MethodsX 13 (2024) 102839

Contents lists available at ScienceDirect

MethodsX

journal homepage: www.elsevier.com/locate/methodsx

Transformer-based decoder of melanoma classification using hand-crafted texture feature fusion and Gray Wolf Optimization algorithm

Hemant Kumar[a,*], Abhishek Dwivedi[b], Abhishek Kumar Mishra[c], Arvind Kumar Shukla[d], Brajesh Kumar Sharma[e], Rashi Agarwal[f], Sunil Kumar[a]

[a] Department of Information Technology, School of Engineering & Technology, Chhatrapati Shahu Ji Maharaj University, Kanpur, India
[b] Department of Computer Applications, School of Engineering & Technology, Chhatrapati Shahu Ji Maharaj University, Kanpur, India
[c] Department of Computer Science & Engineering, IFTM University, Moradabad, India
[d] Department of Computer Science & Applications, IFTM University, Moradabad, India
[e] Department of Computer Science & Engineering, Sir Padampat Singhania University, Udaipur, India
[f] Department of Computer Science & Engineering, Harcourt Butler Technical University, Kanpur, India

ARTICLE INFO

Method name:
Melanoma classification using Feature fusion, Gray Wolf Optimization and Transformer based encoder

Keywords:
Melanoma
Transformer
Multi-head attention
Scale-dot product
Gray Wolf Optimization
Gray level concurrence matrix (GLCM)
Linear binary pattern (LBP)

ABSTRACT

Melanoma is a type of skin cancer that poses significant health risks and requires early detection for effective treatment. This study proposing a novel approach that integrates a transformer-based model with hand-crafted texture features and Gray Wolf Optimization, aiming to enhance efficiency of melanoma classification. Preprocessing involves standardizing image dimensions and enhancing image quality through median filtering techniques. Texture features, including GLCM and LBP, are extracted to capture spatial patterns indicative of melanoma. The GWO algorithm is applied to select the most discriminative features. A transformer-based decoder is then employed for classification, leveraging attention mechanisms to capture contextual dependencies. The experimental validation on the HAM10000 dataset and ISIC2019 dataset showcases the effectiveness of the proposed methodology. The transformer-based model, integrated with hand-crafted texture features and guided by Gray Wolf Optimization, achieves outstanding results. The results showed that the proposed method performed well in melanoma detection tasks, achieving an accuracy and F1-score of 99.54% and 99.11% on the HAM10000 dataset, and an accuracy of 99.47%, and F1-score of 99.25% on the ISIC2019 dataset.

- We use the concepts of LBP and GLCM to extract features from the skin lesion images.
- The Gray Wolf Optimization (GWO) algorithm is employed for feature selection.
- A decoder based on Transformers is utilized for melanoma classification.

Specifications table

Subject area:	Computer Science
More specific subject area:	Deep Learning, Biomedical Imaging
Name of your method:	Melanoma classification using Feature fusion, Gray Wolf Optimization and Transformer based encoder
Name and reference of original method:	An Identification Method of Feature Interpretation for Melanoma Using Machine Learning https://www.mdpi.com/2076-3417/13/18/10076
Resource availability:	Dataset: https://challenge.isic-archive.com/landing/2018/ https://www.kaggle.com/datasets/andrewmvd/isic-2019/code

* Corresponding author.
E-mail address: hemantkumar@csjmu.ac.in (H. Kumar).

https://doi.org/10.1016/j.mex.2024.102839
Received 12 February 2024; Accepted 2 July 2024
Available online 3 July 2024

Method details

Introduction

Melanoma, a highly aggressive type of skin cancer, presents considerable diagnostic difficulties due to its intricate and diverse characteristics. The American Cancer Society predicted that 609,360 deaths and approximately 1.9 million new cancer cases will occur in 2022. Currently, skin cancer is a highly dangerous and prevalent disease. Skin cancer can manifest in various forms, including basal cell carcinoma, melanoma, and squamous cell carcinoma. Among these, melanoma is considered the most unpredictable. Timely detection and precise diagnosis play a crucial role in enhancing patient outcomes and survival rates [1].

Deep learning models, specifically Convolutional Neural Networks (CNNs), have demonstrated potential in image-based medical diagnoses by effectively extracting pertinent features from raw image data [2,3]. Nevertheless, convolutional neural networks (CNNs) have certain limitations, such as the requirement for extensive annotated datasets and significant computational resources. The Transformer architecture, initially designed for natural language processing, has shown promise in computer vision tasks by effectively modeling long-range dependencies and capturing intricate patterns within the data [4]. Transformers, equipped with attention mechanisms, present a strong alternative to conventional CNNs, demonstrating enhanced performance in diverse image classification tasks [5].

The integration of hand-crafted texture features with deep learning models in melanoma classification can improve performance by combining domain-specific knowledge with deep learning's strong feature extraction capabilities. Texture features, such as those obtained from the Gray-Level Co-occurrence Matrix (GLCM), are able to capture significant spatial relationships and patterns in skin lesion images. These features offer valuable diagnostic information [6]. The combination of hand-crafted features and deep learning representations can enhance classification accuracy by leveraging the strengths of both approaches. Feature fusion poses challenges in feature selection and optimizing their combination. The Gray Wolf Optimization (GWO) algorithm is a meta-heuristic optimization technique that addresses complex, multi-dimensional problems. It is inspired by the social hierarchy and hunting behavior of gray wolves [7]. The balance between exploration and exploitation in GWO makes it suitable for optimizing the feature fusion process in melanoma classification systems.

This study presents a novel approach for melanoma classification using a transformer-based decoder. We enhance the approach by fusing hand-crafted texture features and optimizing the model using the Gray Wolf Optimization algorithm. Using the attention mechanisms of the Transformer model and texture features from GLCM, the proposed method creates a strong system for finding melanoma. We use the GWO algorithm to optimize feature selection and fusion, maximizing classification performance by utilizing the most relevant features.

