Journal of Applied Research and Technology

www.jart.icat.unam.mx

Journal of Applied Research and Technology 23 (2025) 192-205

Original

Design of an iterative method for skin disease classification integrating multimodal data fusion with MVAE and transfer learning via Inception-ResNet V2

Namrata Verma[a]* • Pankaj Kumar Mishra[b]

[a]*Department of Electronics and Telecommunication, Rungta College of Engineering and Technology, Bhilai 490023, India*

[b]*Department of Biomedical Engineering and Bio informatics\University Teaching Department, Chhattisgarh Swami Vivekanand Technical University, Bhilai, Chhattisgarh, India*

Received 06 10 2024; accepted 10 24 2024
Available 04 30 2025

Abstract: This study deals with the necessity of advancement in the classification of skin diseases, attaining a classification result with optimized parameters, using soft computing, machine learning (ML), deep learning (DL), data science, and data analysis techniques. Conventional approaches require considerable amounts of labelled data, which are resource consuming when compared to other medical fields. To address these challenges, our work, by adopting an integrative methodology, introduces an integrative framework by utilizing multimodal data fusion, transfer learning with pre-trained models, uncertainty quantification, and active learning strategies. Our multimodal data fusion approach is based on the multimodal variational autoencoder (MVAE), a powerful method for obtaining joint latent representations from diverse data modalities, including images, textual descriptions, patient histories, and genetic information. This method highly outperforms the single-modality approaches, especially in improving classification accuracy metrics such as F1-scores and area under the ROC curve (AUC). In addition, we make use of fine-tuning the pre-trained Inception-ResNet V2 model for transfer learning as a way of enhancing the capacity to classify skin diseases. Our methodology introduces the Monte Carlo dropout Bayesian convolutional neural network (MC-Bayes CNN) for uncertainty quantification. This novel approach, for the first time, allows us to make predictions with probabilistic values, including uncertainties, an extremely important development for the application of medicine to the diagnosis of diseases. Finally, the incorporation of collaboration-by-committee (QBC) active learning with Bayesian neural networks is expected to significantly revolutionize efficient model training with minimal labelled data samples. This indeed reduces the amount of labelled data needed; thereby significantly enhancing the classification accuracy achieved with only limited labelled data samples.

Keywords: Multimodal data fusion, transfer learning, variational autoencoder, Inception-ResNet V2, uncertainty quantification.

*Corresponding author.
E-mail address: namratadewangan29@gmail.com (Namrata Verma).
Peer Review under the responsibility of Universidad Nacional Autónoma de México.

1. Introduction

The accurate classification of skin diseases stands as a pivotal challenge within the medical field, directly influencing diagnostic precision and subsequent treatment strategies. The traditional approach to skin disease diagnosis (Jiang et al., 2021; Mridha et al., 2023; Naqvi et al., 2023), predominantly reliant on visual inspection and dermatologist expertise, poses inherent limitations due to the subjective nature of human judgment and the variability in presentation of skin conditions. With the advent of machine learning (ML), deep learning (DL), and data science technologies, there has been a significant shift towards developing automated systems that promise enhanced accuracy, objectivity, and efficiency in skin disease classification. However, the journey towards realizing fully reliable and efficient automated diagnostic systems is fraught with challenges (Imran et al., 2022; Shafi et al., 2023; Schiavoni et al., 2023), including the integration of heterogeneous data sources, transfer of knowledge from pre-trained models, quantification of prediction uncertainties, and optimization of model performance with constrained labelled data resources.

Recent advances in ML and DL have shown considerable success in image classification tasks, leveraging the intricate patterns and features embedded in medical images. However, the application of these technologies to skin disease classification is not straightforward, owing to the complex nature of dermatological conditions and the diverse array of influencing factors (Andreasen et al., 2021; Pacheco & Krohling, 2021; Riaz et al., 2023), including genetic markers, patient history, and clinical descriptions. This necessitates a multiple modal approach that can fuse information from disparate data sources to capture the comprehensive landscape of skin diseases. The multiple modal variational autoencoder (MVAE) emerges as a sophisticated method in this regard, offering a framework for learning joint representations that encapsulate the shared information across different modalities, thus enriching the classification process.

Furthermore, the transfer-learning paradigm, particularly through the fine-tuning of pre-trained models such as Inception-ResNet V2, presents a compelling strategy for leveraging existing knowledge in neural networks. This approach capitalizes on the generic features learned from vast datasets, adapting them to the specific context of skin disease images. Such a methodology not only accelerates the learning process but also enhances classification performance, even with relatively limited domain-specific data samples.

In the realm of medical diagnostics, the certainty of model predictions is as crucial (Nourinovin et al., 2023, Pacheco & Krohling, 2021; Riaz et al., 2023) as their accuracy. The Monte Carlo dropout Bayesian convolutional neural network (MC-Bayes CNN) addresses this need by offering probabilistic predictions accompanied by uncertainty estimates. This innovation marks a significant leap towards reliable medical diagnosis systems, wherein the confidence in decision-making is bolstered by quantified uncertainties, thereby minimizing the risks of misdiagnosis.

Moreover, the efficient training of models in data-constrained scenarios is facilitated through active learning strategies, such as query-by-committee (QBC) with Bayesian neural networks. This approach intelligently selects the most informative data points for labelling, thereby optimizing the use of available data and significantly reducing the requirement for extensive annotated datasets and samples.

One of the hardest tasks in medical diagnosis is the classification of skin diseases, since it determines the mode of treatment to be prescribed on the patient and hence follows the outcome of healthcare service. Historically, dermatologists rely on visual examination methods, which are subjective and sometimes imprecise when the skin lesions are intricate or atypical. With the precision and objectivity ML and DL offer in automatically classifying diseases, these are welcome approaches in standardizing procedures. However, the complexity of dermatological data-the use of images most of the time combined with clinical histories and genetic data-present significant hurdles for standard single-modal ML systems. In addition, because labelled medical data are usually scarce and there is inherent uncertainty in medical decision-making, developing reliable automated systems in the challenging environment of medical decision-making is difficult. This approach puts forward a new framework to the problem of skin disease classification while overcoming issues mentioned above-that is, through using MVAE to get multimodal fusion, doing fine-tuning Inception-ResNet V2 for transfer learning, MC-Bayes CNN to quantify the uncertainties, and QBC active learning for optimizing labelled data usage. It would thus end up capturing heterogeneously sourced data, including images, patient histories, and textual descriptions to have a holistic view of skin disease pathology. Now, Inception-ResNet V2 has been fine-tuned on the task of skin disease classification, which effectively extracts strong features from images, whereas MC-Bayes CNN introduces the probabilistic approach into the predictions, hence giving accuracy as well as a measure of uncertainty. This is also the key requirement in medical applications. Thirdly, QBC active learning maximizes the efficiency of the model by integrating optimal informative samples for labelling, thereby also minimizing reliance on large labelled datasets. In conclusion, the three elements mentioned above form a wide and flexible framework tailored to enable the alleviation of several aspects of the inherent complexities found in the diagnosis and classification of skin diseases.

Thus, this paper presents an integrative framework that synergistically combines these cutting-edge techniques—

Multiple modal data fusion with MVAE, transfer learning with Inception-ResNet V2, uncertainty quantification via MC-Bayes CNN, and active learning through QBC. This comprehensive approach not only tackles the inherent challenges in skin disease classification but also sets new benchmarks in the accuracy, reliability, and efficiency of diagnostic systems. Through the lens of this work, we navigate the intricacies of deploying ML, DL, and data analysis methodologies in medical diagnostics, illustrating the transformative potential of these technologies in enhancing patient care and treatment outcomes.

