EC-EBMs: Skin Lesion Classification with EBMs

Quyen Van Vo [1,2], An Cong Tran [1,3], and Hiep Xuan Huynh [1,3,*]

[1] College of Information and Communication Technology, Can Tho University, Can Tho, Vietnam
[2] Department of Science and Technology, Can Tho University of Medicine and Pharmacy, Can Tho, Vietnam
[3] CTU Leading Research Team on Automation, Artificial Intelligence, inforMation tEchnology and Digital Transformation (CTU-AIMED), Can Tho University, Can Tho, Vietnam
Email: vovanquyen@ctump.edu.vn (Q.V.V.); tcan@ctu.edu.vn (A.C.T.); hxhiep@ctu.edu.vn (H.X.H.)
*Corresponding author

***Abstract*—Skin lesion classification presents significant challenges in the medical field, primarily due to the increasing complexity and diversity of datasets. Traditional methods often struggle with issues like class imbalance and unseen distribution of disease types, which are common in medical data. In this paper, we introduce a novel approach: the integration of the Energy Correlation (EC) method into Energy-Based Models (EBMs), hereafter abbreviated as EC-EBMs, which combines EBMs with EC to enhance the classification of skin lesion images. This method effectively addresses the challenges of class imbalance and data heterogeneity by optimizing the relationship between features and labels, minimizing intraclass energy, and maximizing interclass differences. Experimental results from two datasets demonstrate the effectiveness of EC-EBMs. On the ISIC 2020 dataset, our approach achieved an accuracy of 98.24% in distinguishing between malignant melanoma and benign non-melanoma lesions. Additionally, it showed promising results on the ISIC 2019 dataset, successfully classifying eight diagnostic labels with an accuracy of 72.33%, sensitivity of 55.09%, and specificity of 95.18%. These findings highlight the potential of EC-EBMs to improve classification performance and provide a new solution to complex challenges in medical classification.**

***Keywords*—Energy-Based Models (EBMs), Energy Correlation (EC), skin lesion classification, ISIC datasets, nonlinear relationships, loss function optimization**

I. Introduction

Skin cancer is a common and potentially life-threatening condition characterized by the uncontrolled growth of abnormal skin cells [1]. Timely detection and accurate classification of skin cancer are essential for improving treatment effectiveness and reducing both the incidence and mortality rates associated with these prevalent skin cancers [2]. Traditional methods of diagnosing skin cancer, based on dermatologists' clinical observations, can result in diagnostic errors. In today's context, the use of advanced technologies, particularly the development of deep learning and AI-based classification models, is becoming increasingly widespread [3]. Many previous studies have focused on the application of Deep Learning and Machine Learning methods to improve the classification of skin lesion images [4–9]. Convolutional Neural Networks (CNNs) have been widely employed in medical image classification and have achieved encouraging results [10–13]. A notable study utilizing CNNs on the ISIC dataset reached an accuracy of up to 86.6% in skin lesion classification [14]. However, traditional CNN-based methods require large amounts of data and often struggle when dealing with heterogeneous datasets, such as the imbalance between different types of lesions in medical data [4]. This leads to difficulties in achieving high accuracy on datasets with a high degree of variability. Moreover, these methods usually focus on optimizing the neural network architecture without thoroughly considering the correlation between features and classification labels, which limits the effectiveness of the classification process.

In this paper, we propose a new approach to integrate the Energy Correlation method into Energy-Based Models (EC-EBMs) to optimize the classification of skin lesions. While previous works have explored deep learning for skin lesion classification, they primarily rely on supervised learning with standard loss functions. In contrast, our method introduces an energy-based learning paradigm that enhances class separation by minimizing the energy correlation between different lesion types. To the best of our knowledge, this is the first application of EBMs with Energy Correlation Loss in dermatological image classification. Energy Correlation improves the correlation between features and class labels, allowing the model to more accurately differentiate among various classes, even when the data is imbalanced. The experiments were conducted on the ISIC 2019 and ISIC 2020 datasets [15], aiming to enhance the model's accuracy in classifying complex skin lesions. Despite these advantages, the proposed model also faces limitations regarding computation time due to the precise tuning of the loss function and model structure needed to achieve optimal performance. The main contributions of this paper include the introduction of Energy Correlation Loss (L_{EC}) for improved feature-label correlation and the use of EC-EBMs to handle imbalanced medical datasets.

The rest of this paper is organized as follows: Section II presents the skin lesion image representation. Section III describes Loss Function according to Energy Correlation.

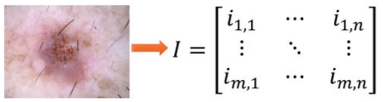
Manuscript received January 24, 2025; revised February 21, 2025; accepted July 14, 2025; published November 7, 2025.

Section IV introduces the proposed EC-EBMs tool for skin lesion classification. Section V details our experimental setup and results. Finally, Section VI concludes the paper and suggests future research directions.

II. Skin Lesion Image Representation

A. Representation of Skin Lesion Images as a Pixel Matrix

From a mathematical and machine learning perspective, skin lesion images can be represented in formats that enable computational models to process and analyze them. One common representation is the weight matrix, where each value corresponds to the intensity of a pixel at a specific location. Suppose a skin lesion image has dimensions $m \times n$, where m represents the number of rows (horizontal) and n represents the number of columns (vertical). In this case, the image can be represented as a matrix $I \in \mathrm{R}^{m \times n}$, as shown in Fig. 1.



$$I = \begin{bmatrix} i_{1,1} & \cdots & i_{1,n} \\ \vdots & \ddots & \vdots \\ i_{m,1} & \cdots & i_{m,n} \end{bmatrix}$$

Fig. 1. Skin lesion image [15]: Weight matrix I with $i_{m,n}$ being the pixel value at coordinates *(m, n)*.