Abbreviations

Variable	Description	Variable	Description
GLCM	Gray-Level Co-occurrence Matrix	W_{emb}	embedding matrix
LBP	Local Binary Pattern	e_{pos}	positional encoding
Gwo	Gray-wolf optimization	W^q, W^k and W^v	Weight matrix
I_p, I_c	pth neighbor pixel intensity, Center pixel intensity	Q^i, K^i and V^i	query, key and value
$\mathcal{X}_p$	prey's location vector	h_i	Attention head
$\mathcal{X}$	wolves' position in d-dimensional space	MHSA	Multi-Head Self Attention
AKIEC	Actinic Keratosis	VASC	Vascular Lesion
BCC	Basal Cell Carcinoma	NV	Melanocytic Nevus
BKL	Benign Keratosis	MEL	Melanoma
DF	Skin Fibroma		

Literature review

Abdar et al. [8] employed a three-way decision-based Bayesian deep learning approach for automated boundary detection and shape retrieval. This was followed by the use of GLCM and the Euclidean distance transform for texture feature computation. Qureshi and Roos [9] investigated the use of deep neural network ensembles for transfer learning in order to improve skin cancer detection in scenarios with imbalanced datasets. Araaf et al. [10] proposed a SCDC system that utilized GLCM features at different angles as well as a KNN classifier. However, they restricted their study to binary classification. The authors proposed enhancing feature extraction techniques and augmenting machine learning models with advanced deep learning methods.

Tiwari et al. [11] proposed a technique for melanoma detection using HOSMI-LBP features. The authors' method uses a combination of a Convolutional Neural Network (CNN) and a Neural Network (NN) for classification, along with the Sine Cosine Algorithm-Gray Wolf Optimization (SLI-GWO) algorithm.

H. Ding et al. [12] introduced the Archimedes Optimization Algorithm (AOA), which utilizes a dynamic transfer operator for the purpose of feature selection and classification. The proposed approach combines Gray-Level Co-occurrence Matrix (GLCM) features with the Dynamic Archimedes Optimization Algorithm (DAOA) and Support Vector Machines (SVM). The method achieves an accuracy of 88%, a sensitivity of 96%, and a specificity of 81%.

Mahmoud and Soliman [13] proposed an artificial intelligence (AI) system to detect skin cancer at an early stage. They utilized the Gray Level Co-occurrence Matrix (GLCM) for texture analysis and the ABCDE rules for feature extraction. The authors utilized

Adaptive Segmentation (AS) and Region Growing (RG) algorithms for segmentation and employed Neural Network (NN) and Support Vector Machine (SVM) algorithms for classification. The Artificial Neural Network (ANN) achieved an accuracy of 94%, precision of 96%, specificity of 95.83%, and sensitivity of 92.30%.

Midasala et al. [14] introduced MFEUsLNet, an AI system for skin cancer detection that utilizes unsupervised learning and multi-level feature extraction. The system demonstrated excellent performance on the ISIC-2020 dataset. The proposed model incorporates USL-KMC segmentation, GLCM, RDWT feature extraction techniques, and an RNN classifier.

A study by Magdy et al. [15] introduced two classification methods for distinguishing between benign and malignant tumors in dermoscopic images. These models achieved accuracy rates exceeding 99% in certain cases. The authors employed GLCM and texture features to classify the data. They evaluated different machine learning and deep learning methods, such as artificial neural networks, KNN, SVM, Naïve Bayes, decision trees, and pretrained CNNs like AlexNet and VGG-16.

Manoj et al. proposed a YOLOv3-DCNN architecture for the detection and classification of malignant melanoma. Their methodology entails utilizing YOLOv3 to generate feature maps, QuadHistogram to extract color features, and GLCM and RCT to extract texture features. These features are subsequently combined and pass into a DCNN for classification [16].

Materials and methods

The melanoma classification approach encompasses preprocessing, texture feature extraction (combining GLCM and LBP), feature selection using the Gray Wolf Optimization (GWO) algorithm, and a transformer-based decoder for classification. Fig. 1, depicts the proposed methodology of the melanoma classification.

Preprocessing

In this section, we preprocess the dermoscopic skin images in the following manner: First, it is necessary to resize the skin image to a fixed dimension of 256 × 256. Next, applying a median filter will help enhance the overall image quality. To ensure accuracy, it is crucial to remove any hair from the images using morphological operations. The next step involves applying the k-means clustering algorithm [17] to segment the skin lesion and normal skin regions. Finally, the segmented skin lesion regions can be cropped for further analysis. Fig. 2 illustrates the preprocessing process.

Texture feature extraction

Gray-level Co-occurrence matrix (GLCM)

This research employs textural features extracted from the gray-level co-occurrence matrix (GLCM) [18]. GLCM is a robust feature descriptor that effectively captures the spatial relationships between pixel intensities in images. The matrix $I(i, j, d, \theta)$ measures the frequency of pixel pairings with the same brightness levels, which are separated by distance d and angle θ. The angles $0°, 45°, 90°$, and $135°$, correspond to different spatial orientations, and the distance between pixels is set to one unit [6].

$$I(i, j, d, \theta) = \frac{\text{Number of occurrances of } (i, j) \text{ at distance d and orientation } \theta}{\text{Total occurrences at distance } d \text{ and orientation } \theta} \tag{1}$$

After the GLCM matrix was acquired, several texture characteristics (such as contrast, entropy, homogeneity, energy, and correlation) [19] were extracted to describe the spatial patterns in pictures of skin lesions [20].

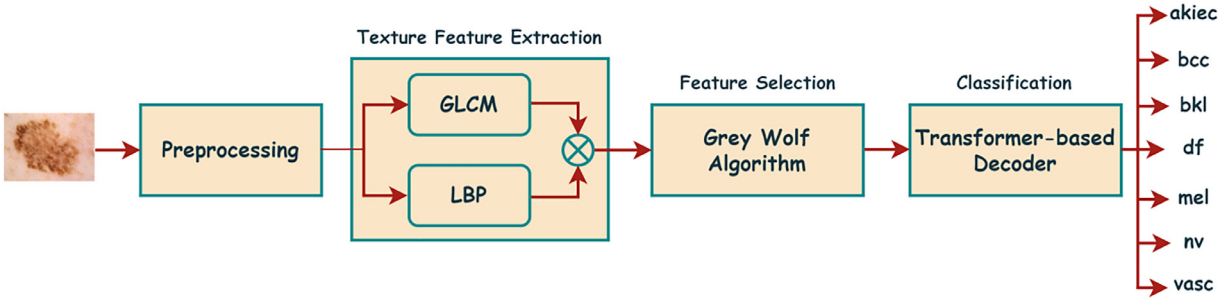

Fig. 1. Proposed Method of Melanoma classification.