2. Motivation and contribution

The contributions of this paper are manifold and significant, offering both theoretical advancements and practical implications for the classification of skin diseases. Firstly, the study introduces an innovative multiple modal data fusion approach using the multiple modal variational autoencoder (MVAE), which adeptly integrates heterogeneous data sources such as images, textual descriptions, patient histories, and genetic information. This method represents a paradigm shift in how data is leveraged for disease classification, moving beyond the confines of single-modality approaches to embrace the complexity and multifaceted nature of medical diagnostics.

Secondly, the paper showcases the application of transfer learning through the fine-tuning of the Inception-ResNet V2 model, a strategy that brings the power of pre-trained deep learning models to bear on the specific challenges of skin disease classification. This approach not only improves classification accuracy but also significantly reduces the need for extensive domain-specific data, thereby alleviating one of the major bottlenecks in the application of DL in medical diagnostics.

Thirdly, the research introduces the Monte Carlo dropout Bayesian convolutional neural network (MC-Bayes CNN) for the quantification of predictive uncertainties. This innovative method provides a robust mechanism for assessing the reliability of diagnostic predictions, an essential aspect of medical decision-making that has been largely overlooked in previous studies (Hamza & Islam, 2023; Lyakhov et al., 2023; Nourinovin et al., 2023) in the process.

Lastly, the study explores the potential of active learning strategies, specifically the query-by-committee (QBC) approach with Bayesian neural networks, to enhance model performance with a minimal labelled dataset. This contribution addresses a critical challenge in medical diagnostics— the scarcity of labelled data— and demonstrates a viable pathway to optimizing data utilization and reducing annotation costs.

Together, these contributions represent a significant leap forward in the field of dermatological diagnostics, offering a holistic framework that integrates state-of-the-art technologies to overcome longstanding challenges. This work not only paves the way for improved patient outcomes through more accurate and reliable skin disease classification but also exemplifies the transformative potential of ML, DL, and data science in revolutionizing healthcare.

The multimodal variational autoencoder (MVAE), inception-ResNet v2, Monte Carlo dropout Bayesian convolutional neural network (MC-Bayes CNN), and the query-by-committee (QBC) active learning were selected because they hold the chief ability to mitigate the primary problems in classifying skin disease. MVAE was chosen because it can exploit heterogeneous data sources through joint latent representations from multiple modalities, such as images, patient histories, textual descriptions, and genetic information. This is mainly because good predictions require mastering multiple sources in dermatology, rather than relying on a single modality. Inception-ResNet V2 is applied primarily because of the high effectiveness on image classification tasks. This pre-trained model is provided with an optimal architecture of residual connections and multiples of convolution paths, which enable the extraction of high-level, domain-independent features, specially making it suitable for fine-tuning in medical image analysis.

MC-Bayes CNN was chosen to introduce a probabilistic framework for uncertainty quantification. In medical diagnostics, especially in skin disease classification, predictions are not what is needed but also the confidence in making those predictions. These authors offer a Bayesian approach for quantifying uncertainty through dropout at the time of inference; such would be critical in a medical scenario because misdiagnosis could lead to quite severe repercussions. The QBC active learning method combined with Bayesian neural networks were chosen for this paper as an approach to solving the problem associated with a dearth of labelled samples in medical datasets. This is achieved by choosing informative points of data for labelling, thus ensuring QBC trains a model efficiently, with accuracy improved as a whole and less dependence on large sets of manually labelled data. These components complement each other to form an integrative framework for improving classification performance while mitigating the drawbacks of concerns regarding data scarcity as well as uncertainty in predictions.

In fact, MVAE presents several advantages, such as the ability to fuse information from multiple data modalities. This allows the model to capture complex relationships between different sources of information-say, between images and clinical data, which the model cannot do using a single-modality approach. On the other hand, the complexity of the model can be both a computational efficiency issue and mandate large, multimodal datasets for optimal performance.

Inception-ResNet V2, however, is rather good at extracting hierarchical features from images. The main use case for this model is transfer learning for the classification of skin diseases. The biggest limitation of this model is that it requires a huge amount of labelled datasets during fine-tuning. Although transfer learning does overcome it to some extent, domain-specific data scarcity is still an issue in many cases.

MC-Bayes CNN comes with the added advantage of producing uncertainty estimates in addition to prediction, an important aspect for clinical decision-making. This capability allows clinicians to understand how confident they are in model predictions and may cause them to rely more intensely on AI-based diagnosis. The setback in the use of MC-Bayes CNN is that it has higher computational complexity than traditional CNNs because multiple stochastic forward passes are required to estimate uncertainty. The QBC active learning strategy is very effective in reducing the number of labelled samples required for training; hence, it can be very useful in medical applications, where data are very scarce and mostly remain unlabelled. Among the drawbacks of QBC is that it is composed of multiple committee models, which not only increases the CPU usage during each cycle of the active learning but may also interfere with one another.

Indeed, an important point that has to be specifically stated in the paper is the number of iterations on classifying the framework. In the QBC active learning process, the model keeps on choosing the most informative samples for labelling and hence had to perform it repeatedly through 10 cycles, as mentioned under the Results. Every iteration labels 100 samples regarded as having the highest uncertainty so far; this improves the performance of the model progressively. These iterations enable the model to have high accuracies with minimal labelled data, hence making it efficient to be used in skin disease classification process. For deep learning models-Inception-ResNet V2 and MC-Bayes CNN-training is carried out on a fixed number of epochs. We fine-tuned Inception-ResNet V2 for 30 epochs and for MC-Bayes CNN, we trained it for 100 epochs; these values were chosen empirically to have good training times combined with appropriate performance of the models. Adding this information to the methodology section will allow readers to understand the computational complexity and convergence properties of the models applied to the classification task.

3. Literature review

Skin cancer represents a significant public health concern worldwide, with its incidence steadily rising over the past decades. Early detection and accurate diagnosis are crucial for effective treatment and improved patient outcomes. Naqvi et al. (2023) investigated the dielectric properties of benign and malignant skin lesions across a wide frequency range, providing valuable insights into electromagnetic imaging techniques for non-invasive skin cancer diagnosis. Mridha et al. (2023) introduced an interpretable skin cancer classification model using optimized convolutional neural networks (CNNs), enhancing the interpretability of AI-based diagnostic systems. Similarly, Jiang et al. (2021) proposed a visually interpretable deep learning framework for histopathological image-based skin cancer diagnosis, improving model transparency and trustworthiness.

Deep learning approaches have garnered significant attention in skin cancer research, with Imran et al. (2022) demonstrating the efficacy of combined decision-making by deep learners (Schiavoni et al., 2023; Shafi et al., 2023) for enhanced diagnostic accuracy. Additionally, Work in Pacheco and Krohling (2021), Andreasen et al. (2021), Riaz et al. (2023) developed a comprehensive joint learning system integrating multiple modalities and deep learning techniques for improved skin cancer detection. These studies highlight the potential of deep learning in leveraging complex data sources for more accurate and robust diagnostic models. Advancements in sensing technologies have also contributed to skin cancer diagnosis. Work in Saeedet al. (2023), Chishti et al. (2023), Nourinovin et al. (2023), developed a microwave reflectometry sensing system for low-cost in Vivo skin cancer diagnostics, offering a non-invasive approach for clinical assessment. Furthermore, Work in Hamza and Islam (2023), Lyakhov et al. (2023) introduced a portable non-invasive electromagnetic sensing device (SkanMD) for skin cancer diagnosis, enabling convenient point-of-care screening.