B. Representation of Skin Lesion Images as Feature Vectors

To represent a skin lesion image as a feature vector, the first step is to extract key characteristics from the image and encode them into a mathematical vector. Let I be the skin lesion image, and let F represent the feature vector, where each element f_i in F corresponds to a specific feature extracted from the image.

$$F = \{f_1, f_2, \ldots f_k\} \tag{1}$$

where F is a feature vector with k elements, each element f_i represents a feature extracted from the skin lesion image.

C. The Key Characteristics of Skin Lesion Images

1) Morphology

- Perimeter (p): The perimeter refers to the total length of the lesion's boundary. Malignant lesions typically exhibit larger perimeters due to their irregular and uneven borders, in contrast to the smoother edges commonly observed in benign lesions.

$$p = \int_i ds \tag{2}$$

where i is the boundary of the lesion, ds is the infinitesimal arc length along the boundary.

- Area (a): The area represents the total number of pixels within the region of the lesion. Larger lesions are often associated with a higher risk of malignancy.

$$a = \sum_{(x,y) \in Region} 1 \tag{3}$$

- Circularity (c): Circularity ranges from 0 to 1, where a value closer to 1 indicates a shape more similar to a perfect circle. Malignant lesions tend to have lower circularity due to their irregular shapes.

$$c = \frac{4\pi a}{P^2} \tag{4}$$

- Aspect Ratio (ar): This metric describes the overall shape of the lesion. Higher values indicate elongated lesions, which are often associated with atypical characteristics.

$$ar = \frac{\text{Major Axis Length}}{\text{Minor Axis Length}} \tag{5}$$

- Maximum Diameter (d): a diameter greater than 6 mm is considered a warning sign for malignant lesions.

$$d = max\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \tag{6}$$

where (x_i,y_i) and (x_j,y_j) are any two points on the lesion boundary

2) Color features

The Hue-Saturation-Value (HSV) color space offers several advantages over the Red-Green-Blue (RGB) color space for analyzing skin lesion images. The HSV system is calculated from normalized RGB values (R,G,B$\in$[0,1]) using the following formulas:

Hue (H) represents the precise color of the skin lesion.

$$H = \begin{cases} 60° \times \dfrac{G - B}{V - \min(R,G,B)} & \text{If } V = R \\ 60° \times \left(2 + \dfrac{B - R}{V - \min(R,G,B)}\right) & \text{If } V = G \\ 60° \times \left(2 + \dfrac{R - G}{V - \min(R,G,B)}\right) & \text{If } V = B \end{cases} \tag{7}$$

where:

Significance: Identifies the primary color of a pixel (e.g., red, yellow, green, blue).

Applications: Detects characteristic colors such as black (indicative of deep pigmentation), blue (subdermal pigmentation), or red (vascular structures) and facilitates lesion segmentation based on color thresholds.

Saturation (S) represents the intensity of color saturation.

$$S = \begin{cases} 0 & \text{If } V = 0 \\ \dfrac{V - \min(R,G,B)}{V} & \text{If } V \neq 0 \end{cases} \tag{8}$$

where:

Significance: Measures the purity of a color, where S = 0 indicates white, gray, or black color.

Applications: Identifies areas with faded or absent coloration, often indicative of malignant cancer, and differentiates pure pigmented regions from faint or inflamed areas.

Value (V) represents the brightness of the lesion area.

$$V = \max(R,G,B) \tag{9}$$

where:

Significance: Indicates the overall brightness of a color.

Applications: Analyzes light and shadow to distinguish lesions from the background and eliminates noise caused by uneven lighting conditions in the image.

3) ABCDE features

- Asymmetry (*A*): Benign skin lesions typically exhibit symmetrical shapes, whereas malignant lesions are often asymmetrical.

$$A = \frac{|A_L - A_R|}{A_L + A_L} \tag{10}$$

where:

A_L, A_R*:* The area of the two halves of the lesion when divided along the vertical or horizontal axis.

- Abnormal Indicator: If a lesion cannot be divided into two symmetrical halves, it is indicative of malignancy.
- Border Irregularity (*B*): Irregular or jagged borders are indicative of skin cancer. Benign lesions typically have smooth, round, or oval borders.

$$B = \frac{P}{2\sqrt{\pi A}} \tag{11}$$

where

P: Perimeter of the skin lesion

A: Area of the skin lesion.

- Abnormal Indicator: Higher values indicate irregular lesion borders, a common characteristic of melanoma.
- Color Variation (*C*): Benign lesions typically have a uniform color (light brown or black). Malignant lesions may exhibit a variety of colors, including Black, Dark brown, Red, Blue, and White (indicative of necrosis or depigmentation).

$$C = \frac{1}{N}\sum_{i=1}^{N}(I_c(x_i, y_i) - \mu_c)^2 \tag{12}$$

where:

$I_c(x_i, y_i)$*:* color value at point (x_i, y_i)

μ_c *:* Average color value across the lesion

- Abnormal Indicator: The presence of multiple colors or uneven distribution is a sign of malignant lesions.
- Diameter (*D*): Malignant lesions are typically larger in size compared to benign lesions. The general rule is that if the diameter D > 6 mm, further examination is required.

$$D = \max\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \tag{13}$$

where:

(x_i, y_i) and (x_j, y_j): any two points on the lesion border.

