Inf. Process. Syst. 33, 21271–21284 (2020)
43. He, K., et al.: Momentum contrast for unsupervised visual representation learning. arXiv e-prints, art. arXiv preprint arXiv:1911.05722 (2019)
44. Oord, A.v.d., Li, Y., Vinyals, O.: Representation learning with contrastive predictive coding. arXiv preprint arXiv:1807.03748 (2018)
45. Zbontar, J., et al.: Barlow twins: self-supervised learning via redundancy reduction. In: International Conference on Machine Learning, pp. 12310–12320. PMLR (2021)

How to cite this article: Rahman, M.A., et al.: Advancing skin cancer detection integrating a novel unsupervised classification and enhanced imaging techniques. CAAI Trans. Intell. Technol. 1–20 (2025). https://doi.org/10.1049/cit2.12410