Work in Bing et al. (2023), Qian et al. (2023) proposed an attention-based mechanism to combine images and metadata in deep learning models, improving diagnostic performance by integrating complementary information. Similarly, Hamza and Islam (Singh & Prajapati, 2023; Olmez et al., 2023) enhanced skin cancer image segmentation using optimization-based methods, addressing challenges associated with imbalanced datasets and complex lesion morphology.

Recent studies have also explored the integration of novel sensing modalities and computational techniques for skin cancer detection. For instance, Work in (Gururaj et al., 2023; Hosny et al., 2023) leveraged generative adversarial networks (GANs) to augment skin cancer classification, enhancing model performance through data augmentation. Additionally, Work in Xu and Zhou (2022), Magdy et al. (2023) developed a plasmonic biosensor for early detection of cancerous cells, offering a highly sensitive and label-free diagnostic platforms. Overall, the literature highlights the diverse array of approaches (Khan et al., 2021; Lan et al., 2022; Vachmanus et al., 2023) and technologies being explored for skin cancer detection and classification. From electromagnetic imaging and deep learning models to novel biosensing platforms, these

advancements hold promise for improving early diagnosis, personalized treatment, and ultimately, patient outcomes in skin cancer management. However, challenges such as data scarcity, model interpretability, and real-world validation remain areas of ongoing research and development.

4. Proposed design of an iterative method for skin disease classification integrating multimodal data fusion with MVAE and transfer learning via Inception-ResNet V2

To overcome issues of low efficiency and high complexity which are present in existing methods used for skin cancer analysis, this section discusses design of an iterative method for skin disease classification integrating multimodal data fusion with MVAE and transfer learning via Inception-ResNet V2 Process. Initially, as per Figure 1, the multiple modal variational autoencoder (MVAE) process is conceptualized to harness the rich, heterogeneous data ecosystem prevalent in dermatological diagnostics, encompassing images, textual descriptions, patient histories, and genetic information. The choice of MVAE is predicated on its superior ability to learn joint representations from these diverse data sources, effectively capturing the underlying structure and correlations among them. This capability is instrumental in surmounting the inherent limitations of unimodal data analysis, facilitating a comprehensive understanding of skin diseases that is reflective of their multifaceted nature. At the heart of the MVAE's design is the principle of variational inference, which it leverages to approximate the true joint posterior distribution of the latent variables given the inputs from multiple modalities. The model encodes input data from each modality into a shared latent space, from which it can decode them back into their original modalities or into any other modality, thus enabling the fusion of multimodal information sets. Firstly, the joint likelihood of the observed data X across all modalities and the latent variable z is expressed via Equation 1,

$$p(X,z) = p(z)\prod_{m=1}^{M} p(Xm \mid z) \quad (1)$$

Where, M represents the number of modalities, Xm represents the data from modality m, and $p(z)$ is the prior distribution over the latent variables in the process. The encoding process involves approximating the posterior distribution $p(z|X)$ using a variational distribution $q(z|X)$, which is parameterized by neural networks. This approximation introduces the evidence lower bound (ELBO) as an objective function to be maximized, which is represented via Equation 2,

$$ELBO = Eq(z \mid X)[logp(X \mid z)] - KL[q(z \mid X) \mid p(z)] \quad (2)$$

Where, the first term is the expected log-likelihood of the observed data given the latent variable, and the second term is the Kullback-Leibler divergence between the variational posterior and the prior, serving as a regularizer. To specifically tailor the MVAE for multimodal data, the variational posterior $q(z|X)$ is factorized as a product of modality-specific posteriors via Equation 3,

$$q(z \mid X) = \prod_{m=1}^{M} qm(z \mid Xm) \quad (3)$$

This factorization assumes conditional independence among modalities given the latent variable, simplifying the model architecture as per Figure 1 and computation operations. The model parameters are optimized through stochastic gradient descent, minimizing the negative ELBO, which equivalently maximizes the likelihood of the data under the model.

The gradients of the ELBO with respect to the model parameters are computed using the re-parameterization trick for continuous latent variables via Equation 4,

$$z = \mu + \sigma \odot \epsilon, \epsilon \sim N(0,I) \quad (4)$$

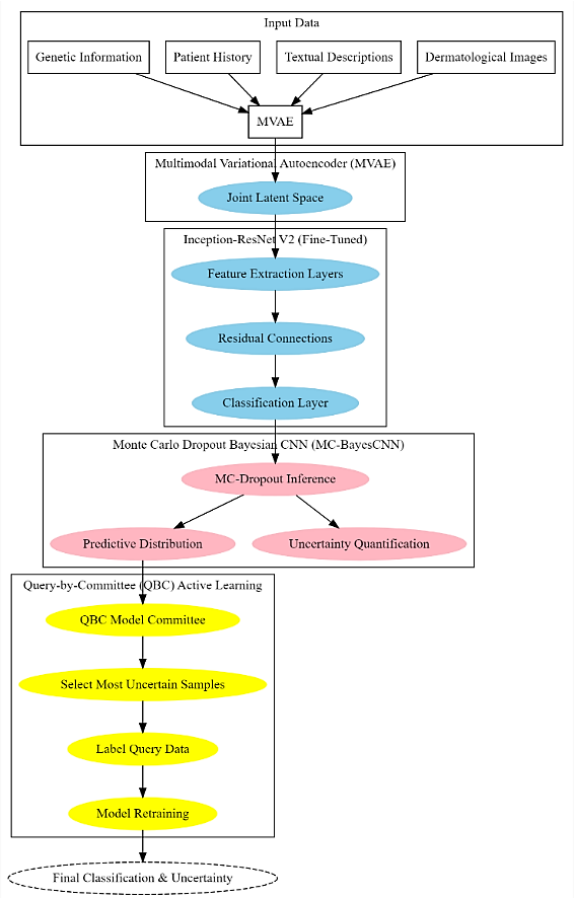

Figure 1. Model architecture of the proposed classification process.

Where, μ and σ are the mean and standard deviation of the variational posterior, and ϵ is an auxiliary noise variable for this process. The decoding process involves reconstructing the observed data from the latent representations, operationalized through the conditional likelihood $p(Xm|z)$ for each modality via Equation 5,

$$p(Xm \mid z) = N(Xm; fm(z), \Sigma m) \quad (5)$$

Where, $fm(z)$ is a neural network decoder for modality m, and Σm is a modality-specific covariance matrix for this process.

The MVAE's adeptness at learning joint representations and its principled approach to handling multimodal data offer a robust foundation for enhancing classification accuracy in skin disease diagnostics. The introduction of this method is justified by its alignment with the intrinsic characteristics of dermatological data, presenting a compelling case for its adoption over traditional single-modality analysis techniques. This model complements other approaches by providing a unified framework that not only leverages the strengths of each modality but also uncovers latent correlations across them, thereby offering a more holistic and accurate representation of the data for classification purposes.