Abnormal Indicator: a diameter greater than 6 mm is often considered suspicious for cancer.

- Evolving (E): Malignant lesions often change over time in size, color, or shape, making this the most critical factor in identifying malignancy.
  - Area Change:

$$\Delta A = A_{t+1} - A_t \tag{14}$$

  - Perimeter Change:

$$\Delta P = P_{t+1} - P_t \tag{15}$$

  - Color Evolution:

$$\Delta C = \|C_{t+1} - C_t\| \tag{16}$$

where:

A_t, P_t, C_t*:* area, perimeter, and color at time *t*

A_{t+1}, P_{t+1}, C_{t+1}*:* values at time *t*+1.

- Abnormal Indicator: If a lesion exhibits rapid changes in any characteristic (size, border, or color), immediate medical evaluation is necessary.

III. Loss Function According to Energy Correlation

A. Energy Correlation (EC)

Energy Correlation is a concept used in statistical analysis and machine learning, particularly in classification or regression tasks. It is used to measur the correlation between variables based on the energy distance, rather than linear correlation. This approach can capture more complex relationships than traditional methods like Pearson correlation.

Suppose there are two data distributions from skin lesion images, *P* and *Q*, with feature vectors sampled from two different probability distributions. We can calculate the Energy Correlation between these two distributions using the formula:

$$EC(P, Q) = \frac{ED(P, Q)}{\sqrt{ED(P, P) \cdot ED(Q, Q)}} \tag{17}$$

where Energy Distance (ED) [16] is a distance measure based on the energy function between probability distributions [17]:

$$\begin{aligned} ED(P, Q) = 2 \cdot Ex_p \sim P, x_q \sim Q[d(x_p, x_q)] - Ex_p, x'_p \\ \sim P[d(x_p, x'_p)] - Ex_q, x'_q \\ \sim Q[d(x_q, x'_q)] \end{aligned} \tag{18}$$

$$d(x_p, x_q) = \| x_p - x_q \|_2 = \sqrt{\sum_{i=1}^{d}(x_p^{(i)} - x_q^{(i)})^2} \tag{19}$$

ED(P,Q) is the energy distance between *P* and *Q*; *E* denotes the mathematical expectation; and $d(x_p, x_q)$ is a distance function between two data points x_p, x_q.

When *EC(P,Q)* is equal to 1, the two variables are perfectly correlated; when it is close to 0, the two variables are independent of each other.

Loss function in EBMs can be customized to include Energy Correlation to optimize the correlation between features and classes.

B. Energy Correlation Loss (L_{EC})

EBMs learns to minimize the energy level for the correct class while maintaining a distinction between unrelated classes, which helps improve classification accuracy:

$$L_{EC} = \sum_{i=1}^{N} (E(x_i, y_i) - \beta \cdot EC(x_i, y_i)) \quad (20)$$

where $E(x_i, y_i)$ is the predicted energy of layer y_i with feature x_i; $EC(x_i, y_i)$ is the energy correlation function between x_i and y_i; β is the weight adjustment factor of Energy Correlation during the optimization process.

IV. Proposed EC-EBMs Tool for Skin Lesion Classification

A. Energy-Based Model (EBMs) in Skin Lesion Classification

Energy-Based Model (EBMs) is a machine learning approach that does not rely on conditional probabilities but instead models the data distribution through an energy function [18]. Rather than computing probabilities directly, EBMs defines an energy measure for each input state and uses it to make predictions.

$$p(y|x;\theta) = \frac{e^{-E(x,y;\theta)}}{Z(x;\theta)} \quad (21)$$

where $E(x, y; \theta)$ is the energy function, representing the compatibility between input x (feature vector) and output y (classification label) [19], computed as:

$$E(X, y; \theta) = -logP(y|X;\theta) \quad (22)$$

$Z(x; \theta)$ is the partition function, computed as:

$$Z(x, \theta) = \sum_{y\prime} e^{-E(x,y\prime;\theta)} \quad (23)$$

x is the input feature vector, and y is the output label. θ represents the model parameters, which are optimized during training. The parameters θ of the model are updated using the Gradient Descent method [20, 21] to minimize the loss function L_{EC}:

$$\theta \leftarrow \theta - \mu \nabla_\theta L_{EC}(\theta) \quad (24)$$

η is learning rate. $\nabla_\theta L_{EC}(\theta)$ is the gradient of the loss function with respect to the parameters θ.

The goal of the model is to minimize energy $E(x, y; \theta)$ for correct data pairs while maximizing the energy for incorrect ones [22].

This process continues until the model converges to the optimal value of the parameters, meaning when the energy difference between samples of the same class and different classes is clearly distinguished.

Classification calculation: After the model is trained, classification for a new skin lesion image X^* is performed by selecting the class y such that the energy function $E(X^*, y; \theta^*)$ is minimized:

$$y^* = \arg\min E(X^*, y; \theta^*) \quad (25)$$

The expected outcome is an efficient classification system that utilizes the Loss Function L_{EC} with Energy Correlation to achieve high accuracy in classifying different types of skin lesions.

B. Overview Diagram

The overview diagram of the system shown is Fig. 2 illustrates the entire process from data input to result evaluation. This diagram is divided into four main functional blocks: Image Acquisition, Image Preprocessing and Feature Extraction, Energy Modeling, and Skin Lesion Classification. The steps in the diagram are sequentially connected, ensuring that the input data is processed, modeled, and classified accurately and efficiently.