Fig. 2. Preprocessing pipeline.

- *Correlation feature* quantifies the linear relationship between pixel intensities

$$Correlation = \sum_{i,j} I(i,j)\ \frac{(i-\mu_i)(j-\mu_j)}{\sigma_i\ \sigma_j} \tag{2}$$

- *Contrast* measure quantifies the local variations in pixel intensities. It is obtained by calculating the weighted sum of the squared difference between pixel intensities in the GLCM.

$$Contrast = \sum_{i,j} I(i,j)(i-j)^2 \tag{3}$$

- *Homogeneity* reflects the proximity of pixel intensities in GLCM.

$$Homogeneity = \sum_{i,j} \frac{I(i,j)}{1+|i-j|} \tag{4}$$

- *Angular Second Moment* measures the uniformity of pixel pairs in an image, specifically focusing on GLCM texture features.

$$Angular\ Second\ Moment = \sum_{i,j} (I(i,j))^2 \tag{5}$$

- *Entropy* quantifies the level of randomness and uncertainty present in an image.

$$Entropy = -\sum_{i,j} I(i,j)\ \log(I(i,j)+\epsilon) \tag{6}$$

where: i, j are the pixel intensities; $I(i,j)$ is the joint probability at certain distance and orientation; μ_i, μ_j are the means of intensities in *i*th and *j*th distance, respectively; σ_i, σ_j are the standard deviations of the intensities in the *i*th and *j*th distance, respectively; ϵ is a small positive constant.

Local binary pattern (LBP)

LBP is a texture descriptor that encodes the relationship between pixel intensities and their neighbors to characterize local patterns [21]. The process entails comparing the central pixel's intensity with its surrounding neighbors in a circular region, resulting in a binary code that represents the local texture pattern [22,23]. The creation of the LBP feature vector involves the following steps:

- *LBP Operator:* Given an image pixel $I(x,y)$ at coordinates (x,y), the LBP operator is

$$LBP_{P,R}(x,y) = \sum_{p=0}^{P-1} s\left(I_p - I_c\right)\ 2^p \tag{7}$$

where P is number of sampling points on circle; R is radius of circle; I_c is intensity of the center pixel at (x,y); I_p is intensity of the *p*th neighbor pixel.; $s(z) = \begin{cases} 1, & z \geq 0 \\ 0, & z < 0 \end{cases}$ is a step function.

The LBP code is generated by comparing the intensities of neighboring pixels with the center pixel and applying a threshold. It indicates if the neighboring pixel is brighter or darker than the center.

- *LBP Histogram:* Subsequently, the LBP codes are computed for each pixel in the image, and a histogram H is generated to represent the distribution of various LBP patterns in the image.

$$H(i) = \text{number of occurrences of LBP code } i \tag{8}$$

- *LBP Uniform Patterns:* LBP codes can be classified into two categories: uniform and non-uniform patterns. *Uniform patterns* are patterns that have a maximum of two transitions (0–1 or 1–0) in their binary representation.
- *Rotation Invariance:* The LBP operator was updated to accommodate circularly shifted patterns for rotation invariance. The rotation with the lowest LBP code.

$$LBP^{ri}_{P,\ R}(x,y) = \min\left\{LBP_{P,R}(x,y,r) | r = 0,\ 1,\ \ldots,\ P-1\right\} \tag{9}$$

where $LBP_{P,R}(x,y,r)$ is the LBP code at rotation r.

Feature selection

A total of 328 texture features, consisting of 256 features based on LBP and 72 features based on GLCM, were extracted from the skin lesion image. This study employs the Gray Wolf Optimization (GWO) algorithm [24] to reduce the dimensionality of the feature vector while maintaining the data quality.

Gray-Wolf optimization

GWOs [7] resemble the hunting strategy employed by a pack of gray wolves. The behavior of gray wolves living in packs of five to twelve wolves is influenced by their social intelligence. This model mimics a leadership system with four levels: α, β, δ, and ω. The significant role of the alpha (α) is to make decisions; the beta (β) assists the α alpha in making decisions and provides feedback. Delta (δ) is a hunter, and omega (ω) wolves must follow all wolves [25]. The circling behavior of GWOs is

$$\mathcal{X}^{t+1} = \mathcal{X}_p^t + \mathcal{A}\ D \tag{10}$$

Here $\mathcal{A}$ is coefficient vectors, $\mathcal{X}_p$ is prey's location vector, $\mathcal{X}$ represents the of wolves' position in d-dimensional space, t is the number of iterations. D is defined as

$$D = \left| C\ \mathcal{X}_p^t - \mathcal{X}^t \right| \tag{11}$$

here, $\mathcal{A}$ and C are evaluated as

$$\mathcal{A} = 2\mathcal{a}\ r_1 - \mathcal{a} \tag{12}$$

$$C = 2\ r_2 \tag{13}$$

The vectors $r_1$and r_2 are randomly generated within the range of $[0, 1]$. The encircling coefficient a decreases from 2 to 0 as the number of rounds increases. Gray wolves observe the alpha as the most suitable candidate for the task, while the beta and delta are expected to possess knowledge about the location of the prey. Until a particular iteration, the top three answers are retained, compelling entities such as omega to modify their positions in the decision space.

The updated locations are

$$\mathcal{X}_{t+1} = \left(x_1 + x_2 + x_3\right)/3 \tag{14}$$

$$x_1 = \mathcal{X}_\alpha - \mathcal{A}_1\ \mathcal{D}_\alpha\ ;\ x_2 = \mathcal{X}_\beta - \mathcal{A}_2\ \mathcal{D}_\beta;\ x_3 = \mathcal{X}_\delta - \mathcal{A}_3\ \mathcal{D}_\delta \tag{15}$$

where, x_1, x_2 and x_3 are the three optimal wolves at an iteration t. $\mathcal{A}_1$, $\mathcal{A}_2$ and $\mathcal{A}_3$ are evaluated using Eq. (12). $\mathcal{D}_\alpha$, $\mathcal{D}_\beta$ and $\mathcal{D}_\delta$ are computed using Eq. (11)

$$\mathcal{D}_\alpha = \left|C_1\ \mathcal{X}_\alpha - \mathcal{X}\right|;\ \mathcal{D}_\beta = \left|C_2\ \mathcal{X}_\beta - \mathcal{X}\right|;\ \mathcal{D}_\delta = \left|C_3\ \mathcal{X}_\delta - \mathcal{X}\right| \tag{16}$$

Where C_1, C_2 and C_3 are evaluated using Eq. (13).