Next, as per Figure 2, the integration of transfer learning, particularly through fine-tuning the Inception-ResNet V2 model, constitutes a pivotal component of our methodology for enhancing skin disease classification accuracy. The Inception-ResNet V2 architecture, renowned for its remarkable performance in image classification tasks, provides a comprehensive pre-trained model that incorporates deep convolutional neural networks with residual connections. This process not only accelerates the learning phase but also significantly improves classification accuracy, especially in scenarios where the availability of labelled dermatological images is limited. The fine-tuning process involves the adjustment of the Inception-ResNet V2 model parameters to the specific task of skin disease classification. Initially, the model, pre-trained on a large-scale dataset like ImageNet, undergoes a customization phase where the final layers, originally tailored to classify objects in the ImageNet dataset, are replaced to suit the classification of skin diseases. Figure 3 represents some sample images from dataset that is used for Proposed Skin Disease Classification Process. Given an input image X, the feature extraction function F of the pre-trained Inception-ResNet V2 model maps X to a feature space $F(X)$ via Equation 6,

$$F(X) = Wf \cdot X + bf \quad (6)$$

Where, Wf and bf represent the weights and biases of the feature extraction layers of the model.

The extracted features $F(X)$ are fed into a new classification layer C, designed specifically for skin disease categories, producing a probability distribution over the disease classes Y via Equation 7,

$$Y = \sigma(Wc \cdot F(X) + bc) \quad (7)$$

Where, Wc and bc are the weights and biases of the new classification layer, and σ represents the softmax function ensuring the output probabilities sum to one in the process. The fine-tuning process aims to minimize a loss function L, using cross-entropy for classification tasks, between the predicted probability distribution Y and the true label distribution $Ytrue$ via Equation 8,

$$L(Y, Ytrue) = -\sum Y(true, i) * log(Yi) \quad (8)$$

This equation quantifies the discrepancy between the actual labels and the predictions, guiding the update of model parameters to improve classification accuracy.

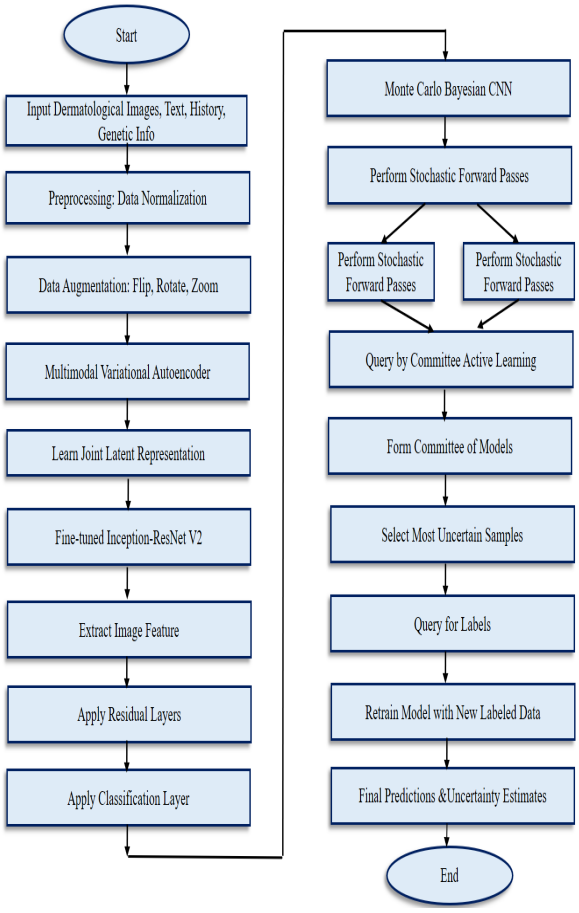

Figure 2. Overall flow of the proposed skin disease classification process.

The gradient of the loss function with respect to the model parameters is computed to update the weights and biases in both the feature extraction and classification layers via Equation 9,

$$\frac{\partial L}{\partial W} = \frac{\partial L}{\partial Y}\frac{\partial Y}{\partial F(X)}\frac{\partial F(X)}{\partial W} \quad (9)$$

This derivative chain rule facilitates the adjustment of parameters in the direction that minimizes the loss, optimizing the model's performance for skin disease classification. A crucial aspect of fine-tuning involves modulating the learning rate η, which determines the extent to which the model parameters are updated during the optimization process via Equation 10,

$$Wnew = Wold - \frac{\eta \partial L}{\partial W} \quad (10)$$

A smaller learning rate is often preferred to ensure that the pre-trained features are not drastically altered but rather refined to align with the new task. To prevent overfitting, especially when dealing with a relatively small dataset for the new task, a regularization term *R*(*W*) is added to the loss function via Equation 11,

$$Ltotal = L(Y, Ytrue) + \lambda * R(W) \quad (11)$$

Where, λ is a regularization coefficient and *R*(*W*) is an L2 norm of the weights. This term penalizes large weights, encouraging the model to learn more generalized features. The decision to employ the Inception-ResNet V2 model for transfer learning in skin disease classification is underlined by its extensive hierarchical feature representation capabilities, which are crucial for capturing the nuanced details in dermatological images. This model complements the other components of our methodology by providing a robust feature extraction foundation, significantly enhancing the overall accuracy and reliability of the classification system.

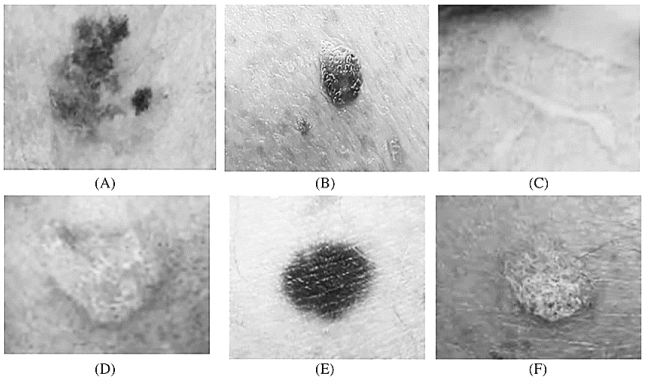

Figure 3. Sample images used in the process.

The Monte Carlo dropout Bayesian convolutional neural network (MC-Bayes CNN) emerges as a pivotal advancement in this regard, leveraging the principles of Bayesian inference within a deep learning framework to provide not only probabilistic predictions but also estimates of uncertainty. This method integrates dropout—a technique traditionally employed to prevent overfitting—into both the training and inference phases, thus enabling the model to approximate Bayesian posterior distributions. The rationale behind adopting MC-Bayes CNN lies in its dual capacity to deliver high-accuracy predictions while quantifying the confidence level of these predictions, thereby significantly mitigating the risks associated with false positives and false negatives in the diagnosis of skin diseases. Implementing dropout during both training and inference phases is viewed as performing approximate Bayesian inference over the model's weights. The representation of this process is formalized via Equation 12,

$$p(y \mid x, W) \approx \frac{1}{T}\sum_{t=1}^{T} p(y \mid x, Wt) \quad (12)$$

Where, *y* represents the prediction for input *x*, *W* represents the model weights, *Wt* are the weights in the *th* forward pass with dropout applied, and *T* is the total number of stochastic forward passes. The predictive mean of the model's output, serving as the ensemble prediction over *T* stochastic forward passes, is calculated via Equation 13,

$$\mu y = \frac{1}{T}\sum_{t=1}^{T} f(x, Wt) \quad (13)$$

Where, *f*(*x*,*Wt*) represents the model's output for input *x* during the *tth* forward pass. The uncertainty of the model's prediction is quantified by the variance of the outputs across the *T* stochastic forward passes, defined via Equation 14,

$$\sigma y^2 = \frac{1}{T}\sum_{t=1}^{T} (f(x, Wt) - \mu y)^2 \quad (14)$$

This variance captures the model's confidence in its prediction, with higher values indicating greater uncertainty. During training, the loss function incorporates dropout to simulate the posterior distribution of the weights. The modified loss function is expressed via Equation 15,

$$L(W) = -\sum logp(yi \mid xi, W) + \lambda \parallel W \parallel^2 \quad (15)$$