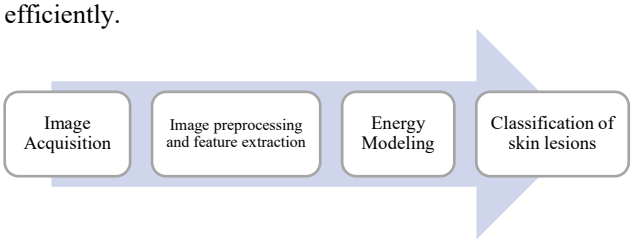


Fig. 2. EC-EBMs overview diagram.

C. Main Functions and Algorithms

Feature extraction from skin lesion images is a critical step that directly impacts the performance of the proposed tool. The extracted features provide essential information about the lesion's attributes, with the goal of transforming raw skin lesion image data into structured formats. This process focuses on identifying significant features while eliminating irrelevant information to serve as inputs for the model, ultimately enhancing the accuracy of the proposed model's outcomes.

Algorithms 1. Feature Extraction

```
Begin
CLASS Features_Extract:
  def Extract_Morphology_Features(image):
    return [aspect_ratio, area, perimeter]
  def Extract_Color_Features(image):
    return [mean_hue, mean_saturation, mean_value]
  def Extract_ABCDE_Features(image):
    return [A, B, C, D, E]
  features_list = []
  combined_features = Extract_Morphology_Features +
  Extract_Color_Features + Extract_ABCDE_Features
  features_list.append(combined_features)
  return feature_list
End
```

Using Energy Correlation Loss to optimize the EBM model in classifying skin lesion images. Energy Correlation enhances the distinction between classes and optimizes the correlation between features and classification labels.

Algorithms 2. Definition of the Energy Correlation Loss function (L_{EC})

```
Begin
  CLASS EnergyCorrelationLoss:
    def initialize class_weights
    def forward:
      energy_loss = (Energy compatibility value
    between predicted label and actual label)
  return energy_loss
End
```

Algorithms 3. Skin lesion classification with EC-EBMs

Input: The Feature Vector was extracted from Features_Extract Class
Output: - Energy Score: EBMs calculates the energy for each input image).
- Label Output: The result is a predicted label.

```
Begin
# Train_EBM
  FUNCTION Train EBM with epoch = 100:
    Create empty list # to save loss history
    FOR epoch FROM 1 TO epochs DO:
      FOR each batch (images, labels) IN dataloader
      DO:
        loss = Cal EnergyCorrelationLoss
        Execute optimizer(), backward() # to calculate gradient of loss
        total_loss += value of loss
        predicted = Get predicted label with highest energy value from output
        correct += Total number of correct predicted labels compared to actual labels
        total += Number of labels in current batch
      average_loss = total_loss / Number of batches
      accuracy = correct / total * 100
    add average_loss to loss_history
  Print "Epoch [epoch/epochs], Loss: average_loss, Accuracy: accuracy%"
End
```

Starting with a dataset that consists of feature vectors extracted from skin lesion images of the ISIC 2019 and ISIC 2020 datasets, along with the corresponding labels for various types of skin lesions. The model computes the Energy Correlation Loss between the feature vectors, defined as the ratio of interclass energy to intraclass energy, and adjusts the parameters to minimize the Energy Correlation Loss. The proposed model consists of two fully connected layers arranged in the architecture of 3072→100→8, totaling 308,108 parameters. In this configuration, fc1 has 307,300 parameters, while fc2 contains 808 parameters. The average training time per epoch on the GPU is about 20 to 30 s, resulting in a total training time of approximately 50 to 60 min on the GPU for 100 epochs. In the result, the model computes the energy values for each classification label, and the label with the lowest energy value is selected as the predicted label, as it has the highest probability.

V. Experiment and Results

A. Data Used

The ISIC 2019 and ISIC 2020 skin image datasets [15] are used in model experiments. The ISIC 2020 dataset consists of nearly 33,000 dermatological images captured using devices like dermoscopes, provided by the International Skin Imaging Collaboration (ISIC). The dataset focuses on classifying melanoma, a dangerous type of skin cancer, with the goal of distinguishing between melanoma (malignant) and non-melanoma (benign) cases. The number of melanoma labels accounts for only about 1.76%, while non-melanoma constitutes the majority, resulting in a significant class imbalance, as shown in Fig. 3. The primary challenge in classification is addressing this imbalance, requiring algorithms to effectively learn the rare melanoma features without being overwhelmed by the abundance of non-melanoma cases.

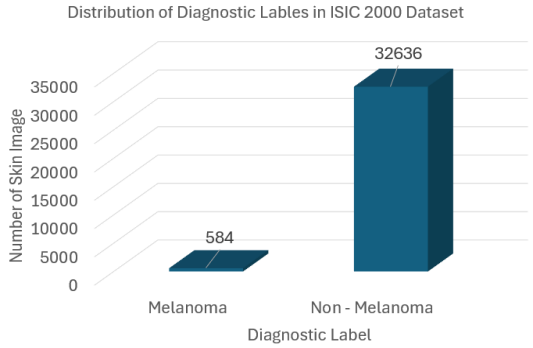


Fig. 3. Image distribution between melanoma and non-melanoma in ISIC 2020.