The GWO utilizes the top three choices to update the location of each wolf. Omega wolves comprise a large population and exhibit lower fitness compared to alpha, beta, and delta wolves. Realigning the less dominant wolves can enhance GWO's ability to diversify and yield improved outcomes. Fig. 3 illustrates the flow diagram of GWO.

Classification using transformer-based decoder

In the classification stage, we utilizes the transformer [5,26] based decoder approach, shows in Fig. 4. The input features of the transformer are selected using GWO algorithm.

The input embedding is obtained by multiplication of input matrix X and embedding matrix W_{emb}, and output is added to the positional encoding e_{pos}.

$$X^i = X\ W_{emb} + e_{pos} \tag{17}$$

Matrices W^q, W^k and W^v with X^i, to generate three vectors Q^i, K^i and V^i are query, key and value, respectively.

$$Q^i = W^q\ X^i;\ K^i = W^k\ X^i;\ V^i = W^v\ X^i \tag{18}$$

The attention output for each token is computed by applying the SoftMax function to the scaled dot product of Q^i and K^i. The scaled dot product is multiplied by V^i to create a refined representation that captures the contextual importance of each token in the input. Fig. 5, shows the attention score calculation and scale dot product.

$$Attention\left(Q^i, K^i, V^i\right) = SoftMax\left(\frac{Q^i\ \left(K^i\right)^T}{\sqrt{d_k}}\right) \cdot V^i \tag{19}$$

After obtaining the attention output, we concatenate the outputs from all attention heads using MHSA layer.

$$MHSA = Concat\left(h_1, h_2, \ldots, h_i, \ldots, h_H\right) W^o \tag{20}$$

where, $h_i = Attention(Q^i\ W_i^Q, K^i\ W_i^K,\ V^i\ W_i^V)$; We are using three head in our model.

The output generated by the MHSA layer is then sent to the position wise feed-forward network (FFN). Two linear transformations are sequentially executed in this network, with an activation function inserted between them.

$$FFN(x) = \sigma\left(xW_1 + b_1\right)W_2 + b_2 \tag{21}$$

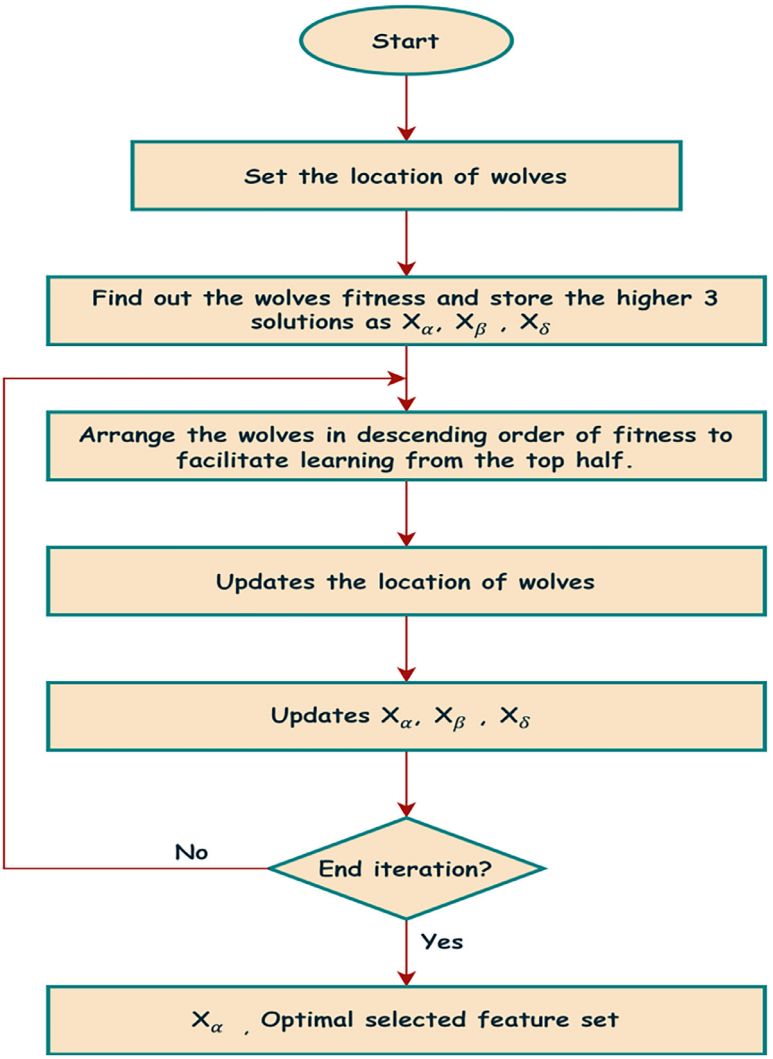

Fig. 3. Gray Wolf Optimization Algorithm.

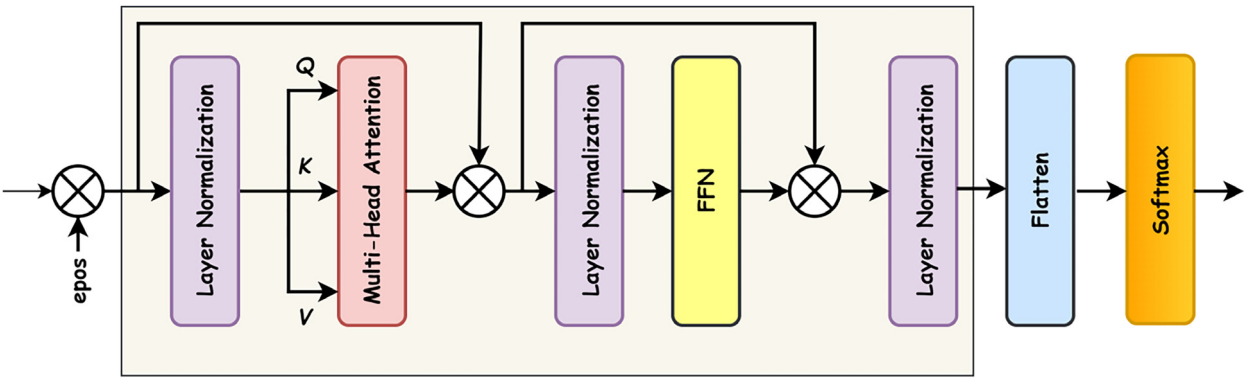

Fig. 4. Transformer-based Decoder.