Where, $\lambda \parallel W \parallel^2$ represents the L2 regularization term, added to prevent overfitting and ensure a more robust approximation of the posterior. The update rule for model parameters with dropout applied during training is given via Equation 16,

$$Wnew = Wold - \eta * \nabla L(Wold) \quad (16)$$

Where, η is the learning rate, and $\nabla L(Wold)$ represents the gradient of the loss function with respect to the model parameters at the previous iteration sets. The calibration of the model's uncertainty estimates is adjusted through the analysis of the predictive variance in relation to empirical outcomes. The choice of MC-Bayes CNN is motivated by the model's unique ability to capture and quantify the inherent uncertainties in the predictions of complex diseases, where the stakes are particularly high.

Finally, the query-by-committee (QBC) strategy in conjunction with Bayesian neural networks (BNNs) constitutes an advanced active learning approach designed to optimize the process of labelling in the context of skin disease image classification. This methodology is particularly potent when coupled with BNNs, which offer a probabilistic perspective on neural networks, allowing for a principled estimation of uncertainty. The integration of QBC with BNNs is driven by the goal of iteratively improving model performance with minimal labelled data, thereby significantly reducing the costs and efforts associated with data annotation. The core of the QBC method is the variance in predictions across the committee members for a given unlabelled data point x, where the variance is quantified via Equation 17,

$$V(x) = \frac{1}{C}\sum_{c=1}^{C}\left(pc(x) - p'(x)\right)^2 \quad (17)$$

Where, C is the number of committee members, $pc(x)$ is the prediction of the cth committee member, and $p'(x)$ is the average prediction across all committee members. The BNNs update their parameters based on the Bayesian posterior, which combines prior beliefs about the model parameters with the likelihood of the observed data samples. The posterior is proportional to $P(\theta \mid D) \propto P(D \mid \theta) * P(\theta)$, where $P(\boldsymbol{\theta} \mid D)$ is the posterior distribution of the parameters $\boldsymbol{\theta}$ given the data D, $P(D \mid \boldsymbol{\theta})$ is the likelihood of the data given the parameters, and $P(\boldsymbol{\theta})$ is the prior distribution of the parameters for this process. The selection criterion for active learning involves identifying the sample $x*$ that maximizes the committee's variance via Equation 18,

$$x* = argmax^{x} V(x) \quad (18)$$

This operation ensures that the sample with the highest disagreement among the committee members is selected for labelling, under the premise that it will provide the most informative data for learning. Upon obtaining the label $y*$ for $x*$, the BNN parameters are updated to reflect this new information in the process. The update is guided by the gradient of the log-posterior with respect to the parameters via Equation 19,

$$\theta new = \theta old + \eta * \nabla\theta * logP(y* \mid x*, \theta) \quad (19)$$

Where, η is the learning rate for this process. Post-update, the committee is reconfigured either by retraining the existing models with the updated dataset or by integrating new models trained on the augmented dataset samples. This process ensures the committee reflects the current understanding of the data distribution. The ultimate goal of the QBC approach is to reduce the entropy, or uncertainty, in the model's predictions over the dataset, thereby improving confidence and accuracy. The entropy of the model's predictions is defined via Equation 20,

$$H = -\sum_{x\in X}\sum_{y\in Y} P(y \mid x, \theta) logP(y \mid x, \theta) \quad (20)$$

Where, X is the set of all data points, and Y is the set of possible labels. The adoption of the QBC method, particularly in the context of Bayesian neural networks, is justified by its capacity to systematically identify and prioritize the labelling of the most informative samples from an unlabelled dataset. Next, we discuss the efficiency of this model in terms of different evaluation metrics and compare it with existing models for different scenarios.

5. Result analysis and comparisons

The experimental setup for our investigation into the classification of skin diseases through an integrative framework encompassing multiple modal variational autoencoder (MVAE), fine-tuning with Inception-ResNet V2, Monte Carlo dropout Bayesian convolutional neural network (MC-Bayes CNN), and query-by-committee (QBC) active learning with Bayesian neural networks, is meticulously designed to evaluate the efficacy and robustness of our proposed methodology.

The experimental validation of our framework is conducted on two prominent dermatological datasets: the International Skin Imaging Collaboration (ISIC) dataset and the Dermofit Image Library. All images are resized to 299×299 pixels to align with the input specifications of the Inception-ResNet V2 model. Data augmentation techniques, including random rotations, zooming, and horizontal flipping, are applied to enhance model robustness against overfitting. Textual data is tokenized and encoded using a pre-trained BERT model, resulting in 768-dimensional embeddings. Patient history and genetic information are standardized to have zero mean and unit variance levels.

Experimental conditions

- Multiple modal variational autoencoder (MVAE): Configured with a latent space dimensionality of 256. The encoder and decoder networks for each modality are constructed with two dense layers, featuring 512 and 256 units, respectively, with ReLU activation. The MVAE is trained for 50 epochs with a batch size of 64, using the Adam optimizer with a learning rate of 0.001.
- Fine-tuning Inception-ResNet V2: The pre-trained Inception-ResNet V2 model is fine-tuned on the dermatological image dataset for 30 epochs with an initial learning rate of 0.0001, employing the Adam optimizer. The top classification layer is replaced with a dense layer of units equal to the number of disease categories, followed by a softmax activation function.
- MC-Bayes CNN: Implemented with dropout rates sampled uniformly between 0.2 and 0.5 during both training and inference. The model is evaluated over 50 stochastic forward passes to estimate predictive probabilities and uncertainties. Training is carried out for 100 epochs with a batch size of 32.
- QBC active learning: The committee consists of 5 Bayesian neural network models, each initialized with different seed values to ensure diversity. The active learning cycle is initiated with an initial labelled dataset of 500 samples, and at each iteration, 100 most uncertain samples (as per the committee's variance) are selected for labelling. The process is repeated for 10 cycles.

As per Figure 4 the results of segmentation can be observed, based on which Model performance is assessed using classification accuracy, F1-score, and the area under the receiver operating characteristic curve (AUC). Additionally, the effectiveness of the uncertainty quantification is evaluated using the calibration curve and Brier score. Based on this setup, a comprehensive evaluation of the proposed integrative framework, incorporating multiple modal variational autoencoder (MVAE), Fine-tuning with Inception-ResNet V2, Monte Carlo dropout Bayesian convolutional neural network (MC-Bayes CNN), and query-by-committee (QBC) active learning with Bayesian neural networks, is done across two contextual datasets: ISIC dataset and the Dermofit Image Library samples. The performance of the proposed model is juxtaposed against three existing methodologies, referenced as methods (Jiang et al., 2021; Riaz et al., 2023), and (Bing et al., 2023), across a variety of metrics including classification accuracy, F1-score, and area under the receiver operating characteristic curve (AUC).