The ISIC 2019 dataset comprises over 25,000 high-quality dermatological images, focusing on skin lesion classification. It includes eight types of lesions: Melanoma, Nevus, Basal cell carcinoma, Actinic keratosis, Benign keratosis, Dermatofibroma, Vascular lesion, and Squamous cell carcinoma. A summary of this dataset is presented in Table I. The label distribution within this dataset is significantly imbalanced, with categories like Nevus and Benign Keratosis being predominant, while Dermatofibroma and Vascular lesions have very few samples, resulting in a notable class imbalance. This label distribution is shown in Fig. 4. The main challenge in classification is addressing this imbalance, requiring models to effectively learn the features of rare classes without being overwhelmed by the more common ones. Overall, this dataset serves as an essential foundation for developing deep learning algorithms in dermatology.

TABLE I. Summary of the ISIC_2019 Dataset [15]

Diagnostic Label	Abbreviation	Number of Skin Image
Melanoma	MEL	4,522
Melanocytic Nevus	NV	12,722
Basal Cell Carcinoma	BCC	3,452
Actinic Keratosis / Bowen's Disease	AKIEC	867
Benign Keratosis	BKL	2,694
Dermatofibroma	DF	239
Vascular Lesion	VASC	253
Squamous Cell Carcinoma	SCC	628
Unknown	UNK	954

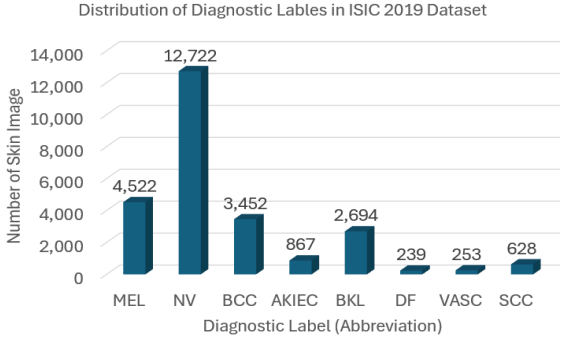


Fig. 4. Image distribution across 8 labels in ISIC 2019.

B. Tools Used

EC-EBMs is a tool developed in Python, using PyTorch, TensorFlow/Keras, and Energy libraries. This tool allows for efficient model training, experimentation, and deployment of the EBM model, encompassing processes such as data preprocessing, feature extraction, energy correlation calculations, model building, skin lesion classification, and result evaluation. Experiment Scenario for Skin Lesion Classification using EBMs with Energy Correlation

1) Scenario 1: EC-EBMs in the classification of melanoma (malignant) and non-melanoma (benign) on ISIC 2020 dataset

The results of skin lesion classification using the EBMs approach with Energy Correlation were experimented with the EC-EBMs tool on the ISIC 2020 dataset.

The skin lesion images in the ISIC 2019 and ISIC 2020 datasets are converted into structured feature vectors through the process of feature extraction. Besides general image characteristics such as texture, shape, and intensity, special attention is paid to domain-specific features that are particularly relevant to dermatological analysis. These include morphological attributes (e.g., shape, border irregularity), color variations, and clinical criteria based on the widely recognized ABCDE rule (Asymmetry, Border, Color, Diameter, and Evolution). These extracted features play a crucial role in enhancing the model's ability to accurately classify and evaluate skin lesions.

During the training of the EBM model on the ISIC_2020 dataset, the value of the loss function (L_{EC}) gradually decreased across epochs, indicating that the model was converging and learning the distinguishing features between different types of skin lesions. In the initial epoch, the loss dropped sharply, showing that the model was learning quickly. After around the 40th epoch, the loss began to decrease more slowly and stabilized after approximately 100 epochs, as shown in Fig. 5.

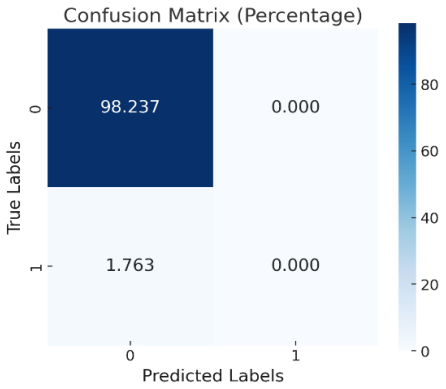


Fig. 5. Confusion matrix on ISIC 2020 dataset.

Classification results: Experiments conducted on the ISIC 2020 dataset demonstrated the model's capability in detecting and classifying skin lesions, particularly malignant ones, achieving an accuracy of 98.237%.

Based on the results, it demonstrates that the proposed EC-EBMs tool performs well on the imbalanced dataset of melanoma and non-melanoma in the ISIC 2020 dataset.

2) Scenario 2: EC-EBMs in the classification of 8 skin lesion types on ISIC 2019 dataset

The results of skin lesion classification using the EBMs approach with Energy Correlation were experimented with the EC-EBMs tool on the ISIC 2019 dataset.

Fig. 6 illustrates the training results along with the L_{EC} chart across the epochs.

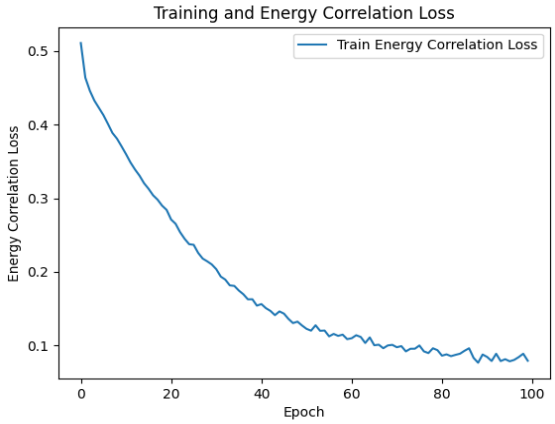


Fig. 6. Training EBMs with Energy Correlation Loss (L_{EC}) on ISIC 2019.