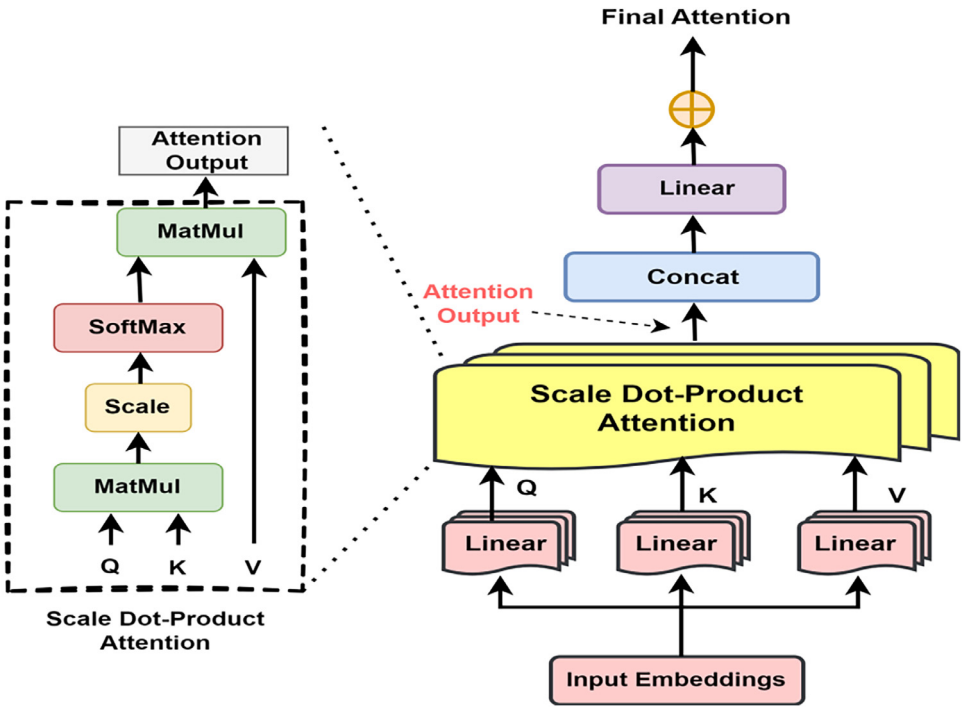

Fig. 5. Attention calculation and scale-dot product.

where, W_1, W_2 are the weights of networks; x is $Output_{MHSA} = MHSA^1, \ldots, MHSA^i, \ldots, MHSA^n$; b_1 b_2 are the bias parameters; σ is the $GeLU(\cdot)$ transfer function. Finally, the result of the FFN layer is processed by an additive normalization layer, generating the final outcome from the single-layer encoder.

$$y_{transformer} = flatten(FFN(x)) \tag{22}$$

Finally, $SoftMax$ activation function predicts the probability distribution of each word y_i in the response.

$$p(y_i|y_0, \ldots, y_{i-1}) = SoftMax(W\ y_{transformer}) \tag{23}$$

Method validation

Dataset

The proposed method was primarily trained on the HAM10000 dataset and its effectiveness was also confirmed on the ISIC2019 dataset. The HAM10000 dataset, consisting of 10,015 skin images categorized into seven lesion types: actinic keratosis (AKIEC), melanocytic nevus (NV), basal cell carcinoma (BCC), melanoma (MEL), vascular lesion (VASC), benign keratosis (BKL), and skin fibroma (DF) [27,28], which is publicly available. The images are sized at 600 × 450 pixels. In Fig. 6, sample images representing each class from the HAM10000 dataset are depicted. The dataset is split into a training set and a test set, following an 80:20 ratio. The ISIC2019 dataset consist of 25,331 images belonging to eight categories: AKIEC, BCC, BKL, DF, MEL, NV, VASC, and squamous cell carcinoma (SCC). The training/test set ratio of 80:20. Training set is 20,269 images and test set are 5062 images. Here is the breakdown of each category in the dataset, as shown in Tables 1 and 2.

Table 1
Dataset distribution in training and test splits of HAM10000 dataset.

Types	AKIEC	NV	BCC	MEL	VASC	BKL	DF	Total
Train	261	5364	411	890	113	879	92	8010
Test	66	1341	103	223	29	220	23	2005
Total	327	6707	514	1113	142	1099	115	10,015

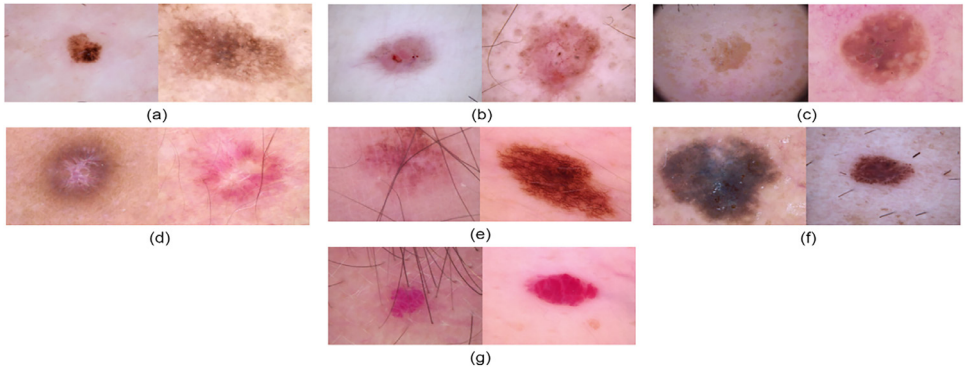

Fig. 6. (a) Actinic Keratosis, (b) Basal Cell Carcinoma, (c) Benign Keratosis, (d) Skin Fibroma, (e) Melanocytic Nevus, (f) Melanoma, (g) Vascular Lesion.