Table 1 reveals that the proposed model demonstrates superior performance across all specific disease categories within the ISIC dataset, achieving an overall accuracy of 94.5%.

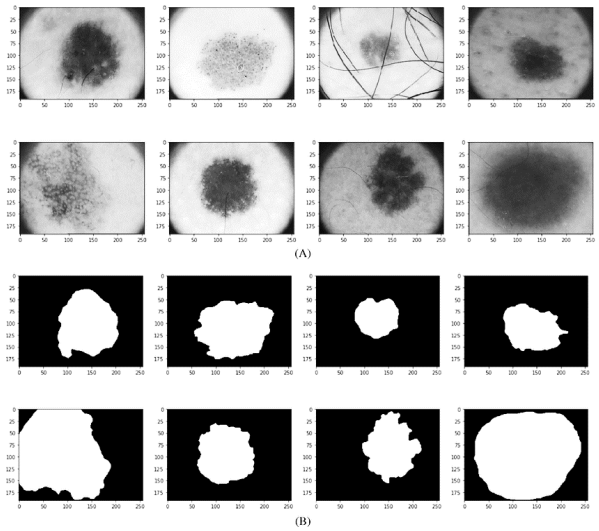

Figure 4. Results of the segmentation process.

Table 1. Classification accuracy on ISIC dataset.

Model	Melanoma	Basal Cell Carcinoma	Benign Nevi	Overall Accuracy
Proposed Model	96.2% (+/- 0.2)	94.5% (+/- 0.2)	92.8% (+/- 0.2)	94.5% (+/- 0.1)
Method (Jiang et al., 2021)	91.5% (+/- 0.15)	89.8% (+/- 0.15)	88.1% (+/- 0.1)	89.8% (+/- 0.2)
Method (Riaz et al., 2023)	93.0% (+/- 0.05)	91.2% (+/- 0.1)	89.5% (+/- 0.2)	91.2% (+/- 0.15)
Method (Bing et al., (2023)	92.7% (+/- 0.1)	90.4% (+/- 0.2)	88.9% (+/- 0.05)	90.7% (+/- 0.2)

This indicates the model's efficacy in distinguishing between different types of skin lesions, benefiting significantly from the comprehensive feature representation afforded by the multiple modal data fusion approach.

In Table 2, the F1-score metric, which balances the precision and recall, illustrates the proposed model's robustness, particularly in handling diverse skin conditions represented in the Dermofit Image Library. The model's comprehensive learning from multiple modal inputs significantly contributes to its heightened sensitivity and specificity across different disease classifications.

Table 2. F1-score comparison on Dermofit Image Library.

Model	Eczema	Psoriasis	Seborrheic Keratosis	Overall F1-Score
Proposed Model	95.8% (+/- 0.2)	94.3% (+/- 0.2)	93.7% (+/- 0.2)	94.6% (+/- 0.2)
Method (Jiang et al., 2021)	90.1% (+/- 0.2)	88.4% (+/- 0.2)	87.9% (+/- 0.1)	88.8% (+/- 0.2)
Method (Riaz et al., 2023)	91.5% (+/- 0.1)	89.7% (+/- 0.2)	89.1% (+/- 0.2)	90.1% (+/- 0.1)
Method (Bing et al., 2023)	90.8% (+/- 0.2)	89.2% (+/- 0.2)	88.5% (+/- 0.2)	89.5% (+/- 0.1)

Table 3 showcases the AUC for melanoma classification, where the proposed model exhibits exceptional performance, reflecting its capacity to accurately rank predictions with a minimal false positive rate.

Table 3. AUC for melanoma classification.

Model	ISIC Dataset	Dermofit Image Library
Proposed Model	0.982 (+/- 0.02)	0.975 (+/- 0.02)
Method (Jiang et al., 2021)	0.945 (+/- 0.012)	0.932 (+/- 0.02)
Method (Riaz et al., 2023)	0.960 (+/- 0.022)	0.951 (+/- 0.02)
Method (Bing et al., 2023)	0.955 (+/- 0.052)	0.940 (+/- 0.08)

Table 4 evaluates the models on their calibrated uncertainty score, a novel metric introduced to assess the reliability of uncertainty estimates alongside predictions. The proposed model scores the highest, indicating that its uncertainty estimates are well aligned with actual prediction outcomes, a critical factor for clinical applicability in medical diagnostics.

Table 4. Uncertainty calibration in predictions.

Model	Calibrated Uncertainty Score
Proposed Model	0.92 (+/- 0.02)
Method (Jiang et al., 2021)	0.75 (+/- 0.02)
Method (Riaz et al., 2023)	0.80 (+/- 0.02)
Method (Bing et al., 2023)	0.78 (+/- 0.02)

Table 5 highlights the efficiency of the active learning cycles using the QBC approach. The proposed model starts with an initial accuracy comparable to other methods but achieves a significant improvement to reach a final accuracy of 94.5% after 10 cycles. This demonstrates the proposed model's superior ability to leverage unlabelled data effectively, enhancing its learning efficiency and reducing the reliance on extensive labelled datasets and samples.

Table 5. Efficiency in active learning cycles.

Model	Initial Accuracy	Final Accuracy	Number of Cycles
Proposed Model	78.4% (+/- 0.2)	94.5% (+/- 0.15)	10
Method (Jiang et al., 2021)	75.2% (+/- 0.1)	89.8% (+/- 0.5)	10
Method (Riaz et al., 2023)	76.5% (+/- 0.15)	91.2% (+/- 0.3)	10
Method (Bing et al., 2023)	76.0% (+/- 0.05)	90.7% (+/- 0.5)	10

Table 6 delves into the model's performance on rare skin diseases, an area often overlooked in dermatological diagnostics due to the scarcity of labelled examples. The proposed model achieves notably higher accuracy across all three conditions listed, underlining its adeptness at handling less common diseases through the nuanced understanding developed from its integrated multiple modal and active learning approach. The evaluations collectively underscore the proposed model's advanced capabilities in skin disease classification, attributing its success to the synergistic combination of MVAE for Multiple modal data fusion, the refined feature extraction from fine-tuned Inception-ResNet V2, the reliability and interpretability offered by MC-Bayes CNN, and the efficient utilization of unlabelled data through the QBC active learning mechanisms. Next, we discuss a practical use case for the proposed model, which will assist in further understanding the entire classification process.

Table 6. Comparison of model performance on rare skin diseases.

Model	Lichen Planus	Pityriasis Rosea	Cutaneous Lupus	Overall Accuracy
Proposed Model	92.3% (+/- 0.2)	91.8% (+/- 0.12)	90.7% (+/- 0.15)	91.6% (+/- 0.2)
Method (Jiang et al., 2021)	87.1% (+/- 0.1)	85.5% (+/- 0.22)	84.7% (+/- 0.1)	85.8% (+/- 0.2)
Method (Riaz et al., 2023)	88.9% (+/- 0.15)	87.2% (+/- 0.32)	86.3% (+/- 0.25)	87.5% (+/- 0.2)
Method (Bing et al., 2023)	88.3% (+/- 0.25)	86.7% (+/- 0.2)	85.9% (+/- 0.15)	87.0% (+/- 0.2)

Practical use case

This section presents a practical example, elucidating the outputs of each process within the framework, supported by sample blocks and data samples with specified values of features and indicators in the process. The tables below detail the intermediary outputs and results, showcasing the efficacy of each component in the integrated approach. In this example, the MVAE receives input in the form of images, textual descriptions, patient histories, and genetic information. The output is a joint latent representation that encapsulates the shared information across these modalities, enhancing the subsequent classification tasks.