From the chart in Fig. 6, the training process of the EC-EBMs model using Energy Correlation Loss over 100 epochs is illustrated. At the beginning of training, the loss value is relatively high (above 0.5), indicating that the model has not yet captured the correlation between features and classification labels. However, as training progresses, the loss decreases rapidly during the first 20 epochs, demonstrating that the model effectively learns the most distinguishable patterns in the early stage. From epoch 20 to epoch 60, the loss continues to decline steadily, showing that the model is gradually optimizing its internal representation of more complex feature-label relationships. After epoch 60, the loss converges to a low and stable value around 0.07, with minor fluctuations, indicating that the model has reached convergence and is no longer significantly improving.

This downward trend in Energy Correlation Loss suggests that the model successfully minimizes intraclass energy while maximizing interclass energy differences, thus enhancing the correlation between features and class labels. The smooth convergence also implies that the training process is stable and not prone to overfitting. Overall, the results demonstrate that the EC-EBMs model effectively learns and generalizes well on the training data.

3) Classification results

The results in Table II indicate that the proposed EC-EBMs model achieved outstanding performance on the ISIC 2020 dataset (Scenario 1), with a Precision of 96.51%, Recall of 98.24%, Accuracy of 98.24%, and an F1-Score of 97.36%. These results demonstrate the model's effectiveness in distinguishing between melanoma (malignant) and non-melanoma (benign) cases, as well as its robustness in handling class imbalance. In contrast, for the ISIC 2019 dataset (Scenario 2), the multiclass classification task yielded a Precision of

72.13%, Recall of 72.33%, Accuracy of 72.33%, and an F1-Score of 72.15%. The lower performance in this scenario can be attributed to greater diversity and significant class imbalance in the dataset, particularly for classes with very limited samples. Overall, EC-EBMs exhibits superior performance in binary classification and maintains competitive results in multiclass classification, highlighting its potential applicability in clinical diagnostic support systems.

TABLE II. PRECISION, RECALL, ACCURACY AND F1-SCORE OF THE CLASSIFICATION

Experiment Scenario	Precision	recall	Accuracy	F1-Score
Scenario 1 (ISIC 2020)	0.9651	0.9824	0.9824	0.9736
Scenario 2 (ISIC 2019)	0.7213	0.7233	0.7233	0.7215

The results from Fig. 7 indicate that the ROC curve effectively illustrates the model's performance in classifying skin lesions using the One-vs-Rest (OvR) approach. The AUC values range from 0.95 to 0.99, demonstrating the model's strong ability to distinguish between classes. Class 2 achieves the highest AUC (0.99), indicating high classification confidence for this class. In contrast, Class 5 has the lowest AUC (0.95), which is still a strong result but may require further improvement. Most ROC curves lie well above the diagonal reference line, confirming that the model performs significantly better than random classification.

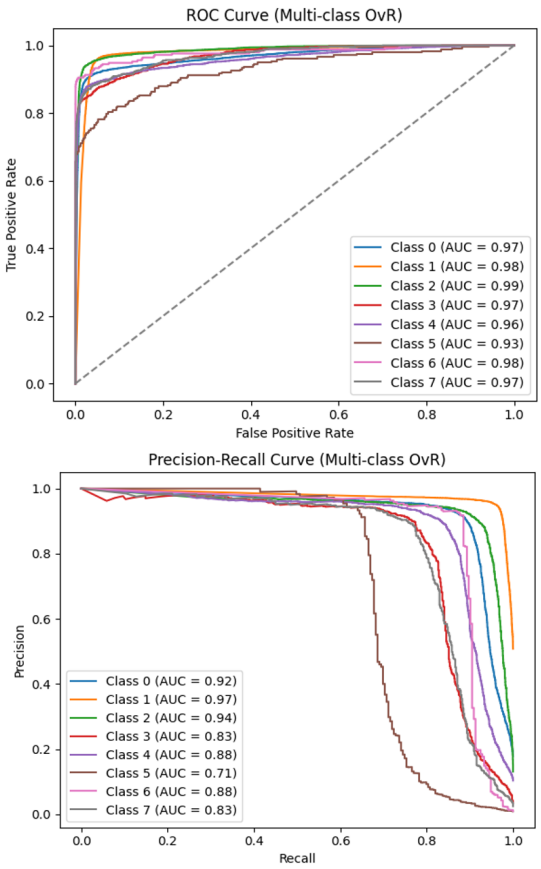


Fig. 7. ROC curve and precision-recall curve.

Additionally, the Precision-Recall (PR) curve evaluates the model's ability to correctly identify positive samples, which is particularly important in cases of imbalanced data. Class 1 (AUC = 0.97) and Class 2 (AUC = 0.93) exhibit the highest precision, indicating that the model accurately detects skin lesions in these categories. However, Class 5 (AUC = 0.71) and Class 7 (AUC = 0.78) have the lowest Precision-Recall values, suggesting that the model struggles to distinguish these classes. Some PR curves drop sharply as recall increases, which may indicate a higher number of false positives for these classes.

Overall, the model demonstrates strong performance, with high classification accuracy for most classes (AUC > 0.95), confirming its reliability in detecting skin lesions.