Table 2
Dataset distribution in training and test splits of ISIC2019 dataset.

Types	AKIEC	NV	BCC	MEL	VASC	BKL	DF	SCC	Total
Train	694	10,300	2659	3618	203	2100	192	503	20,269
Test	173	2575	664	904	50	524	47	125	5062
Total	867	12,875	3323	4522	253	2624	239	628	25,331

Experimental details

We utilized PyTorch in our experimental setup on Kaggle. We utilize a fusion of hand-crafted texture features, such as GLCM and LBP. These features have been meticulously chosen based on GWOs, ensuring their suitability for the task. Lastly, we employ a transformer-based classifier for melanoma detection. The training method used categorical cross-entropy loss, and weight updates were performed using the Adam optimizer, which improves model performance by adjusting weight values. The training consisted of 100 epochs and utilized a learning rate scheduler to optimize the learning rate. Table 3, provides a thorough overview of the hyperparameters.

Evaluation metrics

Four generally used measures are utilized to objectively evaluate the effectiveness of the suggested methodology for detecting tomato diseases: precision, recall, F1 score, and accuracy.

$$Accuracy = \frac{P_T + N_T}{P_T + P_F + N_T + N_F} \tag{24}$$

$$Precision = \frac{P_T}{P_T + P_F} \tag{25}$$

$$Recall = \frac{P_T}{P_T + N_F} \tag{26}$$

$$F1 - score = \frac{2 \times (P \times R)}{(P + R)} \tag{27}$$

Where P_T, N_T, P_F, and N_F are true positives, true negatives, false positives and false negative respectively.

Table 3
Parameter for Transformer based classifier.

Parameter	Value
Learning rate	0.0001
Batch size	128
Optimizer	Adam
Epoch	100
Scheduler	Learning Rate Scheduler
Loss function	Categorical cross entropy

Results & discussion

This section validates the effectiveness of the Gray Wolf Optimization (GWO) method for feature selection. We assess the algorithm's performance through experimental testing on the HAM10000 and ISIC2019 datasets. We evaluate the GWO method against Particle Swarm Optimization (PSO) [29], Whale Optimization (WO) [30], Bat Method (BAT) [31], Multi-Verse Optimization (MVO) [32], and Firefly Optimization (FFO) [33]. We use a variety of measures to evaluate the effectiveness of various optimization methods in addressing complex numerical problems. Table 4 displays the parameters associated with each approach. There are 50 search agents and 500 iterations.

The parameters of MVO (Weighted Euclidean Projection and Total Deviation Ratio) require careful adjustment, which can be especially difficult in applications involving feature selection. On the other hand, GWO's linear reduction of variable "a" is more direct and often more efficient. Although WOA's distinctive processes have advantages, GWO's leadership hierarchy provides a more organized approach to selecting the best feature subsets. PSO necessitates meticulous calibration of many parameters (w, c1, c2), which might pose challenges in feature spaces with a high number of dimensions. The advantage of GWO's simplified parameter adjustment is evident. While BAT efficiently utilizes loudness and pulse rate, GWO's balance between exploration and exploitation is more direct and often more effective for feature selection. FFA is good at multimodal optimization, but GWO is better at feature selection tasks because the top solutions and adaptive parameters interact in a way that makes the system more balanced and stable.

Table 5 summarizes the results of the feature selection algorithms after training with the HAM10000 and ISIC2019 datasets. The findings of the HAM10000 dataset illustrate the efficacy of GWO in attaining the highest values for accuracy, precision, and F1-score. MVO and FFA rank as the second and third most effective, respectively. The second dataset shows that GWO is the most stable, followed by MVO, WO, and FFA.

By using GWOs as a feature selector, it achieves impressive performance metrics scores: 99.54% for accuracy, 99.44% for precision, 98.79% for recall, and 99.11% for f1-score for using HAM10000 dataset. In other dataset the performance metrices score: 99.47% for accuracy, 99.37% for precision, 99.13% for recall, and 99.25% for f1-score. Fig. 7 displays the confusion matrix of the proposed model for melanoma detection using the HAM10000 dataset.

We conduct a comprehensive assessment of the efficiency of our model by comparing it to other cutting-edge techniques for melanoma detection using the HAM10000 dataset, as shown in Table 6. The transformer model we have developed demonstrates excellent performance, with a remarkable accuracy rate of 99.43%. Additionally, it exhibits exceptional levels of accuracy (99.54%) and f1-score (99.11%). The AlexGWO based model has shown an exceptional recall rate (100%) and precision (99.47%) with notable accuracy and f1-score of 99% and 98.63%, respectively. The YOLO-v3-DCNN based model has good performance in terms of recall.

Table 4
Parameter values of each approach.

Algorithm	Parameter Values
PSO	w (Inertia weight) = 0.7, c1 (Cognitive coefficient) =1.5, c2 (Social coefficient) = 1.5
WO	a = Linearly decreases from 2 to 0, A = Random value in [-a, a], C = Random value in [0, 2]
BAT	A (Loudness) = 0.5, r (Pulse rate) = 0.5, Qmin = 0, Qmax = 2
MVO	WEP (Wormhole Existence Probability) = Linearly increases from 0.2 to 1, TDR (Travel Distance Rate) = Linearly decreases from 1 to 0
FFA	α (Randomization parameter) = 0.25, $\beta 0$ (Attractiveness at $r = 0$) = 1, γ (Absorption coefficient) = 1
GWO	a = Linearly decreases from 2 to 0

Table 5
All experimental runs of each algorithm.

Algorithm	HAM10000				ISIC2019			
	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)
PSO	98.33	98.10	97.90	98.00	98.67	98.45	98.25	98.35
WO	98.86	98.71	98.51	98.61	98.92	98.74	98.54	98.64
BAT	97.85	97.66	97.18	97.42	98.54	98.32	98.15	98.23
MVO	99.21	99.01	**98.84**	98.92	99.31	99.21	99.03	99.12
FFA	99.11	98.79	98.56	98.67	98.85	98.72	98.53	98.62
GWO	**99.54**	**99.44**	98.79	**99.11**	**99.47**	**99.37**	**99.13**	**99.25**

Table 6
Comparison with other state-of-art methods.