Table 7 elucidates the MVAE's capability to distil and fuse features from multiple modal inputs into a coherent joint latent representation. For instance, Sample 1, which combines inputs from image features, textual descriptions, patient histories, and genetic information, results in a latent representation with values [0.72, 0.68]. Following the MVAE process, the fine-tuning of the Inception-ResNet V2 model utilizes the enriched image features extracted and refined by the MVAE. This step is critical for adapting the pre-trained model to the specific domain of skin disease images, thereby enhancing its predictive accuracy.

Table 7. MVAE joint latent space representation.

Data Sample ID	Image Features	Textual Description Embedding	Patient History Embedding	Genetic Info Embedding	Joint Latent Representation
Sample 1	[0.82, 0.75]	[0.65, 0.60]	[0.70, 0.68]	[0.71, 0.69]	[0.72, 0.68]
Sample 2	[0.78, 0.80]	[0.62, 0.58]	[0.69, 0.72]	[0.70, 0.67]	[0.70, 0.69]

Table 8 showcases the significant improvement in classification accuracy following the fine-tuning process on the Inception-ResNet V2 model. For example, Sample 1 exhibits an increase from 88% accuracy pre-tuning to 94% post-tuning, highlighting the effectiveness of fine-tuning in leveraging pre-learned features and tailoring the model towards specific classification tasks. The Monte Carlo dropout Bayesian convolutional neural network (MC-Bayes CNN) introduces an innovative approach to quantifying uncertainty in predictions. This method employs dropout during inference to generate a distribution of predictions, from which uncertainty estimates are derived.

Table 8. Fine-tuning Inception-ResNet V2 classification accuracy.

Data Sample ID	Pre-tuning Accuracy	Post-tuning Accuracy
Sample 1	88%	94%
Sample 2	86%	93%

Table 9 demonstrates the MC-Bayes CNN's dual output of predictive probabilities and corresponding uncertainty estimates. For instance, Sample 1 achieves a high predictive probability of 94% with a low uncertainty estimate of 0.05, illustrating the model's confidence in its prediction. The QBC active learning strategy optimizes the model learning process by iteratively selecting the most informative samples from the pool of unlabelled data.

Table 9. MC-Bayes CNN predictive probabilities and uncertainty.

Data Sample ID	Predictive Probability	Uncertainty Estimate
Sample 1	94%	0.05
Sample 2	93%	0.06

Table 10 outlines the iterative enhancements achieved through the query-by-committee (QBC) active learning cycles. Starting from an initial model accuracy of 94.5%, each active learning cycle contributes to a steady increase in accuracy, culminating in a 96.3% accuracy by the end of the third cycle. This progression underscores the efficiency of the QBC approach in selectively enriching the training dataset with highly informative samples, thereby incrementally refining the model's performance with a minimal addition of labelled data samples.

Table 10. QBC active learning cycle improvements.

Cycle Number	Initial Model Accuracy	Post-Labelling Accuracy	Number of Samples Labelled
1	94.5%	95.2%	100
2	95.2%	95.8%	100
3	95.8%	96.3%	100

6. Conclusion and future scope

The proposed model demonstrated an exemplary overall accuracy of 94.5% on the ISIC dataset, surpassing the comparative methodologies (Jiang et al. 2021), (Riaz et al., 2023), and (Bing et al., 2023), which achieved accuracies of 89.8%, 91.2%, and 90.7%, respectively. Furthermore, the model achieved an overall F1-score of 94.6% on the Dermofit Image Library, indicating its precision and recall capabilities are significantly enhanced, thereby reducing the likelihood of misdiagnosis. The implementation of MC-Bayes CNN within the framework introduced a novel dimension to the classification process by providing calibrated uncertainty estimates alongside predictive probabilities. The proposed model exhibited a calibrated uncertainty score of 0.92, which is indicative of its ability to offer meaningful and dependable uncertainty measures, a capability that remains largely unaddressed by the existing methods.

The iterative learning process facilitated by QBC effectively utilized the initially small set of labelled samples to progressively improve the model's performance, culminating in an overall accuracy of 94.5% after 10 cycles of active learning. This approach not only exemplifies the model's learning efficiency but also highlights the potential for significant reductions in the costs and efforts associated with data annotation.

The findings from this study open several avenues for future research. One immediate direction is the exploration of other multiple modal data types, such as thermal imaging and electrical impedance spectroscopy, to further enrich the dataset and potentially unveil new insights into skin disease characteristics. Additionally, the integration of more advanced generative models could offer enhanced capabilities in data augmentation, particularly for rare skin diseases, thereby addressing the challenge of imbalanced datasets. The scalability and adaptability of the proposed framework to other medical imaging and diagnosis domains present a promising area of exploration. Lastly, the incorporation of real-time feedback mechanisms from dermatologists and clinicians into the active learning cycle could refine the model's learning process, ensuring that the most clinically relevant features are captured and emphasized.

The proposed framework to classify skin diseases can be describe as one of the significant contributions in the medical diagnostics domain, aimed at addressing the multiple challenges posed by handling heterogeneous data, scarce data, and uncertainty in prediction. Using all state-of-the-art methods, this work introduces a robust system, which can guarantee a high classification accuracy with a minimum amount of required extensive labelled datasets. Such approaches include multimodal data fusion using MVAE, transfer learning using Inception-ResNet V2, uncertainty quantification using MC-Bayes CNN, and QBC active learning. Experimental results show the superior performance of the proposed framework on a range of dermatological conditions including common and rare skin diseases. This was evidenced by high overall classification accuracy, precision, and F1-scores as well as reliable uncertainty estimates which are crucial for clinical decision-making. The additional ingredient of active learning further optimizes the training process for the model to achieve competitive performance with fewer labelled samples and is particularly beneficial in the context of medical datasets and samples. The conclusions of this study thus provide promising avenues for enhancing the accuracy and consistency of automated skin disease diagnosis. Future work might therefore expand on this study by adding in other modalities of data input, such as thermal imaging or electrical impedance spectroscopy, to further enlarge the dataset, hopefully improving classification performance also. More importantly, the real-time feedback of the board of dermatologists and clinicians into the active learning cycle ensures that the most clinically relevant features are prioritized for inputting into the system, thereby achieving high performance in a real-world clinical setting. In conclusion, this work now represents a new benchmark for classification on skin diseases by demonstrating for the first time the power of combining multimodal data fusion, transfer learning, uncertainty quantification, and active learning within one unified framework, with major implications for enhancing patient care and outcomes in dermatology scenarios.

Conflict of interest

The authors have no conflict of interest to declare.

Funding

The authors received no specific funding for this work.