The result from Fig. 8 indicates that the proposed EC-EBMs tool was able to classify all diagnostic labels with high accuracy. Some labels were classified with particularly good results, achieving high accuracy rates, such as Nevus (85.94%) and Basal Cell Carcinoma (69.32%). This suggests that the model accurately learned the distinguishing features of these types of lesions. However, for certain lesions like Dermatofibroma and Benign Keratosis, the model encountered more difficulty in accurate classification, showing a high misclassification rate.

Melanoma (MEL): The model correctly classified 59.07% of melanoma cases. However, a significant 24.89% of melanoma cases were misclassified as NV (nevus), which indicates confusion between these two classes. This confusion might be due to similar visual features between melanoma and nevus lesions, which can sometimes appear alike in dermoscopic images.

Melanocytic Nevus (NV): The model performed well with 85.94% accuracy in classifying nevi, which is a good result, showing that the model distinguishes this class relatively well. There are minor misclassifications, such as 6.72% of nevus cases being incorrectly classified as MEL.

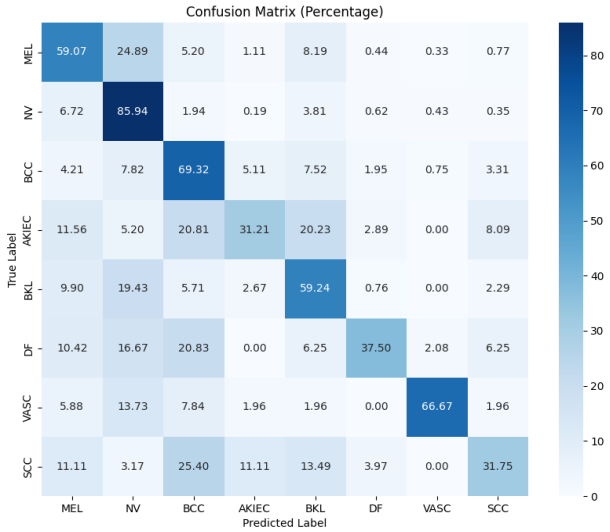


Fig. 8. Confusion matrix on ISIC 2019 dataset.

Basal Cell Carcinoma (BCC): The accuracy for basal cell carcinoma was 69.32%, which shows a moderate performance. Notable confusion is seen with NV (7.82%)

and BKL (7.52%), suggesting that some features overlap between these classes.

Actinic Keratoses and Intraepithelial Carcinoma (AKIEC): This class had a relatively low correct classification rate of 31.21%, with significant misclassifications into BCC at 20.81%, showing that the model struggles with distinguishing these two classes. Misclassifications into other classes, such as BKL and BCC, also occurred.

Benign Keratosis (BKL): The model correctly classified 59.24% of BKL cases, but a large percentage (19.43%) were misclassified as NV. This confusion suggests that the features of benign keratosis might overlap with those of NV, leading to these misclassifications.

Dermatofibroma (DF): The accuracy was low (37.50%), with significant confusion with BCC (20.83%) and NV (16.67%). The overlap in features between DF and these other malignant classes might be the cause of these misclassifications.

Squamous Cell Carcinoma (SCC): For squamous cell carcinoma, the accuracy was low (31.75%), with significant confusion with BCC (25.40%) and BKL (13.49%). The overlap in features between SCC and these other malignant classes might be the cause of these misclassifications.

Vascular Lesions (VASC): The model performed reasonably well with 66.67% correct classification, but there was still confusion with other classes, particularly NV and BCC.

One of the main reasons why certain diagnostic labels, such as MEL, AKIEC, DF, and SCC, have a low percentage is the high similarity in visual characteristics among different skin lesions, making it difficult to distinguish between classes. Many lesions share similar shapes, colors, and textures, which can lead the model to misclassify them, especially when clear visual differences are not apparent. Additionally, the small number of samples for some labels in the dataset further affects classification performance.

VI. Conclusion

The proposed EC-EBMs tool is a novel approach to skin lesion classification using the Energy-Based Model with an Energy Correlation method. By utilizing the L_{EC} loss function, the model not only achieved strong multiclass classification performance but also effectively addressed challenges related to data imbalance, a common issue in the medical field. Experiments on the ISIC 2019 and ISIC 2020 datasets demonstrated that the EC-EBMs tool effectively identifies melanoma (malignant) and non-melanoma (benign) images with an accuracy of up to 98.24% on the ISIC 2020 dataset. And it performs well in classifying 8 diagnostic labels on the ISIC 2019 dataset, achieving an accuracy of 72.33%, a sensitivity of 55.09%, and a specificity of 95.18%. Additionally, the research opens numerous avenues for developing medical diagnostic support systems, ranging from expanding datasets to improving modeling techniques, preprocessing, and feature extraction. Deploying the model in real-world clinical applications to ensure transparency and decision accuracy represents a critical next step in translating this research into practice. In addition to proposing a novel energy-based model leveraging the Energy Correlation approach, we acknowledge the importance of benchmarking its performance against related studies to comprehensively evaluate the effectiveness and competitiveness of the proposed method. Although hyperparameter tuning in this study was conducted manually through empirical observation, we recognize that integrating automated optimization frameworks such as KerasTuner or Optuna could significantly enhance both model performance and training stability. Accordingly, future work will focus on comparative evaluations with state-of-the-art skin lesion classification methods on standardized datasets, alongside the adoption of automated hyperparameter optimization techniques to further improve the model's efficiency and robustness.