Ref	Feature Extraction	Classification	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)
[16]	GLCM + CLCM + RCT	YOLO-v3-DCNN	95	94	98.6	–
[15]	GLCM	AlexNet + GWO	99	**99.47**	**100**	98.63
[34]	GLCM + Color + Shape	SVM	81.35	–	–	–
Our	GLCM + LBP + GWO	Transformer	**99.54**	99.44	98.79	**99.11**

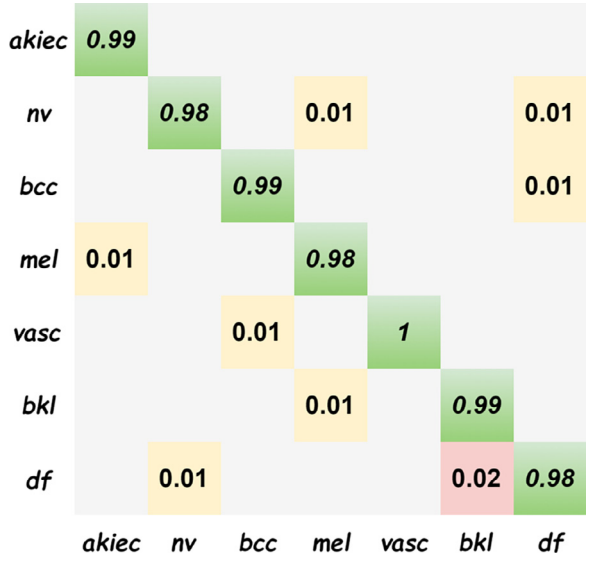

Fig. 7. Confusion Matrix.

Conclusion

This study presents a robust and effective method for detecting melanoma from skin lesion images. It highlights the significance of combining features and utilizing a cutting-edge transformer-based algorithm. The approach encompasses preprocessing, texture feature extraction (combining GLCM and LBP), feature selection using the Gray Wolf Optimization (GWO) algorithm, and a transformer-based decoder for classification. The experimental results on the HAM10000 dataset demonstrate the effectiveness of the proposed model, achieving an impressive accuracy of 99.54% and an exceptional f1-score of 99.11%. The findings underscore the promise of our method for use in dermatology, offering a valuable tool for detecting melanoma. In the future, further investigation of additional hand-crafted and deep learning-based features, such as wavelet transformations or pre-trained representations, might improve the model's performance and usefulness. Furthermore, integrating the system with clinical platforms and developing real-time apps may broaden its practical relevance, making it more accessible for regular dermatological diagnostics. Longitudinal studies that look at how well the model works in clinical settings over long periods of time would also give us useful information about its long-term benefits and usefulness. Similarly, making the transformer-based model easier to understand and interpret could help clinicians use it. However, shortcomings include dependence on hand-crafted characteristics, computational complexity, and the model's opaque nature, which may impede clinical acceptability in the absence of a clear decision-making procedure. Addressing these limits, as well as verifying the model in a variety of datasets and clinical contexts, are critical steps towards increasing its melanoma diagnostic reliability.

Ethics statements

Publicly available dataset, thus no permission required.

CRediT author statement

Hemant Kumar: Conceptualization, Investigation, Methodology, Resources, Writing - original draft, Writing - review and editing. **Abhishek Dwivedi:** Formal analysis, Investigation, Visualization. **Abhishek Kumar Mishra:** Data curation, Formal analysis, Visualization. **Arvind Kumar Shukla:** Resources, Investigation, Validation, Writing - review and editing. **Brajesh Kumar Sharma:** Formal analysis, Methodology, Supervision. **Rashi Agarwal:** Supervision, Conceptualization, Validation, Investigation. **Sunil Kumar:** Visualization, Writing - original draft, Writing - review and editing, Resources.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