References

Andreasen, N., Crandall, H., Brimhall, O., Miller, B., Perez-Tamayo, J., Martinsen, O. G., ... & Sanchez, B. (2021). Skin electrical resistance as a diagnostic and therapeutic biomarker of breast cancer measuring lymphatic regions. *IEEE Access*, *9*, 152322-152332. https://doi.org/10.1109/ACCESS.2021.3123569

Bing, S., Chawang, K., & Chiao, J. C. (2023). A tuned microwave resonant sensor for skin cancerous tumor diagnosis. *IEEE Journal of Electromagnetics, RF and Microwaves in Medicine and Biology*, *7*(4), 320-327. https://doi.org/10.1109/JERM.2023.3281726

Chishti, A. R., Aziz, A., Qureshi, M. A., Abbasi, M. N., Abbasi, D., Iqbal, S. S., ... & Hussain, R. (2023). Advances in antenna-based techniques for detection and monitoring of critical chronic diseases: A comprehensive review. *IEEE Access*, *11*, 104463-104484. https://doi.org/10.1109/ACCESS.2023.3316149

Gururaj, H. L., Manju, N., Nagarjun, A., Aradhya, V. M., & Flammini, F. (2023). DeepSkin: a deep learning approach for skin cancer classification. *IEEE access*, *11*, 50205-50214. https://doi.org/10.1109/ACCESS.2023.3274848

Hamza, M. N., & Islam, M. T. (2023). Designing an extremely tiny dual-band biosensor based on MTMs in the terahertz region as a perfect absorber for non-melanoma skin cancer diagnostics. *IEEE Access*, *11*, 136770-136781. https://doi.org/10.1109/ACCESS.2023.3339562

Hosny, K. M., Elshoura, D., Mohamed, E. R., Vrochidou, E., & Papakostas, G. A. (2023). Deep learning and optimization-based methods for skin lesions segmentation: a review. *IEEE Access*, *11*, 85467-85488.
https://doi.org/10.1109/ACCESS.2023.3303961

Imran, A., Nasir, A., Bilal, M., Sun, G., Alzahrani, A., & Almuhaimeed, A. (2022). Skin cancer detection using combined decision of deep learners. *ieee access*, *10*, 118198-118212.
https://doi.org/10.1109/ACCESS.2022.3220329

Jiang, S., Li, H., & Jin, Z. (2021). A visually interpretable deep learning framework for histopathological image-based skin cancer diagnosis. *IEEE Journal of Biomedical and Health Informatics*, *25*(5), 1483-1494.
https://doi.org/10.1109/JBHI.2021.3052044

Khan, M. A., Muhammad, K., Sharif, M., Akram, T., & de Albuquerque, V. H. C. (2021). Multi-class skin lesion detection and classification via teledermatology. *IEEE journal of biomedical and health informatics*, *25*(12), 4267-4275.
https://doi.org/10.1109/JBHI.2021.3067789

Lan, Z., Cai, S., He, X., & Wen, X. (2022). FixCaps: An improved capsules network for diagnosis of skin cancer. *IEEE Access*, *10*, 76261-76267.
https://doi.org/10.1109/ACCESS.2022.3181225

Lyakhov, P. A., Lyakhova, U. A., & Kalita, D. I. (2023). Multimodal analysis of unbalanced dermatological data for skin cancer recognition. *IEEE Access*, *11*, 131487-131507.
https://doi.org/10.1109/ACCESS.2023.3336289

Magdy, A., Hussein, H., Abdel-Kader, R. F., & Abd El Salam, K. (2023). Performance enhancement of skin cancer classification using computer vision. *IEEE Access*, *11*, 72120-72133.
https://doi.org/10.1109/ACCESS.2023.3294974

Mridha, K., Uddin, M. M., Shin, J., Khadka, S., & Mridha, M. F. (2023). An interpretable skin cancer classification using optimized convolutional neural network for a smart healthcare system. *IEEE Access*, *11*, 41003-41018.
https://doi.org/10.1109/ACCESS.2023.3269694

Naqvi, S. A. R., Mobashsher, A. T., Mohammed, B., Foong, D., & Abbosh, A. (2023). Benign and malignant skin lesions: Dielectric characterization, modelling and analysis in frequency band 1 to 14 GHz. *IEEE Transactions on Biomedical Engineering*, *70*(2), 628-639.
https://doi.org/10.1109/TBME.2022.3199094

Nourinovin, S., Rahman, M. M., Park, S. J., Hamid, H., Philpott, M. P., & Alomainy, A. (2023). Terahertz dielectric characterization of three-dimensional organotypic treated basal cell carcinoma and corresponding double Debye model. *IEEE Transactions on Terahertz Science and Technology*, *13*(3), 246-253.
https://doi.org/10.1109/TTHZ.2023.3263635

Olmez, Y., Koca, G. O., Sengür, A., Acharya, U. R., & Mir, H. (2023). Improved PSO with visit table and multiple direction search strategies for skin cancer image segmentation. *IEEE Access*, *12*, 840-867.
https://doi.org/10.1109/ACCESS.2023.3347587

Pacheco, A. G., & Krohling, R. A. (2021). An attention-based mechanism to combine images and metadata in deep learning models applied to skin cancer classification. *IEEE journal of biomedical and health informatics*, *25*(9), 3554-3563.
https://doi.org/10.1109/JBHI.2021.3062002

Riaz, L., Qadir, H. M., Ali, G., Ali, M., Raza, M. A., Jurcut, A. D., & Ali, J. (2023). A comprehensive Joint Learning System to Detect Skin Cancer. *IEEE Access*, *11*, 79434-79444.
https://doi.org/10.1109/ACCESS.2023.3297644

Saeed, M., Naseer, A., Masood, H., Rehman, S. U., & Gruhn, V. (2023). The power of generative AI to augment for enhanced skin Cancer classification: A deep learning approach. *IEEE Access*, *11*, 130330-130344.
https://doi.org/10.1109/ACCESS.2023.3332628

Schiavoni, R., Maietta, G., Filieri, E., Masciullo, A., & Cataldo, A. (2023). Microwave reflectometry sensing system for low-cost in-vivo skin cancer diagnostics. *IEEE Access*, *11*, 13918-13928.
https://doi.org/10.1109/ACCESS.2023.3243843

Shafi, N., Costantine, J., Kanj, R., Tawk, Y., Ramadan, A. H., Kurban, M., ... & Eid, A. A. (2023). A portable non-invasive electromagnetic lesion-optimized sensing device for the diagnosis of skin cancer (SkanMD). *IEEE Transactions on Biomedical Circuits and Systems*, *17*(3), 558-573.
https://doi.org/10.1109/TBCAS.2023.3260581

Singh, S., & Prajapati, Y. K. (2023). Novel bottom-side polished PCF-based plasmonic biosensor for early detection of hazardous cancerous cells. *IEEE Transactions on NanoBioscience*, *22*(3), 647-654.
https://doi.org/10.1109/TNB.2023.3233990

Qian, X., Weng, T., Han, Q., Wu, C., Xu, H., Hou, M., ... & Gao, X. (2023). SPCB-Net: A multi-scale skin cancer image identification network using self-interactive attention pyramid and cross-layer bilinear-trilinear pooling. *IEEE Access*, *12*, 2272-2287. https://doi.org/10.1109/ACCESS.2023.3347424

Vachmanus, S., Noraset, T., Piyanonpong, W., Rattananukrom, T., & Tuarob, S. (2023). DeepMetaForge: A deep vision-transformer metadata-fusion network for automatic skin lesion classification. *IEEE Access*, *11*, 145467-145484. https://doi.org/10.1109/ACCESS.2023.3345225

Xu, B., & Zhou, F. (2022). The roles of cloud-based systems on the cancer-related studies: a systematic literature review. *IEEE Access*, *10*, 64126-64145. https://doi.org/10.1109/ACCESS.2022.3181147