In conclusion, the proposed method not only contributes to the classification of skin lesions but also broadens the potential applications of Energy-Based Models in both the medical and machine learning fields.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Quyen Van Vo: Conceived and designed the research framework; contributed to the development of the methodology, supervised the overall progress of the study, and supported the coding of the EC-EBMs model; coordinated project administration and manuscript revision. An Cong Tran: Implemented the experiments; performed feature extraction, data preprocessing, and result analysis; contributed to manuscript drafting. Hiep Xuan Huynh: Led the theoretical formulation of Energy Correlation in Energy-Based Models; validated the experimental results; drafted and critically revised the manuscript; served as the corresponding author. All authors have read and approved the final version.

References

[1] N. H. Khan, M. Mir, L. Qian *et al.*, "Skin cancer biology and barriers to treatment: Recent applications of polymeric micro/nanostructures," *J. Adv. Res.*, vol. 36, pp. 223–247, 2022. http://dx.doi.org/10.1016/j.jare.2021.06.014

[2] M. Ragab, H. Choudhry, M. Al-Rabia, S. Binyamin, A. Aldarmahi, and R. Mansour, "Early and accurate detection of melanoma skin cancer using hybrid level set approach," *Front Physiol*., vol. 13, 965630, 2022. http://dx.doi.org/10.3389/fphys.2022.965630

[3] H. Stafford *et al.*, "Non-melanoma skin cancer detection in the age of advanced technology: A review," *Cancers*, vol. 15, no. 12, 3094, 2023.

[4] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Trans. on Knowledge and Data Engineering*, vol. 21, no. 9, pp. 1263–1284, Sept. 2009.

[5] M. Naqvi *et al*., "Skin cancer detection using deep learning—A review," *Diagnostics (Basel, Switzerland)*, vol. 13, no. 11, 1911, May 2023.

[6] T. G. Debelee, "Skin lesion classification and detection using machine learning techniques: A systematic review," *Diagnostics*, vol. 13, no. 19, 3147, 2023.

[7] S. Muthukumar and S. R. Prakash, "Deep learning techniques in skin cancer detection: A survey," in *Proc. the 1st International*

Conference on Computational Science and Technology (ICCST), Chennai, India, 2022, pp. 378–381.
[8] K. M. Hosny, M. A. Kassem, and M. M. Foaud, "Classification of skin lesions using transfer learning and augmentation with Alex-net," *PLOS ONE*, vol. 14, no. 5, e0217293, 2019.
[9] A. Jha and V. S. Ananthanarayana, "An interpretable deep learning model for skin lesion classification," in *Proc. the International Conference on Advances in Computing and Data Sciences, ICACDS 2023*, vol. 1848, Springer, Cham, 2023, pp. 44–52.
[10] M. M. Musthafa *et al.*, "Enhanced skin cancer diagnosis using optimized CNN architecture and checkpoints for automated dermatological lesion classification," *BMC Med. Imaging*, vol. 24, 201, 2024.
[11] Y. Nie, P. Sommella, M. Carratù, M. O'Nils, J. Lundgren, "A deep CNN transformer hybrid model for skin lesion classification of dermoscopic images using focal loss," *Diagnostics*, vol. 13, no. 1, p. 72, 2023.
[12] J. Höhn, A. Hekler, E. Krieghoff-Henning *et al*., "Integrating patient data into skin cancer classification using convolutional neural networks: Systematic review," *J. Med. Internet Res*., vol. 23, no. 7, e20708, 2021. doi: 10.2196/20708
[13] M. M. Forootan, I. Larki, R. Zahedi, and A. Ahmadi, "Machine learning and deep learning in energy systems: A review," *Sustainability*, vol. 14, no. 8, 4832, 2022.
[14] C. K. Viknesh, P. N. Kumar, R. Seetharaman, and D. Anitha, "Detection and classification of melanoma skin cancer using image processing technique," *Diagnostics*, vol. 13, no. 21, 3313, 2023.
[15] International Skin Imaging Collaboration (ISIC). ISIC 2019: Skin Lesion Analysis Towards Melanoma Detection. [Online]. Available: https://www.isic-archive.com/
[16] M. L. Rizzo and G. J. Székely, "Energy distance," *WIREs Comput. Stat.*, vol. 8, pp. 27–38, 2016.
[17] G. J. Székely and M. L. Rizzo, "Energy statistics: A class of statistics based on distances," *Journal of Statistical Planning and Inference*, vol. 143, no. 8, pp. 1249–1272, 2013. doi: 10.1016/j.jspi.2013.03.018
[18] Y. LeCun, S. Chopra, M. Hadsell, M. Ranzato, and F. Huang, *A Tutorial on Energy-Based Learning*, MIT Press, 2006.
[19] W. Grathwohl, K. Wang, J. Jacobsen, D. Duvenaud, and K. Swersky, "Your classifier is secretly an energy-based model, and you should treat it like one," in *Proc. Int. Conf. Learning Representations (ICLR)*, 2020.
[20] F. Grignaffini, F. Barbuto, L. Piazzo *et al.*, "Machine learning approaches for skin cancer classification from dermoscopic images: A systematic review," *Algorithms*, vol. 15, no. 11, 438, 2022.
[21] S. Ruder, "An overview of gradient descent optimization algorithms," arXiv preprint, arXiv:1609.04747, 2016.
[22] Y. Song and S. Ermon, "Generative modeling by estimating gradients of the data distribution," in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 32, 2019.