I have shared a link of data which is used in my research

Acknowledgment

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

[1] R.L. Siegel, K.D. Miller, N.S. Wagle, A. Jemal, Cancer statistics, 2023, CA. Cancer J. Clin. 73 (1) (2023) 17–48, doi:10.3322/caac.21763.
[2] A. Esteva, et al., Dermatologist-level classification of skin cancer with deep neural networks, Nature 542 (7639) (2017) 115–118, doi:10.1038/nature21056.
[3] I. Haq, et al., YOLO and residual network for colorectal cancer cell detection and counting, Heliyon 10 (2) (2024) e24403, doi:10.1016/j.heliyon.2024.e24403.
[4] A. Dosovitskiy et al., "An image is worth 16x16 words: transformers for image recognition at scale," 2020, [Online]. Available: http://arxiv.org/abs/2010.11929
[5] A. Vaswani, et al., Attention is all you need, in: NIPS'17: *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Curran Associates Inc., Red Hook, NY, USA, 2017, pp. 6000–6010.
[6] R.M. Haralick, K. Shanmugam, I. Dinstein, Textural features for image classification, IEEE Trans. Syst. Man. Cybern. SMC-3 (6) (1973) 610–621, doi:10.1109/TSMC.1973.4309314.
[7] S. Mirjalili, S.M. Mirjalili, A. Lewis, Grey Wolf Optimizer, Adv. Eng. Softw. 69 (2014) 46–61, doi:10.1016/j.advengsoft.2013.12.007.
[8] M. Abdar, et al., Uncertainty quantification in skin cancer classification using three-way decision-based Bayesian deep learning, Comput. Biol. Med. 135 (2021) 104418, doi:10.1016/j.compbiomed.2021.104418.
[9] A.S. Qureshi, T. Roos, Transfer learning with ensembles of deep neural networks for skin cancer detection in imbalanced data sets, Neural Process. Lett. 55 (4) (2023) 4461–4479, doi:10.1007/s11063-022-11049-4.
[10] M.A. Araaf, K. Nugroho, D.R.I.M. Setiadi, Comprehensive analysis and classification of skin diseases based on image texture features using K-Nearest neighbors algorithm, J. Comput. Theor. Appl. 1 (1) (2023) 31–40, doi:10.33633/jcta.v1i1.9185.
[11] A.K. Tiwari, M.K. Mishra, A.R. Panda, B. Panda, HOSMI-LBP-based feature extraction for Melanoma detection using hybrid deep learning models, J. Mech. Med. Biol. 21 (3) (2021), doi:10.1142/S0219519421500299.
[12] H. Ding, Q. Huang, A. Alkhayyat, A computer aided system for skin cancer detection based on developed version of the archimedes optimization algorithm, Biomed. Signal Process. Control 90 (2024), doi:10.1016/j.bspc.2023.105870.
[13] N.M. Mahmoud, A.M. Soliman, Early automated detection system for skin cancer diagnosis using artificial intelligent techniques, Sci. Rep. 14 (1) (2024) 9749, doi:10.1038/s41598-024-59783-0.
[14] V.D. Midasala, B. Prabhakar, J.K. Chaitanya, K. Sirnivas, D. Eshwar, P.M. Kumar, MFEUsLNet: skin cancer detection and classification using integrated AI with multilevel feature extraction-based unsupervised learning, Eng. Sci. Technol. Int. J. 51 (2024), doi:10.1016/j.jestch.2024.101632.
[15] A. Magdy, H. Hussein, R.F. Abdel-Kader, K.A. El Salam, Performance enhancement of skin cancer classification using computer vision, IEEE Access 11 (2023) 72120–72133, doi:10.1109/ACCESS.2023.3294974.
[16] S.O. Manoj, K.R. Abirami, A. Victor, M. Arya, Automatic detection and categorization of skin lesions for early diagnosis of skin cancer using YOLO-v3 - DCNN architecture, Image Anal. Stereol. 42 (2) (2023) 101–117, doi:10.5566/ias.2773.
[17] C.-C. Chang, Y.-Z. Li, H.-C. Wu, M.-H. Tseng, Melanoma detection using XGB classifier combined with feature extraction and K-means SMOTE techniques, Diagnostics 12 (7) (2022), doi:10.3390/diagnostics12071747.
[18] Z. Abbas, M. Rehman, S. Najam, S.M. Danish Rizvi, An efficient Gray-Level Co-Occurrence Matrix (GLCM) based approach towards classification of skin lesion, in: 2019 Amity International Conference on Artificial Intelligence (AICAI), 2019, pp. 317–320, doi:10.1109/AICAI.2019.8701374.
[19] A. Rehman, M.A. Khan, Z. Mehmood, T. Saba, M. Sardaraz, M. Rashid, Microscopic melanoma detection and classification: a framework of pixel-based fusion and multilevel features reduction, Microsc. Res. Tech. 83 (4) (2020) 410–423, doi:10.1002/jemt.23429.
[20] A. Eleyan, H. Demirel, Co-occurrence matrix and its statistical features as a new approach for face recognition, Turkish J. Electr. Eng. Comput. Sci. 19 (1) (2011) 97–107, doi:10.3906/elk-0906-27.
[21] A. Hadid, The local binary pattern approach and its applications to face analysis, in: *2008 First Workshops on Image P*rocessing *Theory, Tools* and *Applications*, 2008, pp. 1–9, doi:10.1109/IPTA.2008.4743795.
[22] S. Kumar, H. Kumar, Classification of COVID-19 X-ray images using transfer learning with visual geometrical groups and novel sequential convolutional neural networks, MethodsX 11 (2023) 102295, doi:10.1016/j.mex.2023.102295.
[23] X. Qian, X.-S. Hua, P. Chen, L. Ke, PLBP: an effective local binary patterns texture descriptor with pyramid representation, Pattern Recognit. 44 (10) (2011) 2502–2515, doi:10.1016/j.patcog.2011.03.029.
[24] C. Shen, K. Zhang, Two-stage improved Grey Wolf Optimization algorithm for feature selection on high-dimensional classification, Complex Intell. Syst. 8 (4) (2022) 2769–2789, doi:10.1007/s40747-021-00452-4.
[25] S. Mirjalili, "Design and Optimization of Optical Devices Using Artificial Intelligence Techniques," no. June, p. 122, 2021, [Online]. Available: https://spectrum.library.concordia.ca/id/eprint/988867/
[26] Z. Wang, H. Lu, H. Yan, H. Kan, L. Jin, Vison transformer adapter-based hyperbolic embeddings for multi-lesion segmentation in diabetic retinopathy, Sci. Rep. 13 (1) (2023) 11178, doi:10.1038/s41598-023-38320-5.
[27] N.C.F. Codella, 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 (ISBI2018), 2018, pp. 168–172, doi:10.1109/ISBI.2018.8363547.
[28] P. Tschandl, C. Rosendahl, H. Kittler, The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions, Sci. Data 5 (1) (2018) 180161, doi:10.1038/sdata.2018.161.
[29] A.G. Gad, Particle swarm optimization algorithm and its applications: a systematic review, Arch. Comput. Methods Eng. 29 (5) (2022) 2531–2561, doi:10.1007/s11831-021-09694-4.
[30] S. Mirjalili, A. Lewis, The whale optimization algorithm, Adv. Eng. Softw. 95 (2016) 51–67, doi:10.1016/j.advengsoft.2016.01.008.
[31] M. Shehab, et al., A comprehensive review of bat inspired algorithm: variants, applications, and hybridization, Arch. Comput. Methods Eng. 30 (2) (2023) 765–797, doi:10.1007/s11831-022-09817-5.
[32] S. Mirjalili, S.M. Mirjalili, A. Hatamlou, Multi-verse optimizer: a nature-inspired algorithm for global optimization, Neural Comput. Appl. 27 (2) (2016) 495–513, doi:10.1007/s00521-015-1870-7.
[33] I. Fister, I. Fister, X.-S. Yang, J. Brest, A comprehensive review of firefly algorithms, Swarm Evol. Comput. 13 (2013) 34–46, doi:10.1016/j.swevo.2013.06.001.
[34] Z. Li, Q. Ji, X. Yang, Y. Zhou, S. Zhi, An identification method of feature interpretation for melanoma using machine learning, Appl. Sci. 13 (18) (2023), doi:10.3390/app131810076.